COUNTY SANITATION DISTRICTS
OF LOS ANGELES COUNTY

1955 Workman Mill Road, Whittier, CA 90601-1400
Mailing Address: P.O. Box 4998, Whittier, CA 90607-4998
Telephone: (562) 699-7411, FAX: (562) 699-5422
www.lacsd.org

GRACE ROBINSON HYDE
Chief Engineer and General Manager

March 30, 2017
File No. 31-370.40.4A

Via Electronic Mail

Dr. Jun Zhu
California Regional Water Quality Control Board
Los Angeles Region
320 West 4th Street, Suite 200
Los Angeles, CA 90013

Dear Dr. Zhu:

Comments on the February 2017 Proposed 2016 Los Angeles Region
<u>Clean Water Act Section 303(d) List of Impaired Waters</u>

The Sanitation Districts of Los Angeles County (Sanitation Districts) appreciate the opportunity to comment on the February 2017 proposed 2016 Los Angeles Region Clean Water Act Section 303(d) List of Impaired Waters (Draft List) prepared by the California Regional Water Quality Control Board, Los Angeles Region (Regional Board). The Sanitation Districts are a consortium of 24 independent special districts serving the wastewater and solid waste management needs of over five million people and 3,300 industries in Los Angeles County, California. The Sanitation Districts currently operate and maintain over 1,400 miles of trunk sewers and 11 wastewater treatment plants that collectively treat over 450 million gallons per day of wastewater. Of the 11 wastewater treatment plants, nine are located in the Los Angeles Region. Seven of the these treatment plants discharge to inland surface waters in the San Gabriel River, Santa Clara River, and Rio Hondo watersheds; one discharges to the Pacific Ocean; and one does not discharge to surface waters but instead solely supplies recycled water for irrigation.

The Sanitation Districts commend Regional Board staff for their diligent implementation of the State Water Resources Control Board's (State Board's) Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (Listing Policy) to produce a Draft List that is generally well-documented and scientifically valid. In addition, the Sanitation Districts greatly appreciate the efforts of the Regional Board staff to make the listing process more transparent, particularly by making the data used to assess listings available on the Regional Board's website and by producing clear fact sheets on each water body/pollutant combination. Staff were also very helpful in addressing questions and meeting with us during the preparation of these comments and their assistance was greatly appreciated.

However, the Sanitation Districts have concerns on some aspects of the Draft List, particularly where the listing thresholds used in the Staff Report appear to differ from receiving water quality objectives contained in the Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) or other regulatory programs. Additionally, there appear to be data errors that impact some listing decisions. General comments relating to these concerns are provided below and detailed specific comments for each listing are provided in Attachment 1 and appendices to this letter.

1. Data Were Incorrectly Attributed to Some Reaches

The Draft List contains a number of newly proposed listings based, in part, on data collected from incorrect reaches. Specific listings where this appears to have occurred include the benthic community and toxicity listings for Santa Clara River Reach 5; the temperature listing for Santa Clara River Reach 6; the toxicity, DO, and iron listings for Rio Hondo Reach 2; and the toxicity listing for San Jose Creek Reach 2.

2. Not All of the Data Submitted for Listing Consideration Were Used in Making the Listing Decision

The Draft List contains a number of newly proposed listings where only a subset of the data submitted for listing consideration were evaluated; these data are included in the data files appended to the Staff Report but were not used in the listing analysis. Specific listings where this appears to have occurred include the toxicity listing for Santa Clara River Reach 5 and the temperature listing for Santa Clara River Reach 6.

3. The Draft List Includes Inappropriate Impairment Listings for "Benthic-Macroinvertebrate Bioassessments"

The Draft February 2017 version of the 2016 303(d) List contains a number of newly proposed listings for "Benthic-Macroinvertebrate Bioassessments." The proposed listings are based on application of the Southern California Coastal Index of Biological Integrity (SCIBI) and, in some cases, the California Stream Condition Index (CSCI). These include listings for Santa Clara River Reach 5, Los Angeles River Reach 3, and Medea Creek Reach 1. The Sanitation Districts believe these proposed listings should be removed, for the reasons listed below.

> <u>Listings Based on the SCIBI and CSCI Are Inconsistent With State Policy</u>. The Water Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (Listing Policy) indicates that water bodies should only be listed for degradation of biological populations if they have significant degradation <u>**relative to reference sites**</u> [emphasis added]. Although the scientists that developed the SCIBI attempted to incorporate reference conditions into the index itself, the reference conditions used to develop the index did not include any low elevation, low gradient locations in Los Angeles County similar to the Los Angeles River and the Santa Clara River reaches of concern.[1] Although the CSCI at least partially addresses some of the problems with the SCIBI by employing a modeled reference condition as opposed to the regional reference pool used by the SCIBI, the lack of any reference sites in large watersheds, low gradient, and low elevation systems still limits the identification of appropriate thresholds using the CSCI.
>
> Section 6.1.5.8 of the Listing Policy also states that when "evaluating biological data and information, RWQCBs shall evaluate all readily available data and information and shall…<u>**evaluate physical habitat data**</u> and other water quality data, when available, to support conclusions about the status of the water segment." [Emphasis added.] All of the reaches mentioned in this comment letter represent reaches that have undergone various levels of physical habitat modifications and there is no indication that an evaluation of the physical habitat was conducted. It is well recognized by the scientific community that a single standard or threshold is not applicable to all waterbodies of the State due to unmanageable

[1] Ode, P.R., A.C. Rehn, J.T. May. 2005. A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams. Environmental Management Vol. 35, No 4, pp. 493-504.

non-pollutant physical habitat alterations that would preclude many streams from ever having biological assemblages similar to reference. The threshold used as the listing criterion for these reaches is therefore likely inappropriate for these modified waterbodies.

<u>Appropriate Thresholds for Interpretation of the CSCI Have Not Yet Been Determined.</u> The State Board has not yet developed any recommended thresholds for the CSCI. The proposed threshold of 0.79 used in the Draft List is the 10th percentile of the reference pool and was used as an arbitrary point of reference for a regional monitoring program with no regulatory vetting. Use of this threshold for impairment listings would result in 10% of the unimpaired reference streams being erroneously listed as impaired. Additionally, it is well recognized by the scientific community that a single standard or threshold will not be applicable to all waterbodies of the State since unmanageable non-pollutant features such as habitat condition/modifications are likely to preclude many streams from ever having biological assemblages similar to reference.

The Sanitation Districts believe that it is inappropriate to make impairment decisions using the SCIBI and premature to rely on the improved, but still limited CSCI for making impairment decisions, particularly in reaches where surrounding development and instream physical habitat limitations are recognized. Therefore, the Sanitation Districts respectfully recommend that the Regional Board delay making decisions regarding benthic macroinvertebrate community impairments in this listing cycle, and instead continue to work with stakeholders, scientists, and the State Board that are currently engaged in efforts to address these and other issues as part of the Biointegrity/Bio-stimulatory Policy.

4. The Draft List Includes Inappropriate Impairment Listings for Temperature

The Draft List contains a number of newly proposed listings for temperature. The Sanitation Districts believe the proposed temperature listings for San Gabriel River Reach 2, San Jose Creek Reach 1, and San Gabriel River Reach 1 should be removed because the impairment listings are inconsistent with the Basin Plan water quality objective for temperature, which states, "at no time shall these WARM-designated waters be raised above 80^{o}F **<u>as a result of waste discharges.</u>**" [Emphasis added.] This water quality objective clearly distinguishes between exceedance of the 80^{o}F standard caused by "waste discharges" and those associated with other causes. Evidence indicates that summertime excursions greater than the 80^{o}F are not caused by wastes discharged but are likely due to elevated ambient air temperature, conductive and radiative heating associated with hardened landscapes, a lack of riparian cover, and increased ambient temperatures related to climate change. Additionally, the Draft List does not contain any analysis or evidence indicating that the elevated temperatures occurred as result of wastes discharged.

Additionally, the Sanitation Districts believe that the proposed temperature listing for Santa Clara River Reach 6 is inappropriate. Measurements for this listing were taken immediately downstream of the Saugus Water Reclamation Plant (WRP), where tertiary treated effluent is discharged along one bank of the Santa Clara River bed. The flow remains isolated from the main channel of the Santa Clara River and percolates rapidly into the soil; groundwater resurfaces downstream near Reach 5 of the Santa Clara River. The predominant natural condition of this stretch of river is dry and would not be expected to support aquatic life without the Saugus WRP discharge; therefore, application of the 80^{o}F water quality objective is unnecessary and inappropriate. The only reasonable alternative for meeting the water quality objective would be to eliminate the discharge flows; however, the California Department of Fish and Wildlife would likely prohibit that option, due to the effluent's contribution to the groundwater and subsequent downstream flows. Upon resurfacing near Reach 5, the water temperature averages 69^{o}F, demonstrating that elevated temperatures in this isolated discharge area are not detrimental to beneficial uses in reaches where water occurs naturally in the river. Finally, elevated ambient temperatures regularly exceed 90 ^{o}F during the summer months, and heavily influence both the Saugus WRP discharge and the immediate downstream receiving water location. As indicated for the other temperature listings, the water quality objective for temperature in the Los Angeles Region Basin Plan clearly distinguishes between temperature exceedances caused by "waste discharges" and those associated with other causes. However, the Draft List does not contain any analysis to distinguish the relative contributions by the temperatures of the ambient air and wastes discharged on the receiving water.

5. Thresholds Used For Toxicity Impairment Listings Are Inconsistent With Basin Plan Objectives

The Draft List contains a number of newly proposed listings for toxicity that include San Gabriel River Estuary, San Gabriel River Reach 3, Rio Hondo Reach 2, and Santa Clara River Reach 5. These listings should be removed for the reasons below.

<u>The Acute Toxicity Impairment Criterion is Inconsistent With the Basin Plan Water Quality Objective for Acute Toxicity</u>
The Staff Report fact sheets for the specific listings mentioned above state that "<100% survival (acute) was considered an exceedance." However, the Basin Plan states that "the acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board." Therefore, a single-test threshold of less than 70% survival should be used to determine impairments; even a threshold of less than 90% survival would still be more conservative than Basin plan objective.

<u>The Chronic Toxicity Impairment Criterion is Inconsistent With Water Quality Objective Interpretations Provided in NPDES Permits</u>

The Staff Report fact sheets for the specific listings mentioned above indicate that a single NOEC result of less than 100% receiving water represents an exceedance of the water quality objective. Although the Basin Plan provides no numeric chronic toxicity objectives, recently adopted Los Angeles Region NPDES permits do provide very specific direction on interpretation of the narrative water quality objectives for chronic toxicity. In a number of these permits, a footnote associated with the Receiving Water Monitoring Requirements Table of the Monitoring and Reporting Program states; "The <u>**median monthly summary result is a threshold value**</u> for a determination of meeting the narrative receiving water objective and shall be reported as 'Pass' or 'Fail'."[2] [Emphasis added.]

In addition to aligning with the NPDES permit language, use of a monthly median will also address concerns regarding false positive error rates. The USEPA has determined that the expected false positive error rate for chronic toxicity testing using the NOEC is 5%. With this error rate, on average, one in 20 individual chronic toxicity tests will be erroneously identified as "toxic" using the NOEC, and there is a nearly 34% probability that 2 or more individual chronic toxicity test exceedances would be observed within a set of 24 discrete measurements in a completely non-toxic stream reach. When there are two or more exceedances out of 24

[2] Pomona WRP - ORDER R4-2014-0212-A01 NPDES NO. CA0053619, Long Beach WRP - ORDER NO. R4-2015-0123 NPDES NO. CA0054119, Los Coyotes - ORDER NO. R4-2015-0124 NPDES NO. CA0054011, San Jose Creek WRP - ORDER R4-2015-0070 NPDES NO. CA0053911, Saugus WRP- ORDER R4-2015-0072 NPDES NO. CA0054313, Valencia WRP- ORDER R4-2015-0071 NPDES NO. CA0054216, Whittier Narrows WRP ORDER R4-2014-0213-A01 NPDES NO. CA0053716

measurements, the Listing Policy specifies that a reach be listed as impaired. Therefore, using single chronic toxicity exceedances as the 303(d) criterion would eventually result in more and more non-toxic stream reaches being erroneously listed over time. However, using a monthly median chronic toxicity exceedance threshold would reduce the likelihood of inappropriate reach listings due to false positive chronic toxicity results to less than 1%.

6. Specific Comments on Individual Reach/pollutant Listing Decisions

In addition to these general comments, the Sanitation Districts have comments on some specific listing decisions. As stated above, detailed comments are provided in the appendices to this letter. Because the implications of erroneous listings are substantial, the Sanitation Districts urge the Regional Board to consider this information in making the appropriate changes to the Draft List.

In conclusion, the Sanitation Districts would like to thank the Regional Board for its efforts up to this point in revising the proposed 2016 303(d) List. We urge the Regional Board to consider the information and analysis contained in this letter to complete the development of a scientifically and legally defensible list with a sound and consistent basis. If you have any questions regarding our comments or the information and data we are providing to you, please contact Phil Markle at (562) 908-4288, extension 2808, pmarkle@lacsd.org.

Very truly yours,

Philip L. Friess

Philip L. Friess
Department Head
Technical Services

PLF:PJM:nm
Attachments

cc: LB Nye, Jun Zhu, Kangshi Wang, Regional Board, Los Angeles Region

!űűI·ŎKϒŜyű м

Sanitation Districts of Los Angeles County
2016 303(d) List Fact Sheets

Fact Sheet #1

Water Body:	**San Gabriel River Estuary**
Pollutant:	**Toxicity**
Listing:	**List on 303(d) List (TMDL Required List)**
Comment & Recommendation:	**Do Not List – Water Quality Objectives Being Achieved**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is currently proposing that a new listing for toxicity be made to the 303(d) list for the San Gabriel River Estuary, based on one line of evidence: 14 of 113 samples exceeded the objective. The Districts believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved, for the reasons listed below; supporting evidence is provided in the sections that follow.

- Appendix A of this letter contains the full set of data applicable to this listing from Appendix G of the Regional Board Draft Staff Report. Using the temporal range indicated (June 2006 through May 2010), only six of 120 samples failed the thresholds specified in the fact sheet. According to Table 3.1 of the California Clean Water Act 303(d) Listing Policy (Listing Policy), an impairment listing is appropriate if 11 or more exceedances are observed when 120 samples are available.

- Although the Staff Report fact sheet states that “<100% survival (acute) was considered an exceedance,” the Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that “the acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board.” Therefore, a single-test threshold of less than 70% survival should be used to determine impairments; even a threshold of less than 90% survival would still be more conservative than Basin Plan objective. Applying a 90% threshold, none of the 120 samples would have exceeded the water quality objective. Therefore, this reach fails to meet the listing criteria for toxicity.

- The full set of data appended to Appendix G of the Staff Report, including those that fell outside the indicated temporal range, contain a total 151 discrete toxicity tests. Sixteen failed the <100% acute survival threshold. Using a conservative 90% acute survival threshold, there are no toxicity exceedances, and the number of measured exceedances is insufficient to place this water segment on the section 303(d) list.

Fact Sheet #1

Use of a <100% Survival Water Quality Objective Threshold Is Inappropriate and Unsupported.

All of the San Gabriel River toxicity "exceedances" indicated in the Regional Board Staff Report were acute toxicity results of <100% survival in undiluted receiving water; the lowest result in the data tables was a percent survival of 90%. However, as described in the subsections below, the 100% threshold is inconsistent with the Basin Plan and other documentation supplied by the Regional Board, with criteria used for other acute toxicity listings, and with the results of statistical testing.

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with the Basin Plan and Other Documentation from the Regional Board.

In the Basin Plan and Sanitation Districts NPDES permits, the narrative acute toxicity receiving water quality objective is numerically defined: "the average survival in the undiluted receiving water for any three (3) consecutive 96-hour static, static-renewal, or continuous flow bioassay tests shall be at least 90%, and (ii) no single test producing less than 70% survival." Furthermore, the Water Quality Objective/Criterion reference provided in the Staff Report indicates that "the power to detect differences drops quickly below 15%, therefore care should be taken when declaring samples less than 15% different from the control as toxic." Following this reference, an exceedance of the water quality objective would be potentially questionable if the survival response was greater than 85%. Finally, the Sanitation Districts NPDES permits specify the use of a laboratory method with minimum test acceptability criteria of 90% for non-toxic control survival in the freshwater fish acute test, indicating that percent survival in undiluted receiving water of 90% or greater would be consistent with an expected response in a non-toxic sample.

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with Criteria Used for Other Acute Toxicity Listings.

Regional Boards across California use a variety of thresholds to determine is acute toxicity water quality objectives are being met. However, based on a review of approved listing decisions from across California, a threshold of less than 100% was never used. Below is summary of criteria utilized to evaluate percent effect/response acute data:

Region 2

"Acute toxicity is defined as a median of less than 90 percent survival, or less than 70 percent survival, 10 percent of the time, of test organisms in a 96-hour static or continuous flow test." (Water Quality Control Plan (Basin Plan) for the San Francisco Bay Basin, Section 3.3.18).
http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/planningtmdls/basinplan/web/bp_ch3.shtml#3.3.18

"Statistical evaluation and a default threshold of 80% of the control value were used to establish whether the sediment exhibited significant toxicity adversely impacting aquatic organisms." (Proposed 2016 Section 303(d) and 305(b) Integrated Report, Region 2, Appendix G, Line of Evidence (LOE) for Decision ID 43948, Toxicity, Lagunitas Creek)
http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/2016_303d/00653.shtml#43948

Region 3

"Statistical evaluation (alpha = 0.05) and a default threshold of 80% of the control value were used to establish whether water exhibited significant toxicity adversely impacting aquatic organisms." (Final 2012 California Integrated Report (Clean Water Act Section 303(d) List / 305(b) Report, Region 3, Line of Evidence (LOE) for Decision ID 28270, Toxicity, Kirker Creek).
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01826.shtml#28270

Fact Sheet #1

Region 4

"There shall be no acute toxicity in ambient waters, including mixing zones. The acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board."
Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties, Page 3-38.
http://www.waterboards.ca.gov/losangeles/water_issues/programs/basin_plan/electronics_documents/Final%20Chapter%203%20Text.pdf

"Non-toxic if greater than or equal to 80% survival; moderately toxic if between 50 to 80% survival; and highly toxic if less than 50% survival." (Draft 2016 Section 303(d) and 305(b) Integrated Report, Region 4, Line of Evidence (LOE) for Decision ID 43062, Toxicity, Dominguez Channel Estuary (unlined portion below Vermont Ave).
http://www.waterboards.ca.gov/losangeles/water_issues/programs/303d/2016/Appendix_G/00134.shtml#43062

"Toxicity was defined as a reduction of the NOEC below 100% and was considered significant if the effect on the sample exposure was greater than 25%." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 4, Line of Evidence (LOE) for Decision ID 28344, Toxicity, Dominguez Channel (lined portion above Vermont Ave)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01077.shtml#28344

Region 5

"Significant toxicity is defined as a statistically significant (p<0.5) increase in mortality (>=20%) compared to the laboratory control." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 5, Line of Evidence (LOE) for Decision ID 26730, Unknown Toxicity, Feather River, Lower (Lake Oroville Dam to Confluence with Sacramento River)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01204.shtml#26730

Region 8

"Survival of organisms during toxicity bioassays no less than 80%." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 8, Line of Evidence (LOE) for Decision ID 27875, Sediment Toxicity, Elsinore, Lake)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/00489.shtml#27875

Region 9

"Samples were found to exhibit toxicity when the No Observed Effect Concentration (NOEC) or median lethal concentration (LC50) for any given species was estimated to be less than 100% of the test sample concentration." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 9, Line of Evidence (LOE) for Decision ID 28361, Toxicity, Agua Hedionda Creek)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01602.shtml#28361

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with the Results of Statistical Testing.

Although the data summary provided as part of the 303(d) data submission included only percent survival results, control data were included when the data were submitted as part of routine NPDES compliance reports. Using the control data, the Sanitation Districts staff conducted statistical analyses for the acute toxicity data included in the Staff Report and found that no statistically significant differences were observed between the control and undiluted receiving water samples.

Fact Sheet #1

The Water Quality Objective/Threshold for Chronic Toxicity Should Be a Monthly Median.

When a sample fails a chronic toxicity test, additional samples may be collected during the same calendar month in an effort to confirm the result and identify potential toxicants. Application of a single test failure chronic toxicity water quality objective provides a disincentive for this type of proactive monitoring and is at odds with the intent of chronic toxicity testing.

Chronic toxicity is intended to assess potential aquatic life impacts associated with long-term exposures. Therefore, it is analogous to the Criterion Continuous Concentration (CCC) used for estimating "safe" chemical concentrations for long-term exposure as opposed to the Criterion Maximum Concentration (CMC) that is intended to protect against short-term exposure. In EPA's Region 9 and 10 Toxicity Training Tool, the CCC is defined as "the highest in-stream concentration of a toxic or an effluent to which organisms can be exposed indefinitely without causing unacceptable effects such as the exceedance of a chronic water quality criterion."[1] This same document also recommends "direct application of 1.0 TUc as the monthly compliance level for NPDES discharges without a mixing zone or dilution allowance. In conjunction and limited to this discharge situation, because: (1) there are no values below 1.0 TUc and (2) an arithmetic average is sensitive to extremely large and small values, the median is favored as the better measure of central tendency for the monthly compliance level."

Although the Basin Plan provides no numeric chronic toxicity objectives, recently adopted Los Angeles Region NPDES permits do provide very specific direction on interpretation of the narrative water quality objectives for chronic toxicity. In the Long Beach WRP NPDES permit, footnote 25 of Table E-6 of the Monitoring and Reporting Program states; "The **median monthly summary result is a threshold value** for a determination of meeting the narrative receiving water objective and shall be reported as 'Pass' or 'Fail.'"[2] [Emphasis added.]

In addition to aligning with the NPDES permit language, use of a monthly median will also address concerns regarding false positive error rates. The USEPA has determined that the expected false positive error rate for chronic toxicity testing using the NOEC is 5%. With this error rate, on average, one in 20 individual chronic toxicity tests will be erroneously identified as "toxic" using the NOEC, and there is a nearly 34% probability that 2 or more individual chronic toxicity test exceedances would be observed within a set of 24 discrete measurements in a completely non-toxic stream reach. When there are two or more exceedances out of 24 measurements, the Listing Policy specifies that a reach be listed as impaired. Therefore, using single chronic toxicity exceedances as the 303(d) criterion would, over time, result in more and more non-toxic stream reaches being erroneously listed. However, using a monthly median chronic toxicity exceedance threshold would reduce the likelihood of inappropriate reach listings due to false positive chronic toxicity results to less than 1%.

Since the data set used for assessing the San Gabriel River Estuary does not include multiple tests conducted during the same month, the individual test result is the also the monthly median. Therefore, this correction will have no impact on the listing decision for toxicity in the San Gabriel River Estuary, but may have a significant impact in other reaches.

---

[1] Denton DL, Miller JM, Stuber RA. 2007. EPA Regions 9 and 10 toxicity training tool (TTT). November 2007. San Francisco, CA: https://www.epa.gov/sites/production/files/documents/ToxTrainingTool10Jan2010.pdf

[2] Pomona WRP - ORDER R4-2014-0212-A01 NPDES NO. CA0053619, Long Beach WRP - ORDER NO. R4-2015-0123 NPDES NO. CA0054119, Los Coyotes - ORDER NO. R4-2015-0124 NPDES NO. CA0054011, San Jose Creek WRP - ORDER R4-2015-0070 NPDES NO. CA0053911, Saugus WRP- ORDER R4-2015-0072 NPDES NO. CA0054313, Valencia WRP- ORDER R4-2015-0071 NPDES NO. CA0054216, Whittier Narrows WRP ORDER R4-2014-0213-A01 NPDES NO. CA0053716

Fact Sheet #2

Water Body: San Gabriel River Reach 3
Pollutant: Toxicity
Listing: List on 303(d) List (TMDL Required List)
Comment & Recommendation: Do Not List – Water Quality Objectives Being Achieved

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for toxicity be made to the 303(d) list for Reach 3 of the San Gabriel River, based on one line of evidence using two datasets: 2 of 38 samples exceeded the objective in a dataset related to a previously conducted TMDL study and 13 of 75 samples exceeded the objective in a second dataset comprised of routine receiving water tests conducted as part of an NPDES permit. The Sanitation Districts of Los Angeles County (Sanitation Districts) believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved, for the reasons listed below; supporting evidence is provided in the sections that follow.

- Appendix A of this letter contains the full set of data applicable to this listing from Appendix G of the Regional Board Draft Staff Report. No data related to the TMDL study were provided; therefore, the number of tests and exceedances reported (2 of 38) could not be independently verified and were assumed to be accurate. For the dates indicated (June 2006 through May 2010), 13 exceedances were associated with only 66 samples. Combining the two datasets resulted seven acute and eight chronic toxicity exceedances out of 104 samples.

- Although the Staff Report fact sheet states that "<100% survival (acute) was considered an exceedance," the Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that "the acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board." Therefore, a single-test threshold of less than 70% survival should be used to determine impairments; even a threshold of less than 90% survival would still be more conservative than Basin plan objective. Applying a 90% threshold, no acute toxicity samples in the dataset exceeded the water quality objective and 8 of 104 total samples exceeded the objective. According to Table 3.1 of the California Clean Water Act 303(d) Listing Policy (Listing Policy), an impairment listing is appropriate if 9 or more exceedances are observed when 104 samples are available. Therefore, this reach fails to meet the listing criteria for toxicity.

- The Staff Report considered each chronic toxicity test result as an independent data point, even when multiple bioassays were conducted within a single month. However, the San Jose Creek (SJCWRP) and Whittier Narrows Water Reclamation Plant (WNWRP) permits state that the water quality objective for chronic toxicity is based on a monthly median; therefore, all tests within a single month should be considered part of a monthly median, rather than independent tests. Based on appropriate application of the monthly median as the chronic water quality objective (and a 90% acute toxicity threshold), there were 6 toxicity exceedances out of a total of 96 tests. According to Table 3.1 of the California Clean Water Act 303(d) Listing Policy (Listing Policy), an impairment listing is appropriate if 9 or more exceedances are observed when 96 samples are available. Therefore, this reach fails to meet the listing criteria for toxicity.

- The full set of data (sets 1 and 2) appended to Appendix G of the Staff Report for all dates, including those outside the indicated temporal range, contain a total of 119 discrete toxicity tests. Using a conservative 90% acute survival threshold and appropriate monthly median chronic threshold, there are no acute exceedances and 6 chronic exceedances out of 110 results. This total

Fact Sheet #2

does not meet the minimum number of measured exceedances needed to place a water segment on the section 303(d) list.

Use of a <100% Survival Effect Water Quality Objective Threshold Is Inappropriate and Unsupported for Acute Toxicity Testing.

Seven of 15 San Gabriel River toxicity "exceedances" indicated in the Regional Board Staff Report were based on acute toxicity results of <100% survival in undiluted receiving water; the lowest result in the data tables was 97.5% survival. However, as described in the subsections below, the 100% threshold is inconsistent with the Basin Plan and other documentation supplied by the Regional Board, with criteria used for other acute toxicity listings, and with the results of statistical testing.

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with the Basin Plan and Other Documentation from the Regional Board.

In the Basin Plan and Sanitation Districts NPDES permits, the narrative acute toxicity receiving water quality objective is numerically defined: "the average survival in the undiluted receiving water for any three (3) consecutive 96-hour static, static-renewal, or continuous flow bioassay tests shall be at least 90%, and (ii) no single test producing less than 70% survival." Furthermore, the Water Quality Objective/Criterion reference provided in the Staff Report indicates that "the power to detect differences drops quickly below 15%, therefore care should be taken when declaring samples less than 15% different from the control as toxic." Following this reference, an exceedance of the water quality objective would be potentially questionable if the survival response was greater than 85%. Finally, the Sanitation Districts NPDES permits specify the use of a laboratory method with minimum test acceptability criteria of 90% for non-toxic control survival in the freshwater fish acute test, indicating that percent survival in undiluted receiving water of 90% or greater would be consistent with an expected response in a non-toxic sample.

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with Criteria Used for Other Acute Toxicity Listings.

Regional Boards across California use a variety of thresholds to determine if acute toxicity water quality objectives are being met. However, based on a review of approved listing decisions from across California, a threshold of less than 100% was never used. Below is summary of criteria utilized to evaluate percent effect/response acute data:

Region 2

"Acute toxicity is defined as a median of less than 90 percent survival, or less than 70 percent survival [in a single test], 10 percent of the time, of test organisms in a 96-hour static or continuous flow test."

(Water Quality Control Plan (Basin Plan) for the San Francisco Bay Basin, Section 3.3.18).
http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/planningtmdls/basinplan/web/bp_ch3.shtml#3.3.18

"Statistical evaluation and a default threshold of 80% of the control value were used to establish whether the sediment exhibited significant toxicity adversely impacting aquatic organisms." (Proposed 2016 Section 303(d) and 305(b) Integrated Report, Region 2, Appendix G, Line of Evidence (LOE) for Decision ID 43948, Toxicity, Lagunitas Creek)
http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/2016_303d/00653.shtml#43948

Fact Sheet #2

Region 3

"Statistical evaluation (alpha = 0.05) and a default threshold of 80% of the control value were used to establish whether water exhibited significant toxicity adversely impacting aquatic organisms." (Final 2012 California Integrated Report (Clean Water Act Section 303(d) List / 305(b) Report, Region 3, Line of Evidence (LOE) for Decision ID 28270, Toxicity, Kirker Creek).
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01826.shtml#28270

Region 4

"There shall be no acute toxicity in ambient waters, including mixing zones. The acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board."
Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties, Page 3-38.
http://www.waterboards.ca.gov/losangeles/water_issues/programs/basin_plan/electronics_documents/Final%20Chapter%203%20Text.pdf

"Non-toxic if greater than or equal to 80% survival; moderately toxic if between 50 to 80% survival; and highly toxic if less than 50% survival." (Draft 2016 Section 303(d) and 305(b) Integrated Report, Region 4, Line of Evidence (LOE) for Decision ID 43062, Toxicity, Dominguez Channel Estuary (unlined portion below Vermont Ave).
http://www.waterboards.ca.gov/losangeles/water_issues/programs/303d/2016/Appendix_G/00134.shtml#43062

"Toxicity was defined as a reduction of the NOEC below 100% and was considered significant if the effect on the sample exposure was greater than 25%." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 4, Line of Evidence (LOE) for Decision ID 28344, Toxicity, Dominguez Channel (lined portion above Vermont Ave)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01077.shtml#28344

Region 5

"Significant toxicity is defined as a statistically significant ($p<0.5$) increase in mortality (>=20%) compared to the laboratory control." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 5, Line of Evidence (LOE) for Decision ID 26730, Unknown Toxicity, Feather River, Lower (Lake Oroville Dam to Confluence with Sacramento River)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01204.shtml#26730

Region 8

"Survival of organisms during toxicity bioassays no less than 80%." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 8, Line of Evidence (LOE) for Decision ID 27875, Sediment Toxicity, Elsinore, Lake)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/00489.shtml#27875

Region 9

"Samples were found to exhibit toxicity when the No Observed Effect Concentration (NOEC) or median lethal concentration (LC50) for any given species was estimated to be less than 100% of the test sample concentration." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 9, Line of Evidence (LOE) for Decision ID 28361, Toxicity, Agua Hedionda Creek)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01602.shtml#28361

Fact Sheet #2

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with the Results of Statistical Testing.

Although the data summary provided as part of the 303(d) data submission included only percent survival results, control data were included when the data were submitted as part of routine NPDES compliance reports. Using the control data, the Sanitation Districts staff conducted statistical analyses for the acute toxicity data included in the Staff Report and found that no statistically significant differences were observed between the control and undiluted receiving water samples.

The Water Quality Objective/Threshold for Chronic Toxicity Should Be a Monthly Median.

When a sample fails a chronic toxicity test, additional samples may be collected during the same calendar month in an effort to confirm the result and identify potential toxicants. Application of a single test failure chronic toxicity water quality objective provides a disincentive for this type of proactive monitoring and is at odds with the intent of chronic toxicity testing.

Chronic toxicity is intended to assess potential aquatic life impacts associated with long-term exposures. Therefore, it is analogous to the Criterion Continuous Concentration (CCC) used for estimating “safe” chemical concentrations for long-term exposure as opposed to the Criterion Maximum Concentration (CMC) that is intended to protect against short-term exposure. In EPA’s Region 9 and 10 Toxicity Training Tool, the CCC is defined as “the highest in-stream concentration of a toxic or an effluent to which organisms can be exposed indefinitely without causing unacceptable effects such as the exceedance of a chronic water quality criterion.”[1] This same document also recommends “direct application of 1.0 TUc as the monthly compliance level for NPDES discharges without a mixing zone or dilution allowance. In conjunction and limited to this discharge situation, because: (1) there are no values below 1.0 TUc and (2) an arithmetic average is sensitive to extremely large and small values, the median is favored as the better measure of central tendency for the monthly compliance level.”

Although the Basin Plan provides no numeric chronic toxicity objectives, recently adopted Los Angeles Region NPDES permits do provide very specific direction on interpretation of the narrative water quality objectives for chronic toxicity. In the Whittier Narrows WRP NPDES permit, footnote 22 of Table E-5a of the Monitoring and Reporting Program states of the Monitoring and Reporting Program states; “The **median monthly summary result is a threshold value** for a determination of meeting the narrative receiving water objective and shall be reported as ‘Pass’ or ‘Fail.’”[2] [Emphasis added.]

In addition to aligning with the NPDES permit language, use of a monthly median will also address concerns regarding false positive error rates. The USEPA has determined that the expected false positive error rate for chronic toxicity testing using the NOEC is 5%. With this error rate, on average, one in 20 individual chronic toxicity tests will be erroneously identified as “toxic” using the NOEC, and there is a nearly 34% probability that 2 or more individual chronic toxicity test exceedances would be observed within a set of 24 discrete measurements in a completely non-toxic stream reach. When there are two or

[1] Denton DL, Miller JM, Stuber RA. 2007. EPA Regions 9 and 10 toxicity training tool (TTT). November 2007. San Francisco, CA: https://www.epa.gov/sites/production/files/documents/ToxTrainingTool10Jan2010.pdf

[2] Pomona WRP - ORDER R4-2014-0212-A01 NPDES NO. CA0053619, Long Beach WRP - ORDER NO. R4-2015-0123 NPDES NO. CA0054119, Los Coyotes - ORDER NO. R4-2015-0124 NPDES NO. CA0054011, San Jose Creek WRP - ORDER R4-2015-0070 NPDES NO. CA0053911, Saugus WRP- ORDER R4-2015-0072 NPDES NO. CA0054313, Valencia WRP- ORDER R4-2015-0071 NPDES NO. CA0054216, Whittier Narrows WRP ORDER R4-2014-0213-A01 NPDES NO. CA0053716

Fact Sheet #2

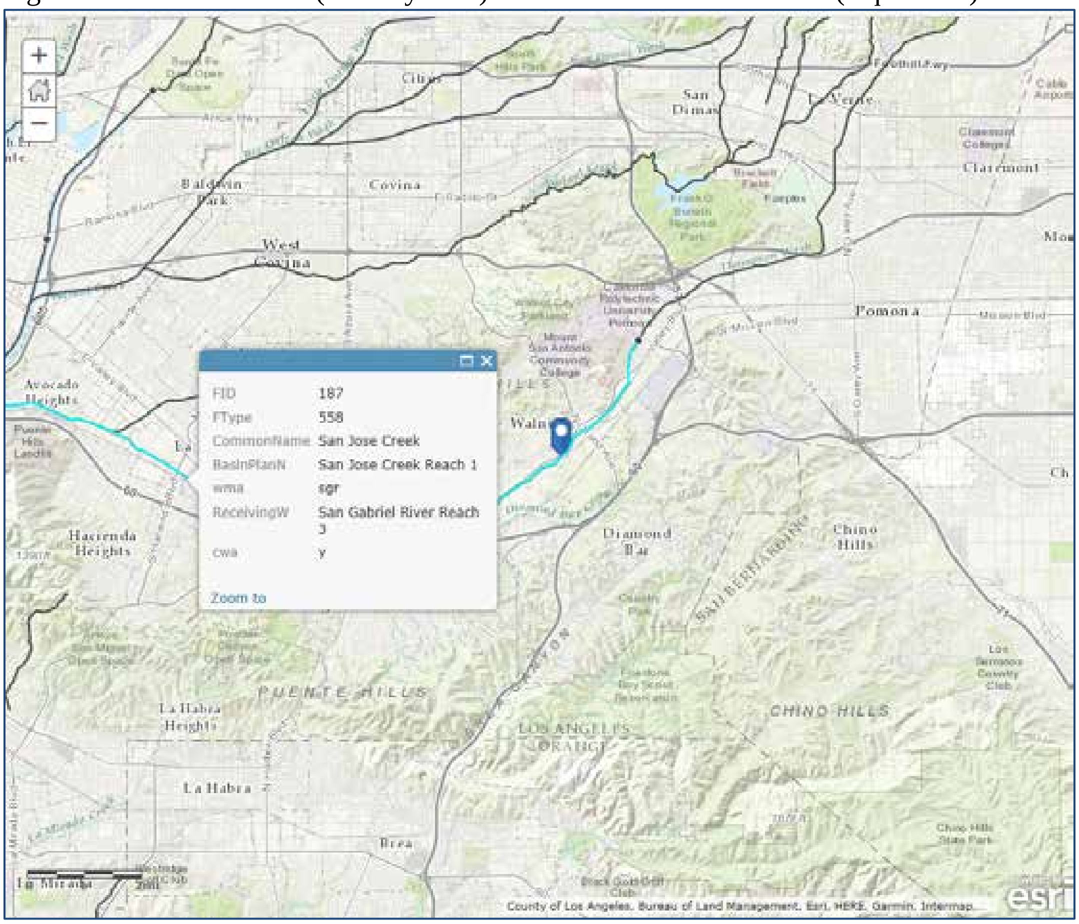
more exceedances out of 24 measurements, the Listing Policy specifies that a reach be listed as impaired. Therefore, using single chronic toxicity exceedances as the 303(d) criterion would, over time, result in more and more non-toxic stream reaches being erroneously listed. However, using a monthly median chronic toxicity exceedance threshold would reduce the likelihood of inappropriate reach listings due to false positive chronic toxicity results to less than 1%.

Fact Sheet #3

Water Body:	**San Jose Creek Reach 2**
Pollutant:	**Toxicity**
Listing:	**List on 303(d) List (TMDL Required List)**
Comment & Recommendation:	**Apply Data to Reach 1**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for toxicity be made to the 303(d) list for Reach 2 of the San Jose Creek, based on one line of evidence: 8 of 24 samples exceeded the objective. The Sanitation Districts believe this proposed listing is inappropriate and should be moved to Reach 1. All cited toxicity data is from receiving water station RC (N 34° 01' 8.6" W 117° 50' 27.7") for the Pomona Water Reclamation Plant, which is located in Reach 1 of San Jose Creek (Figure 1). This reach is already listed for toxicity under section 303(d).

Figure 1. Station Pom-RC (Blue Symbol) and San Jose Creek Reach 1 (Aqua Line)



Fact Sheet #4

Water Body:	**Rio Hondo Reach 2**
Pollutant:	**Toxicity**
Listing:	**List on 303(d) List (TMDL Required List)**
Comment & Recommendation:	**Do Not List – Water Quality Objectives Being Achieved**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for toxicity be made to the 303(d) list for Reach 2 of the Rio Hondo, based on one line of evidence: 5 of 31 samples exceeded the objective. The Districts believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved, for the reasons listed below; supporting evidence is provided in the sections that follow.

- Appendix A of this letter contains the full set of data applicable to this listing from Appendix G of the Regional Board Draft Staff Report. All cited toxicity data are from receiving water station RD1 for the Whittier Narrows Water Reclamation Plant (WNWRP). This sampling location (N 34° 02' 26.5" W 118° 04' 27") is in Reach 3 of the Rio Hondo, not Reach 2 (Figure 1).

- Using the data for the temporal range indicated (June 2006 through May 2010), 7 of 33 samples failed the thresholds specified in the fact sheet.

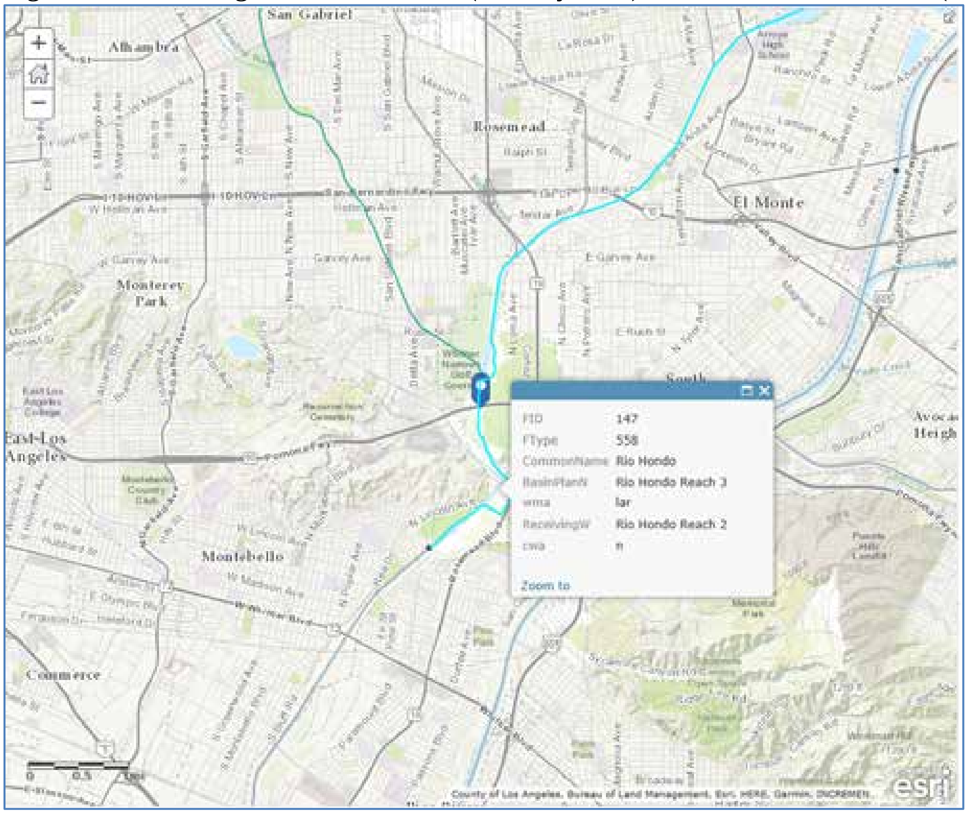
- Although the Staff Report fact sheet states that "<100% survival (acute) was considered an exceedance," the Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that "the acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board." Therefore, a single-test threshold of less than 70% survival should be used to determine impairments; even a threshold of less than 90% survival would still be more conservative than Basin plan objective. Applying a 90% threshold, no samples exceeded the acute toxicity water quality objective.

- The Staff Report considered each chronic toxicity test result as independent data, even when multiple bioassays were conducted within a single month. However, the WNWRP permit states that the water quality objective for chronic toxicity is based on a monthly median; therefore, all tests within a single month should be considered part of a monthly median, rather than independent tests. Based on appropriate application of the monthly median as the chronic water quality objective (and a 90% acute toxicity threshold), there were 2 toxicity exceedances out of 31 tests. According to Table 3.1 of the California Clean Water Act 303(d) Listing Policy (Listing Policy), an impairment listing is appropriate if 3 or more exceedances are observed when 31 samples are available. Therefore, Reach 2 of the Rio Hondo fails to meet the listing criteria for toxicity.

- The full set of data appended to Attachment G of the Staff Report, including those that fell outside the indicated temporal range, contains a total 38 discrete toxicity tests. Using a conservative 90% acute survival threshold and appropriate monthly median chronic threshold, there are no acute exceedances and 2 chronic exceedances out of 36 results. This total does not meet the minimum number of measured exceedances needed to place a water segment on the section 303(d) list.

Fact Sheet #4

Figure 1. Monitoring Station WN-RD1 (Blue Symbol) and Rio Hondo Reach 3 (Aqua Line)



Use of a <100% Survival Water Quality Objective Threshold Is Inappropriate and Unsupported.

Three of the Rio Hondo toxicity "exceedances" indicated in the Regional Board Staff Report were acute toxicity results of <100% survival in undiluted receiving water; the lowest result in the data tables was a percent survival of 90%. However, as described in the subsections below, the 100% threshold is inconsistent with the Basin Plan and other documentation supplied by the Regional Board, with criteria used for other acute toxicity listings, and with the results of statistical testing.

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with the Basin Plan and Other Documentation from the Regional Board.

In the Basin Plan and Sanitation Districts NPDES permits, the narrative acute toxicity receiving water quality objective is numerically defined: "the average survival in the undiluted receiving water for any three (3) consecutive 96-hour static, static-renewal, or continuous flow bioassay tests shall be at least 90%, and (ii) no single test producing less than 70% survival." Furthermore, the Water Quality Objective/Criterion reference provided in the Staff Report indicates that "the power to detect differences drops quickly below 15%, therefore care should be taken when declaring samples less than 15% different from the control as toxic." Following this reference, an exceedance of the water quality objective would be potentially questionable if the survival response was greater than 85%. Finally, the Sanitation Districts NPDES permits specify the use of a laboratory method with minimum test acceptability criteria of 90% for non-toxic control survival in the freshwater fish acute test, indicating that percent survival in undiluted receiving water of 90% or greater would be consistent with an expected response in a non-toxic sample.

Fact Sheet #4

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with Criteria Used for Other Acute Toxicity Listings.

Regional Boards across California use a variety of thresholds to determine is acute toxicity water quality objectives are being met. However, based on a review of approved listing decisions from across California, a threshold of less than 100% was never used. Below is summary of criteria utilized to evaluate percent effect/response acute data:

Region 2

"Acute toxicity is defined as a median of less than 90 percent survival, or less than 70 percent survival, 10 percent of the time, of test organisms in a 96-hour static or continuous flow test." (Water Quality Control Plan (Basin Plan) for the San Francisco Bay Basin, Section 3.3.18). http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/planningtmdls/basinplan/web/bp_ch3.shtml#3.3.18

"Statistical evaluation and a default threshold of 80% of the control value were used to establish whether the sediment exhibited significant toxicity adversely impacting aquatic organisms." (Proposed 2016 Section 303(d) and 305(b) Integrated Report, Region 2, Appendix G, Line of Evidence (LOE) for Decision ID 43948, Toxicity, Lagunitas Creek)
http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/2016_303d/00653.shtml#43948

Region 3

"Statistical evaluation (alpha = 0.05) and a default threshold of 80% of the control value were used to establish whether water exhibited significant toxicity adversely impacting aquatic organisms." (Final 2012 California Integrated Report (Clean Water Act Section 303(d) List / 305(b) Report, Region 3, Line of Evidence (LOE) for Decision ID 28270, Toxicity, Kirker Creek).
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01826.shtml#28270

Region 4

"There shall be no acute toxicity in ambient waters, including mixing zones. The acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board."
Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties, Page 3-38.
http://www.waterboards.ca.gov/losangeles/water_issues/programs/basin_plan/electronics_documents/Final%20Chapter%203%20Text.pdf

"Non-toxic if greater than or equal to 80% survival; moderately toxic if between 50 to 80% survival; and highly toxic if less than 50% survival." (Draft 2016 Section 303(d) and 305(b) Integrated Report, Region 4, Line of Evidence (LOE) for Decision ID 43062, Toxicity, Dominguez Channel Estuary (unlined portion below Vermont Ave).
http://www.waterboards.ca.gov/losangeles/water_issues/programs/303d/2016/Appendix_G/00134.shtml#43062

"Toxicity was defined as a reduction of the NOEC below 100% and was considered significant if the effect on the sample exposure was greater than 25%." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 4, Line of Evidence (LOE) for Decision ID 28344, Toxicity, Dominguez Channel (lined portion above Vermont Ave)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01077.shtml#28344

Region 5

"Significant toxicity is defined as a statistically significant ($p<0.5$) increase in mortality (>=20%) compared to the laboratory control." (Final California 2012 Integrated Report (303(d) List/305(b)

Fact Sheet #4

Report), Region 5, Line of Evidence (LOE) for Decision ID 26730, Unknown Toxicity, Feather River, Lower (Lake Oroville Dam to Confluence with Sacramento River)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01204.shtml#26730

Region 8
"Survival of organisms during toxicity bioassays no less than 80%." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 8, Line of Evidence (LOE) for Decision ID 27875, Sediment Toxicity, Elsinore, Lake)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/00489.shtml#27875

Region 9
"Samples were found to exhibit toxicity when the No Observed Effect Concentration (NOEC) or median lethal concentration (LC50) for any given species was estimated to be less than 100% of the test sample concentration." (Final California 2012 Integrated Report (303(d) List/305(b) Report), Region 9, Line of Evidence (LOE) for Decision ID 28361, Toxicity, Agua Hedionda Creek)
http://www.waterboards.ca.gov/water_issues/programs/tmdl/2012state_ir_reports/01602.shtml#28361

Use of a <100% Survival Water Quality Objective Threshold Is Inconsistent with the Results of Statistical Testing.

Although the data summary provided as part of the 303(d) data submission included only percent survival results, control data were included when the data were submitted as part of routine NPDES compliance reports. Using the control data, the Sanitation Districts staff conducted statistical analyses for the acute toxicity data included in the Staff Report and found that no statistically significant differences were observed between the control and undiluted receiving water samples.

The Water Quality Objective/Threshold for Chronic Toxicity Should Be a Monthly Median.

When a sample fails a chronic toxicity test, additional samples may be collected during the same calendar month in an effort to confirm the result and identify potential toxicants. Application of a single test failure chronic toxicity water quality objective provides a disincentive for this type of proactive monitoring and is at odds with the intent of chronic toxicity testing.

Chronic toxicity is intended to assess potential aquatic life impacts associated with long-term exposures. Therefore, it is analogous to the Criterion Continuous Concentration (CCC) used for estimating "safe" chemical concentrations for long-term exposure as opposed to the Criterion Maximum Concentration (CMC) that is intended to protect against short-term exposure. In EPA's Region 9 and 10 Toxicity Training Tool, the CCC is defined as "the highest in-stream concentration of a toxic or an effluent to which organisms can be exposed indefinitely without causing unacceptable effects such as the exceedance of a chronic water quality criterion."[1] This same document also recommends "direct application of 1.0 TUc as the monthly compliance level for NPDES discharges without a mixing zone or dilution allowance. In conjunction and limited to this discharge situation, because: (1) there are no values below 1.0 TUc and (2) an arithmetic average is sensitive to extremely large and small values, the median is favored as the better measure of central tendency for the monthly compliance level."

[1] Denton DL, Miller JM, Stuber RA. 2007. EPA Regions 9 and 10 toxicity training tool (TTT). November 2007. San Francisco, CA: https://www.epa.gov/sites/production/files/documents/ToxTrainingTool10Jan2010.pdf

Fact Sheet #4

Although the Basin Plan provides no numeric chronic toxicity objectives, recently adopted Los Angeles Region NPDES permits do provide very specific direction on interpretation of the narrative water quality objectives for chronic toxicity. In the Whittier Narrows WRP NPDES permit, footnote 22 of Table E-5 of the Monitoring and Reporting Program states; "The **median monthly summary result is a threshold value** for a determination of meeting the narrative receiving water objective and shall be reported as 'Pass' or 'Fail.'"[2] [Emphasis added.]

In addition to aligning with the NPDES permit language, use of a monthly median will also address concerns regarding false positive error rates. The USEPA has determined that the expected false positive error rate for chronic toxicity testing using the NOEC is 5%. With this error rate, on average, one in 20 individual chronic toxicity tests will be erroneously identified as "toxic" using the NOEC, and there is a nearly 34% probability that 2 or more individual chronic toxicity test exceedances would be observed within a set of 24 discrete measurements in a completely non-toxic stream reach. When there are two or more exceedances out of 24 measurements, the Listing Policy specifies that a reach be listed as impaired. Therefore, using single chronic toxicity exceedances as the 303(d) criterion would, over time, result in more and more non-toxic stream reaches being erroneously listed. However, using a monthly median chronic toxicity exceedance threshold would reduce the likelihood of inappropriate reach listings due to false positive chronic toxicity results to less than 1%.

[2] Pomona WRP - ORDER R4-2014-0212-A01 NPDES NO. CA0053619, Long Beach WRP - ORDER NO. R4-2015-0123 NPDES NO. CA0054119, Los Coyotes - ORDER NO. R4-2015-0124 NPDES NO. CA0054011, San Jose Creek WRP - ORDER R4-2015-0070 NPDES NO. CA0053911, Saugus WRP- ORDER R4-2015-0072 NPDES NO. CA0054313, Valencia WRP- ORDER R4-2015-0071 NPDES NO. CA0054216, Whittier Narrows WRP ORDER R4-2014-0213-A01 NPDES NO. CA0053716

Fact Sheet #5

Water Body:	**Santa Clara River Reach 5**
Pollutant:	**Toxicity**
Listing:	**List on 303(d) List (TMDL Required List)**
Comment & Recommendation:	**Do Not List – Water Quality Objectives Being Achieved**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for toxicity be made to the 303(d) list for Reach 5 of the Santa Clara River, based on one line of evidence: 2 of 2 samples exceeded the objective. The Sanitation Districts of Los Angeles County (Sanitation Districts) believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved, for the reasons listed below; supporting evidence is provided in the sections that follow.

- Inappropriate data were utilized. Toxicity results were reported for sites SCR 1272 and SCR 14156. However, SCR 14156 is in Reach 6 of the Santa Clara River and should not be included in an evaluation of Reach 5 (Figure 1).

- Incomplete data were utilized. The ”Data for Various Pollutants in Various Water Bodies in Sanitation Districts of Los Angeles County 2005-2010” dataset should be included in this analysis as it was provided in response to the call for data, readily available, and used in other current listing recommendations. Appendix A of this letter contains the full set of data applicable to this listing from Appendix G of the Regional Board Draft Staff Report.

- The Los Angeles Region Basin Plan states, “the acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board.” Therefore, a single-test threshold of less than 70% survival should be used to determine impairments. Applying this threshold (or even a more conservative 90% threshold) to the appropriate and complete dataset that excludes site SCR 14156 and includes Sanitation Districts data, there were five chronic toxicity exceedances out of 90 valid toxicity tests. This total does not meet the minimum number of measured exceedances needed to place a water segment on the section 303(d) list.

Fact Sheet #5

Figure 1. Santa Clara River Reach 5 and RWB4 Stormwater Monitoring Council CY2008 CY2009 Sampling Locations

Fact Sheet #5

The Los Angeles Region Basin Plan Establishes Acute Toxicity Thresholds

The Los Angeles Region Basin Plan states, "the acute toxicity objective for discharges dictates that the average survival in undiluted effluent for any three consecutive 96-hour static or continuous flow bioassay tests shall be at least 90%, with no single test having less than 70% survival when using an established USEPA, State Board, or other protocol authorized by the Regional Board." Therefore, a single-test threshold of less than 70% survival should be used to determine impairments. However, even if a more conservative 90% survival single-test threshold were to be applied, no tests in the data set would exceed this threshold.

The Water Quality Objective/Threshold for Chronic Toxicity Should Be a Monthly Median.

When a sample fails a chronic toxicity test, additional samples may be collected during the same calendar month in an effort to confirm the result and identify potential toxicants. Application of a single test failure chronic toxicity water quality objective provides a disincentive for this type of proactive monitoring and is at odds with the intent of chronic toxicity testing.

Chronic toxicity is intended to assess potential aquatic life impacts associated with long-term exposures. Therefore, it is analogous to the Criterion Continuous Concentration (CCC) used for estimating "safe" chemical concentrations for long-term exposure as opposed to the Criterion Maximum Concentration (CMC) that is intended to protect against short-term exposure. In EPA's Region 9 and 10 Toxicity Training Tool, the CCC is defined as "the highest in-stream concentration of a toxic or an effluent to which organisms can be exposed indefinitely without causing unacceptable effects such as the exceedance of a chronic water quality criterion."[1] This same document also recommends "direct application of 1.0 TUc as the monthly compliance level for NPDES discharges without a mixing zone or dilution allowance. In conjunction and limited to this discharge situation, because: (1) there are no values below 1.0 TUc and (2) an arithmetic average is sensitive to extremely large and small values, the median is favored as the better measure of central tendency for the monthly compliance level."

Although the Basin Plan provides no numeric chronic toxicity objectives, recently adopted Los Angeles Region NPDES permits do provide very specific direction on interpretation of the narrative water quality objectives for chronic toxicity. In the Valencia WRP permit, footnote 30 associated with the Receiving Water Monitoring Requirements Table (Table E-5a) of the Monitoring and Reporting Program states; "The **median monthly summary result is a threshold value** for a determination of meeting the narrative receiving water objective and shall be reported as 'Pass' or 'Fail.'"[2] [Emphasis added.]

In addition to aligning with the NPDES permit language, use of a monthly median will also address concerns regarding false positive error rates. The USEPA has determined that the expected false positive error rate for chronic toxicity testing using the NOEC is 5%. With this error rate, on average, one in 20 individual chronic toxicity tests will be erroneously identified as "toxic" using the NOEC, and there is a nearly 34% probability that 2 or more individual chronic toxicity test exceedances would be observed within a set of 24 discrete measurements in a completely non-toxic stream reach. When there are two or

---

[1] Denton DL, Miller JM, Stuber RA. 2007. EPA Regions 9 and 10 toxicity training tool (TTT). November 2007. San Francisco, CA: https://www.epa.gov/sites/production/files/documents/ToxTrainingTool10Jan2010.pdf

[2] Pomona WRP - ORDER R4-2014-0212-A01 NPDES NO. CA0053619, Long Beach WRP - ORDER NO. R4-2015-0123 NPDES NO. CA0054119, Los Coyotes - ORDER NO. R4-2015-0124 NPDES NO. CA0054011, San Jose Creek WRP - ORDER R4-2015-0070 NPDES NO. CA0053911, Saugus WRP- ORDER R4-2015-0072 NPDES NO. CA0054313, Valencia WRP- ORDER R4-2015-0071 NPDES NO. CA0054216, Whittier Narrows WRP ORDER R4-2014-0213-A01 NPDES NO. CA0053716

Fact Sheet 5

more exceedances out of 24 measurements, the Listing Policy specifies that a reach be listed as impaired. Therefore, using single chronic toxicity exceedances as the 303(d) criterion would, over time, result in more and more non-toxic stream reaches being erroneously listed. However, using a monthly median chronic toxicity exceedance threshold would reduce the likelihood of inappropriate reach listings due to false positive chronic toxicity results to less than 1%.

Fact Sheet #6

Water Body: **Santa Clara River Reach 5**
Pollutant: **Benthic Community Effects**
Listing: **List on 303(d) List (TMDL Required List)**
Comment & Recommendation: **Do Not List – Water Quality Objectives Being Achieved**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is currently proposing that a new listing for benthic community effects be made to the 303(d) list for Reach 5 of the Santa Clara River, based on two lines of evidence: Southern Coastal California Index of Biotic integrity (SCIBI) and California Stream Condition Index (CSCI) scores. The Districts believe this proposed listing is inappropriate and recommend not listing for the reasons listed below; supporting evidence is provided in the sections that follow.

- The SCIBI-based analysis has been demonstrated to be inadequate for use in low gradient/low elevation watersheds similar to the reaches in the upper Santa Clara River. For this and other reasons, the State Water Resources Control Board (State Board) has rejected use of the SCIBI in favor of the technically superior CSCI scoring tool.

- Although the CSCI at least partially addresses some of the problems with the SCIBI by employing a modeled reference condition as opposed to the regional reference pool used by the SCIBI, the lack of any reference sites in large watersheds, low gradient, and low elevation systems still limits the identification of appropriate thresholds using the CSCI. Specifically, several Santa Clara River sites have been shown to fall outside the experience of the CSCI model.

- Appropriate water quality thresholds for the CSCI have not been established. Although examples of approaches for developing CSCI thresholds have been published (e.g., by the Southern California Coastal Water Research Project), it is well recognized by the scientific community that a single standard should not be applicable to all water bodies because unmanageable non-pollutant features such as habitat condition are likely to preclude many streams from ever having biological assemblages similar to reference. The State Board is currently investing considerable resources to develop thresholds and should be allowed to complete the process before determination of impairment and listings.

- The CSCI analysis for this listing used data from both Reach 5 and Reach 6 of the Santa Clara River. The <u>CSCI analysis of the data collected from the Reach 5 location actually met the 0.79 threshold proposed by the Regional Board.</u>

- Physical habitat was not assessed, as required by the State Board Water Quality Control Board Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (Listing Policy). Historically unmanaged or unmanageable stressors (e.g. channel/habitat modifications) are well documented as precluding sites from achieving reference conditions.

- The proposed listing fails to associate the alleged impairment with other pollutants, namely toxicity and iron, which were listed as co-occurring.

Fact Sheet #6

SCIBI Is an Inadequate Metric for Assessing Low Gradient, Low Elevation Streams.

Section 3.9 of the Listing Policy states:

> "A water segment shall be placed on the section 303(d) list if the water segment exhibits significant degradation in biological populations and/or communities **as compared to reference site(s)** and is associated with water or sediment concentrations of pollutants including but not limited to chemical concentrations, temperature, dissolved oxygen, and trash." [Emphasis added.]

While it is commonly assumed that the SCIBI inherently accounted for reference conditions, the reference conditions used to develop the SCIBI were not representative of the low elevation/low gradient streams commonly found in the alluvial plains of the Los Angeles Region.[1,2] It was developed using data from 275 sites, ranging from Monterey County to the Mexican border but not a single reference location represented low elevation and low gradient streams. Santa Clara River Reach 5 is an extremely low gradient (less than 0.5%), low elevation, large coastal water body; therefore, the reference pool used for development of the SCIBI is not representative of natural conditions relevant to this reach. As described in more detail below, technical experts have acknowledged the limitations of the SCIBI (and other IBIs) and indicated that it is critical that reference conditions represent the full range of environmental gradients where an index will be used.[3] Consequently, the State Water Board has supported and funded the development of the CSCI scoring tool; this new, predictive index represents a substantial increase in the applicability of indices.

The lead scientist for development of the SCIBI, Dr. Peter Ode, has acknowledged the limitations on application of the SCIBI. In a peer-reviewed published paper, he concluded that the SCIBI did not adequately address reference conditions in low elevation sites and was "not completely effective at controlling for an elevation gradient."[4] Dr. Ode was also the co-author of a March 2009 report on recommendations for development and maintenance of a network of reference sites to support biological assessment of California's wadeable streams, which notes that, "A crucial component to the development of assessment tools is understanding biological expectations at reference sites that consist of natural, undisturbed systems.[5] These reference systems set the biological condition benchmarks for comparisons to the site(s) being evaluated." They also clearly note that adequate reference sites have not been identified in southern California, stating, "human-dominated landscapes can be so pervasive in locations such as urban southern California and the agriculturally dominated Central Valley that no undisturbed reference sites may currently exist in these regions. A statewide framework for consistent selection of reference sites must account for this complexity."

---

[1] Ode, P.R., A.C. Rehn, J.T. May. 2005. A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams. Environmental Management Vol. 35, No 4, pp. 494, Figure 1. Copy included in Appendix B.

[2] Carter, J.L. and V.H. Resh. (2005). Pacific Coast Rivers of the Coterminous United States. pp. 541-590 *in*: A.C. Benke and C.E. Cushing (eds.), Rivers of North America. Elsevier Academic Press. Boston, MA.

[3] Mazor, R.D., A.C. Rehn, P.R. Ode, M. Engeln, K.C. Schiff, E.D. Stein, D.J. Gillett, D.B. Herbst, and C.P. Hawkins. (2016). Bioassessment in complex environments: Designing an index for consistent meaning in different settings. Freshwater Science 35(1): 249-271. Copy included in Appendix B.

[4] Ode, P.R., C.P. Hawkins, R.D. Mazor, Comparability of Biological Assessments Derived from Predictive Models and Multimetric Indices of Increasing Geographic Scope, J. N. Am. Benthol. Soc., 2008, 27(4):967-985.p. 982. Copy included in Appendix B.

[5] Ode, P.R., K. Schiff. Recommendations for the Development and Maintenance of a Reference Condition Management Program to Support Biological Assessment of California's Wadeable Streams: Report to the Surface Water Ambient Monitoring Program. Southern California Coastal Water Research Project, Technical Report 581. March 2009. Copy included in Appendix B.

Fact Sheet #6

Furthermore, a memorandum prepared by Jerry Diamond of Tetra Tech, one of the leading national technical experts on bioassessments, confirmed that adequate reference sites are not available to assess benthic macroinvertebrate populations for low gradient and low elevation streams in the LA Region.[6] Dr. Diamond is the author of several technical reports prepared for the LA Regional Board on tiered aquatic life uses (TALU) based on bioassessments.[7,8] Dr. Diamond states that there is "high uncertainty regarding appropriate reference conditions for low gradient and low elevation streams in this region [Southern California]," and that "low elevation streams lacked a clear reference conditions in this region [Southern California]." He further states that a technical advisory committee for a US EPA-funded project on TALU "identified a lack of appropriate reference sites for low elevation/low gradient streams as a critical data gap." The technical advisory committee consisted of regional experts from California Fish & Wildlife, State Water Board, other Regional Boards, US EPA Region 9, and universities. Dr. Diamond also worked with SCCWRP and the LA Regional Board in facilitating two workshops on TALU for Southern California. Dr. Diamond states, "In the most recent stakeholder workshop…there was agreement that low gradient (rather than low elevation) was perhaps the most critical factor distinguishing stream biology in the region and that the reference condition for low gradient streams (many but not all of which occur at low elevation) is a critical data gap...."[9]

Other scientific experts concur with Dr. Diamond's conclusions. As part of a 2009 study examining low gradient streams in California, including sites within Reach 5 of the Santa Clara River, Raphael D. Mazor of SCCWRP stated, "Several biomonitoring efforts in California specifically target low-gradient streams, as these habitats are subject to numerous impacts and alterations…even though the applicability of assessment tools created and validated in high-gradient streams have not been tested."[10] The study found that "As a consequence of these differences [substrate material, bed morphology, and distribution of microhabitats], traditional bioassessment approaches in California that were developed in high-gradient streams with diverse microhabitats have limited applications in low-gradient reaches,"[10] and "Caution should be used when applying sampling methods for assessment tools that were calibrated for specific habitat types (e.g., high gradient streams) to new habitats (e.g., low gradient streams)."[10] The study also concluded that "…observation of the sites in this study suggests that the lack of stable microhabitats (e.g., riffles and vegetated margins) may account for the reduced number of macroinvertebrates, as few species are adapted to the shifting sandy substrate found in most low gradient streams in California."[10] Moreover, the State Water Board, Surface Water Ambient Monitoring Program, California Department of Fish and Game, and others recognize the limitations of the SCIBI regarding reference sites. They have identified application of TALU and the selection of more representative/appropriate regional reference locations as being necessary components to the state's bioassessment program.[5,7] This sentiment was shared in an evaluation of California's bioassessment program. Specifically, "The National Research Council's review makes clear that all states need better biological endpoints, adequate monitoring and assessment, and

---

[6] Diamond, Jerry. Reference Conditions and Bioassessments in Southern California Streams. July 31, 2009. Memorandum to Phil Markle of the Sanitation Districts. Copy included in Appendix B.

[7] Schiff, K. and Diamond, J., Identifying Barriers to Tiered Aquatic Life Uses (TALU) in Southern California, Southern California Coastal Water Research Project, Technical Report 590. June 2009. Copy included in Appendix B.

[8] Tetra Tech, Revised Analyses of Biological Data to Evaluate Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams. Prepared for EPA Region 9 and California Regional Water Quality Control Board, Los Angeles Region. 2006. Tetra Tech, Inc., Owings Mills, MD. Copy included in Appendix B.

[9] For a report summarizing the outcome of the workshops, see Schiff, K. and Diamond, J., Identifying Barriers to Tiered Aquatic Life Uses (TALU) in Southern California, Southern California Coastal Water Research Project, Technical Report 590. June 2009. Copy included in Appendix B.

[10] Mazor, Raphael D.; Schiff, Kenneth; Ritter, Kerry; Rehn, Andy; and Ode, Peter; Bioassessment Tools in Novel Habitats: An Evaluation of Indices and Sampling Methods in Low-Gradient Streams in California, Environ. Monit. Assess., DOI 10.1007/s10661-009-1033-3. Copy included in Appendix B.

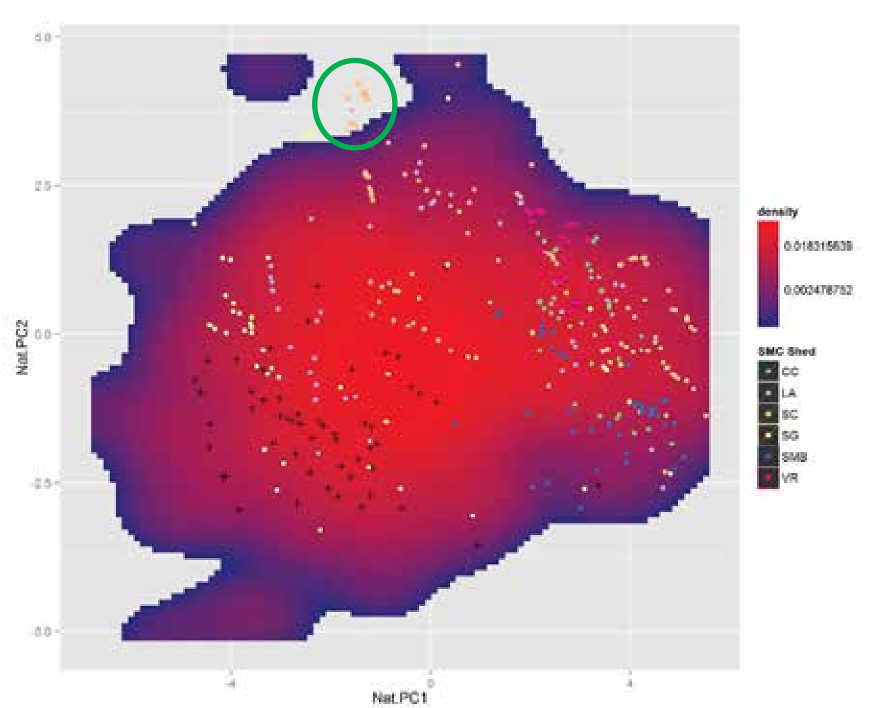
Fact Sheet #6

tiered aquatic life uses (TALU) in order to develop and refine appropriate and effective water quality standards that result in more accurate and appropriate protection for biological resources."[11]

CSCI Improves on the SCIBI But Some Limitations Remain

The State Board is developing the CSCI scoring tool that is intended to replace the flawed IBI scoring tools statewide. The CSCI at least partially addresses some of the problems with the SCIBI by employing a modeled reference condition as opposed to the regional reference pool used by the SCIBI. Reliance upon this modeled reference condition has significantly improved the applicability and resolution of the bioassessment scoring tools; however, the lack of any reference sites in large watersheds, low gradient, and low elevation systems still limits the identification of appropriate thresholds using the CSCI. A number of these environmental gradients exist, alone or in combination. Figure 1 shows the use of a data density approach to quantify the availability of data to determine reference conditions; red areas indicate a higher density of reference locations, darker/blue areas indicate fewer reference locations, and gray indicates sites that may be outside the experience of the CSCI.[12] Several of the Santa Clara River sites (orange symbols circled in Figure 1) fall outside of CSCI reference conditions. In these situations, it has been suggested that the CSCI could be used in conjunction with an alternative (i.e., non-threshold based) assessment option (i.e., upstream-downstream comparison).[13]

Figure 1. CSCI Reference Density Cloud (Santa Clara River Sites Within Green Circle).



[11] Yoder, C.O. and Plotnikoff, R. (2009). Evaluation of the California State Water Resource Control Board's Bioassessment Program. Final Report to EPA-OST and Region IX. Copy included in Appendix B.

[12] Ode, P.R., Rehn, A.C., Mazor, R.D., and Schiff, K.C. (2012) Building the Technical Foundation for Biological Objectives. Presentation to the California Aquatic Bioassessment Workgroup, November 7, 2012. Copy included in Appendix B.

[13] California Biological Objectives Science Advisory Panel. (2012). Science Advisory Panel Response, October 18, 2012. Copy included in Appendix B.

Fact Sheet #6

Appropriate Water Quality Standards (i.e. Biocriteria) Have Not Been Established

The State Board is determining numeric translators to address the narrative biological objective that includes "bioassessment." However, the State Board has not yet developed any recommended thresholds for the CSCI. The proposed threshold of 0.79 is the 10th percentile of the reference pool and was used as a point of reference for a regional monitoring program. However, by this definition, 10% of California's identified reference pool would be considered impaired.

Furthermore, it is well recognized by the scientific community that a single standard or threshold should not be applicable to all waterbodies of the State since unmanageable non-pollutant features such as habitat condition are likely to preclude many streams from ever having biological assemblages similar to reference.[14,15] In fact, as part of California's biological objectives program, SCCWRP has developed a workplan to identify and evaluate the constraint these traditionally unmanageable features may place on biological indices.[16] For example, the Southern California Stormwater Monitoring Coalition (SMC) found that engineered (i.e. modified) channels appear to be in worse ecological health than natural channels based on macroinvertebrate and algae assemblages and that tradeoffs between ecological health and flood protection may be unavoidable.[15] This impact of unmanageable stressors is not limited to engineered channels; studies have shown that hydrological alterations attributable to reaches with as little as 5% coverage by impervious surfaces in the localized watershed are associated with unhealthy biological communities.[17] These and many other examples clearly illustrate the infeasibility of a single criterion with statewide applicability.

As such, utilization of an undeveloped and unsupported standard (e.g. CSCI <0.79) is premature. Given the substantial resources the State is investing in the development of numerical translators, Regional Boards should allow the State to complete the process before determination of impairment and listings, as appropriate.

CSCI Data from Within Reach 5 of the Santa Clara River Show No Impairment

The proposed listing cites one dataset for the CSCI. This dataset inappropriately aggregates two stations (SCR 14156 and SCR 1272) that are approximately 5 miles from each other (Figure 2). Section 6.1.5.2 of the Listing Policy states:

> "Samples collected within 200 meters of each other should be considered samples from the same station or location. However, samples less than 200 meters apart may be considered to be spatially independent samples if justified in the water body fact sheet."

These two stations are too far apart to justify aggregation. Furthermore, SCR 14156 is in Reach 6 of the Santa Clara River and should not be considered as a line of evidence in any proposed Reach 5 listing. The single station with Reach 5, SCR 1272, had a CSCI score of 0.91. **Thus, the only CSCI score in this Reach is above the proposed threshold of impairment**.

---

[14] Waite, I.R. J.G. Kennen, J.T. May, L.R. Brown, T.F. Cuffney, K.A. Jones, and J.L. Orlando. (2012). Comparison of stream invertebrate response models for bioassessment metrics. Journal of the American Water Resources Association 48. Copy included in Appendix B.

[15] Southern California Stormwater Monitoring Coalition (SMC). 2017. 2015 Report on the Stormwater Monitoring Coalition Regional Stream Survey. SCCWRP Technical Report 963. Southern California Coastal Water Research Project. Costa Mesa, CA. Copy included in Appendix B.

[16] Mazor, R., Sutula, M., Stein, E., Rehn, A., and Ode, P. (2017) Draft Work Plan. Predicting Biological Integrity of Streams Across a Gradient of Development in California Landscapes. Copy included in Appendix B.

[17] Stein, E.D., Sengupta, A., Mazor, R.D., and McCune, K. (2016). Application of Regional Flow-ecology to Inform Management Decision in the San Diego River Watershed. Southern California Coastal Water Research Project Technical Report #948. Copy included in Appendix B.

Fact Sheet #6

Figure 2. Santa Clara River Reach 5 and Monitoring Stations Used in Listing

Fact Sheet #6

The Proposed Listing Fails to Evaluate Physical Habitat Data

Section 6.1.5.8 of the listing policy states:

> "When evaluating biological data and information, RWQCBs shall evaluate all readily available data and information and shall
> - Evaluate bioassessment data from other sites, and compare to reference condition.
> - Evaluate physical habitat data and other water quality data, when available, to support conclusions about the status of the water segment."

EPA's causal assessment manual cites physical habitat as a leading cause of impairment in streams on 303d lists and recommends that, in all cases where physical habitat is evaluated, stream size and channel dimensions, channel gradient, channel substrate size and type, habitat complexity and cover, vegetation cover and structure, and channel-riparian interactions should all be considered before making a decision.[18] Likewise, the SMC identified habitat stressors among the highest priority for evaluation in relation to depressed benthic community assemblages.[15] The need to consider physical habitat is apparent in the low gradient Santa Clara River where sediment and leaf litter/detritus loads are naturally deposited in the channel, filling up the available spaces between rocks. These habitats support a much different population of invertebrates (more detritus feeders and fewer predators) than the rocky/sandy reference conditions, and do not necessarily indicate an "impaired" population.

The Proposed Listing Fails to Associate the Alleged Impairment with Other Pollutants

The Listing Policy states:

> "A water segment shall be placed on the section 303(d) list if the water segment exhibits significant degradation in biological populations and/or communities as compared to reference site(s) and **is associated with** water or sediment concentrations of pollutants including but not limited to chemical concentrations, temperature, dissolved oxygen, and trash." [Emphasis added.]

In the fact sheets supporting its impairment decisions for each of these listings, the LA Regional Board stated that the alleged impairment in benthic community composition in Reach 5 was justified by being "associated" with impairments for two pollutants, iron and toxicity simply because these constituents co-occurred. However, based on further investigations, it is apparent that these constituents would not be associated with benthic community impairment because the iron would not be bioavailable and no impairment exists for toxicity.

- Iron
  - The 1.0 ppm iron criterion used as the basis for the proposed iron impairment in this reach is a 4-day average threshold taken from the 1976 USEPA "Red Book" and was updated using the 1985 Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. However, iron was detected only sporadically at levels above 1.0 ppm, and concentrations below the point source discharge were consistently low, suggesting that the 4-day average threshold of 1.0 mg/L is likely achieved.

---

[18] U.S. EPA (Environmental Protection Agency). (2010). Causal Analysis/Diagnosis Decision Information System (CADDIS). Office of Research and Development, Washington, DC. Available online at https://www.epa.gov/caddis. Last updated September 23, 2010

Fact Sheet #6

- Furthermore, the bioavailable form of iron is ferrous iron and only exists at low pH levels. The pH in Reach 5 averages 7.9 with a 5th percentile pH of 7.5. In ambient waters with sufficient dissolved oxygen and a pH above 7.0, iron will rapidly oxidize to a non-bioavailable form and would not be responsible for impacts to aquatic life.In fact, the Red Book includes a disclaimer that "data obtained under laboratory conditions suggest a greater toxicity for iron than that obtained in natural ecosystems."

- Toxicity
  - SCCWRP has concluded that sub-lethal water column toxicity is a poor indicator of benthic community impairment.[19] Furthermore, the data do not support a toxicity listing (Fact Sheet #5). Station SCR 14156 is in Reach 6 of the Santa Clara River and should not be included in this analysis. Conversely, the readily accessible ”Data for Various Pollutants in Various Water Bodies in Sanitation Districts of Los Angeles County 2005-2010” dataset should be included in this analysis. Using the complete and appropriate dataset, six of 91 Santa Clara River Reach 5 toxicity tests exceed the objective, which fails to meet the minimum number of measured exceedances needed to place a water segment on the section 303(d) list for toxicants as specified in table 3.1 of the Listing Policy.

[19] Southern California Stormawater Monitoring Coalition (SMC). 2015. Bioassessment of Perennial Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition’s Regional Stream Survey. SCCWRP Technical Report 844. Southern California Coastal Water Research Project. Costa Mesa, CA. Copy included in Appendix B.

Fact Sheet #7

Water Body:	**Los Angeles River Reach 3**
Pollutant:	**Benthic Community Effects**
Listing:	**List on 303(d) List (TMDL Required List)**
Comment & Recommendation:	**Do Not List – Water Quality Objectives Being Achieved**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for benthic community effects be made to the 303(d) list for Reach 3 of the Los Angeles River, based on a weight of evidence approach using Southern Coastal California Index of Biotic integrity (SCIBI) scores. The Districts believe this proposed listing is inappropriate and recommend not listing for the reasons listed below; supporting evidence is provided in the sections that follow.

· The SCIBI-based analysis has been demonstrated to be inadequate for use in low gradient/low elevation watersheds similar to Los Angeler River Reach 3. For this, and other reasons, the State Water Resources Control Board (State Board) has rejected use of the SCIBI in favor of the technically superior CSCI scoring tool. No CSCI results have been used for this listing, but a more detailed assessment of the CSCI can be found in Fact Sheet #6.

· Physical habitat was not assessed, as required by the State Board Water Quality Control Board Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (Listing Policy). Historically unmanaged or unmanageable stressors (e.g. channel/habitat modifications) are well documented as precluding sites from achieving reference conditions.

SCIBI Is an Inadequate Metric for Assessing Low Gradient, Low Elevation Streams.

Section 3.9 of the Listing Policy states:

> "A water segment shall be placed on the section 303(d) list if the water segment exhibits significant degradation in biological populations and/or communities **as compared to reference site(s)** and is associated with water or sediment concentrations of pollutants including but not limited to chemical concentrations, temperature, dissolved oxygen, and trash." [Emphasis added.]

While it is commonly assumed that the SCIBI inherently accounted for reference conditions, the reference conditions used to develop the SCIBI were not representative of the low elevation/low gradient streams commonly found in the alluvial plains of the Los Angeles Region.[1,2] It was developed using data from 275 sites, ranging from Monterey County to the Mexican border but not a single reference location represented low elevation and low gradient streams. Los Angeles River Reach 3 is a low gradient, low elevation, large coastal water body; therefore, the reference pool used for development of the SCIBI is not representative of natural conditions relevant to this reach. As described in more detail below, technical experts have acknowledged the limitations of the SCIBI (and other IBIs) and indicated that it is critical that reference conditions represent the full range of environmental gradients where an index will be used.[3]

---

[1] Ode, P.R., A.C. Rehn, J.T. May. 2005. A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams. Environmental Management Vol. 35, No 4, pp. 494, Figure 1. Copy included in Appendix B.
[2] Carter, J.L. and V.H. Resh. (2005). Pacific Coast Rivers of the Coterminous United States. pp. 541-590 *in*: A.C. Benke and C.E. Cushing (eds.), Rivers of North America. Elsevier Academic Press. Boston, MA.
[3] Mazor, R.D., A.C. Rehn, P.R. Ode, M. Engeln, K.C. Schiff, E.D. Stein, D.J. Gillett, D.B. Herbst, and C.P. Hawkins. (2016). Bioassessment in complex environments: Designing an index for consistent meaning in different settings. Freshwater Science 35(1): 249-271. Copy included in Appendix B.

Fact Sheet #7

Consequently, the State Water Board has supported and funded the development of the CSCI scoring tool; this new, predictive index represents a substantial increase in the applicability of indices.

The lead scientist for development of the SCIBI, Dr. Peter Ode, has acknowledged the limitations on application of the SCIBI. In a peer-reviewed published paper, he concluded that the SCIBI did not adequately address reference conditions in low elevation sites and was "not completely effective at controlling for an elevation gradient."[4] Dr. Ode was also the co-author of a March 2009 report on recommendations for development and maintenance of a network of reference sites to support biological assessment of California's wadeable streams, which notes that, "A crucial component to the development of assessment tools is understanding biological expectations at reference sites that consist of natural, undisturbed systems".[5] These reference systems set the biological condition benchmarks for comparisons to the site(s) being evaluated." They also clearly note that adequate reference sites have not been identified in southern California, stating, "human-dominated landscapes can be so pervasive in locations such as urban southern California and the agriculturally dominated Central Valley that no undisturbed reference sites may currently exist in these regions. A statewide framework for consistent selection of reference sites must account for this complexity."

Furthermore, a memorandum prepared by Jerry Diamond of Tetra Tech, one of the leading national technical experts on bioassessments, confirmed that adequate reference sites are not available to assess benthic macroinvertebrate populations for low gradient and low elevation streams in the LA Region.[6] Dr. Diamond is the author of several technical reports prepared for the LA Regional Board on tiered aquatic life uses (TALU) based on bioassessments.[7,8] Dr. Diamond states that there is "high uncertainty regarding appropriate reference conditions for low gradient and low elevation streams in this region [Southern California]," and that "low elevation streams lacked a clear reference conditions in this region [Southern California]." He further states that a technical advisory committee for a US EPA-funded project on TALU "identified a lack of appropriate reference sites for low elevation/low gradient streams as a critical data gap." The technical advisory committee consisted of regional experts from California Fish & Wildlife, State Water Board, other Regional Boards, US EPA Region 9, and universities. Dr. Diamond also worked with SCCWRP and the LA Regional Board in facilitating two workshops on TALU for Southern California. Dr. Diamond states, "In the most recent stakeholder workshop…there was agreement that low gradient (rather than low elevation) was perhaps the most critical factor distinguishing stream biology in the region and that the reference condition for low gradient streams (many but not all of which occur at low elevation) is a critical data gap...."[9]

---

[4] Ode, P.R., C.P. Hawkins, R.D. Mazor, Comparability of Biological Assessments Derived from Predictive Models and Multimetric Indices of Increasing Geographic Scope, J. N. Am. Benthol. Soc., 2008, 27(4):967-985.p. 982. Copy included in Appendix B.

[5]Ode, P.R., K. Schiff. Recommendations for the Development and Maintenance of a Reference Condition Management Program to Support Biological Assessment of California's Wadeable Streams: Report to the Surface Water Ambient Monitoring Program. Southern California Coastal Water Research Project, Technical Report 581. March 2009. Copy included in Appendix B.

[6] Diamond, Jerry. Reference Conditions and Bioassessments in Southern California Streams. July 31, 2009. Memorandum to Phil Markle of the Sanitation Districts. Copy included in Appendix B.

[7]Schiff, K. and Diamond, J., Identifying Barriers to Tiered Aquatic Life Uses (TALU) in Southern California, Southern California Coastal Water Research Project, Technical Report 590. June 2009. Copy included in Appendix B.

[8] Tetra Tech, Revised Analyses of Biological Data to Evaluate Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams. Prepared for EPA Region 9 and California Regional Water Quality Control Board, Los Angeles Region. 2006. Tetra Tech, Inc., Owings Mills, MD. Copy included in Appendix B.

[9]For a report summarizing the outcome of the workshops, see Schiff, K. and Diamond, J., Identifying Barriers to Tiered Aquatic Life Uses (TALU) in Southern California, Southern California Coastal Water Research Project, Technical Report 590. June 2009. Copy included in Appendix B.

Fact Sheet #7

Other scientific experts concur with Dr. Diamond's conclusions. As part of a 2009 study examining low gradient streams in California, Raphael D. Mazor of SCCWRP stated, "Several biomonitoring efforts in California specifically target low-gradient streams, as these habitats are subject to numerous impacts and alterations…even though the applicability of assessment tools created and validated in high-gradient streams have not been tested."[10] The study found that "As a consequence of these differences [substrate material, bed morphology, and distribution of microhabitats], traditional bioassessment approaches in California that were developed in high-gradient streams with diverse microhabitats have limited applications in low-gradient reaches,"[10] and "Caution should be used when applying sampling methods for assessment tools that were calibrated for specific habitat types (e.g., high gradient streams) to new habitats (e.g., low gradient streams)."[10] The study also concluded that "…observation of the sites in this study suggests that the lack of stable microhabitats (e.g., riffles and vegetated margins) may account for the reduced number of macroinvertebrates, as few species are adapted to the shifting sandy substrate found in most low gradient streams in California."[10] Moreover, the State Water Board, Surface Water Ambient Monitoring Program, California Department of Fish and Game, and others recognize the limitations of the SCIBI regarding reference sites. They have identified application of TALU and the selection of more representative/appropriate regional reference locations as being necessary components to the state's bioassessment program.[5,7] This sentiment was shared in an evaluation of California's bioassessment program. Specifically, "The National Research Council's review makes clear that all states need better biological endpoints, adequate monitoring and assessment, and tiered aquatic life uses (TALU) in order to develop and refine appropriate and effective water quality standards that result in more accurate and appropriate protection for biological resources."[11]

CSCI Improves on the SCIBI But Some Limitations Remain

The State Board is developing the CSCI scoring tool that is intended to replace the flawed IBI scoring tools statewide. The CSCI at least partially addresses some of the problems with the SCIBI by employing a modeled reference condition as opposed to the regional reference pool used by the SCIBI. Reliance upon this modeled reference condition has significantly improved the applicability and resolution of the bioassessment scoring tools; however, the lack of any reference sites in large watersheds, low gradient, and low elevation systems still limits the identification of appropriate thresholds using the CSCI. A number of these environmental gradients exist, alone or in combination. In these situations, it has been suggested that the CSCI could be used in conjunction with an alternative (i.e., non-threshold based) assessment option (i.e., upstream-downstream comparison).[12]

The Proposed Listing Fails to Evaluate Physical Habitat Data

Section 6.1.5.8 of the listing policy states:

> "When evaluating biological data and information, RWQCBs shall evaluate all readily available data and information and shall
>
> - Evaluate bioassessment data from other sites, and compare to reference condition.

[10] Mazor, Raphael D.; Schiff, Kenneth; Ritter, Kerry; Rehn, Andy; and Ode, Peter; Bioassessment Tools in Novel Habitats: An Evaluation of Indices and Sampling Methods in Low-Gradient Streams in California, Environ. Monit. Assess., DOI 10.1007/s10661-009-1033-3. Copy included in Appendix B.

[11] Yoder, C.O. and Plotnikoff, R. (2009). Evaluation of the California State Water Resource Control Board's Bioassessment Program. Final Report to EPA-OST and Region IX. Copy included in Appendix B.

[12] California Biological Objectives Science Advisory Panel. (2012). Science Advisory Panel Response, October 18, 2012. Copy included in Appendix B.

Fact Sheet #7

- Evaluate physical habitat data and other water quality data, when available, to support conclusions about the status of the water segment."

EPA's causal assessment manual cites physical habitat as a leading cause of impairment in streams on 303d lists and recommends that, in all cases where physical habitat is evaluated, stream size and channel dimensions, channel gradient, channel substrate size and type, habitat complexity and cover, vegetation cover and structure, and channel-riparian interactions should all be considered before making a decision.[13] Likewise, the SMC identified habitat stressors among the highest priority for evaluation in relation to depressed benthic community assemblages.[14] The need to consider physical habitat is evident in low gradient engineered channels such as the Los Angeles River Reach 3, an environment that experts agree is unlikely to have biological assemblages similar to reference regardless of water quality.

---

[13] U.S. EPA (Environmental Protection Agency). (2010). Causal Analysis/Diagnosis Decision Information System (CADDIS). Office of Research and Development, Washington, DC. Available online at https://www.epa.gov/caddis. Last updated September 23, 2010

[14] Southern California Stormawater Monitoring Coalition (SMC). 2015. Bioassessment of Perennial Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey. SCCWRP Technical Report 844. Southern California Coastal Water Research Project. Costa Mesa, CA. Copy included in Appendix B.

Fact Sheet #8

Water Body:	**Medea Creek Reach 1**
Pollutant:	**Benthic Community Effects**
Listing:	**List on 303(d) List (TMDL Required List)**
Comment & Recommendation:	**Do Not List – Water Quality Objectives Being Achieved**

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for benthic community effects be made to the 303(d) list for Reach 1 of the Medea Creek, based on a weight of evidence approach using California Stream Condition Index (CSCI) and Southern Coastal California Index of Biotic integrity (SCIBI) scores. The Districts believe this proposed listing is inappropriate and recommend not listing for the reasons listed below; supporting evidence is provided in the sections that follow.

- Appropriate water quality thresholds for the CSCI have not been established. Although examples of approaches for developing CSCI thresholds have been published (e.g., by the Southern California Coastal Water Research Project), it is well recognized by the scientific community that a single standard should not be applicable to all water bodies because unmanageable non-pollutant features such as habitat condition are likely to preclude many streams from ever having biological assemblages similar to reference. The State Board is currently investing considerable resources to develop thresholds and should be allowed to complete the process before determination of impairment and listings.

- Physical habitat was not assessed, as required by the State Board Water Quality Control Board Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (Listing Policy). Historically unmanaged or unmanageable stressors (e.g. channel/habitat modifications) are well documented as precluding sites from achieving reference conditions.

- The proposed listing fails to associate the alleged impairment with other pollutants, namely trash and selenium, which were listed as co-occurring.

Appropriate Water Quality Standards (i.e. Biocriteria) Have Not Been Established

The State Board is developing the CSCI scoring tool and numeric translators to address the narrative biological objective that includes "bioassessment." However, the State Board has not yet developed any recommended thresholds for the CSCI. The proposed threshold of 0.79 is the 10th percentile of the reference pool and was used as a point of reference for a regional monitoring program. However, by this definition, 10% of California's identified reference pool would be considered impaired.

Furthermore, it is well recognized by the scientific community that a single standard or threshold should not be applicable to all waterbodies of the State since unmanageable non-pollutant features such as habitat condition are likely to preclude many streams from ever having biological assemblages similar to reference. For example, the Southern California Stormwater Monitoring Coalition (SMC) found that engineered (i.e. modified) channels appear to be in worse ecological health than natural channels based on macroinvertebrate and algae assemblages and that tradeoffs between ecological health and flood protection may be unavoidable.[1] This impact of unmanageable stressors is not only limited to engineered channels; studies have shown that hydrological alterations attributable to reaches with as little as 5% coverage by impervious surfaces in the localized watershed is associated with unhealthy biological

[1] Southern California Stormwater Monitoring Coalition (SMC). 2017. 2015 Report on the Stormwater Monitoring Coalition Regional Stream Survey. SCCWRP Technical Report 963. Southern California Coastal Water Research Project. Costa Mesa, CA. Copy included in Appendix B.

Fact Sheet #8

communities.[2] These and many other examples clearly illustrate the infeasibility of a single criterion with statewide applicability.

As such, utilization of an undeveloped and unsupported standard (e.g. CSCI <0.79) is premature. Given the substantial resources the State is investing in the development of numerical translators, Regional Boards should allow the State to complete the process before determination of impairment and listings, as appropriate.

The Proposed Listing Fails to Evaluate Physical Habitat Data

Section 6.1.5.8 of the Listing Policy states:

> "When evaluating biological data and information, RWQCBs shall evaluate all readily available data and information and shall
> - Evaluate bioassessment data from other sites, and compare to reference condition.
> - Evaluate physical habitat data and other water quality data, when available, to support conclusions about the status of the water segment."

EPA's causal assessment manual cites physical habitat as a leading cause of impairment in streams on 303d lists and recommends that, in all cases where physical habitat is evaluated, stream size and channel dimensions, channel gradient, channel substrate size and type, habitat complexity and cover, vegetation cover and structure, and channel-riparian interactions should all be considered before making a decision.[3] Likewise, the SMC identified habitat stressors among the highest priority for evaluation in relation to depressed benthic community assemblages.[4] These stressors include features such as channel alteration, impervious surface proliferation in the watershed, and unique geological conditions. Medea Creek is impacted by at least two of these three examples. The channel is shored (Figure 1) and much of the watershed has unique geological conditions, which may impact the benthic community.[5]

---

[2] Stein, E.D., Sengupta, A., Mazor, R.D., and McCune, K. (2016). Application of Regional Flow-ecology to Inform Management Decision in the San Diego River Watershed. Southern California Coastal Water Research Project Technical Report #948. Copy included in Appendix B.

[3] U.S. EPA (Environmental Protection Agency). (2010). Causal Analysis/Diagnosis Decision Information System (CADDIS). Office of Research and Development, Washington, DC. Available online at https://www.epa.gov/caddis. Last updated September 23, 2010

[4] Southern California Stormawater Monitoring Coalition (SMC). 2015. Bioassessment of Perennial Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey. SCCWRP Technical Report 844. Southern California Coastal Water Research Project. Costa Mesa, CA. Copy included in Appendix B.

[5] U.S. EPA Region 9 (Environmental Protection Agency). (2013). Malibu Creek & Lagoon TMDL for Sedimentation and Nutrients Impacting Benthic Community, Technical Appendices. Available at https://www3.epa.gov/region9/water/tmdl/malibu/2013-07-02-malibu-creek-lagoon-tmdl-appendices.pdf**.**

Fact Sheet #8

Figure 1. Medea Creek Channel Modifications

The Proposed Listing Fails to Associate the Alleged Impairment with Other Pollutants

The Listing Policy states:

> "A water segment shall be placed on the section 303(d) list if the water segment exhibits significant degradation in biological populations and/or communities as compared to reference site(s) and **is associated with** water or sediment concentrations of pollutants including but not limited to chemical concentrations, temperature, dissolved oxygen, and trash." [Emphasis added.]

In the fact sheets supporting its impairment decisions for each of these listings, the LA Regional Board stated that the alleged impairment in benthic community composition in Reach 1 was justified by being "associated" with impairments for two pollutants, trash and selenium, simply because these constituents co-occurred.

- Trash listings address non-contact recreation, not aquatic life, beneficial uses. Furthermore, the most common routes of harm to aquatic organisms by trash are due to ingestion and entanglement – problems unlikely to impact benthic macroinvertebrate larvae.
- Much of the Malibu Creek watershed is listed as impaired for selenium. However, EPA has recognized that "Sulfate and selenium concentrations are present in excess of water quality criteria, **apparently due to natural geologic background**."[5] [Emphasis Added.] As such, this should not be associated as a pollutant.

Fact Sheet #9

Water Body: San Jose Creek Reach 1
Pollutant: Temperature, Water
Listing: List on 303(d) List (TMDL Required List)
Comment & Recommendation: Do Not List – Meets Water Quality Objective

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for impairment due to water temperature be made to the 303(d) list for Reach 1 of San Jose Creek. The Sanitation Districts of Los Angeles County (Sanitation Districts) believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved.

Failure to Meet Water Quality Objectives Has Not Been Demonstrated

The Water Quality Control Plan: Los Angeles Region Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that:

> "At no time shall these WARM-designated waters be raised above 80ºF **as a result of waste discharges**." [Emphasis added.]

This water quality objective clearly distinguishes between exceedance of the 80ºF standard caused by "waste discharges" and those associated with other causes. Evidence indicates that summertime excursions greater than the 80ºF are not caused by waste discharges but are likely due to elevated ambient air temperature, conductive and radiative heating associated with hardened landscapes, a lack of riparian cover, and increased ambient temperatures related to climate change (details below). Additionally, the Draft List does not contain any analysis or evidence indicating that the elevated temperatures occurred as result of wastes discharged.

The Regional Board Fact Sheet states that a single line of evidence was used in the assessment of temperature. Specifically, 42 of 301 samples from Pom-RD, Pom-RC, SJC-C1, and SJC-C2 exceeded the objective from July 2005 to November 2010 using the "Data for Various Pollutants in Various Water Bodies in Sanitation Districts of Los Angeles County, 2005-2010" dataset. Appendix A of this letter contains the full set of data applicable to this listing from Appendix G of the Regional Board Draft Staff Report.

Based on a review of the dataset utilized for the listing evaluation, the Sanitation Districts identified 339 discrete temperature measurements, not 301. The dataset contains 368 results (Appendix 1); however, 29 samples were duplicates. Of the 339 unique temperature measurements, 46 exhibited a temperature that exceeded 80 ºF, not 42. However, 14 of the 46 temperature exceedances were demonstrably caused by conduction and radiation (details below), not waste discharges. Conduction and radiative heating likely also caused the remaining 32 exceedances out of 339 measurements; this total does not meet the minimum number of measured exceedances needed to place a water segment on the section 303(d) list.

Pom-RC and Pom-RD Excursions Above 80 ºF Are Demonstrably Not a Result of Waste Discharges

Tertiary treated water from the Pomona Water Reclamation Plant is discharged to the south fork of San Jose Creek and flows into Reach 1. Receiving water stations Pom-RC, Pom-RD, and SJC-C1 are located approximately 3, 12, and 12.5 miles from the upstream border of Reach 1, respectively. Reach 1 is fully lined in concrete from the upstream border to just upstream of SJC-C1 (Figure 1).

As observed by Sanitation Districts staff and corroborated by EPA staff, groundwater exudes from relief structures distributed throughout the concrete-lined bottom, even in mid-summer (August) after several

Fact Sheet #9

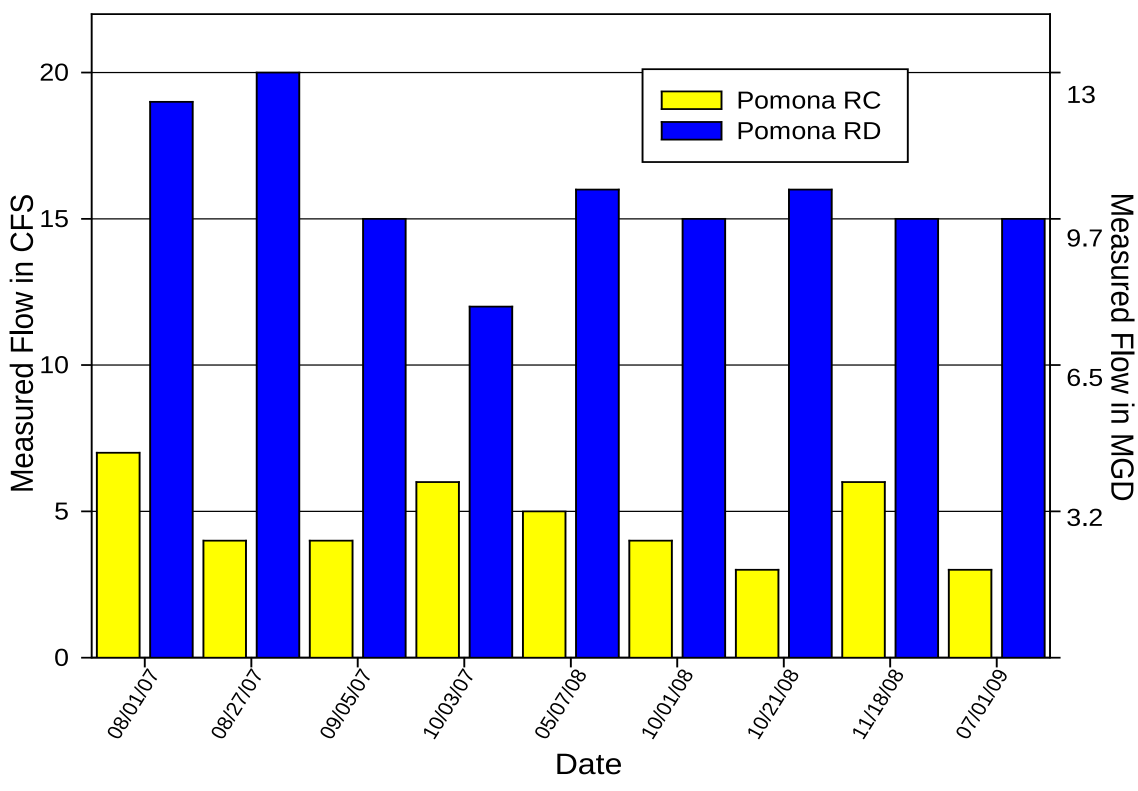
years of drought (Figure 2).[1] In the absence of discharge from the Pomona Water Reclamation Plant or other observed discharges, flows in SJC between Pom-RC and Pom-RD increase by 200% to greater than 400% (Figure 3) due to the release of this groundwater, which has a localized average temperature of approximately 67 ºF.[2] As this groundwater-dominated flow travels downstream, the temperature naturally rises (Figure 4) due to heat conduction through the warm concrete lining and solar radiation exposure in the unshaded channel (Figure 5 shows ambient air temperature as a proxy for solar radiation[3]). When the concrete channel ends upstream of SJC-C1, the water leaves the heat source (concrete channel) and mixes with additional groundwater, resulting in consistently cooler temperatures. The observed spatial and temporal temperature profile, coupled with no identifiable waste discharges and substantial groundwater contributions, clearly demonstrates that the temperature excursions in Reach 1 of San Jose Creek are not a result of waste discharges.

[1] Fleming, Terrence. 2009. Selenium Data from San Jose Creek. Email to Phil Markle. Copy included in Appendix 1.
[2] https://www3.epa.gov/ceampubl/learn2model/part-two/onsite/ex/jne_henrys_map.html
[3] PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu/explorer/, created 24 Feb 2017.

Fact Sheet #9

Figure 2. Manhole Exuding Groundwater into San Jose Creek

Figure 3. Measured Flow at Pom-RC and Pom-RD in the Absence of Discharge from Pomona WRP



Fact Sheet #9

Figure 4. Monthly Average Water Temperatures Between July 2005 and November 2010 in the Absence of Discharge from the Pomona WRP at

- Pom-RC: Upstream Location in the Concrete-Lined Portion of the Reach
- Pom-RD: Downstream Location in the Concrete-Lined Portion of the Reach
- SJC-C1: Unlined Portion of the Reach

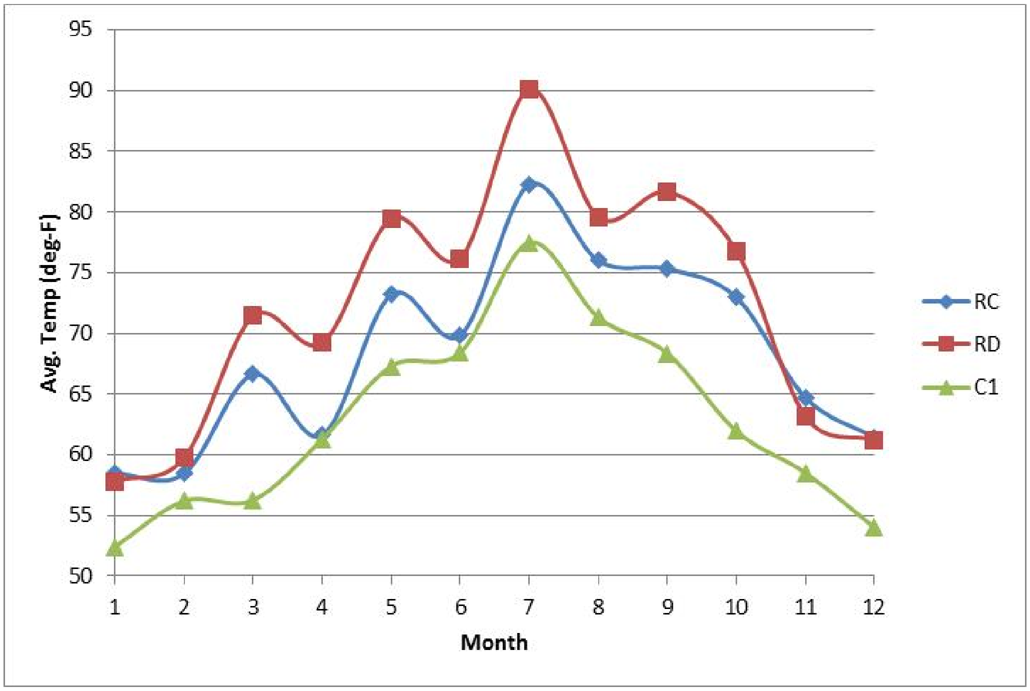
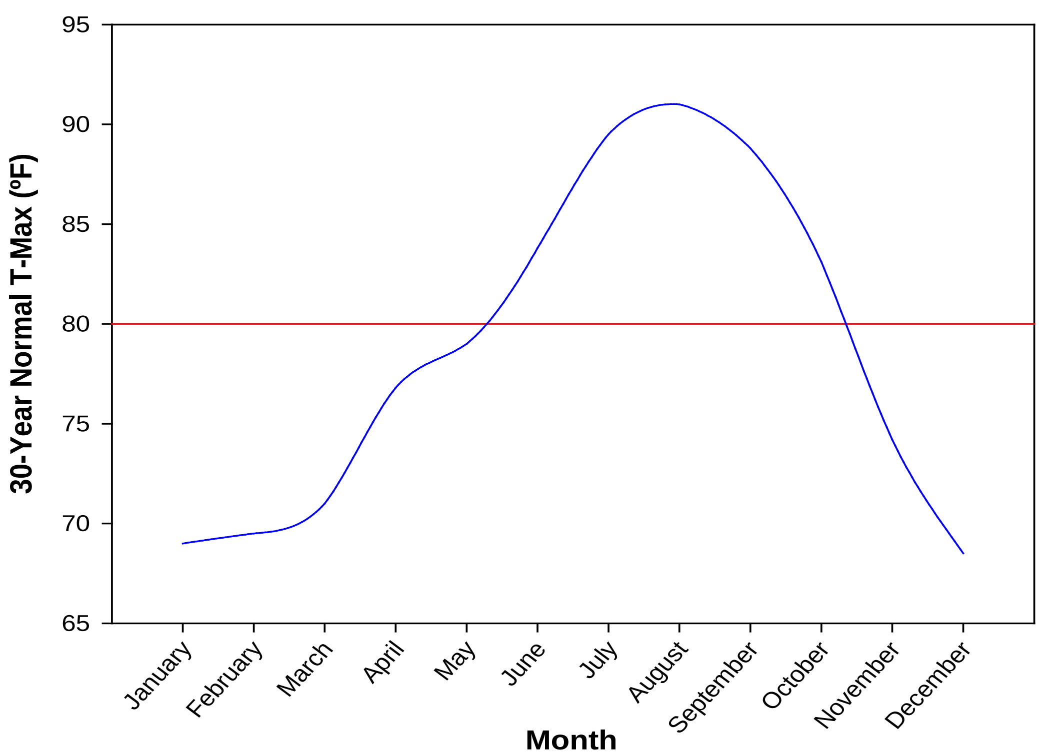


Figure 5. 30-Year Normal Monthly Maximum Air Temperature at Pom-RD[3]



Fact Sheet #10

Water Body: San Gabriel River Reach 1
Pollutant: Temperature, Water
Listing: List on 303(d) List (TMDL Required List)
Comment & Recommendation: Do Not List – Meets Water Quality Objective

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for impairment due to water temperature be made to the 303(d) list for Reach 1 of the San Gabriel River. The Sanitation Districts of Los Angeles County (Sanitation Districts) believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved.

Failure to Meet Water Quality Objectives Has Not Been Demonstrated

The Water Quality Control Plan: Los Angeles Region Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that:

> "At no time shall these WARM-designated waters be raised above 80ºF **as a result of waste discharges**." [Emphasis added.]

This water quality objective clearly distinguishes between exceedance of the 80ºF standard caused by "waste discharges" and those associated with other causes. Evidence indicates that summertime excursions greater than the 80 ºF are not caused by waste dischargefs but are likely due to elevated ambient air temperature, conductive and radiative heating associated with hardened landscapes, a lack of riparian cover, and increased ambient temperatures related to climate change (details below). Additionally, the Draft List does not contain any analysis or evidence indicating that the elevated temperatures occurred as result of wastes discharged.

The Regional Board Fact Sheet states that a single line of evidence was used in the assessment of temperature. Specifically, 93 of 234 samples from LC-R4, R3-1, and R3-1b exceeded the objective from July 2005 to November 2009 using the "Data for Various Pollutants in Various Water Bodies in Sanitation Districts of Los Angeles County, 2005-2010" dataset.

Based on a review of the entire dataset utilized for the listing evaluation,[1] the Sanitation Districts identified 288 discrete temperature measurements, 117 of which exhibited a temperature that exceeded 80ºF. However, these temperature exceedances were not as a result of waste discharges, but were directly associated with high elevated ambient air temperatures as well as conduction and radiation (details below). Therefore, under the definition in the Basin Plan, no exceedances of the water quality objective were observed.

San Gabriel River Reach 1 Excursions Above 80 ºF Are a Result of Radiative and Conductive Heating

Tertiary treated water from the San Jose Creek and Los Coyotes Water Reclamation Plants (WRPs) is discharged to the main stem of the San Gabriel River. Reach 1 is a fully lined concrete channel from approximately 0.25 miles downstream of the San Jose Creek WRP discharge point 001 to the San Gabriel River estuary. As explained in Fact Sheet #9, elevated temperatures in Reach 1 of San Jose Creek occurred even in the absence of observable waste discharges and were caused by conductive heating through the concrete lining and solar radiation exposure. Although a comprehensive assessment of flows, in the absence of WRP discharge, cannot be conducted along the San Gabriel River, the same conditions

---

[1] Data available from Los Angeles Regional Board at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.zip. Accessed 03/21/2017.

Fact Sheet #10

associated with the radiative and conductive heating exist in San Gabriel River Reach 1. This is supported by a significant correlation between ambient air temperature and receiving water temperature ($R^2 = 0.61$) and the fact that 90% of excursions above 80ºF in the receiving water environment occurred during summer months, between June and September. The weight of evidence supports the contention that receiving water temperatures above 80ºF were a result of ambient and environmental conditions (i.e., summer weather and a concrete channel) and not waste discharges.

Fact Sheet #11

Water Body: San Gabriel River Reach 2
Pollutant: Temperature, Water
Listing: List on 303(d) List (TMDL Required List)
Comment & Recommendation: Do Not List – Meets Water Quality Objective

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for impairment due to water temperature be made to the 303(d) list for Reach 2 of the San Gabriel River. The Sanitation Districts of Los Angeles County (Sanitation Districts) believe this proposed listing is inappropriate and recommend not listing due to water quality objectives being achieved.

Failure to Meet Water Quality Objectives Has Not Been Demonstrated

The Water Quality Control Plan: Los Angeles Region Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that:

> "At no time shall these WARM-designated waters be raised above 80ºF **as a result of waste discharges**." [Emphasis added.]

This water quality objective clearly distinguishes between exceedance of the 80ºF standard caused by "waste discharges" and those associated with other causes. Evidence indicates that summertime excursions greater than the 80ºF are not caused by waste discharges but are likely due to elevated ambient air temperature, conductive and radiative heating associated with hardened landscapes, a lack of riparian cover, and increased ambient temperatures related to climate change (details below). Additionally, the Draft List does not contain any analysis or evidence indicating that the elevated temperatures occurred as result of wastes discharged.

The Regional Board Fact Sheet states that a single line of evidence was used in the assessment of temperature. Specifically, 81 of 224 samples from SJC-R2 and SJC-R12 exceeded the objective from July 2005 to November 2009 using the "Data for Various Pollutants in Various Water Bodies in Sanitation Districts of Los Angeles County, 2005-2010" dataset.

Based on a review of the entire dataset utilized for the listing evaluation,[1] the Sanitation Districts identified 81 excursions above 80 ºF out of 232 discrete temperature measurements, not 224. However, these temperature exceedances were not as a result of waste discharges, but were directly associated with high elevated ambient air temperatures as well as conduction and radiation (details below). Therefore, under the definition in the Basin Plan, no exceedances of the water quality objective were observed.

San Gabriel River Reach 2 Excursions Above 80 ºF Are a Result of Radiative and Conductive Heating

Tertiary treated water from the San Jose Creek Water Reclamation Plant (WRP) is discharged to the main stem of the San Gabriel River. The uppermost ¼ mile of Reach 2 is a fully lined concrete channel, containing the R2 receiving water station. Data from this station represents 215 of 232 data points. As explained in Fact Sheet #9, elevated temperatures in Reach 1 of San Jose Creek occurred even in the absence of observable waste discharges and were caused by conductive heating through the concrete lining and solar radiation exposure. Although a comprehensive assessment of flows, in the absence of

---

[1] Data available from Los Angeles Regional Board at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.zip. Accessed 03/21/2017.

Fact Sheet #11

WRP discharge, cannot be conducted along the San Gabriel River, the same conditions associated with the radiative and conductive heating exist in San Gabriel River Reach 2. This is supported the fact that 99% of excursions above 80 ºF in the receiving water environment occurred during summer months, between June and October. The weight of evidence supports the contention that receiving water temperatures above 80 ºF were a result of ambient and environmental conditions (i.e., summer weather and a concrete channel) and not waste discharges.

Fact Sheet #12

Water Body: Santa Clara River Reach 6
Pollutant: Temperature, Water
Listing: List on 303(d) List (TMDL Required List)
Comment & Recommendation: Do Not List

The California Regional Water Quality Control Board, Los Angeles Region (Regional Board) is proposing that a new listing for impairment due to water temperature be made to the 303(d) list for Reach 6 of Santa Clara River. The Sanitation Districts of Los Angeles County (Sanitation Districts) believe this proposed listing is inappropriate and recommend not listing.

Incorrect Datasets Were Used for Listing

The Regional Board Fact Sheet states that a single line of evidence was used in the assessment of temperature. Specifically, 40 of 152 samples from Sa-RA, Sa-RB, and SCR-14 exceeded the objective from June 2005 to October 2010 using the “Data for Various Pollutants in Various Water Bodies in Sanitation Districts of Los Angeles County, 2005-2010” dataset.

Temperature data from location SCR-14 (34.42833333N 118.5394444W) was evaluated as part of Reach 6 of the Santa Clara River. However, SCR-14 is located on Bouquet Canyon Creek, which is recognized as a distinct waterbody by the Region 4 Basin Plan. Figure 1 utilizes a reach delineation layer provided to the Sanitation Districts by Regional Board staff that clearly places SCR-14 in the Bouquet Canyon Creek Reach and not Reach 6. Therefore, temperature measurements from SCR-14 should not be included in the Reach 6 evaluation.

Figure 1. Stations Sa-RB (1), Sa-RA (2), SCR-14 (14), and Bouquet Canyon Creek (Aqua Line)

Fact Sheet #12

Locations Sa-RA and Sa-RB were correctly associated with Reach 6, but results were averaged in the listing evaluation based on the assessment that they were “not spatially independent.” However, as highlighted in Figure 2, Sa-RA is located within the main channel of the Santa Clara River and is typically dry; all 25 temperature measurements at Sa-RA utilized in the Staff Report were associated with upstream dewatering activities or extreme storm events. Sa-RB is located in an isolated pool at the southern edge of the Reach 6 channel that receives recycled water discharges from the Saugus Water Reclamation Plant (WRP). Surface flows from this location travel less than a half-mile downstream in a disconnected side channel before percolating into the dry riverbed. Therefore, even though the two locations are relatively close to each other, Sa-RA is hydrologically isolated from Sa-RB except during extreme rainfall events. Consequently, the two locations would be expected to have very different temperature profiles and should therefore be considered spatially independent, with no averaging of results.

Figure 2. Satellite Imagery of Saugus WRP Ambient Monitoring Stations

The 80°F Water Quality Temperature Objective Is Unnecessary and Inappropriate for Santa Clara River Reach 6

The only dry weather surface flows within this stretch of Reach 6 are associated with recycled water discharges from the Saugus WRP, which percolate into the dry riverbed and eventually resurface downstream near the Reach 5 boundary. At the point of resurfacing, the water temperature averages 69°F and this perennial surface flow supports a diverse aquatic life community in Reach 5.[1] However, the predominant natural condition of Reach 6 is dry and would not be expected to support any aquatic life

---

[1] Hovey, T. (2007) Update: Convict cichlids (*Archocentrus nigrofasciatus*) in the Santa Clara River. Copy included in Appendix B.

Fact Sheet #12

without the Saugus WRP discharge. In addition, the cool temperatures in the water that resurfaces near the Reach 5 boundary demonstrate that elevated temperatures in the isolated discharge area are not detrimental to beneficial uses. Therefore, application of the 80^{o}F water quality objective in Santa Clara Reach 6 is unnecessary and inappropriate, as the presence of water exceeding the 80^{o}F water quality objective would not result in any impairment to naturally occurring aquatic life.

Mitigating the Elevated Temperature at Sa-RB Is Not Feasible

The only reasonable alternative to address the temperature water quality objective below the Saugus WRP at location Sa-RB during dry weather would be to eliminate the discharge. However, it is highly unlikely that the California Department of Fish and Wildlife would support any discharge reductions or elimination, because this action would remove all dry weather surface flows in that stretch of Santa Clara Reach 6 and could potentially reduce the amount of resurfacing groundwater flows that actually support a diverse aquatic community in Santa Clara River Reach 5.

An Evaluation of the Relative Contribution of Radiative and Convective Heating Was Not Conducted

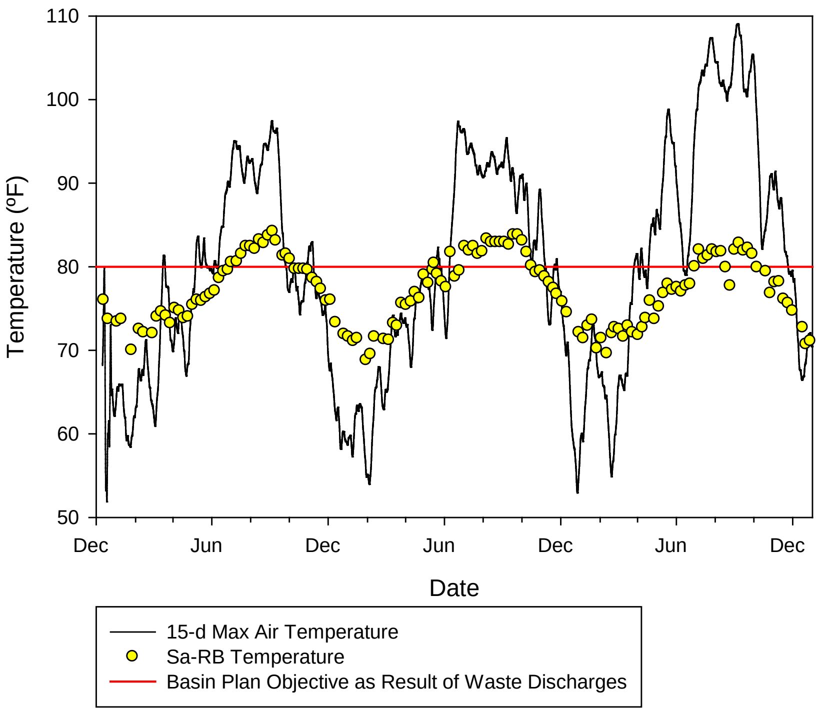
Los Angeles Region Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties (Basin Plan) states that:

> "At no time shall these WARM-designated waters be raised above 80ºF **as a result of waste discharges**." [Emphasis added.]

This objective clearly distinguishes between temperature exceedances caused by "waste discharges" and those associated with other causes. Both the Saugus WRP discharge and the immediate downstream receiving water location (Sa-RB) are heavily influenced by ambient air temperature. Figure 3 includes a plot of the 15-day average values of the maximum air temperature along with the individual water temperature measurements collected at the Sa-RB location. Nearly all of the 80^{o}F temperature exceedances were associated with the higher summer time air temperatures and the two have a statistically significant correlation (R^2 = 0.76). Because exceedances of the Basin Plan temperature objective are limited to those "as a result of waste discharges," an evaluation of the contribution of ambient air temperature to the receiving water should have been conducted before identifying receiving water excursions above 80^{o}F as exceedances of the objective.

Fact Sheet #12

Figure 3. Sa-RB Temperature vs. Maximum Ambient Air Temperature (15-Day Average Value)



Appendix A: Data Tables

Data used in LACSD analysis of San Gabriel River Estuary Toxicity listing.
Accessed via Fact Sheet at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.

Within Data Set but Excluded From Water Board Analysis (n=31)
Single Test/Monthly Median Chronic Toxcity Exceedances (NOEC <100%) (n=0)
Acute Test with <100% Survival (n=6)

Date	ID	Location	Test Name	Single Test Result	Unit
20050601	SJ30206	R9W	Cerio. Chronic-Survival	100	%EFFL
20050601	SJ30206	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20050718	SJ33266	R9W	Cerio. Chronic-Survival	100	%EFFL
20050718	SJ33266	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20050718	SJ33295	R9W	Fathead Acute (Pimphales Prome	100	%SURV
20050801	SJ34396	R9W	Cerio. Chronic-Survival	100	%EFFL
20050801	SJ34396	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20050823	SJ35945	RA2	Topsmelt Acute	97.5	%SURV
20050823	SJ35944	R6	Topsmelt Acute	95	%SURV
20050823	SJ35943	R7	Topsmelt Acute	95	%SURV
20050823	SJ35942	R8	Topsmelt Acute	97.5	%SURV
20050907	SJ36944	R9W	Cerio. Chronic-Survival	100	%EFFL
20050907	SJ36944	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20051102	SJ41229	R9W	Cerio. Chronic-Survival	100	%EFFL
20051102	SJ41229	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20051201	SJ43223	R9W	Cerio. Chronic-Survival	100	%EFFL
20051201	SJ43223	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060105	SJ50383	R9W	Cerio. Chronic-Survival	100	%EFFL
20060105	SJ50383	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060105	SJ50285	R9W	Fathead Acute (Pimphales Prome	100	%SURV
20060118	SJ51126	RA2	90 Menidia Acute	90	%SURV
20060118	SJ51129	R8	90 Menidia Acute	97.5	%SURV
20060125	SJ51588	RA2	Menidia-Survival	100	%EFFL
20060125	SJ51588	RA2	Menidia-Growth	100	%EFFL
20060125	SJ51587	R6	Menidia-Survival	100	%EFFL
20060125	SJ51587	R6	Menidia-Growth	100	%EFFL
20060125	SJ51589	R7	Menidia-Survival	100	%EFFL
20060125	SJ51589	R7	Menidia-Growth	100	%EFFL
20060125	SJ51590	R8	Menidia-Survival	100	%EFFL
20060125	SJ51590	R8	Menidia-Growth	100	%EFFL
20060131	SJ52355	RA2	90 Menidia Acute	97.5	%SURV
20060131	SJ52356	R6	90 Menidia Acute	97.5	%SURV
20060131	SJ52357	R7	90 Menidia Acute	95	%SURV
20060131	SJ52358	R8	90 Menidia Acute	97.5	%SURV
20060202	SJ52190	R9W	Cerio. Chronic-Survival	100	%EFFL
20060202	SJ52190	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060306	SJ55042	R9W	Cerio. Chronic-Survival	100	%EFFL
20060306	SJ55042	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060410	SJ57296	R9W	Cerio. Chronic-Survival	100	%EFFL
20060410	SJ57296	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060428	SJ58518	R6	Topsmelt Chronic Survival	100	%EFFL
20060428	SJ58518	R6	Topsmelt Chronic Growth	100	%EFFL
20060428	SJ58517	R7	Topsmelt Chronic Survival	100	%EFFL
20060428	SJ58517	R7	Topsmelt Chronic Growth	100	%EFFL
20060428	SJ58516	R8	Topsmelt Chronic Survival	100	%EFFL
20060428	SJ58516	R8	Topsmelt Chronic Growth	100	%EFFL
20060429	SJ58519	RA2	Topsmelt Chronic Survival	100	%EFFL

20060429	SJ58519	RA2	Topsmelt Chronic Growth	100	%EFFL
20060503	SJ58913	R9W	Cerio. Chronic-Survival	100	%EFFL
20060503	SJ58913	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060705	SJ62086	R9W	Cerio. Chronic-Survival	100	%EFFL
20060705	SJ62086	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060710	SJ62487	R9W	90 Fathead Acute	100	%EFFL
20060717	SJ64093	R6	Topsmelt Chronic Survival	100	%EFFL
20060717	SJ64093	R6	Topsmelt Chronic Growth	100	%EFFL
20060717	SJ63725	R7	Topsmelt Acute	100	%SURV
20060717	SJ64092	R7	Topsmelt Chronic Survival	100	%EFFL
20060717	SJ64092	R7	Topsmelt Chronic Growth	100	%EFFL
20060717	SJ63724	R8	Topsmelt Acute	100	%SURV
20060717	SJ64091	R8	Topsmelt Chronic Growth	100	%EFFL
20060717	SJ63723	RA2	Topsmelt Acute	100	%SURV
20060717	SJ64094	RA2	Topsmelt Chronic Survival	100	%EFFL
20060717	SJ64094	RA2	Topsmelt Chronic Growth	100	%EFFL
20060814	SJ64571	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20060814	SJ64571	R9W	Cerio. Chronic-Survival	100	%EFFL
20060911	SJ66183	R9W	Cerio. Chronic-Survival	100	%EFFL
20060911	SJ66183	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20061011	SJ68407	R9W	Cerio. Chronic-Survival	100	%EFFL
20061011	SJ68407	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20061101	SJ69224	R9W	Cerio. Chronic-Survival	100	%EFFL
20061101	SJ69224	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20061204	SJ70941	R9W	Cerio. Chronic-Survival	100	%EFFL
20061204	SJ70941	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070108	SJ81063	R6	Topsmelt Chronic Survival	100	%EFFL
20070108	SJ81063	R6	Topsmelt Chronic Growth	100	%EFFL
20070108	SJ81066	R6	Topsmelt Acute	100	%SURV
20070108	SJ81062	R7	Topsmelt Chronic Survival	100	%EFFL
20070108	SJ81062	R7	Topsmelt Chronic Growth	100	%EFFL
20070108	SJ81065	R7	Topsmelt Acute	100	%SURV
20070108	SJ81061	R8	Topsmelt Chronic Survival	100	%EFFL
20070108	SJ81061	R8	Topsmelt Chronic Growth	100	%EFFL
20070108	SJ81074	R8	Topsmelt Acute	100	%SURV
20070108	SJ80358	R9W	Cerio. Chronic-Survival	100	%EFFL
20070108	SJ80358	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070108	SJ80649	R9W	90 Fathead Acute	100	%SURV
20070108	SJ81064	RA2	Topsmelt Chronic Survival	100	%EFFL
20070108	SJ81064	RA2	Topsmelt Chronic Growth	100	%EFFL
20070108	SJ81067	RA2	Topsmelt Acute	100	%SURV
20070226	SJ83442	R9W	Cerio. Chronic-Survival	100	%EFFL
20070226	SJ83442	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070312	SJ83843	R9W	Cerio. Chronic-Survival	100	%EFFL
20070312	SJ83843	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070411	SJ85623	R9W	Cerio. Chronic-Survival	100	%EFFL
20070411	SJ85623	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070430	SJ88059	R6	Topsmelt Chronic Survival	100	%EFFL
20070430	SJ88059	R6	Topsmelt Chronic Growth	100	%EFFL
20070430	SJ88058	R7	Topsmelt Chronic Survival	100	%EFFL
20070430	SJ88058	R7	Topsmelt Chronic Growth	100	%EFFL
20070430	SJ88057	R8	Topsmelt Chronic Survival	100	%EFFL
20070430	SJ88057	R8	Topsmelt Chronic Growth	100	%EFFL
20070430	SJ88060	RA2	Topsmelt Chronic Survival	100	%EFFL
20070430	SJ88060	RA2	Topsmelt Chronic Growth	100	%EFFL

A4

20070611	SJ88456	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070611	SJ88456	R9W	Cerio. Chronic-Survival	100	%EFFL
20070709	SJ90526	R6	Topsmelt Chronic Survival	100	%EFFL
20070709	SJ90526	R6	Topsmelt Chronic Growth	100	%EFFL
20070709	SJ90594	R6	Topsmelt Acute	100	%SURV
20070709	SJ90525	R7	Topsmelt Chronic Survival	100	%EFFL
20070709	SJ90525	R7	Topsmelt Chronic Growth	100	%EFFL
20070709	SJ90593	R7	Topsmelt Acute	100	%SURV
20070709	SJ90524	R8	Topsmelt Chronic Survival	100	%EFFL
20070709	SJ90524	R8	Topsmelt Chronic Growth	100	%EFFL
20070709	SJ90592	R8	Topsmelt Acute	100	%SURV
20070709	SJ89637	R9W	Cerio. Chronic-Survival	100	%EFFL
20070709	SJ89637	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070709	SJ89890	R9W	90 Fathead Acute	**95**	%SURV
20070709	SJ90527	RA2	Topsmelt Chronic Survival	100	%EFFL
20070709	SJ90527	RA2	Topsmelt Chronic Growth	100	%EFFL
20070709	SJ90596	RA2	Topsmelt Acute	100	%SURV
20070801	SJ91186	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070801	SJ91186	R9W	Cerio. Chronic-Survival	100	%EFFL
20070912	SJ93350	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20070912	SJ93350	R9W	Cerio. Chronic-Survival	100	%EFFL
20071008	SJ95043	R6	Fathead Chronic-Survival	100	%EFFL
20071008	SJ95043	R6	Fathead Chronic-Growth	100	%EFFL
20071008	SJ95042	R7	Fathead Chronic-Survival	100	%EFFL
20071008	SJ95042	R7	Fathead Chronic-Growth	100	%EFFL
20071008	SJ95041	R8	Fathead Chronic-Survival	100	%EFFL
20071008	SJ95041	R8	Fathead Chronic-Growth	100	%EFFL
20071008	SJ94573	R9W	Cerio. Chronic-Survival	100	%EFFL
20071008	SJ94573	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20071008	SJ95040	RA2	Fathead Chronic-Survival	100	%EFFL
20071008	SJ95040	RA2	Fathead Chronic-Growth	100	%EFFL
20080102	SJ00349	R6	Topsmelt Acute	100	%SURV
20080102	SJ00348	R7	Topsmelt Acute	100	%SURV
20080102	SJ00347	R8	Topsmelt Acute	100	%SURV
20080102	SJ00151	R9W	90 Fathead Acute	100	%SURV
20080102	SJ00350	RA2	Topsmelt Acute	100	%SURV
20080109	SJ00527	R9W	Cerio. Chronic-Reproduction	100	%EFFL
20080109	SJ00527	R9W	Cerio. Chronic-Survival	100	%EFFL
20080114	SJ01165	R6	Topsmelt Chronic Survival	100	%EFFL
20080114	SJ01165	R6	Topsmelt Chronic Growth	100	%EFFL
20080114	SJ01166	R7	Topsmelt Chronic Survival	100	%EFFL
20080114	SJ01166	R7	Topsmelt Chronic Growth	100	%EFFL
20080114	SJ01167	R8	Topsmelt Chronic Survival	100	%EFFL
20080114	SJ01167	R8	Topsmelt Chronic Growth	100	%EFFL
20080114	SJ01164	RA2	Topsmelt Chronic Survival	100	%EFFL
20080114	SJ01164	RA2	Topsmelt Chronic Growth	100	%EFFL
20080414	SJ06491	R6	Topsmelt Chronic Survival	100	%EFFL
20080414	SJ06491	R6	Topsmelt Chronic Growth	100	%EFFL
20080414	SJ06490	R7	Topsmelt Chronic Survival	100	%EFFL
20080414	SJ06490	R7	Topsmelt Chronic Growth	100	%EFFL
20080414	SJ06489	R8	Topsmelt Chronic Survival	100	%EFFL
20080414	SJ06489	R8	Topsmelt Chronic Growth	100	%EFFL
20080414	SJ06097	R9W	Fathead Chronic-Survival	100	%EFFL
20080414	SJ06097	R9W	Fathead Chronic-Growth	100	%EFFL
20080414	SJ06492	RA2	Topsmelt Chronic Survival	100	%EFFL

20080414	SJ06492	RA2	Topsmelt Chronic Growth	100	%EFFL
20080707	SJ10883	R6	Topsmelt Acute	100	%SURV
20080707	SJ11389	R6	Topsmelt Chronic Survival	100	%EFFL
20080707	SJ11389	R6	Topsmelt Chronic Growth	100	%EFFL
20080707	SJ10885	R7	Topsmelt Acute	100	%SURV
20080707	SJ11390	R7	Topsmelt Chronic Survival	100	%EFFL
20080707	SJ11390	R7	Topsmelt Chronic Growth	100	%EFFL
20080707	SJ10884	R8	Topsmelt Acute	100	%SURV
20080707	SJ11391	R8	Topsmelt Chronic Survival	100	%EFFL
20080707	SJ11391	R8	Topsmelt Chronic Growth	100	%EFFL
20080707	SJ10641	R9W	90 Fathead Acute	**95**	%SURV
20080707	SJ10656	R9W	Fathead Chronic-Survival	100	%EFFL
20080707	SJ10656	R9W	Fathead Chronic-Growth	100	%EFFL
20080707	SJ10886	RA2	Topsmelt Acute	100	%SURV
20080707	SJ11388	RA2	Topsmelt Chronic Survival	100	%EFFL
20080707	SJ11388	RA2	Topsmelt Chronic Growth	100	%EFFL
20081013	SJ16365	R6	Topsmelt Chronic Survival	100	%EFFL
20081013	SJ16365	R6	Topsmelt Chronic Growth	100	%EFFL
20081013	SJ16366	R7	Topsmelt Chronic Survival	100	%EFFL
20081013	SJ16366	R7	Topsmelt Chronic Growth	100	%EFFL
20081013	SJ16367	R8	Topsmelt Chronic Survival	100	%EFFL
20081013	SJ16367	R8	Topsmelt Chronic Growth	100	%EFFL
20081013	SJ15771	R9W	Fathead Chronic-Survival	100	%EFFL
20081013	SJ15771	R9W	Fathead Chronic-Growth	100	%EFFL
20081013	SJ16364	RA2	Topsmelt Chronic Survival	100	%EFFL
20081013	SJ16364	RA2	Topsmelt Chronic Growth	100	%EFFL
20090112	SJ20906	R6	Topsmelt Acute	100	%SURV
20090112	SJ20987	R6	Topsmelt Chronic Survival	100	%EFFL
20090112	SJ20987	R6	Topsmelt Chronic Growth	100	%EFFL
20090112	SJ20907	R7	Topsmelt Acute	100	%SURV
20090112	SJ20989	R7	Topsmelt Chronic Survival	100	%EFFL
20090112	SJ20989	R7	Topsmelt Chronic Growth	100	%EFFL
20090112	SJ20908	R8	Topsmelt Acute	100	%SURV
20090112	SJ20990	R8	Topsmelt Chronic Survival	100	%EFFL
20090112	SJ20990	R8	Topsmelt Chronic Growth	100	%EFFL
20090112	SJ20583	R9W	Fathead Chronic-Survival	100	%EFFL
20090112	SJ20583	R9W	Fathead Chronic-Growth	100	%EFFL
20090112	SJ20725	R9W	90 Fathead Acute	100	%EFFL
20090112	SJ20905	RA2	Topsmelt Acute	100	%SURV
20090112	SJ20988	RA2	Topsmelt Chronic Survival	100	%EFFL
20090112	SJ20988	RA2	Topsmelt Chronic Growth	100	%EFFL
20090406	SJ25094	R9W	Fathead Chronic-Survival	100	%EFFL
20090406	SJ25094	R9W	Fathead Chronic-Growth	100	%EFFL
20090408	SJ25626	R6	Topsmelt Chronic Survival	100	%EFFL
20090408	SJ25626	R6	Topsmelt Chronic Growth	100	%EFFL
20090408	SJ25627	R7	Topsmelt Chronic Survival	100	%EFFL
20090408	SJ25627	R7	Topsmelt Chronic Growth	100	%EFFL
20090408	SJ25628	R8	Topsmelt Chronic Survival	100	%EFFL
20090408	SJ25628	R8	Topsmelt Chronic Growth	100	%EFFL
20090408	SJ25625	RA2	Topsmelt Chronic Survival	100	%EFFL
20090408	SJ25625	RA2	Topsmelt Chronic Growth	100	%EFFL
20090713	SJ29873	R6	Topsmelt Acute	100	%SURV
20090713	SJ30167	R6	Topsmelt Chronic Survival	100	%EFFL
20090713	SJ30167	R6	Topsmelt Chronic Growth	100	%EFFL
20090713	SJ29874	R7	Topsmelt Acute	100	%SURV

20090713	SJ30168	R7	Topsmelt Chronic Survival	100	%EFFL
20090713	SJ30168	R7	Topsmelt Chronic Growth	100	%EFFL
20090713	SJ29875	R8	Topsmelt Acute	100	%SURV
20090713	SJ30169	R8	Topsmelt Chronic Survival	100	%EFFL
20090713	SJ30169	R8	Topsmelt Chronic Growth	100	%EFFL
20090713	SJ29601	R9W	90 Fathead Acute	**97.5**	%SURV
20090713	SJ29685	RA2	Topsmelt Acute	100	%SURV
20090713	SJ30166	RA2	Topsmelt Chronic Survival	100	%EFFL
20090713	SJ30166	RA2	Topsmelt Chronic Growth	100	%EFFL
20090715	SJ29589	R9W	Fathead Chronic-Survival	100	%EFFL
20091019	SJ34121	R9W	Fathead Chronic-Survival	100	%EFFL
20091019	SJ34121	R9W	Fathead Chronic-Growth	100	%EFFL
20100111	10011200149	R6	Survival NOEC	100	%EFFL
20100111	10011200150	R7	Topsmelt Acute	97.5	%SURV
20100111	10011200151	R8	Topsmelt Acute	97.5	%SURV
20100111	10011100410	R9W	90 Fathead Acute	97.5	%SURV
20100111	10011200148	RA2	Survival NOEC	100	%EFFL
20100129	10012900379	R6	Topsmelt Chronic Survival	100	%EFFL
20100129	10012900380	R7	Topsmelt Chronic Survival	100	%EFFL
20100129	10012900381	R8	Topsmelt Chronic Survival	100	%EFFL
20100129	10012900370	R9W	90 Fathead Acute	100	%EFFL
20100129	10012900378	RA2	Topsmelt Chronic Survival	100	%EFFL
20100416	10041600443	R9W	Reproduction NOEC	100	%EFFL
20100416	10041600443	R9W	Survival NOEC	100	%EFFL
20100421	10042200126	R6	Topsmelt Chronic Survival	100	%EFFL
20100421	10042200127	R7	Topsmelt Chronic Survival	100	%EFFL
20100421	10042200128	R8	Topsmelt Chronic Survival	100	%EFFL
20100421	10042200125	RA2	Topsmelt Chronic Survival	100	%EFFL

Data used in LACSD analysis of San Gabriel River Reach 3 Toxicity listing.
Accessed via Fact Sheet at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.zip

Within Data Set but Excluded from Water Board Analysis (n=14 median values, 15 single tests)
Single Test Chronic Toxcity Exceedances (NOEC <100%) (n=6)
Monthly Median Chronic Toxicity Exceedance (n=4)
Acute Test with <100% Survival (n=7)

Date	ID	Location	Test Name	Symbol	Single Test Result	UNIT	Final Result	UNIT
20050808	SJ34856	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20050808	SJ34856	RA	Fathead Chronic-Growth		100	%EFFL		
20050808	SJ34889	RA	90 Fathead Acute		100	%SURV	100	%SURV
20050808	SJ34892	R11	90 Fathead Acute		100	%SURV	100	%SURV
20050815	SJ35409	R11	90 Fathead Acute		100	%SURV	100	%SURV
20050826	SJ36208	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20050826	SJ36208	R11	Fathead Chronic-Growth		100	%EFFL		
20051102	SJ42240	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20051102	SJ42240	R11	Fathead Chronic-Growth		100	%EFFL		
20051114	SJ42613	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20051114	SJ42613	RA	Fathead Chronic-Growth		100	%EFFL		
20060201	SJ52180	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060201	SJ52180	RA	Fathead Chronic-Growth		100	%EFFL		
20060201	SJ52182	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060201	SJ52182	R11	Fathead Chronic-Growth		100	%EFFL		
20060206	SJ52435	RA	90 Fathead Acute		100	%SURV	100	%SURV
20060206	SJ52437	R11	90 Fathead Acute		100	%SURV	100	%SURV
20060306	SJ54863	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060306	SJ54863	RA	Fathead Chronic-Growth		100	%EFFL		
20060314	SJ55356	RA	Fathead Chronic-Survival		100	%EFFL		
20060314	SJ55356	RA	Fathead Chronic-Growth		100	%EFFL		
20060510	SJ59336	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060510	SJ59336	RA	Fathead Chronic-Growth		100	%EFFL		
20060525	SJ60114	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060525	SJ60114	R11	Fathead Chronic-Growth		100	%EFFL		
20060807	SJ64179	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060807	SJ64179	RA	Fathead Chronic-Growth		100	%EFFL		
20060807	SJ64176	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20060807	SJ64176	R11	Fathead Chronic-Growth		100	%EFFL		
20060807	SJ64235	RA	90 Fathead Acute		100	%SURV	100	%SURV
20060807	SJ64231	R11	90 Fathead Acute		97.5	%SURV	97.5	%SURV
20061108	SJ69865	RA	Fathead Chronic-Survival		100	%EFFL	100	100
20061108	SJ69865	RA	Fathead Chronic-Growth		100	%EFFL		
20061108	SJ69858	R11	Fathead Chronic-Survival		100	%EFFL	100	100
20061108	SJ69858	R11	Fathead Chronic-Growth		100	%EFFL		
20070205	SJ81874	R11	Fathead Chronic-Survival		100	%EFFL	100	100
20070205	SJ81874	R11	Fathead Chronic-Growth		100	%EFFL		
20070205	SJ81875	RA	Cerio. Chronic-Survival		100	%EFFL	100	100
20070205	SJ81875	RA	Cerio. Chronic-Reproduction		100	%EFFL		

20070205	SJ81878	R11	Cerio. Chronic-Survival	100	%EFFL	100	100
20070205	SJ81878	R11	Cerio. Chronic-Reproduction	100	%EFFL		
20070205	SJ81819	RA	90 Fathead Acute	100	%SURV	100	%SURV
20070205	SJ81822	R11	90 Fathead Acute	100	%SURV	100	%SURV
20070502	SJ86664	R11	Fathead Chronic-Survival	100	%EFFL	100	100
20070502	SJ86664	R11	Fathead Chronic-Growth	100	%EFFL		
20070502	SJ86669	R11	Cerio. Chronic-Survival	100	%EFFL	100	100
20070502	SJ86669	R11	Cerio. Chronic-Reproduction	100	%EFFL		
20070808	SJ91575	R11	Fathead Chronic-Survival	100	%EFFL	100	100
20070808	SJ91575	R11	Fathead Chronic-Growth	100	%EFFL		
20070808	SJ91310	RA	90 Fathead Acute	97.5	%SURV	97.5	%SURV
20070808	SJ91312	R11	90 Fathead Acute	100	%SURV	100	100
20070925	SJ93977	RA	Cerio. Chronic-Survival	100	%EFFL	100	100
20070925	SJ93977	RA	Cerio. Chronic-Reproduction	100	%EFFL		
20070925	SJ93976	R11	Cerio. Chronic-Survival	100	%EFFL	100	100
20070925	SJ93976	R11	Cerio. Chronic-Reproduction	100	%EFFL		
20071105	SJ95889	RA	Cerio. Chronic-Survival	100	%EFFL	100	100
20071105	SJ95889	RA	Cerio. Chronic-Reproduction	100	%EFFL		
20071105	SJ95891	R11	Cerio. Chronic-Survival	100	%EFFL	100	100
20071105	SJ95891	R11	Cerio. Chronic-Reproduction	100	%EFFL		
20071105	SJ95898	R11	Fathead Chronic-Survival	100	%EFFL	80	%EFFL
20071105	SJ95898	R11	Fathead Chronic-Growth	80	%EFFL		
20071116	SJ96726	R11	Fathead Chronic-Survival	100	%EFFL		
20071116	SJ96726	R11	Fathead Chronic-Growth	100	%EFFL		
20071126	SJ97079	R11	Fathead Chronic-Survival	100	%EFFL		
20071126	SJ97079	R11	Fathead Chronic-Growth	80	%EFFL		
20071211	SJ97715	R11	Fathead Chronic-Survival	100	%EFFL	100	%EFFL
20071211	SJ97715	R11	Fathead Chronic-Growth	100	%EFFL		
20071226	SJ98390	R11	Fathead Chronic-Survival	100	%EFFL		
20071226	SJ98390	R11	Fathead Chronic-Growth	100	%EFFL		
20080109	SJ00538	R11	Fathead Chronic-Survival	100	%EFFL	100	%EFFL
20080109	SJ00538	R11	Fathead Chronic-Growth	100	%EFFL		
20080116	SJ00997	R11	Fathead Chronic-Survival	40	%EFFL		
20080116	SJ00997	R11	Fathead Chronic-Growth	40	%EFFL		
20080131	SJ01357	R11	Fathead Chronic-Survival	100	%EFFL		
20080131	SJ01357	R11	Fathead Chronic-Growth	100	%EFFL		
20080206	SJ02088	R10	Fathead Chronic-Survival	100	%EFFL	<100	%EFFL
20080206	SJ02088	R10	Fathead Chronic-Growth	< 100	%EFFL		
20080206	SJ02096	R11	Fathead Chronic-Survival	100	%EFFL	100	%EFFL
20080206	SJ02096	R11	Fathead Chronic-Growth	100	%EFFL		
20080206	SJ02090	R11	Cerio. Chronic-Survival	100	%EFFL	100	%EFFL
20080206	SJ02090	R11	Cerio. Chronic-Reproduction	100	%EFFL		
20080213	SJ02600	R11	Fathead Chronic-Survival	100	%EFFL		
20080213	SJ02600	R11	Fathead Chronic-Growth	100	%EFFL		
20080227	SJ02986	R11	Fathead Chronic-Survival	100	%EFFL		
20080227	SJ02986	R11	Fathead Chronic-Growth	100	%EFFL		
20080206	SJ02084	RA	Cerio. Chronic-Survival	100	%EFFL	100	%EFFL
20080206	SJ02084	RA	Cerio. Chronic-Reproduction	100	%EFFL		
20080206	SJ01675	RA	90 Fathead Acute	100	%SURV	100	%SURV
20080206	SJ01679	R11	90 Fathead Acute	97.5	%SURV	97.5	%SURV
20080305	SJ03503	R11	Fathead Chronic-Survival	100	%EFFL	100	%EFFL
20080305	SJ03503	R11	Fathead Chronic-Growth	100	%EFFL		
20080312	SJ04454	R11	Fathead Chronic-Survival	100	%EFFL		
20080312	SJ04454	R11	Fathead Chronic-Growth	100	%EFFL		
20080505	SJ06886	R11	Fathead Chronic-Survival	100	%EFFL	100	%EFFL

A9

20080505	SJ06886	R11	Fathead Chronic-Growth		100	%EFFL	100	%EFFL
20080609	SJ08468	RA	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
20080609	SJ08468	RA	Cerio. Chronic-Reproduction		100	%EFFL		
20080609	SJ08470	R11	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
20080609	SJ08470	R11	Cerio. Chronic-Reproduction		100	%EFFL		
20080804	SJ11926	R10	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20080804	SJ11926	R10	Fathead Chronic-Growth		100	%EFFL		
20080804	SJ11917	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20080804	SJ11917	R11	Fathead Chronic-Growth		100	%EFFL		
20080804	SJ11923	R11	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
20080804	SJ11923	R11	Cerio. Chronic-Reproduction		100	%EFFL		
20080804	SJ12066	RA	90 Fathead Acute		100	%SURV	100	%SURV
20080804	SJ12076	R10	90 Fathead Acute		100	%SURV	100	%SURV
20080804	SJ12070	R11	90 Fathead Acute		97.5	%SURV	97.5	%SURV
20081112	SJ17321	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20081112	SJ17321	R11	Fathead Chronic-Growth		100	%EFFL		
20081112	SJ17324	RA	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
20081112	SJ17324	RA	Cerio. Chronic-Reproduction		100	%EFFL		
20081112	SJ17327	R11	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
20081112	SJ17327	R11	Cerio. Chronic-Reproduction		100	%EFFL		
20090202	SJ21563	RA	90 Fathead Acute		97.5	%SURV	97.5	%SURV
20090202	SJ21566	R11	90 Fathead Acute		97.5	%SURV	97.5	%SURV
20090303	SJ22953	RA	Fathead Chronic-Survival	<	100	%EFFL	<100	%EFFL
20090303	SJ22953	RA	Fathead Chronic-Growth	<	100	%EFFL		
20090303	SJ22951	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20090303	SJ22951	R11	Fathead Chronic-Growth		100	%EFFL		
20090511	SJ26787	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20090511	SJ26787	R11	Fathead Chronic-Growth		100	%EFFL		
20090527	SJ27141	RA	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20090527	SJ27141	RA	Fathead Chronic-Growth		100	%EFFL		
20090810	SJ31132	RA	Fathead Chronic-Survival		100	%EFFL	<100	%EFFL
20090810	SJ31132	RA	Fathead Chronic-Growth	<	100	%EFFL		
20090810	SJ31129	R11	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
20090810	SJ31129	R11	Fathead Chronic-Growth		100	%EFFL		
20090810	SJ30785	RA	90 Fathead Acute		97.5	%SURV	97.5	%SURV
20090810	SJ30781	R11	90 Fathead Acute		100	%SURV	100	%SURV
20091109	09110900445	R11	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
20100216	10021600440	R11	90 Fathead Acute		100	%SURV	100	%SURV
20100216	10021600447	RA	90 Fathead Acute		100	%SURV	100	%SURV
20100310	10031000460	RA	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL
20100310	10031000460	RA	Cerio. Chronic-Survival		100	%EFFL		
20100324	10032400521	R11	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL
20100324	10032400521	R11	Cerio. Chronic-Survival		100	%EFFL		
20100510	10051100160	R11	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL
20100510	10051100160	R11	Cerio. Chronic-Survival		100	%EFFL		

*Final result is the monthly median value for chronic toxicity and the single test value for acute toxicity.

Data used in LACSD analysis of Rio Hondo Reach 2 Toxicity listing.
Accessed via Fact Sheet at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.zip

Within Data Set but Excluded from Water Board Analysis (n=5)
Single Test Chronic Toxcity Exceedances (NOEC <100%) (n=4)
Monthly Median Chronic Toxicity Exceedance (n=2)
Acute Test with <100% Survival (n=3)

Date	ID	Location	Test Name	Symbol	Single Test Result	UNIT	Final Result*	UNIT
8/8/2005	SJ34891	RD1	90 Fathead Acute		100	%SURV	100	%SURV
11/2/2005	SJ41223	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
11/2/2005	SJ41223	RD1	Fathead Chronic-Growth		100	%EFFL		
2/1/2006	SJ52185	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
2/1/2006	SJ52185	RD1	Fathead Chronic-Growth		100	%EFFL		
2/6/2006	SJ52434	RD1	90 Fathead Acute		100	%SURV	100	%SURV
5/10/2006	SJ59344	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
5/10/2006	SJ59344	RD1	Fathead Chronic-Growth		100	%EFFL		
8/7/2006	SJ64238	RD1	90 Fathead Acute		100	%SURV	100	%SURV
8/7/2006	SJ64415	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
8/7/2006	SJ64415	RD1	Fathead Chronic-Growth		100	%EFFL		
11/20/2006	SJ70373	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
11/20/2006	SJ70373	RD1	Fathead Chronic-Growth		100	%EFFL		
2/5/2007	SJ81821	RD1	90 Fathead Acute		100	%SURV	100	%SURV
2/5/2007	SJ81873	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
2/5/2007	SJ81873	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
5/2/2007	SJ86670	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
5/2/2007	SJ86670	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
8/8/2007	SJ91309	RD1	90 Fathead Acute		100	%SURV	100	%SURV
9/25/2007	SJ93974	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
9/25/2007	SJ93974	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
11/5/2007	SJ95892	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
11/5/2007	SJ95892	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
2/6/2008	SJ01678	RD1	90 Fathead Acute		95	%SURV	95	%SURV
2/6/2008	SJ02089	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
2/6/2008	SJ02089	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
6/9/2008	SJ08471	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
6/9/2008	SJ08471	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
8/4/2008	SJ11922	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
8/4/2008	SJ11922	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
8/4/2008	SJ12068	RD1	90 Fathead Acute		100	%SURV	100	%SURV
11/12/2008	SJ17326	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
11/12/2008	SJ17326	RD1	Cerio. Chronic-Reproduction		100	%EFFL		
2/2/2009	SJ21565	RD1	90 Fathead Acute		95	%SURV	95	%SURV
3/3/2009	SJ22952	RD1	Fathead Chronic-Survival	<	100	%EFFL	<100	%EFFL
3/3/2009	SJ22952	RD1	Fathead Chronic-Growth	<	100	%EFFL		
3/11/2009	SJ23743	RD1	Fathead Chronic-Survival		100	%EFFL		
3/11/2009	SJ23743	RD1	Fathead Chronic-Growth	<	100	%EFFL		
3/23/2009	SJ24217	RD1	Fathead Chronic-Survival	<	100	%EFFL		
3/23/2009	SJ24217	RD1	Fathead Chronic-Growth	<	100	%EFFL		
5/11/2009	SJ26795	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
5/11/2009	SJ26795	RD1	Fathead Chronic-Growth		100	%EFFL		
8/10/2009	SJ30779	RD1	90 Fathead Acute		100	%SURV	100	%SURV
8/10/2009	SJ31130	RD1	Fathead Chronic-Survival		100	%EFFL	100	%EFFL
8/10/2009	SJ31130	RD1	Fathead Chronic-Growth		100	%EFFL		
11/09/2009	09110900470	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%SURV
11/09/2009	09110900478	RD	Cerio. Chronic-Survival		100	%EFFL	100	%SURV
02/16/2010	10021600426	RD1	90 Fathead Acute		100	%SURV	100	%SURV
02/16/2010	10021600430	RD	90 Fathead Acute		97.5	%SURV	97.5	%SURV
03/10/2010	10031000461	RD1	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL

03/10/2010	10031000461	RD1	Cerio. Chronic-Survival		100	%EFFL	100	%EFFL
03/10/2010	10031000462	RD	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL
03/10/2010	10031000462	RD	Cerio. Chronic-Survival		100	%EFFL		
05/10/2010	10051100166	RD	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL
05/10/2010	10051100166	RD	Cerio. Chronic-Survival		100	%EFFL		
05/10/2010	10051100168	RD1	Cerio. Chronic-Reproduction		100	%EFFL	100	%EFFL
05/10/2010	10051100168	RD1	Cerio. Chronic-Survival		100	%EFFL		
03/10/2010	10031000467	RDB	Survival TUc		1.0	TUc	>1.0	TUc
03/10/2010	10031000467	RDB	Reprod TUc	>	1.0	TUc		
05/10/2010	10051100165	RDB	Survival TUc		1.0	TUc	1.0	TUc
05/10/2010	10051100165	RDB	Reprod TUc		1.0	TUc		
11/09/2009	09110900469	RDB	Survival TUc		1.0	TUc	1.0	TUc
11/09/2009	09110900469	RDB	Growth TUc		1.0	TUc		

*Final result is the monthly median value for chronic toxicity and the single test value for acute toxicity.

A12

Data used in LACSD analysis of Santa Clara River Reach 5 Toxicity listing.
Accessed via Fact Sheet at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.zip

Single Test/Monthly Median Chronic Toxcity Exceedances (NOEC <100%) (n=5)
Acute Test with <100% Survival (n=10)

Date	ID	Location	Test Name	Symbol	Single Test Result	UNIT
7/11/2005	SJ32723	RC	Fathead Chronic-Survival	<	100	%EFFL
7/11/2005	SJ32723	RC	Fathead Chronic-Growth	<	100	%EFFL
7/11/2005	SJ32724	RD	Fathead Chronic-Survival		100	%EFFL
7/11/2005	SJ32724	RD	Fathead Chronic-Growth		100	%EFFL
7/11/2005	SJ32728	RC	90 Fathead Acute		97.5	%SURV
7/11/2005	SJ32729	RD	90 Fathead Acute		100	%SURV
7/11/2005	SJ32730	RE	90 Fathead Acute		97.5	%SURV
7/25/2005	SJ33924	RE	Fathead Chronic-Survival		100	%EFFL
7/25/2005	SJ33924	RE	Fathead Chronic-Growth		100	%EFFL
10/3/2005	SJ38690	RE	Fathead Chronic-Survival		100	%EFFL
10/3/2005	SJ38690	RE	Fathead Chronic-Growth		100	%EFFL
10/3/2005	SJ38691	RC	Fathead Chronic-Survival		100	%EFFL
10/3/2005	SJ38691	RC	Fathead Chronic-Growth		100	%EFFL
10/3/2005	SJ38692	RD	Fathead Chronic-Survival		100	%EFFL
10/3/2005	SJ38692	RD	Fathead Chronic-Growth		100	%EFFL
1/9/2006	SJ50891	RE	90 Fathead Acute		100	%SURV
1/9/2006	SJ50892	RD	90 Fathead Acute		100	%SURV
1/9/2006	SJ50893	RC	90 Fathead Acute		100	%SURV
1/9/2006	SJ51551	RC	Fathead Chronic-Survival		100	%EFFL
1/9/2006	SJ51551	RC	Fathead Chronic-Growth		100	%EFFL
1/9/2006	SJ51552	RD	Fathead Chronic-Survival		100	%EFFL
1/9/2006	SJ51552	RD	Fathead Chronic-Growth		100	%EFFL
1/9/2006	SJ51553	RE	Fathead Chronic-Survival		100	%EFFL
1/9/2006	SJ51553	RE	Fathead Chronic-Growth		100	%EFFL
4/17/2006	SJ57771	RC	Fathead Chronic-Survival		100	%EFFL
4/17/2006	SJ57771	RC	Fathead Chronic-Growth		100	%EFFL
4/17/2006	SJ57772	RE	Fathead Chronic-Survival		100	%EFFL
4/17/2006	SJ57772	RE	Fathead Chronic-Growth		100	%EFFL
4/17/2006	SJ57776	RD	Fathead Chronic-Survival		100	%EFFL
4/17/2006	SJ57776	RD	Fathead Chronic-Growth		100	%EFFL
7/5/2006	SJ62082	RE	Fathead Chronic-Survival		100	%EFFL
7/5/2006	SJ62082	RE	Fathead Chronic-Growth		100	%EFFL
7/5/2006	SJ62083	RD	Fathead Chronic-Survival		100	%EFFL
7/5/2006	SJ62083	RD	Fathead Chronic-Growth		100	%EFFL
7/5/2006	SJ62084	RC	Fathead Chronic-Survival		100	%EFFL
7/5/2006	SJ62084	RC	Fathead Chronic-Growth		100	%EFFL
7/10/2006	SJ62478	RC	90 Fathead Acute		100	%SURV
7/10/2006	SJ62479	RD	90 Fathead Acute		100	%SURV
7/10/2006	SJ62480	RE	90 Fathead Acute		100	%SURV
10/16/2006	SJ68391	RD	Fathead Chronic-Survival		100	%EFFL
10/16/2006	SJ68391	RD	Fathead Chronic-Growth		100	%EFFL
10/16/2006	SJ68392	RC	Fathead Chronic-Survival		100	%EFFL
10/16/2006	SJ68392	RC	Fathead Chronic-Growth		100	%EFFL
10/16/2006	SJ68393	RE	Fathead Chronic-Survival		100	%EFFL
10/16/2006	SJ68393	RE	Fathead Chronic-Growth		100	%EFFL
1/3/2007	SJ80157	RC	Fathead Chronic-Survival		100	%EFFL

1/3/2007	SJ80157	RC	Fathead Chronic-Growth		100	%EFFL
1/3/2007	SJ80160	RE	Fathead Chronic-Survival		100	%EFFL
1/3/2007	SJ80160	RE	Fathead Chronic-Growth	<	100	%EFFL
1/3/2007	SJ80161	RD	Fathead Chronic-Survival		100	%EFFL
1/3/2007	SJ80161	RD	Fathead Chronic-Growth		100	%EFFL
1/8/2007	SJ80643	RE	90 Fathead Acute		100	%SURV
1/8/2007	SJ80644	RD	90 Fathead Acute		97.5	%SURV
1/8/2007	SJ80645	RC	90 Fathead Acute		100	%SURV
4/2/2007	SJ85062	RD	Fathead Chronic-Survival		100	%EFFL
4/2/2007	SJ85062	RD	Fathead Chronic-Growth		100	%EFFL
4/2/2007	SJ85063	RE	Fathead Chronic-Survival		100	%EFFL
4/2/2007	SJ85063	RE	Fathead Chronic-Growth		100	%EFFL
4/2/2007	SJ85064	RC	Fathead Chronic-Survival		100	%EFFL
4/2/2007	SJ85064	RC	Fathead Chronic-Growth		100	%EFFL
7/16/2007	SJ90059	RC	90 Fathead Acute		100	%SURV
7/16/2007	SJ90060	RD	90 Fathead Acute		100	%SURV
7/16/2007	SJ90061	RE	90 Fathead Acute		100	%SURV
7/16/2007	SJ90118	RC	Fathead Chronic-Survival		100	%EFFL
7/16/2007	SJ90118	RC	Fathead Chronic-Growth		100	%EFFL
7/16/2007	SJ90119	RD	Fathead Chronic-Survival		100	%EFFL
7/16/2007	SJ90119	RD	Fathead Chronic-Growth		100	%EFFL
7/16/2007	SJ90120	RE	Fathead Chronic-Survival		100	%EFFL
7/16/2007	SJ90120	RE	Fathead Chronic-Growth		100	%EFFL
10/15/2007	SJ95013	RC	Fathead Chronic-Survival		100	%EFFL
10/15/2007	SJ95013	RC	Fathead Chronic-Growth	<	100	%EFFL
10/15/2007	SJ95014	RD	Fathead Chronic-Survival		100	%EFFL
10/15/2007	SJ95014	RD	Fathead Chronic-Growth		100	%EFFL
10/15/2007	SJ95015	RE	Fathead Chronic-Survival		100	%EFFL
10/15/2007	SJ95015	RE	Fathead Chronic-Growth		100	%EFFL
1/9/2008	SJ00535	RD	Fathead Chronic-Survival		100	%EFFL
1/9/2008	SJ00535	RD	Fathead Chronic-Growth		100	%EFFL
1/9/2008	SJ00536	RC	Fathead Chronic-Survival		100	%EFFL
1/9/2008	SJ00536	RC	Fathead Chronic-Growth		100	%EFFL
1/9/2008	SJ00537	RE	Fathead Chronic-Survival		100	%EFFL
1/9/2008	SJ00537	RE	Fathead Chronic-Growth		100	%EFFL
1/9/2008	SJ00567	RC	90 Fathead Acute		100	%SURV
1/9/2008	SJ00568	RD	90 Fathead Acute		95	%SURV
1/9/2008	SJ00569	RE	90 Fathead Acute		90	%SURV
4/7/2008	SJ05704	RD	Fathead Chronic-Survival		100	%EFFL
4/7/2008	SJ05704	RD	Fathead Chronic-Growth		100	%EFFL
4/7/2008	SJ05707	RC	Fathead Chronic-Survival		100	%EFFL
4/7/2008	SJ05707	RC	Fathead Chronic-Growth		100	%EFFL
4/7/2008	SJ05708	RE	Fathead Chronic-Survival		100	%EFFL
4/7/2008	SJ05708	RE	Fathead Chronic-Growth		100	%EFFL
7/14/2008	SJ10962	RC	90 Fathead Acute		97.5	%SURV
7/14/2008	SJ10963	RD	90 Fathead Acute		95	%SURV
7/14/2008	SJ10964	RE	90 Fathead Acute		97.5	%SURV
7/14/2008	SJ10993	RE	Fathead Chronic-Survival		100	%EFFL
7/14/2008	SJ10993	RE	Fathead Chronic-Growth		100	%EFFL
7/14/2008	SJ10997	RC	Fathead Chronic-Survival		100	%EFFL
7/14/2008	SJ10997	RC	Fathead Chronic-Growth		100	%EFFL
7/14/2008	SJ10998	RD	Fathead Chronic-Survival		100	%EFFL
7/14/2008	SJ10998	RD	Fathead Chronic-Growth		100	%EFFL
10/6/2008	SJ15483	RC	Fathead Chronic-Survival		100	%EFFL
10/6/2008	SJ15483	RC	Fathead Chronic-Growth		100	%EFFL

10/6/2008	SJ15484	RD	Fathead Chronic-Survival		100	%EFFL
10/6/2008	SJ15484	RD	Fathead Chronic-Growth		100	%EFFL
10/6/2008	SJ15485	RE	Fathead Chronic-Survival		100	%EFFL
10/6/2008	SJ15485	RE	Fathead Chronic-Growth		100	%EFFL
1/5/2009	SJ20232	RC	Fathead Chronic-Survival		100	%EFFL
1/5/2009	SJ20232	RC	Fathead Chronic-Growth		100	%EFFL
1/5/2009	SJ20233	RD	Fathead Chronic-Survival		100	%EFFL
1/5/2009	SJ20233	RD	Fathead Chronic-Growth		100	%EFFL
1/5/2009	SJ20240	RC	90 Fathead Acute		100	%SURV
1/5/2009	SJ20241	RD	90 Fathead Acute		100	%SURV
1/6/2009	SJ20234	RE	Fathead Chronic-Survival		100	%EFFL
1/6/2009	SJ20234	RE	Fathead Chronic-Growth	<	100	%EFFL
1/6/2009	SJ20242	RE	90 Fathead Acute		100	%SURV
4/13/2009	SJ25146	RC	Fathead Chronic-Survival		100	%EFFL
4/13/2009	SJ25146	RC	Fathead Chronic-Growth		100	%EFFL
4/13/2009	SJ25148	RE	Fathead Chronic-Survival		100	%EFFL
4/13/2009	SJ25148	RE	Fathead Chronic-Growth		100	%EFFL
4/20/2009	SJ25586	RD	Fathead Chronic-Survival		100	%EFFL
4/20/2009	SJ25586	RD	Fathead Chronic-Growth		100	%EFFL
7/6/2009	SJ29167	RC	Fathead Chronic-Survival		100	%EFFL
7/6/2009	SJ29167	RC	Fathead Chronic-Growth	<	100	%EFFL
7/6/2009	SJ29167	RC	90 Fathead Acute		100	%SURV
7/6/2009	SJ29168	RD	Fathead Chronic-Survival		100	%EFFL
7/6/2009	SJ29168	RD	Fathead Chronic-Growth		100	%EFFL
7/6/2009	SJ29168	RD	90 Fathead Acute		100	%SURV
7/6/2009	SJ29169	RE	Fathead Chronic-Survival		100	%EFFL
7/6/2009	SJ29169	RE	Fathead Chronic-Growth		100	%EFFL
7/6/2009	SJ29169	RE	90 Fathead Acute		100	%SURV
10/5/2009	SJ33437	RD	Fathead Chronic-Survival		100	%EFFL
10/5/2009	SJ33437	RD	Fathead Chronic-Growth		100	%EFFL
10/5/2009	SJ33438	RC	Fathead Chronic-Survival		100	%EFFL
10/5/2009	SJ33438	RC	Fathead Chronic-Growth		100	%EFFL
10/5/2009	SJ33439	RE	Fathead Chronic-Survival		100	%EFFL
10/5/2009	SJ33439	RE	Fathead Chronic-Growth		100	%EFFL
01/04/2010	10010400421	RC	90 Fathead Acute		97.5	%SURV
01/04/2010	10010400422	RD	90 Fathead Acute		100	%SURV
01/04/2010	10010400423	RE	90 Fathead Acute		97.5	%SURV
02/16/2010	10021600412	RC	Survival NOEC		100	%EFFL
02/16/2010	10021600413	RD	Survival NOEC		100	%EFFL
02/16/2010	10021600414	RE	Survival NOEC		100	%EFFL
04/19/2010	10041900436	RC	Survival NOEC		100	%EFFL
04/19/2010	10041900438	RD	Survival NOEC		100	%EFFL
04/19/2010	10041900440	RE	Survival NOEC		100	%EFFL

A15

Data used in LACSD analysis of San Jose Creek Reach 1 Temperature listing.
Accessed via Fact Sheet at http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2010/ref3966.zip

Duplicate Sample - Removed from Analyses (n=29)
Discrete Sample with T> 80 ºF Not Attributable to Waste Discharges (n=14)
Discrete Sample with T> 80 ºF, Possibly Due to Waste Discharges (n=32)
Discrete Samples with T<80 ºF (n=293)

SDATE	Month	JOB	LOC	SUBLOC	Test Name	S	VALUE
20050706	7	SJ32503	SG	C2	Temperature		72.5
20050713	7	SJ32895	SG	C2	Temperature		79.6
20050713	7	SJ32894	SG	C1	Temperature		73.1
20050713	7	SJ32894	SG	C1	Temperature		73.1
20050713	7	SJ32895	SG	C2	Temperature		79.6
20050719	7	SJ33430	POM	RC	Temperature		84.4
20050719	7	SJ33431	POM	RD	Temperature		87.6
20050720	7	SJ33620	SG	C2	Temperature		80.4
20050727	7	SJ34108	SG	C2	Temperature		79.2
20050803	8	SJ34650	SG	C2	Temperature		78.8
20050810	8	SJ35094	SG	C1	Temperature		70.2
20050810	8	SJ35094	SG	C1	Temperature		70.2
20050810	8	SJ35095	SG	C2	Temperature		75
20050810	8	SJ35095	SG	C2	Temperature		75
20050816	8	SJ35444	POM	RC	Temperature		73.8
20050816	8	SJ35445	POM	RD	Temperature		75.9
20050817	8	SJ35605	SG	C2	Temperature		73.9
20050824	8	SJ35973	SG	C2	Temperature		66.4
20050831	8	SJ36563	SG	C2	Temperature		69.4
20050907	9	SJ36923	SG	C2	Temperature		69.1
20050914	9	SJ37388	SG	C1	Temperature		63.5
20050914	9	SJ37388	SG	C1	Temperature		63.5
20050914	9	SJ37389	SG	C2	Temperature		67.1
20050914	9	SJ37389	SG	C2	Temperature		67.1
20050923	9	SJ38029	SG	C2	Temperature		75.9
20050927	9	SJ38223	POM	RC	Temperature		68
20050927	9	SJ38224	POM	RD	Temperature		75.2
20050928	9	SJ38352	SG	C2	Temperature		68.5
20051005	10	SJ38887	SG	C2	Temperature		66
20051012	10	SJ39320	SG	C2	Temperature		67.1
20051025	10	SJ40108	POM	RC	Temperature		66.4
20051025	10	SJ40109	POM	RD	Temperature		64.9
20051026	10	SJ40280	SG	C1	Temperature		58.8
20051026	10	SJ40280	SG	C1	Temperature		58.8
20051026	10	SJ40281	SG	C2	Temperature		69.4
20051026	10	SJ40281	SG	C2	Temperature		69.4
20051102	11	SJ40946	SG	C2	Temperature		76.1
20051109	11	SJ41478	SG	C2	Temperature		72
20051115	11	SJ41779	POM	RC	Temperature		70

20051115	11	SJ41780	POM	RD	Temperature	67.7
20051116	11	SJ41944	SG	C1	Temperature	57.4
20051116	11	SJ41945	SG	C2	Temperature	72.1
20051121	11	SJ42189	SG	C2	Temperature	58.1
20051130	11	SJ42673	SG	C2	Temperature	55.6
20051207	12	SJ43176	SG	C2	Temperature	57.7
20051213	12	SJ43482	POM	RC	Temperature	56.1
20051213	12	SJ43483	POM	RD	Temperature	56.7
20051214	12	SJ43677	SG	C2	Temperature	58.8
20051221	12	SJ44026	SG	C1	Temperature	55.8
20051221	12	SJ44026	SG	C1	Temperature	55.8
20051221	12	SJ44027	SG	C2	Temperature	64
20051221	12	SJ44027	SG	C2	Temperature	64
20051228	12	SJ44249	SG	C2	Temperature	63
20060105	1	SJ50229	SG	C2	Temperature	55.6
20060111	1	SJ50626	SG	C1	Temperature	52.5
20060111	1	SJ50626	SG	C1	Temperature	52.5
20060111	1	SJ50627	SG	C2	Temperature	62.4
20060111	1	SJ50627	SG	C2	Temperature	62.4
20060117	1	SJ50934	POM	RC	Temperature	55.8
20060117	1	SJ50935	POM	RD	Temperature	56.7
20060118	1	SJ51099	SG	C2	Temperature	51.1
20060125	1	SJ51604	SG	C2	Temperature	50.7
20060201	2	SJ52119	SG	C1	Temperature	54.9
20060201	2	SJ52119	SG	C1	Temperature	54.9
20060201	2	SJ52120	SG	C2	Temperature	60.8
20060201	2	SJ52120	SG	C2	Temperature	60.8
20060208	2	SJ52741	SG	C2	Temperature	58.3
20060215	2	SJ53448	SG	C2	Temperature	64
20060221	2	SJ53771	POM	RC	Temperature	54.5
20060221	2	SJ53772	POM	RD	Temperature	60.4
20060222	2	SJ54012	SG	C2	Temperature	51.5
20060227	2	SJ54354	SG	C2	Temperature	62.4
20060309	3	SJ55095	SG	C2	Temperature	68.5
20060315	3	SJ55542	SG	C1	Temperature	54.5
20060315	3	SJ55542	SG	C1	Temperature	54.5
20060315	3	SJ55543	SG	C2	Temperature	67.3
20060315	3	SJ55543	SG	C2	Temperature	67.3
20060323	3	SJ56066	SG	C2	Temperature	60.4
20060323	3	SJ56091	POM	RC	Temperature	76.6
20060323	3	SJ56092	POM	RD	Temperature	79.5
20060327	3	SJ56406	SG	C2	Temperature	62.4
20060403	4	SJ56845	SG	C2	Temperature	61.3
20060412	4	SJ57490	SG	C2	Temperature	62.2
20060418	4	SJ57802	POM	RC	Temperature	65.7
20060418	4	SJ57803	POM	RD	Temperature	72.5
20060419	4	SJ57896	SG	C1	Temperature	60.8
20060419	4	SJ57896	SG	C1	Temperature	60.8

A17

20060419	4	SJ57897	SG	C2	Temperature	70.7
20060419	4	SJ57897	SG	C2	Temperature	70.7
20060426	4	SJ58301	SG	C2	Temperature	64.2
20060503	5	SJ58739	SG	C2	Temperature	69.3
20060510	5	SJ59211	SG	C2	Temperature	73.6
20060517	5	SJ59631	SG	C1	Temperature	68.9
20060517	5	SJ59632	SG	C2	Temperature	70.1
20060525	5	SJ60106	SG	C2	Temperature	72.3
20060530	5	SJ60233	POM	RC	Temperature	73.4
20060530	5	SJ60234	POM	RD	Temperature	90
20060531	5	SJ60319	SG	C2	Temperature	68.9
20060607	6	SJ60703	SG	C1	Temperature	68.9
20060607	6	SJ60703	SG	C1	Temperature	68.9
20060607	6	SJ60704	SG	C2	Temperature	76.1
20060607	6	SJ60704	SG	C2	Temperature	76.1
20060614	6	SJ61114	SG	C2	Temperature	61.9
20060620	6	SJ61419	POM	RC	Temperature	75.9
20060620	6	SJ61420	POM	RD	Temperature	86.5
20060621	6	SJ61520	SG	C2	Temperature	66.2
20060628	6	SJ61849	SG	C2	Temperature	74.7
20060705	7	SJ62025	SG	C2	Temperature	76.8
20060712	7	SJ62387	SG	C1	Temperature	89.8
20060712	7	SJ62387	SG	C1	Temperature	89.8
20060712	7	SJ62388	SG	C2	Temperature	74.5
20060712	7	SJ62388	SG	C2	Temperature	74.5
20060718	7	SJ62668	POM	RC	Temperature	82.9
20060718	7	SJ62669	POM	RD	Temperature	92.7
20060719	7	SJ62848	SG	C2	Temperature	81.9
20060726	7	SJ63466	SG	C2	Temperature	78.9
20060802	8	SJ63967	SG	C2	Temperature	77.6
20060809	8	SJ64371	SG	C2	Temperature	76.1
20060816	8	SJ64621	SG	C1	Temperature	71.2
20060816	8	SJ64621	SG	C1	Temperature	71.2
20060816	8	SJ64622	SG	C2	Temperature	74.3
20060816	8	SJ64622	SG	C2	Temperature	74.3
20060823	8	SJ65105	SG	C2	Temperature	75
20060823	8	SJ65121	POM	RC	Temperature	75
20060823	8	SJ65122	POM	RD	Temperature	85.3
20060830	8	SJ65518	SG	C2	Temperature	74.3
20060906	9	SJ65870	SG	C2	Temperature	82.1
20060913	9	SJ66242	SG	C1	Temperature	69.2
20060913	9	SJ66242	SG	C1	Temperature	69.2
20060913	9	SJ66243	SG	C2	Temperature	78.5
20060913	9	SJ66243	SG	C2	Temperature	78.5
20060920	9	SJ66792	SG	C2	Temperature	49.5
20060927	9	SJ67302	POM	RC	Temperature	69.6
20060927	9	SJ67303	POM	RD	Temperature	78.4
20061004	10	SJ67774	POM	RC	Temperature	69.8

A18

20061004	10	SJ67775	POM	RD	Temperature	75.2
20061011	10	SJ68030	SG	C1	Temperature	62.1
20061011	10	SJ68030	SG	C1	Temperature	62.1
20061011	10	SJ68031	SG	C2	Temperature	69.3
20061011	10	SJ68031	SG	C2	Temperature	69.3
20061018	10	SJ68515	SG	C2	Temperature	71.8
20061101	11	SJ69193	SG	C2	Temperature	66.2
20061101	11	SJ69178	POM	RC	Temperature	64.8
20061101	11	SJ69179	POM	RD	Temperature	65.7
20061108	11	SJ69584	SG	C1	Temperature	60.4
20061108	11	SJ69585	SG	C2	Temperature	73
20061115	11	SJ70125	SG	C2	Temperature	65.3
20061122	11	SJ70521	SG	C2	Temperature	67.5
20061129	11	SJ70705	SG	C2	Temperature	69.6
20061206	12	SJ71055	SG	C2	Temperature	54.5
20061206	12	SJ71157	POM	RC	Temperature	56.8
20061206	12	SJ71158	POM	RD	Temperature	59.7
20061213	12	SJ71334	SG	C1	Temperature	53.1
20061213	12	SJ71334	SG	C1	Temperature	53.1
20061213	12	SJ71335	SG	C2	Temperature	70.7
20061213	12	SJ71335	SG	C2	Temperature	70.7
20061220	12	SJ71763	SG	C2	Temperature	66.8
20070103	1	SJ80095	SG	C2	Temperature	56.5
20070103	1	SJ80110	POM	RC	Temperature	56.1
20070103	1	SJ80111	POM	RD	Temperature	56.8
20070110	1	SJ80416	SG	C1	Temperature	52.9
20070110	1	SJ80417	SG	C2	Temperature	67.8
20070124	1	SJ81144	SG	C2	Temperature	67.1
20070125	1	SJ81163	SG	C2	Temperature	67.6
20070207	2	SJ81998	SG	C2	Temperature	64.9
20070207	2	SJ82019	POM	RC	Temperature	63.3
20070207	2	SJ82020	POM	RD	Temperature	66
20070214	2	SJ82488	SG	C1	Temperature	53
20070214	2	SJ82489	SG	C2	Temperature	67.8
20070221	2	SJ82923	SG	C2	Temperature	70.5
20070307	3	SJ83632	SG	C2	Temperature	65.5
20070307	3	SJ83637	POM	RC	Temperature	59.7
20070307	3	SJ83638	POM	RD	Temperature	63.2
20070314	3	SJ83902	SG	C1	Temperature	58.3
20070314	3	SJ83903	SG	C2	Temperature	64.2
20070328	3	SJ84732	SG	C2	Temperature	58.7
20070404	4	SJ85120	SG	C2	Temperature	66.9
20070404	4	SJ85123	POM	RC	Temperature	59.5
20070404	4	SJ85124	POM	RD	Temperature	63.9
20070411	4	SJ85397	SG	C1	Temperature	62.4
20070411	4	SJ85398	SG	C2	Temperature	66.4
20070418	4	SJ85835	SG	C2	Temperature	68.4
20070425	4	SJ86212	SG	C2	Temperature	72.8

20070502	5	SJ86560	SG	C2	Temperature	73.2
20070502	5	SJ86582	POM	RC	Temperature	75.7
20070502	5	SJ86583	POM	RD	Temperature	76.5
20070509	5	SJ86895	SG	C1	Temperature	67.9
20070509	5	SJ86896	SG	C2	Temperature	71.6
20070523	5	SJ87637	SG	C2	Temperature	74.5
20070530	5	SJ87913	SG	C2	Temperature	75.9
20070606	6	SJ88214	SG	C2	Temperature	76.1
20070606	6	SJ88229	POM	RC	Temperature	65.3
20070606	6	SJ88230	POM	RD	Temperature	77.2
20070613	6	SJ88488	SG	C1	Temperature	71.1
20070613	6	SJ88489	SG	C2	Temperature	79.5
20070620	6	SJ88910	SG	C2	Temperature	79.5
20070627	6	SJ89280	SG	C2	Temperature	79.7
20070704	7	SJ89460	SG	C2	Temperature	80.8
20070705	7	SJ89484	POM	RC	Temperature	82.8
20070705	7	SJ89485	POM	RD	Temperature	91.2
20070711	7	SJ89585	SG	C1	Temperature	71.8
20070711	7	SJ89582	SG	C2	Temperature	81
20070718	7	SJ90193	SG	C2	Temperature	81.3
20070725	7	SJ90634	SG	C2	Temperature	82.9
20070801	8	SJ90951	SG	C2	Temperature	81.9
20070801	8	SJ90976	POM	RC	Temperature	74.1
20070801	8	SJ90977	POM	RD	Temperature	79.2
20070808	8	SJ91407	SG	C1	Temperature	72.9
20070808	8	SJ91408	SG	C2	Temperature	80.8
20070815	8	SJ91972	SG	C2	Temperature	84.6
20070815	8	SJ91973	SG	C2	Temperature	85.5
20070822	8	SJ92330	SG	C2	Temperature	82.4
20070827	8	SJ92507	POM	RC	Temperature	71.2
20070827	8	SJ92508	POM	RD	Temperature	80.4
20070829	8	SJ92740	SG	C2	Temperature	80.8
20070905	9	SJ92936	SG	C2	Temperature	82.9
20070905	9	SJ92943	POM	RC	Temperature	79.2
20070905	9	SJ92944	POM	RD	Temperature	79.8
20070912	9	SJ93279	SG	C1	Temperature	70.9
20070912	9	SJ93280	SG	C2	Temperature	80.2
20070919	9	SJ93686	SG	C2	Temperature	79.2
20070926	9	SJ94081	SG	C2	Temperature	78.8
20071003	10	SJ94410	SG	C2	Temperature	78.3
20071003	10	SJ94412	POM	RC	Temperature	72.7
20071003	10	SJ94413	POM	RD	Temperature	79
20071010	10	SJ94624	SG	C1	Temperature	59.2
20071010	10	SJ94625	SG	C2	Temperature	71.1
20071017	10	SJ94990	SG	C2	Temperature	70.3
20071024	10	SJ95294	SG	C2	Temperature	79.2
20071031	10	SJ95708	SG	C2	Temperature	76.5
20071107	11	SJ96093	SG	C2	Temperature	61.9

A20

20071107	11	SJ96091	POM	RC	Temperature	63
20071107	11	SJ96092	POM	RD	Temperature	59
20071114	11	SJ96307	SG	C1	Temperature	60.3
20071114	11	SJ96308	SG	C2	Temperature	76.3
20071204	12	SJ97384	POM	RC	Temperature	65.8
20071204	12	SJ97385	POM	RD	Temperature	64.9
20071212	12	SJ97779	SG	C1	Temperature	48.4
20071212	12	SJ97780	SG	C2	Temperature	64.2
20071226	12	SJ98314	SG	C2	Temperature	65.5
20080102	1	SJ00117	SG	C2	Temperature	69.8
20080102	1	SJ00122	POM	RC	Temperature	60.5
20080102	1	SJ00123	POM	RD	Temperature	66.4
20080109	1	SJ00352	SG	C1	Temperature	50.9
20080109	1	SJ00353	SG	C2	Temperature	65.7
20080116	1	SJ00889	SG	C2	Temperature	67.3
20080131	1	SJ01361	SG	C2	Temperature	64.2
20080206	2	SJ01656	SG	C2	Temperature	61
20080206	2	SJ01703	POM	RC	Temperature	62.2
20080206	2	SJ01704	POM	RD	Temperature	59.5
20080213	2	SJ02118	SG	C1	Temperature	57.4
20080213	2	SJ02119	SG	C2	Temperature	69.4
20080227	2	SJ02888	SG	C2	Temperature	70
20080305	3	SJ03381	SG	C2	Temperature	64
20080305	3	SJ03428	POM	RC	Temperature	65.3
20080305	3	SJ03429	POM	RD	Temperature	70.7
20080312	3	SJ03805	SG	C1	Temperature	59.2
20080312	3	SJ03806	SG	C2	Temperature	70.2
20080319	3	SJ04440	SG	C2	Temperature	70.9
20080326	3	SJ05156	SG	C2	Temperature	72.3
20080402	4	SJ05419	SG	C2	Temperature	66.6
20080402	4	SJ05441	POM	RC	Temperature	57.3
20080402	4	SJ05442	POM	RD	Temperature	66.4
20080409	4	SJ05732	SG	C1	Temperature	57.6
20080409	4	SJ05733	SG	C2	Temperature	69.1
20080416	4	SJ06108	SG	C2	Temperature	79.2
20080423	4	SJ06371	SG	C2	Temperature	70.5
20080430	4	SJ06571	SG	C2	Temperature	73.1
20080507	5	SJ07052	SG	C2	Temperature	72.8
20080507	5	SJ07013	POM	RC	Temperature	62.6
20080507	5	SJ07014	POM	RD	Temperature	63.5
20080514	5	SJ07254	SG	C1	Temperature	63.5
20080514	5	SJ07255	SG	C2	Temperature	74.6
20080528	5	SJ07929	SG	C2	Temperature	73.2
20080604	6	SJ08240	SG	C2	Temperature	75
20080604	6	SJ08248	POM	RC	Temperature	64.4
20080604	6	SJ08249	POM	RD	Temperature	67.8
20080611	6	SJ08511	SG	C1	Temperature	69.8
20080611	6	SJ08512	SG	C2	Temperature	78.6

A21

20080618	6	SJ09029	SG	C2	Temperature	81.1
20080625	6	SJ09614	SG	C2	Temperature	79.5
20080702	7	SJ10352	SG	C2	Temperature	79
20080702	7	SJ10294	POM	RC	Temperature	84.9
20080702	7	SJ10295	POM	RD	Temperature	90.5
20080709	7	SJ10667	SG	C1	Temperature	70.2
20080709	7	SJ10668	SG	C2	Temperature	79.9
20080716	7	SJ11014	SG	C2	Temperature	81.5
20080723	7	SJ11371	SG	C2	Temperature	81.1
20080730	7	SJ11620	SG	C2	Temperature	81.9
20080806	8	SJ11951	SG	C2	Temperature	82.8
20080806	8	SJ11853	POM	RC	Temperature	85.8
20080806	8	SJ11854	POM	RD	Temperature	77
20080813	8	SJ12243	SG	C1	Temperature	71.4
20080813	8	SJ12244	SG	C2	Temperature	76.6
20080903	9	SJ12915	POM	RC	Temperature	84.4
20080903	9	SJ12916	POM	RD	Temperature	93.2
20080910	9	SJ13466	SG	C1	Temperature	71.8
20080910	9	SJ13467	SG	C2	Temperature	76.6
20080924	9	SJ14592	SG	C2	Temperature	73.7
20081001	10	SJ15115	POM	RC	Temperature	82.9
20081001	10	SJ15116	POM	RD	Temperature	88
20081008	10	SJ15453	SG	C1	Temperature	70
20081008	10	SJ15454	SG	C2	Temperature	74.7
20081112	11	SJ16988	SG	C1	Temperature	55.9
20081112	11	SJ16989	SG	C2	Temperature	58.4
20081113	11	SJ17173	POM	RC	Temperature	68.9
20081113	11	SJ17174	POM	RD	Temperature	59.7
20081118	11	SJ17400	POM	RC	Temperature	56.8
20081118	11	SJ17401	POM	RD	Temperature	63.6
20081119	11	SJ17511	SG	C2	Temperature	69.9
20081124	11	SJ17664	SG	C2	Temperature	66.8
20081201	12	SJ17793	SG	C1	Temperature	56.3
20081201	12	SJ17794	SG	C2	Temperature	68.3
20081202	12	SJ18127	SG	C1	Temperature	59.9
20081202	12	SJ18126	SG	C2	Temperature	58.7
20081203	12	SJ18407	SG	C2	Temperature	69.8
20081203	12	SJ18362	POM	RC	Temperature	66.9
20081203	12	SJ18363	POM	RD	Temperature	63.5
20081210	12	SJ18872	SG	C1	Temperature	49.5
20081210	12	SJ18873	SG	C2	Temperature	54.5
20081231	12	SJ19536	SG	C2	Temperature	64
20090102	1	SJ20073	SG	C2	Temperature	62.6
20090107	1	SJ20253	SG	C2	Temperature	64
20090107	1	SJ20188	POM	RC	Temperature	61.5
20090107	1	SJ20189	POM	RD	Temperature	51.4
20090114	1	SJ20517	SG	C1	Temperature	53.4
20090114	1	SJ20518	SG	C2	Temperature	57.3

A22

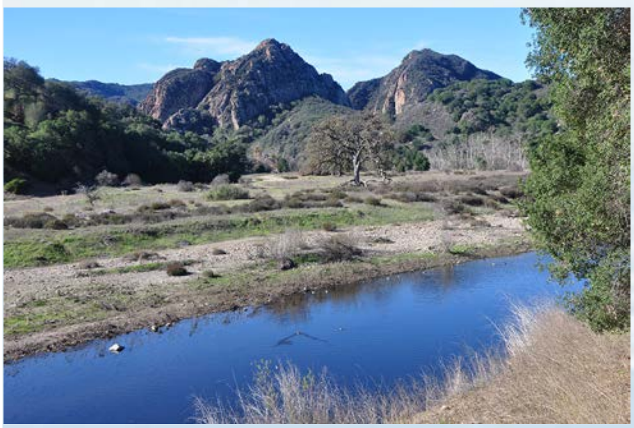
20090121	1	SJ20883	SG	C2	Temperature	68.7
20090128	1	SJ21192	SG	C2	Temperature	64.6
20090204	2	SJ21489	SG	C2	Temperature	67.3
20090204	2	SJ21422	POM	RC	Temperature	53.8
20090204	2	SJ21423	POM	RD	Temperature	53.1
20090212	2	SJ21816	SG	C2	Temperature	64.8
20090223	2	SJ22242	SG	C1	Temperature	60.8
20090223	2	SJ22243	SG	C2	Temperature	61.3
20090311	3	SJ23588	SG	C1	Temperature	54.8
20090311	3	SJ23589	SG	C2	Temperature	57.6
20090317	3	SJ22677	POM	RC	Temperature	64.9
20090317	3	SJ22678	POM	RD	Temperature	72.3
20090318	3	SJ23943	SG	C2	Temperature	69.8
20090401	4	SJ24369	POM	RC	Temperature	63.9
20090401	4	SJ24370	POM	RD	Temperature	74.1
20090408	4	SJ24822	SG	C1	Temperature	64.8
20090408	4	SJ24823	SG	C2	Temperature	63.9
20090506	5	SJ25996	POM	RC	Temperature	81.1
20090506	5	SJ25997	POM	RD	Temperature	87.8
20090513	5	SJ26389	SG	C1	Temperature	68.9
20090513	5	SJ26390	SG	C2	Temperature	73.4
20090603	6	SJ27385	SG	C2	Temperature	75
20090603	6	SJ27304	POM	RC	Temperature	73.8
20090603	6	SJ27305	POM	RD	Temperature	73.3
20090610	6	SJ27760	SG	C1	Temperature	63.5
20090610	6	SJ27761	SG	C2	Temperature	74.1
20090617	6	SJ28224	SG	C2	Temperature	78.6
20090624	6	SJ28547	SG	C2	Temperature	76.5
20090701	7	SJ28828	SG	C2	Temperature	78.6
20090701	7	SJ28781	POM	RC	Temperature	76.3
20090701	7	SJ28782	POM	RD	Temperature	88.5
20090708	7	SJ29042	SG	C1	Temperature	74.6
20090708	7	SJ29043	SG	C2	Temperature	80.3
20090715	7	SJ29483	SG	C2	Temperature	82
20090722	7	SJ29806	SG	C2	Temperature	82.2
20090812	8	SJ30821	SG	C1	Temperature	72.1
20090812	8	SJ30822	SG	C2	Temperature	81.5
20090909	9	SJ31905	SG	C1	Temperature	70.4
20090909	9	SJ31906	SG	C2	Temperature	82.7
20091021	10	SJ34087	SG	C1	Temperature	62.5
20091021	10	SJ34088	SG	C2	Temperature	77.5

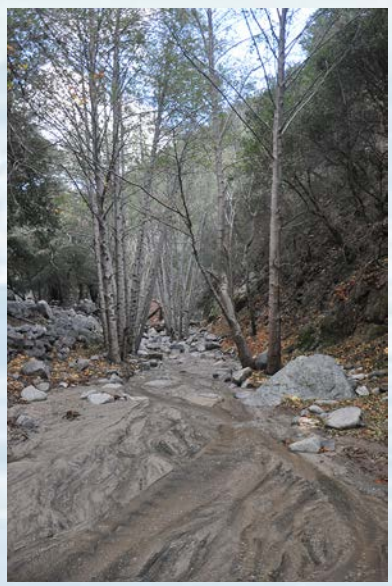
Appendix B: References

Appendix 1

B3

Bioassessment of Perennial Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey





Raphael D. Mazor

Southern California Coastal Water Research Project

SCCWRP Technical Report 844

SMC
Southern California
STORMWATER MONITORING
Coalition

SWAMP
Surface Water
Ambient Monitoring
Program

SCCWRP
Established 1969

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

Ventura Countywide
Stormwater Quality
Management Program

LOS ANGELES COUNTY FLOOD CONTROL DISTRICT

Council for
Watershed Health

COUNTY OF SAN BERNARDINO
DEPARTMENT OF PUBLIC WORKS

COUNTY OF ORANGE CALIFORNIA

Riverside County Flood Control
Water Conservation District

THE NOBLEST MOTIVE IS THE PUBLIC GOOD
COUNTY OF SAN DIEGO
MDCCCLI

CALIFORNIA
DEPARTMENT OF
FISH &
WILDLIFE

CALIFORNIA REGIONAL WATER QUALITY CONTROL BOARD
LOS ANGELES REGION

STATE OF CALIFORNIA
SANTA ANA REGION
REGIONAL WATER QUALITY CONTROL BOARD

CALIFORNIA REGIONAL WATER QUALITY CONTROL BOARD
SAN DIEGO REGION (9)

B5

A REGIONAL APPROACH TO EVALUATING THE BIOLOGICAL CONDITION OF SOUTHERN CALIFORNIA'S WADEABLE STREAMS

2009-2013: THE FIRST FIVE YEARS OF THE STORMWATER MONITORING COALITION'S REGIONAL MONITORING PROGRAM

SMC Southern California Stormwater Monitoring Coalition

OVERVIEW

In 2009, the Southern California Stormwater Monitoring Coalition embarked on an ambitious effort to evaluate the biological condition of 4,300 miles of wadeable streams in the region's coastal watersheds. Over the ensuing five years, the coalition's participating agencies conducted extensive survey and sampling work at more than 500 randomly selected sites encompassing 15 major watersheds in California's South Coast region. Monitoring efforts that had historically been done with minimal coordination were unified around a cohesive, shared vision for the first time, generating high-quality data sets that have painted a powerful picture of regional stream condition. The SMC survey is a regional enhancement of the statewide Perennial Stream Assessment.

The mature riparian plants and biological complexity observed in upper portions of Trabuco Creek in the Santa Ana Mountains reflect a stream that is in good biological condition. 25% of wadeable stream-miles in Southern California were found to be in good condition in the five-year survey.

Caballero Creek, a channelized, algae -filled tributary to the Los Angeles River, reflects severe habitat degradation and impacts of elevated nutrient concentrations. The survey found that both types of stressors were widespread in Southern California streams.

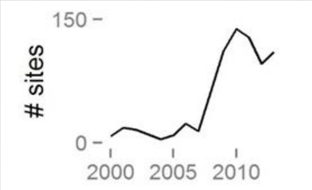
PROGRAM BENEFITS AND IMPACTS

- **Relevant to managers**: Comprehensive data sets inform decisions about priorities and resource allocation, and identify opportunities for causal assessment follow-up studies.
- **Cost-effective:** Each participant realizes approximately 10 times the data value relative to costs.
- **More influential**: Regional collaborations provide more data to inform statewide policymaking, and highlight local concerns.
- **Conversation-altering**: Provides a starting point for developing innovative management strategies that consider and go beyond water chemistry.

KEY FINDINGS

25% of the region's wadeable stream-miles are in good biological condition, including:

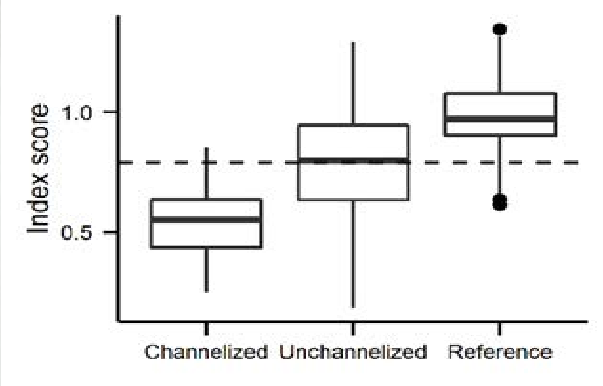
- **60%** of stream-miles in open-space
- **9%** in agricultural areas
- **2%** in urban areas



The Regional Monitoring Program stream survey, which began in 2009, significantly increased the number of stream sites sampled in the region.

HIGH-PRIORITY STRESSORS ON WADEABLE STREAMS

Stressors affecting more than 25% of stream-miles	Stressors affecting 10% to 25% of stream-miles
• Nutrients (Nitrogen and Phosphorus) • Physical habitat degradation • Sulfates • Total dissolved solids	• Chloride • Total suspended solids • pH



Index scores based on benthic macroinvertebrates were lower in channelized streams than non-channelized and reference streams; however, high scores for algal indices were observed in channelized streams where water quality was good. These findings provide a basis for regulators and stormwater agencies to discuss management strategies for channelized streams.

FOUR ECOLOGICAL CONDITION INDICATORS

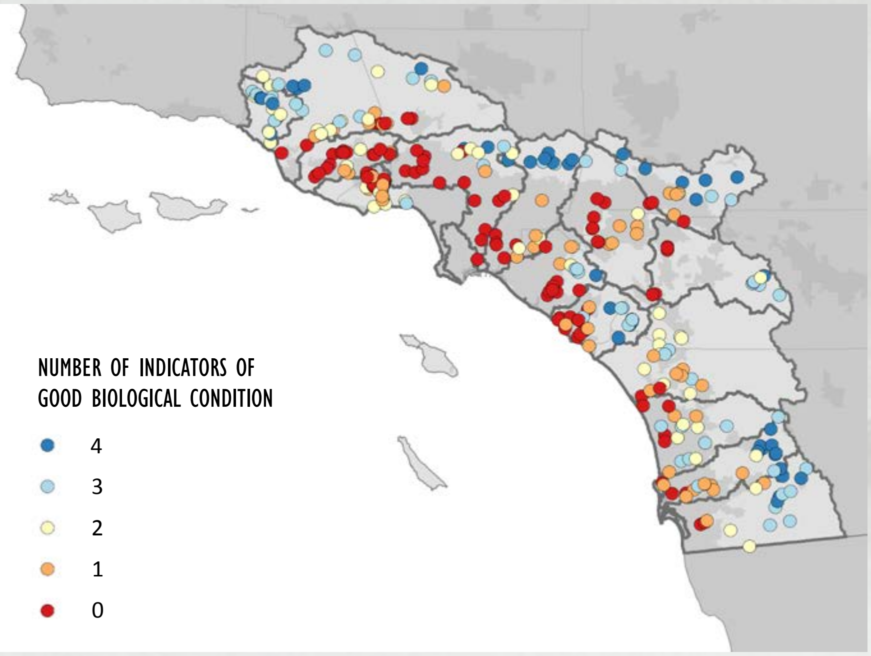
The biological condition of streams was assessed by collecting data for four biological indicators. Each indicator is sensitive to a unique combination of stream stressors, allowing it to provide different types of information about a stream's overall health. Collectively, the four indicators provide comprehensive, direct evidence of a stream's capacity to support aquatic life, a more revealing approach than measuring the chemical concentrations of pollutants.

1 **Benthic macro-invertebrates**, such as aquatic insects, snails, and worms, respond to changes in habitat or water quality over their lifespans.

2 **Soft algae**, such as *Vaucheria*, may form clumps or filaments on submerged rocks. Some species proliferate when nutrients are elevated, while others thrive when nutrients are scarce.

3 **Diatoms**, such as *Navicula*, respond strongly to changes in water chemistry and sedimentation.

4 **Riparian habitats**, which support both terrestrial and in-stream wildlife, may be degraded by habitat alteration, upstream discharges, and hydrologic modification.



At the 500+ randomly selected sampling sites in the stream survey, anywhere from 0 to all 4 biological indicators indicated that a site was in good biological condition. The four indicators – benthic macroinvertebrates, diatoms, soft algae, and riparian habitat condition – collectively were used to assess a site's biological condition.

WATERSHEDS WITH MANY STREAMS IN GOOD CONDITION

- Ventura River
- Upper Santa Ana River
- Tijuana + Sweetwater + Otay Rivers

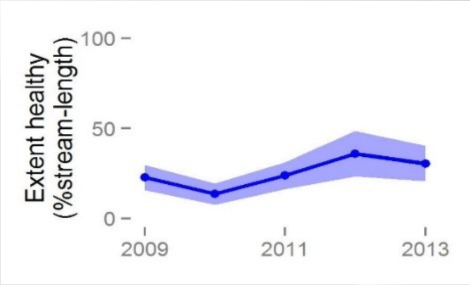
WATERSHEDS WITH FEW STREAMS IN GOOD CONDITION

- Calleguas Creek
- Lower Santa Ana River
- San Dieguito River + Carlsbad Hydrologic Unit

NO APPARENT TRENDS

Although there was some year-over-year variability, the survey did not find a change in the health of the streams over the five-year sampling period, from 2009 to 2013.

Urban streams tended to be in consistently poor biological condition, whereas open-space and agricultural streams tended to experience greater year-to-year variability.



The portion of healthy stream-miles fluctuated over the five-year sampling period, but overall showed no clear trends in either direction. The blue shading represents the 95% confidence interval.

A NEW SURVEY UNDERWAY

The success of the SMC's Regional Monitoring Program has paved the way for a second round of the program, which began in spring 2015. The first five-year survey will serve as a baseline for detecting trends over time.

The second cycle includes nonperennial streams, a critical habitat that makes up more than half of the region's stream-miles, and will seek to clarify the linkage between stressors and biotic integrity.

STORMWATER MONITORING COALITION MEMBERS

County of Los Angeles Department of Public Works, County of Orange Public Works, County of San Diego Department of Public Works, Riverside County Flood Control and Water Conservation District, San Bernardino County Flood Control District, Ventura County Watershed Protection District, City of Long Beach Public Works Department, City of Los Angeles Department of Public Works, California Regional Water Quality Control Board—Santa Ana Region, Los Angeles Region, and San Diego Region, State Water Resources Control Boards, California Department of Transportation, Southern California Coastal Water Research Project (SCCWRP). Collaborating organization: U.S. Environmental Protection Agency Office of Research and Development | www.socalsmc.org

DEVELOPED IN COLLABORATION WITH THE SURFACE WATER AMBIENT MONITORING PROGRAM

SWAMP Surface Water Ambient Monitoring Program

B7

Bioassessment of Perennial Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey

Stormwater Monitoring Coalition Bioassessment Workgroup

Prepared by Raphael D. Mazor
Southern California Coastal Water Research Project

May 2015

Technical Report 844

ACKNOWLEDGEMENTS

This survey was jointly funded by the California State Water Resources Control Board, the Stormwater Monitoring Coalition, and the SMC member agencies. The authors thank the following labs, scientists, and agencies for providing data for this project: the California Surface Water Ambient Monitoring Program; Ventura Countywide Stormwater Quality Management Program; Los Angeles County Flood Control District; Council for Watershed Health; Orange County Public Works; San Bernardino County Stormwater Program; Riverside County Flood Control and Water Conservation District; San Diego County Public Works; San Francisco Estuary Institute; California Regional Water Quality Control Boards of the Los Angeles, Santa Ana, and San Diego Regions; State Water Resources Control Board; and California Department of Fish and Wildlife Aquatic Bioassessment Laboratory; as well as, Aquatic Bioassay and Consulting Laboratories; AMEC Environment & Infrastructure; Nautilus Environmental; Weston Solutions, Inc.; Drs. Dessie Underwood and Richard Gossett at California State University, Long Beach; EcoAnalysts, Inc.; Dr. Patrick Kociolek at University of Colorado, Boulder; Drs. Robert Sheath and Rosalina Hristova at California State University, San Marcos; Moss Landing Marine Laboratories; CRG Marine Laboratories; and Edward S. Babcock Laboratories, Inc. Additional technical support was provided by the Los Angeles County Sanitation District, the City of San Diego, and the Geographic Information Center at California State University, Chico.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

Acronym	Definition
AFDM	Ash-free dry mass
CI	Confidence interval
CMAP	California Monitoring and Assessment Program
CRAM	California Rapid Assessment Method
CSCI	California Stream Condition Index
CTR	California Toxics Rule
D18	Diatom Index of Biotic Integrity
EMAP	Environmental Monitoring and Assessment Program
EPA	Environmental Protection Agency
IBI	Southern and Central California Index of Biotic Integrity
NHD	National Hydrography Dataset
NRSA	National Rivers and Streams Assessment
O/E	Ratio of Observed to Expected Taxa
PCT_BIGR	% large substrate (>128 mm)
PCT_CPOM	% cover by coarse particulate organic matter
PCT_FAST	% fast-water habitat
PCT_MAP	% macroalgae cover
PCT_MCP	% macrophyte cover
PCT_MIAT1	% cover by thick (>1 mm) microalgae
PCT_SAFN	% sands and fines (≤2 mm)
pMMI	Predictive Multi-Metric Index
PSA	Perennial Stream Assessment
S2	Soft Algae Index of Biotic Integrity
SD	Standard Deviation
SMC	Stormwater Monitoring Coalition
TDS	Total Dissolved Solids
TN	Total Nitrogen
TP	Total Phosphorous
XEMBED	Mean % cobble embeddedness
XFC_NAT_SWAMP	Natural fish cover
XMIATP	Mean microlagae thickness (where present)

EXECUTIVE SUMMARY

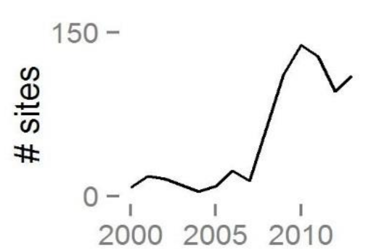
Streams are important natural resources in the South Coast of California, a region that extends from Ventura to San Diego counties. Competing needs for aquatic resources are intense and growing. Assessing the biological condition of these streams has been the focus of considerable monitoring activity. However, until 2009 these efforts were minimally coordinated and provided only limited information about the health of streams in the region, as a result of an emphasis on end-of-watershed monitoring. The Stormwater Monitoring Coalition (SMC) regional perennial stream survey was created in response to the need for a more holistic and coordinated approach. This report provides the results of a five-year probability-based bioassessment of southern California's perennial wadeable streams and represents one of the most comprehensive assessments of stream conditions in the United States.

The five-year survey was designed to answer key questions that are essential to watershed management:

1) What is the biological condition of perennial streams in the region?

2) What stressors are associated with poor condition?

3) Are conditions changing over time?



The Stormwater Monitoring Coalition has greatly increased the number of sites sampled in southern California.

Answering these questions at the regional scale provides resource managers with the ability to contextualize their programs and improve understanding of the effectiveness of management actions, prioritization of streams most in need of protection, and identification of stressors that are likely to pose the greatest risk to stream health.

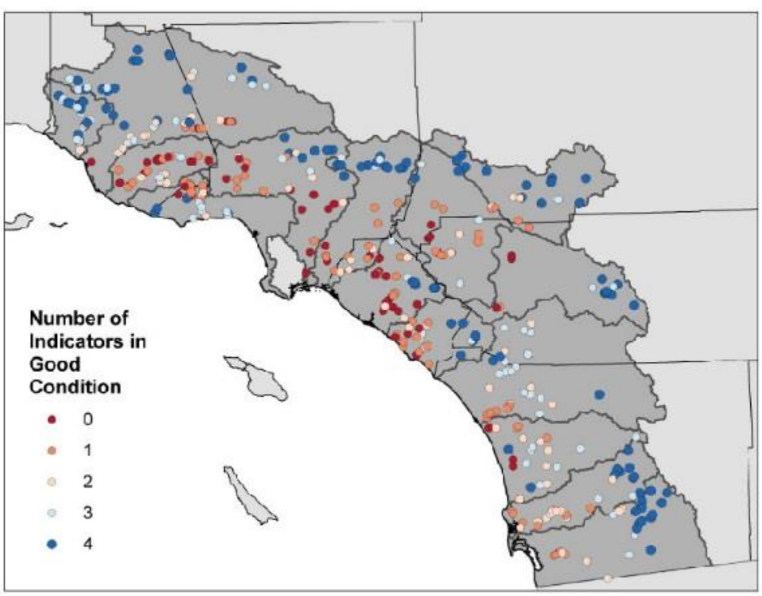
Prior to the initiation of the SMC perennial stream survey, bioassessment efforts in southern California had a limited ability to answer any of these questions. Lead monitoring agencies worked with little coordination, typically addressing site-specific problems with sometimes incomparable methodologies and rarely sharing data. Targeted monitoring mandated by permits did not provide the regional context needed to inform management decisions. Earlier probabilistic sampling efforts in southern California were limited (rarely more than a handful of sites per year), and were conducted as a small part of a statewide or national assessment.

Since the initiation of the SMC perennial stream survey in 2009, stormwater agencies have been able to coordinate their monitoring efforts with regulatory agencies, reallocate resources, and generate the needed data in a cost-neutral way, while simultaneously allowing regulated agencies to fulfill their permit obligations. This survey serves as the regional component of the statewide Perennial Stream Assessment, allowing both the SMC and the State Water Resources Control Board to leverage resources and support each other's surveys.

To answer key management questions, over 500 sites were sampled for four key indicators of biological condition: benthic macroinvertebrates, diatoms, soft algae, and riparian wetlands. These indicators were used to assess the biological health of over 7000 km of streams. In addition, water chemistry, water column toxicity, and physical habitat were examined in order to identify stressors affecting biological conditions in the region. Furthermore, because the survey spanned five years, initial estimates of regional trends are now possible.

Key Findings

Biologically healthy perennial streams are a scarce resource, comprising only 25% of perennial wadeable stream-miles in the region. Based on four biological indicators (i.e., benthic macroinvertebrates, diatoms, soft algae, and riparian wetlands), perennial streams in good biological condition (i.e., scores above the 10th percentile of reference sites) were largely confined to undeveloped portions of watersheds; most indicators identified slightly better conditions at agricultural streams relative to urban streams. Ventura, Santa Clara, Upper Santa Ana, and Southern San Diego watersheds were in better condition than other watersheds for most indicators, whereas perennial streams in poor condition (i.e., scores below the 10th percentile of reference sites) were most extensive in Calleguas, Los Angeles, San Gabriel, and Lower and Middle Santa Ana watersheds.



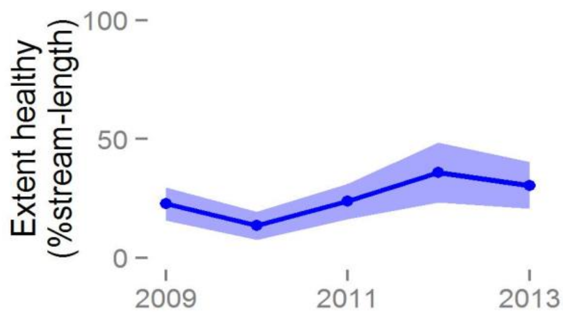
Perennial stream condition was evaluated with four biological indicators: benthic macroinvertebrates, diatoms, soft algae, and riparian condition. In general, these components of the stream community rarely indicated good health in developed portions of watersheds.

Nutrients, sulfates, and habitat degradation were extensive, high-risk stressors associated with poor biological condition. Future investigations should consider these possible candidate stressors as potential causes of poor biological condition. In contrast, metals, pyrethroids, and toxicity were either rarely above threshold or weakly associated with biological condition.

A large extent of the South Coast region was at risk from physical habitat degradation, elevated nutrients, and major ions. Pyrethroids and metals were either weakly or rarely associated with poor health.

Very high priority **(Affects more than 25% of region)**	**High priority** **(Affects more than 10% of region)**	**Moderate or low priority** **(Limited extent or low risk)**
Nitrogen Phosphorus Physical habitat Sulfates Dissolved solids	Chloride Suspended solids pH	Pyrethroids Metals Biomass Toxicity

No changes in biological condition were detected. Although mean condition estimates fluctuated from year to year, conditions in 2013 were similar to those observed in 2009; fluctuations were primarily driven by variability in undeveloped streams, as urban streams were consistently in poor condition, varying little from year to year. At no time during the survey were more than 35% or less than 14% of streams estimated to be intact for all indicators. Moving forward, the ability to detect trends could be improved by minor changes to the study design, such as revisiting sites over several years and by extending the survey for additional years.



Extent of perennial streams in good biological condition for all four indicators (benthic macroinvertebrates, diatoms, soft algae, and riparian condition) fluctuated from year to year, but was always limited to less than 35% of perennial stream-miles in the region. The band indicates the 95% confidence interval.

How can this survey support management decisions?

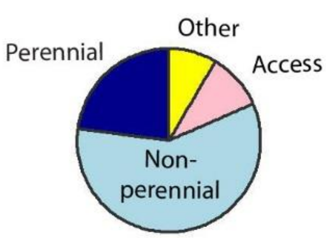
Evaluate steps to protect healthy streams and improve unhealthy streams. Given the small extent of healthy perennial stream-miles in the southern California, protecting such streams may be a priority for resource managers. Additionally, the relatively large extent of stream-miles in poor condition suggests that managers will need to prioritize actions to address stressors affecting unhealthy streams. Prioritization should focus on likelihood of success, achievability of objectives, breadth of impact, and costs associated with management activities, as well as local objectives and needs for each waterbody. Although most of the actions required will be site-specific, a regionally coordinated approach will aid in priority ranking and enable leveraging of efforts across sites or watersheds.

Use regional context in site-specific evaluations. The primary application of survey data is to provide context in evaluating site-specific questions. Comparing the condition of a specific site to conditions at sites with similar land use within the region may provide more useful benchmarks for management objectives than comparison to reference sites, which may not provide an achievable management objective.

Use survey data in causal assessments to identify candidate stressors. Because of the breadth of information collected at each site, the comparability of methods used, and the diversity of sites sampled, data from this survey are well suited to causal assessment applications. With some investment in tool development, regional watershed managers will be able to overcome the data limitations (such as difficulties in identifying comparison sites with information on stressors) that often hinder effective causal assessments.

Recommendations for future monitoring

Although this survey successfully produced preliminary answers to key questions, important knowledge gaps remain. Continuing the survey with modifications will address these gaps.



Include stream types that were previously excluded from the survey. The chief limitation of this survey is that it was restricted to perennial, wadeable streams, 2nd order and higher. The condition of nonperennial and headwater streams represents the largest gap in our regional assessment. Perennial streams account for only 25% of stream-miles in the region as a whole, and as little as 5% in certain watersheds; this variation is caused by both natural factors (such as climate) and land use. Because perennial and higher-order streams are more abundant in developed regions, it is likely that the surveyed portion of the region is in worse condition than the region as a whole. Expanding the survey to include assessment of nonperennial streams (approximately 59% of stream-miles in the region), and exploring ways to map them will help fill these knowledge gaps. Existing assessment tools may be appropriate to assess condition of nonperennial streams, and new tools should be developed as needed.

Improve trend detection through site revisits. Probabilistic sites that are revisited for several years can be used to estimate the extent of improving, degrading, or stable streams in the region. Additionally, management practices associated with changes in conditions can be identified.

Use survey data and special studies to support causal assessments and investigate high-priority stressors. Stressor prioritizations are strictly associative and cannot identify with certainty causal relationships between stressors and biological condition. In some cases, stressors that were identified as high priority (e.g., nutrients) might not directly affect biological condition. Instead, the high risk may reflect a correlation with an unmeasured stressor. The frequent co-occurrence of multiple stressors can make it difficult to disentangle the relationships between individual stressors and biological condition. The SMC can address these limitations in several ways:

- Analyze existing data to explore the diagnostic potential of biological indicators to identify specific stressors.
- Enhance the stream survey with new indicators related to habitat degradation (e.g., hydromodification indicators) or nutrient enrichment (e.g., continuous water quality loggers, algae biomass), or other stressors of emerging concern (e.g., sediment pyrethroids).
- Conduct special studies to distinguish biological constraints imposed by habitat degradation, channel engineering, water chemistry, and natural factors.

Survey Overview

This survey provides the best estimate of the extent of perennial (e.g., Big Tujunga Creek, upper photo) and nonperennial streams (e.g., San Juan Creek, lower photo) in the South Coast region.

Introduction

Southern California's coastal watersheds contain important aquatic resources that support a variety of ecological functions and environmental values. Comprising over 7,000 stream-kilometers, both humans and wildlife depend on these watersheds for habitat, drinking water, agriculture, and industrial uses. In order to assess the health of streams in these watersheds, the Stormwater Monitoring Coalition (SMC), a coalition of multiple state, federal, and local agencies, initiated a regional monitoring program in 2009. Using multiple indicators of ecological health, including benthic macroinvertebrates, benthic algae, riparian wetland condition, water chemistry, water column toxicity, and physical habitat, the SMC has led the first comprehensive assessment of southern California's watersheds based on a probabilistic survey design. Through the re-allocation of permit-required monitoring efforts, the SMC has developed a cooperative sampling program that is efficient and cost-effective for participants. This report represents a summary of data collected in the first five years of the SMC's stream survey. Data from previous surveys, such as the Environmental Protection Agency (EPA) Environmental Monitoring and Assessment Program (EMAP) and California's Perennial Stream Assessment (PSA), are included as well.

The SMC monitoring program was designed to address three main questions:

1) What is the biological condition of perennial streams in the region?
2) What stressors are associated with poor condition?
3) Are conditions changing over time?

The first question is addressed by estimating the extent of biologically intact streams, as determined by key biological indicators. The second question is addressed by estimating the extent of streams with stressors above key thresholds, and by associating stress levels with biological indicators through correlation and relative risk analyses (Van Sickle *et al.* 2006). The third question is addressed by comparing condition across years of the survey.

Regional assessments provide critical information to complement site-specific monitoring at sites of interest. Regional surveys that use a probabilistic design provide statistically valid and unbiased assessments of large geographic areas (Gibson *et al.* 1996). Crucially, regional assessments provide context to site-specific problems and allow sites to be prioritized for protection or restoration (Barbour *et al.* 1996). Furthermore, regional assessments provide a comprehensive perspective on reference conditions (Reynoldson *et al.* 1997). Although regional programs do not replace the need for monitoring at sites of interest (such as below discharges or within sensitive wildlife areas), the context provided by a regional assessment is essential for effective watershed management (Barbour *et al.* 1996, Gibson *et al.* 1996).

Methods

Study Area

Coastal southern California (i.e., the South Coast) is a semi-arid region with a Mediterranean climate, which experiences nearly all of its precipitation as rainfall during winter months. Lower elevations are characterized by chaparral, oak woodlands, grasslands, and coastal sage scrub. The region is bordered by the Transverse Ranges to the North, and the Peninsular Ranges to the East, and continues to the Mexican border to the South. Both Transverse and Peninsular ranges contain peaks that exceed 10,000 feet and regularly experience snow, although contributions to stream flow are limited. Much of the higher elevations are undeveloped and remain protected in a network of national, state, and county parks and forests. The lower elevations have been largely urbanized or converted to agriculture. Wildfires and drought are frequent in the region, with extensive fires occurring in 2007, 2009, and 2013 throughout much of the area. By area, the overall region is 59% undeveloped open space, 28% urban, and 13% agricultural (National Oceanic and Atmospheric Administration (NOAA) 2001).

Survey Design

The target population of the survey was defined as perennial, wadeable second-order and higher streams located in the six southern California counties draining into the Southern California Bight. The study area was divided into fifteen management units (hereafter referred to as watersheds) based on a combination of hydrologic and political boundaries (Table S-1, Figure S-1). The National Hydrography Dataset Plus stream network (NHD Plus; US Geological Survey and US Environmental Protection Agency 2005) was used as the sample frame. Stream segments in the NHD Plus typically represent lengths of streams between two confluences, although particularly long reaches are often split into shorter lengths. In order to assign land-use to each segment of the NHD Plus frame, a 500-m buffer was drawn around each stream segment and overlain in a GIS onto a landcover layer (NOAA 2001). If the buffer was more than 75% natural or open land, the segment was considered open space; if not, it was considered urban or agricultural, depending on which land use was relatively more dominant. Very short segments were occasionally hand corrected if the buffers were too small to adequately capture the adjacent land use; these corrections were most typically used for segments representing individual channels in complex braided systems, such as the mainstem of the Santa Clara River.

The study employed the “master list” approach to integrate sampling efforts by multiple agencies and to facilitate collaboration with other monitoring programs (Larsen *et al.* 2008). A master list was generated, containing over 50,000 sites randomly distributed across the entire stream network using a spatially balanced generalized random-tessellation design (Stevens and Olsen 2004). Sites were then assigned to a watershed using a geographic information system (GIS). Sites were attributed with Strahler stream order from the NHD Plus dataset, and with land use based on the designation of the stream segment, as described above. Sites were then attributed with watershed, stream order, and land-use of the corresponding stream segment of the sample frame. First order streams were excluded from the survey, because these sites typically have a higher rejection rate based on nonperenniality or inaccessibility in mountainous regions. A target sample of 30 sites was selected from each watershed, with heavier representation in relatively uncommon strata (e.g., agricultural streams) to improve balance among the sampled stream types. Large oversamples (ranging from 5x to 20x) were selected as well because of high rejection rates in certain strata. Sites in the sample draw and oversamples were distributed to field crews for evaluation for sampling suitability.

Sites were evaluated for sampling using both desktop and field reconnaissance. Field crews attempted to locate a reach suitable for sampling within 300 m of the target coordinates. Sites with no nearby suitable reaches were rejected for sampling. Reasons for rejection included nonperenniality (see box below), inaccessibility (defined as sites that cannot be safely reached and sampled within one day), refusal or lack of response from landowners, map errors (e.g., no channel near the target coordinates), nonwadeability (i.e., >1 m deep for at least 50% of the reach) and inappropriate waterbody types (e.g., tidally influenced, impounded, etc.). Sites with temporary accessibility or permission issues (e.g., road closures, late responses from landowners) were re-evaluated for sampling in subsequent years.

Defining and Determining Perennial Streams

Perennial streams were defined as those with continuous flow that lasts until the end of the hydrologic year (i.e., September 30) in most years. Determining if a site met these criteria required that field crews find the best available data, including stream gauges, field indicators, historical imagery, consultation with local experts, and best professional judgment. Although all reasonable efforts were made to confirm the perenniality of the sampled sites, it is likely that some of them do not meet the survey's criteria for perennial streams during the years of the study. Therefore, the survey reflects the condition of a mixture in unknown proportions of perennial and long-lasting nonperennial streams. Development of an objective tool to characterize hydrologic regimes remains a priority research area for the SMC.

Sampling Methods

Biological Indicators

Benthic Macroinvertebrates

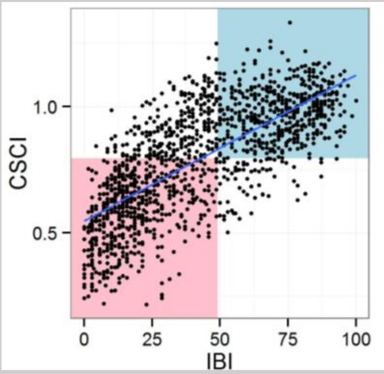
Benthic macroinvertebrates were collected using protocols described by Ode (2007). At each transect established for physical habitat sampling, a sample was collected using a D-frame kicknet at 25, 50, or 75% of the stream width. A total of 11 ft^2 (~1.0 m^2) of streambed was sampled. This method was identical to the Reach-Wide Benthos method used by EMAP (Peck *et al.* 2006). However, in low-gradient streams (i.e., gradient <1%), sampling locations were adjusted to 0, 50, and 100% of the stream width, because traditional sampling methods fail to capture sufficient organisms for bioassessment indices in these types of streams (Mazor *et al.* 2010). Benthic macroinvertebrates were collected and preserved in 95% ethanol (final concentration 70%), and sent to one of five labs for identification. At all labs, a target number of at least 600 organisms were removed from each sample and identified to the highest taxonomic resolution that can be consistently achieved (i.e., SAFIT Level 2 in Richards and Rogers 2011); in general, most taxa were identified to species and Chironomidae (i.e., midges) were identified to genus. Benthic macroinvertebrate data was used to calculate the California Stream Condition Index (CSCI; Mazor et al. In Press). Samples from streams in reference condition are expected to have a mean CSCI score of 1.

CSCI vs. IBI

Like the Southern and Central California Index of Biotic Integrity (IBI), the CSCI was designed to measure the biological condition of streams, as indicated by benthic macroinvertebrate assemblage structure. The CSCI characterizes benthic macroinvertebrate assemblage structure in two ways: 1) As the ratio of observed-to-expected taxa (an O/E index), and 2) as a multi-metric index (MMI), where biological metrics related to important ecological attributes (e.g., number of sensitive taxa) are compared with expected values. Both components are compared to expectations that vary from site to site, and these expectations are derived from reference sites in similar environmental settings.

The CSCI was developed specifically to address some of the shortcomings of traditional indices like the IBI and provides a better measure of stream health than its predecessor because of two key features. First, the CSCI was developed with a much larger, more representative data set. For example, 473 reference sites were used to calibrate the CSCI (including 27 from lower elevation South Coast xeric sites), versus 88 for the IBI (of which only 9 were from South Coast xeric regions). More importantly, the CSCI sets biological benchmarks for a site based on its environmental setting (determined by environmental factors, like climate, geology, watershed area, and elevation) whereas the IBI makes minimal adjustments for natural environmental influences on stream communities.

Overall, the CSCI and IBI have similar performance, and samples that score high for one index usually score high for the other (Pearson's $r^2 = 0.54$). In general, the CSCI is more accurate, and is less likely than the IBI to give false indications of nonreference condition. However, it is also less sensitive, and is less likely to indicate nonreference conditions at severely stressed sites. If a threshold based on the 10^{th} percentile of reference sites is applied to both indices (i.e., 0.79 for the CSCI and 49 for the IBI), approximately one-third of streams below the IBI threshold would be above the CSCI threshold; in contrast, only 2% of streams below the CSCI threshold would be above the IBI threshold.



Correlation between IBI and CSCI scores for sites in southern California. The pink area indicates sites where both indices suggest likely altered biological condition (i.e., Class 3 and 4), and the blue area indicates sites where both indices suggest intact or possibly altered biological condition (i.e., Class 1 and 2). The blue line represents a linear regression between the two indices.

Benthic Algae

Benthic algae samples were collected using the protocols of Fetscher *et al.* (2009), approximately 1 foot upstream of each location where benthic macroinvertebrates were collected. Diatom samples were preserved in formalin, and soft algae samples were preserved in glutaraldehyde. Unpreserved, qualitative soft algae samples were also collected to produce fruiting bodies that facilitate identification of soft algae species. Benthic algae samples were identified to the best taxonomic resolution possible, which was typically species. Benthic algae was assessed using two indices from Fetscher *et al.* (2014): a soft algae index (S2), and a diatom index (D18). Calculations were completed using custom scripts in the statistical software R. Samples from streams in reference condition are expected to have a mean D18 score of 79 and a mean S2 score of 69. Although these indices are not "predictive" like the CSCI score, little bias from natural gradients was evident at reference sites (Fetscher *et al.*2014).

Riparian Wetlands

Riparian wetland condition was assessed using the California Rapid Assessment Method (CRAM; Collins *et al.* 2008). Briefly, the CRAM method assesses four attributes of wetland condition: buffer and landscape, hydrologic connectivity, physical structure, and biotic structure. Each of these attributes is comprised of a number of metrics and submetrics that are evaluated in the field for a prescribed assessment area. Streams in reference condition are expected to have a mean CRAM score of 84.

Water Chemistry

Field crews measured pH, specific conductance, dissolved oxygen, salinity, and alkalinity at each site visit using digital field sensors (or by collecting samples for lab analyses, where appropriate). In addition, samples of stream water were collected for measurements of 36 different analytes, including: total suspended solids, total hardness (as $CaCO_3$), silica, sulfate and other major ions, nutrients, dissolved and total metals, and pyrethroid pesticides. Analytical methods and quality assurance protocols are described in SWAMP QAT (2008).

Toxicity

At each site, ~4 L of water were collected for toxicity assays, primarily using the water flea *Ceriodaphnia dubia*. Six to eight day exposures to undiluted field-collected stream water were conducted, and both survival (acute toxicity as percent mortality) and reproduction (chronic toxicity as young per female) endpoints were recorded. In samples with specific conductivity ≥2500 µS/cm, a 10-day survival assay using the amphipod *Hyalella azteca* was used instead, with no reproductive endpoint (USEPA 2002, SWAMP QAT 2008).

Physical Habitat

At each site, physical habitat was evaluated using a physical habitat assessment as specified in Ode (2007) and Fetscher *et al.* (2009), which were adapted from EMAP (Peck *et al.* 2006). Briefly, a 150-m reach (250-m for streams over 10 m wide) was divided into 11 equidistant transects, with 10 inter-transects located halfway between them. At each transect, the following parameters were measured: bank dimensions, wetted width, water depth in five locations, substrate size, cobble embeddedness, bank stability, microalgae thickness, presence of coarse particulate organic matter, presence of attached or unattached macroalgae, presence of macrophytes, riparian vegetation, instream habitat complexity, canopy cover using a densiometer, human influence, and flow habitats. A subset of these variables were measured at each inter-transect as well. The slope of the water surface was measured across the entire reach at each site. Metrics based on physical habitat data were calculated using custom scripts in R, based on those presented in Kaufmann *et al.* (1999).

Challenges in Assessing Physical Habitat

Although many studies point to a crucial role for physical habitat in supporting healthy streams, assessing the condition of physical habitat remains a challenge for bioassessments. There are four parts to this challenge: 1) measuring the right variables, 2) calculating meaningful metrics from these variables, 3) comparing these metrics to benchmarks that are appropriate for the environmental setting of a site, and 4) ensuring that the metrics are comprehensive enough to characterize important aspects of habitat degradation. To some extent, the first two problems have been addressed. The protocol developed by SWAMP, based on methods developed by the EPA (Peck *et al.* 2006), encompasses over 1000 individual measurements per site, and these measurements are converted into more than 150 metrics that characterize the physical habitat, again based on earlier efforts of the EPA (Kaufmann *et al.* 1999). However, most of these metrics vary widely among reference sites, based on environmental factors like climate and watershed size. Predictive models to set reference-based expectations for physical habitat metrics are in development, but are not yet available. Once such models are developed, a remaining challenge will be to select which metrics (and in which combinations) are most useful in characterizing the overall condition of the physical habitat of a site.

Landscape Variables

Landscape variables were calculated for three purposes: CSCI calculation (see Mazor *et al.* In review), reference site screening (see Ode *et al.* In review), and biological relationships. Using a GIS, watersheds were delineated for each site from 30-m digital elevation models (USGS 1999), and visually corrected to reflect local conditions. For sites draining ambiguous watersheds with minimal topography, delineations were modified using CALWATER boundaries (California Department of Forestry and Fire Protection 2004) or by consulting local experts. Watersheds were clipped at 5 km and 1 km to evaluate local conditions, creating a total of three scales (abbreviated as WS, 5k, and 1k). A fourth scale (i.e., point), based only on the site location, was used to calculate distance-based metrics. These delineations were then used to calculate metrics from source layers relating to landcover (NOAA 2001), transportation (CDFG custom roads layer, P. Ode, unpublished data), geology (J. Olson and C. Hawkins, unpublished data), and hydrology (National Inventory of Dams and NHD Plus). For sites sampled in 2013, only variables related to the CSCI were calculated.

Summary of Data from Other Surveys

Data from other surveys were included in this report, where possible. In order to be included, these surveys had to meet the several criteria: 1) benthic macroinvertebrates were collected using similar protocols (e.g., EMAP), 2) benthic macroinvertebrates were identified to equivalent taxonomic resolution, 3) survey design documentation (including stratifications) and site evaluation data were available, and 4) compatible sample frames were used for survey design (specifically, the NHD Plus or its predecessor RF3). These surveys are summarized in Table S-2. Note that some sites, although selected for sampling for a probabilistic survey, were revisited under other programs (such as reference sampling, fire studies, or other targeted designs), and these data were included in the current assessment as well. With few exceptions, limited data types (generally, benthic macroinvertebrates and physical habitat) were collected for these surveys.

Climate Data

Monthly rainfalls for stations throughout the region were downloaded from The National Oceanic and Atmospheric Administration's California and Nevada River Forecast Center (www.cnrfc.noaa.gov/rainfall_data.php). Annual totals were then calculated and plotted to evaluate the conditions during the study period relative to longer term trends. Three representative stations were selected for plotting (i.e., downtown Los Angeles, Big Bear Lake, and Lindbergh Field).

Data Analysis

Weighted Magnitudes and Extent Estimates

Adjusted sample weights were calculated for each site. Because multiple surveys with different designs were included in analysis, weights needed to be recalculated for each site. Stratification approaches from all surveys were combined to create "cross-strata" in which all evaluated sites have an equal probability of being sampled. Adjusted weights were recalculated as the total stream length within each strata, divided by the number of sites evaluated in that stratum. Strata with no evaluations were excluded from analysis. Because these strata comprised less than 2% of the total stream length, these exclusions are unlikely to affect condition estimates. These weights were used to estimate distribution points for selected variables and extents (e.g., % of stream-length in classes of interest) using the Horvitz-Thompson estimator (Horvitz-Thompson 1952). These estimates were calculated for reporting units of interest, including watersheds, land use classes, and (for trend estimates) years. Confidence intervals (CIs) were based on local neighborhood variance estimators (Stevens and Olsen 2004). All calculations were conducted using the spsurvey package (Kincaid and Olsen 2013) in R version 3.0.3 (R Core Team 2012).

Extent Estimates

When surveys use a probabilistic design, the data they produce can be used to make inferences about the region as a whole, and not just about sampled sites. Therefore, statements about the extent of perennial wadeable streams, or about the average CSCI score in a watershed can be made. Probabilistic surveys provide context about ambient condition, which can be used to compare against sites of interest.

The key benefit of a probabilistic survey is its ability to estimate the true extent of a resource of interest, such as perennial, wadeable streams. Sites sampled under a targeted design provide valuable information about local conditions, but cannot be used to estimate the condition of the region as a whole. Because targeted studies are typically designed to assess known impacts (e.g., downstream of discharges), the sites may be in worse condition than the average site in the region; therefore, estimates of regional condition from targeted sites may be biased.

When sites are sampled according to a probabilistic design, measurements represent not just local conditions, but also reflect conditions of a much larger population. The condition of each probabilistic site therefore contributes to condition estimates of the region as a whole. The weight (i.e., the contribution to regional estimates) of each site varies; sites in large, sparsely sampled regions (e.g., open streams) make a larger contribution to regional estimates than sites in small or densely sampled regions (e.g., agricultural streams).

Results

A total of 760 probabilistic sites were sampled in the South Coast region, of which 515 were sampled by the SMC or affiliated programs (Table S-2). To attain this sample size, 4330 unique sites were evaluated, yielding a rejection rate of 82%. The most common cause for rejecting a site was nonperenniality (75% of rejected sites), followed by physical barriers (9% of rejected sites). Determinations of nonperenniality were made during both office and field reconnaissance. Other causes for rejection (e.g., map errors, inappropriate waterbody types, nonwadeability) were infrequently encountered (≤5% of rejected sites; Table S-3; Figure S-2).

Analysis of rejected sites indicated large differences in the extent of perennial streams by watershed and land use. For example, perennial streams made up 53% of stream-miles in the Los Angeles watershed, but only 6% of the San Jacinto watershed (median watershed extent: 26%). Land-use was strongly associated with perenniality, as 35% of urban stream-length, but 12% of agricultural stream-length and 16% of open stream-length were perennial (Figures S-2, S-3, S-4).

Overall, the survey occurred in a drier than normal period. Rainfall during 2011 was slightly above average, although most other years were well below normal. Notably, the survey occurred shortly after one of the driest years on record (i.e., 2007), when even the rainier weather stations (e.g., Big Bear Lake) reported extremely low precipitation (Figure S-5).

Discussion

Perennial wadeable streams are a small component of the region, and protecting this limited resource may be a high priority for watershed managers, particularly because of their importance to a variety of beneficial uses (such as fisheries, wildlife, and swimming). At the same time, the need to expand attention to nonperennial streams is apparent: A comprehensive assessment of the coastal watersheds of southern California should not exclude the large extent of nonperennial streams. Ongoing research in the region addresses the question of whether the condition indices used in this survey are valid in nonperennial streams. However, it is likely that assessment tools currently available to watershed managers are adequate to include at least some portion of nonperennial streams in future surveys.

The observed extents of perennial streams in urban and agricultural areas are probably elevated by imported water sources (either as wastewater effluent or as runoff). Because nonperennial streams are so extensive in undeveloped areas, it is likely that this survey excludes many of the healthiest, least disturbed streams in the region. Therefore, although this survey provides an unbiased assessment of the perennial portion of southern California streams, extrapolation to the nonperennial portion may lead to incorrect conclusions about the health of the region as a whole.

Climatic trends may have also influenced the extent and location of perennial streams. Frequently, field crews were unable to sample reaches that were historically perennial, suggesting that long-term drought or changes in water management may have converted some perennial streams to nonperennial. The variability of flow regimes in southern California streams has been documented in special studies commissioned by the SMC (e.g., Mazor *et al.* 2014), and this variability underscores the need for a flexible approach towards characterizing stream hydrology.

The widespread conversion of streams from nonperennial to perennial (and vice versa) presents a question about setting appropriate ecological objectives. Should a converted stream be compared to perennial reference streams? Or is it more appropriate to compare them to their historical conditions? This survey used the former approach, although in certain applications, such as setting restoration objectives, different goals may be appropriate.

However objectives are set for streams with altered hydrology, managing flows may be an important tool in supporting their ecological health. The causes of elevated water flows were not investigated in this survey. In major tributaries and mainstems of large rivers, elevated flows may be driven by effluent from treatment plants managed by sanitation districts. In smaller streams, runoff may be an important driver, where flood control agencies manage stream flows. Diversions and groundwater extraction are particularly important in streams in agricultural areas. Therefore, if flow regime management needs to change to support ecological health, coordination among several agencies working under different permits may be required.

Table S-1. Characteristics of each watershed.

Watersheds	Stream Order	Area (km^2)	Total Stream Length (km)	Land Use (%)		
				Open	Agricultural	Urban
Ventura	6	642	236	68	15	17
Santa Clara	7	4327	1429	81	14	6
Calleguas	5	891	315	28	35	36
Santa Monica Bay	4	1171	200	73	2	25
Los Angeles	5	2160	519	41	1	59
San Gabriel	5	1758	487	50	0	50
Santa Ana River	6	7092	1708	49	15	36
–Lower Santa Ana	6	1253	298	36	10	53
–Middle Santa Ana	6	2135	519	38	14	48
–Upper Santa Ana	5	1721	523	64	12	24
–San Jacinto	4	1984	367	55	24	21
San Juan	4	1019	337	66	5	29
Northern San Diego	6	3640	1055	58	28	14
Central San Diego	5	1725	430	38	12	51
Mission Bay/San Diego River	5	1270	322	64	4	32
Southern San Diego	5	2355	535	80	6	14
Entire Region	7	28051	7574	59	13	28

Table S-2. Probabilistic surveys included in the study. Note that the SMC program includes sites sampled under nested programs that used the same master sample draw, such as the San Gabriel River Regional Monitoring Program, the Los Angeles Watershed Monitoring Program, and Region 4 Probabilistic Sampling; sites from these surveys were included only if they were part of the SMC's target population of second-order or higher perennial, wadeable streams.

Survey	Years	Sites
Environmental Monitoring and Assessment Program (EMAP)	2000 to 2003	42
California Monitoring and Assessment Program (CMAP)	2004 to 2007	12
National Rivers and Streams Assessment (NRSA)	2009 and 2013	1
Perennial Streams Assessment (PSA)	2008	11
Stormwater Monitoring Coalition (SMC)	2008 through 2013	515
Region 8 Trend Monitoring (R8T)	2006 through 2013	102

Table S-3. Extent (in percent stream-miles) of perennial and non-perennial streams by subpopulation.

Subpopulation	Perennial, sampled (n sampled)	Perennial, not sampled	Rejected		
			Nonperennial	Physical Barrier	Other
South Coast	20.7 (682)	2.3	58.5	10.0	8.4
Land Use					
Agricultural	11.9 (92)	4.0	70.7	1.2	12.3
Open	15.9 (306)	1.4	61.1	16.3	5.3
Urban	35.3 (284)	3.4	47.2	0.8	13.4
Watershed					
Region 4					
Ventura	25.3 (37)	0.8	62.6	7.1	4.3
Santa Clara	16.2 (94)	2.1	55.2	24.0	2.6
Calleguas	30.2 (38)	6.0	48.2	3.0	12.6
Santa Monica Bay	23.6 (72)	2.1	52.7	9.6	11.9
Los Angeles	47.1 (44)	5.6	25.3	13.2	8.8
San Gabriel	43.7 (39)	1.1	23.0	16.6	15.5
Region 8					
Lower Santa Ana	16.3 (45)	3.1	46.6	8.2	25.8
Middle Santa Ana	13.1 (57)	4.0	61.3	4.7	16.9
Upper Santa Ana	25.1 (67)	2.8	44.6	22.2	5.3
San Jacinto	5.3 (28)	0.7	77.5	8.6	7.9
Region 9					
San Juan	27.5 (30)	1.0	68.0	1.1	2.5
Northern San Diego	7.1 (36)	0.7	81.0	1.5	9.6
Central San Diego	37.1 (35)	3.1	54.3	0.5	5.2
Mission Bay and San Diego River	14.5 (29)	2.8	74.6	1.3	6.8
Southern San Diego	8.3 (31)	0.8	83.7	0.8	6.3

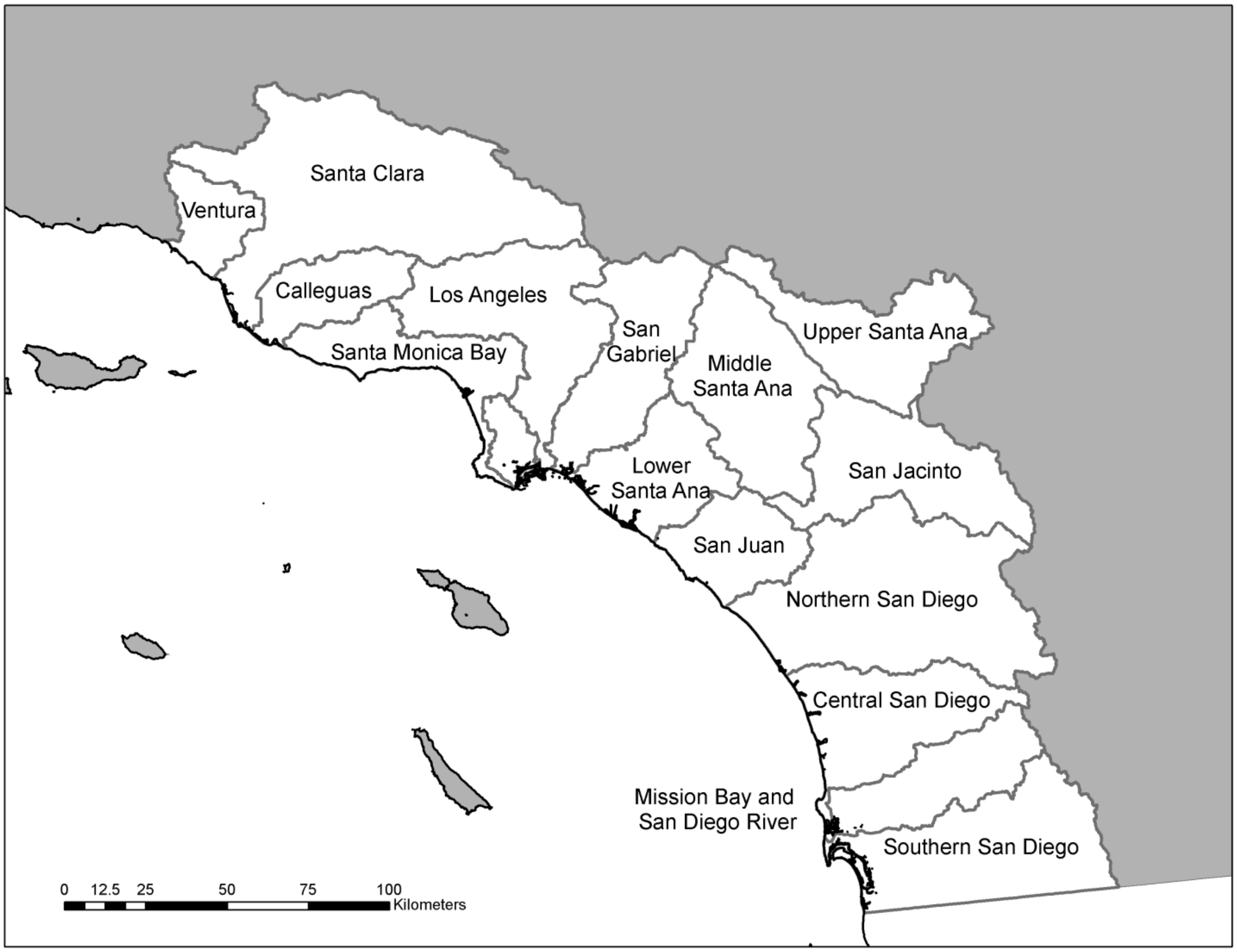


Figure S-1. Major watersheds in the South Coast survey area.

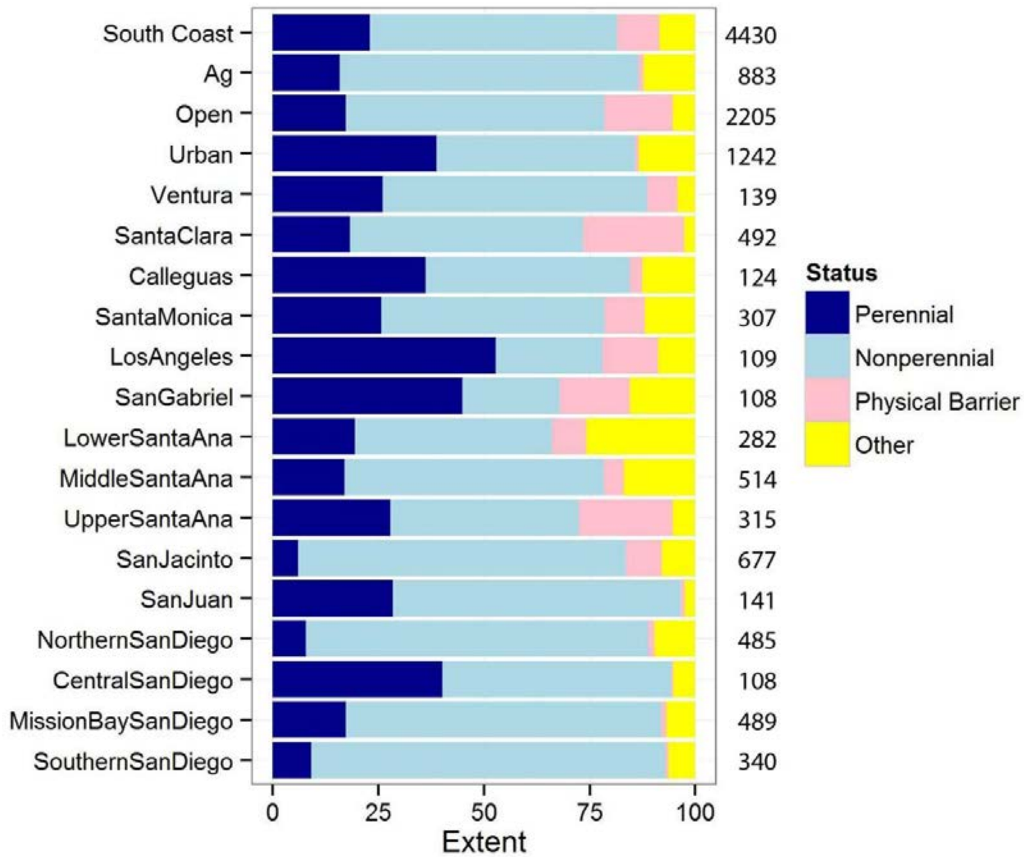


Figure S-2. Site evaluation results by watershed or land use. Numbers to the right of each bar represent the total number of sites evaluated for inclusion in the SMC and other survey.

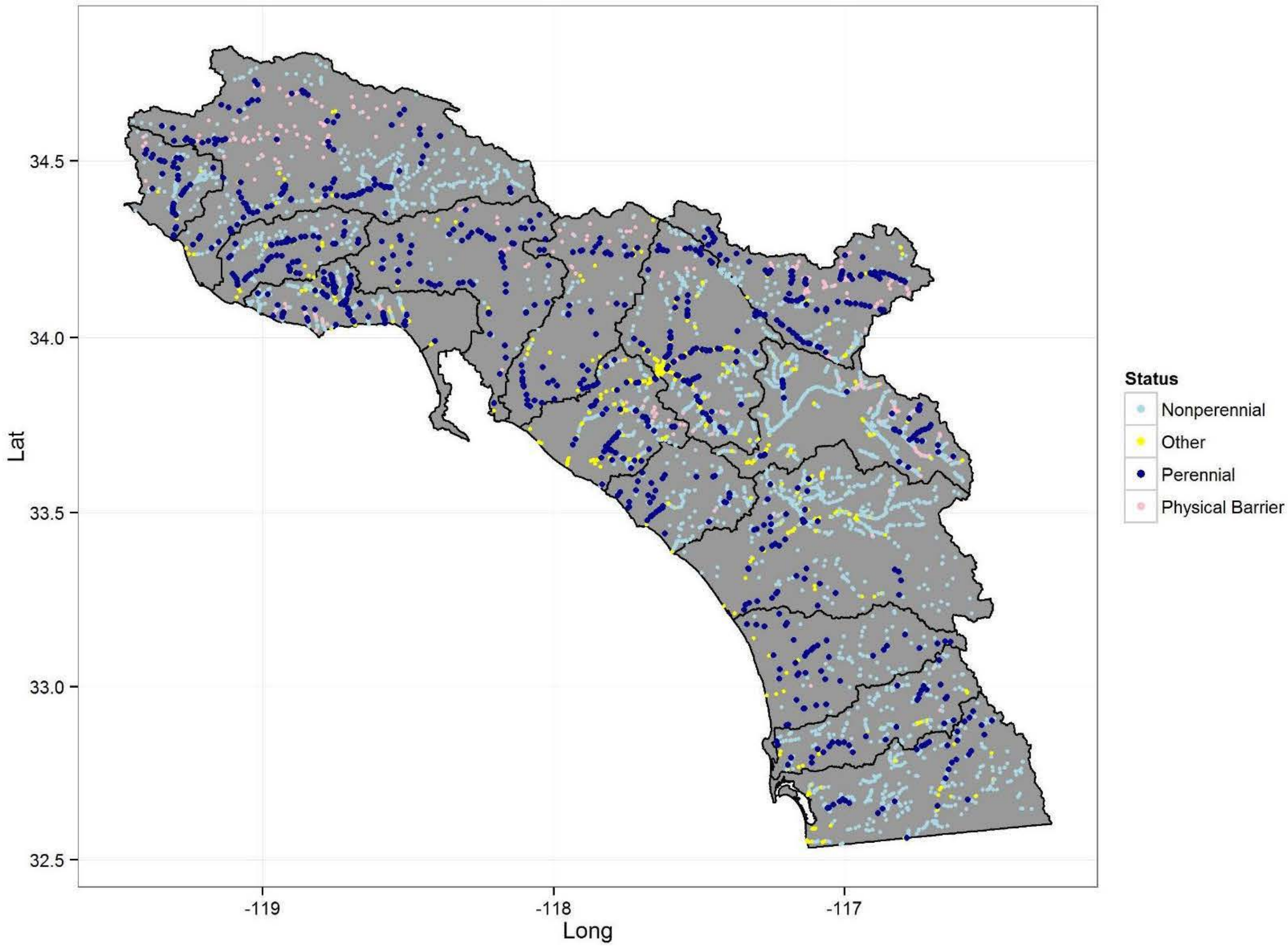


Figure S-3. Map of site evaluation results.

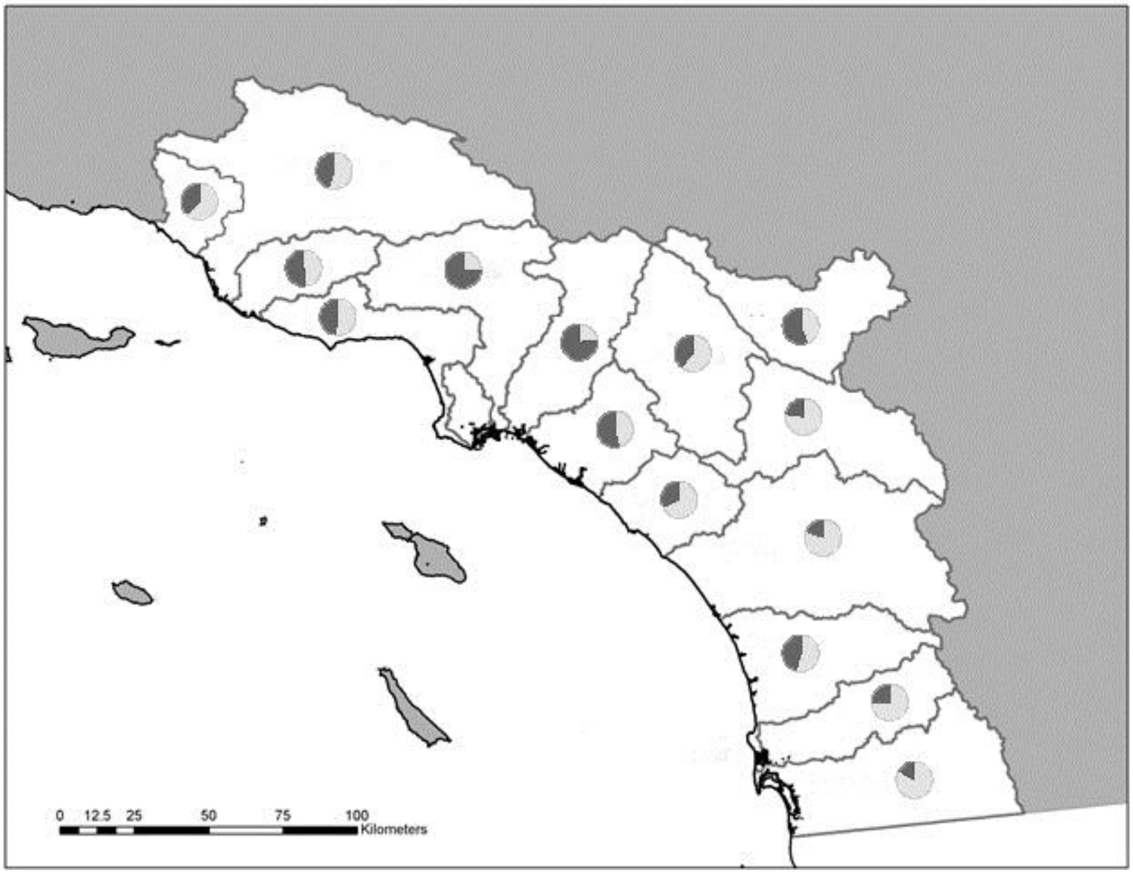


Figure S-4. Percent of nonperennial stream-miles (shown in light gray) for each watershed.

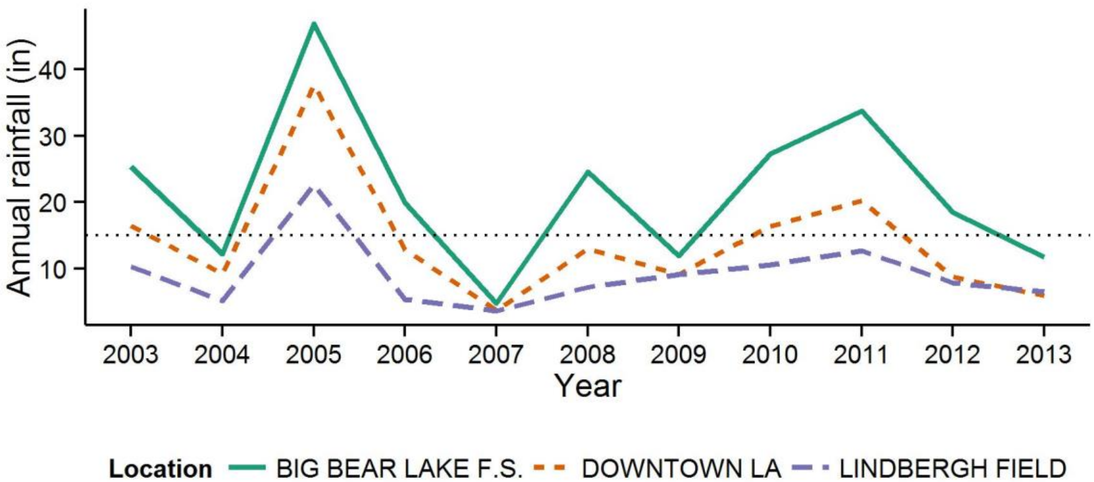


Figure S-5. Annual precipitation at three weather stations in the South Coast. The horizontal line reflects the average for downtown Los Angeles between 1877 and 2012.

QUESTION 1: WHAT IS THE BIOLOGICAL CONDITION OF PERENNIAL STREAMS IN THE SOUTH COAST REGION?

Healthy perennial streams, like this site on the North Fork of the San Jacinto River, are a scarce resource in the South Coast region.

Introduction

Surveys of ambient biological condition provide essential context for watershed management. At larger geographic scales, ambient surveys allow watershed managers to identify regional priorities. At local scales, ambient surveys allow managers to compare sites of interest to typical ranges in the region. This context informs decisions about which sites need protection or rehabilitation.

The biological condition of perennial streams was assessed by sampling four key biological indicators (i.e., benthic macroinvertebrates, diatoms, soft algae, and CRAM) at sites throughout the region, and comparing them to thresholds benchmarked to the distribution of scores at reference sites. These biological indicators provide a direct measurement of ecological health, and are an effective tool to determine if streams are supporting aquatic life or other beneficial uses. Additionally, their ability to integrate multiple stressors across both time and space make them a superior measure of biological condition to direct measures of stressors.

Methods

Data Collection

Data were collected as described in the Survey Overview.

Data Aggregation

Where multiple biological samples were collected at a single site within a year, data were aggregated as the maximum value within a site (with the assumption that index scores may be spuriously low, but not spuriously high). Multi-year mean values for each site were then calculated from these aggregated values if sites were revisited in multiple years. Missing values were ignored for all relevant analyses, where appropriate.

Thresholds

Biological indicators were compared to the 30th, 10th, and 1st percentile of reference sites (Table 1-1); these percentiles correspond to different probabilities that a score is from a site in reference condition. This approach creates four biological condition-classes that may be interpreted as indicating a stream's biology is likely intact (Class 1), possibly altered (Class 2), likely altered (Class 3), and very likely altered (Class 4). These percentiles were selected to reflect a range of conditions. Because this approach is consistent across indicators, it is possible to compare results from one index to another. Means and standard deviations were from published sources (CSCI: Mazor et al. In review; algae IBIs: Fetscher *et al.* 2014) or unpublished data (CRAM). Each threshold has an associated error rate; for example, 10% of reference sites are in Class 3 or 4, despite the fact that they are, by definition, intact.

Integrating Multiple Indicators

In order to determine a stream's overall condition, the four biological indicators were evaluated together to provide a comprehensive assessment of ecological health. To be considered intact for multiple indicators, all four indicators need to suggest that a stream is in reference condition. A single indicator below this threshold suggests that a stream is not in reference condition. To maintain an overall error rate of 10%, a site had to have scores above the 2.5^{th} percentile of reference sites for each indicator (Table 1-1).

Weighted Magnitudes and Extent Estimates

Adjusted sample weights were calculated for each site. Because multiple surveys with different designs were included in analysis, weights needed to be recalculated for each site. Stratification approaches from all surveys were combined to create "cross-strata" in which all evaluated sites have an equal probability of being sampled. Adjusted weights were recalculated as the total stream length within each strata, divided by the number of sites evaluated in that stratum. Strata with no evaluations were excluded from analysis. Because these strata comprised less than 2% of the total stream length, these exclusions are unlikely to affect condition estimates. These weights were used to estimate distribution points for selected variables and extents for selected categories using the Horvitz-Thompson estimator (Horvitz-Thompson 1952). These estimates were calculated for reporting units of interest, including watersheds, land use classes, and (for trend estimates) years. Confidence intervals (CIs) were based on local neighborhood variance estimators (Stevens and Olsen 2004). All calculations were conducted using the spsurvey package (Kincaid and Olsen 2013) in R version 3.0.3 (R Core Team 2012).

Results

All data used in this report can be downloaded from
ftp.sccwrp.org/pub/download/SMCReport/SMCDataFor5yearReport.zip.

Benthic Macroinvertebrates

Biological indicators suggested that most stream-kilometers in the survey's target population (i.e., perennial wadeable streams in southern coastal California) do not support healthy biology (Table 1-2a to c; Figures 1-1 and 1-2). For example, the mean CSCI score for the region was 0.77 and only 29% of stream-miles were in the top biological condition class for this indicator. Of the two components of the CSCI, the pMMI (which measures ecological structure) was more sensitive; the pMMI indicated that only 22% of South Coast stream-miles were in Class 1, whereas the O/E (which measures taxonomic completeness) indicated 46% were in Class 1.

The CSCI indicated that open streams were in better condition than agricultural streams, which were in turn better than urban streams. In fact, at open sites, mean CSCI scores were close to reference (i.e., 0.93), and only 5% of open stream-miles was in Class 4 (i.e., the worst condition class). In contrast, 31% of agricultural streams and 58% of urban streams were in Class 4.

Although this ranking of land use classes was evident with both components of the CSCI, the O/E generally categorized agricultural streams as intermediate between open and urban classes, whereas the difference was small when examined with the pMMI.

The watersheds with the greatest proportion of streams in Class 1 were located, roughly, in the northern and southern ends of the region, while the middle portions of the region had streams in poorer health. For example, the greatest extent of Class 1 stream-miles was located in the Ventura watershed (68%), followed by Southern San Diego (65%). These watersheds, along with the Santa Clara, all had mean CSCI scores greater than 0.9. The smallest extents of Class 1 stream-miles were observed in the Calleguas (9%), Central San Diego (10%), Lower Santa Ana (11%) and Middle Santa Ana (11%) watersheds.

Benthic Algae

In general, the algae indices showed similar patterns of regional stream condition as the CSCI (Table 1-2d and e; Figures 1-1 and 1-2). For example, the diatom index (D18) showed that 27% of stream-miles were in Class 1, while the soft algae index (S2) showed that 25% were in this class; these numbers are only slightly less than the estimate for the CSCI (i.e., 29%).

In contrast with the CSCI, algae-based indices only weakly differentiated between urban and agricultural streams, and estimated both to be in far worse condition than open streams. For example, D18 rarely identified developed streams as Class 1 (Agricultural: 11%; Urban: 2%). Uniquely, S2 scores were generally lower at agricultural streams (mean: 26) than urban streams (mean: 32). In contrast, mean D18 scores were similar in both urban (43) and agricultural (45) streams.

Although there were some differences among the two algae indices, they both showed that the watersheds in the northern portions of the region had the greatest extent of streams in Class 1. For example, D18 indicated the greatest extent of streams in Class 1 in the Ventura (84%) and Upper Santa Ana (63%,) watersheds, whereas S2 indicated the greatest extent of stream-miles in Class 1 in the Upper Santa Ana (47%) and Santa Clara (46%) watersheds. Depending on the index used, Class 1 streams were rarely or never observed in the Calleguas, Santa Monica Bay, Lower Santa Ana, San Juan, and Central San Diego watersheds.

Riparian Condition

Most streams in southern California did not support healthy riparian communities, as only 30% of stream-miles in the region had CRAM scores in the top condition class (i.e., a CRAM score ≥ 79), and the mean CRAM score (64) was much lower than the reference mean (i.e., 84). However, the extent of stream-miles in Class 1 was greater for individual attributes (e.g., 40% for the landscape and buffer attribute), indicating that different attributes limit overall riparian condition at different sites (Table 1-2f; Figures 1-1 and 1-2).

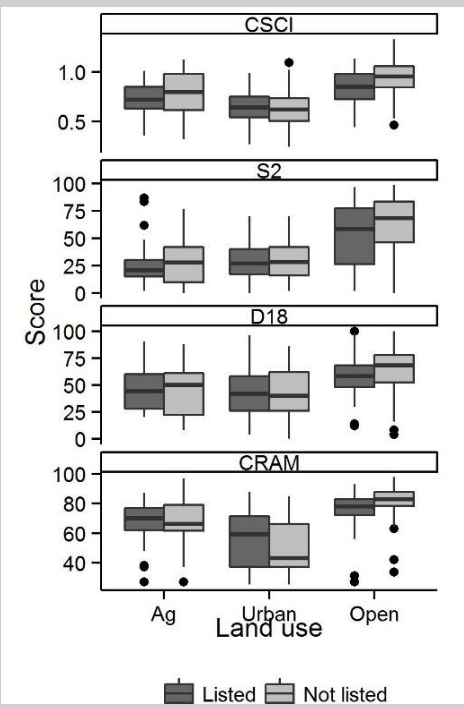
Land use was strongly associated with CRAM scores, even more so than with other indicators. For example, Class 1 CRAM scores were observed at 65% of open stream-miles (mean: 81), but only 20% of agricultural streams (mean: 68) and 7% of urban stream-miles (mean: 51). This contrast was particularly strong at the attribute level (especially the buffer and landscape attribute). For example, hydrologic conditions were in the top class at 57% of open stream-miles, but only 17% of agricultural stream-miles and 17% of urban stream-miles.

Class 1 riparian conditions were observed at the majority of stream-miles within five watersheds that were geographically dispersed across the region, with the greatest extents in the San Jacinto (63%) and Northern San Diego (57%) watersheds, followed by Ventura (54%) and Southern San Diego (52%). Streams with Class 1 riparian condition were scarce in the Calleguas (3%) and Los Angeles (14%) watersheds. Across the four attributes, four watersheds ranked among the worst in terms of the extent of streams in Class 4: Los Angeles, San Gabriel, Lower Santa Ana and Middle Santa Ana. All attributes were in the worst condition class for at least 50% of these watersheds (Table 1-2g to j) with the exception of the biotic structure attribute in the Lower Santa Ana (36% in Class 4).

303(d)-Listed Streams

The State Water Resources Control Board has designated approximately 2000 stream-kilometers in southern California as impaired for water quality pursuant to Section 303(d) of the Clean Water Act. Streams are usually listed as "impaired" due to exceedances of a chemical water quality standard. The potential relationship between designated impairments and instream biological condition was evaluated by comparing biological index scores from streams listed as impaired to streams from comparable land use categories that are not listed. Listed streams were obtained from the State Water Board 303(d) list; in Ventura and Riverside counties, agency staff modified this list by reclassifying listings believed to be unrelated to aquatic life uses (e.g., bacteria) as "not listed" for this analysis.

Land use was more strongly associated with scores than with status on the 303(d) list. For example, scores at urban and agricultural sites were lower than scores at open sites, whether or not the sites were included on the 303(d) list. There was no significant difference in scores between listed and unlisted streams at urban or agricultural sites. Scores at open listed sites were slightly lower than at open unlisted sites; however, this difference was small, and the proportion of Class 3 or 4 sites was no greater at open listed sites than open unlisted sites.



Index scores based on benthic macroinvertebrates (CSCI), soft algae (S2), diatoms (D18) and riparian condition (CRAM) for 303(d)-listed and unlisted streams, by land use.

Condition of Engineered Channels: Exploring options for alternative thresholds

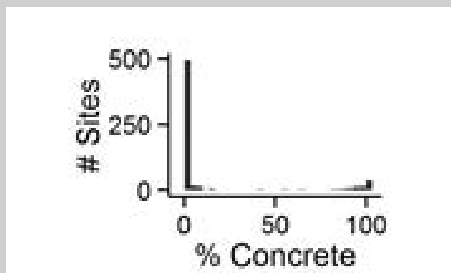
Many of the streams in this survey have been engineered to some degree for flood management purposes, and these engineered features may constrain biological condition. Therefore, we estimated the biological condition of streams with engineered channels relative to those with natural channels. The best condition observed in engineered channels may be a more realistic threshold than a reference-based threshold, assuming that the effects of channel engineering cannot be mitigated. If the best observed condition in engineered channels is substantially below a reference-based threshold, an alternative threshold may be appropriate.

Because consistently derived region-wide maps identifying the location of engineered channels are not available, habitat data was used to classify streams as likely concrete-lined (i.e., at least 5% concrete in the streambed), or likely non-concrete lined (i.e., less than 5% concrete in the streambed). This approach overlooks forms of engineered channels that do not use concrete, such as ungrouted rock, while also misclassifying streams affected by other types of concrete structures, such as road crossings. It also ignores the substantial variation of channel forms in engineered systems, which may affect biological condition. But despite these shortcomings, this approach represents a useful starting point until better data are available about engineered channels.

Overall, approximately 26% of perennial stream-miles were estimated to be concrete-lined. About half of urban streams were concrete lined and 13% of agricultural streams, but only 2% of open streams. Concrete-lined streams comprised a majority of stream-miles in the Los Angeles and San Gabriel watersheds, but none were sampled in the Northern and Southern San Diego watersheds.

Extent of concrete channels in southern California

Subpopulation	Concrete-Lined Channels	
	# sites	% stream-miles
South Coast	130	26
Land use		
Urban	107	53
Open	10	2
Agricultural	13	13
Watershed		
Los Angeles Region		
Ventura	2	4
Santa Clara	3	3
Calleguas	12	29
Santa Monica Bay	13	19
Los Angeles	22	51
San Gabriel	23	69
Santa Ana Region		
Lower Santa Ana	11	26
Middle Santa Ana	22	41
Upper Santa Ana	1	2
San Jacinto	5	19
Northern San Diego		
San Juan	6	24
Northern San Diego	0	0
Mission Bay and San Diego River	6	24
Central San Diego	4	14
Southern San Diego	0	0



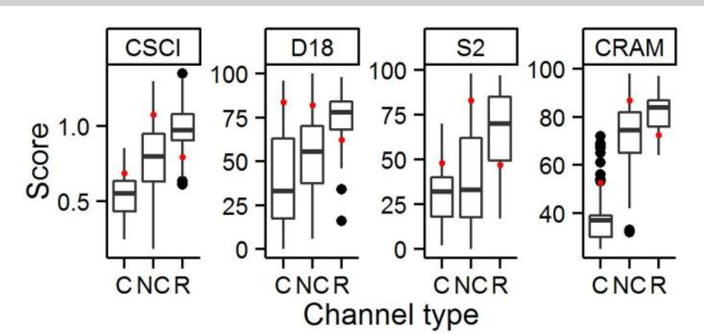
% concrete substrate at each sampled site. Concrete was absent from most sites, but comprised nearly 100% for a small handful of sites. Intermediate values were rarely observed.

Condition of Engineered Channels (Continued)

To investigate the constraints concrete lining imposes on biological condition, sites were divided into three classes: concrete-lined, no concrete, and reference. The range of index scores within each class was examined by creating boxplots. For indices where the 90th percentile of concrete-lined channels is less than the 10th percentile of reference streams, lower thresholds may be appropriate.

In general, scores of all indices were lower in concrete-lined channels than in reference streams, suggesting that these streams were typically in poor condition. For most indices the highest scores in concrete-lined channels were lower than lowest scores observed at reference sites (estimated at the 90th and 10th percentiles, respectively). For example, the 90th percentile of CSCI scores was 0.69 (i.e., "Class 3"), suggesting that an alternative threshold may reflect a more attainable management objective than the 10th percentile of reference sites. Additional data and analyses (particularly on channel type) are needed if alternative thresholds for concrete-lined channels are used for regulatory purposes.

In contrast, this analysis did not support alternative thresholds for algae indices. High scores were frequently observed in concrete-lined channels. In fact, the 90th percentile of D18 scores in concrete-lined channels was 84, which is substantially higher than the threshold based on the 10th percentile of reference sites (i.e., 62). Therefore, it is probable that low D18 and S2 scores in concrete-lined channels are attributable to impacts not directly related to channelization, and may instead be related to water quality impacts.



Distribution of scores at concrete-lined channels (C), nonconcrete-lined channels (NC), and reference streams (R). The red dot represents the 90th percentile of scores of concrete- and nonconcrete-lined channels and the 10th percentile of reference streams.

Options for setting thresholds in concrete-lined channels. A traditional approach is based on the distribution of scores at reference sites, whereas an alternative approach is based on the distribution of scores at concrete-lined channels. These numbers reflect preliminary analyses.

Index	Option 1: Threshold based on reference	Option 2: Threshold based on concrete-lined channels
CSCI	0.79	0.68
D18	62	84
S2	47	48
CRAM	72	53

Multiple indicators

Only 25% of streams-miles in the region were intact for all four indices, and these conditions were almost exclusively observed at streams with undeveloped watersheds (Table 1-3, Figures 1-3 and 1-4). Overall, 60% of open stream-miles were in this category. Streams with index scores above the multi thresholds were absent from the Calleguas watershed and scarce in Santa Monica Bay, Los Angeles, Middle Santa Ana, and Central San Diego watersheds. In contrast, a majority of stream-miles were intact for multiple indicators in the Upper Santa Ana (62%), Southern San Diego (61%), San Jacinto (53%) and Ventura (50%) watersheds.

Most commonly, streams were limited (i.e., below the "multi" threshold) for multiple indicators, and all four indicators were identified as limiting for 15% of stream-miles region-wide (Table 1-3; Figures 1-3 and 1-4). More than a quarter of stream-miles were limited for all indicators in certain watersheds (specifically, Calleguas, Los Angeles, Lower Santa Ana, and San Jacinto watersheds) and in urban streams, but this situation was rare in other watersheds (specifically, Ventura, Upper Santa Ana, Northern San Diego, and Mission Bay and San Diego watersheds), and in open streams. Streams limited for single indicators were more extensive in these open streams, and algae indices (D18, S2, or both) were most commonly the only limiting indicator. For example, 41% of stream-miles in the Northern San Diego and 37% in the Ventura watersheds were limited for D18 or S2, but not CRAM or CSCI.

Discussion

The scarcity of streams with intact biology may prompt managers to evaluate ways to protect these streams, or improve the condition of streams where indicators suggest altered biological condition. The emphasis may vary from protection in one part of the region to rehabilitation in another, depending on local needs and interests. However, many watershed managers in southern California would benefit from a coordinated approach towards prioritizing local objectives, given the extent of streams with altered biology. Uncoordinated efforts to address pervasive challenges have historically met with little success (Bernstein and Schiff 2002).

Multiple indicators proved valuable for several reasons. 1) Redundancy improves precision and guards against incorrect conclusions from sampling error or natural variability. 2) The different life histories of each indicator provided a broader assessment of ecosystem function. 3) The unique properties of the indices increase overall sensitivity to different stressors. 4) The different responsiveness of the indices allows better discrimination among condition-classes along the biological condition gradient.

The identification of "limiting indicators" may provide initial steps towards diagnosing stressors or prioritizing sites for rehabilitation. The fact that so many streams were limited for multiple indicators (frequently all four indicators used in the survey) suggests that pressures on many streams are diverse, severe, or both, and fixing these streams may be major challenge. But 19% of the region was limited for a single indicator, and this may indicate that pressures are less severe or more similar in action; rehabilitating these streams may be a more surmountable challenge than streams with fewer indicators in intact condition.

Table 1-1. Thresholds for identifying non-reference condition for biological indicators. Ref mean: Mean of reference sites. Ref SD: Standard deviation of reference sites. Numbers in parentheses refer to the percentiles used to set boundaries between classes. "Multi" refers to the threshold used in multiple-indicator analyses (i.e., the 2.5th percentile); samples with scores above all "multi" thresholds are considered to be in reference condition, with a 10% error rate.

Index	Ref N	Ref mean	Ref SD	Class 1 (≥30th Intact)	Class 2 (10th to 30th)	Class 3 (1st to 10th)	Class 4 (<1st Altered)	Multi
Benthic Macroinvertebrates								
CSCI	479	1.00	0.16	≥0.92	0.79 to 0.92	0.63 to 0.79	<0.63	0.69
-pMMI	479	1.00	0.18	≥0.91	0.77 to 0.91	0.58 to 0.77	<0.58	--
-OE	479	1.00	0.19	≥0.90	0.76 to 0.90	0.56 to 0.76	<0.56	--
Benthic Algae								
D18	122	79	13	≥72	62 to 72	49 to 62	<49	54
S2	122	69	17	≥60	47 to 60	29 to 47	<29	69
CRAM								
Overall Score	86	84	9	≥79	72 to 79	63 to 72	<63	66
Buffer and Landscape	86	95	10	≥90	82 to 90	72 to 82	<72	--
Hydrologic Connectivity	86	81	13	≥74	64 to 74	51 to 64	<51	--
Physical Structure	86	81	16	≥73	60 to 73	44 to 60	<44	--
Biotic Structure	86	75	16	≥67	54 to 67	38 to 54	<38	--

Table 1-2a: Mean CSCI scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	682	0.76	0.24	29	16	23	31
Land Use							
Agricultural	92	0.74	0.19	20	17	31	31
Open	306	0.93	0.17	59	21	15	5
Urban	284	0.59	0.16	2	11	30	58
Watershed							
Region 4							
Ventura	37	0.95	0.15	68	17	15	0
Santa Clara	94	0.91	0.21	54	20	15	11
Calleguas	38	0.65	0.15	9	3	38	49
Santa Monica Bay	72	0.70	0.20	18	9	31	43
Los Angeles	44	0.70	0.23	15	23	29	33
San Gabriel	39	0.62	0.25	17	11	15	57
Region 8							
Lower Santa Ana	45	0.59	0.21	11	14	10	65
Middle Santa Ana	57	0.64	0.23	11	16	30	43
Upper Santa Ana	67	0.88	0.20	49	16	26	10
San Jacinto	28	0.72	0.19	14	24	31	31
Region 9							
San Juan	30	0.72	0.18	15	20	27	38
Northern San Diego	36	0.83	0.19	55	11	13	21
Central San Diego	35	0.72	0.17	10	17	37	35
Mission Bay and San Diego	29	0.78	0.27	33	9	25	33
Southern San Diego	31	0.91	0.16	65	19	5	11

Table 1-2b. Mean pMMI scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	682	0.68	0.25	22	10	24	44
Land Use							
Agricultural	92	0.62	0.17	4	16	36	45
Open	306	0.87	0.20	47	19	27	7
Urban	284	0.49	0.12	0	1	18	81
Watershed							
Region 4							
Ventura	37	0.83	0.22	32	26	27	15
Santa Clara	94	0.86	0.22	49	16	25	11
Calleguas	38	0.54	0.09	0	0	32	68
Santa Monica Bay	72	0.64	0.19	13	13	24	50
Los Angeles	44	0.61	0.23	10	1	35	53
San Gabriel	39	0.57	0.25	15	9	6	70
Region 8							
Lower Santa Ana	45	0.50	0.18	0	12	19	68
Middle Santa Ana	57	0.59	0.21	9	9	24	58
Upper Santa Ana	67	0.86	0.23	39	19	34	8
San Jacinto	28	0.62	0.19	12	10	27	51
Region 9							
San Juan	30	0.56	0.22	13	4	6	76
Northern San Diego	36	0.72	0.21	32	14	21	33
Central San Diego	35	0.60	0.18	10	2	34	54
Mission Bay and San Diego	29	0.72	0.27	27	10	11	52
Southern San Diego	31	0.81	0.19	41	33	9	18

Table 1-2c. Mean O/E scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	682	0.85	0.27	46	20	17	18
Land Use							
Agricultural	92	0.86	0.24	47	14	29	10
Open	306	1.00	0.21	71	18	7	4
Urban	284	0.69	0.23	20	23	24	33
Watershed							
Region 4							
Ventura	37	1.09	0.15	94	3	3	0
Santa Clara	94	0.96	0.23	67	15	11	6
Calleguas	38	0.76	0.23	21	20	45	15
Santa Monica Bay	72	0.77	0.24	28	20	35	17
Los Angeles	44	0.80	0.27	31	36	5	28
San Gabriel	39	0.68	0.28	19	25	17	39
Region 8							
Lower Santa Ana	45	0.68	0.27	22	15	32	31
Middle Santa Ana	57	0.70	0.29	28	17	21	34
Upper Santa Ana	67	0.91	0.26	60	15	8	17
San Jacinto	28	0.82	0.27	46	11	24	19
Region 9							
San Juan	30	0.87	0.18	42	33	18	7
Northern San Diego	36	0.96	0.24	70	7	17	6
Central San Diego	35	0.83	0.23	51	10	21	17
Mission Bay and San Diego	29	0.85	0.28	38	29	19	15
Southern San Diego	31	1.01	0.18	75	14	11	0

Table 1-2d. Mean D18 and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	525	53	25	27	13	18	42
Land Use							
Agricultural	70	45	23	11	15	27	47
Open	221	67	21	47	19	16	18
Urban	234	43	24	12	9	18	62
Watershed							
Region 4							
Ventura	35	79	11	84	11	4	2
Santa Clara	63	59	18	28	16	31	25
Calleguas	38	34	16	0	1	19	80
Santa Monica Bay	54	45	18	3	12	36	48
Los Angeles	40	41	26	15	13	12	60
San Gabriel	32	69	23	52	9	19	21
Region 8							
Lower Santa Ana	33	39	23	3	19	12	66
Middle Santa Ana	30	63	25	41	17	14	28
Upper Santa Ana	27	72	23	63	14	7	16
San Jacinto	21	58	25	24	37	10	29
Region 9							
San Juan	30	41	25	10	16	17	57
Northern San Diego	33	58	19	30	23	17	30
Central San Diego	29	46	23	16	8	14	62
Mission Bay and San Diego	30	56	27	28	18	17	37
Southern San Diego	30	58	22	21	32	19	28

Table 1-2e. Mean S2 scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	524	44	25	25	16	27	32
Land Use							
Agricultural	71	26	18	5	6	27	61
Open	217	62	24	59	13	15	12
Urban	236	32	16	2	19	35	43
Watershed							
Region 4							
Ventura	36	49	25	39	4	33	24
Santa Clara	60	58	27	46	16	23	15
Calleguas	38	26	15	0	13	28	59
Santa Monica Bay	54	37	24	20	19	15	46
Los Angeles	41	41	20	21	11	35	33
San Gabriel	32	49	21	26	23	27	24
Region 8							
Lower Santa Ana	33	32	22	11	10	26	53
Middle Santa Ana	30	36	16	8	13	46	33
Upper Santa Ana	26	53	28	47	10	19	23
San Jacinto	21	54	24	51	10	21	19
Region 9							
San Juan	30	45	29	27	6	35	32
Northern San Diego	33	45	26	36	15	12	37
Central San Diego	30	33	19	4	31	22	43
Mission Bay and San Diego	30	49	31	39	11	22	29
Southern San Diego	30	57	27	41	21	21	17

Table 1-2f. Mean CRAM and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	529	64	21	30	13	16	41
Land Use							
Agricultural	77	68	15	20	19	29	32
Open	203	81	10	65	20	12	2
Urban	249	51	18	7	7	16	70
Watershed							
Region 4							
Ventura	32	79	9	54	19	25	2
Santa Clara	69	76	11	48	24	16	12
Calleguas	31	57	18	3	22	17	59
Santa Monica Bay	67	64	19	25	15	22	38
Los Angeles	41	50	19	14	4	16	66
San Gabriel	37	52	22	24	6	2	68
Region 8							
Lower Santa Ana	33	56	18	11	12	20	57
Middle Santa Ana	29	52	23	24	6	4	67
Upper Santa Ana	23	74	10	34	19	30	17
San Jacinto	18	79	13	63	10	10	16
Region 9							
San Juan	31	66	21	38	6	11	45
Northern San Diego	31	81	10	57	19	21	4
Central San Diego	29	63	17	17	14	28	41
Mission Bay and San Diego	30	70	21	50	13	13	25
Southern San Diego	28	76	15	52	19	13	16

Table 1-2g. Mean CRAM Buffer and Landscape attribute scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	529	75	24	40	10	11	39
Land Use							
Agricultural	77	81	18	44	13	21	21
Open	203	92	13	81	12	4	4
Urban	249	62	22	10	8	14	67
Watershed							
Region 4							
Ventura	32	91	12	71	16	11	2
Santa Clara	69	91	12	70	13	10	7
Calleguas	31	65	21	7	15	27	52
Santa Monica Bay	67	72	26	38	8	21	34
Los Angeles	41	67	23	26	9	5	61
San Gabriel	37	68	21	27	5	0	68
Region 8							
Lower Santa Ana	33	59	26	11	12	14	62
Middle Santa Ana	29	53	28	16	0	14	69
Upper Santa Ana	23	86	23	69	8	0	23
San Jacinto	18	79	23	43	13	16	27
Region 9							
San Juan	31	71	24	33	6	10	52
Northern San Diego	31	93	8	74	12	12	2
Central San Diego	29	71	24	29	13	26	31
Mission Bay and San Diego	30	77	24	50	8	7	35
Southern San Diego	28	87	21	67	11	10	12

Table 1-2h. Mean CRAM Hydrologic structure attribute scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	529	63	21	25	18	24	33
Land Use							
Agricultural	77	66	15	17	28	34	22
Open	203	81	15	57	22	18	3
Urban	249	51	17	4	15	26	55
Watershed							
Region 4							
Ventura	32	80	15	52	26	19	4
Santa Clara	69	74	13	35	30	28	7
Calleguas	31	54	16	8	9	32	51
Santa Monica Bay	67	63	17	25	16	30	30
Los Angeles	41	52	22	20	6	22	52
San Gabriel	37	53	24	22	8	9	61
Region 8							
Lower Santa Ana	33	53	20	12	6	28	53
Middle Santa Ana	29	50	20	11	6	26	57
Upper Santa Ana	23	75	19	48	12	31	10
San Jacinto	18	76	22	58	19	0	23
Region 9							
San Juan	31	65	21	18	30	17	35
Northern San Diego	31	79	15	44	28	25	2
Central San Diego	29	65	15	12	28	41	19
Mission Bay and San Diego	30	69	19	30	28	20	22
Southern San Diego	28	78	16	46	25	22	7

Table 1-2i. Mean CRAM Physical structure attribute scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	529	56	25	38	12	15	35
Land Use							
Agricultural	77	59	20	32	23	20	25
Open	203	75	17	71	14	10	4
Urban	249	43	22	16	9	17	58
Watershed							
Region 4							
Ventura	32	76	21	65	15	16	4
Santa Clara	69	73	17	60	22	13	5
Calleguas	31	52	25	31	7	21	41
Santa Monica Bay	67	63	22	46	23	13	19
Los Angeles	41	39	20	17	1	18	64
San Gabriel	37	44	26	21	13	2	64
Region 8							
Lower Santa Ana	33	49	26	29	10	5	56
Middle Santa Ana	29	40	22	18	2	17	63
Upper Santa Ana	23	55	18	22	26	32	20
San Jacinto	18	59	22	50	0	24	26
Region 9							
San Juan	31	66	25	58	5	15	22
Northern San Diego	31	71	16	63	21	8	9
Central San Diego	29	55	20	28	11	29	32
Mission Bay and San Diego	30	64	24	50	23	2	25
Southern San Diego	28	67	17	63	10	17	10

Table 1-2j. Mean CRAM Biotic structure attribute scores and extent estimates for each condition class. n: number of sites used in the analysis. SD: Standard deviation. Class 1: % of streams with scores above the 30% percentile of reference sites. Class 2: % of streams with scores between the 10th and 30th percentiles of reference sites. Class 3: % of streams with scores between the 1st and 10th percentiles of reference sites. Class 4: % of streams with scores below the 1st percentile of reference sites.

Subpopulation	n	Mean	SD	Class 1	Class 2	Class 3	Class 4
South Coast	529	57	24	42	17	11	30
Land Use							
Agricultural	77	63	19	46	27	13	13
Open	203	72	17	69	19	8	4
Urban	249	45	22	22	15	13	50
Watershed							
Region 4							
Ventura	32	66	12	50	29	18	2
Santa Clara	69	66	16	53	24	17	6
Calleguas	31	55	20	35	30	7	28
Santa Monica Bay	67	59	19	42	28	14	16
Los Angeles	41	41	22	19	14	6	61
San Gabriel	37	42	24	24	6	9	62
Region 8							
Lower Santa Ana	33	51	23	33	8	23	36
Middle Santa Ana	29	43	26	21	13	10	56
Upper Santa Ana	23	58	24	38	25	16	22
San Jacinto	18	75	21	73	12	4	11
Region 9							
San Juan	31	63	23	52	7	16	26
Northern San Diego	31	81	13	84	14	2	0
Central San Diego	29	62	19	41	32	15	13
Mission Bay and San Diego	30	69	23	74	4	0	22
Southern San Diego	28	70	16	70	16	6	8

Table 1-3. Percent of stream-miles intact for multiple indicators, or limiting for specific indicators, for each subpopulation. Note that, in contrast to Table 1-2, these results are based on an adjusted "multi" threshold in Table 1-1, which reduces the error associated with multiple comparisons. CI: Confidence interval.

Subpopulation	n	% Intact			Indicators of Poor Condition						
		Estimate	95% CI		CSCI Alone	D18 Alone	S2 Alone	D18 or S2	All Benthic Indicators	CRAM Alone	All Four Indicators
South Coast	453	25	21	28	2	6	7	18	4	3	15
Land Use											
Agricultural	66	9	4	15	1	6	15	29	6	3	22
Open	172	60	51	68	4	11	10	25	1	6	0
Urban	215	2	0	4	1	3	4	10	6	0	25
Watershed											
Region 4											
Ventura	31	50	31	69	9	5	32	37	0	0	0
Santa Clara	51	43	30	55	5	17	6	25	1	3	7
Calleguas	30	0	0	0	0	0	12	29	11	0	32
Santa Monica Bay	47	10	3	16	5	10	6	25	10	0	12
Los Angeles	33	13	5	21	0	0	4	4	0	10	34
San Gabriel	31	28	19	37	0	0	3	3	0	3	7
Region 8											
Lower Santa Ana	32	15	7	23	0	3	6	15	0	0	46
Middle Santa Ana	25	5	0	13	4	0	7	12	3	1	13
Upper Santa Ana	19	62	42	82	0	0	0	5	9	8	0
San Jacinto	14	53	35	70	13	0	0	0	7	0	27
Region 9											
San Juan	29	18	9	27	10	7	6	13	7	0	16
Northern San Diego	31	33	4	62	2	8	23	41	9	2	0
Central San Diego	25	6	0	15	0	15	9	28	4	0	19
Mission Bay and San Diego	29	32	22	41	0	10	0	14	13	0	0
Southern San Diego	26	61	53	70	0	10	0	20	0	2	2

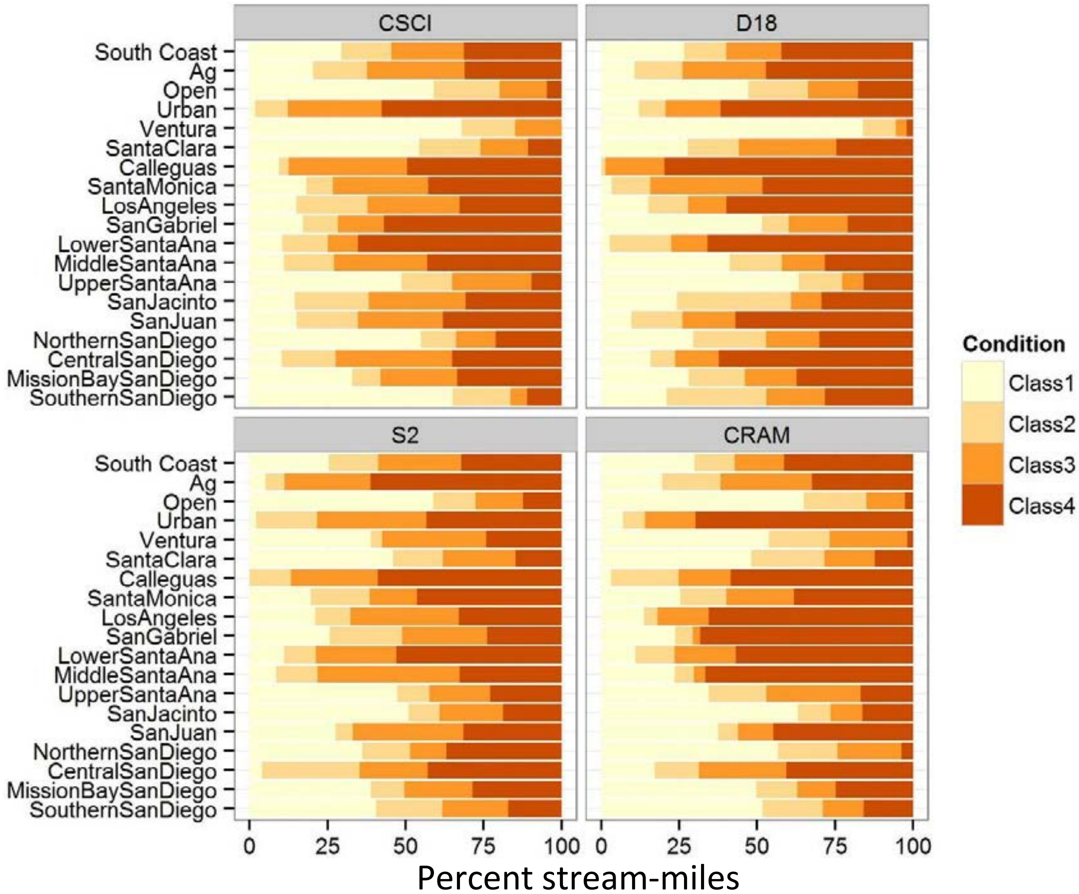


Figure 1-1. Percent of stream-miles in each condition class for each indicator by subpopulation.

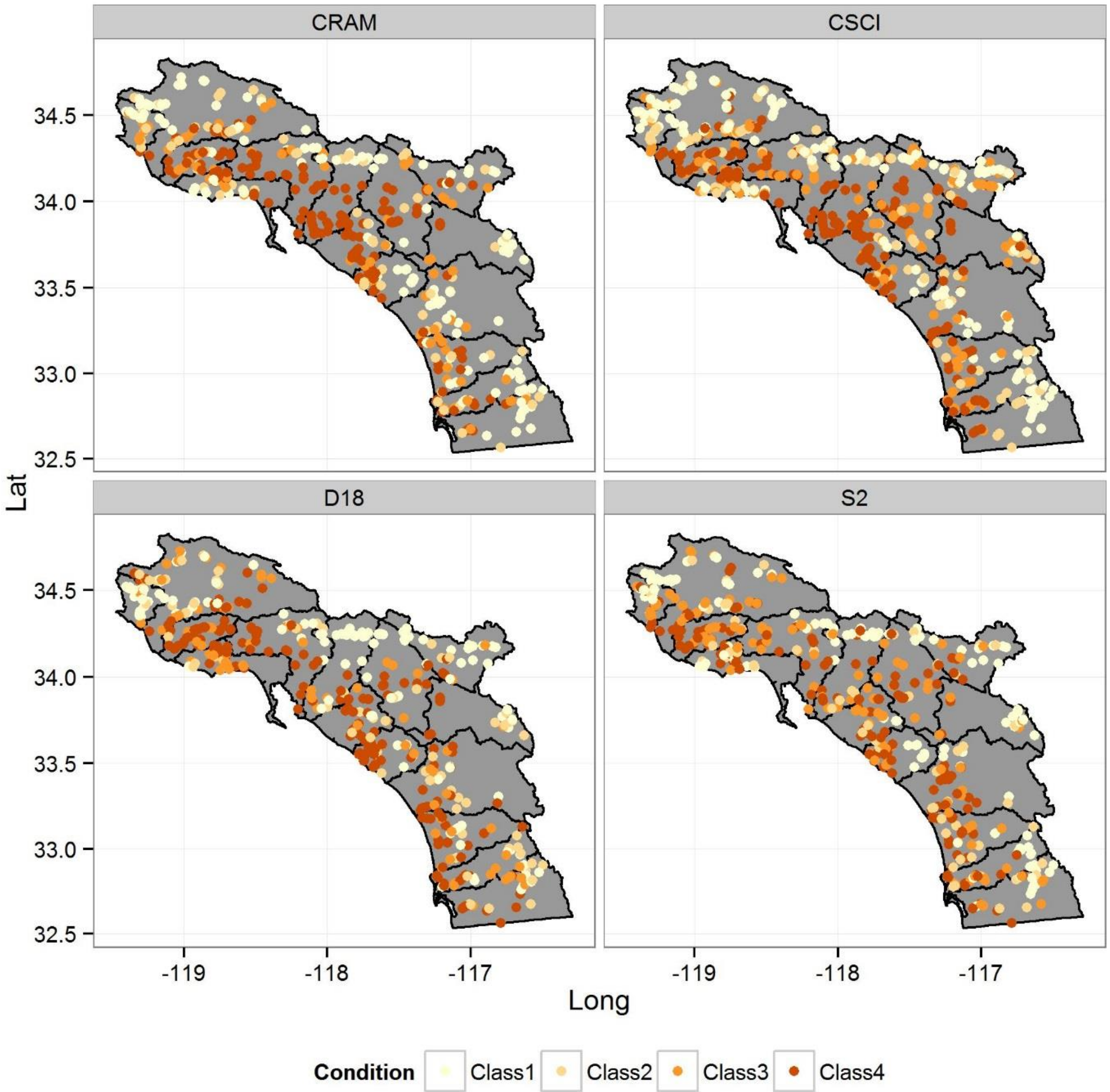


Figure 1-2. Map of scores for key indicators.

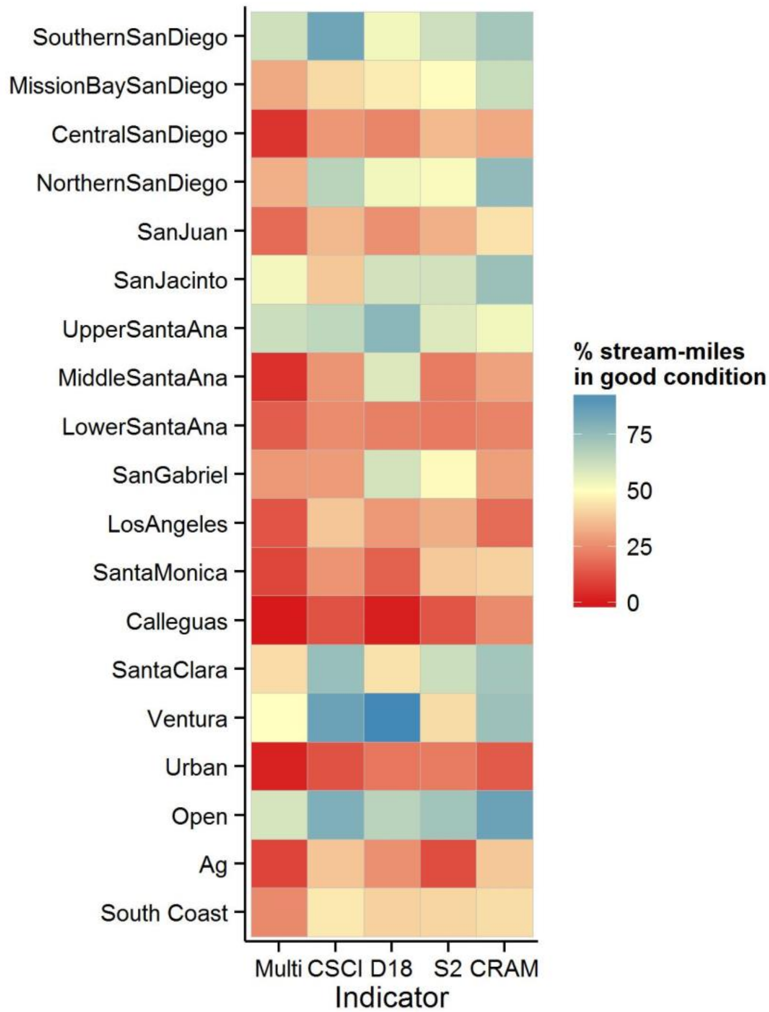


Figure 1-3. Percent of stream-miles in good condition by subpopulation. For the "multi" column, the number reflects the percent of stream-miles with scores for all indicators above the 2.5th percentile of reference sites; all other columns reflect the percent of stream-miles with scores above the 10th percentile of reference sites.

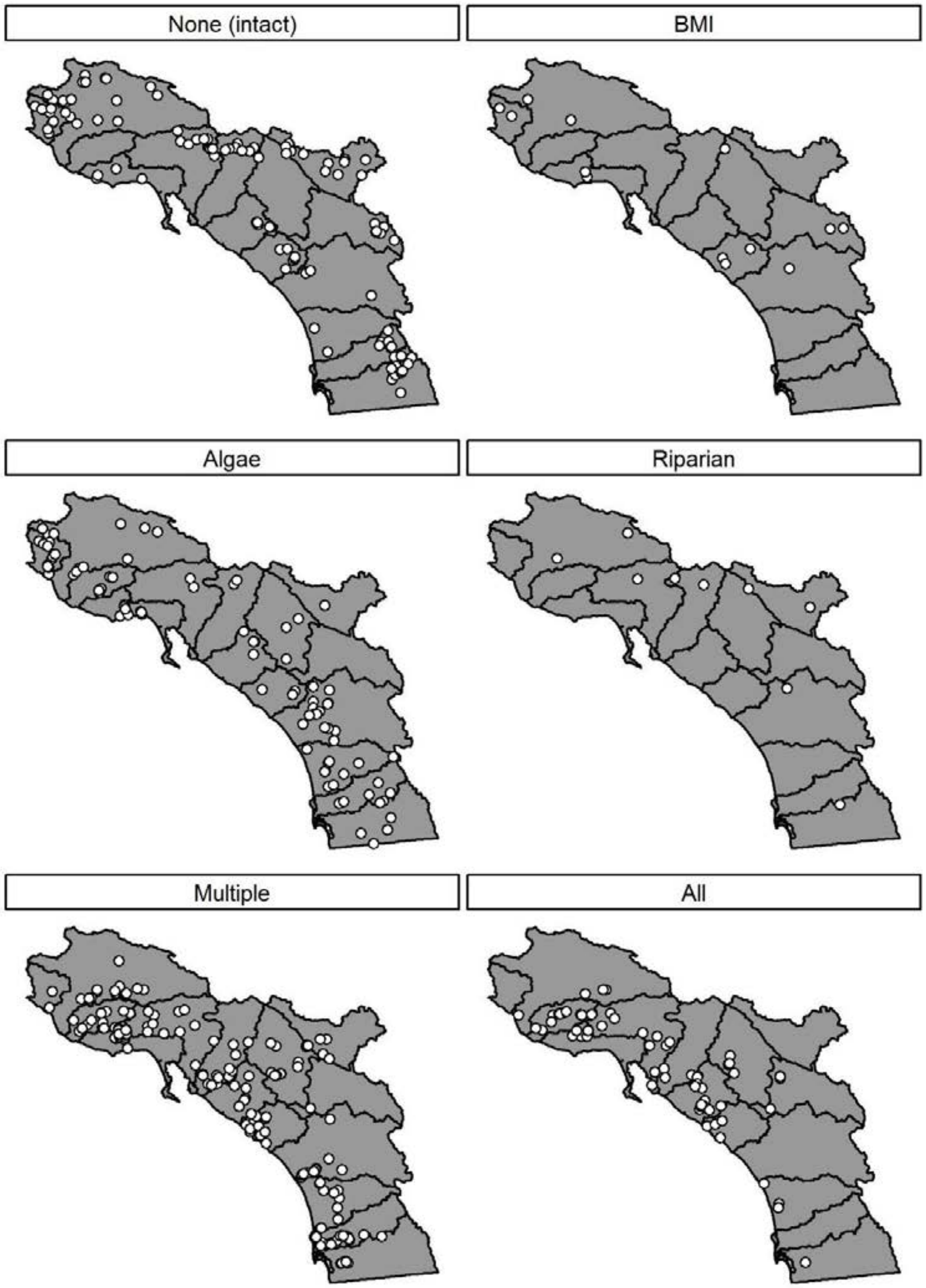


Figure 1-4. Map of limiting indicators. In the top left panel, points represent sites where scores for all four indicators above the 2.5th percentile of reference sites. For all other panels, points represent sites where scores for the specified indicator or indicators were below the 2.5th percentile of reference sites.

QUESTION 2: WHICH STRESSORS ARE ASSOCIATED WITH POOR BIOLOGICAL CONDITION?

Caballero Creek, in the Los Angeles watershed, exemplifies both the severe habitat alteration and nutrient enrichment that affects many streams in southern California.

Introduction

Although the direct measurement of stressors cannot determine the ecological health of a stream, it is essential in determining which factors may limit its health, and provides essential data to inform causal assessment at degraded sites. The SMC stream survey took a notably broad approach towards assessing stressors, measuring nutrients, total and dissolved metals, major ions, water column toxicity, and physical habitat. For some constituents, this survey represents the first unbiased estimate of the extent and magnitude of stressors in aquatic systems. By assessing the extent of these stressors and assessing their associations with biological condition, this survey allows the prioritization of stressors of regional interest, which can then inform local management decisions.

Methods

Data Collection

Data were collected as described in the Survey Overview.

Data Aggregation

Where multiple samples were collected at a single site within a year, data were aggregated as the maximum value within a site. Multi-year mean values for each site were then calculated from these aggregated values if sites were revisited in multiple years. Missing values were ignored for all relevant analyses, where appropriate.

Thresholds

Our goal in setting stressor thresholds was to prioritize stressors in terms of their associated risks to biological condition, as opposed to validating the adequacy of existing regulatory thresholds or assessing compliance with permit requirements. Therefore, the best threshold for this goal is one that is associated with the biggest change in biological condition. Stressor thresholds do not necessarily reflect the most appropriate water quality standards for a given site, which may vary based on site-specific conditions. Therefore, exceeding one of the stressor thresholds used in this analysis may not necessarily indicate impairment or noncompliance with permit requirements.

Stressor thresholds were derived from values published in relevant literature or regulations, where possible (Tables 2-1, 2-2). For chemical nutrients and for most habitat metrics (which are occur naturally and do not have regionally applicable regulatory thresholds), thresholds were established at the 90th or 10th percentile of the distribution among reference sites (as per Ode *et al.* In review). For pyrethroids without published thresholds, a threshold of zero was used.

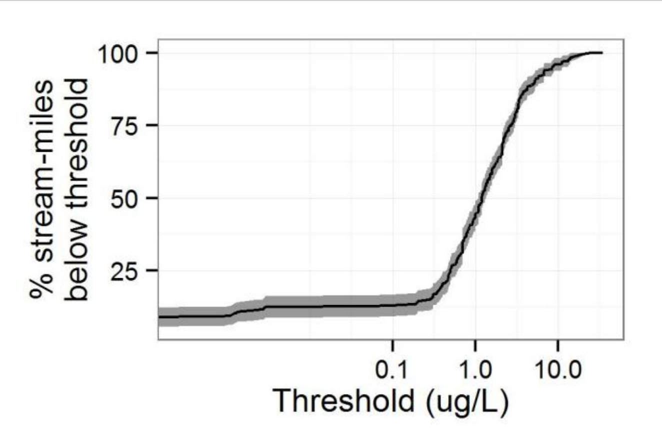
Toxicity tests were compared against controls. If endpoints were significantly different from controls and had values that were 80% of control values or lower, the samples were considered toxic. Toxic survival endpoints were given precedence over nonlethal endpoints (e.g., depressed reproduction).

Reference-Based Thresholds

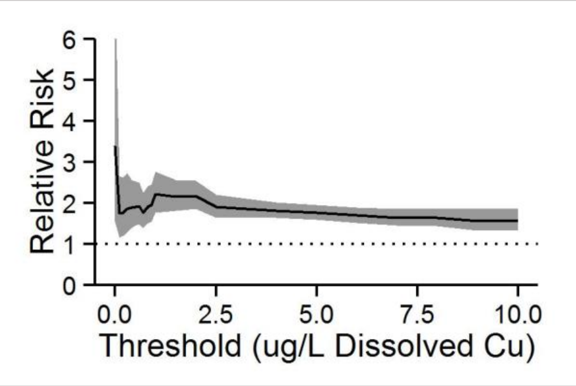
Reference-based thresholds, while appropriate for assessing whether biological indices reflect reference condition, may not be appropriate for water chemistry or physical habitat variables, as they may be excessively stringent. Because of uncertainty about the applicability of certain water chemistry thresholds, a number of alternative thresholds recommended by participating agencies were evaluated.

Copper

To evaluate the impacts of metals on stream condition, this survey used hardness-adjusted thresholds from the California Toxics Rule (EPA 2000). These thresholds are intended to prevent toxic effects on a variety of aquatic species based on the concentration of bio-available toxicants. However, because many of these metals have natural geological sources in the region (e.g., Yoon and Stein 2008), a reference-based threshold, such as those used for nutrients, would better identify sites that exceed natural concentrations. Therefore, a reference-based threshold for copper was calculated as the 90th percentile of concentrations at reference sites within the South Coast region (i.e., 3.4 ug/L), and the extent of stream-miles below this threshold was estimated. Whereas 96% of stream-miles across the region were below the hardness-adjusted threshold for total copper, only 67% were below the reference-based threshold. The difference was even greater for dissolved Copper: 99% of stream-miles were below the hardness-adjusted CTR threshold, whereas only 39% were below the reference threshold of 0.8. Relative risk estimates were only marginally affected (e.g., risk to CSCI scores went up from 1.7 to 1.9 for dissolved copper). However, attributable risks increased considerably (e.g., from 0.004 to 0.360), reflecting the larger number of stream-miles exceeding the reference-based threshold, which would have increased the priority given to this stressor.



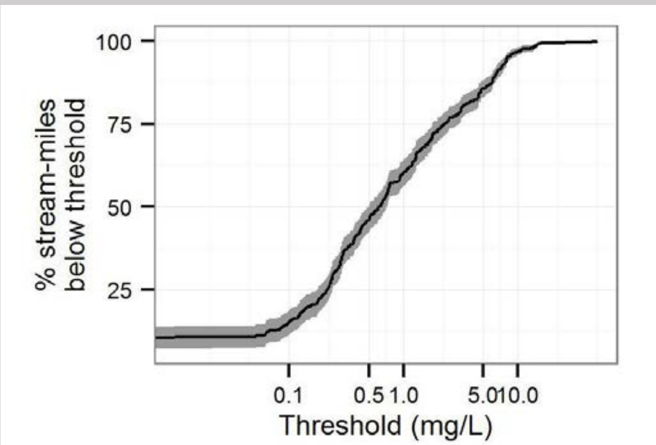
Effects of varying thresholds on the percent of perennial stream-miles below threshold for dissolved copper. The gray band indicates the 95% confidence interval.



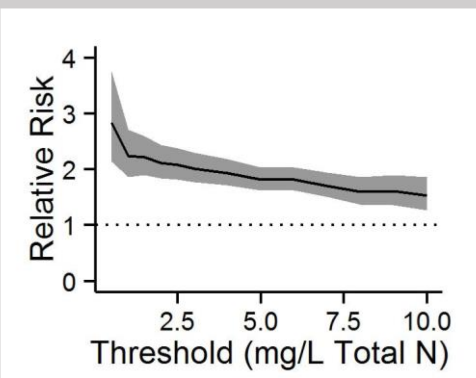
Risk to CSCI scores remain high at all levels of dissolved copper analyzed. The gray band represents the 95% confidence interval. Relative risks greater than 1 (represented by the dotted line) indicate that the stressor is associated with poor biological condition.

Reference-Based Thresholds (Continued)

Total Nitrogen
This study and others (see Herlihy and Sifneos 2008) have shown a strong association between nutrient concentrations and poor biological condition. However, the reference based thresholds used here are much lower than those used in basin plans or TMDLs throughout the region. For example, the reference-based threshold for total nitrogen (TN) was 0.37 mg/L, whereas the San Diego Basin Plan specifies a threshold of 1 mg/L. The Los Angeles Basin Plan sets a much higher threshold of 10 mg/L (although this threshold is explicitly linked to risks to human health and municipal water uses, not aquatic life). Although 39% of stream-miles across the region were below the reference threshold, this number increased to 60% if a threshold of 1 mg/L was used, and to 98% if a threshold of 10 mg/L was used.



Effect of varying thresholds on the percent of perennial stream-miles below threshold for total nitrogen. The gray band indicates the 95% confidence interval.



Risk to CSCI scores remain high at all levels of total N analyzed. The gray band represents the 95% confidence interval. Relative risks greater than 1 (represented by the dotted line) indicate that the stressor is associated with poor biological condition.

Stressor Extent Estimates

Extent estimates and related distribution points were calculated as described in the Survey Overview. These estimates were calculated for land use classes and for the region as a whole, but not for individual watersheds.

Stressor Associations and Prioritization

Relative risk analysis was used to estimate the likelihood of poor biological condition given the presence of a stressor, relative to the likelihood in the absence of a stressor (Van Sickle *et al.* 2006). Attributable risk analysis was then used to estimate the proportion of streams in the region where biological condition may improve if a stressor were removed. Biological condition was determined as described in the section on Question 1, except that Class 1 and 2 streams (Table 1-1) were both treated as "good", and Class 3 and 4 streams were both treated as "poor".

Stressors were then designated as very high priority (attributable risk > 25% of the region for any indicator), high priority (attributable risk between 10% and 25% for any indicator), moderate (attributable risk <10%, but relative risk > 1 for any indicator), and low (relative risk <1 for all indicators).

Relative and Attributable Risk

Relative risk assessment is statistical method of associating the increased risk associated with a stressor (Van Sickle *et al.* 2006). Originally developed for public health studies, relative risk analysis has become popular in environmental assessment because it facilitates prioritization of stressors by identifying which ones are most strongly associated with poor condition. Relative risk compares the odds of observing poor biological condition when a stressor is present to the odds of observing it when the stressor is absent:

$$\text{Relative risk} = \frac{Proportion\ of\ stressed\ stream\text{–}miles\ in\ poor\ condition}{Proportion\ of\ unstressed\ stream\text{–}miles\ in\ poor\ condition}$$

Stressors with relative risks greater than 1 are considered to be associated with poor condition; larger relative risks indicate stronger associations, although any stressor with a risk greater than 1 is a good candidate for further study (e.g., causal analysis).

Relative risk analysis can be extended through attributable risk analysis, which accounts for the fact that low-risk but extensive stressors may be higher regional priorities than high-risk stressors that affect few stream miles (Van Sickle and Paulsen 2008). Attributable risk is calculated as follows:

$$\text{Attributable risk} = \frac{(Proportion\ of\ stressed\ stream\text{–}miles)\times(Relative\ risk-1)}{1+(Proportion\ of\ stressed\ stream\text{–}miles)\times(Relative\ risk-1)}$$

Thus, the attributable risk of a stressor is large if a stressor is extensive and has a relative risk greater than 1. If one assumes a perfect causal relationship between the stressor and poor condition, the attributable risk represents the proportion of the region that would be improved if the stressor were eliminated (Van Sickle and Paulsen 2008). But even when this assumption is violated, attributable risk is a useful metric for ranking stressors by regional importance because it accounts for both stressor extent and strength of association with biological condition.

Both relative risk and attributable risk require stressor thresholds for calculation, and modifying the threshold may alter estimates of risk. If stressor thresholds are set too high, relative risk estimates will go down as the proportion of unstressed stream-miles in poor condition increases. Similarly, if stressor thresholds are set too low, relative risk estimates will also go down as the proportion of stressed stream-miles in poor condition decreases. Ideally, stressor thresholds are set at the level where streams are most likely to switch from poor to good condition (or vice-versa), thereby allowing more direct comparisons of risk across stressors.

Results

All data used in this report can be downloaded from ftp.sccwrp.org/pub/download/SMCReport/SMCDataFor5yearReport.zip.

Stressor Extents

Regional results for all analytes are presented, but only subpopulations where at least 5% of the stream-miles exceeded the threshold are included.

Water Chemistry

In general, nutrients and sulfate exceeded the threshold in extensive portions of the region, while exceedances of pyrethroids and metals were rare (Table 2-3a, Figures 2-1 and 2-2). For example, total Nitrogen exceeded the reference benchmark of 0.37 mg/L in 61% of stream-miles across the region, and sulfates exceeded the benchmark of 250 mg/L in 45% of stream-miles. In contrast, Bifenthrin, the most commonly detected pyrethroid, exceeded the benchmark of 0.0006 ug/L in only 16% of stream-miles, and Selenium exceeded the threshold of 5 ug/L in only 13% of stream-miles. Even within urban areas, pyrethroid and metal exceedances were observed in fewer than 24% of stream-miles (Table 2-3b). Several analytes (e.g., Alkalinity, Arsenic, Nickel, and Zinc) were within thresholds at all sites in the survey. Nonetheless, exceedances of certain constituents were extensive in individual watersheds (Table 2-3c). For example, Bifenthrin exceeded the benchmark in 35% of stream-miles in the Santa Monica Bay watershed, and 30% of the Lower Santa Ana, whereas Selenium exceeded its threshold in 40% of the Calleguas and 55% of the Santa Monica Bay watersheds. Geographic clustering of exceedances was evident for both Selenium and Chloride (Figure 2-2), suggesting a localized (perhaps geological) source for these constituents. Exceedances of the reference-based threshold for total dissolved solids (TDS; i.e., 498 mg/L) were also widespread, affecting 76% of stream-miles region-wide, and nearly all agricultural (97%) and urban (99%) stream-miles. However, a large extent (50%) of open stream-miles also exceeded this threshold, as did 100% of certain watersheds (i.e., Calleguas, Santa Monica, and Lower Santa Ana).

With the exception of Ammonia (whose threshold is based on its toxicity to aquatic invertebrates), nutrients frequently exceeded their benchmarks, based on concentrations observed at reference sites, and these extents were closely related to land use. For example, 71% of open streams were below the threshold for total nitrogen (TN), yet only 12% of urban and 13% of agricultural streams had similarly low concentrations of nitrogen. Exceedances for TN were relatively limited in the Ventura (26%) and Santa Clara (30%) watersheds, but pervasive within the Calleguas (94%). and Lower Santa Ana (90%) watersheds. Total phosphorous (TP) exceedances exhibited similar patterns. For example, 57% of stream-miles exceeded the reference-based benchmark of 0.03 mg/L. As with nitrogen, phosphorous exceedances were pervasive in urban (83% of stream-miles) and agricultural (72%) land uses, and were relatively common in open streams (29%).

Toxicity

Toxicity was detected in surprising geographic patterns. Sublethal toxicity (i.e., depressed reproduction) was somewhat common (evident in 25% of stream-length), and was more extensive in open (33%) than agricultural (30%) or urban (19%) streams (Table 2-4, Figure 2-3). Sublethal toxicity was particularly extensive in the Los Angeles (57%) and Santa Clara (49%) watersheds, but rare within neighboring watersheds, like the San Gabriel (6%) and Calleguas (8%) watersheds. In contrast, toxicity to survival endpoints was evident in only 6% of streams region-wide, and was less extensive in open streams (2%) than urban (8%) or agricultural (15%). Lethal toxicity was most extensive in the Central San Diego watershed (26%), but was fairly limited (extent <10%) in most other watersheds.

Physical habitat

Region-wide, the majority of stream-miles were within the reference distribution for all habitat variables examined, although the more aggregated measures of habitat condition tended to show the most extensive alteration (Table 2-5). For example, the three diversity metrics (i.e., Shannon_Flow, Shannon_Habitat, and Shannon_Substrate), as well as the fish cover metric (i.e., XFC_NAT_SWAMP) were depressed for more than 25% of stream-miles in the region (Figures 2-4 and 2-5).

With the exception of algal biomass variables, the extent of open streams exceeding a benchmark was typically close to the expected distribution at reference sites (i.e., 10%). For example, the Shannon flow metric was outside threshold in 32% of urban stream-miles, 20% of agricultural stream-miles, and only 7% of open stream-miles. This pattern, with the greatest extent of streams exceeding thresholds in urban, followed by agricultural streams, was typical of most habitat variables. A notable exception includes variables directly related to fine sediment (e.g., % sands and fines (PCT_SAFN) and % cobble embeddedness (XEMBED)) were more extensively above threshold in agricultural streams than in urban streams; these metrics may reflect channelization or other flood-control activities that reduce particulate substrates (such as cobbles and sand grains) in urban streams.

Biomass variables frequently exceeded reference-based thresholds across different land-use types, including undeveloped streams. For example, macroalgae cover (i.e., PCT_MAP) exceeded the threshold in 42% of urban streams, 31% of agricultural streams, and 17% of open streams. In contrast, variables related to habitat complexity or riparian vegetation showed a more familiar pattern across land use types.

The extent of altered habitat varied widely by watershed. For example, the extent of exceedances of biomass thresholds was about a third or less for most watersheds, with the notable exception of benthic Chlorophyll a and ash-free dry mass, where exceedances affected nearly two-thirds of the Santa Monica Bay watershed. The exceedances of the Shannon habitat metric affected 3% or less of the Ventura and Northern San Diego watersheds, but more than half of the Los Angeles, San Gabriel, and Middle Santa Ana watersheds. In fact, exceedances affected more than 50% of these three watersheds for many habitat variables.

Stressor prioritization

Nutrients, variables related to ionic concentration (e.g., TDS, sulfates), and several habitat variables were classified as very high priority stressors, having both high relative and attributable risks for several indicators (Tables 2-6 and 2-7, Figure 2-6). For example, TN had an attributable risk of 0.51 for the CSCI. Total dissolved solids and sulfate were also high priority because of their high attributable risk for the CSCI and S2. In contrast, metals and pyrethroids were typically classified as moderate priority. Some, like Bifenthrin or copper, had comparatively high relative risks (>1.5), but because of their limited extents, were estimated to affect less than 10% of the region. Variables related to biomass were also classified as moderate, but for the opposite reason: low risk, but extensive exceedances of threshold contributed to elevated attributable risks.

While there was general agreement among indices, risks were overall greater for the CSCI, followed by S2, with D18 showing the lowest risks. The same five stressors (TDS, PCT_BIGR, W1_HALL, TP, and TN) had the highest attributable risk for all indices. Copper and XEMBED had relatively high attributable risk for the algae indices, compared to the CSCI, which in turn had higher risk for several habitat complexity measures (e.g., Shannon_Substrate, XPCMG).

Discussion

Nutrients, altered physical habitat, and major ions were both widespread and strongly associated with altered biology. Although metals and pyrethroids may be important stressors at specific sites, they should be considered a lower priority for regional programs (generally because they affected only a limited extent of streams).

Although physical habitat was repeatedly identified as a high-risk stressor, it was not possible to characterize these impacts in a precise, unbiased manner. Many physical habitat variables show large site-to-site variability within undisturbed areas, reflecting the influence of environmental gradients, like watershed size, climate, and geology. Establishing site-specific benchmarks based on environmental setting would probably yield a more accurate assessment of physical habitat. Data collected at reference sites could be used to develop models that can set these benchmarks for different stream types. Additionally, integrating multiple physical habitat variables into one or more indices would probably provide a more comprehensive characterization of habitat condition than the metric-by-metric approach used here.

Why were nutrients so strongly associated with poor biology if elevated biomass, the presumed mechanism of impact, had only a moderately high risk? This apparent conflict could result from several possible reasons: 1) timing of sampling, which may miss peak algae biomass; 2) co-occurrence with other stressors (such as habitat alteration; Bernal *et al.* 2013), or 3) other mechanisms of impact, such as cyanotoxins or microsystins (e.g., Aboal *et al.* 2002). Because nutrients are such a high priority for the region, further investigation of these explanations may be warranted.

Table 2-1. Analyte threshold by category. Asterisks indicate thresholds that were used when hardness data were unavailable.

Category	Analyte	Threshold	Unit	Source
Ions	Alkalinity as CaCO3	20000	mg/L	EPA (1986)
Ions	Chloride	260	mg/L	EPA (1986)
Ions	Sulfate	250	mg/L	EPA (1986)
Field	pH	6.5 and 8.5		EPA (1986)
Field	Turbidity	3.8	NTU	Ref (n=47)
Field	Specific conductance	878	uS/cm	Ref (n=77)
Solids	Suspended solids	9.5	mg/L	Ref (n=65)
Solids	Dissolved solids	498	mg/L	Ref (n=19)
Metals	Arsenic	150	ug/L	EPA (2000)
Metals	Cadmium	2.2	ug/L	EPA (2000)
Metals	Copper	9*	ug/L	EPA (2000)
Metals	Nickel	2.5*	ug/L	EPA (2000)
Metals	Lead	52*	ug/L	EPA (2000)
Metals	Selenium	5	ug/L	EPA (2000)
Metals	Zinc	120*	ug/L	EPA (2000)
Nutrients	TN	0.42	mg/L	Ref (n=65)
Nutrients	Ammonia-N	1.71	mg/L	EPA 2000
Nutrients	TP	0.03	mg/L	Ref (n=64)
Pyrethroids	Allethrin	0	ug/L	Detection
Pyrethroids	Bifenthrin	0.0006	ug/L	Central Valley draft TMDL (2014)
Pyrethroids	Cyfluthrin	0.00005	ug/L	Central Valley draft TMDL (2014)
Pyrethroids	Cyhalothrin Lambda	0.0005	ug/L	Central Valley draft TMDL (2014)
Pyrethroids	Cypermethrin	0.0002	ug/L	Central Valley draft TMDL (2014)
Pyrethroids	Deltamethrin/Tralomethrin	0	ug/L	Detection
Pyrethroids	Esfenvalerate/Fenvalerate	0.003	ug/L	Central Valley draft TMDL (2014)
Pyrethroids	Permethrin	0.002	ug/L	Central Valley draft TMDL (2014)

Table 2-2. Thresholds for physical habitat variables. n: number of reference sites used to estimate reference distribution. Ref: estimated from reference distribution. RCMP: Reference Condition Monitoring Program, from Ode *et al.* (In review).

Variable	Description	Direction	Threshold	Units	n	Source
Biomass						
Chlorophyll_a	Benthic chlorophyll a	Increase	56	ug/cm^2	66	Ref
AFDM	Benthic ash-free dry mass	Increase	37	mg/cm^2	64	Ref
PCT_MAP	% macro-algae cover	Increase	41	%	49	Ref
XMIATP	Mean microalgae thickness (where present)	Increase	1.0	mm	53	Ref
PCT_MIAT1	% thick (>1 mm) microalgae cover	Increase	18	%	53	Ref
PCT_MCP	% macrophyte cover	Increase	37	%	49	Ref
PCT_CPOM	% coarse particulate organic matter cover	Increase	71	%	60	Ref
Instream habitat						
XFC_NAT_SWAMP	Natural fish cover	Decrease	18	%	73	Ref
Shannon_Habitat	Fish cover diversity	Decrease	1.1		73	Ref
Shannon_Flow	Flow habitat diversity	Decrease	2.4		61	Ref
PCT_FAST	% fast-water habitat	Decrease	7	%	61	Ref
Riparian						
XCDENMID	% shading	Decrease	17	%	72	Ref
XCMG	Mean riparian vegetation cover	Decrease	32	%	62	Ref
XPCMG	Proportion of reach with all three layers present	Decrease	0.09	Proportion	62	Ref
XPMGVEG	Mean vegetative cover	Decrease	0.23	Proportion	73	Ref
W1_HALL_SWAMP	Human activity metric	Decrease	1.5		60	RCMP
Substrate						
PCT_BIGR	% large substrate (>128 mm)	Decrease	27	%	73	Ref
PCT_SAFN	% sands and fines (<2 mm)	Increase	57	%	73	Ref
Shannon_Substrate	Substrate diversity	Decrease	0.53		73	Ref
XEMBED	% cobble embeddedness	Increase	55	%	73	Ref

Table 2-3a. Regional extent and distributions for chemical stressors.

Stressor	n	% Below Threshold			Concentration		
		Estimate	95% CI		Median	Mean	SD
Ions							
Alkalinity as CaCO3	558	100	100	100	200	217	100
Chloride	513	81	77	84	108	182	316
Sulfate	507	55	51	59	228	294	327
Metals (dissolved)							
Arsenic (d)	443	100	100	100	1.9	2.3	2.7
Copper (d)	443	99	99	100	1.2	2.3	3.3
Nickel (d)	443	100	100	100	2.2	4.3	15.4
Lead (d)	443	100	100	100	0.00	0.05	0.17
Selenium (d)	469	89	86	91	0.99	2.59	6.51
Zinc (d)	486	100	100	100	2.0	4.1	7.2
Metals (total)							
Arsenic (t)	458	100	100	100	2.3	2.9	7.5
Copper (t)	458	96	94	98	2.0	5.2	9.6
Nickel (t)	458	100	100	100	2.6	5.9	18.1
Lead (t)	458	95	93	97	0.08	1.57	3.85
Selenium (t)	458	87	84	89	1.20	3.33	13.24
Zinc (t)	458	100	100	100	3.9	15.8	31.1
Nutrients							
TN	503	39	35	43	0.6	2.2	4.1
Ammonia-N	516	99	97	100	0.01	1.58	19.52
TP	513	43	39	47	0.05	3.91	65.11
Pyrethroids							
Bifenthrin	430	84	81	88	0	0.8	4.2
Cyfluthrin	430	93	90	96	0	0.2	1.6
Cyhalothrin lambda	430	95	92	97	0	0.022	0.228
Cypermethrin	430	92	88	95	0	0.20	1.32
Deltamethrin	169	89	84	94	0	0.0001	0.0022
Esfenvalerate/Fenvalerate	406	98	97	100	0	0.0282	0.3271
Permethrin	430	97	95	99	0	0.146	1.769
Solids							
Suspended solids	528	75	71	79	4	16	57
Dissolved solids	226	24	19	28	856	1034	774
Field							
pH	645	85	82	88	8.05	8.07	0.62
Turbidity	418	76	72	81	1.7	7.9	48.7
Specific conductance	656	75	72	78	1034	1259	1210

Table 2-3b. Extent and distributions for chemical stressors in each land use class. Only analytes with extents greater than 5% exceeding a threshold are shown.

Stressor	n	% Below Threshold			Concentration		
		Estimate	95% CI		Median	Mean	SD
Agricultural							
Ions							
Chloride	73	84	77	90	133	209	280
Sulfate	74	31	22	39	324	424	344
Metals (dissolved)							
Selenium	68	74	64	85	3.06	6.23	12.00
Metals (total)							
Selenium	67	77	66	88	3.31	6.34	11.83
Nutrients							
TN	72	13	7	20	2.5	6.5	9.9
TP	73	28	21	35	0.08	0.50	0.78
Pyrethroids							
Bifenthrin	62	90	82	97	0	0.2	1.1
Cyfluthrin	62	95	86	100	0	0.2	0.7
Cypermethrin	62	90	80	100	0	0.08	0.45
Esfenvalerate/Fenvalerate	58	89	78	99	0	0.31	1.07
Solids							
Suspended solids	73	79	69	89	5	43	144
Dissolved solids	25	3	0	9	983	1037	383
Field							
pH	87	94	91	97	7.98	8.03	0.45
Turbidity	56	70	58	81	2.4	45.0	159.0
Specific conductance	87	69	61	78	1322	1542	888
Open							
Ions							
Sulfate	220	73	68	77	71	170	214
Metals (total)							
Lead	178	93	89	97	0.03	1.40	3.37
Selenium	178	92	88	96	0.78	1.52	2.22
Nutrients							
TN	219	71	65	77	0.2	0.5	1.2
TP	225	71	66	76	0.02	0.09	0.43
Pyrethroids							
Bifenthrin	163	95	92	98	0	0.0	0.1
Deltamethrin	74	92	86	97	0	0	0
Solids							
Suspended solids	227	89	85	93	2	4	7
Dissolved solids	108	50	42	58	493	678	490

Stressor	n	% Below Threshold			Concentration		
		Estimate	95% CI		Median	Mean	SD
Field							
Turbidity	187	87	83	92	0.9	2.3	6.8
Specific conductance	291	91	88	94	478	672	570
Urban							
Ions							
Chloride	223	66	60	72	190	303	397
Sulfate	213	42	35	48	289	391	369
Metals (dissolved)							
Selenium	207	84	80	89	1.20	3.27	7.36
Metals (total)							
Selenium	213	84	80	88	1.30	4.17	17.41
Nutrients							
TN	212	12	6	19	1.5	3.0	3.4
TP	215	17	11	22	0.11	8.35	96.04
Pyrethroids							
Bifenthrin	205	76	69	83	0	1.4	5.7
Cyfluthrin	205	90	85	95	0	0.4	2.2
Cyhalothrin lambda	205	93	88	97	0	0.041	0.313
Cypermethrin	205	88	82	93	0	0.36	1.79
Deltamethrin	74	85	75	95	0	0	0
Solids							
Suspended solids	228	61	54	69	8	22	56
Dissolved solids	93	1	0	3	1093	1388	885
Field							
pH	272	72	66	79	8.17	8.24	0.69
Turbidity	175	65	57	74	2.3	7.2	19.8
Specific conductance	278	62	56	67	1397	1800	1439

Table 2-3c. Extent and distributions for chemical stressors in each watershed. Only analytes with extents greater 5% exceeding a threshold are shown. Physical habitat variable abbreviations are provided in Table 2-2.

Stressor		n	% Below Threshold			Concentration		
			Estimate	95% CI		Median	Mean	SD
Region 4								
Ventura								
Ions	Sulfate	38	36	23	50	270	262	66
Nutrients	TN	38	74	64	84	0.1	0.5	1.0
Nutrients	TP	36	92	87	97	0	0.02	0.06
Pyrethroids	Bifenthrin	35	93	86	100	0	0.0	0.0
Solids	Dissolved solids	5	50	4	97	477	560	96
Field	Turbidity	8	76	39	100	0.5	1.9	1.7
Santa Clara								
Ions	Sulfate	75	59	50	68	221	305	333
Metals (dissolved)	Selenium	70	92	86	97	0.81	1.69	3.25
Metals (total)	Copper	59	91	85	98	0.8	6.3	16.1
Metals (total)	Lead	59	91	86	97	0.01	2.17	4.21
Metals (total)	Selenium	59	90	83	97	0.89	3.15	12.17
Nutrients	TN	70	70	61	78	0.2	0.9	2.4
Nutrients	TP	73	82	75	89	0.02	0.10	0.41
Pyrethroids	Bifenthrin	53	93	86	99	0	0.0	0.0
Pyrethroids	Cyfluthrin	53	92	85	100	0	0.1	0.6
Pyrethroids	Cypermethrin	53	94	87	100	0	0.00	0.02
Pyrethroids	Deltamethrin	33	84	73	95	0	0	0
Pyrethroids	Esfenvalerate/Fenvalerate	50	94	87	100	0	0.1178	0.6286
Solids	Suspended solids	73	91	84	98	2	16	83
Solids	Dissolved solids	45	28	15	42	667	751	467
Field	Turbidity	72	89	84	94	1.5	16.9	98.9
Calleguas								
Ions	Chloride	34	86	70	100	182	193	54
Ions	Sulfate	40	25	13	38	419	484	347
Metals (dissolved)	Selenium	38	60	46	74	4.16	7.14	11.46
Metals (total)	Selenium	37	60	47	74	4.18	7.12	11.01
Nutrients	TN	38	6	0	14	4.4	6.7	9.9
Nutrients	Ammonia-N	35	95	87	100	0.06	0.23	0.70
Nutrients	TP	37	23	6	39	0.13	0.83	1.02
Pyrethroids	Bifenthrin	37	86	76	97	0	0.2	1.0
Pyrethroids	Cypermethrin	37	92	82	100	0	0.15	0.53
Pyrethroids	Esfenvalerate/Fenvalerate	31	94	87	100	0	0.1290	0.7575
Solids	Suspended solids	33	72	56	88	6	27	89
Field	pH	34	86	75	98	7.94	8.04	0.47

Stressor		n	% Below Threshold			Concentration		
			Estimate	95% CI		Median	Mean	SD
Field	Turbidity	9	73	43	100	1.4	2.9	3.0
Field	Specific conductance	34	60	43	77	1691	1785	597
Santa Monica Bay								
Ions	Chloride	47	86	80	93	190	199	72
Ions	Sulfate	54	8	4	12	884	954	570
Metals (dissolved)	Selenium	53	41	34	49	6.61	13.76	20.47
Metals (total)	Selenium	54	45	38	53	5.33	21.80	58.27
Nutrients	TN	50	30	22	39	0.6	1.3	2.0
Nutrients	TP	49	18	11	24	0.10	0.15	0.18
Pyrethroids	Bifenthrin	42	65	52	78	0	3.5	15.6
Pyrethroids	Cyfluthrin	42	89	81	97	0	1.0	4.7
Pyrethroids	Cyhalothrin lambda	42	74	62	86	0	0.237	1.083
Pyrethroids	Cypermethrin	42	83	73	93	0	0.42	1.73
Pyrethroids	Deltamethrin	24	71	56	87	0	0	0
Pyrethroids	Esfenvalerate/Fenvalerate	42	93	86	100	0	0.0291	0.1146
Pyrethroids	Permethrin	42	86	76	95	0	1.119	4.593
Solids	Suspended solids	47	88	81	96	2	10	44
Field	Turbidity	65	70	61	80	1.8	10.9	46.0
Field	Specific conductance	69	59	52	67	1640	1899	1265
Los Angeles								
Ions	Sulfate	32	86	76	96	84	137	152
Metals (total)	Copper	26	82	67	98	7.0	10.4	10.1
Metals (total)	Lead	26	92	82	100	0.65	1.60	2.26
Nutrients	TN	31	34	19	49	1.1	2.5	2.6
Nutrients	TP	22	18	0	36	0.17	0.20	0.16
Pyrethroids	Bifenthrin	26	73	57	89	0	0.5	1.1
Pyrethroids	Cypermethrin	26	92	80	100	0	0.55	1.90
Solids	Suspended solids	19	63	43	84	5	22	35
Solids	Dissolved solids	9	28	4	52	653	1061	837
Field	pH	42	66	53	78	8.25	8.45	0.79
Field	Turbidity	8	67	33	100	0.4	7.6	11.7
Field	Specific conductance	44	91	83	100	570	838	561
San Gabriel								
Ions	Chloride	29	89	76	100	146	127	97
Ions	Sulfate	28	79	59	99	168	151	115
Metals (total)	Copper	27	94	86	100	2.7	7.0	11.4
Metals (total)	Lead	27	91	81	100	0.16	2.04	5.39
Metals (total)	Selenium	27	88	80	97	1.29	2.16	2.00
Nutrients	TN	29	36	20	52	0.6	1.6	2.1
Nutrients	TP	30	44	26	62	0.06	0.12	0.24

Stressor		n	% Below Threshold			Concentration		
			Estimate	95% CI		Median	Mean	SD
Pyrethroids	Bifenthrin	24	87	72	100	0	1.7	6.3
Pyrethroids	Cyfluthrin	24	87	72	100	0	0.8	2.9
Pyrethroids	Cyhalothrin lambda	24	87	72	100	0	0.105	0.371
Pyrethroids	Cypermethrin	24	87	72	100	0	0.82	3.10
Solids	Suspended solids	30	69	51	86	8	37	96
Solids	Dissolved solids	14	13	5	22	859	823	262
Field	pH	33	59	42	76	8.25	8.39	0.65
Field	Turbidity	17	67	44	90	2.1	4.3	4.3
Region 8								
Lower Santa Ana								
Ions	Chloride	29	81	68	94	179	186	91
Ions	Sulfate	24	40	22	58	300	372	248
Metals (dissolved)	Selenium	28	86	76	97	1.30	5.38	10.38
Metals (total)	Selenium	28	86	76	97	1.40	5.37	10.17
Nutrients	TN	24	10	0	20	2.2	3.4	3.5
Nutrients	TP	27	20	8	31	0.12	157.2	398.9
Pyrethroids	Bifenthrin	27	70	55	85	0	0.9	2.0
Pyrethroids	Cyhalothrin lambda	27	93	86	100	0	0.000	0.000
Pyrethroids	Permethrin	27	87	75	99	0	0.121	0.727
Solids	Suspended solids	36	63	52	75	6	11	15
Field	pH	41	87	80	94	7.98	7.97	0.64
Field	Turbidity	36	87	79	95	1.9	2.7	3.6
Field	Specific conductance	41	68	57	80	1408	1587	580
Middle Santa Ana								
Metals (dissolved)	Copper	10	89	70	100	3.1	3.9	3.5
Metals (total)	Copper	15	93	80	100	3.7	5.1	4.4
Nutrients	TN	23	16	2	30	2.0	4.1	4.4
Nutrients	TP	33	14	7	21	0.19	0.52	0.59
Solids	Suspended solids	35	72	62	83	5	8	8
Field	pH	55	65	54	75	8.20	8.29	0.90
Field	Turbidity	23	63	45	81	3.1	5.4	6.2
Field	Specific conductance	55	78	68	88	935	866	416
Upper Santa Ana								
Metals (dissolved)	Copper	12	93	81	100	0.9	1.8	2.8
Nutrients	TN	31	50	37	64	0.3	0.6	0.9
Nutrients	Ammonia-N	43	91	77	100	0.01	23.61	75.91
Nutrients	TP	42	54	42	67	0.02	0.29	0.66
Pyrethroids	Bifenthrin	15	90	77	100	0	0.0	0.0
Solids	Suspended solids	44	75	62	88	3	9	20

Stressor		n	% Below Threshold			Concentration		
			Estimate	95% CI		Median	Mean	SD
Field	pH	67	83	75	91	7.98	7.66	0.96
Field	Turbidity	32	88	77	99	0.4	1.5	2.6
San Jacinto								
Ions	Chloride	16	83	73	94	16	90	142
Nutrients	TN	14	53	41	65	0.3	0.8	1.1
Nutrients	TP	17	18	2	36	0.08	0.17	0.23
Solids	Suspended solids	17	82	70	95	2	6	9
Field	pH	27	81	73	89	7.48	7.67	0.84
Field	Turbidity	6	66	32	99	2.3	38.1	57.4
Field	Specific conductance	27	84	75	94	192	451	568
Region 9								
San Juan								
Ions	Chloride	31	65	51	79	151	205	149
Ions	Sulfate	31	43	31	56	289	450	432
Metals (dissolved)	Selenium	30	76	62	90	1.96	5.00	6.85
Metals (total)	Lead	30	94	88	100	0.00	1.83	2.68
Metals (total)	Selenium	30	75	61	89	1.99	5.10	6.75
Nutrients	TN	30	56	40	71	0.3	0.7	1.1
Nutrients	TP	27	29	18	41	0.06	1.26	4.27
Pyrethroids	Bifenthrin	30	77	64	90	0	0.7	2.0
Pyrethroids	Cyfluthrin	30	86	75	97	0	0.2	0.6
Pyrethroids	Cyhalothrin lambda	30	92	84	100	0	0.017	0.097
Pyrethroids	Cypermethrin	30	86	75	97	0	0.08	0.24
Pyrethroids	Deltamethrin	13	92	84	100	0	0	0
Solids	Suspended solids	30	87	76	97	3	7	12
Solids	Dissolved solids	30	27	18	37	1193	1331	1061
Field	Turbidity	29	83	70	96	0.9	1.8	2.6
Field	Specific conductance	31	59	47	71	1394	1690	1191
Northern San Diego								
Ions	Chloride	31	74	61	87	120	161	141
Ions	Sulfate	31	58	41	75	220	203	190
Nutrients	TN	31	16	1	31	1.2	2.3	3.3
Nutrients	TP	29	51	36	67	0.03	0.07	0.10
Solids	Suspended solids	33	86	76	97	4	6	11
Solids	Dissolved solids	7	22	0	51	780	767	268
Field	Turbidity	28	83	70	96	0.7	6.0	17.5
Field	Specific conductance	33	63	49	77	834	1046	772
Central San Diego								
Ions	Chloride	36	42	29	55	289	507	631
Ions	Sulfate	36	23	13	32	330	359	273

Stressor		n	% Below Threshold			Concentration		
			Estimate	95% CI		Median	Mean	SD
Metals (dissolved)	Selenium	31	89	78	100	1.09	1.65	1.85
Metals (total)	Selenium	31	89	78	100	1.14	1.74	2.03
Nutrients	TN	33	16	6	25	1.3	3.5	4.3
Nutrients	TP	29	12	3	21	0.09	0.10	0.06
Pyrethroids	Bifenthrin	31	77	62	92	0	2.2	5.7
Pyrethroids	Cyfluthrin	31	88	75	100	0	0.2	0.5
Pyrethroids	Cyhalothrin lambda	31	93	83	100	0	0.007	0.029
Pyrethroids	Cypermethrin	31	87	75	99	0	0.01	0.03
Pyrethroids	Deltamethrin	21	83	68	99	0	0	0
Pyrethroids	Permethrin	31	94	85	100	0	0.114	0.462
Solids	Suspended solids	35	52	36	67	9	15	23
Solids	Dissolved solids	9	16	0	38	1306	1112	517
Field	pH	36	95	86	100	7.89	7.90	0.32
Field	Turbidity	30	63	45	80	2.6	8.6	17.1
Field	Specific conductance	37	25	14	35	2112	2469	2151
Mission Bay and San Diego								
Ions	Chloride	30	37	32	42	447	398	332
Ions	Sulfate	30	41	35	46	314	345	334
Metals (dissolved)	Selenium	30	93	84	100	0.77	1.25	1.71
Metals (total)	Selenium	30	93	84	100	0.82	1.34	1.74
Nutrients	TN	28	28	19	37	1.1	2.2	3.4
Nutrients	TP	28	35	21	49	0.05	0.11	0.13
Pyrethroids	Bifenthrin	30	86	75	97	0	0.0	0.2
Pyrethroids	Cyfluthrin	30	93	86	100	0	0.0	0.1
Pyrethroids	Cyhalothrin lambda	30	91	83	99	0	0.004	0.020
Pyrethroids	Cypermethrin	30	89	79	99	0	0.00	0.02
Pyrethroids	Deltamethrin	19	94	85	100	0	0	0
Solids	Suspended solids	31	66	52	80	4	11	14
Solids	Dissolved solids	9	88	72	100	333	450	368
Field	pH	30	93	86	100	7.95	7.94	0.40
Field	Turbidity	26	64	50	77	2.5	4.8	5.1
Field	Specific conductance	30	39	32	47	2385	1933	1532
Southern San Diego								
Ions	Chloride	33	78	72	84	60	308	538
Ions	Sulfate	33	81	75	87	68	128	145
Metals (total)	Lead	30	93	84	100	0.09	1.21	2.36
Nutrients	TN	33	60	49	70	0.3	0.9	1.7
Nutrients	TP	30	38	22	54	0.04	0.23	0.83
Pyrethroids	Bifenthrin	30	98	95	100	0	0.0	0.1
Pyrethroids	Cypermethrin	30	98	95	100	0	0.00	0.03

Stressor		n	% Below Threshold			Concentration		
			Estimate	95% CI		Median	Mean	SD
Solids	Suspended solids	33	83	71	94	4	6	10
Solids	Dissolved solids	10	63	40	86	479	510	219
Field	Turbidity	29	71	55	86	1.6	3.5	4.1
Field	Specific conductance	33	54	42	65	671	1500	1911

Table 2-4. Extent of toxicity by subpopulation.

Subpopulation	n	% stream-miles with toxicity to survival	% stream-miles with toxicity to reproduction	% stream-miles with no toxicity
South Coast	431	6	25	67
Land Use				
Agricultural	67	15	30	55
Open	171	2	33	61
Urban	193	8	19	73
Watershed				
Region 4				
Ventura	34	1	15	77
Santa Clara	56	8	42	45
Calleguas	36	1	8	91
Santa Monica	38	7	33	60
Los Angeles	34	2	57	42
San Gabriel	26	1	6	90
Region 8				
Lower Santa Ana	28	0	26	67
Middle Santa Ana	22	0	4	96
Upper Santa Ana	14	11	12	77
San Jacinto	14	0	12	88
Region 9				
San Juan	25	8	23	69
Northern San Diego	30	3	23	74
Central San Diego	24	26	12	61
Mission Bay and San Diego River	26	4	31	65
Southern San Diego	24	13	11	76

Table 2-5a. Extent and mean values of selected physical habitat variables within the region. Abbreviations are provided in Table 2-2.

Variable	n	% Within Threshold			Median	Mean	SD
		Estimate	95% CI				
Biomass							
AFDM	526	82	78	85	7	652	2877
Chlorophyll a	531	83	79	87	10	165	880
PCT_CPOM	599	90	88	92	28	33	26
PCT_MAP	481	69	65	74	26	30	25
PCT_MCP	481	89	86	92	5	13	18
PCT_MIAT1	519	92	90	94	0	4	11
XMIATP	519	91	89	94	0.10	0.32	0.63
Instream habitat							
PCT_FAST	601	75	72	79	28	37	33
Shannon_Flow	601	80	76	83	2.7	2.7	0.3
Shannon_Habitat	634	68	65	72	1.4	1.2	0.5
XFC_NAT_SWAMP	634	73	69	76	51	54	41
Riparian							
W1_HALL_SWAMP	597	55	52	59	1.2	1.8	1.9
XCDENMID	617	69	66	73	43	45	35
XCMG	602	68	65	72	80	80	60
XPCMG	602	71	68	74	0.65	0.53	0.42
XPMGVEG	634	70	67	73	0.75	0.59	0.41
Substrate							
PCT_BIGR	634	49	45	52	25	30	28
PCT_SAFN	634	78	75	81	25	33	27
Shannon_Substrate	634	73	69	77	1.0	0.9	0.5
XEMBED	485	89	86	92	35	36	18

Table 2-5b. Extent and mean values of selected physical habitat variables by land use. Only variables with exceedances greater than 5% of a subpopulation are shown.

Variable		n	% Within Threshold			Median	Mean	SD
			Estimate	95% CI				
Agricultural								
Biomass	AFDM	75	72	62	81	13	703	2427
Biomass	Chlorophyll a	75	74	64	84	20	486	1837
Biomass	PCT_CPOM	76	86	79	94	36	38	27
Biomass	PCT_MAP	69	69	60	79	28	30	22
Biomass	PCT_MCP	69	88	81	95	12	18	18
InstreamHab	PCT_FAST	76	71	61	81	24	37	33
InstreamHab	Shannon_Flow	76	80	72	89	2.6	2.6	0.3
InstreamHab	Shannon_Habitat	81	80	73	87	1.4	1.3	0.4
InstreamHab	XFC_NAT_SWAMP	81	79	72	87	49	61	47
Riparian	W1_HALL_SWAMP	76	70	63	78	0.6	1.0	1.2
Riparian	XCDENMID	76	58	49	67	23	35	35
Riparian	XCMG	76	80	72	88	104	94	59
Riparian	XPCMG	76	76	68	84	0.79	0.61	0.41
Riparian	XPMGVEG	81	85	78	91	0.81	0.70	0.35
Substrate	PCT_BIGR	81	24	16	32	9	18	21
Substrate	PCT_SAFN	81	40	30	49	63	60	27
Substrate	Shannon_Substrate	81	78	69	88	0.8	0.9	0.4
Substrate	XEMBED	54	81	71	92	40	41	22
Open								
Biomass	AFDM	224	82	77	87	11	173	672
Biomass	Chlorophyll a	227	85	80	90	12	62	201
Biomass	PCT_CPOM	261	88	85	92	34	38	25
Biomass	PCT_MAP	203	83	78	88	14	21	21
Biomass	PCT_MCP	203	87	83	91	7	14	16
Biomass	PCT_MIAT1	217	94	90	97	0	4	9
Biomass	XMIATP	217	94	90	97	0.10	0.26	0.49
InstreamHab	PCT_FAST	263	92	89	95	40	46	29
InstreamHab	Shannon_Flow	263	93	90	95	2.8	2.8	0.3
InstreamHab	Shannon_Habitat	290	90	87	93	1.5	1.4	0.3
InstreamHab	XFC_NAT_SWAMP	290	93	89	97	71	72	35
Riparian	W1_HALL_SWAMP	261	91	87	94	0.2	0.5	0.8
Riparian	XCDENMID	289	85	82	89	61	58	31
Riparian	XCMG	264	93	89	98	108	106	45
Riparian	XPCMG	264	93	90	95	0.86	0.70	0.33
Riparian	XPMGVEG	290	93	90	96	0.91	0.78	0.29
Substrate	PCT_BIGR	290	82	78	86	54	49	24

Variable		n	% Within Threshold			Median	Mean	SD
			Estimate	95% CI				
Substrate	PCT_SAFN	290	88	84	91	24	29	21
Substrate	Shannon_Substrate	290	92	87	96	1.2	1.2	0.4
Substrate	XEMBED	276	91	88	94	35	36	16
Urban								
Biomass	AFDM	227	83	77	89	5	1089	3944
Biomass	Chlorophyll a	229	83	77	89	7	206	991
Biomass	PCT_CPOM	262	92	89	96	17	27	27
Biomass	PCT_MAP	209	58	50	65	37	38	27
Biomass	PCT_MCP	209	91	87	95	2	12	20
Biomass	PCT_MIAT1	232	90	86	94	0	5	12
Biomass	XMIATP	232	89	84	93	0.11	0.37	0.68
InstreamHab	PCT_FAST	262	62	55	69	14	30	34
InstreamHab	Shannon_Flow	262	68	62	74	2.6	2.6	0.2
InstreamHab	Shannon_Habitat	263	44	38	50	0.9	0.9	0.6
InstreamHab	XFC_NAT_SWAMP	263	50	45	56	19	34	36
Riparian	W1_HALL_SWAMP	260	23	87	98	2.9	3.0	1.8
Riparian	XCDENMID	252	54	48	60	20	32	35
Riparian	XCMG	262	44	39	49	22	54	62
Riparian	XPCMG	262	51	46	57	0.10	0.37	0.42
Riparian	XPMGVEG	263	44	39	49	0.09	0.37	0.42
Substrate	PCT_BIGR	263	20	15	24	1	13	21
Substrate	PCT_SAFN	263	74	69	79	25	33	30
Substrate	Shannon_Substrate	263	53	47	59	0.6	0.6	0.5
Substrate	XEMBED	155	86	80	92	35	35	20

Table 2-5c. Extent and mean values of selected physical habitat variables by watershed. Only variables with exceedances greater than 5% of a subpopulation are shown.

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
Region 4								
Ventura								
Biomass	AFDM	37	89	79	100	4	786	3883
Biomass	Chlorophyll a	37	89	79	100	5	88	384
Biomass	PCT_MAP	24	78	60	96	19	25	22
Biomass	PCT_MCP	24	93	85	100	1	7	14
InstreamHab	PCT_FAST	36	95	90	99	36	45	26
Riparian	W1_HALL_SWAMP	36	93	87	98	0.5	0.6	0.6
Riparian	XCDENMID	37	87	75	98	58	59	32
Riparian	XPMGVEG	38	90	78	100	0.69	0.66	0.23
Substrate	PCT_BIGR	38	86	79	94	62	62	22
Santa Clara								
Biomass	AFDM	73	78	70	86	23	153	917
Biomass	Chlorophyll a	75	83	75	92	18	64	200
Biomass	PCT_CPOM	72	73	63	83	54	54	26
Biomass	PCT_MAP	66	75	66	84	28	29	21
Biomass	PCT_MCP	66	84	76	92	18	19	18
Biomass	PCT_MIAT1	70	93	87	99	0	3	8
Biomass	XMIATP	70	91	84	98	0.02	0.24	0.43
InstreamHab	PCT_FAST	72	87	80	94	28	37	27
InstreamHab	Shannon_Flow	72	92	86	98	2.8	2.8	0.3
InstreamHab	Shannon_Habitat	83	86	78	93	1.5	1.4	0.3
InstreamHab	XFC_NAT_SWAMP	83	94	89	99	61	69	34
Riparian	W1_HALL_SWAMP	72	92	87	97	0.0	0.4	0.8
Riparian	XCDENMID	83	72	63	80	37	44	32
Riparian	XCMG	72	93	90	97	112	108	44
Riparian	XPCMG	72	89	84	94	0.86	0.69	0.35
Riparian	XPMGVEG	83	94	91	98	0.90	0.81	0.25
Substrate	PCT_BIGR	83	74	67	81	47	44	24
Substrate	PCT_SAFN	83	83	77	90	30	35	23
Substrate	Shannon_Substrate	83	92	86	98	1.3	1.2	0.4
Substrate	XEMBED	75	87	81	93	34	36	17
Calleguas								
Biomass	AFDM	40	73	59	88	9	1435	3373
Biomass	Chlorophyll a	40	68	53	83	23	1035	2807
Biomass	PCT_MAP	27	61	43	80	37	36	22
InstreamHab	PCT_FAST	37	84	73	94	30	37	25
InstreamHab	Shannon_Flow	37	89	80	98	2.7	2.7	0.2

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
InstreamHab	Shannon_Habitat	39	73	60	86	1.4	1.2	0.5
InstreamHab	XFC_NAT_SWAMP	39	74	62	86	41	38	27
Riparian	W1_HALL_SWAMP	37	28	14	43	2.7	2.6	1.3
Riparian	XCDENMID	39	60	47	72	25	33	30
Riparian	XCMG	37	67	54	81	58	56	40
Riparian	XPCMG	37	71	58	83	0.25	0.42	0.38
Riparian	XPMGVEG	39	67	55	79	0.40	0.44	0.36
Substrate	PCT_BIGR	39	27	14	41	8	18	23
Substrate	PCT_SAFN	39	62	49	76	43	42	29
Substrate	Shannon_Substrate	39	69	57	80	0.8	0.8	0.5
Substrate	XEMBED	26	89	80	99	32	38	19
Santa Monica Bay								
Biomass	AFDM	53	36	25	47	55	59	40
Biomass	Chlorophyll a	54	39	28	49	67	107	109
Biomass	PCT_CPOM	66	43	33	53	77	71	24
Biomass	PCT_MAP	60	53	42	63	40	40	26
Biomass	PCT_MCP	60	91	85	97	6	13	17
Biomass	PCT_MIAT1	60	91	85	98	0	5	13
Biomass	XMIATP	60	94	89	100	0.08	0.40	1.19
InstreamHab	PCT_FAST	66	77	70	85	17	21	17
InstreamHab	Shannon_Flow	66	86	79	93	2.7	2.8	0.3
InstreamHab	Shannon_Habitat	66	86	80	92	1.6	1.5	0.4
InstreamHab	XFC_NAT_SWAMP	66	90	85	95	84	82	44
Riparian	W1_HALL_SWAMP	66	69	61	78	0.6	1.1	1.3
Riparian	XCDENMID	66	88	82	94	83	71	31
Riparian	XCMG	66	86	81	92	138	124	54
Riparian	XPCMG	66	91	87	96	0.98	0.85	0.31
Riparian	XPMGVEG	66	85	79	91	0.95	0.81	0.34
Substrate	PCT_BIGR	66	70	63	76	43	44	28
Substrate	PCT_SAFN	66	92	87	96	17	24	20
Substrate	Shannon_Substrate	66	88	83	93	1.3	1.2	0.5
Los Angeles								
Biomass	AFDM	31	80	67	92	4	907	2294
Biomass	Chlorophyll a	31	74	61	87	7	133	364
Biomass	PCT_MAP	33	67	52	82	28	33	23
InstreamHab	PCT_FAST	44	77	65	89	53	51	37
InstreamHab	Shannon_Flow	44	72	61	83	2.6	2.6	0.2
InstreamHab	Shannon_Habitat	47	49	39	60	0.9	0.9	0.6
InstreamHab	XFC_NAT_SWAMP	47	45	33	57	14	32	36
Riparian	W1_HALL_SWAMP	44	45	33	56	2.8	2.7	2.4
Riparian	XCDENMID	47	58	45	70	21	31	34

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
Riparian	XCMG	44	32	20	43	16	32	36
Riparian	XPCMG	44	53	40	65	0.09	0.26	0.35
Riparian	XPMGVEG	47	37	26	48	0.00	0.27	0.38
Substrate	PCT_BIGR	47	40	30	50	1	21	26
Substrate	Shannon_Substrate	47	52	38	65	0.5	0.6	0.5
San Gabriel								
Biomass	AFDM	28	72	53	92	5	1758	3644
Biomass	Chlorophyll a	28	75	57	94	6	279	550
Biomass	PCT_MAP	28	52	35	68	36	40	33
InstreamHab	PCT_FAST	40	62	46	77	27	42	39
InstreamHab	Shannon_Flow	40	69	54	83	2.5	2.6	0.3
InstreamHab	Shannon_Habitat	40	39	28	50	0.7	0.8	0.6
InstreamHab	XFC_NAT_SWAMP	40	42	29	55	14	33	40
Riparian	W1_HALL_SWAMP	38	26	19	34	3.2	3.0	1.9
Riparian	XCDENMID	40	50	38	61	11	28	33
Riparian	XCMG	40	35	25	44	9	36	45
Riparian	XPCMG	40	39	28	49	0.00	0.27	0.39
Riparian	XPMGVEG	40	29	19	40	0.00	0.24	0.35
Substrate	PCT_BIGR	40	28	18	39	0	21	30
Substrate	PCT_SAFN	40	91	82	100	6	15	20
Substrate	Shannon_Substrate	40	47	34	60	0.5	0.6	0.6
Substrate	XEMBED	24	91	77	100	34	33	18
Region 8								
Lower Santa Ana								
Biomass	AFDM	29	91	82	99	4	193	754
Biomass	Chlorophyll a	29	91	82	99	9	89	354
Biomass	PCT_MAP	27	57	43	71	39	36	18
InstreamHab	PCT_FAST	38	57	43	71	16	23	28
InstreamHab	Shannon_Flow	38	59	45	74	2.5	2.5	0.3
InstreamHab	Shannon_Habitat	38	66	55	77	1.3	1.2	0.4
InstreamHab	XFC_NAT_SWAMP	38	71	60	82	53	53	45
Riparian	W1_HALL_SWAMP	38	17	7	26	2.3	2.7	1.5
Riparian	XCDENMID	38	50	36	65	18	36	38
Riparian	XCMG	38	46	31	60	27	47	41
Riparian	XPCMG	38	52	38	66	0.10	0.34	0.40
Riparian	XPMGVEG	38	53	38	68	0.28	0.45	0.42
Substrate	PCT_BIGR	38	35	22	47	7	21	25
Substrate	PCT_SAFN	38	69	55	82	48	45	27
Substrate	Shannon_Substrate	38	78	67	88	0.8	0.8	0.4
Substrate	XEMBED	28	87	73	100	37	37	20
Middle Santa Ana								

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
Biomass	AFDM	28	91	79	100	3	11	17
Biomass	PCT_CPOM	52	95	90	100	21	23	21
Biomass	PCT_MAP	32	87	79	95	15	21	20
Biomass	PCT_MCP	32	89	80	98	0	9	14
Biomass	PCT_MIAT1	32	77	63	91	1	13	21
Biomass	XMIATP	32	77	63	91	0.37	0.98	1.66
InstreamHab	PCT_FAST	53	42	31	53	2	22	32
InstreamHab	Shannon_Flow	53	39	29	50	2.3	2.4	0.4
InstreamHab	Shannon_Habitat	54	29	19	40	0.9	0.8	0.6
InstreamHab	XFC_NAT_SWAMP	54	41	32	49	11	28	35
Riparian	W1_HALL_SWAMP	52	49	40	58	1.6	2.0	1.7
Riparian	XCDENMID	54	39	29	48	2	27	36
Riparian	XCMG	53	54	46	62	42	51	49
Riparian	XPCMG	53	47	37	57	0.00	0.36	0.42
Riparian	XPMGVEG	54	58	51	66	0.41	0.46	0.43
Substrate	PCT_BIGR	54	23	17	29	0	17	29
Substrate	PCT_SAFN	54	63	56	69	31	41	40
Substrate	Shannon_Substrate	54	43	33	53	0.4	0.6	0.5
Substrate	XEMBED	28	94	86	100	33	33	20
Upper Santa Ana								
Biomass	PCT_MAP	27	90	82	98	3	13	19
Biomass	PCT_MCP	27	93	85	100	1	8	16
Biomass	XMIATP	27	94	87	100	0.14	0.28	0.52
InstreamHab	PCT_FAST	47	93	87	99	81	66	33
InstreamHab	Shannon_Flow	47	69	57	81	2.6	2.6	0.2
InstreamHab	Shannon_Habitat	52	58	47	68	1.2	1.1	0.5
InstreamHab	XFC_NAT_SWAMP	52	88	81	95	58	63	39
Riparian	W1_HALL_SWAMP	47	96	91	100	0.2	0.4	0.5
Riparian	XCDENMID	52	68	58	78	66	55	38
Riparian	XCMG	47	75	65	86	73	79	54
Riparian	XPCMG	47	63	51	74	0.68	0.51	0.42
Riparian	XPMGVEG	52	79	70	89	0.72	0.64	0.37
Substrate	PCT_BIGR	52	82	74	90	60	55	24
Substrate	PCT_SAFN	52	92	87	98	25	29	17
Substrate	Shannon_Substrate	52	92	86	99	1.1	1.1	0.4
Substrate	XEMBED	49	88	80	96	38	41	11
San Jacinto								
Biomass	AFDM	17	91	79	100	12	19	24
Biomass	PCT_MAP	22	88	76	99	5	13	15
Biomass	PCT_MCP	22	77	62	92	16	20	20
InstreamHab	PCT_FAST	26	44	31	58	5	20	28

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
InstreamHab	Shannon_Flow	26	53	38	69	2.4	2.5	0.2
InstreamHab	Shannon_Habitat	27	72	58	85	1.3	1.2	0.4
Riparian	W1_HALL_SWAMP	26	65	52	77	1.0	1.4	1.4
Riparian	XCDENMID	27	81	74	89	85	69	33
Riparian	XCMG	26	95	87	100	80	93	49
Riparian	XPCMG	26	79	67	91	0.77	0.67	0.39
Riparian	XPMGVEG	27	90	81	99	0.86	0.75	0.30
Substrate	PCT_BIGR	27	65	55	74	39	34	26
Substrate	PCT_SAFN	27	70	56	84	44	46	26
Substrate	Shannon_Substrate	27	83	71	95	1.1	1.0	0.4
Substrate	XEMBED	23	90	79	100	41	41	9
<u>Region 9</u>								
San Juan								
Biomass	AFDM	31	76	62	90	6	1916	7004
Biomass	Chlorophyll a	31	75	60	90	18	123	333
Biomass	PCT_MAP	28	48	31	65	42	41	25
Biomass	PCT_MCP	28	92	85	99	3	10	14
Biomass	PCT_MIAT1	30	82	70	93	0	7	12
Biomass	XMIATP	30	85	75	95	0.04	0.45	0.95
InstreamHab	PCT_FAST	31	83	72	94	31	36	26
InstreamHab	Shannon_Habitat	31	76	63	90	1.4	1.2	0.5
InstreamHab	XFC_NAT_SWAMP	31	74	59	88	46	43	29
Riparian	W1_HALL_SWAMP	31	46	33	58	2.1	2.5	2.1
Riparian	XCDENMID	31	77	62	91	53	50	29
Riparian	XCMG	31	71	56	87	77	74	53
Riparian	XPCMG	31	79	67	92	0.57	0.54	0.36
Riparian	XPMGVEG	31	69	54	83	0.72	0.57	0.42
Substrate	PCT_BIGR	31	54	39	69	29	29	23
Substrate	PCT_SAFN	31	88	80	97	39	36	21
Substrate	Shannon_Substrate	31	69	54	83	0.7	0.8	0.4
Substrate	XEMBED	25	90	80	99	34	34	14
Northern San Diego								
Biomass	AFDM	36	91	84	99	4	12	18
Biomass	Chlorophyll a	36	94	88	100	4	13	26
Biomass	PCT_CPOM	31	90	79	100	41	45	16
Biomass	PCT_MAP	29	76	63	89	13	21	23
Biomass	PCT_MCP	29	79	66	92	15	21	19
InstreamHab	PCT_FAST	31	73	61	85	26	25	24
InstreamHab	Shannon_Flow	31	82	68	96	2.7	2.6	0.3
Riparian	W1_HALL_SWAMP	31	96	91	100	0.1	0.4	0.5
Riparian	XCDENMID	29	93	87	100	70	71	25

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
Substrate	PCT_BIGR	33	55	38	72	28	31	26
Substrate	PCT_SAFN	33	45	20	70	58	57	24
Substrate	Shannon_Substrate	33	84	72	96	1.1	1.0	0.4
Substrate	XEMBED	21	75	54	97	35	40	21
Central San Diego								
Biomass	PCT_CPOM	27	78	62	94	55	55	22
Biomass	PCT_MAP	26	87	76	98	21	22	22
Biomass	PCT_MCP	26	86	74	99	12	20	26
Biomass	PCT_MIAT1	26	78	62	93	9	13	14
Biomass	XMIATP	26	69	51	88	0.66	0.76	0.64
InstreamHab	PCT_FAST	27	70	53	88	12	21	25
InstreamHab	Shannon_Flow	27	85	74	97	2.7	2.7	0.2
InstreamHab	Shannon_Habitat	31	74	58	91	1.5	1.4	0.5
InstreamHab	XFC_NAT_SWAMP	31	80	68	93	70	62	38
Riparian	W1_HALL_SWAMP	27	28	12	44	2.1	2.1	1.1
Riparian	XCMG	28	94	87	100	137	132	55
Riparian	XPCMG	28	90	78	100	0.90	0.77	0.34
Riparian	XPMGVEG	31	94	89	100	0.95	0.88	0.24
Substrate	PCT_BIGR	31	27	13	41	13	18	20
Substrate	PCT_SAFN	31	43	27	59	62	56	29
Substrate	Shannon_Substrate	31	80	65	95	1.1	1.0	0.5
Substrate	XEMBED	23	77	61	93	42	42	23
Mission Bay and San Diego								
Biomass	AFDM	30	95	87	100	4	10	13
Biomass	PCT_CPOM	27	90	82	97	47	48	18
Biomass	PCT_MAP	27	81	68	94	12	21	21
Biomass	PCT_MCP	27	72	58	86	15	22	18
Biomass	PCT_MIAT1	27	77	63	91	2	12	18
Biomass	XMIATP	27	77	62	91	0.44	0.71	0.68
InstreamHab	PCT_FAST	27	66	54	77	17	29	30
InstreamHab	Shannon_Flow	27	78	66	90	2.8	2.8	0.3
InstreamHab	Shannon_Habitat	27	84	75	94	1.5	1.4	0.4
InstreamHab	XFC_NAT_SWAMP	27	88	81	95	82	75	39
Riparian	W1_HALL_SWAMP	27	52	42	62	0.4	1.6	1.8
Riparian	XCDENMID	23	85	76	94	66	53	29
Riparian	XCMG	27	84	75	94	131	110	54
Riparian	XPCMG	27	84	75	94	0.86	0.69	0.35
Riparian	XPMGVEG	27	92	84	100	0.99	0.78	0.31
Substrate	PCT_BIGR	27	51	37	65	28	29	26
Substrate	PCT_SAFN	27	66	51	82	40	44	26
Substrate	Shannon_Substrate	27	88	81	95	1.1	1.1	0.5

Variable		n	% within Threshold			Median	Mean	SD
			Estimate	95% CI				
Substrate	XEMBED	21	91	82	100	39	38	18
Biomass	AFDM	32	76	62	90	5	23	35
Southern San Diego								
Biomass	PCT_CPOM	25	76	62	90	49	50	23
Biomass	PCT_MAP	25	66	50	82	10	24	28
Biomass	PCT_MCP	25	56	36	76	35	34	23
Biomass	PCT_MIAT1	25	89	80	99	4	8	9
Biomass	XMIATP	25	92	82	100	0.51	0.55	0.37
InstreamHab	PCT_FAST	26	85	77	92	29	32	21
InstreamHab	Shannon_Flow	26	94	87	100	2.9	2.8	0.2
InstreamHab	Shannon_Habitat	28	85	74	96	1.4	1.4	0.3
InstreamHab	XFC_NAT_SWAMP	28	94	85	100	60	67	36
Riparian	W1_HALL_SWAMP	25	93	87	99	0.3	0.5	0.6
Riparian	XCDENMID	24	90	77	100	53	58	28
Riparian	XPCMG	26	91	80	100	0.76	0.64	0.33
Substrate	PCT_BIGR	28	48	30	66	25	28	22
Substrate	PCT_SAFN	28	51	33	68	51	52	23
Substrate	Shannon_Substrate	28	95	88	100	1.1	1.0	0.3
Substrate	XEMBED	20	86	72	100	37	39	16

Table 2-6. Relative (RR) and attributable (AR) risks for selected indicators. n: number of sites included in the analysis. 95% CI: 95% confidence interval around estimate. (t) indicates that the total fraction of metals were used in the analysis. (d) indicates that the dissolved fraction of metals were used in the analysis. VH: Very high priority (i.e., attributable risk ≥ 0.25 for at least 1 indicator). H: High priority (i.e., attributable risk ≥ 0.1 for at least 1 indicator). M: Moderate priority (i.e., relative risk > 1). L: Low priority (relative risk ≤ 1). Physical habitat variable abbreviations are provided in Table 2-2. *Some chemistry variables are excluded because they had too few exceedances of thresholds to permit relative risk analysis.

Stressor	Priority	CSCI							D18							S2						
		RR	95% CI		AR	95% CI		n	RR	95% CI		AR	95% CI		n	RR	95% CI		AR	95% CI		n
Chemistry																						
Nutrients																						
TP	VH	2.8	2.1	3.7	0.51	0.39	0.61	469	2.4	1.8	3.1	0.46	0.34	0.56	411	2.1	1.7	2.6	0.08	0.06	0.11	411
TN	VH	2.7	2.0	3.8	0.51	0.36	0.63	473	1.7	1.4	2.2	0.32	0.18	0.43	439	2.7	1.9	3.8	0.53	0.37	0.65	439
NH4	M	1.1	0.5	2.5	0.00	0.00	0.01	473	1.0	0.5	2.4	0.00	0.00	0.01	412	0.6	0.1	2.9	0.00	0.00	0.00	412
Metals																						
Se (d)	M	1.8	1.6	2.0	0.08	0.05	0.11	454	1.5	1.4	1.7	0.06	0.04	0.09	437	1.5	1.3	1.8	0.06	0.03	0.09	438
Cu (d)	M	1.7	1.6	1.8	0.00	0.00	0.01	428	1.6	1.5	1.7	0.00	0.00	0.00	435	1.7	1.5	1.8	0.00	0.00	0.00	437
Se (t)	M	1.5	1.3	1.7	0.06	0.03	0.09	441	1.4	1.2	1.6	0.05	0.02	0.08	450	1.4	1.2	1.6	0.05	0.02	0.08	452
Cu (t)	M	1.4	1.1	1.8	0.02	0.00	0.04	441	1.2	0.9	1.7	0.01	0.00	0.03	450	1.6	1.4	1.9	0.02	0.01	0.04	452
Pb (t)	L	0.8	0.5	1.3	0.00	0.00	0.01	441	0.6	0.4	1.1	0.00	0.00	0.00	450	1.0	0.7	1.4	0.00	0.00	0.02	452
Pyrethroids																						
Bifenthrin	M	1.6	1.4	1.9	0.09	0.05	0.13	415	1.4	1.2	1.7	0.06	0.03	0.10	423	1.5	1.2	1.7	0.07	0.03	0.10	425
Delta/ Tralomethrin	M	1.6	1.1	2.3	0.05	0.00	0.11	162	1.1	0.7	1.5	0.01	0.00	0.04	168	0.4	0.2	0.9	0.00	0.00	0.00	168
Cypermethrin	M	1.5	1.3	1.8	0.04	0.01	0.07	415	1.2	0.9	1.6	0.01	0.00	0.04	423	1.4	1.1	1.8	0.03	0.00	0.06	425
Cyfluthrin	M	1.4	1.2	1.8	0.03	0.00	0.06	415	1.3	1.0	1.7	0.02	0.00	0.04	423	1.3	0.9	1.7	0.02	0.00	0.04	425
Cyhalothrin	M	1.3	1.0	1.6	0.01	0.00	0.03	415	1.1	0.8	1.6	0.01	0.00	0.03	423	1.0	0.7	1.5	0.00	0.00	0.02	425
Esfenvalerate/ Fenvalerate	M	1.3	0.8	2.1	0.01	0.00	0.02	391	1.2	0.8	2.0	0.00	0.00	0.01	399	1.2	0.7	2.0	0.00	0.00	0.01	401
Permethrin	M	1.1	0.7	1.6	0.00	0.00	0.02	415	1.6	1.5	1.7	0.02	0.01	0.03	423	0.8	0.5	1.4	0.00	0.00	0.01	425
Other chemistry																						
TDS	VH	5.2	2.1	12.6	0.76	0.44	0.90	221	1.8	1.3	2.6	0.38	0.16	0.55	222	3.1	1.9	5.3	0.62	0.39	0.76	222
pH	H	1.9	1.7	2.1	0.12	0.08	0.16	593	1.2	1.0	1.5	0.03	0.00	0.07	492	1.6	1.4	1.8	0.08	0.05	0.12	491
Cl	H	1.9	1.6	2.1	0.14	0.09	0.19	489	1.3	1.1	1.5	0.05	0.01	0.09	436	1.1	0.9	1.3	0.02	0.00	0.06	437
SO4	VH	1.8	1.5	2.1	0.26	0.17	0.34	489	1.5	1.3	1.7	0.19	0.11	0.26	459	1.4	1.2	1.7	0.17	0.08	0.24	459
SpCond	H	1.7	1.5	1.9	0.14	0.10	0.18	603	1.5	1.3	1.7	0.13	0.08	0.18	494	1.5	1.3	1.8	0.13	0.08	0.18	493
TSS	H	1.7	1.4	2.0	0.14	0.08	0.19	485	1.3	1.1	1.6	0.07	0.03	0.12	422	1.2	1.0	1.4	0.04	0.00	0.10	423

Stressor	Priority	CSCI							D18							S2						
		RR	95% CI		AR	95% CI		n	RR	95% CI		AR	95% CI		n	RR	95% CI		AR	95% CI		n
Turbidity	H	1.5	1.2	1.8	0.10	0.04	0.16	379	1.2	1.0	1.5	0.06	0.00	0.12	292	0.9	0.7	1.2	0.00	0.00	0.05	289
PHAB																						
Biomass																						
PCT_MAP	H	1.5	1.3	1.8	0.15	0.08	0.21	433	1.3	1.1	1.5	0.08	0.02	0.14	432	1.5	1.3	1.7	0.14	0.08	0.19	431
PCT_CPOM	M	1.2	1.0	1.5	0.02	0.00	0.04	534	1.1	0.9	1.4	0.01	0.00	0.04	494	1.0	0.8	1.2	0.00	0.00	0.02	493
Chl a	M	1.2	0.9	1.4	0.03	0.00	0.07	495	1.2	1.0	1.4	0.03	0.00	0.06	480	1.3	1.1	1.5	0.05	0.02	0.09	479
PCT_MIAT1	M	1.1	0.9	1.5	0.01	0.00	0.04	470	0.9	0.7	1.2	0.00	0.00	0.01	469	0.8	0.6	1.2	0.00	0.00	0.01	468
XMIATP	M	1.1	0.9	1.5	0.01	0.00	0.04	470	0.9	0.7	1.2	0.00	0.00	0.01	469	1.0	0.7	1.3	0.00	0.00	0.02	468
AFDM	M	1.0	0.8	1.3	0.01	0.00	0.05	490	1.1	0.9	1.3	0.02	0.00	0.06	477	1.2	1.0	1.4	0.04	0.00	0.08	476
PCT_MCP	L	0.9	0.7	1.2	0.00	0.00	0.02	433	0.9	0.7	1.2	0.00	0.00	0.02	432	0.8	0.6	1.1	0.00	0.00	0.00	431
Substrate																						
PCT_BIGR	VH	3.1	2.5	3.9	0.51	0.42	0.59	568	2.0	1.7	2.4	0.34	0.26	0.42	494	2.0	1.7	2.4	0.35	0.26	0.42	493
Shannon_Substrate	VH	2.4	2.1	2.7	0.27	0.21	0.32	568	1.4	1.2	1.7	0.11	0.05	0.16	494	1.6	1.4	1.8	0.14	0.09	0.19	493
XEMBED	M	1.3	0.9	1.9	0.04	0.00	0.08	432	1.5	1.3	1.9	0.04	0.01	0.07	374	1.7	1.3	2.3	0.04	0.02	0.07	372
PCT_SAFN	H	1.3	1.1	1.5	0.06	0.02	0.10	568	1.5	1.3	1.7	0.11	0.07	0.14	494	1.3	1.1	1.5	0.06	0.02	0.10	493
Instream habitat																						
XFC_NAT	VH	2.5	2.2	2.9	0.30	0.24	0.35	568	1.3	1.1	1.5	0.07	0.02	0.12	494	1.6	1.4	1.9	0.15	0.10	0.20	493
Shannon_Habitat	VH	2.3	2.0	2.6	0.28	0.22	0.34	568	1.3	1.1	1.5	0.09	0.04	0.15	494	1.6	1.4	1.9	0.17	0.11	0.22	493
PCT_FAST	H	1.7	1.4	1.9	0.14	0.09	0.19	536	1.3	1.1	1.5	0.07	0.02	0.11	494	1.3	1.1	1.5	0.07	0.02	0.11	493
Shannon_Flow	H	1.6	1.4	1.9	0.11	0.07	0.16	536	1.3	1.1	1.5	0.05	0.01	0.09	494	1.4	1.2	1.7	0.07	0.03	0.11	493
Riparian																						
W1_HALL	VH	3.0	2.5	3.6	0.47	0.40	0.54	534	1.8	1.5	2.1	0.25	0.18	0.32	494	1.8	1.6	2.1	0.26	0.19	0.33	493
XCMG	VH	2.4	2.1	2.7	0.30	0.25	0.36	537	1.4	1.2	1.6	0.11	0.06	0.16	494	1.5	1.3	1.8	0.14	0.09	0.20	493
XPMGVEG	VH	2.1	1.9	2.5	0.25	0.19	0.30	568	1.4	1.3	1.7	0.12	0.07	0.17	494	1.5	1.3	1.7	0.14	0.08	0.19	493
XPCMG	H	2.0	1.8	2.3	0.23	0.17	0.28	537	1.3	1.1	1.5	0.07	0.02	0.12	494	1.4	1.2	1.6	0.11	0.06	0.15	493
XCDENMID	H	1.9	1.7	2.3	0.22	0.16	0.28	551	1.2	1.0	1.4	0.05	0.00	0.10	478	1.3	1.1	1.5	0.08	0.03	0.14	477
Toxicity																						
Toxicity (lethal)	M	1.3	1.0	1.7	0.02	0.00	0.04	420	1.2	1.0	1.6	0.02	0.00	0.03	437	1.3	1.1	1.7	0.02	0.00	0.04	438
Toxicity (all endpoints)	M	1.0	0.8	1.2	0.00	0.00	0.05	420	1.2	1.0	1.4	0.05	0.00	0.11	437	1.0	0.8	1.2	0.01	0.00	0.06	438

Table 2-7. Summary of stressor prioritization.

Very high (AR > 0.25)	High (AR 0.1 to 0.25)	Moderate (RR >1)	Low (RR <1)
Water Chemistry	Water Chemistry	Water Chemistry	Water Chemistry
Nutrients	*Other chemistry*	*Nutrients*	*Metals*
TP	Cl	NH4	Pb (t)
TN	pH	*Metals*	Habitat
Habitat	TSS	As (t)	*Biomass*
Instream habitat	SpCond	Se (t, d)	PCT_MCP
XFC_NAT	Habitat	Cu (t, d)	
Shannon_Habitat	*Biomass*	*Pyrethroids*	
Substrate	PCT_MAP	Delta/Tralomethrin	
Shannon_Substrate	*Instream habitat*	Esfenvalerate/Fenvalerate	
PCT_BIGR	Shannon_Flow	Permethrin	
Riparian	PCT_FAST	Cyhalothrin	
XPMGVEG	*Substrate*	Cyfluthrin	
XCMG	PCT_SAFN	Cypermethrin	
W1_HALL	*Riparian*	Bifenthrin	
	XCDENMID	Habitat	
	XPCMG	*Biomass*	
		PCT_MIAT1	
		XMIATP	
		PCT_CPOM	
		AFDM	
		Chl a	
		Substrate	
		XEMBED	
		Toxicity	
		Reproduction	
		Survival	

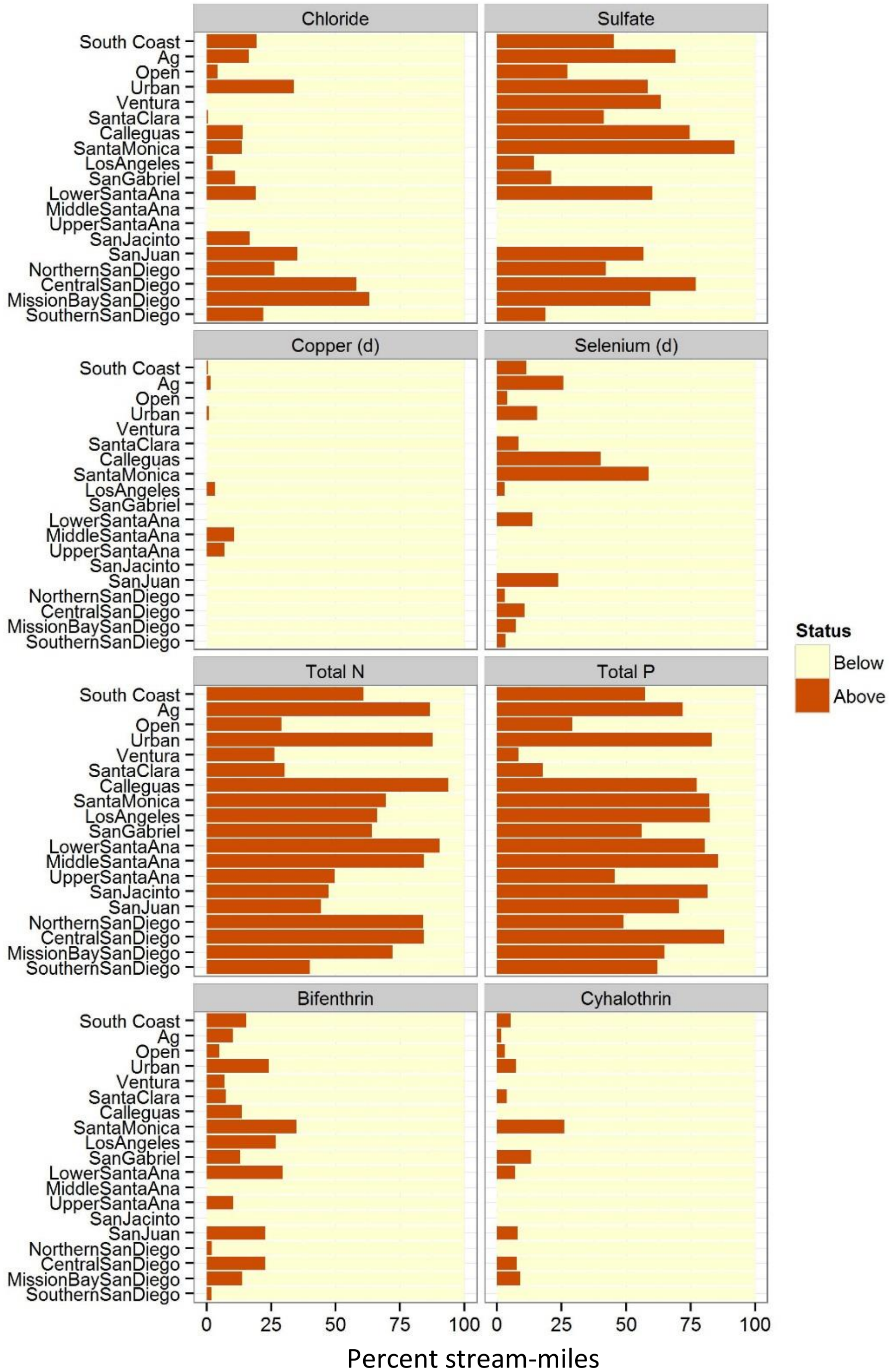


Figure 2-1. Extents of selected water-chemistry variables exceeding thresholds.

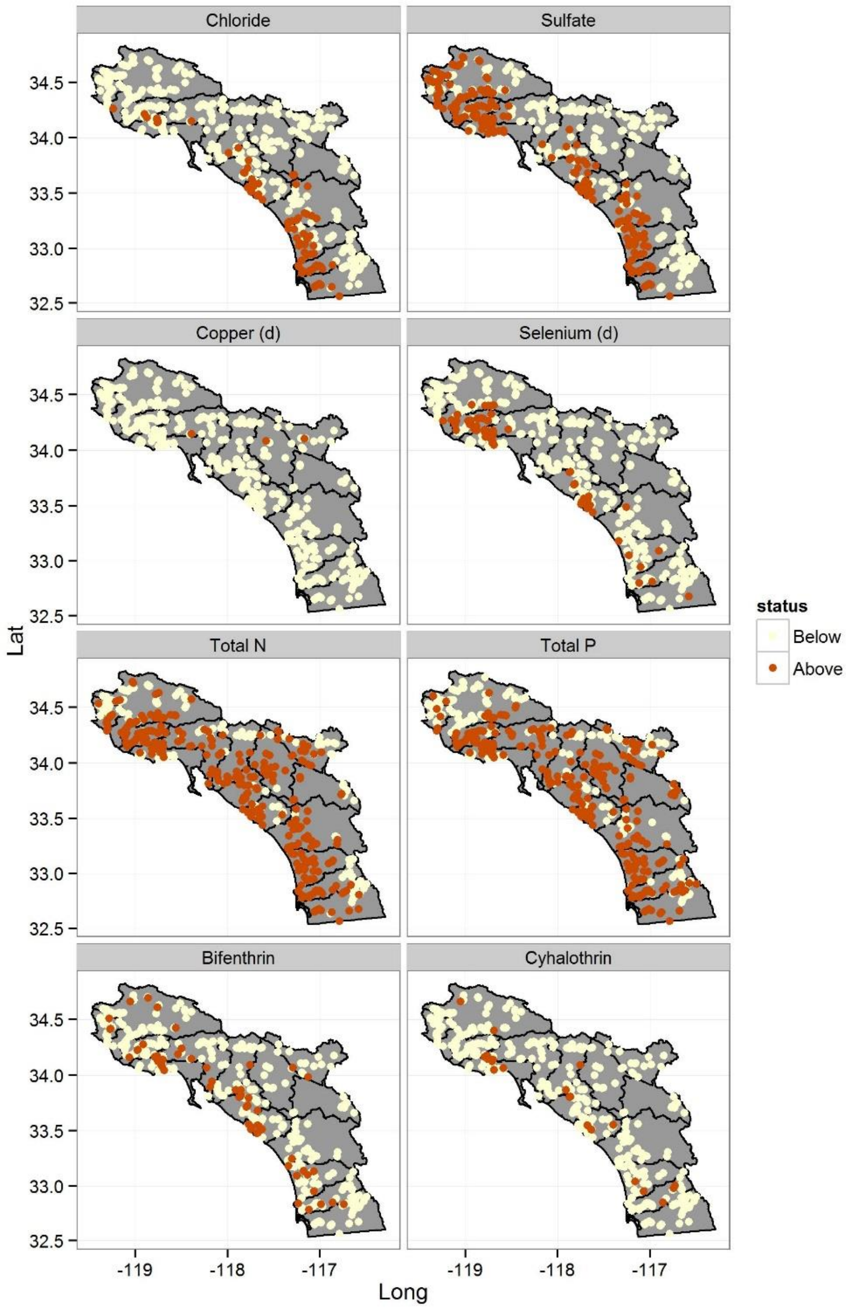


Figure 2-2. Maps of selected water-chemistry variables

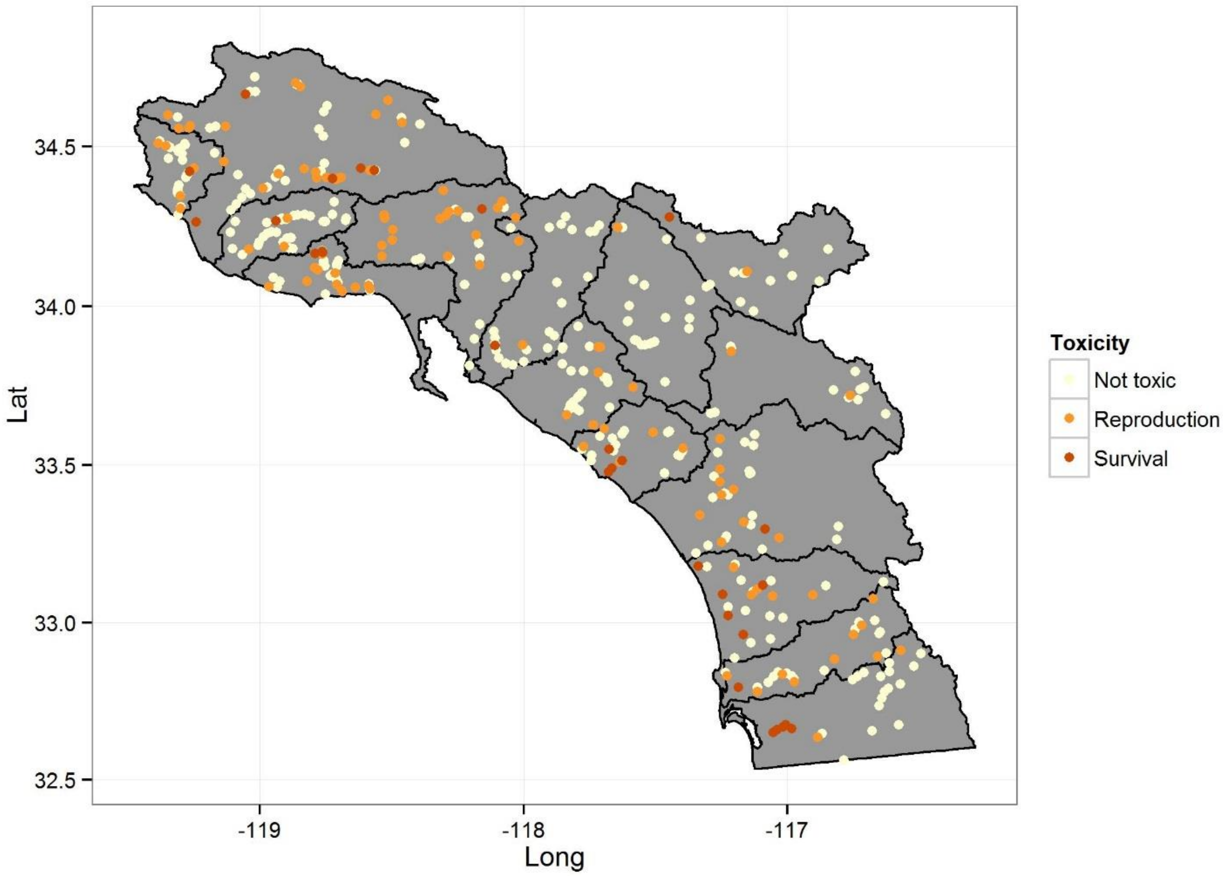


Figure 2-3. Map of toxicity.

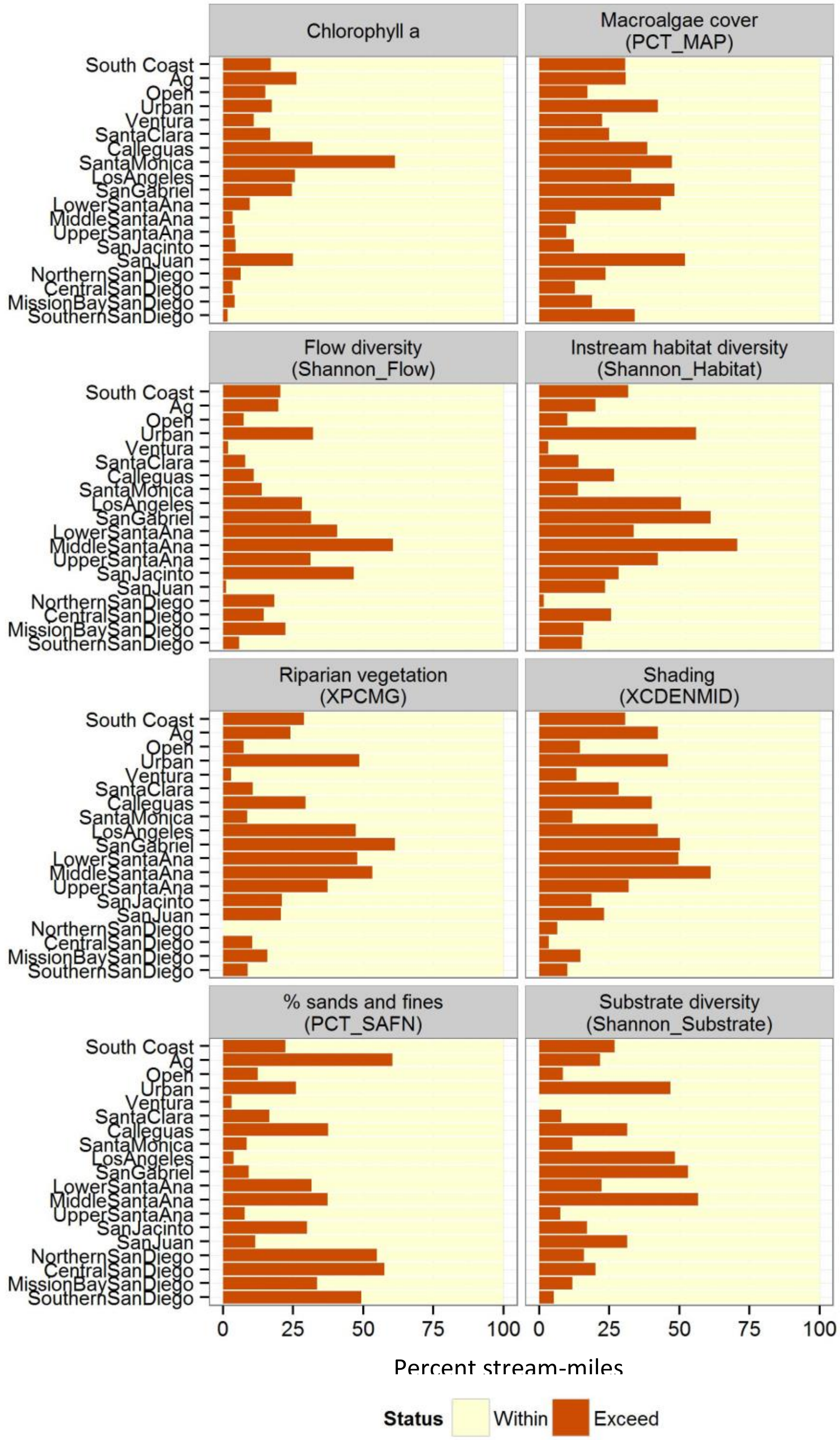


Figure 2-4. Extents of selected physical habitat variables.

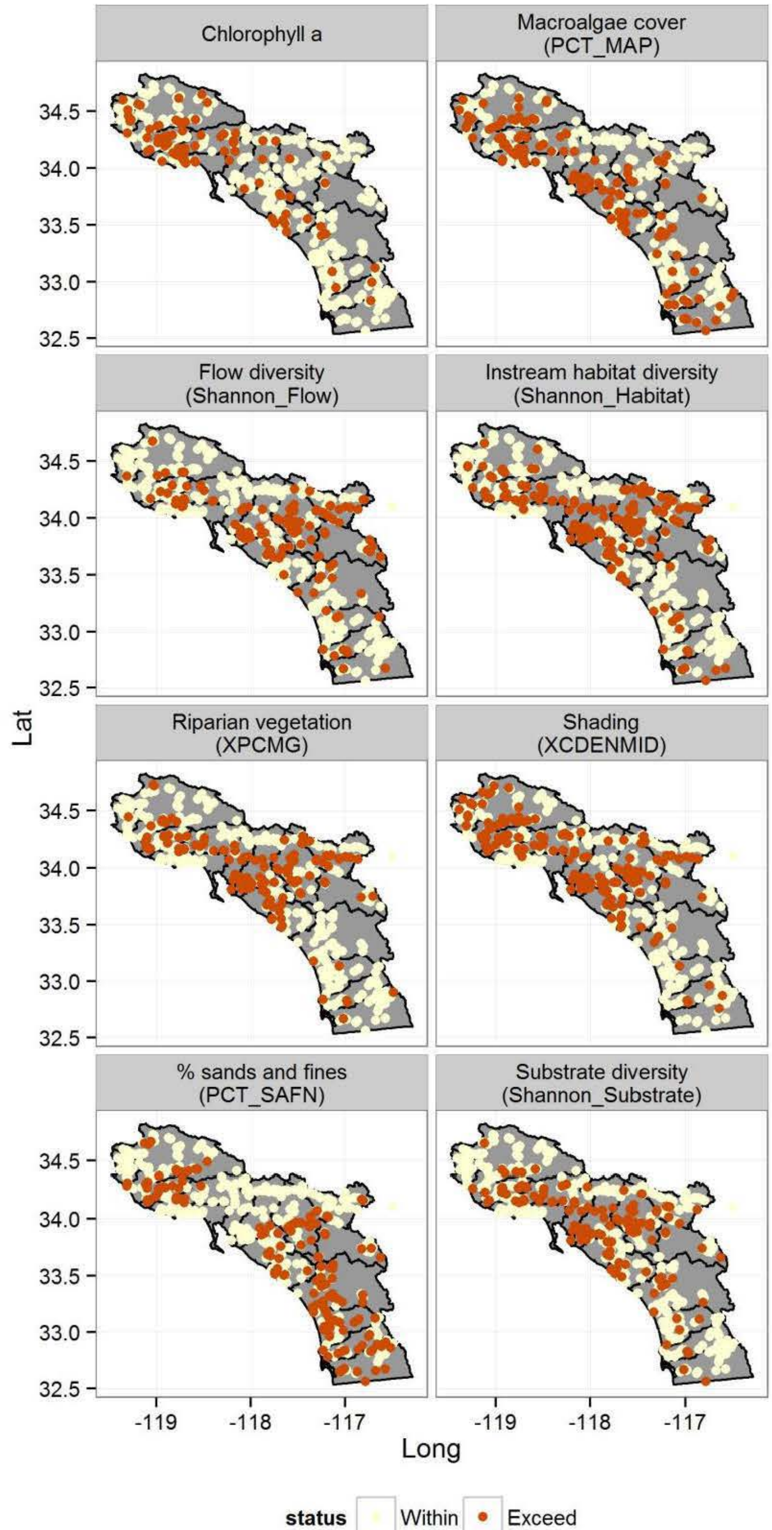


Figure 2-5. Map of selected physical habitat variables.

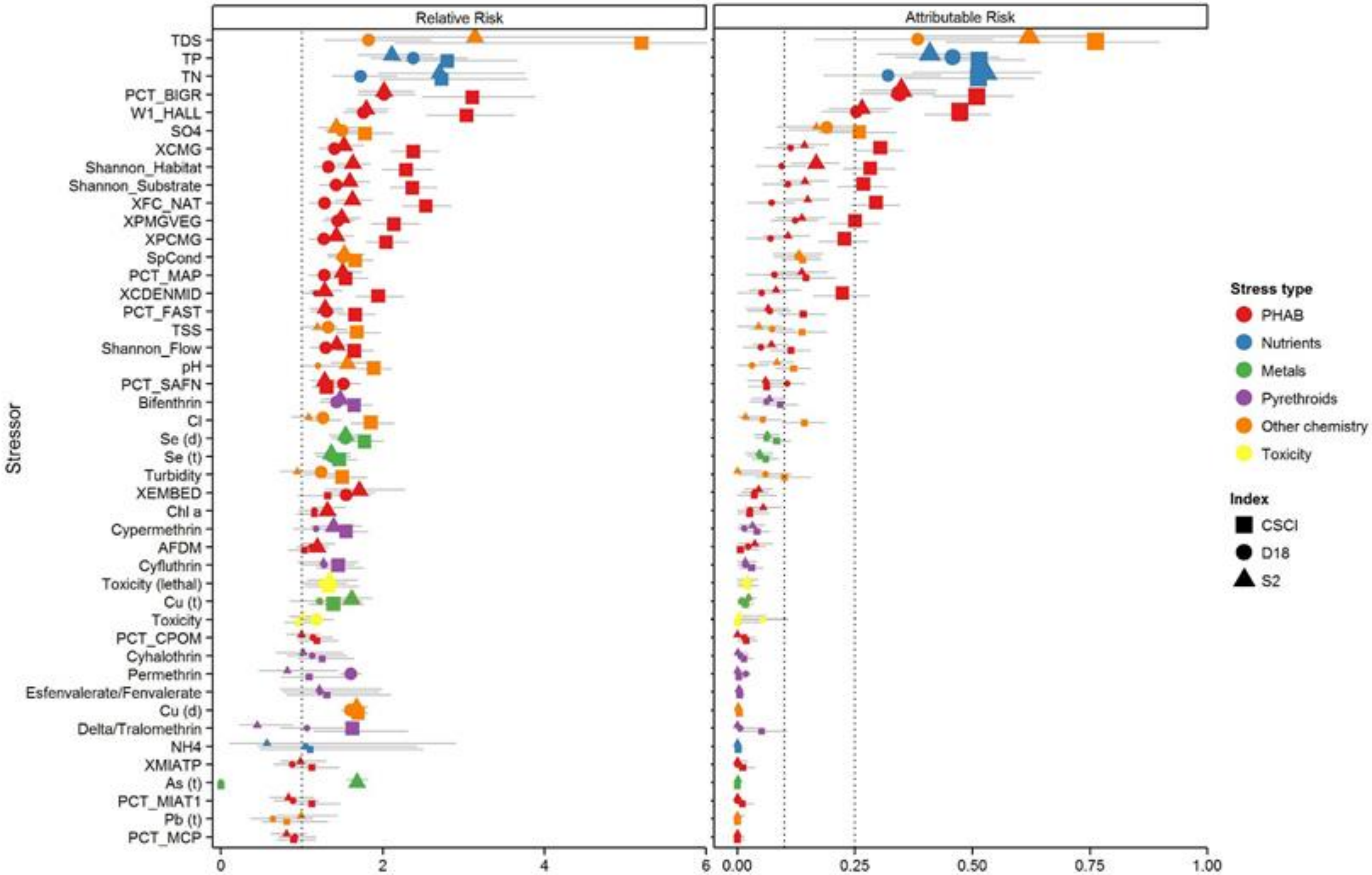


Figure 2.6. Relative and attributable risks. The horizontal lines represent the 95% confidence interval around each estimate. The dotted vertical lines represent the thresholds used to prioritize stressors.

QUESTION 3: HOW ARE BIOLOGICAL CONDITIONS CHANGING OVER TIME?

Murrieta Creek, Fall 2003

Murrieta Creek, Spring 2004

Changes in land use, such as the installation of a sand mining operation, can profoundly alter the habitat and degrade biological condition.
Photos by Scott Johnson.

Introduction

Analysis of trends allows managers to assess the effects of policies that have been implemented during the study period, the influence of disturbances like wildfire, or other activities that might change the biological condition of streams in the region. Changes observed in the region provide context to understanding site specific changes. For example, if conditions deteriorate in less disturbed areas (such as open streams), then degradation observed at an urban site might be attributable to regional stressors, such as climate change or atmospheric deposition of nutrients, rather than to management activities.

Methods

Data Collection

Data were collected as described in the Survey Overview.

Data Aggregation

Where multiple samples were collected at a single site within a year, data were aggregated as the maximum value within a site. Missing values were ignored for all relevant analyses, where appropriate.

Thresholds

Thresholds were applied as described in the section on Question 1.

Weighted Magnitudes and Extent Estimates

Weighted estimates were calculated as described in the section on Question 1, using each year (or year within land use class) as a stratum. Extents of streams in each condition class were estimated for the CSCI, S2, D18, and CRAM. In addition, the extent of streams intact for all indicators was estimated as well.

Results

All data used in this report can be downloaded from ftp.sccwrp.org/pub/download/SMCReport/SMCDataFor5yearReport.zip.

Since 2009, no obvious trends were evident for any indicator, although all indicators showed a slight depression in scores in the year 2010 (Tables 3-1 and 3-2; Figure 3-1). The median score for the CSCI, S2, and CRAM fluctuated between Class 2 and 3, while D18 fluctuated between Class 3 and 4. The percent of streams that were intact for all four indicators was highest (at 36%) in 2012, but was only 14% in 2010 (Figure 3-2). Most of the fluctuations in score affected the open streams, while the extent of healthy agricultural and urban streams remained low throughout the survey (Table 3-1, 3-2). Extent estimates were particularly imprecise for agricultural streams, as in some years very few of these sites were sampled (e.g., 5 agricultural sites were sampled for all indicators in 2011 and 2012), leading to erratic confidence intervals

(Figure 3-1). Although CSCI scores were generally high in the earlier years of the survey, these estimates were based on very small sample sizes (<25 sites in any year), and should be interpreted with caution.

Discussion

We were unable to detect trends in condition. Our inability to detect trends stems from the relatively short time frame of the survey (i.e., 5 years), as well as a study design that did not include site revisits over multiple years. These two characteristics of the survey make it difficult to distinguish trends from natural variation driven by climate or other factors. Given that a different set of sites was examined each year, the regional focus of the program, and that only five years of data are presented, it is not surprising that no distinct trends were observed. For a trend at this regional scale to be evident, a longer time period would be required and/or site revisits. It is possible that site-specific management activities affecting stream health were within the sample frame, but may have been obscured by the overall regional focus. Revisiting sites sampled in early years of this survey would provide site-specific trend estimates, which could then provide a better estimate of trends across the region. Additionally, we would be able to explore potential drivers of any observed trends.

Table 3-1. Medians for key indicators by year.

Subpopulation	2009	2010	2011	2012	2013
South Coast					
CSCI	0.71	0.70	0.81	0.80	0.65
D18	55	50	54	59	57
S2	37	34	39	43	50
CRAM	71	62	72	69	67
Agricultural					
CSCI	0.70	0.74	0.79	0.79	0.71
D18	49	49	67	61	37
S2	25	17	17	41	38
CRAM	64	66	66	74	72
Open					
CSCI	0.95	0.77	0.93	0.95	0.96
D18	75	67	68	71	75
S2	83	75	52	68	61
CRAM	82	78	83	82	84
Urban					
CSCI	0.65	0.52	0.61	0.67	0.53
D18	52	41	41	39	35
S2	33	26	27	33	48
CRAM	56	45	40	37	52

Table 3-2. Percent of stream-miles within the 10th percentile of scores at reference sites for each year

Subpopulation	2009	2010	2011	2012	2013
South Coast					
CSCI	41	28	56	52	36
D18	41	35	38	45	43
S2	34	41	36	44	59
CRAM	46	34	50	48	39
Multiple indicators	23	14	24	36	31
Agricultural					
CSCI	42	39	47	35	39
D18	28	19	61	33	42
S2	15	4	19	28	17
CRAM	25	36	35	77	51
Multiple indicators	2	8	0	40	22
Open					
CSCI	84	46	88	87	82
D18	70	62	60	71	79
S2	70	86	54	84	72
CRAM	87	70	91	85	89
Multiple indicators	57	34	51	83	79
Urban					
CSCI	8	12	19	17	7
D18	20	24	17	26	20
S2	11	23	19	12	58
CRAM	23	13	12	15	11
Multiple indicators	1	4	0	1	3

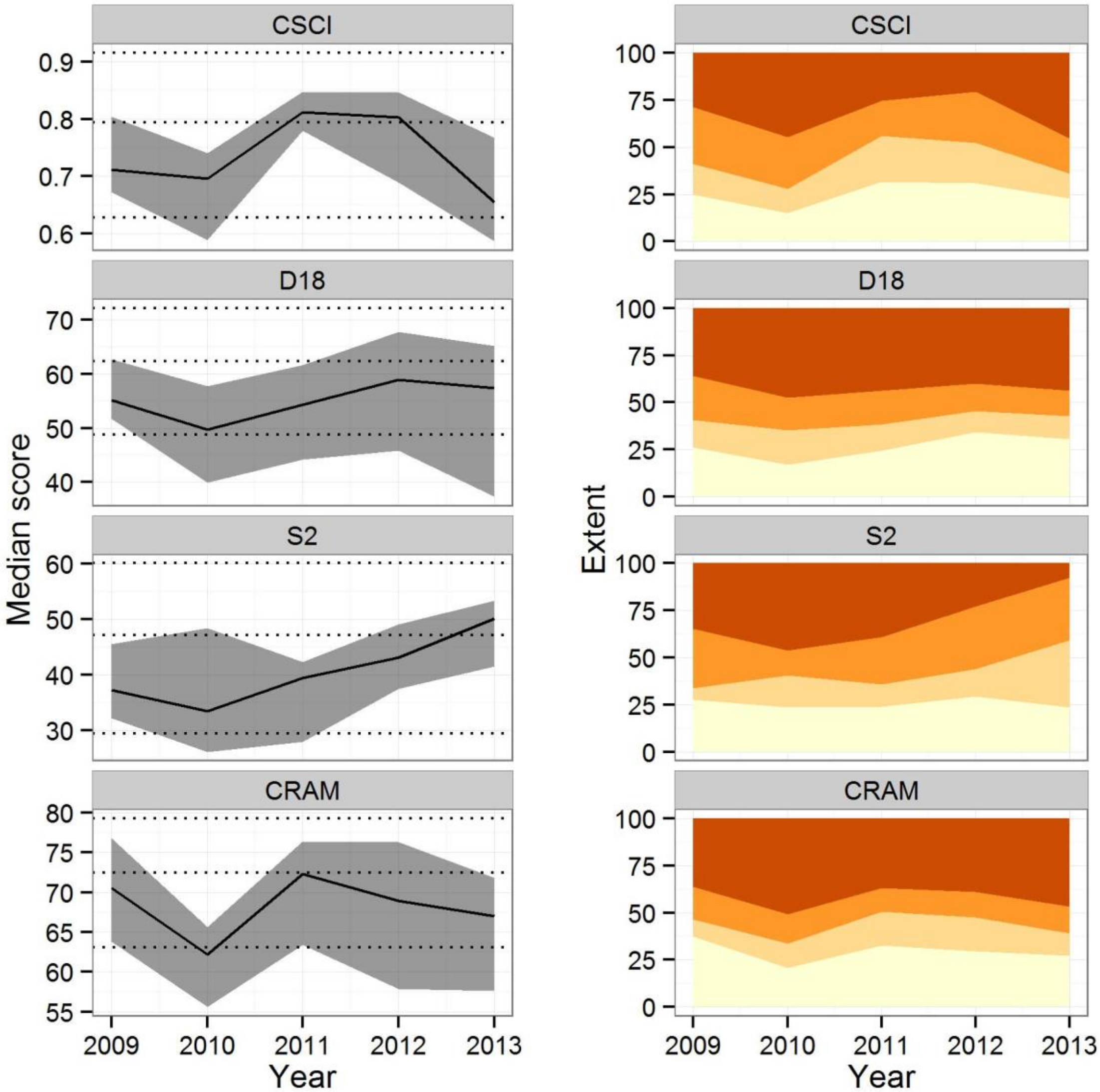


Figure 3-1. Median score and extent of condition classes by year for each indicator. The gray band in the left panel indicates the 95% confidence interval. Color in the right panel indicates condition class; lighter colors indicate better condition.

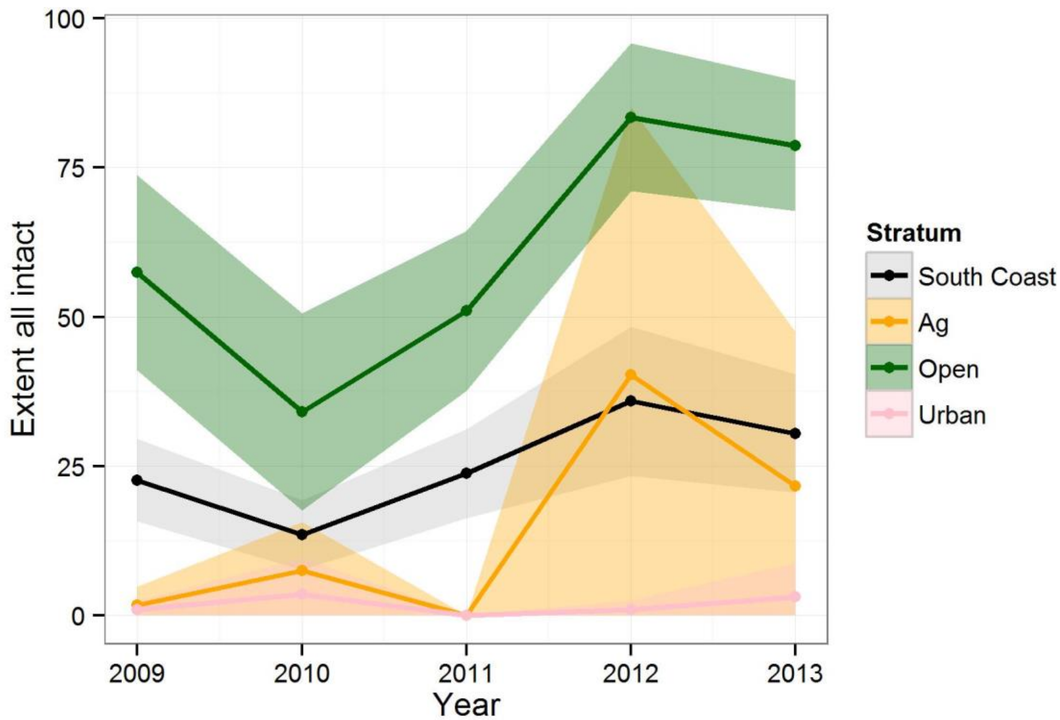


Figure 3-2. Percent of stream-miles that were intact for all four indicators

RECOMMENDATIONS

- Continue the survey for another five years, focusing on key biological indicators of stream condition, as well as high-priority stressors.
- Expand the survey to include nonperennial streams.
- Improve trend estimates by revisiting previously sampled probabilistic sites.
- Continue to investigate high priority stressors, such as habitat degradation and nutrient enrichment.
- Support studies that identify constraints on biological condition imposed by natural factors, channel engineering, water chemistry, and habitat degradation.

LITERATURE CITED

Aboal, M., M.A. Puig, P. Mateo and E. Perona. 2002. Implications of cyanophyte toxicity on biological monitoring of calcareous streams in north-east Spain. *Journal of Applied Phycology* 14:49-56

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White and M.L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15:185-211.

Bernal, S., D. Von Schiller, F. Sabater and E. Marti. 2013. Hydrological extremes modulate nutrient dynamics in Mediterranean climate streams across different spatial scales. *Hydrobiologia* 719:31-42.

Bernstein, B., and K. Schiff. 2002. Stormwater Research Needs in Southern California. SCCWRP Technical Report 358. Costa Mesa, CA. Available from http://www.socalsmc.org/Docs/StormwaterResearchNeedsinSouthernCalifornia.pdf

California Department of Forestry and Fire Protection. 2004. The California Interagency Watershed Map of 1999, version 2.2.1.

Central Valley Regional Water Quality Control Board. 2014. Central Valley Pyrethroid TMDL and Basin Plan Ammendment [Draft]. Available from http://www.waterboards.ca.gov/centralvalley/water_issues/tmdl/central_valley_projects/central_valley_pesticides/pyrethroid_tmdl_bpa/index.shtml.

Collins, J.N., E.D. Stein, M. Sutula, R. Clark, A.E. Fetscher, L. Grenier, C. Grosso, and A. Wiskind. 2008. California Rapid Assessment (CRAM) for Wetlands, v5.0.2. 157 pp. San Francisco Estuary Institute. Oakland, CA.

Environmental Protection Agency (EPA). 1986. Quality Criteria for Water. EPA 440/5-86-001. US Environmental Protection Agency, Office of Water. Washington, DC.

EPA. 2000. Establishment of numeric criteria for priority toxic pollutants for the state of California. *Federal Register* 65:31682-31719.

Fetscher, A.E., L.B. Busse and P.R. Ode. 2009. Standard Operating Procedures for Collecting Stream Algae Samples and Associated Physical Habitat and Chemical Data for Ambient Bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 002. (updated May 2010). SWAMP. Sacramento, CA.

Fetscher, A.E., R. Stancheva, J.P. Kociolek, R.G. Sheath, E.D. Stein, R.D. Mazor, P.R. Ode and L.B. Busse. 2014. Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination. *Journal of Applied Phycology* 26:433-450.

Gibson, J.R., M.T. Barbour, J.B. Stribling, J. Gerritsen, and J.R. Karr. 1996. Biological Criteria: Technical Guidance for Streams and Rivers (revised edition). EPA 822-B-96-001. US Environmental Protection Agency, Office of Water. Washington, DC.

Herlihy, A.T. and J.C. Sifneos. 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. *Journal of the North American Benthological Society* 27:932-948.

Horvitz, D.G. and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47:663-685.

Kaufmann, P.R., P. Levine, E.G. Robinson, C. Seeliger and D.V. Peck. 1999. Quantifying Physical Habitat in wadeable streams. EPA/620/R-99/003. US Environmental Protection Agency, Research Ecology Branch. Corvallis, OR.

Kincaid, T.M., and A.R. Olsen. 2013. Spsurvey: Spatial survey design and analysysis. R package version 2.6. R Foundation for Statistical Computing. Vienna, Austria.

Larsen, D.P., A.R. Olsen and D.L. Stevens. 2008. Using a master sample to integrate stream monitoring programs. *Journal of Agricultural, Biological, and Environmental Statistics* 13:243-254.

Mazor, R.D., K. Schiff, K. Ritter, A. Rehn and P. Ode. 2010. Bioassessment tools in novel habitats: An evaluation of indices and sampling methods in low-gradient streams in California. *Environmental Monitoring and Assessment* 167:91-104.

Mazor, R.D., A. Rehn, P.R. Ode, M. Engeln, K. Schiff, E.D. Stein, D. Gillett, D.B. Herbst and C.P. Hawkins. In Press. Bioassessment in complex environments: Designing an index for consistent meaning in different settings. Freshwater Science.

Mazor, R.D., E.D. Stein, P.R. Ode and K. Schiff. 2014. Integrating intermittent streams into watershed assessments: Applicability of an index of biotic integrity. *Freshwater Science* 33:459-474.

National Oceanic and Atmospheric Administration (NOAA). 2001. Coastal Change Analysis Program (C-CAP): Guidance for Regional Implementation. Technical Report NMFS 123. Department of Commerce. Available from http://www.csc.noaa.gov/crs/lca/pdf/protocol.pdf.

Ode, P.R. 2007. Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Surface Water Ambient Monitoring Program. Sacramento, CA.

Ode, P.R., A.C. Rehn, R.D. Mazor, K.C. Schiff, E.D. Stein, J.T. May, L.R. Brown, D.B. Herbst, D. Gillett, K. Lunde and C.P. Hawkins. In review. Evaluating the adequacy of a reference site pool for the ecological assessment of streams in environmentally complex regions.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchek, F.H. McCormick, S.A. Peterson, S.A. Ringold, T. Magee and M. Cappaert. 2006. Environmental Monitoring and Assessment Program - Surface Waters Western Pilot Study: field operations manual for wadeable streams. EPA/620/R-06/003. US Environmental Protection Agency, Office of Research and Development. Corvallis, OR.

R Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day and D.M. Rosenberg. 1997. The reference condition: A comparison of multimetric and multivariate approaches to assess water quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* 16:833-852.

Richards, A.B. and D.C. Rogers. 2011. List of Freshwater Macroinvertebrate Taxa from California and Adjacent States Including Standard Taxonomic Effort Levels. Southwest Association of Freshwater Invertebrate Taxonomists. Chico, CA.

Stevens, D.L. and A.R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association* 99:262-278.

SWAMP QAT. 2008. Quality Assurance Program Plan. Surface Water Ambient Monitoring Program Quality Assurance Team. Sacramento, CA.

United State Environmental Protection Agency (USEPA). 2002. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. Fourth Edition. EPA-821-R-02-013. USEPA, Office of Water. Washington, DC.

United States Geological Survey (USGS). 1999. National Elevation Database. Sioux Falls, SD. Accessed from: http://ned.usgs.gov.

United States Geological Survey and United States Environmental Protection Agency. 2005. National Hydrography Dataset Plus. Reston, VA.

Van Sickle, J., J.L. Stoddard, S.G. Paulsen and A.R. Olsen. 2006. Using relative risk to compare the effects of aquatic stressors at a regional scale. *Environmental Management* 38:1020-1030.

Van Sickle, J. and S.G. Paulsen. 2008. Assessing the attributable risks, relative risks, and regional extents of aquatic stressors. *Journal of the North American Benthological Society* 27:920-931.

Yoon, V.K. and E.D. Stein. 2008. Natural catchments as sources of background levels of stormwater metals, nutrients, and solids. *Journal of Environmental Engineering* 134:961-973

Appendix 2

Appendix 7-B

DOI: 10.1007/s00267-004-0035-8

A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams

PETER R. ODE*
Aquatic Bioassessment Laboratory
Water Pollution Control Laboratory
Department of Fish and Game
2005 Nimbus Road
Rancho Cordova, California 95670, USA

ANDREW C. REHN
Aquatic Bioassessment Laboratory
Chico State University Research Foundation
2005 Nimbus Road
Rancho Cordova, California 95670, USA

JASON T. MAY
3421 I Street 1
Sacramento, California, 95816, USA

ABSTRACT / We developed a benthic macroinvertebrate index of biological integrity (B-IBI) for the semiarid and populous southern California coastal region. Potential reference sites were screened from a pool of 275 sites, first with quantitative GIS landscape analysis at several spatial scales and then with local condition assessments (in-stream and riparian) that quantified stressors acting on study reaches. We screened 61 candidate metrics for inclusion in the B-IBI based on three criteria: sufficient range for scoring, responsiveness to watershed and reach-scale disturbance gradients, and minimal correlation with other responsive metrics. Final metrics included: percent collector-gatherer + collector-filterer individuals, percent noninsect taxa, percent tolerant taxa, Coleoptera richness, predator richness, percent intolerant individuals, and EPT richness. Three metrics had lower scores in chaparral reference sites than in mountain reference sites and were scored on separate scales in the B-IBI. Metrics were scored and assembled into a composite B-IBI, which was then divided into five roughly equal condition categories. PCA analysis was used to demonstrate that the B-IBI was sensitive to composite stressor gradients; we also confirmed that the B-IBI scores were not correlated with elevation, season, or watershed area. Application of the B-IBI to an independent validation dataset (69 sites) produced results congruent with the development dataset and a separate repeatability study at four sites in the region confirmed that the B-IBI scoring is precise. The SoCal B-IBI is an effective tool with strong performance characteristics and provides a practical means of evaluating biotic condition of streams in southern coastal California.

Assemblages of freshwater organisms (e.g., fish, macroinvertebrates, and periphyton) are commonly used to assess the biotic condition of streams, lakes, and wetlands because the integrity of these assemblages provides a direct measure of ecological condition of these water bodies (Karr and Chu 1999). Both multimetric (Karr and others 1986; Kerans and Karr 1994; McCormick and others 2001; Klemm and others 2003) and multivariate (Wright and others 1983; Hawkins and others 2000; Reynoldson and others 2001) methods have been developed to characterize biotic condition and to establish thresholds of ecological impairment. In both approaches, the ability to recognize degradation at study sites relies on an understanding of the organismal assemblages expected in the absence of disturbance. Thus, the adoption of a consistent and quantifiable method for defining reference condition is fundamental to any biomonitoring program (Hughes 1995).

Southern California faces daunting challenges in the conservation of its freshwater resources due to its aridity, its rapidly increasing human population, and its role as one of the world's top agricultural producers. In recent years, several state and federal agencies have become increasingly involved in developing analytical tools that can be used to assess the biological and physical condition of California's streams and rivers. For example, the US Environmental Protection Agency (EPA), the US Forest Service (USFS), and California's state and regional Water Quality Control Boards (WQCBs) have collected fish, periphyton and benthic macroinvertebrates (BMIs) from California streams and rivers as a critical component of regional water

KEY WORDS: Benthic macroinvertebrates; B-IBI; Biomonitoring; Mediterranean climate

Published online May 10, 2005.

*Author to whom correspondence should be addressed; *email:* pode@ospr.dfg.ca.gov

quality assessment and management programs. Together, these agencies have sampled BMIs from thousands of sites in California, but no analysis of BMI assemblage datasets based on comprehensively defined regional reference conditions has yet been undertaken. In the only other large-scale study within the state, Hawkins and others (2000) developed a predictive model of biotic integrity for third- to fourth-order streams on USFS lands in three montane regions in northern California. This ongoing effort (Hawkins unpublished) is an important contribution to bioassessment in the state, but the emphasis of this work has been concentrated on logging impacts within USFS lands. The lack of a broadly defined context for interpretation of BMI-based bioassessment remains the single largest impediment to the development of biocriteria for the majority of California streams and rivers. This article presents a benthic index of biotic integrity (B-IBI) for wadeable streams in southern coastal California assembled from BMI data collected in the region by the USFS, EPA, and state and regional WQCBs between 2000 and 2003.

Methods

Study Area

The Southern Coastal California B-IBI (SoCal B-IBI) was developed for the region bounded by Monterey County in the north, the Mexican border in the south, and inland by the eastern extent of the southern Coast Ranges (Figure 1). This Mediterranean climate region comprises two Level III ecoregions (Figure 1; Omernik 1987) and shares a common geology (dominated by recently uplifted and poorly consolidated marine sediments) and hydrology (precipitation averages 10–20 in./year in the lower elevations and 20–30 in./year in upper elevations, reaching 30–40 in./year in the highest elevations and in some isolated coastal watersheds (Spatial Climate Analysis Service, Oregon State University, www.climatesource.com). The human population in the region was approximately 20 million in 2000 and is projected to exceed 28 million by 2025 (California Department of Finance, Demographic Research Unit, www.dof.ca.gov).

Field Protocols and Combining Datasets

The SoCal B-IBI is based on BMI and physical habitat data collected from 275 sites (Figure 1) using the 3 protocols described in the following subsections. Sites were sampled during base flow periods between April and October of 2000–2003.

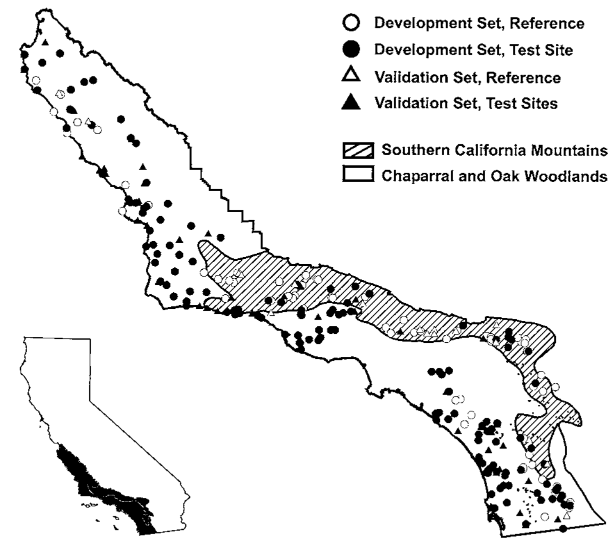


Figure 1. Map of study area showing the location of the study area within California, the distribution of test and reference sites and development and validation sites, and the boundaries of the two main ecoregions in the study area.

California Stream Bioassessment Protocol (CSBP, 144 sites). Several of the regional WQCBs in southern coastal California have implemented biomonitoring programs in their respective jurisdictions and have collected BMIs according to the CSBP (Harrington 1999). At CSBP sites, three riffles within a 100-m reach were randomly selected for sampling. At each riffle, a transect was established perpendicular to the flow, from which three separate areas of 0.18 m^2 each were sampled upstream of a 0.3-m-wide D-frame net and composited by transect. A total of 1.82 m^2 of substrate was sampled per reach and 900 organisms were subsampled from this material (300 organisms were processed separately from each of 3 transects). Water chemistry data were collected in accordance with the protocols of the different regional WQCBs (Puckett 2002) and qualitative physical habitat characteristics were measured according to Barbour and others (1999) and Harrington (1999).

USFS (56 sites). The USFS sampled streams on national forest lands in southern California in 2000 and 2001 using the targeted riffle protocol of Hawkins and others (2001). All study reaches were selected nonrandomly as part of a program to develop an interpretive (reference) framework for the results of stream biomonitoring studies on national forests in California. BMIs were sampled at study reaches (containing at least four fast-water habitat units) by disturbing two separate 0.09-m^2 areas of substrate upstream of a 0.3-m-wide D-frame net in each of four separate fast-water units; a total of 0.72 m^2 was disturbed and all sample material from a reach was composited. Field crews used a combination of qualitative and quantitative measures to collect physical habitat and water chemistry data (Hawkins and others 2001). A 500-organism subsample was processed from the composite sample and identified following methods described by Vinson and Hawkins (1996).

Environmental Monitoring and Assessment Program (EMAP, 75 sites). The EPA sampled study reaches in southern coastal California from 2000 through 2003 as part of its Western EMAP pilot project. A sampling reach was defined as 40 times the average stream width at the center of the reach, with a minimum reach length of 150-m and maximum length of 500-m. A BMI sample was collected at each site using the USFS methodology described earlier (Hawkins and others 2001) in addition to a standard EMAP BMI sample (not used in this analysis). A 500-organism subsample was processed in the laboratory according to EMAP standard taxonomic effort levels (Klemm and others 1990). Water chemistry samples were collected from the midpoint of each reach and analyzed using EMAP protocols (Klemm and Lazorchak 1994). Field crews recorded physical habitat data using EPA qualitative methods (Barbour and others 1999) and quantitative methods (Kaufmann and others 1999).

As part of a methods comparison study, 77 sites were sampled between 2000 and 2001 with both the CSBP and USFS protocols. The two main differences between the methods are the area sampled and the number of organisms subsampled (discussed earlier). To determine the effect of sampling methodology on assessment of biotic condition, we compared the average difference in a biotic index score between the two methods at each site. Biotic index scores were computed with seven commonly used biotic metrics (taxonomic richness, Ephemoptera, Plecoptera, and Trichoptera (EPT) richness, percent dominant taxon, sensitive EPT individuals, Shannon diversity, percent intolerant taxa, and percent scraper individuals) according to the following equation:

$$Score = \sum (x_i - \bar{x})/sem_i$$

where x_i is the site value for the *i*th metric, x is the overall mean for the *i*th metric, and SEM_i is the standard error of the mean for the *i*th metric. A score of zero is the mean value.

Because USFS-style riffle samples were collected at all EMAP sites, only two field methods were combined in this study. All EMAP and CSBP samples were collected and processed by the California Department of Fish and Game's Aquatic Bioassessment Laboratory (ABL) and all USFS samples were processed by the US Bureau of Land Management's Bug Lab in Logan, Utah. Taxonomic data from both labs were combined in an MS Access© database application that standardized BMI taxonomic effort levels and metric calculations, allowing us to minimize any differences between the two labs that processed samples. Taxonomic effort followed standards defined by the California Aquatic Macroinvertebrate Laboratory Network (CAMLnet 2002; www.dfg.ca.gov/cabw/camlnetste.pdf). Sites with fewer than 450 organisms sampled were omitted from the analyses.

Screening Reference Sites

We followed an objective and quantitative reference site selection procedure in which potential reference sites were first screened with quantitative Geographical Information System (GIS) land-use analysis at several spatial scales and then local condition assessments (instream and riparian) were used to quantify stressors acting within study reaches. We calculated the proportions of different land-cover classes and other measures of human activity upstream of each site at four spatial scales that give unique information about potential stressors acting on each site: (1) within polygons delimiting the entire watershed upstream of each sampling site, (2) within polygons representing local regions (defined as the intersection of a 5-km-radius circle around each site and the primary watershed polygon), (3) within a 120-m riparian zone on each side of all streams within each watershed, and (4) within a 120-m riparian zone in the local region. We used the ArcView® (ESRI 1999) extension ATtILA (Ebert and Wade 2002) to calculate the percentage of various land-cover classes (urban, agriculture, natural, etc.) and other measures of human activity (population density, road density, etc.) in each of the four spatial areas defined for each site. Two satellite imagery datasets from the mid-1990s were combined for the land-cover analyses: California Land Cover Mapping & Monitoring Program (LCMMP) vegetation data (CalVEG) and a recent dataset produced by the Central Coast Watershed Group (Newman and Watson 2003). Population data were derived from the 2000 migrated TIGER dataset (California Department of Forestry and Fire Protection, www.cdf.ca.gov). Stream layers were obtained from the US Geological Survey (USGS) National Hydrography Dataset (NHD). The road network was obtained from the California Spatial Information Library (CaSIL, gis.ca.gov) and elevation was based on the USGS National Elevation Dataset (NED). Frequency histograms of land-use percentages for all sites were used to establish subjective thresholds for elim-

Table 1. List of minimum or maximum landuse thresholds used for rejecting potential reference sites

Stressor metric	Definition	Threshold
N_index_L	Percentage of natural land use at the local scale	≤ 95%
Purb_L	Percental of urban land use at the local scale	> 3%
Pagt_L	Percentage of total agriculture at the local scale	> 5%
Rddens_L	Road density at the local scale	> 2.0 km/km^2
PopDens_L	Population density (2000 census) at the local scale	> 150 indiv./km^2
N_index_W	Percentage of natural landuse at the watershed scale	≤ 95%
Purb_W	Percentage of urban landuse at the watershed scale	> 5%
Pagt_W	Percentage of total agriculture at the watershed scale	> 3%
Rddens_W	Road density at the watershed scale	> 2.0 km/km^2
PopDens_W	Population density (2000 census) at the watershed scale	> 150 indiv./km^2

inating sites from the potential reference pool (Table 1). Sites were further screened from the reference pool on the basis of reach-scale conditions (obvious bank instability or erosion/ sedimentation problems, evidence of mining, dams, grazing, recent fire, recent logging).

Eighty-eight sites passed all the land-use and local condition screens and were selected as reference sites, leaving 187 sites in the test group. We randomly divided the full set of sites into a development set (206 sites total: 66 reference/140 test) and a validation set (69 sites total: 22 reference/47 test). The development set was used to screen metrics and develop scoring ranges for component B-IBI metrics; the validation set was used for an independent evaluation of B-IBI performance.

Screening Metrics and Assembling the B-IBI

Sixty-one metrics were evaluated for possible use in the SoCal B-IBI (Table 2). A multistep screening process was used to evaluate each metric for (1) sufficient range to be used in scoring, (2) responsiveness to watershed-scale and reach-scale disturbance variables, and (3) lack of correlation with other responsive metrics.

Pearson correlations between all watershed-scale and reach-scale disturbance gradients were used to define the smallest suite of independent (nonredundant) disturbance variables against which to test biological metric response. Disturbance variables with correlation coefficients $|r| \geq 0.7$ were considered redundant. Responsiveness was assessed using visual inspection of biotic metric versus disturbance gradient scatterplots and linear regression coefficients. Metrics were selected as responsive if they showed either a linear or a "wedge-shaped" relationship with disturbance gradients. Biological metrics often show a "wedge-shaped" response rather than a linear response to single disturbance gradients because the single gradient only defines the upper boundary of the biological response; other independent disturbance gradients and natural limitations on species distributions might result in lower metric values than expected from response to the single gradient. Biotic metrics and disturbance gradients were log-transformed when necessary to improve normality and equalize variances. Metrics that passed the range and responsiveness tests were tested for redundancy. Pairs of metrics with product–moment correlation coefficients $|r| \geq 0.7$ were considered redundant and the least responsive metric of the pair was eliminated.

Scoring ranges were defined for each metric using techniques described in Hughes and others (1998), McCormick and others (2001), and Klemm and others (2003). Metrics were scored on a 0–10 scale using statistical properties of the raw metric values from both reference and nonreference sites to define upper and lower thresholds. For positive metrics (those that increase as disturbance decreases), any site with a metric value equal to or greater than the 80th percentile of reference sites received a score of 10; any site with a metric value equal to or less than the 10th percentile of the nonreference sites received a score of 0; these thresholds were reversed for negative metrics (20th percentile of reference and 90th percentile of nonreference). In both cases, the remaining range of intermediate metric values was divided equally and assigned scores of 1 through 9. Before assembling the B-IBI, we tested whether any of the final metrics were significantly different between chaparral and mountain reference sites in the southern California coastal region, in which case they would require separate scoring ranges in the B-IBI. Finally, an overall B-IBI score was calculated for each site by summing the constituent metric scores and adjusting the B-IBI to a 100-point scale.

Table 2. The 61 BMI metrics screened for use in the SoCal IBI

Candidate metrics	Disturbance variables										
	U_index_W	Pagt_W	Purb_L	RdDens_L	Channel Alteration	Bank Stability	Percent Fines	Total Dissolved Solids	Total Phosphorus	Total Nitrogen	Range Test
Taxonomic group metrics											
Coleoptera richness*	M	w	M	S	S	—	—	—	—	—	P
Crustacea + Mollusca richness	—	—	—	—	—	—	—	—	—	—	F
Diptera richness	—	—	—	—	—	—	—	—	—	—	P
Elmidae richness	w	—	w	M	S	M	S	M	—	—	F
Ephemerellidae richness					S	M	M	S	—	M	F
Ephemeroptera richness	S	S	M	S	w	M	S	S	—	S	P
EPT richness*	S	S	S	S	S	S	S	S	—	S	P
Hydropsychidae richness	—	—	w	—	S	—	—	—	—	—	F
Percent Amphipoda individuals	—	—	—	—	—	—	—	—	—	—	P
Percent Baetidae individuals	—	—	—	—	w	—	—	—	—	—	P
Percent Chironomidae individuals	—	—	—	—	—	—	—	—	M	—	P
Percent Corbicula individuals	—	—	—	—	—	—	—	—	—	—	P
Percent Crustacea individuals	—	—	—	—	—	—	—	—	—	—	P
Percent Diptera individuals	—	w	—	—	—	—	—	—	—	—	P
Percent Elmidae individuals	—	—	—	w	M	S	S	w	—	M	P
Percent Ephemeroptera individuals	—	w	w	—	M	w	—	—	—	—	P
Percent EPT individuals	—	—	M	M	M	M	—	—	—	—	P
Percent Gatropoda individuals	—	—	—	w	—	—	—	—	—	—	P
Percent Glossosomatidae individuals	—	—	—	—	w	—	—	—	—	M	F
Percent Hydropsychidae individuals	—	—	—	M	w	M	—	—	—	—	P
Percent Hydroptilidae individuals	—	—	—	M	—	w	—	—	—	—	F
Percent Mollusca individuals	—	—	—	w	w	—	—	—	—	—	P
Percent non-*Baetis/Fallceon* Ephemeroptera individuals	w	w	—	M	w	M	—	—	w	—	P
Percent non-*Hydropsyche* Hydropsychidae individuals	—	—	—	M	w	w	—	—	—	—	F
Percent non-*Hydropsyche/Cheumatopsyche* Trichoptera individuals	w	w	—	M	w	M	M	w	—	—	P
Percent non-insect Taxa*	M	w	M	M	w	—	—	—	w	M	F
Percent Oligochaeta individuals	—	—	—	—	w	—	—	—	—	—	P
Percent Perlodidae individuals	—	—	—	w	w	—	—	—	—	—	F
Percent Plecoptera individuals	—	—	—	M	M	M	M	M	w	S	P
Percent Rhyacophilidae individuals	—	—	—	w	S	S	w	—	—	M	F
Percent Simuliidae individuals	—	w	—	w	S	w	—	—		—	P
Percent Trichoptera	w	—	—	M	M	M	M	w	w	—	P
Plecoptera richness	M	S	w	M	w	w	M	S	—	S	F
Total taxa richness	M	M	w	S	w	w	w	w	w	M	P
Trichoptera richness	S	S	S	S	S	M	S	w	—	w	P

Table 2. Continued.

Candidate metrics	U_index_W	Pagt_W	Purb_L	RdDens_L	Channel Alteration	Bank Stability	Percent Fines	Total Dissolved Solids	Total Phosphorus	Total Nitrogen	Range Test
	Disturbance variables										
Functional feeding metrics											
Collector (filterers) richness	w	—	M	S	S	M	w	—	—	—	F
Collector (gatherers) richness	—	—	—	—	—	—	—	—	—	w	P
Percent collector (filterer) + collector (gatherer) individuals*	M	—	—	S	—	w	—	M	w	M	P
Percent collector (filterer) individuals	—	—	—	w	M	M	w	—	—	—	P
Percent collector (gatherer) individuals	—	—	—	w	M	—	—	w	M	w	P
Percent predator individuals	—	—	—	w	M	—	—	—	—	—	P
Percent scraper individuals	w	w	—	M	M	w	w	—	—	—	P
Percent scraper minus snails individuals	—	—	—	w	—	w	—	—	—	—	P
Percent shredder individuals	—	—	—	w	w	—	—	—	—	—	P
Predator richness*	S	S	w	M	w	—	—	S	—	M	P
Scraper richness	S	M	M	S	S	S	S	S	—	S	P
Shredder richness	M	M	—	M	S	—	—	—	—	M	F
Tolerance metrics											
Average tolerance value	M	w	w	S	w	—	M	—	—	w	P
Intolerant EPT richness	M	w	w	M	S	—	S	S	—	S	P
Intolerant taxa richness	M	w	w	M	S	M	S	S	—	S	P
Percent intolerant Diptera individuals	—	—	—	—	—	—	—	—	—	—	F
Percent intolerant individuals*	M	w	—	M	S	M	M	S	—	M	P
Percent intolerant scraper individuals	—	—	—	w	M	w	w	w	—	—	P
Percent of intolerant Ephemeroptera individuals	—	—	—	w	w	—	w	w	—	—	P
Percent of intolerant Trichoptera individuals	—	w	—	—	w	w	w	w	—	—	P
Percent sensitive EPT individuals	w	w	—	M	M	M	M	M	w	M	P
Percent tolerant individuals	—	—	—	—	—	—	w	w	—	—	P
Percent tolerant taxa*	w	—	w	M	—	—	—	w	—	M	P
Tolerant taxa richness	—	—	—	—	—	M	—	—	—	—	P
Others											
Percent dominant taxon	—	—	—	—	—	—	—	—	—	—	P
Shannon Diversity Index	w	w	w	M	M	w	—	w	w	w	P

Note: Each metric is indicated as having either no response (—), weak response (w), moderate response (M), or strong response (S) to each of eleven minimally correlated disturbance variables and whether each metric passed (P) or failed (F) the range test. The final seven minimally correlated metrics are indicated with an asterisk (*).

Table 3. Scoring ranges for seven component metrics in the SoCal B-IBI

Metric score	Coleoptera taxa (all sites)	EPT taxa 6	EPT taxa 8	Predator taxa (all sites)	% Collector individuals 6	% Collector individuals 8	% Intolerant individuals 6	% Intolerant individuals 8	% Noninsect taxa (all sites)	% Tolerant taxa (all sites)
10	>5	>17	>18	>12	0–59	0–39	25–100	42–100	0–8	0–4
9		16–17	17–18	12	60–63	40–46	23–24	37–41	9–12	5–8
8	5	15	16	11	64–67	47–52	21–22	32–36	13–17	9–12
7	4	13–14	14–15	10	68–71	53–58	19–20	27–31	18–21	13–16
6		11–12	13	9	72–75	59–64	16–18	23–26	22–25	17–19
5	3	9–10	11–12	8	76–80	65–70	13–15	19–22	26–29	20–22
4	2	7–8	10	7	81–84	71–76	10–12	14–18	30–34	23–25
3		5–6	8–9	6	85–88	77–82	7–9	10–13	35–38	26–29
2	1	4	7	5	89–92	83–88	4–6	6–9	39–42	30–33
1		2–3	5–6	4	93–96	89–94	1–3	2–5	43–46	34–37
0	0	0–1	0–4	0–3	97–100	95–100	0	0–1	47–100	38–100

Note: Three metrics have separate scoring ranges for the two Omernik Level III ecoregions in southern coastal California region (6 = chaparral and oak woodlands, 8 = Southern California mountains).

Validation of B-IBI and Measurement of Performance Characteristics

To test whether the distribution of B-IBI scores in reference and test sites might have resulted from chance, we compared score distributions in the development set to those in the validation set. We also investigated a separate performance issue that ambient bioassessment studies often neglect: spatial variation at the reach scale. Although our use of a validation dataset tests whether the B-IBI scoring range is repeatable (Fore and others 1996; McCormick and others 2001), we designed a separate experiment to explicitly measure index precision. Four sites were resampled in May 2003. At each site, nine riffles were sampled following the CSBP, and material from randomly selected riffles was combined into three replicates of three riffles each. B-IBI scores were then calculated for each replicate. Variance among these replicates was used to calculate the minimum detectable difference (MDD) between two B-IBI scores based on a two-sample *t*-test model (Zar 1999). The index range can be divided by the MDD to estimate the number of stream condition categories detectable by the B-IBI (Doberstein and others 2000; Fore and others 2001).

Results

Combining Datasets

Unmodified CSBP samples (900 count) had significantly higher biotic condition scores ($t = -6.974$, $P < 0.0001$) than did USFS samples (500 count). However, there was no difference in biotic condition scores between USFS samples and CSBP samples that were randomly subsampled to reduce the 900 count to 500 ($t = -0.817$, $P = 0.416$). Thus, data from both targeted-riffle protocols were combined in B-IBI development.

Selected Metrics

Ten nonredundant stressor gradients were selected for metric screening: percent watershed unnatural, percent watershed in agriculture, percent local watershed in urban, road density in local watershed, qualitative channel alteration score, qualitative bank stability score, percent fine substrates, total dissolved solids, total nitrogen, and total phosphorous. Twenty-three biotic metrics that passed the first two screens (range and dose response) were analyzed for redundancy with Pearson product–moment correlation, and a set of seven minimally correlated metrics was selected for the B-IBI: percent collector-gatherer + collector-filterer individuals (% collectors), percent noninsect taxa, percent tolerant taxa, Coleoptera richness, predator richness, percent intolerant individuals, and EPT richness (Table 3). All metrics rejected as redundant were derived from taxa similar to those of selected metrics, but they had weaker relationships with stressor gradients. Dose–response relationships of the selected metrics to the 10 minimally correlated stressor variables are shown in Figure 2 and reasons for rejection or acceptance of all metrics are listed in Table 2. Regression coefficients were significant at the $P \leq 0.0001$ level among all seven selected metrics and at least two stressor gradients: percent watershed unnatural and road density in local watershed (Table 4). The final seven metrics included several metric types: richness, composition, tolerance measures, and func-

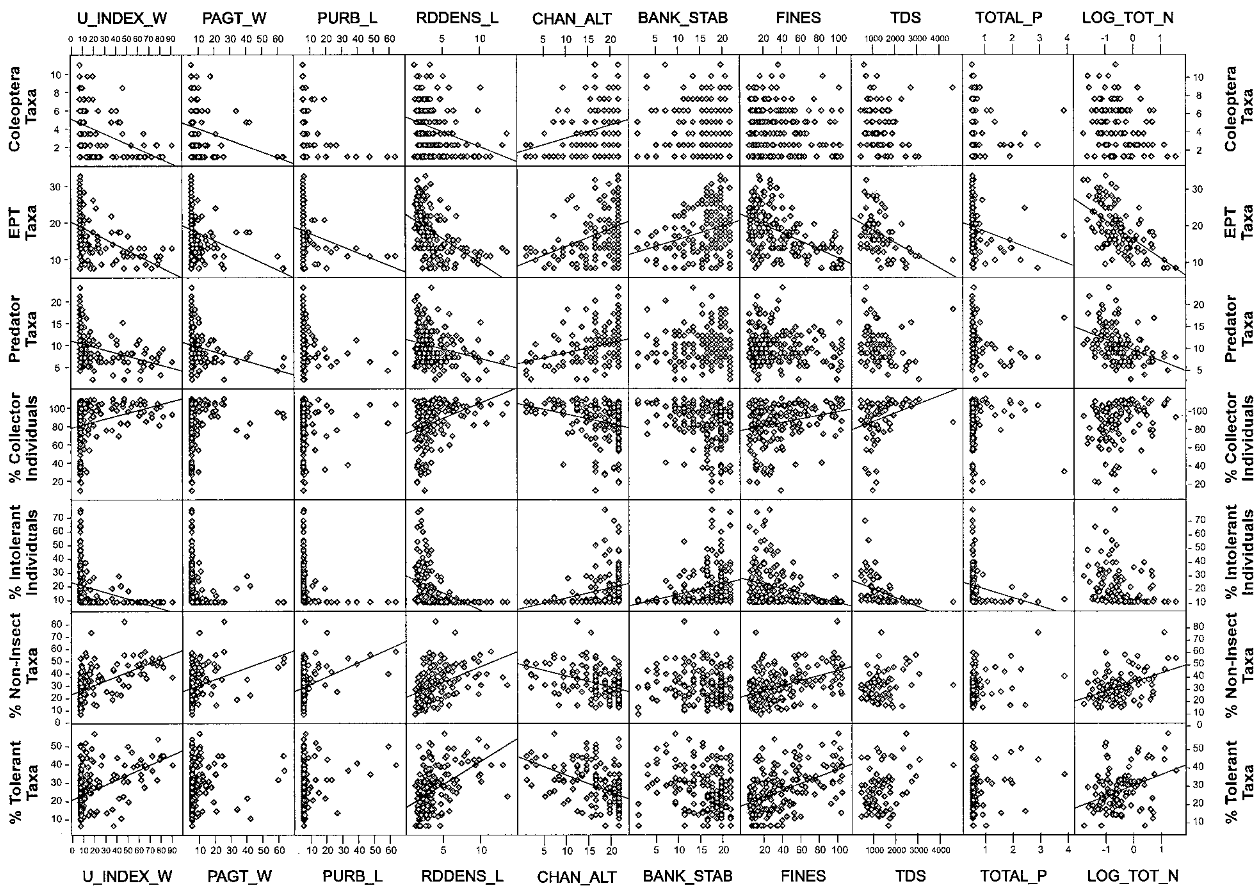


Figure 2. Scatterplots of dose–response relationships among 10 stressor gradients and 7 macroinvertebrate metrics (lines represent linear "best-fit" relationships; see text for abbreviations).

Table 4. Significance levels of linear regression relationships among 10 stressor metrics and 7 biological metrics

Metric	Coleoptera taxa	EPT taxa	Predator taxa	% Collector individuals	% Intolerant individuals	% Noninsect taxa	% Tolerant taxa
Bank Stability	0.813	**<0.0001**	0.3132	0.0009	0.0001	0.1473	0.0013
Fines	0.0017	**<0.0001**	0.0171	**0.0003**	**<0.0001**	**<0.0001**	**<0.0001**
Chan_Alt	**<0.0001**	**<0.0001**	**<0.0001**	**0.0003**	**<0.0001**	**<0.0001**	**<0.0001**
Log_U_Index_W	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
Log_PAgT_W	**0.0007**	**<0.0001**	**0.0004**	0.0054	0.0014	**<0.0001**	0.0012
Log_PUrb_L	0.0367	**0.0007**	0.0344	0.6899	0.0045	**0.0002**	0.0215
Log RdDens_L	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
Log_TDS	0.0094	**<0.0001**	0.0035	**0.0005**	**<0.0001**	0.0271	0.004
Log_Tot_N	0.0019	**<0.0001**	**<0.0001**	0.0078	0.0019	**<0.0001**	**<0.0001**
Log_Tot_P	0.062	**<0.0001**	0.0085	0.0162	**0.0001**	0.0018	0.0059

Note: Significant *P*-values corrected for 70 simultaneous comparisons ($P < 0.0007$) are highlighted in bold. Abbreviations are defined in Table 1 and in the text.

tional feeding groups. Because there are only seven metrics in the B-IBI, final scores calculated using this IBI are multiplied by 1.43 to adjust the scoring range to a 100-point scale.

The B-IBI scores were lower in chaparral reference sites than in mountain reference sites when calculated using unadjusted metric scores (Mann–Whitney *U*-test; $P = 0.02$). Although none of the final seven metrics

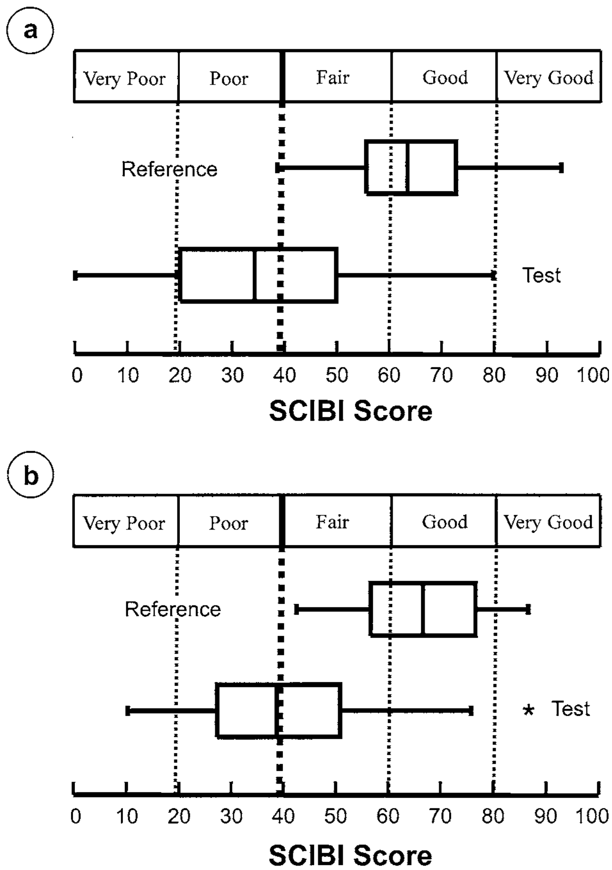


Figure 3. Box plots of B-IBI site scores for reference and test groups showing B-IBI scoring categories: (a) development sites and (b) validation sites. Dotted lines indicate condition category boundaries and heavy dotted lines indicate impairment thresholds.

were significantly different between chaparral reference sites and mountain reference sites at the $P = 0.05$ level ($P < 0.007$ after Bonferroni correction), scores for three metrics (EPT richness, percent collector-gatherer + collector-filterer individuals, and percent intolerant individuals) were substantially lower in chaparral reference sites than in mountain reference sites. We adjusted for this difference by creating separate scoring scales for the three metrics in the two ecoregions (Table 3). There was no difference in B-IBI scores between reference sites in the two ecoregions after the adjustment (Mann–Whitney *U*-test, $P = 0.364$).

Validation of B-IBI and Measurement of Performance Characteristics

The distribution of B-IBI scores at reference and nonreference sites was nearly identical between the development and validation data sets (Figure 3), indicating that our characterization of reference conditions and subsequent B-IBI scoring was repeatable and not likely due to chance. Based on a two-sample *t*-test model (setting $\alpha = 0.05$ and $\beta = 0.20$), the MDD for the SoCal IBI is 13.1. Thus, we have an 80% chance of detecting a 13.1-point difference between sites at the $P = 0.05$ level. Dividing the 100-point B-IBI scoring range by the MDD indicates that the SoCal B-IBI can detect a maximum of seven biological condition categories, a result similar to or more precise than other recent estimates of B-IBI precision (Barbour and others 1999; Fore and others 2001). We used a statistical criterion (two standard deviations below the mean reference site score) to define the boundary between "fair" and "poor" conditions, thereby setting B-IBI = 39 as an impairment threshold. The scoring range below 39 was divided into two equal condition categories, and the range above 39 was divided into three equal condition categories: 0–19 = "very poor", 20–39 = "poor", 40–59 = "fair", 60-79 = "good", and 80–100 = "very good" (Figure 3).

We ran two principle components analyses (PCAs) on the environmental stressor values used for testing metric responsiveness: 1 that included all 275 sites for which we calculated 4 watershed scale stressor values and another based on 124 sites for which we had measurements of 9 of the 10 minimally correlated stressor variables. We plotted B-IBI scores as a function of the first multivariate stressor axis from each PCA. We log-transformed percent watershed unnatural, percent watershed in agriculture, percent local watershed in urban, road density in local watershed, total nitrogen, and total phosphorous. Only PCA Axis 1 was significant in either analysis, having eigenvalues larger than those predicted from the broken-stick model (McCune and Grace 2002). In both PCAs, the B-IBI score decreased with increasing human disturbance (Figure 4) and was correlated (Spearman ρ) with PCA Axis 1 ($r = -0.652$, $P < 0.0001$ for all 275 sites; $r = -0.558$, $P \leq 0.0001$ for 124 sites). In the analysis of all 275 sites, all 4 watershed-scale stressors had high negative loadings, with percent watershed unnatural and local road density being the highest (Figure 5a). In the analysis of 124 sites, percent watershed unnatural, percent watershed in agriculture, and local road density had the highest negative loadings on the first axis, and channel alteration had the highest positive loading (Figure 4b).

Finally, we found no relationship between B-IBI scores and ecoregion (Mann–Whitney U, $P = 0.364$), Julian date ($R^2 = 0.01$, $P = 0.349$), watershed area ($R^2 = 0.002$, $P = 0.711$), or elevation ($R^2 = 0.01$, $P = 0.349$), indicating that the B-IBI scoring is robust with respect to these variables (Figure 5). Our ecoregion scoring adjustment probably corrects for the

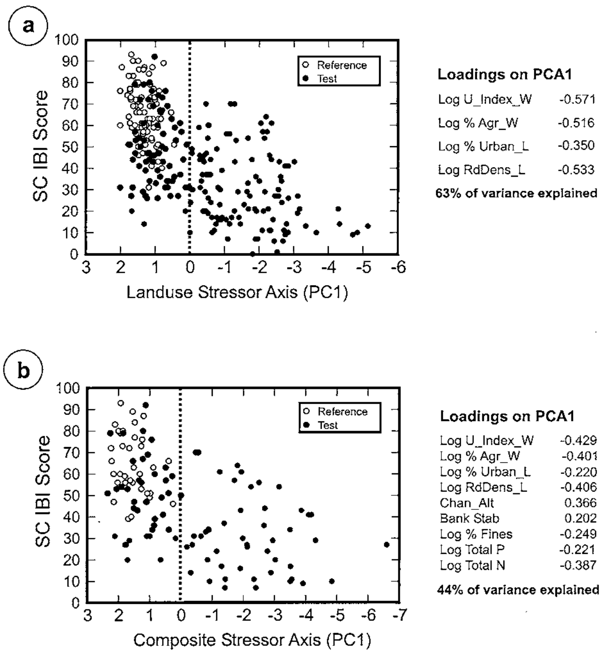


Figure 4. Scatterplots of SoCal B-IBI scores against two composite stressor axes from PCA: (a) values for all 275 sites; composite axis includes 4 land-use gradients; (b) values for 124 sites; composite axis includes 9 local and watershed scale stressor gradients.

strongest elevation effects, but there is no evidence that B-IBI scores are related to elevation differences within each ecoregion.

Discussion

The SoCal B-IBI is the most comprehensive assessment to date of freshwater biological integrity in California. As in other Mediterranean climate regions, the combination of aridity, geology, and high-amplitude cycles of seasonal flooding and drying in southern coastal California makes its streams and rivers particularly sensitive to disturbance (Gasith and Resh 1999). This sensitivity, coupled with the burgeoning human population and vast conversion of natural landscapes to agriculture and urban areas, has made it the focus of both state and federal attempts to maintain the ecological integrity of these strained aquatic resources.

Unfortunately, growing interest in biomonitoring is unmatched by financial resources available for this monitoring. Thus, combination of data among programs is very desirable, although this goal is rarely achieved in practice. We demonstrated that macroinvertebrate bioassessment data from multiple agencies could be successfully combined to produce a regional index that is useful to all agencies involved. This index is easy to apply, its fundamental assumptions are transparent, it provides precise condition assessments, and it is demonstrated to be responsive to a wide range of anthropogenic stressors. The index can also be applied throughout a long index period (mid-spring to mid-fall): Just as biotic factors tend to have more influence on assemblage structure during the summer dry period of Mediterranean climates than during the wet season when abiotic factors dominate (Cooper and others 1986; Gasith and Resh 1999), it is likely that our biotic index is more sensitive to anthropogenic stressors during the summer dry period. Because of these qualities, we expect the SoCal B-IBI to be a practical management tool for a wide range of water quality applications in the region.

This B-IBI is a regional adaptation of an approach to biotic assessment developed by Karr (1981) and subsequently extended and refined by many others (Kerans and Karr 1994; Barbour and others 1996; Fore and others 1996; Hughes and others 1998). We drew heavily upon recent refinements in multimetric index methodology that improve the objectivity and defensibility of these indices (McCormick and others 2001; Klemm and others 2003). A central goal of bioassessment is to select metrics that maximize the detection of anthropogenic stress while minimizing the noise of natural variation. One of the most important recent advances in B-IBI methods is the emphasis on quantitative screening tools for selecting appropriate metrics. We also minimized sources of redundancy in the analysis: (1) between watershed and local-scale stressor gradients for dose–response screening of biotic metrics and (2) in the final selection of metrics. The former guards against a B-IBI that is biased toward a set of highly correlated stressors and is, therefore, of limited sensitivity; the latter assures a compact B-IBI with component metrics that contribute independent information about stream condition. Combined with an assessment of responsiveness to specific regional disturbance gradients, these screening tools minimize the variability of B-IBI scores and improve its sensitivity.

The seven component metrics used in this B-IBI are similar to those selected for other B-IBIs (DeShon 1995; Barbour and others 1995, 1996; Fore and others 1996; Klemm and others 2003), but some of the metrics are either unique or are variations on other commonly used metrics. Like Klemm and others (2003), we found noninsect taxa to be responsive to human stressors, but richness was more responsive than percent of individuals. Some authors have separated the EPT metric into two or three metrics based on its component orders because the orders provided unique signals (Clements 1994; Fore and others 1996; Klemm

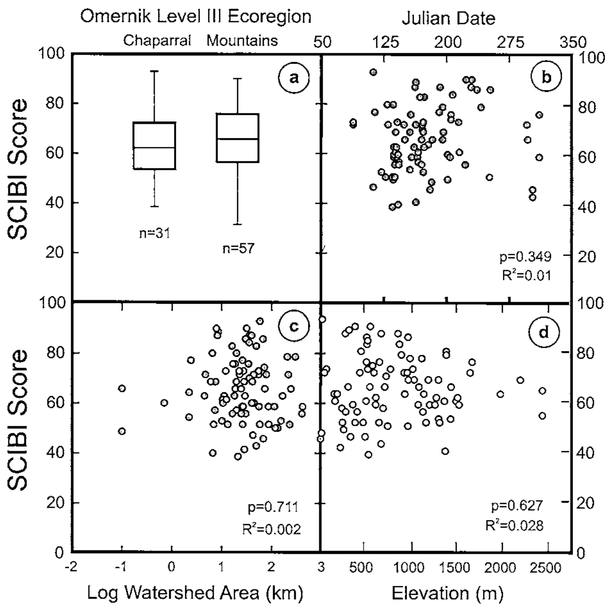


Figure 5. Relationship between B-IBI scores at 88 reference sites and (a) Omernik Level III ecoregion, (b) Julian date, (c) log watershed area, and (d) elevation.

and others 2003), but we found very similar patterns in these orders' response to various stressors we measured. To our knowledge, Coleoptera richness has not previously been included in a B-IBI, but beetle taxa might be a good indicator of the effects of fine sediments at impaired sites in this region (Brown 1973). A recent study of benthic assemblages in North Africa noted a high correspondence between EPT and EPTC (EPT + Coleoptera) (Beauchard and others 2003), but these orders were not highly correlated in our dataset. Feeding groups appear less often in B-IBIs than other metric types (Klemm and others 2003), but they were represented by two metrics in this B-IBI: predator richness and percent collectors (gatherers and filterers combined). Scraper richness was also responsive, but was rejected here because it was highly correlated with EPT richness.

The SoCal IBI should prove useful as a foundation for state and regional ambient water quality monitoring programs. Because the 75 EMAP sites were selected using a probabilistic statistical design, it will also be possible to use those samples to estimate the percentage of stream miles that are in "good", "fair", and "poor" condition in the southern California coastal region. These condition estimates, combined with stressor association techniques, have great potential to serve as a scientifically defensible basis for allocating precious monitoring resources in this region.

Acknowledgments

Major funding for this study was provided by David Gibson of the San Diego WQCB and Joseph Furnish of the USFS, with supplemental funding from other regional WQCBs in southern California. We thank James Harrington and other staff of California Department of Fish and Game's Aquatic Bioassessment Laboratory for efforts that contributed to the success of this project. Biologists Mike Dawson, Shawn McBride, and Jennifer York collected much of the field data; taxonomists Dan Pickard, Doug Post, Brady Richards, and Joe Slusark performed most of the taxonomy, except for the USFS samples, which were processed by the Bug Lab at Utah State University. Dan Heggem (US EPA, Landscape Ecology Branch), Mark Rosenberg (California Department of Forestry), and Mark Angelo (Central Coast WQCB) helped identify appropriate GIS land-use layers, and Glenn Sibbald (CDFG) assisted with watershed delineation. Mary Adams (Central Coast WQCB) collected invertebrate samples and provided local site condition data for sites in the Central Coast region. This manuscript was greatly improved by comments from Larry Brown, Joseph Furnish, Robert Hughes, Leska Fore, and three anonymous reviewers.

Literature Cited

Barbour, M. T., J. B. Stribling, and J. R Karr. 1995. The multimetric approach for establishing biocriteria and measuring biological conditions. Pages 63–77 *in* W. S. Davis, T. Simon (eds.), Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Barbour, M.T., J. Gerritson, G. E. Griffith, R. Frydenborg, E. McCarron, J. S. White, and M. L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15:185–211.

Barbour, M.T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Revision to rapid bioassessment protocols for use in stream and rivers: periphyton, BMIs and fish. EPA 841-D-97-002. US Environmental Protection Agency, Washington, DC.

Beauchard, O., J. Gagneur, and S. Brosse. 2003. Macroinvertebrate richness patterns in North African streams. *Journal of Biogeography* 30:1821–1833.

Brown, H. P. 1973. Survival records for elmid beetles, with notes on laboratory rearing of various dryopoids (Coleoptera). *Entomological News* 84:278–284.

Clements, W. H. 1994. Benthic invertebrate community response heavy metals in the Upper Arkansas River Basin, Colorado. *Journal of the North American Benthological Society* 13:30–44.

Cooper, S .D., T. L. Dudley, and N. Hemphill. 1986. The biology of chaparral streams in southern California. Pages 139–151 *in* Proceedings of the Chaparral Ecosystems Research Conference, J. J. DeVries (ed.). California Water

Resources Center Report 62. University of California, Davis, California, 155 pp.

DeShon, J. E. 1995. Development and application of the invertebrate community index (ICI). Pages 217–243 *in* W. S. Davis, T. P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making. CRC Press, Boca Raton, Florida.

Doberstein, C.P., J. R. Karr, and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology* 44:355–371.

Ebert, D. W., and T. G. Wade. 2002. Analytical tools interface for landscape assessments (ATtILA), Version 3.0. US EPA, Office of Research and Development, Washington, DC.

ESRI. 1999. ArcView GIS, Version 3.2. Spatial Analyst Extension. Environmental Systems Research Institute, Inc., Redlands, CA.

Fore, L.S., J. R. Karr, and R. W. Wisseman. 1996. Assessing invertebrate responses to human activities: Evaluating alternative approaches. *Journal of the North American Benthological Society* 15:212–231.

Fore, L.S., K. Paulsen, and K. O'Laughlin. 2001. Assessing the performance of volunteers in monitoring streams. *Freshwater Biology* 46:109–123.

Gasith, A., and V. H. Resh. 1999. Streams in Mediterranean climate regions: Abiotic influences and biotic responses to predictable seasonal events. *Annual Review of Ecology and Systematics 30* : 51–81.

Harrington, J. M. 1999. California stream bioassessment procedures. California Department of Fish and Game, Water Pollution Control Laboratory. Rancho Cordova, California.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. M. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10:1456–1477.

Hawkins, C. P., J. Ostermiller, and M. Vinson. 2001. Stream invertebrate, periphyton and environmental sampling associated with biological water quality assessments. Unpublished manuscript, Utah State University, Logan, Utah.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31–47 *in* W. S. Davis, T .P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making. CRC Press, Boca Raton, Florida.

Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. *Canadian Journal of Fisheries and Aquatic Sciences* 55:1618–1631.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities.. *Fisheries* 6:21–27.

Karr, J. R., K. D. Fausch, L. Angermeyer, P. R. Yant, and I. J. Schlosser. 1986. Assessment of biological integrity in running waters: A method and its rationale. Illinois Natural History Survey Special Publication No. 5, Illinois Natural History Survey, Urbana Champaign, IL.

Karr, J. R., and E. W. Chu. 1999. Restoring life in running waters: Better biological monitoring. Island Press, Covelo, California.

Kaufmann, P. R., P. Levine, E. G. Robison, C. Seeliger, and D. V. Peck. 1999. Surface waters: Quantifying physical habitat in wadeable streams. EPA/620/R-99/003. US EPA. Office of Research and Development, Washington, DC.

Kerans, B. L., and J. R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications* 4:768–785.

Klemm, D. J., P. A. Lewis, F. A. Fulk, and J. M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA/600/4-90/030. US Environmental Protection Agency, Cincinati, Ohio.

Klemm, D. J., and J. M. Lazorchak. 1994. Environmental monitoring and assessment program, surface water and Region 3 regional monitoring and assessment program, 1994 pilot laboratory methods manual for streams. EPA/62/R-94/003. US EPA, Office of Research and Development, Washington, DC.

Klemm, D. J., K. A. Blocksom, F. A. Fulk, A. T. Herlihy, R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, W. T. Thoeny, M. B. Griffith, and W. S. Davis. 2003. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic Highland streams. *Environmental Management* 31:656–669.

McCormick, F. H., R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, and A. T. Herlihy. 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands Region. *Transactions of the American Fisheries Society* 130:857–877.

McCune, B., and J. B. Grace. 2002. Analysis of ecological communities. MjM Software Design. Gleneden Beach, Oregon.

Newman, W. B., and F. R. G. Watson. 2003. Land use history and mapping in California's Central Coast region. The Watershed Institute, California State University, Monterey Bay, California.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). *Annals of the Association of American Geographers* 77(1):118–125.

Puckett, M. 2002. Quality Assurance Management Plan for the State of California's Surface Water Ambient Monitoring Program (SWAMP). Prepared for California State Water Resources Control Board, Division of Water Quality, Sacramento, CA. First version. December 2002. Available at http://www.swrcb.ca.gov/swamp/qapp.html.

Reynoldson, T. B., D. M. Rosenberg, and V. H. Resh. 2001. Comparison of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchment, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* 58:1395–1410.

Vinson, M. R., and C. P. Hawkins. 1996. Effects of sampling area and subsampling procedures on comparisons of taxonomic richness among streams. *Journal of the North American Benthological Society* 15:392–399.

Wright, J. F., M. T. Furse, and P. D. Armitage. 1983. RIVPACS: A technique for evaluating the biological quality of rivers in the U.K. *European Water Pollution Control* 3:15–25.

Zar, J. H. 1999. Biostatistical analysis, 4th ed. Prentice-Hall, Upper Saddle River, New Jersey.

Appendix 3

Bioassessment in complex environments: designing an index for consistent meaning in different settings

Raphael D. Mazor[1,2,5], Andrew C. Rehn[2,6], Peter R. Ode[2,7], Mark Engeln[1,8], Kenneth C. Schiff[1,9], Eric D. Stein[1,10], David J. Gillett[1,11], David B. Herbst[3,12], and Charles P. Hawkins[4,13]

[1]Southern California Coastal Water Research Project, 3535 Harbor Boulevard, Suite 110, Costa Mesa, California 92626 USA
[2]Aquatic Bioassessment Laboratory, California Department of Fish and Wildlife, 2005 Nimbus Road, Rancho Cordova, California 95670 USA
[3]Sierra Nevada Aquatic Research Laboratory, University of California, 1016 Mt. Morrison Road, Mammoth Lakes, California 93546 USA
[4]Department of Watershed Sciences, Western Center for Monitoring and Assessment of Freshwater Ecosystems, and the Ecology Center, Utah State University, Logan, Utah 84322-5210 USA

Abstract: Regions with great natural environmental complexity present a challenge for attaining 2 key properties of an ideal bioassessment index: 1) index scores anchored to a benchmark of biological expectation that is appropriate for the range of natural environmental conditions at each assessment site, and 2) deviation from the reference benchmark measured equivalently in all settings so that a given index score has the same ecological meaning across the entire region of interest. These properties are particularly important for regulatory applications like biological criteria where errors or inconsistency in estimating site-specific reference condition or deviation from it can lead to management actions with significant financial and resource-protection consequences. We developed an index based on benthic macroinvertebrates for California, USA, a region with great environmental heterogeneity. We evaluated index performance (accuracy, precision, responsiveness, and sensitivity) throughout the region to determine if scores provide equivalent ecological meaning in different settings. Consistent performance across environmental settings was improved by 3 key elements of our approach: 1) use of a large reference data set that represents virtually all of the range of natural gradients in the region, 2) development of predictive models that account for the effects of natural gradients on biological assemblages, and 3) combination of 2 indices of biological condition (a ratio of observed-to-expected taxa [O/E] and a predictive multimetric index [pMMI]) into a single index (the California Stream Condition Index [CSCI]). Evaluation of index performance across broad environmental gradients provides essential information when assessing the suitability of the index for regulatory applications in diverse regions.

Key words: bioassessment, predictive modelling, predictive multimetric index, reference condition

A major challenge for conducting bioassessment in environmentally diverse regions is ensuring that an index provides consistent meaning in different environmental settings. A given score from a robust index should indicate the same biological condition, regardless of location or stream type. However, the performance (e.g., accuracy, precision, responsiveness, and sensitivity) of an index may vary in different settings, complicating its interpretation (Hughes et al. 1986, Yuan et al. 2008, Pont et al. 2009). Effective bioassessment indices should account for naturally occurring variation in aquatic assemblages so that deviations from reference conditions resulting from anthropogenic disturbance are minimally confounded by natural variability (Hughes et al. 1986, Reynoldson et al. 1997). When bioassessment indices are used in regulatory applications, such as measuring compliance with biocriteria (Davis and Simon 1995, Council of European Communities 2000, USEPA 2002, Yoder and Barbour 2009), variable meaning of an index score may lead to poor stream management, particularly if the environmental factors affecting index performance are unrecognized. Those who develop bioassessment indices or the policies that rely on them should evaluate index performance carefully across the different environmental gradients where an index will be applied.

A reference data set that represents the full range of environmental gradients where an index will be used is key for index development in environmentally diverse regions. In addition, reference criteria should be consistently defined so that benchmarks of biological condition are equivalent across environmental settings. Indices based on

E-mail addresses: [5]raphaelm@sccwrp.org; [6]andy.rehn@wildlife.ca.gov; [7]peter.ode@wildlife.ca.gov; [8]mengeln@gmail.com; [9]kens@sccwrp.org; [10]erics@sccwrp.org; [11]davidg@sccwrp.org; [12]david.herbst@lifesci.ucsb.edu; [13]chuck.hawkins@usu.edu

DOI: 10.1086/684130. Received 21 September 2014; Accepted 28 April 2015; Published online 22 October 2015.

benthic macroinvertebrates (BMI) for use in California were developed with reference data sets that used different criteria in different regions (e.g., Hawkins et al. 2000, Herbst and Silldorff 2009, Rehn 2009). For example, several reference sites used to calibrate an index for the highly urbanized South Coast region had more nonnatural land use than any reference site used to develop an index for the rural North Coast region (Ode et al. 2005, Rehn et al. 2005). Furthermore, lower-elevation settings were poorly represented in these reference data sets. In preparation for establishing statewide biocriteria, regulatory agencies and regulated parties desired a new index based on a larger, more consistently defined reference data set that better represented all environmental settings. Considerable effort was invested to expand the statewide pool of reference sites to support development of a new index (Ode et al. 2016). The diversity of stream environments represented in the reference pool necessitated scoring tools that could handle high levels of complexity.

Predictive modeling of the reference condition is an increasingly common way to obtain site-specific expectations for diverse environmental settings (Hawkins et al. 2010b). Predictive models can be used to set biological expectations at test sites based on the relationship between biological assemblages and environmental factors at reference sites. Thus far, predictive modeling has been applied almost exclusively to multivariate indices focused on taxonomic completeness of a sample, such as measured by the ratio of observed-to-expected taxa (O/E) (Moss et al. 1987, Hawkins et al. 2000, Wright et al. 2000), or location of sites in ordination space (e.g., **BE**nthic **A**ssessment of **S**edimen**T** [BEAST]; Reynoldson et al. 1995). Applications of predictive models to multimetric indices (i.e., predictive multimetric indices [pMMIs]) are relatively new (e.g., Cao et al. 2007, Pont et al. 2009, Vander Laan and Hawkins 2014). MMIs include information on the life-history traits observed within an assemblage (e.g., trophic groups, habitat preferences, pollution tolerances), so they may provide useful information about biological condition that is not incorporated in an index based only on loss of taxa (Gerritsen 1995). Predictive models that set site-specific expectations for biological metric values may improve the accuracy, precision, and sensitivity of MMIs when applied across diverse environmental settings (e.g., Hawkins et al. 2010a).

A combination of multiple indices (specifically, a pMMI and an O/E index) into a single index might provide more consistent measures of biological condition than just one index by itself. Variation in performance of an index would be damped by averaging it with a 2nd index, and poor performance in particular settings might be improved. For example, an O/E index may be particularly sensitive in mountain streams that are expected to be taxonomically rich, whereas a pMMI might be more sensitive in lowland areas, where stressed sites may be well represented in calibration data. Moreover, pMMIs and O/E indices characterize assemblage data in fundamentally different ways. Thus, they provide complementary measures of stream ecological condition and may contribute different types of diagnostic information. Taxonomic completeness, as measured by an O/E index, and ecological structure, as measured by a pMMI, are both important aspects of stream communities, and certain stressors may affect these aspects differently. For example, replacement of native taxa with invasive species may reduce taxonomic completeness, even if the invaders have ecological attributes similar to those of the taxa they displaced (Collier 2009). Therefore, measuring both taxonomic completeness and ecological structure may provide a more complete picture of stream health.

Our goal was to construct a scoring tool for perennial wadeable streams that provides consistent interpretations of biological condition across environmental settings in California, USA. Our approach was to design the tool to maximize the consistency of performance across settings, as indicated by evaluations of accuracy, precision, responsiveness, and sensitivity. We first constructed predictive models for both a taxon loss index (O/E) and a pMMI. Second, we compared the accuracy, precision, responsiveness, and sensitivity of the O/E, pMMI, and combined O/E + pMMI index across a variety of environmental settings. Our primary motivation was to develop biological indices to support regulatory applications in the State of California. However, our broader goal was to produce a robust assessment tool that would support a wide variety of bioassessment applications, such as prioritization of restoration projects or identification of areas with high conservation value.

METHODS

Study region

California contains continental-scale environmental diversity within 424,000 km^2 that encompass some of the most extreme gradients in elevation and climate found in the USA. It has temperate rainforests in the North Coast, deserts in the east, and chaparral, oak woodlands, and grasslands with a Mediterranean climate in coastal regions (Omernik 1987). Large areas of the state are publicly owned, but vast regions have been converted to agricultural (e.g., the Central Valley) or urban (e.g., the South Coast and the San Francisco Bay Area) land uses (Sleeter et al. 2011). Forestry, grazing, mining, other resource extraction activities, and intensive recreation occur throughout rural regions of the state, and the fringes of urban areas are undergoing increasing development. For convenience, we divided the state into 6 regions and 10 subregions based on ecoregional (Omernik 1987) and hydrologic boundaries (California State Water Resources Control Board 2013) (Fig. 1).

Compilation of data

We compiled data from >20 federal, state, and regional monitoring programs. Altogether, we aggregated data from

Figure 1. Regions and subregions of California. Thick gray lines indicate regional boundaries, and thin white lines indicate subregional boundaries. NC = North Coast, CHco = Coastal Chaparral, Chin = Interior Chaparral, SCm = South Coast mountains, SCx = South Coast xeric, CV = Central Valley, SNws = Sierra Nevada-western slope, SNcl = Sierra Nevada-central Lahontan, DMmo: Desert/Modoc-Modoc plateau, DMde =Desert/Modoc-deserts.

4457 samples collected from 2352 unique sites between 1999 and 2010 into a single database. We excluded BMI samples with insufficient numbers of organisms or taxonomic resolution (described below) from analyses. We treated observations at sites in close proximity to each other (within 300 m) as repeat samples from a single site. For sites with multiple samples meeting minimum requirements, we randomly selected a single sample for use in all analyses described below, and we withheld repeat samples from all analyses, except where indicated below. We used 1318 sites sampled during probabilistic surveys (e.g., Peck et al. 2006) to estimate the ambient condition of streams (described below).

Biological data

Fifty-five percent of the BMI samples were collected following a reach-wide protocol (Peck et al. 2006), and the other samples were collected with targeted riffle protocols, which produce comparable data (Gerth and Herlihy 2006, Herbst and Silldorff 2006, Rehn et al. 2007). For most samples, taxa were identified to genus, but this level of effort and the total number of organisms/sample varied among samples, necessitating standardization of BMI data. We used different data standardization approaches for the pMMI and the O/E. For the pMMI, we aggregated identifications to 'Level 1' standard taxonomic effort (most insect taxa identified to genus, Chironomidae identified to family) as defined by the Southwest Association of Freshwater Invertebrate Taxonomists (SAFIT; Richards and Rogers 2011) and used computer subsampling to generate 500-count subsamples. We excluded samples with <450 individuals (i.e., not within 10% of target). For the O/E index, we used operational taxonomic units (OTUs) similar to SAFIT Level 1 except that we aggregated Chironomidae to subfamily. We excluded ambiguous taxa (i.e., those identified to a higher level than specified by the OTU). We also excluded samples with >50% ambiguous individuals from O/E development, no matter how many unambiguous individuals remained. We used computer subsampling to generate 400-count subsamples, and we excluded samples with <360 individuals. A smaller subsample size was used for the O/E index than for the pMMI because exclusion of ambiguous taxa often reduced sample size to <500 individuals. A final data set of 3518 samples from 1985 sites met all requirements and was used for development and evaluation of both the O/E and pMMI indices.

Environmental data

We collected environmental data from multiple sources to characterize natural and anthropogenic factors known to affect benthic communities, such as climate, elevation, geology, land cover, road density, hydrologic alteration, and mining (Tables 1, 2). We used geographic information system (GIS) variables that characterized natural, unalterable environmental factors (e.g., topography, geology, climate) as predictors for O/E and pMMI models and variables related to human activity (e.g., land use) to classify sites as reference and to evaluate responsiveness of O/E and pMMI indices to human activity gradients. We calculated most variables related to human activity at 3 spatial scales (within the entire upstream drainage area [watershed], within the contributing area 5 km upstream of a site [5 km], and within the contributing area 1 km upstream of a site [1 km]) so that we could screen sites for local and catchment-scale impacts. We created polygons defining these spatial analysis units using ArcGIS tools (version 9.0; Environmental Systems Research Institute, Redlands, California).

Classification of sites along a human activity gradient

We were unable to measure stress directly with this data set, so instead, we used a human activity gradient under the assumption that it was correlated with stress (Yates and Bailey 2010). We divided sites into 3 sets for development and evaluation of indices: reference (i.e., low activity), moderate-, and high-activity sites. We defined reference

Table 1. Natural gradients and their importance (Gini = mean decrease in Gini index), MSE = % increase in mean squared error) for random-forest models for the observed (O)/expected (E) taxa index and each metric used in the predictive multimetric index (pMMI). Predictors that were evaluated but not selected for any model include % sedimentary geology, nitrogenous geology, soil hydraulic conductivity, soil permeability, S-bearing geology, calcite-bearing geology, and magnesium oxide-bearing geology. Sources: A = National Elevation Dataset (http://ned.usgs.gov/), B = PRISM climate mapping system (http://www.prism.oregonstate.edu), C = generalized geology, mineralogy, and climate data derived for a conductivity prediction model (Olson and Hawkins 2012). Dashes indicate that the predictors were not used to model the metric.

Variable	Description	O/E		Taxonomic richness MSE	% intolerant MSE	# Shredder taxa MSE	Clinger % taxa MSE	Coleoptera % taxa MSE	EPT % taxa MSE	Data source
		Gini	MSE							
Location										
New lat	Latitude	90.5	0.09	18.8	0.0063	1.26	0.0054	0.00079	0.0027	
New long	Longitude	–	–	25.3	0.0058	0.99	0.0030	–	0.0024	
SITE_ELEV	Elevation	89.5	0.11	11.8	–	–	–	0.00231	–	A
Catchment morphology										
LogWSA	Log watershed area	86.6	0.06	–	0.0020	1.23	–	–	–	A
ELEV_RANGE	Elevation range	–	–	2.4	–	–	0.0026	–	–	A
Climate										
PPT	10-y (2000–2009) average precipitation at the sampling point	74.8	0.07	8.4	0.0063	0.92	–	–	0.0016	B
TEMP	10-y (2000–2009) average air temperature at the sampling point	81.9	0.09	9.3	0.0052	–	0.0023	–	0.0019	B
SumAve_P	Mean June to September 1971–2000 monthly precipitation, averaged across the catchment	–	–	5.5	–	–	–	–	0.0033	B
Geology										
BDH_AVE	Average bulk soil density	–	–	5.7	–	–	0.0021	–	–	C
KFCT_AVE	Average soil erodibility factor (k)	–	–	6.2	–	–	0.0027	–	0.0025	C
Log_P_MEAN	Log % P geology	–	–	3.7	–	–	–	–	–	C

Table 2. Stressor and human-activity gradients used to identify reference sites and evaluate index performance. Sites that did not exceed the listed thresholds were used as reference sites. Sources A = National Landcover Data Set (http://www.epa.gov/mrlc/nlcd-2006.html), B = custom roads layer, C = National Hydrography Dataset Plus (http://www.horizon-systems.com/nhdplus), D = National Inventory of Dams (http://geo.usace.army.mil), E = Mineral Resource Data System (http://tin.er.usgs.gov/mrds), F = predicted specific conductance (Olson and Hawkins 2012), G = field-measured variables. WS = watershed, 5 km = watershed clipped to a 5-km buffer of the sampling point, 1 km = watershed clipped to a 1-km buffer of the sampling point, W1_HALL = proximity-weighted human activity index (Kaufmann et al. 1999), Code 21 = landuse category that corresponds to managed vegetation, such as roadsides, lawns, cemeteries, and golf courses. * indicates variable used in the random-forest evaluation of index responsiveness.

	Variable	Scale	Threshold	Unit	Data source
*	% agricultural	1 km, 5 km, WS	<3	%	A
*	% urban	1 km, 5 km, WS	<3	%	A
*	% agricultural + % urban	1 km, 5 km, WS	<5	%	A
*	% Code 21	1 km and 5 km	<7	%	A
*		WS	<10	%	A
*	Road density	1 km, 5 km, WS	<2	km/km^2	B
*	Road crossings	1 km	<5	crossings	B, C
*		5 km	<10	crossings	B, C
*		WS	<50	crossings	B, C
*	Dam distance	WS	<10	km	D
*	% canals and pipelines	WS	<10	%	C
*	Instream gravel mines	5 km	<0.1	mines/km	C, E
*	Producer mines	5 km	0	mines	E
	Specific conductance	Site	99/1[a]	prediction interval	F
	W1_HALL	Reach	<1.5	NA	G
	% sands and fines	Reach		%	G
	Slope	Reach		%	G

[a] The 99th and 1st percentiles of predictions were used to generate site-specific thresholds for specific conductance. The model underpredicted at higher levels of specific conductance (data not shown), so a threshold of 2000 μS/cm was used as an upper bound if the prediction interval included 1000 μS/cm.

sites as 'minimally disturbed' sensu Stoddard et al. (2006) and selected them by applying screening criteria based primarily on landuse variables calculated at multiple spatial scales (i.e., 1 km, 5 km, watershed; Table 2). We calculated some screening criteria at only 1 spatial scale (e.g., in-stream gravel mine density at the 5-km scale and W1_HALL, a proximity-weighted index of human activity based on field observations made within 50 m of a sampling reach; Kaufmann et al. 1999). We excluded sites thought to be affected by grazing or recreation from the reference data set, even if they passed all reference criteria. Identification of high-activity sites was necessary for pMMI calibration (described below) and for performance evaluation of both pMMI and O/E. We defined high-activity sites as meeting any of the following criteria: ≥50% developed land (i.e., % agricultural + % urban) at all spatial scales, ≥5 km/km^2 road density, or W1_HALL ≥ 5. We defined sites not identified as either reference or high-activity as moderate-activity sites. We further divided sites in each set into calibration (80%) and validation (20%) subsets and stratified assignment to calibration and validation sets by subregion to ensure representation of all environmental settings in both sets (Fig. 1). Only 1 reference site was found in the Central Valley, so that region was combined with the Interior Chaparral (whose boundary was within 500 m of the site) for stratification purposes.

Development of the O/E index

Development of an O/E index or pMMI follows the same basic steps: biological characterization, modeling of reference expectations from environmental factors, selection of metrics or taxa, and combining of metrics or taxa into an index. pMMI development has an additional intermediate step to set biological expectations for sites with high levels of activity (Fig. 2). Taxonomic completeness, as measured by O/E, quantifies degraded biological condition as loss of expected native taxa (Hawkins 2006). E represents the number of taxa expected in a specific sample, based on its environmental setting, and O represents the number of those expected taxa that were actually observed. We developed models to calculate the O/E index following the general approach of Moss et al. (1987). First, we defined groups of reference calibration sites based on their

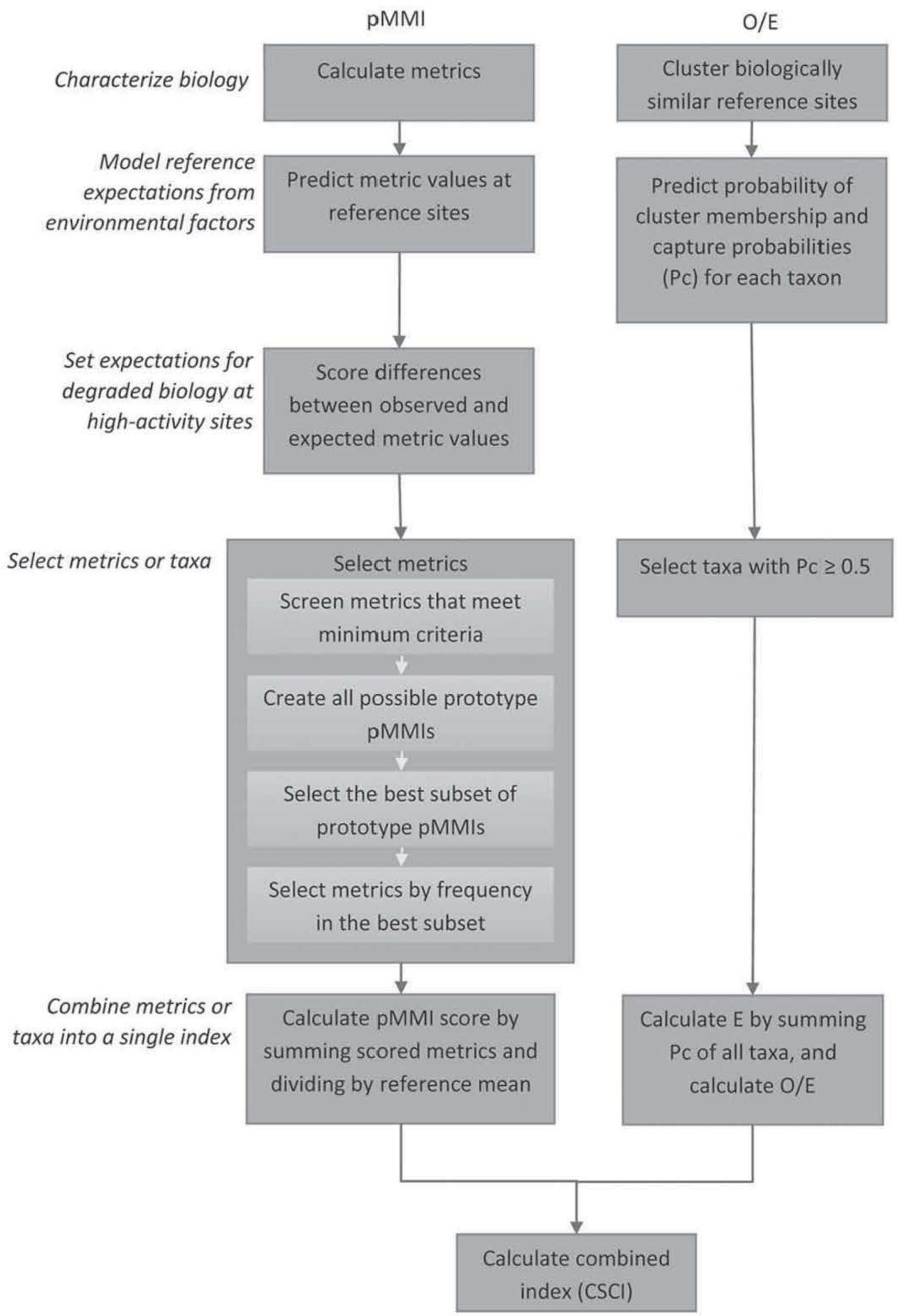


Figure 2. Summary of steps in developing the predictive multimetric index (pMMI) and observed (O)/expected (E) taxa index. Pc = probability of observing a taxon at a site, CSCI = California State Condition Index.

taxonomic similarity. Second, we developed a random-forest model (Cutler et al. 2007) to predict group membership based on naturally occurring environmental factors minimally affected by human activities. We used this model to predict cluster membership for test sites based on their natural environmental setting. The probability of observing a taxon at a test site (i.e., the capture probability) was calculated as the cluster-membership-probability-weighted frequencies of occurrence summed across clusters:

$$Pc_j = \sum_{i=1}^{k}(G_iF_i), \qquad \text{(Eq. 1)}$$

where Pc_j is the probability of observing taxon j at a site, G_i is the probability that a site is a member of group i, F_i is the relative frequency of the taxon in group i, and k is the number of groups used in modeling. The sum of the capture probabilities is the expected number of taxa (E) in a sample from a site:

$$E = \sum_{j=1}^{m}Pc_j, \qquad \text{(Eq. 2)}$$

where m is the number of taxa observed across all reference sites. We used Pc values ≥ 0.5 when calculating O/E because excluding locally rare taxa generally improves precision of O/E indices (Hawkins et al. 2000, Van Sickle et al. 2007). This model was used to predict E at reference and nonreference sites based on their natural environmental setting.

We used presence/absence-transformed BMI data from reference calibration sites to identify biologically similar groups of sites. We excluded taxa occurring in <5% of reference calibration samples from the cluster analysis because inclusion of regionally rare taxa can obscure patterns associated with more common taxa (e.g., Gauch 1982, Clarke and Green 1988, Ostermiller and Hawkins 2004). We created a dendrogram with Sørensen's distance measure and flexible β (β = −0.25) unweighted pair group method with arithmetic mean (UPGMA) as the linkage algorithm in R (version 2.15.2; R Project for Statistical Computing, Vienna, Austria) with the *cluster* package (Maechler et al. 2012) and scripts written by J. Van Sickle (US Environmental Protection Agency, personal communication). We identified groups containing ≥10 sites and subtended by relatively long branches (to maximize differences in taxonomic composition among clusters) by visual inspection of the dendrogram. We retained rare taxa that were excluded from the cluster analysis for other steps in index development.

We constructed a 10,000-tree random-forest model with the *randomForest* package in R (Liaw and Wiener 2002) to predict cluster membership for new test sites. We excluded predictors that were moderately to strongly correlated with one another (|Pearson's r| ≥ 0.7). When we observed correlation among predictors, we selected the predictor that was simplest to calculate (e.g., calculated from point data rather than delineated catchments) as a candidate predictor. We used an initial random-forest model based on all possible candidate predictors to identify those predictors that were most important for predicting new test sites into biological groups as measured by the Gini index (Liaw and Wiener 2002). We evaluated different combinations of the most important variables to identify a final, parsimonious model that minimized the standard deviation (SD) of reference site O/E scores at calibration reference sites with the fewest predictors.

We evaluated O/E index performance in 2 ways. First, we compared index precision with the lowest and highest precision possible given the sampling and sample-processing methods used (Van Sickle et al. 2005). SD of O/E index scores produced by a null model (i.e., all sites are in a single group, and capture probabilities for each taxon are the same for all sites) estimates the lowest precision possible for an O/E index. SD of O/E values based on estimates of variability among replicate samples (SDRS) estimates the highest attainable precision possible for the index. Second, we evaluated the index for consistency by regressing O against E for reference sites. Slopes close to 1 and intercepts close to 0 indicate better performance.

Development of the pMMI

We followed the approach of Vander Laan and Hawkins (2014) to develop a pMMI. In contrast to traditional MMIs, which typically attempt to control for the effects of natural factors on biological metrics via landscape classifications or stream typologies, a pMMI accounts for these effects by predicting the expected (i.e., naturally occurring) metric values at reference sites given their specific environmental setting. A pMMI uses the difference between the observed and predicted metric values when scoring biological condition, whereas a traditional MMI uses the raw metric for scoring. Traditional approaches to MMI development may reduce the effects of natural gradients on metric values through classification (e.g., regionalization or typological approaches; see Ode et al. 2005 for a California example), but they seldom produce site-specific expectations for different environmental settings (Hawkins et al. 2010b).

We developed the pMMI in 5 steps (Fig. 2): 1) metric calculation, 2) prediction of metric values at reference sites, 3) metric scoring, 4) metric selection, and 5) assembly of the pMMI. Apart from step 2, the process for developing a pMMI is comparable to that used for a traditional MMI (e.g., Stoddard et al. 2008). We developed a null MMI based on raw values of the selected metrics to allow us to estimate how much predictive modeling improved pMMI performance. The process was intended to produce a pMMI that was unbiased, precise, responsive, and able to characterize a large breadth of ecological attributes of the BMI assemblage.

Metric calculation We calculated biological metrics that characterized the ecological structure of BMI assemblages for each sample in the data set. We used custom scripts in R and the *vegan* package (Oksanen et al. 2013) to calculate a suite of 48 widely used bioassessment metrics, chosen because they quantify important ecological attributes, such as taxonomic richness or trophic diversity (a subset of which is presented in Table 3). Many of these metrics are widely used in other bioassessment indices (e.g., Royer et al. 2001, Stribling et al. 2008). Different formulations of metrics based on taxonomic composition (e.g., Diptera metrics) or traits (e.g., predator metrics) were assigned to thematic metric groups representing different ecological attributes (Table 3). These thematic groups were used to help ensure that the metrics included in the pMMI were ecologically diverse.

Prediction of metric values at reference sites We used random-forest models to predict values for all 48 metrics at reference calibration sites based on the same GIS-derived candidate variables that were used for O/E devel-

Table 3. Metrics evaluated for inclusion in the predictive multimetric index (pMMI). Only metrics that met all evaluation criteria are shown. EPT = Ephemeroptera, Plecoptera, and Trichoptera; Resp = direction of response; I = metric increases with human-activity gradients; D = metric decreases with human-activity gradients; Var Exp = % variance explained by the random-forest model; r^2 (cal) = squared Pearson correlation coefficient between predicted and observed values at reference calibration sites; r^2 (val) = squared Pearson correlation coefficient between predicted and observed values at reference validation sites; t (null) = t-statistic for the comparison of the raw metric between the reference and high-activity samples within the calibration data set; t (mod) = t-statistic for the comparison of the residual metric between the reference and high-activity samples within the calibration data set; F = F-statistic for an analysis of variance of metric residual values from reference calibration sites among regions shown in Fig. 1; S : N = signal-to-noise ratio; Freq = frequency of the metric among the best-performing combinations of metrics. Tolerance, functional feeding group, and habit data were from CAMLnet (2003). * indicates metric selected for inclusion in the pMMI.

Metric	Resp	Var Exp	r^2 (cal)	r^2 (val)	t (null)	t (mod)	F	S : N	Freq
Taxonomic diversity									
*Taxonomic richness	D	0.27	0.27	0.15	21.6	23.7	1.0	6.7	0.83
Functional feeding group									
Scrapers									
No. Scraper taxa	D	0.40	0.40	0.29	15.3	19.1	1.2	7.6	0.17
Shredders									
% Shredder taxa	D	0.27	0.27	0.46	17.6	10.6	1.0	4.1	0.33
* No. Shredder taxa	D	0.39	0.39	0.35	19.2	15.2	1.9	5.4	0.50
Habit									
Clingers									
* % Clinger taxa	D	0.34	0.34	0.42	21.7	14.6	0.2	4.8	1.00
No. Clinger taxa	D	0.39	0.40	0.32	26.0	25.3	0.5	11.1	0
Taxonomy									
Coleoptera									
* % Coleoptera taxa	D	0.30	0.31	0.22	10.3	15.8	1.0	5.0	0.83
No. Coleoptera taxa	D	0.34	0.34	0.29	13.6	20.9	0.6	6.2	0.17
EPT									
* % EPT taxa	D	0.31	0.32	0.46	30.0	23.1	0.4	6.0	0.67
No. EPT taxa	D	0.40	0.40	0.31	27.8	25.3	1.4	10.0	0.17
Tolerance									
* % Intolerant taxa	D	0.23	0.23	0.15	21.7	15.6	0.5	5.1	0.67
% Intolerant taxa	D	0.51	0.51	0.58	32.7	25.3	1.5	6.9	0.17
No. Intolerant taxa	D	0.52	0.52	0.53	28.4	21.8	1.5	9.6	0
Tolerance value	I	0.22	0.25	0.20	–21.5	–17.0	0.4	5.0	0
% Tolerant taxa	I	0.22	0.24	0.38	–26.1	–22.3	1.4	4.9	0.17

opment (Table 1). Manual refinement was impractical because of the large number of models that were developed, so we used an automated approach (recursive feature elimination [RFE]) to select the simplest model (the model with the fewest predictors) whose root mean square error (RMSE) was ≤2% greater than the RMSE of the optimal model (the model with the lowest RMSE). We considered only models with ≤10 predictors. Limiting the complexity of the model typically reduces overfitting and improves model validation (Strobl et al. 2007). We implemented RFE with the *caret* package in R using the default settings for random-forest models (Kuhn et al. 2012). We used the *randomForest* package (Liaw and Wiener 2002) to create a final 500-tree model for each metric based on the predictors used in the model selected by RFE. We then used these models to predict metric values for all sites. We used out-of-bag predictions for the reference calibration set (an out-of-bag prediction is based only on the subset of trees in which a calibration site was excluded during model training). To evaluate how well each model predicted metric values, we regressed raw observed values against predicted values for reference sites. Slopes close to 1 and intercepts close to 0 indicate better model performance. If the pseudo-R^2 of the model (calculated as 1 – mean squared error [MSE]/variance) was >0.2, we used the model to adjust metric values (i.e., observed – predicted), otherwise we used the observed metric values. Hereafter, 'metric' is used to refer to both raw and adjusted metric values.

Metric scoring Scoring is required for MMIs because metrics have different scales and different responses to stress (Blocksom 2003). Scoring transforms metrics to a standard scale ranging from 0 (i.e., most stressed) to 1 (i.e., identical to reference sites). We scored metrics following Cao et al. (2007). We scored metrics that decrease with human activity as

$$(\text{Observed}-\text{Min})/(\text{Max}-\text{Min}), \quad \text{(Eq. 3)}$$

where Min is the 5$^{\text{th}}$ percentile of high-activity calibration sites and Max is the 95$^{\text{th}}$ percentile of reference calibration sites. We scored metrics that increase with human activity as

$$(\text{Observed}-\text{Max})/(\text{Min}-\text{Max}), \quad \text{(Eq. 4)}$$

where Min is the 5$^{\text{th}}$ percentile of reference calibration sites, and Max is the 95$^{\text{th}}$ percentile of high-activity sites. We trimmed scores outside the range of 0 to 1 to 0 or 1. We used 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles instead of minimum or maximum values because they are more robust estimates of metric range than minima and maxima (Blocksom 2003, Stoddard et al. 2008).

Metric selection We selected metrics in a 2-phase process: 1) based on their individual performance, and 2) based on their frequency in high-performing prototype pMMIs. Evaluating the performance of many prototype pMMIs avoids selection of metrics with spuriously good performance and is preferable to selecting metrics or pMMIs based on performance evaluations conducted 1 metric at a time (Hughes et al. 1998, Roth et al. 1998, Angradi et al. 2009, Van Sickle 2010). Initial elimination of metrics based on their individual performance alleviates the computational challenge of evaluating large numbers of prototype pMMIs.

We used several performance criteria to eliminate metrics from further analysis. We assessed responsiveness to human activity by computing t-statistics based on comparisons of mean metric values at reference sites and sites with high levels of activity and eliminated metrics with a t-statistic < 10. We assessed bias by determining whether metric values varied among predefined geographic regions (Fig. 1). We considered metrics with an F-statistic > 2 derived from analysis of variance (ANOVA) by geographic region to have high regional bias and eliminated them. Other screening criteria were modified from Stoddard et al. (2008). We excluded metrics with >⅔ zero values across samples and richness metrics with range < 5. We also eliminated metrics with a signal-to-noise ratio (ratio of between-site to within-site variance estimated from data collected at sites with multiple samples) < 3.

We further screened metrics by evaluating the performance of all possible combinations as prototype pMMIs and selecting metrics that were frequent among prototypes with the best performance. First, we assembled all nonredundant combinations of metrics that met minimum performance criteria into prototype pMMIs. Limiting the redundancy of metrics increases the number of thematic groups included in prototypes, thereby improving the ecological breadth of the pMMI. Redundant combinations of metrics included those with multiple metrics from a single metric group (e.g., tolerance metrics; Table 3) or correlated metrics (|Pearson's $r \geq |0.7|$). Prototype pMMIs ranged in size from a minimum of 5 to a maximum of 10 metrics, a range that is typical of MMIs used for stream bioassessment (e.g., Royer et al. 2001, Fore and Grafe 2002, Ode et al. 2005, Stoddard et al. 2008, Van Sickle 2010). We calculated scores for these prototype pMMIs by averaging metric scores and rescaling by the mean of reference calibration sites, which allows comparisons among prototype pMMIs.

Subsequently, we ranked prototype pMMIs to identify those with the best responsiveness and precision. Biased metrics already had been eliminated from consideration, and none of the prototypes exhibited geographic bias (results not shown), so we did not use accuracy to rank prototype pMMIs. We estimated responsiveness as the t-statistic based on mean scores at reference and high-activity calibration sites and precision as the SD of scores from reference calibration sites. We identified the best subset of prototype pMMIs as those appearing in the top quartile for both criteria. Therefore, prototype pMMIs in the best subset possessed several desirable characteristics: ecological breadth, high responsiveness, and high precision.

We assembled the final pMMI by selecting metrics in order of their frequency in the best subset of prototype pMMIs. We added metrics in order of decreasing frequency and avoided adding metrics from the same thematic group or correlated (Pearson's $r \geq 0.7$) metrics. We excluded metrics that appeared in <⅓ of the best prototype pMMIs from the final pMMI.

Aggregation of the pMMI We calculated scores for the final pMMI by averaging metric scores and rescaling by the mean of reference calibration sites (as for prototype pMMIs). Rescaling of pMMI scores ensures that pMMI and O/E are expressed in similar scales (i.e., as a ratio of observed to reference expectations) and improves comparability of the 2 indices.

We calculated scores for a combined index (the California Stream Condition Index [CSCI]) by averaging pMMI and O/E scores. We calculated a null combined index by averaging null MMI and null O/E scores.

Performance evaluation Evaluation of index performance focused on accuracy, precision, responsiveness, and sensitivity (Table 4). We compared the performance of each index to that of its null counterpart. Many of our approaches to measuring performance also have been used widely in index development (e.g., Hawkins et al. 2000, 2010a, Clarke et al. 2003, Ode et al. 2008, Cao and Hawkins 2011). We scored all indices on similar scales (i.e., a minimum of 0, with a reference expectation of 1), so no adjustments were required to make comparisons (Herbst and Silldorff 2006, Cao and Hawkins 2011). We conducted all performance evaluations separately on calibration and validation data sets.

We regarded indices as accurate if scores at reference sites were not influenced by environmental setting or time of sampling. Precise indices were those with low variability among reference sites and among samples from repeated visits within sites. Responsive indices were those that showed large decreases in response to human activity. Sensitive indices were those that frequently found nonreference sites to be below an impairment threshold (e.g., 10^{th} percentile of scores at reference sites).

Performance of the indices along a gradient of expected numbers of common taxa (E) The performance of an ideal index should not vary with E. For example, index accuracy should not be influenced by the expected richness of a site. We evaluated the accuracy, precision, and sensitivity of the indices against E by grouping sites into bins that ranged in the number of expected taxa (bin size = 4 taxa). We chose this bin size because it was the smallest number that allowed analysis of a wide range of values of E with large numbers of sites in each bin (i.e., ≥37 sites for accuracy and precision estimates and 15 sites for sensitivity estimates). We measured accuracy as the proportion of reference sites in each bin with scores $\geq 10^{th}$ percentile of reference calibration sites. We measured precision as the SD of reference sites in each bin and sensitivity as the

Table 4. Summary of performance evaluations. SD = standard deviation.

Aspect	Description	Indication of good performance
Accuracy and bias	Scores are minimally influenced by natural gradients	• Approximately 90% of validation reference sites have scores $>10^{th}$ percentile of calibration reference sites • Landscape-scale natural gradients explain little variability in scores at reference sites, as indicated by a low pseudo-R^2 for a 500-tree random-forest model • No visual relationship evident in plots of scores at reference sites against field measurements of natural gradients
Precision	Scores are similar when measured under similar settings	• Low SD of scores among reference sites (1 sample/site) • Low pooled SD of scores among samples at reference sites with multiple sampling events
Responsiveness	Scores change in response to human activity gradients	• Large t-statistic in comparison of mean scores at reference and high-activity sites • Landscape-scale human activity gradients explain variability in scores, as indicated by a high pseudo-R^2 for a 500-tree random-forest model
Sensitivity	Scores indicate poor condition at high-activity sites	• High percentage of high-activity sites have scores $<10^{th}$ percentile of calibration reference sites

proportion of high-activity sites within each bin with scores <10th percentile of reference calibration sites. We repeated all analyses with scores from indices based on null models.

Unlike accuracy and precision, the sensitivity of an ideal index (if measured as described above) may vary with E, but only to the extent that stress levels vary with E. However, how stress levels truly varied with E is unknown because human activity gradients were used to approximate stressor gradients, and direct, quantitative measures of stress levels are not possible. Even direct measures of water chemistry or habitat-related variables are at best incomplete estimates of the stress experienced by stream communities, and these data were not available for many sites in our data set. Therefore, we supplemented analyses of sensitivity against E by evaluating the difference in sensitivity between the pMMI and O/E against E. We calculated the difference as the adjusted Wald interval for a difference in proportions with matched pairs (Agresti and Min 2005) with the *PropCIs* package in R (Scherer 2013). The difference between the indices should be constant if E has no influence on sensitivity, or if E affects both indices in the same way. In the absence of direct measures of stress levels, these analyses provide a good measure of the influence of E on index sensitivity.

Establishment of biological condition classes, and application to a statewide assessment

We created 4 condition classes based on the distribution of scores at reference calibration sites, with a recommended interpretation for each condition class: likely to be intact (>30th percentile of reference calibration site CSCI scores), possibly altered (10th–30th percentiles), likely to be altered (1st–10th percentile), and very likely to be altered (<1st percentile). We used the *qnorm*() function in R to estimate thresholds from the observed mean and SD of reference calibration site CSCI scores. We explored other approaches to setting thresholds, such as varying thresholds by ecoregion or setting thresholds from environmentally similar reference sites, but rejected these approaches because of their added complexity and minimal benefits (Appendix S1).

We applied thresholds to a subset of sites from probabilistic surveys (n = 1318 sites) to provide weighted estimates of stream condition in California and for each major region. We also used the thresholds to make unweighted estimates of reference, moderate-activity, and high-activity sites for each region of the state. We used unweighted estimates because few reference probabilistic samples were available in certain regions. For weighted estimates, we calculated site weights by dividing total stream length in each stratum by the number of sampled sites in that stratum (these strata were defined as the intersections of strata from each contributing survey). All weight calculations were conducted using the *spsurvey* package (Kincaid and Olsen 2013) in R (version 2.15.2). We used site weights to estimate regional distributions for environmental variables using the Horvitz–Thompson estimator (Horvitz and Thomson 1952). Confidence intervals for estimates of the proportion of California's stream length meeting reference criteria were based on local neighborhood variance estimators (Stevens and Olsen 2004).

RESULTS

Biological and environmental diversity of California

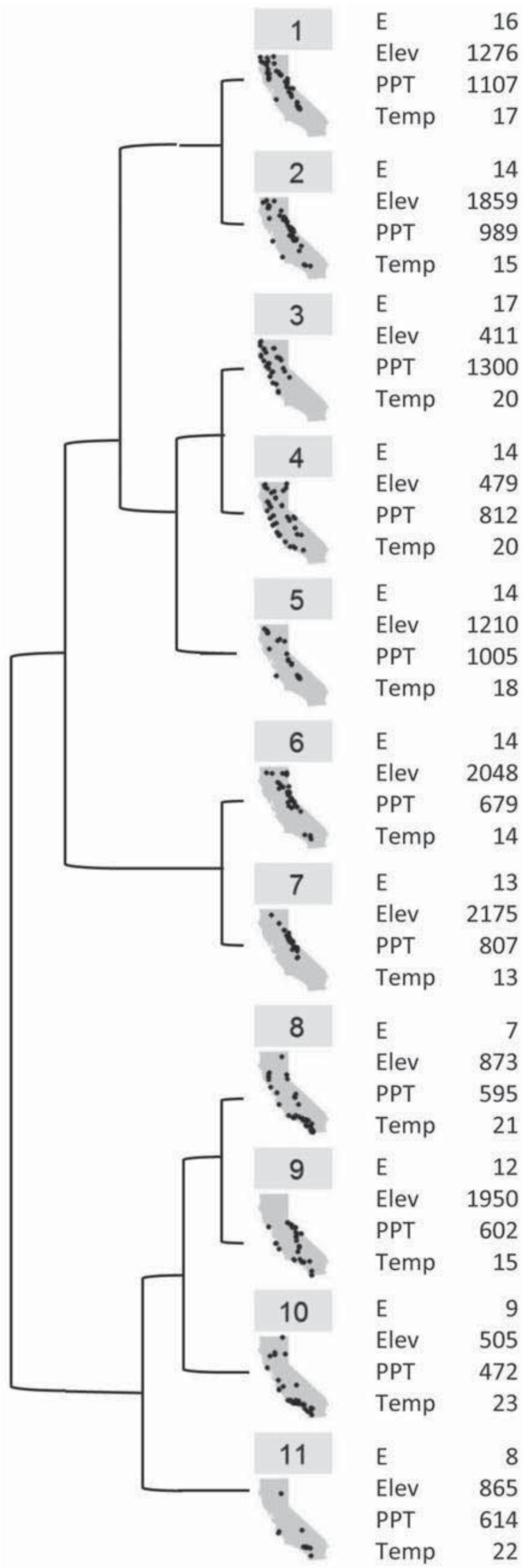
Biological assemblages varied markedly across natural gradients in California, as indicated by cluster analysis. We identified 11 groups that contained 13 to 61 sites (Fig. 3). A few of these groups were geographically restricted, but most were distributed across many regions of the state. For example, sites in group 10 were concentrated in the Transverse Ranges of southern California, and sites in group 7 were entirely within the Sierra Nevada. In contrast, sites in groups 1 and 4 were broadly distributed across the northern ⅔ of California.

Environmental factors differed among several groups. Groups 8 through 11, all in the southern portions of the state, were generally drier and hotter than other groups, whereas groups 1 through 5, predominantly in mountainous and northern regions, were relatively wet and cold. Expected number of taxa also varied across groups. For example, the highest median E (i.e., sum of capture probabilities > 0.5) (17.2) was observed in group 3, whereas the lowest (7) was observed in group 8. The median E was <10 for 3 of the 11 groups (groups 8, 10, and 11). Sites in low-E groups were preponderantly (but not exclusively) in the southern portions of the state.

Development of predictive models

Predicting the number of locally common taxa for the O/E index The random-forest model selected to predict assemblage composition used 5 predictors: latitude, elevation, watershed area, mean annual precipitation, and mean annual air temperature (Table 1). The model explained 74 and 64% of the variation in O at calibration and validation sites, respectively. Regression slopes (1.05 and 0.99 at calibration and validation sites, respectively) and intercepts (–0.36 and 0.52) were similar to those expected from unbiased predictions (i.e., slope = 1 and intercept = 0, p > 0.05). The random-forest model was modestly more precise (SD = 0.19) than the null model (SD = 0.21) but substantially less precise than the best model possible (SD = 0.13).

Predicting metric values and developing the pMMI Predictive models explained >20% of variance in 17 of the 48 metrics evaluated for inclusion in the pMMI (a subset of which are shown in Table 3). For 10 metrics, ≥30% of



the variance was explained, and for 2 metrics (no. intolerant taxa and % intolerant taxa), >50% of the variance was explained. Squared correlation coefficients (r^2) between predicted and observed metric values ranged from near 0 (e.g., Simpson diversity) to >0.5 (no. and % intolerant taxa metrics). Results for validation reference sites were consistent with results for calibration sites, but r^2 values differed markedly between calibration and validation data sets for some metrics (Table 3). In general, models explained the most variance for %-taxa metrics, and the least for %-abundance metrics, but this pattern was not consistent for all groups of metrics.

Metrics selected for the pMMI Of the 48 metrics evaluated, 15 met all acceptability criteria (Table 3). The bias criterion was the most restrictive and eliminated 21 metrics, including all raw metrics and 2 modeled metrics (% climber taxa and % predators). The discrimination criterion eliminated 15 metrics, most of which were already eliminated by the bias criterion. Other criteria eliminated few metrics, all of which were already rejected by other criteria. The 15 acceptable metrics yielded 28,886 possible prototype pMMIs ranging in size from 5 to 10 metrics, but only 234 prototype pMMIs contained uncorrelated metrics or metrics belonging to unique metric groups (data not shown). All of these prototype pMMIs contained ≤7 metrics. Of these 234 prototypes, only 6 were in the top quartile for both discrimination between reference and high-activity calibration samples and for lowest SDs among reference calibration samples.

The final pMMI included 1 metric from each of 6 metric groups (Table 3). Some of the selected metrics (e.g., Coleoptera % taxa) were similar to those used in regional indices previously developed in California (e.g., Ode et al. 2005). However, other widely used metrics (e.g., noninsect metrics) were not selected because they were highly correlated with other metrics that had better performance (pairwise correlations not shown).

The random-forest models varied in how much of the variation in the 6 individual metrics they explained (Pseudo-R^2 range: 0.23–0.39). Regressions of observed on predicted values for reference calibration data showed that several intercepts were significantly different from 0 and slopes were significantly different from 1 (i.e., $p < 0.05$), but these differences were small. The number of predictors used in each of the 6 models ranged from 2 (for no. Coleoptera

Figure 3. Dendrogram and geographic distribution of each group identified during cluster analysis. Numbers next to leaves are median values for expected number of taxa (E), elevation (Elev, m), precipitation (PPT, mm), and air temperature (Temp, °C).

taxa) to 10 (for taxonomic richness) (Table 1). Predictors related to location (e.g., latitude, elevation) were widely used, with latitude appearing in every model. In contrast, predictors related to geology (e.g., soil erodibility) or catchment morphology (e.g., watershed area) were used less often. In general, the most frequently used predictors also had the highest importance in the predictive models, as measured by % increase in mean square error. The least frequently used predictor (i.e., % P geology) was used in 1 model (taxonomic richness).

Performance of predictive models

Effects of predictive modeling on metrics For most metrics, reducing the influence of natural gradients through predictive modeling reduced the calculated difference between high-activity and reference sites, a result suggesting that stressor and natural gradients can have similar and confounded effects on many metric values (Table 3). For example, for 27 of the 48 metrics evaluated, the absolute t-statistic was much higher (difference in $|t| > 1$) for the raw metric than for the residuals. In contrast, the absolute t-statistic for residuals was higher for only 12 metrics.

Performance evaluation of the O/E, pMMI, and combined indices By all measures, predictive indices (whether used alone or combined) performed better than their null counterparts, particularly with respect to accuracy/bias (Table 5). For example, mean regional differences in null index scores at reference sites were large and significant (Fig. 4A, C, E), and responses to natural gradients were strong (Fig. 5A–O). In contrast, all measures of biases were greatly reduced for predictive indices (Fig. 4B, D, F).

Predictive modeling improved several aspects of precision. Variability of scores among reference sites was lower for all predictive indices than for their null counterparts, particularly for the pMMI (Table 5). Regional differences in precision were larger for the pMMI than O/E (both predictive and null models), and combining these 2 indices into the CSCI improved regional consistency in precision (Fig. 4B, D, F). Predictive modeling had a negligible effect on within-site variability (Table 5).

In contrast to precision and accuracy, responsiveness was more affected by index type than whether predictive or null models were used. Both predictive and null MMIs appeared to be slightly more responsive than the combined indices, which in turn were more responsive than O/E indices. This pattern was evident in all measures of responsiveness, such as magnitude of t-statistics, variance explained by multiple human-activity gradients in a random-forest model, and steepness of slopes against individual gradients (Table 5, Fig. 6A–I).

Analysis of sensitivity indicated stronger sensitivity of the pMMI than the O/E, and the combined index had intermediate sensitivity. Overall, 47% of nonreference sites had scores <10th percentile of reference calibration sites for the CSCI, in contrast with 52% of the pMMI and 35% of the O/E. Despite the overall difference between the pMMI and the O/E, agreement was relatively high (76%) when the 10th percentile was used as an impairment threshold (i.e., O/E $\geq$ 0.76 and pMMI $\geq$ 0.77). When the 1st percentile was used to set thresholds (i.e., O/E $\geq$ 0.56 and pMMI $\geq$ 0.58), the agreement rate was 90%.

Table 5. Performance measures to evaluate California State Condition Index (CSCI), MMI = multimetric index, and observed (O)/expected (E) taxa index at calibration (Cal) and validation (Val) sites. For accuracy tests, only reference sites were used. Ref mean = mean score of reference sites (* indicates value is mathematically fixed at 1), F = F-statistic for differences in scores at calibration sites among 5 regions (shown in Fig. 1, Central Valley excluded; residual df = 467), Var = variance in index scores explained by natural gradients at reference sites, among sites = standard deviation of scores at reference sites, within sites = standard deviation of within-site residuals for reference Cal (n = 220 sites) and Val (n = 60) sites with multiple samples, t = t-statistic for difference between mean scores at reference and high-activity sites, var = variance in index scores explained by human-activity gradients at all sites.

		Accuracy						Precision				Responsiveness			
		Ref mean		F		Var		Among sites		Within sites		t		Var	
Index	Type	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
CSCI	Predictive	1.01	1.01	1.3	1.4	−0.08	−0.13	0.16	0.17	0.11	0.1	28.5	13	0.49	0.42
	Null	1*	1	52.9	4.7	0.41	0.12	0.21	0.2	0.11	0.11	28.6	14.8	0.64	0.58
MMI	Predictive	1*	0.98	0.8	1.3	−0.15	−0.09	0.18	0.19	0.12	0.12	30.9	14.4	0.54	0.48
	Null	1*	1	62.2	8.7	0.46	0.2	0.24	0.24	0.12	0.12	29.2	15.3	0.67	0.61
O/E	Predictive	1.02	1.03	1.2	1	0.01	−0.12	0.19	0.2	0.16	0.13	21.0	9.3	0.31	0.25
	Null	1*	1	23.5	0.9	0.23	−0.03	0.21	0.22	0.15	0.13	24.1	11.8	0.48	0.41

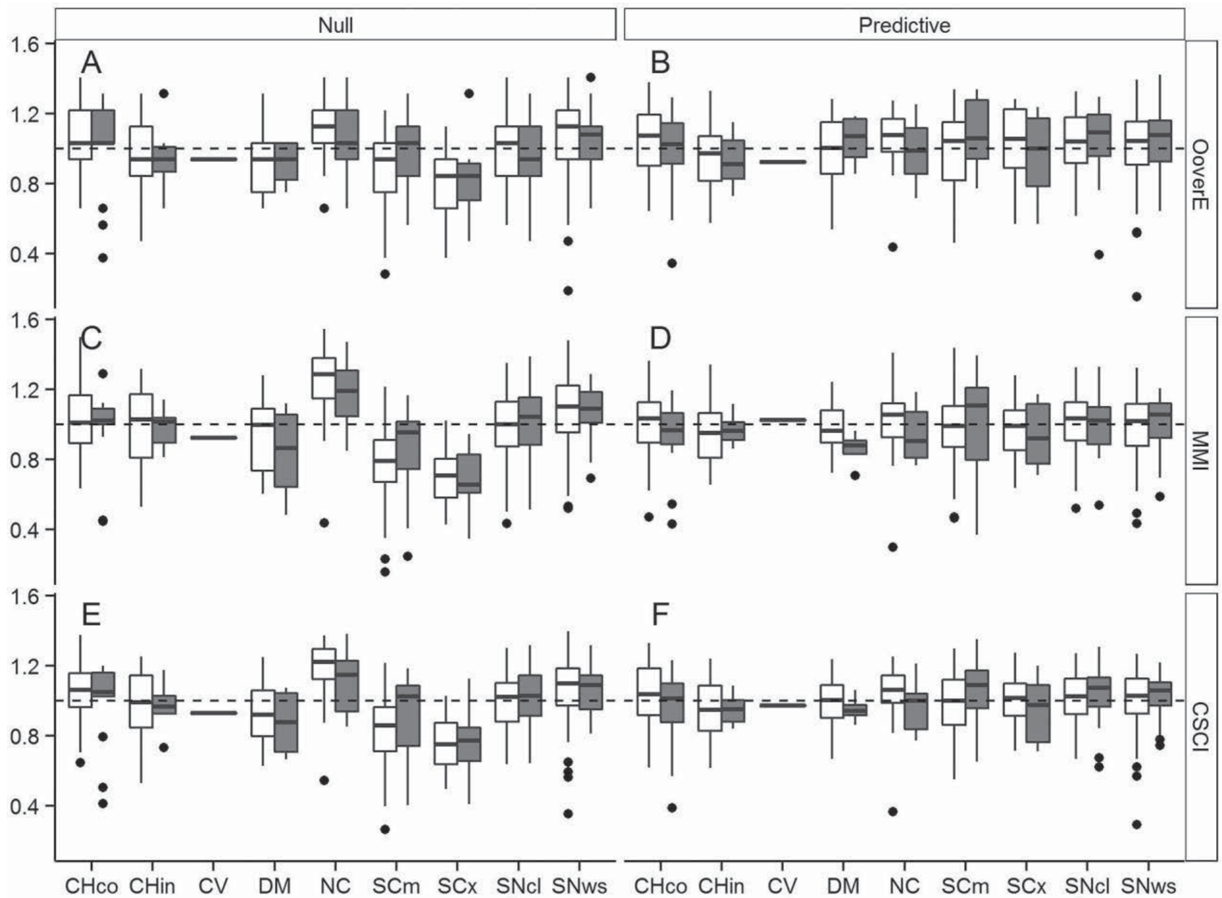


Figure 4. Box-and-whisker plots for distribution of scores for null (A, C, E) and predictive (B, D, F) models for the observed (O)/expected (E) taxon index (A, B), multimetric index (MMI) (C, D), and the combined index (CSCI) (E, F) scores by geographic region (see Fig. 1 for codes). White boxes indicate scores at calibration sites, and gray boxes indicate scores at validation sites. The horizontal dashed lines indicate the expected value at reference sites (= 1). Lines in boxes are medians, box ends are quartiles, whiskers are 1.5× the interquartile range, and dots are outliers (i.e., values >1.5× the interquartile range).

Effect of E on performance By most measures, performance was better at high-E than at low-E sites, but predictive indices were much more consistent than their null equivalents. For example, the accuracy of null indices was very poor at low-E sites (0.46–0.54 at E = 5; Fig. 7A), whereas predictive indices were much more accurate (0.73–0.86 at E = 5; Fig. 7E. At high-E sites, accuracy was >0.90 for both predictive and null indices. Precision was better at high-E sites for the pMMI and O/E index, but the CSCI had better and more consistent precision than the other indices at all values of E (Fig. 7B, F). For example, precision ranged from 0.22 to 0.15 (range = 0.07) for both the pMMI and the O/E, whereas it ranged from 0.18 to 0.14 (range = 0.04) for the CSCI.

In contrast to the weak associations between E and accuracy and precision, E was very strongly associated with sensitivity, as measured by the percentage of high-activity sites with scores <10th percentile threshold (Fig. 7C, G). The pMMI classified a larger proportion of sites as in nonreference condition across nearly all values of E than the O/E index did, but the difference was largest at low-E sites (Fig. 7D, H). For example, at the lowest values of E analyzed (5), the pMMI identified 87% of high-activity sites as biologically different from reference, whereas O/E identified only 47% of sites as in nonreference condition. As E increased, the difference between the 2 indices in proportion of sites classified as nonreference decreased. Wald's interval test indicated significant differences between the indices for values of E up to 13. At low-E sites, the sensitivity of the CSCI was between the 2 indices, but at high-E sites, CSCI was more similar to pMMI. All 3 indices showed that low-E sites were more pervasively in nonreference condition than high-E sites, and the proportion of sites with scores <10th percentile of reference calibration sites decreased as E increased. In contrast to precision and accuracy, sensitivity was more consistent across settings for

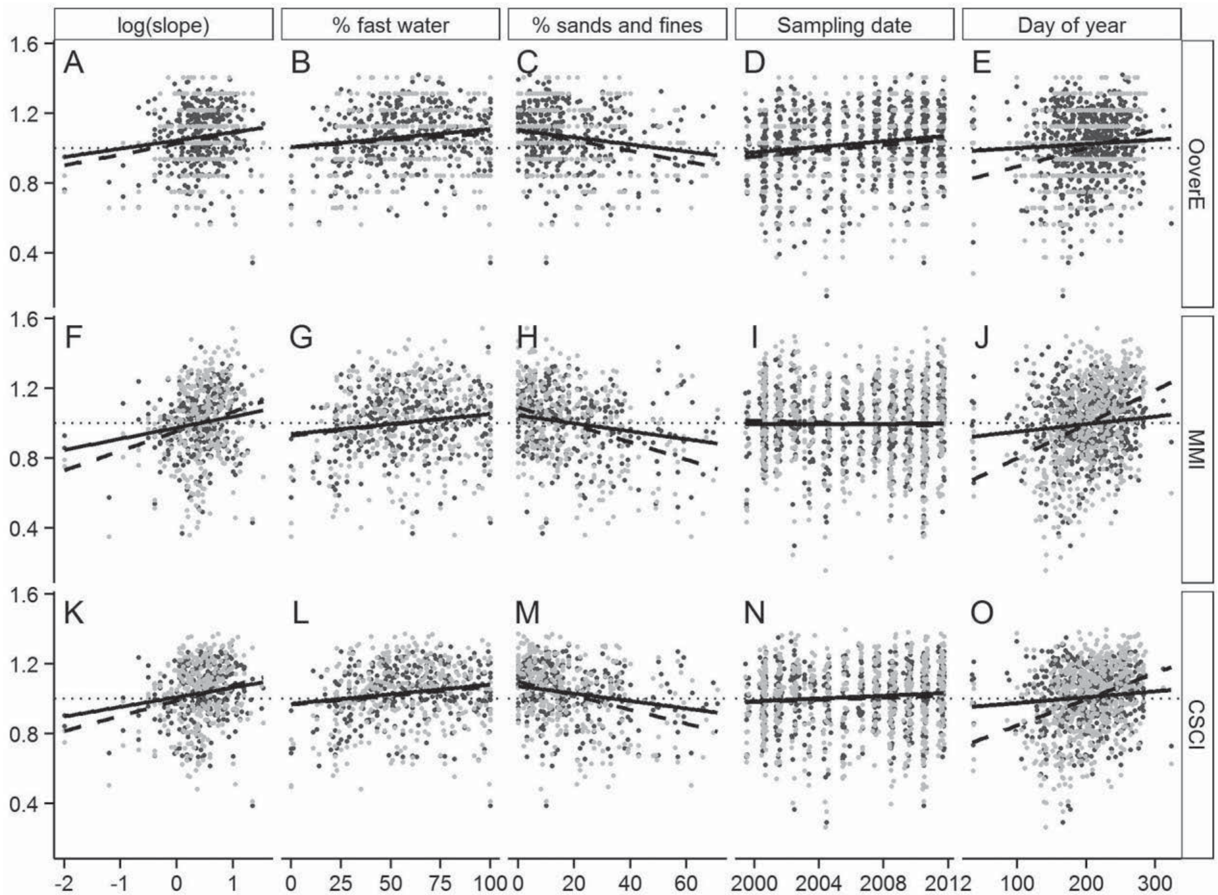


Figure 5. Relationships between observed (O)/expected (E) taxon index (A–E), multimetric index (MMI) (F–J), and the combined index (CSCI) (K–O) scores and slope (A, F, K), % fast water (area of reach with riffle, run, cascade, or rapid microhabitats) (B, G, L), % sand and fines (C, H, M), sampling date (D, I, N), and day of the year (E, J, O) at reference sites for predictive (black symbols, solid lines) and null (gray symbols, dashed lines) indices. The dotted line indicates a perfect relationship without bias.

null than predictive indices. For all analyses of performance relative to E, validation data yielded similar results (not shown).

Establishment of biological condition classes and application to a statewide assessment

We established 4 biological condition classes based on the distribution of CSCI scores at reference calibration sites. Statewide, 52% of streams were likely to be intact (i.e., CSCI ≥ 0.92 [30th percentile of reference calibration sites]). Another 18% were possibly altered (i.e., CSCI ≥ 0.79 [10th percentile]), 11% were likely to be altered (i.e., CSCI ≥ 0.63 [1st percentile]), and 19% were very likely to be altered (i.e., CSCI < 1st percentile) (Table 6). Although many (i.e., 49%) high-activity sites were very likely to be altered, this number varied considerably by region. Few high-activity sites were in this condition class in the more forested regions (e.g., 24% in the North Coast, 15% in the Sierra Nevada), whereas higher numbers were observed in relatively arid regions (e.g., 100% in the Desert/Modoc region and 68% in the Central Valley). In contrast, the percentage of reference sites in the top 2 classes varied much less across regions, from a low of ~85% in the South Coast and Desert/Modoc regions to a high of 98% in the North Coast (Table 6).

DISCUSSION

Our evaluation of index performance across different environmental settings demonstrates that, to the greatest extent possible with existing data, we have designed an index with scores that have comparable meanings for different stream types in an environmentally heterogeneous region of the USA. Each site is benchmarked against appropriate biological expectations anchored by a large and consistently defined reference data set, and deviations from these expectations reflect site condition in a consistent way across environmental settings. Thus, the index can be used to evaluate the condition of nearly all perennial streams in California, despite the region's considerable environmental and biological complexity. Three ele-

B145

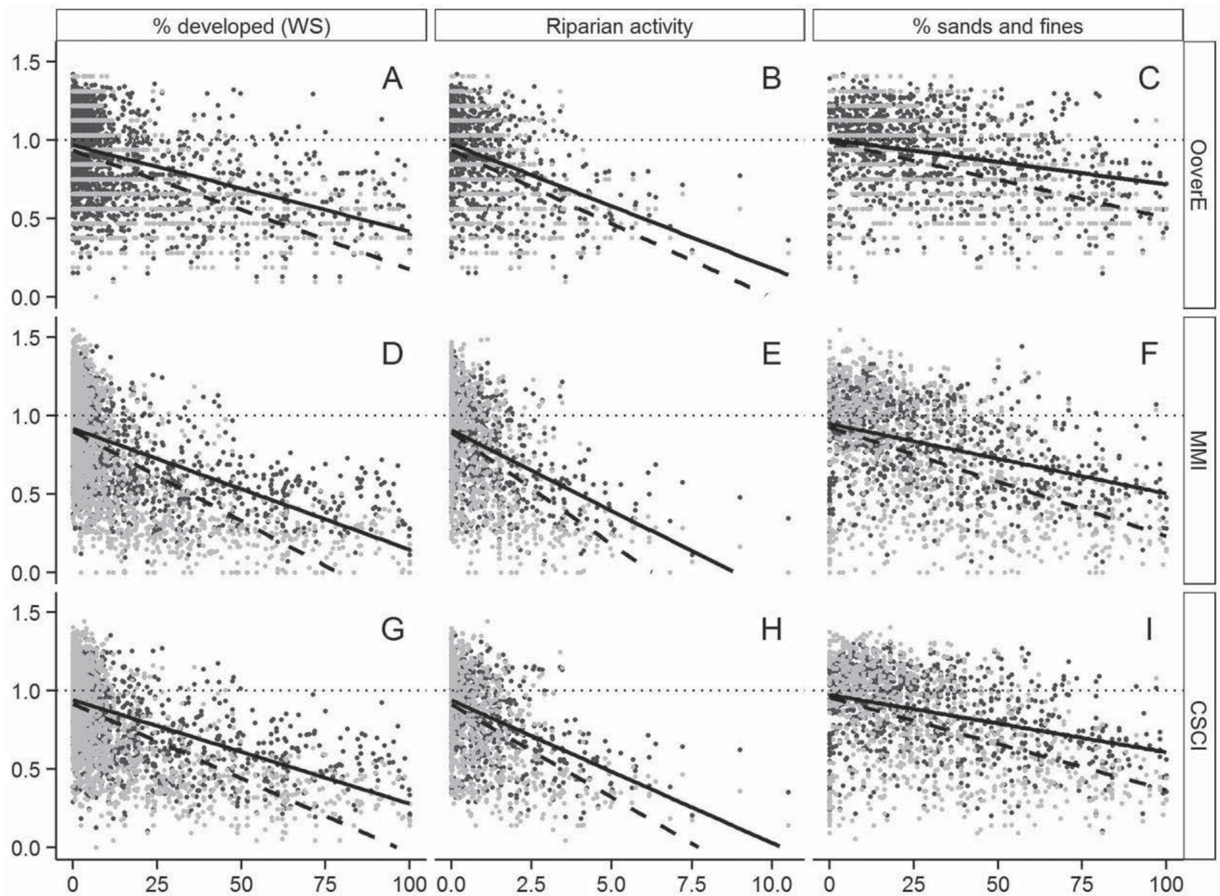


Figure 6. Relationships between observed (O)/expected (E) taxon index (A–C), multimetric index (MMI) (D–F), and the combined index (CSCI) (G–I) scores and % developed area of the watershed (WS) (A, D, G), riparian activity (B, E, H), and % sand and fines (C, F, I) for predictive (black symbols, solid lines) and null indices (gray symbols, dashed lines). The dotted line indicates the reference expectation of 1.

ments of the design process contributed to the utility of this index in an environmentally complex region: a robust reference data set, predictive modeling, and the combination of multiple endpoints into a single index.

Large, representative reference data sets

The 1st element was the large, representative, and rigorously evaluated reference data set (Ode et al. 2016). Natural factors that influence biological assemblages must be adequately accounted for to create an assessment tool that performs well across environmental settings (Cao et al. 2007, Schoolmaster et al. 2013). The strength of relationship between natural factors and biology varies with geographic scale (Mykrä et al. 2008, Ode et al. 2008), and representing locally important factors (such as unusual geology types with limited geographic extent, e.g., Campbell et al. 2009) contributes to the ability of the index to distinguish natural from anthropogenic biological variability in these environmental settings. Our reference data set was spatially representative and encompassed >10 y of sampling. Long-term temporal coverage improves the representation of climatic variability, including El Niño-related storms and droughts. The spatial and temporal breadth of sampling at reference sites provides confidence in the applicability of the CSCI for the vast majority of wadeable perennial streams in California.

Predictive modeling

The 2nd element of the CSCI's design, predictive modeling, enabled the creation of site-specific expectations for 2 indices, and these models created indices superior to those created by null models in nearly every aspect, particularly with respect to bias in certain settings. These results are consistent with a large body of literature showing similar results for indices that measure changes in taxonomic composition (e.g., Reynoldson et al. 1997, Hawkins et al. 2000, Van Sickle et al. 2005, Hawkins 2006, Mazor et al. 2006). However, few studies to date showed that the benefits extend to MMIs (e.g., Bates Prins and Smith 2007, Pont et al. 2009, Hawkins et al. 2010b, Schoolmaster et al. 2013, Vander Laan and Hawkins 2014).

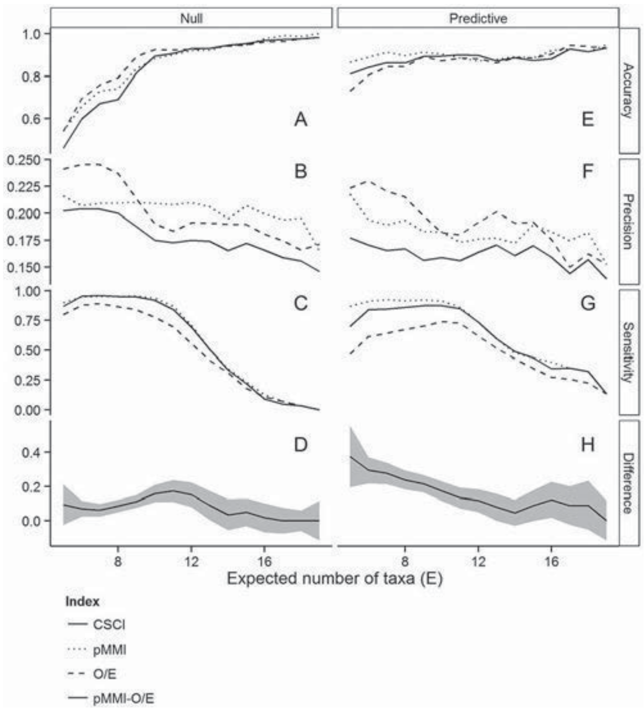


Figure 7. Effect of expected number of taxa (E) on accuracy (A, E), precision (B, F), sensitivity (C, G), and difference in sensitivity between the predictive multimetric index (pMMI) and the observed (O)/expected (E) taxa indices (D, H) for null (A–D) and predictive (E–H) index performance. The gray bands in the bottom panels C and G indicate the 95% confidence interval around the difference. Accuracy = proportion of reference calibration sites in reference condition (i.e., score >10th percentile of reference calibration sites) for each index. Precision = standard deviation of reference calibration sites for each index. Sensitivity = proportion of high-activity sites not in reference condition.

Our preference for predictive over traditional MMIs is not based only on the superior performance the pMMI relative to its null counterpart. The null MMI evaluated in our study was simplistic and did not reflect typical typological approaches to MMI development, which include regionalization in metric selection (e.g., Stoddard et al. 2008), regionalization in scoring (e.g., Ode et al. 2005), or normalization to watershed area (e.g., Klemm et al. 2003) to account for variability across reference sites. However, traditional MMIs based on regionalization usually lack metric and scoring standardization, which complicates interregional comparisons. Even if typological approaches provided equivalent performance to predictive indices, the latter would be preferred because of their ability to set site-specific management goals because predictive indices can better match the true potential of individual sites (Hawkins et al. 2010b). Thus, a watershed manager could take action to maintain a level of diversity a stream can truly support, rather than a level typical of potentially dissimilar reference sites.

Combining multiple indices

The 3rd element of the CSCI's design that contributed to its utility in different stream types was inclusion of both the pMMI and the O/E index. Regulatory agencies expressed a strong preference for a single index to support biocriteria implementation, and we thought that the CSCI was preferable to either the pMMI or O/E index. The different sensitivities of the 2 components should enhance the utility of the CSCI across a broad range of disturbances and settings. Together, they provide multiple lines of evidence about the condition of a stream and provide greater confidence in the results than a single index that might be biased in certain settings. Use of both metric and multivariate indices is widespread in assessments of coastal condition (e.g., the M-AMBI index; Muxika et al. 2007) specifically because the combination takes advantage of the unique sensitivities of each index in different habitat types (Sigovini et al. 2013). Applications of a multiple-index approach in stream assessment programs are uncommon, but the need has been suggested (e.g., Reynoldson et al. 1997, Mykrä et al. 2008, Collier 2009).

The decision to use both the pMMI and O/E index was based, at least partly, on observations that they had different sensitivities in different settings, particularly at low-E sites. The difference between the 2 indices might mean that the O/E index correctly indicates a greater resilience to stress at certain stream types or that the pMMI is more finely tuned to lower levels of stress simply because it was specifically calibrated against high-activity sites in similar settings. Mechanistically, the difference probably occurred because O/E index scores are mainly affected by the loss of common taxa. For example, in low-E sites (which were common in dry, low-elevation environments in southern and central coastal California), the O/E index predicted occurrence of only a small number of highly tolerant taxa (e.g., baetid mayflies) because only these tolerant taxa occur with high probability in these naturally stressful environments. Sensitive taxa also occur at reference sites in drier, low-elevation settings, but they were typically too rare to affect the O/E index (Appendix S2).

The interpretive value of rare, sensitive taxa in estimation of biological integrity of an individual site is unclear, but the ability of a site to support these taxa may be important to the health of a dynamic metacommunity, where rare taxa occupy only a small subset of suitable sites at any one time. Although several investigators have shown that exclusion of rare taxa usually enhances precision of O/E indices (e.g., Ostermiller and Hawkins 2004, Van Sickle et al. 2007), our results suggest that in certain settings, this exclusion may obscure an important response to

Table 6. Percentage of sites in different condition classes by region and site status. Percentiles refer to the distribution of scores at reference calibration (Cal) sites. Overall estimates are based on sites from probabilistic surveys and are not split into Cal or validation (Val) sets. For reference, moderate-, and high-activity sites, numbers in the last 6 columns are percentage of sites. For overall assessments, these numbers are percentage of stream miles. Dashes indicate that no sites were analyzed.

	Total sites		Likely to be intact ≥30th percentile (CSCI ≥ 0.92)		Possibly altered 30th–10th percentile (CSCI ≥ 0.79)		Likely to be altered 1st–10th percentile (CSCI ≥ 0.63)		Very likely to be altered <1st percentile (CSCI < 0.63)	
Region	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
Statewide										
Reference	473	117	75	74	15	16	8	8	1	3
Moderate activity	626	156	53	56	20	20	18	17	8	7
High activity	497	122	13	18	13	14	25	22	49	46
Overall	919		52		18		11		19	
North Coast										
Reference	60	16	85	63	13	31	0	6	2	0
Moderate activity	88	26	58	50	26	15	9	27	7	8
High activity	45	9	29	67	33	33	13	0	24	0
Overall	162		58		23		10		9	
Chaparral										
Reference	74	19	68	63	20	26	9	0	3	11
Moderate activity	146	34	47	65	18	15	29	15	6	6
High activity	126	28	18	21	13	7	18	11	50	61
Overall	147		34		16		17		33	
South Coast										
Reference	96	23	70	70	16	9	14	22	1	0
Moderate activity	202	52	49	52	22	23	19	17	9	8
High activity	241	60	5	10	12	13	32	27	52	50
Overall	387		44		16		16		24	
Sierra Nevada										
Reference	221	55	77	82	14	11	7	5	1	2
Moderate activity	148	35	68	60	20	29	8	9	5	3
High activity	27	8	56	25	11	38	19	13	15	25
Overall	106		70		19		6		5	
Central Valley										
Reference	1	0	100	–	0	–	0	–	0	–
Moderate activity	8	1	0	0	0	0	38	100	63	0
High activity	47	13	0	0	4	8	28	38	68	54
Overall	60		2		8		18		71	
Desert/Modoc										
Reference	21	4	71	75	14	25	14	0	0	0
Moderate activity	34	8	44	63	9	0	29	13	18	25
High activity	5	4	0	50	0	0	0	50	100	0
Overall	57		48		14		9		30	

stress. Including rare taxa in certain environmental settings while excluding them in others may improve the consistency of an O/E index in complex regions, but we did not explore this option. The observation that sensitivity of all indices was lowest where E was highest was unexpected, and may be attributed to several potential causes. Most probably, anthropogenic stress was less severe at high-E than at low-E sites. High-activity sites were identified via indirect measures based on stressor sources (e.g., development in the watershed) rather than direct measures of water or habitat quality, so we could not ensure homogenous levels of disturbance among this set of sites. Alternatively, high-E settings might be more resilient to stress, perhaps because of their greater diversity (Lake 2000). Thus, the indices may have different responses to the same level of stress in different settings, depending on E.

Despite the lower sensitivity of the O/E index at low-E sites, we think that including it in a combined index was preferable to using the more sensitive pMMI by itself. Combining the 2 indices was a simple way to retain high sensitivity at low-E sites, while retaining the advantages of the O/E as a measure of biodiversity (Moss et al. 1987, Hawkins et al. 2000). The ability of the O/E index to measure taxonomic completeness has direct applications to conservation of biodiversity and makes it particularly sensitive to replacement of native fauna by invasive species. Furthermore, because it is calibrated with only reference sites, the O/E index is not influenced by the distribution or quality of high-activity sites. In contrast, we used the pMMI under the assumption that the set of high-activity sites adequately represented the types of stressors that might be encountered in the future. Inclusion of the O/E index in the CSCI provides a degree of insurance against faulty assumptions about the suitability of the high-activity site set for pMMI calibration.

We combined the 2 indices as an unweighted mean for several technical reasons, but primarily because this was the simplest approach to take without stronger support for more complicated methods. As we demonstrated, the CSCI has less variable performance across stream types than its 2 components. Approaches that let the lowest (or highest) score prevail are more appropriate when the components have similar sensitivity, but in our case would be tantamount to using the pMMI alone and muting the influence of the O/E index. Approaches that weight the 2 components based on site-specific factors (e.g., weighting the pMMI more heavily than the O/E index at low-E sites) are worthy of future exploration. Evaluating the pMMI and O/E indices independently to assess biological condition at a site might be useful, particularly at low-E sites, but the combined index is preferred for applications where statewide consistency is important, such as designation of impaired waterbodies.

Unexplained variability

In our study, predictive models were able to explain only a portion of the variability observed at reference sites—sometimes a fairly small portion. For example, the SD of the predictive O/E was only slightly lower than the SD of the null O/E (0.19 vs 0.21) and much larger than that associated with replicate samples (0.13). None of the selected random-forest models explained >39% (for the no. shredder taxa metric) of the variability at reference calibration sites. The unexplained variability may be related to the additional effects of environmental factors that are unsuitable for predicting reference condition (e.g., alterable factors, like substrate composition or canopy cover), environmental factors unrelated to those used for modeling (e.g., temporal gradients, weather antecedent to sampling), field and laboratory sampling error, metacommunity dynamics (Leibold et al. 2004, Heino 2013), or neutral processes in community assembly that are inherently unpredictable (Hubbell 2001, Rader et al. 2012). The relative contribution of these factors is likely to be a fruitful area of bioassessment research. Given the number and breadth of environmental gradients evaluated for modeling, we think it unlikely that additional data or advanced statistical methods will change the performance of these indices.

Setting thresholds

Some investigators have suggested that thresholds for identifying impairment in environmentally complex regions may require different thresholds in different settings based on the variability of reference streams in each setting. For example, Yuan et al. (2008) proposed ecoregional thresholds for an O/E index for the USA based on the observation that index scores at reference sites were twice as variable in some ecoregions as in others. Alternatively, site-specific thresholds could be established based on the variability of a subset of environmentally similar reference sites. We rejected both of these approaches in favor of uniform thresholds based on the variability of all reference calibration sites. We rejected ecoregional thresholds or other typological approaches because the validity of ecoregional classifications may be questionable for sites near boundaries. We rejected site-specific thresholds based on environmentally similar reference sites because they did not improve accuracy or sensitivity relative to a single statewide threshold when predictive indices are used (Appendix S1). These results are consistent with those of Linke et al. (2005), who showed that indices calibrated with environmentally similar reference sites had similar performance to indices based on predictive models that were calibrated with all available reference sites. Other approaches, such as direct modeling of the SD of index scores as a function of natural factors, also might improve comparability of scores across settings (R. Bailey, Cape Breton University, personal communication).

Conclusions and recommended applications

Many recent technical advances in bioassessment have centered on improving the performance of tools used to score the ecological condition of water bodies. Much of the progress in this area has come from regional, national, and international efforts to produce overall condition assessments of streams in particular regions (e.g., Simpson and Norris 2000, Van Sickle et al. 2005, Hawkins 2006, Hering et al. 2006, Stoddard et al. 2006, Paulsen et al. 2008). A key challenge in completing these projects has been incompatibility among scoring tools designed to assess streams in multiple regions, each calibrated for unique and locally important environmental gradients (Cao and Hawkins 2011). This issue has been well documented for large-scale programs in which investigators have attempted to integrate scores from a patchwork of assessment tools built for smaller subregions (Heinz Center 2002, Hawkins 2006, Meador et al. 2008, Pont et al. 2009), but far less attention has been paid to the meaning of index scores at individual stream reaches (Herlihy et al. 2008, Ode et al. 2008). Assessment of CSCI performance across the range of environmental settings in California was essential because the CSCI is intended for use in regulatory applications that affect the management of individual reaches, and consistent meaning of a score was a key requirement of regulatory agencies and stakeholders. We attempted to maximize consistency of the CSCI by using a large and representative reference set and by integrating multiple indices based on predictive models. Consistent accuracy was attained through the use of predictive models, whereas the consistency of precision and sensitivity was improved through the use of multiple endpoints.

The CSCI was designed for condition assessments, but we think it has broad application to many aspects of stream management. For example, it could be used to select comparator sites with similar biological expectations to test sites for use in causal assessments (e.g., CADDIS; USEPA 2010) or to prioritize streams that can support rare or threatened assemblages for restoration or conservation (Linke et al. 2011). The predictions generated by the index can inform management decisions about streams for which no biological data are available. Predictive indices, such as the CSCI, are powerful additions to the stream manager's tool kit, especially in environmentally complex areas. We recognize the challenges in enabling the general public to calculate an index as complex as the one presented here. Fortunately, online automation of many of the steps is possible. For example, much of the GIS analysis can be simplified by using publicly available resources like StreamStats (US Geological Survey 2012). An automated tool is in development, but people who are interested in using the CSCI or examining its component models are encouraged to contact the authors.

ACKNOWLEDGEMENTS

We thank Dave Buchwalter, Rick Hafele, Chris Konrad, LeRoy Poff, John Van Sickle, Lester Yuan, Jason May, Larry Brown, Ian Waite, Kevin Lunde, Marco Sigala, Joseph Furnish, and Betty Fetscher for contributions to the approach described in this paper. Financial support was provided by grants from the US EPA Region IX and the California State Water Resources Control Board. Most of the data used in this analysis were provided by the California Surface Water Ambient Monitoring Program, the California Department of Fish and Wildlife, the Sierra Nevada Aquatic Research Lab, the Stormwater Monitoring Coalition of Southern California, the US Forest Service, and the Geographic Information Center at California State University Chico. We thank Kevin Collier, Bob Bailey, and 1 anonymous referee for their valuable feedback on the manuscript. This paper is dedicated to the memory of Miriam D. Mazor.

LITERATURE CITED

Agresti, A., and Y. Min. 2005. Simple improved confidence intervals for comparing matched proportions. Statistics in Medicine 24:729–740.

Angradi, T. R., M. S. Pearson, D. W. Bolgrein, T. M. Jicha, D. L. Taylor, and B. H. Hill. 2009. Multimetric macroinvertebrate indices for mid-continent US great rivers. Journal of the North American Benthological Society 28:785–804.

Bates Prins, S. C., and E. Smith. 2007. Using biological metrics to score and evaluate sites: a nearest-neighbour reference condition approach. Freshwater Biology 52:98–111.

Blocksom, K. A. 2003. A performance comparison of metric scoring methods for a multimetric index for Mid-Atlantic highland streams. Environmental Management 42:954–965.

California State Water Resources Control Board. 2013. Regional Water Quality Control Board boundaries. California State Water Resources Control Board, Sacramento, California. (Available from: http://www.waterboards.ca.gov/waterboards_map.shtml)

CAMLnet. 2003. List of California macroinvertebrate taxa and standard taxonomic effort. California Department of Fish and Game, Rancho Cordova, California. (Available from: www.safit.org)

Campbell, R. H., T. H. McCulloh, and J. G. Vedder. 2009. The Miocene Topanga group of southern California—a 100-year history of changes in stratigraphic nomenclature. Open-File Report 2007-1285. Version 1.1. US Geological Survey, Reston, Virginia.

Cao, Y., and C. P. Hawkins. 2011. The comparability of bioassessments: a review of conceptual and methodological issues. Journal of the North American Benthological Society 30:680–701.

Cao, Y., C. P. Hawkins, J. Olson, and M. A. Kosterman. 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. Journal of the North American Benthological Society 26:566–585.

Clarke, K. R., and R. H. Green. 1988. Statistical design and analysis for a "biological effects" study. Marine Ecology Progress Series 46:213–226.

Clarke, R. T., J. F. Wright, and M. T. Furse. 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecological Modelling 160:219–233.

Collier, K. J. 2009. Linking multimetric and multivariate approaches to assess the ecological condition of streams. Environmental Monitoring and Assessment 157:113–124.

Council of European Communities. 2000. Establishing a framework for community action in the field of water policy. Directive 2000/60/EC. Official Journal of European Communities. L327(43):1–72.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, and K. T. Hess. 2007. Random forests for classification in ecology. Ecology 88:2783–2792.

Davis, W. S., and T. P. Simon. 1995. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Press, Boca Raton, Florida.

Fore, L. S., and C. Grafe. 2002. Using diatoms to assess the biological condition of large rivers in Idaho (U.S.A.). Freshwater Biology 47:2014–2037.

Gauch, H. G. 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge, UK.

Gerritsen, J. 1995. Additive biological indices for resource management. Journal of the North American Benthological Society 14:451–457.

Gerth, W. J., and A. T. Herlihy. 2006. The effect of sampling different habitat types in regional macroinvertebrate bioassessment surveys. Journal of the North American Benthological Society 25:501–512.

Hawkins, C. P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. Ecological Applications 16:1277–1294.

Hawkins, C. P., Y. Cao, and B. Roper. 2010a. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshwater Biology 55:1066–1085.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10:1456–1477.

Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010b. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society 29:312–343.

Heino, J. 2013. The importance of metacommunity ecology for environmental assessment research in the freshwater realm. Biological Reviews 88:166–178.

Heinz Center (H. John Heinz III Center for Science and the Environment). 2002. The state of the nation's ecosystems: measuring the lands, waters, and living resources of the United States. Cambridge University Press, New York.

Herbst, D. B., and E. L. Silldorff. 2006. Comparison of the performance of different bioassessment methods: similar evaluations of biotic integrity from separate programs and procedures. Journal of the North American Benthological Society 25:513–530.

Herbst, D. B., and E. L. Silldorff. 2009. Development of a benthic macroinvertebrate index of biological integrity (IBI) for stream assessments in the Eastern Sierra Nevada of California. Sierra Nevada Aquatic Research Lab, Mammoth Lakes, California. (Available from: http://www.waterboards.ca.gov/lahontan/water_issues/programs/swamp/docs/east_sierra_rpt.pdf)

Hering, D., R. K. Johnson, S. Kramm, S. Schmutz, K. Szoszkiewicz, and P. F. Verdonschot. 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshwater Biology 51:1757–1785.

Herlihy, A. T., S. G. Paulsen, J. Van Sickle, J. L. Stoddard, C. P. Hawkins, and L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference condition approach on a continental scale. Journal of the North American Benthological Society 27:860–877.

Horvitz, D. G., and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47:663–685.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity. Monographs in Population Biology 32:1–392.

Hughes, R. M., D. P. Larsen, and J. M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management 10:629–635.

Hughes, R. M., P. R. Kauffmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences 55:1618–1631.

Kaufmann, P. R., P. Levine, E. G. Robinson, C. Seeliger, and D. V. Peck. 1999. Surface waters: quantifying physical habitat in wadeable streams. EPA/620/R-99/003. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

Kincaid, T., and A. Olsen. 2013. Spsurvey: spatial survey design and analysis. R package version 105. R Project for Statistical Computing, Vienna, Austria.

Klemm, D. J., K. A. Blocksom, F. A. Fulk, A. T. Herlihy, R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, and W. T. Thoeny. 2003. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic Highlands streams. Environmental Management 31:656–669.

Kuhn, M., J. Wing, S. Weston, A. Williams, C. Keefer, and A. Engelhardt. 2012. caret: classification and regression training. R package, version 5.15-045. R Project for Statistical Computing, Vienna, Austria. (Available from: http://www.epa.gov/nheerl/arm)

Lake, P. S. 2000. Disturbance, patchiness, and diversity in streams. Journal of the North American Benthological Society 19:573–592.

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2:18–22.

Linke, S., R. H. Norris, D. P. Faith, and D. Stockwell. 2005. ANNA: a new prediction method for bioassessment programs. Freshwater Biology 50:147–158.

Linke, S., E. Turak, and J. Nel. 2011. Freshwater conservation planning: the case for systematic approaches. Freshwater Biology 56:6–20.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2012. cluster: cluster analysis basics and extensions. R package version 1.14.3. R Project for Statistical Computing, Vienna, Austria.

Mazor, R. D., T. B. Reynoldson, D. M. Rosenberg, and V. H. Resh. 2006. Effects of biotic assemblage, classification, and assessment method on bioassessment performance. Canadian Journal of Fisheries and Aquatic Sciences 63:394–411.

Meador, M. R., T. R. Whittier, R. M. Goldstein, R. M. Hughes, and D. V. Peck. 2008. Evaluation of an index of biotic integrity approach used to assess biological condition in western US streams and rivers at varying spatial scales. Transactions of the American Fisheries Society 137:13–22.

Moss, D., T. Furse, J. F. Wright, and P. D. Armitage. 1987. The prediction of macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology 17:41–52.

Muxika, I., Á. Borja, and J. Bald. 2007. Using historical data, expert judgment and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. Marine Pollution Bulletin 55:16–29.

Mykrä, H., J. Aroviita, J. Kotanen, H. Hämäläinen, and T. Muotka. 2008. Predicting the stream macroinvertebrate fauna across regional scales: influence of geographical extent on model performance. Journal of the North American Benthological Society 27:705–716.

Ode, P. R., C. P. Hawkins, and R. D. Mazor. 2008. Comparability of biological assessments derived from predictive models and multimetric indices of increasing geographic scope. Journal of the North American Benthological Society 27:967–985.

Ode, P. R., A. C. Rehn, and J. T. May. 2005. A quantitative tool for assessing the integrity of southern coastal California streams. Environmental Management 35:493–504.

Ode, P. R., A. C. Rehn, R. D. Mazor, K. C. Schiff, E. D. Stein, J. T. May, L. R. Brown, D. B. Herbst, D. Gillett, K. Lunde, and C. P. Hawkins. 2016. Evaluating the adequacy of a reference-pool site for ecological assessments in environmentally complex regions. Freshwater Science 35:237–248.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2013. vegan: community ecology package. R package version 2.0-6. R Project for Statistical Computing, Vienna, Austria.

Olson, J. R., and C. P. Hawkins. 2012. Predicting natural baseflow stream water chemistry in the western United States. Water Resources Research 48:W02504.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the Association of American Geographers 77:118–125.

Ostermiller, J. D., and C. P. Hawkins. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society 23:363–382.

Paulsen, S. G., A. Mayio, D. V. Peck, J. L. Stoddard, E. Tarquinio, S. M. Holdsworth, J. Van Sickle, L. L. Yuan, C. P. Hawkins, A. T. Herlihy, P. R. Kaufmann, M. T. Barbour, D. P. Larsen, and A. R. Olsen. 2008. Condition of stream ecosystems in the US: an overview of the first national assessment. Journal of the North American Benthological Society 27:812–821.

Peck, D. V., A. T. Herlihy, B. H. Hill, R. M. Hughes, P. R. Kaufmann, D. J. Klemm, J. M. Lazorchak, F. H. McCormick, S. A. Peterson, S. A. Ringold, T. Magee, and M. Cappaert. 2006. Environmental Monitoring and Assessment Program—Surface Waters Western Pilot study: field operations manual for wadeable streams. EPA/620/R-06/003. Office of Research and Development, US Environmental Protection Agency, Corvallis, Oregon.

Pont, D., R. M. Hughes, T. R. Whittier, and S. Schmutz. 2009. A predictive index of biotic integrity model for aquatic–vertebrate assemblages of western U.S. streams. Transactions of the American Fisheries Society 138:292–305.

Rader, R. B., M. J. Keleher, E. Billman, and R. Larsen. 2012. History, rather than contemporary processes, determines variation in macroinvertebrate diversity in artesian springs: the expansion hypothesis. Freshwater Biology 57:2475–2486.

Rehn, A. C. 2009. Benthic macroinvertebrates as indicators of biological condition below hydropower dams on West Slope Sierra Nevada streams, California, USA. River Research and Applications 25:208–228.

Rehn, A. C., P. R. Ode, and C. P. Hawkins. 2007. Comparison of targeted-riffle and reach-wide benthic macroinvertebrate samples: implications for data sharing in stream-condition assessments. Journal of the North American Benthological Society 26:332–348.

Rehn, A. C., P. R. Ode, and J. T. May. 2005. Development of a benthic index of biotic integrity (B-IBI) for wadeable streams in northern coastal California and its application to regional 305(b) assessment. Report to the State Water Resources Control Board. California Department of Fish, Rancho Cordova, California. (Available from: http://www.waterboards.ca.gov/water_issues/programs/swamp/docs/reports/assess_nocal2005.pdf)

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on **BE**nthic **A**ssessment of **S**edimen**T** (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20:198–219.

Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16:833–852.

Richards, A. B., and D. C. Rogers. 2011. List of freshwater macroinvertebrate taxa from California and adjacent states including standard taxonomic effort levels. Southwest Association of Freshwater Invertebrate Taxonomists, Chico, California. (Available from: www.safit.org)

Roth, N., M. Southerland, J. Chaillou, R. Klauda, P. Kayzak, S. Stranko, S. Weisberg, L. Hall, and R. Morgan. 1998. Maryland biological stream survey: development of a fish index of biotic integrity. Environmental Monitoring and Assessment 51:89–106.

Royer, T. V., C. T. Robinson, and G. W. Minshall. 2001. Development of macroinvertebrate-based index for bioassessment of Idaho rivers. Environmental Management 27:627–636.

Scherer, R. 2013. PropCIs: various confidence interval methods for proportions. R package, version 0.2-4. R Project for Statistical Computing, Vienna, Austria.

Schoolmaster, D. R., J. B. Grace, E. W. Schweiger, B. R. Mitchell, and G. R. Guntenspergen. 2013. A causal examination of the effects of confounding factors on multimetric indices. Ecological Indicators 29:411–419.

Sigovini, M., E. Keppel, and D. Tagliapietra. 2013. M-AMBI-revisited: looking inside a widely-used benthic index. Hydrobiologia 717:41–50.

Simpson, J. C., and R. H. Norris. 2000. Biological assessment of river quality: development of AusRivAS models and outputs. Pages 125–142 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse (editors). Assessing the biological quality of freshwaters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Sleeter, B. M., T. S. Wilson, C. E. Soulard, and J. Liu. 2011. Estimation of late twentieth century land-cover change in

California. Environmental Monitoring and Assessment 173: 251–266.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278.

Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier, and E. Tarquinio. 2008. A process for creating multimetric indices for large-scale aquatic surveys. Journal of the North American Benthological Society 27:878–891.

Stoddard, J. L., D. P. Larsen, C. P., Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267–1276.

Stribling, J. B., B. K. Jessup, and D. L. Feldman. 2008. Precision of benthic macroinvertebrate indicators of stream condition in Montana. Journal of the North American Benthological Society 27:58–67.

Strobl, C., A.-L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. Bias in random forest variable importance measures: illustrations, sources, and a solution. BMC Bioinformatics 8:25.

USEPA (US Environmental Protection Agency). 2002. Summary of biological assessment programs and biocriteria development for states, tribes, territories, and interstate commissions: streams and wadeable rivers. EPA-822-R-02-048. Office of Environmental Information and Office of Water, US Environmental Protection Agency, Washington, DC.

USEPA (US Environmental Protection Agency). 2010. Causal Analysis/Diagnosis Decision Information System (CADDIS). Office of Research and Development, US Environmental Protection Agency, Washington, DC. (Available from: http://www.epa.gov/caddis)

US Geological Survey. 2012. The StreamStats program. US Geological Survey, Reston, Virginia. (Available from: http://streamstats.usgs.gov)

Vander Laan, J. J., and C. P. Hawkins. 2014. Enhancing the performance and interpretation of freshwater biological indices: an application in arid zone streams. Ecological Indicators 36:470–482.

Van Sickle, J. 2010. Correlated metrics yield multimetric indices with inferior performance. Transactions of the American Fisheries Society 139:1802–1817.

Van Sickle, J., C. P. Hawkins, D. P. Larsen, and A. T. Herlihy. 2005. A null model for the expected macroinvertebrate assemblage in streams. Journal of the North American Benthological Society 24:178–191.

Van Sickle, J., D. P. Larsen, and C. P. Hawkins. 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. Journal of the North American Benthological Society 26:319–331.

Wright, J. F., D. W. Sutcliffe, and M. T. Furse. 2000. Assessing the biological quality of freshwaters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Yates, A. G., and R. C. Bailey. 2010. Selecting objectively defined reference stream sites for stream bioassessment programs. Environmental Monitoring and Assessment 170:129–140.

Yoder, C. O., and M. T. Barbour. 2009. Critical elements of state bioassessment programs: a process to evaluate program rigor and comparability. Environmental Monitoring and Assessment 150:31–42.

Yuan, L. L., C. P. Hawkins, and J. Van Sickle. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. Journal of the North American Benthological Society 27:892–905.

Appendix S1. Nearest-neighbor thresholds do not improve performance of predictive indices.

Variable impairment thresholds may be useful when the precision of an index varies greatly across settings (Death and Winterbourn 1994). For example, Yuan et al. (2008) observed 2-fold differences in variability at reference sites across ecoregions in an observed (O)/expected (E) taxa index for the USA, results that justified different thresholds for each region. In such circumstances, a uniform threshold may increase the frequency of errors in the more variable settings. Reference sites with scores below a uniform threshold may be disproportionately common in settings where the index is less precise. A variable threshold that is lower in more variable settings may reduce this error rate (i.e., the reference error rate).

To determine if variable impairment thresholds based on site-specific characteristics could lead to an unbiased distribution of errors across regions, we evaluated 2 approaches to establishing thresholds: 1) a traditional approach, where a single number (based on variability in scores at all reference calibration sites) was used as a threshold, and 2) a site-specific approach, where thresholds were based on only a subset of the most environmentally similar reference calibration sites. In both cases, we considered sites to be in reference condition if their index score was >10th percentile of the relevant set of reference calibration site values. We measured environmental similarity as standard Euclidean distances along all environmental gradients used in predictive models (Table 1). We evaluated several different sizes of reference-site subsets (25, 50, 75, 100, and 200, and the full set of 473). We calculated the error rate for all regions (except for the Central Valley, which had only 1 reference site) as the proportion of sites with scores below the threshold. We plotted these regional error rates against the number of reference sites used to calculate the threshold (Fig. S1) and transformed scores at test sites into percentiles relative to each of these distributions. We used the predictive California Stream Condition Index

(CSCI) and its null equivalent in this analysis.

Variable thresholds greatly reduced the regional bias of the error rate of the null index, but had a negligible effect on the predictive index. For example, the null index had a very high error rate (0.30) in the South Coast when a uniform threshold was used, but this error rate dropped to 0.10 when variable thresholds based on 25 or 50 reference sites were used. In contrast, the regional error rate of the predictive index was always <0.15 and was not highly influenced by the number of reference sites used to establish thresholds.

We recommend a uniform threshold used in conjunction with a predictive index because of the added complexity and minimal benefits provided by the variable, site-specific thresholds.

Literature cited

Death, R. G., and M. J. Winterbourn. 1994. Environmental stability and community persistence: a multivariate perspective. Journal of the North American Benthological Society 13:125–139.

Yuan, L. L., C. P. Hawkins, and J. Van Sickle. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. Journal of the North American Benthological Society 27:892–905.

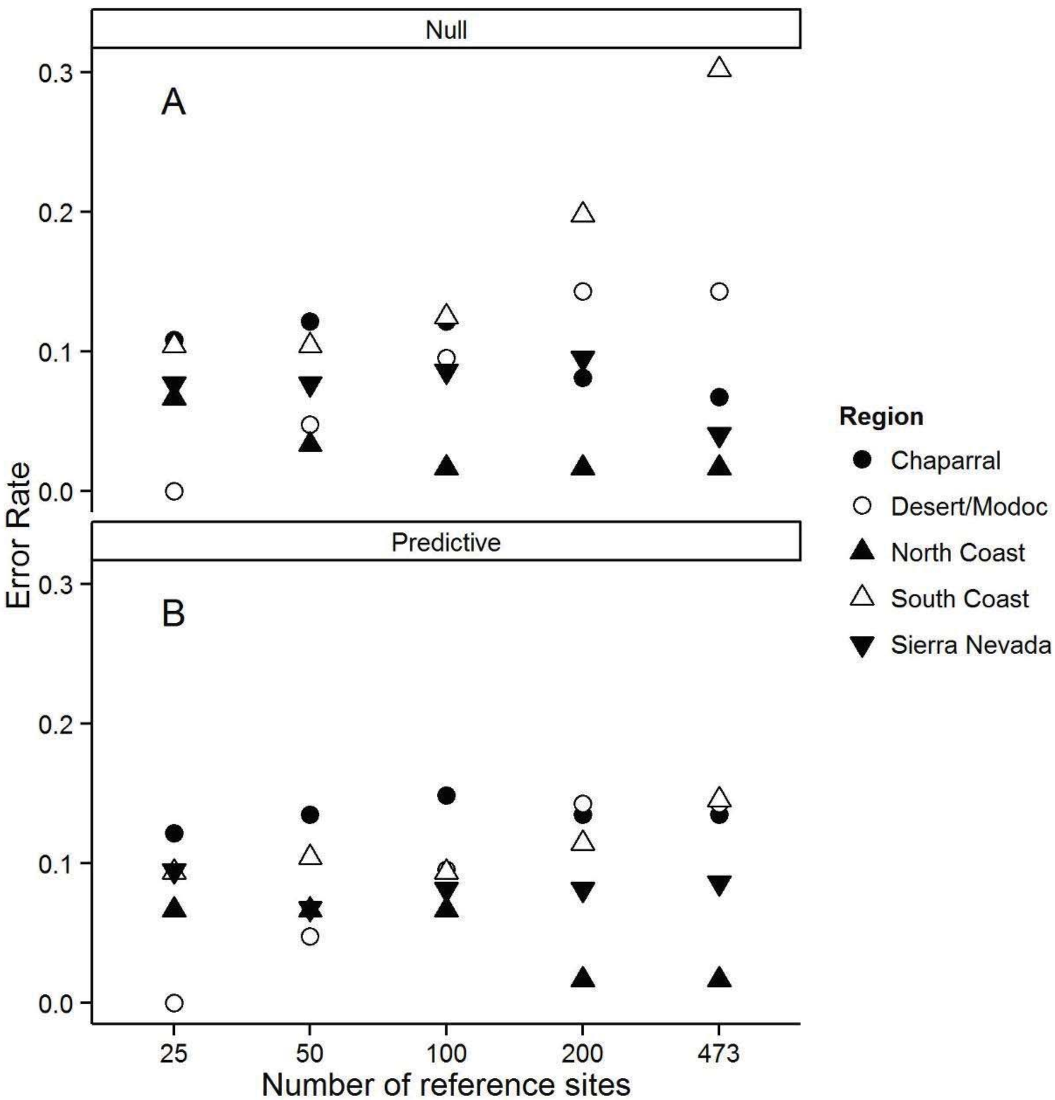


Fig. S1. Effects of nearest neighbor thresholds on error rates, calculated as the proportion of reference calibration sites below the threshold for null (A) and predictive (B) indices. Each point represents a different region. The highest number of reference sites is equivalent to the uniform threshold used in the main study.

Appendix S2. Index responsiveness as a function of predicted % sensitive taxa: a comparison of a predictive metric approach and the observed (O)/expected (E) taxa index.

The responsiveness of a bioassessment index depends on its ability to change in response to stress, and the loss of sensitive taxa is typically one of the strongest responses to stress (Rosenberg and Resh 1993, Statzner et al. 2004). To see if the ability to detect the loss of sensitive taxa depends on number of common taxa (E), we compared the proportion of sensitive taxa expected by an O/E index and a predictive multimetric index (pMMI) under different values of E. For the pMMI, this proportion was calculated as the predicted % intolerant taxa metric, as described in the accompanying manuscript. For the O/E, this proportion was calculated as the % of expected operational taxonomic units (OTUs) that are sensitive (OTUs with tolerance value < 3. For OTUs consisting of multiple taxa with different tolerance values, we used the median tolerance value). CAMLnet (2003) was the source of tolerance values. Estimates from both the O/E and pMMI were plotted against E to see whether the 2 indices allowed consistent ranges of response across values of E. These predictions were compared with the observed % intolerant taxa at reference sites to confirm the validity of these estimates.

At high-E sites (E > 14), both the pMMI and O/E had a consistent capacity to detect loss of sensitive taxa (Fig. S2A, C). Furthermore, both indices estimated similar proportions of sensitive taxa (~40%), suggesting that the 2 indices have similar sensitivity in these settings. Both indices also predicted a decline in the proportion of sensitive taxa at low-E sites, indicating that E affects the sensitivity of the pMMI and O/E. However, at the lowest levels of E, the O/E had no capacity to detect loss of sensitive taxa, whereas the pMMI predicted ~20% sensitive taxa at these sites, preserving a limited capacity to respond to loss of sensitive taxa. This capacity explains why the pMMI was more sensitive than the O/E at low-E sites.

Inspection of the data at reference sites indicates that sensitive taxa were truly present at these low-E sites (Fig. S2B, D) and that modeling the metric directly sets more accurate expectations for sensitive taxa in these settings (metric prediction vs observed $R^2 = 0.80$; O/E prediction vs observed $R^2 = 0.55$). However, these taxa were excluded from the index because of the minimum capture probability (i.e., 50%). Therefore, the predictive metric and not the O/E will be able respond to the loss of sensitive taxa at low-E sites.

Literature Cited

CAMLnet. 2003. List of California macroinvertebrate taxa and standard taxonomic effort. California Department of Fish and Game, Rancho Cordova, California. (Available from: www.safit.org)

Rosenberg, D. M., and V. H. Resh. 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Statzner, B., S. Dolédec, and B. Hugueny. 2004. Biological trait composition of European stream invertebrate communities: assessing the effects of various trait filter types. Ecography 27:470–788.

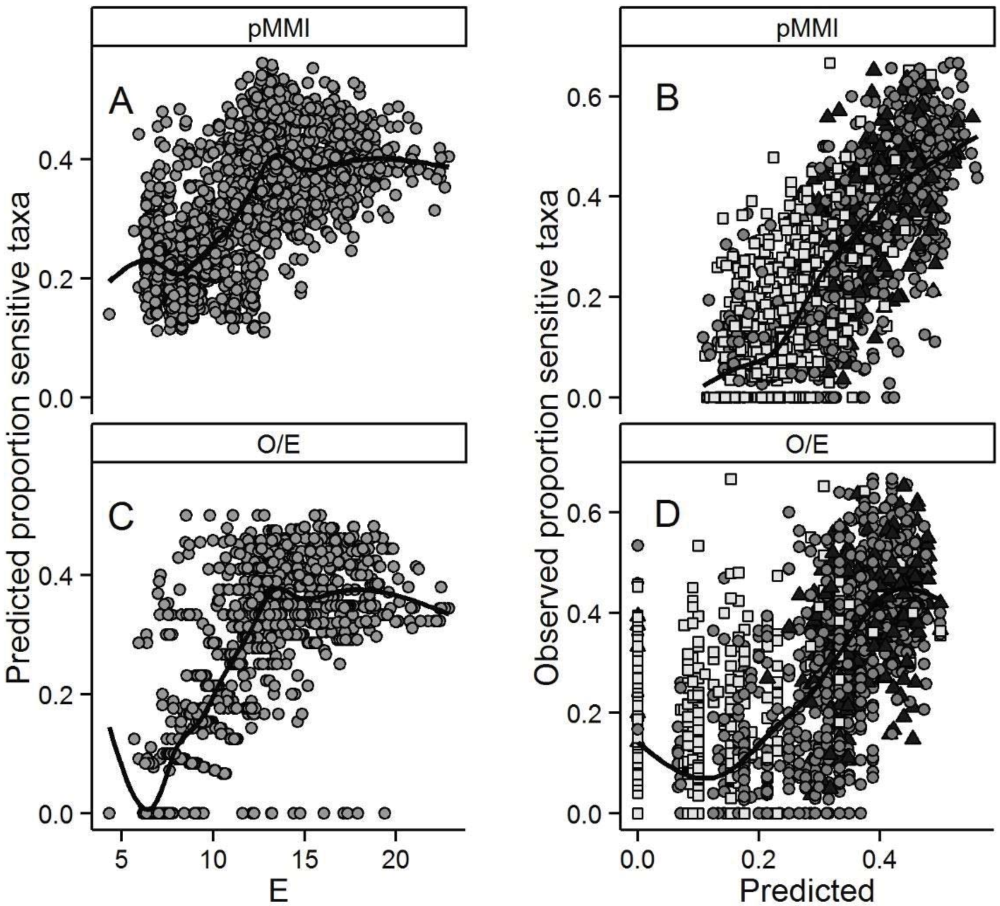


Fig. S2. Proportion of sensitive taxa predicted by a predictive multimetric index (pMMI) (A, B) and an observed (O)/expected (E) taxa index (C, D) at all sites (A, C), or observed at reference calibration (B, D) sites. Dark triangles represent sites with high (>15) numbers of expected taxa, gray circles represent sites with moderate (10–15) numbers of expected taxa, and white squares represent sites with low (<10) numbers of expected taxa. The solid line represents a smoothed fit from a generalized additive model.

Appendix 4

Comparability of biological assessments derived from predictive models and multimetric indices of increasing geographic scope

Peter R. Ode[1], Charles P. Hawkins[2] and Raphael D. Mazor

Abstract

As the use of bioassessment techniques expands, the demand for tools that can score biological condition from aquatic community data has spurred the creation of a large number of predictive models (e.g., observed over expected (O/E) indices) and multimetric indices (MMIs). The geographic and environmental scopes of these indices vary widely and coverages often overlap. If indices developed for large, environmentally heterogeneous regions provide results that are equivalent to those developed for smaller regions, then regulatory entities could adopt indices developed for larger regions rather than fund the development of multiple local indices. This potential was evaluated by comparing the performance (precision, bias, responsiveness, and sensitivity) of benthic macroinvertebrate O/E and MMIs developed for California (CA) with indices developed for two large-scale condition assessments of United States (US) streams: the US Environmental Protection Agency's Western Environmental Monitoring and Assessment Program's (WEMAP) stream project and the western portion of the national Wadeable Streams Assessment (WSA-West). Both WSA-West and WEMAP O/E scores were weakly correlated with CA O/E index scores, had lower precision than the CA index, were influenced by two related natural gradients (percent slope and percent fast water habitat) for which the CA index was not, and disagreed with 21 - 22% of impairment decisions derived from the CA index. The WSA-West O/E index produced many fewer impairment decisions than the CA index. In the MMI comparisons, both WEMAP and WSA-West MMI scores were much more strongly associated with CA MMI scores than those found in the O/E comparisons. However, the WSA-West and WEMAP MMIs produced many fewer impairment determinations than the CA MMI. Because the WEMAP and WSA-West indices were biased and differed in responsiveness compared with CA indices, they could produce different estimates of regional condition compared with indices that are calibrated to local conditions. Furthermore, the lower precision of the WEMAP and WSA-West indices compromises their use in site-specific assessments where both precision and accuracy are important. However, because the magnitude of differences in impairment decisions was very sensitive to the thresholds used to define impaired conditions, it may be possible to adjust for some of the systematic differences among the models, thus rendering the larger models more suitable for local application. Future work should focus on identifying the geographic and environmental scale that optimizes index performance, determining the factors that most strongly influence index performance, and identifying ways of more accurately specifying reference condition from geographically extensive sets of reference site data.

Introduction

The widespread adoption of bioassessment techniques for assessing the ecological condition of waterbodies has generated an abundance of indices available to water resource managers (Reynoldson *et al*. 1997, Hughes *et al*. 1998, Barbour and Yoder

[1] California Department of Fish and Game, Water Pollution Control Laboratory, Aquatic Bioassessment Laboratory, Rancho Cordova, CA
[2] Utah State University, Department of Watershed Sciences, Western Center for Monitoring and Assessment of Freshwater Ecosystems, Logan, UT

2000, Hawkins *et al* 2000a, Van Sickle *et al*. 2005, Bonada *et al*. 2006). Because these tools were generated to meet different needs, their geographic scopes differ widely and often overlap.

As the proliferation of new indices continues, end-users (e.g., regulatory entities developing numeric biocriteria, Yoder and Rankin 1995) will need guidance for selecting among these different indices and evaluating how many different indices a region needs for effective bioassessment. If local and regional assessments based on indices developed for broad geographical areas are equivalent to assessments based on indices developed for smaller areas, then regulatory entities could profit by adopting the large-scale indices and abandoning the development and maintenance of multiple, smaller-scale indices. This potential is attractive because indices that apply to large geographic areas have already been developed for many regions of the world, including: Great Britain (Moss *et al*. 1987), Australia (Simpson and Norris 2000), Europe (Statzner *et al*. 2001), and the United States (Stoddard *et al*. 2006, 2008; Yuan *et al*. 2008). Widespread use of common indices would facilitate consistency in data interpretation among the variety of users of ecological condition indices (Bonada *et al*. 2006, Hawkins 2006).

However, indices developed for large geographic regions may have limitations that could restrict their value for both site and regional assessments. Most notably, such indices must account for natural variation that occurs within large regions. Performance characteristics of both multimetric and predictive model indices are limited by their capacity to account for variability among the reference sites used to develop indices (Moss *et al*. 1987, Hughes 1995, Reynoldson *et al*. 1997, Karr and Chu 1999, Hawkins *et al*. 2000a, Bailey *et al*. 2004, Bonada *et al*. 2006).

It is a central principle of ecology that biological assemblages naturally vary along many environmental gradients (Andrewartha and Birch 1954, Hutchinson 1959, Hynes 1970). The precision and accuracy of any index will therefore depend on how well the mechanics of index calculation account for the effects of these natural gradients on assemblage structure (Johnson *et al*. 2004, Johnson *et al*. 2007, Van Sickle *et al*. 2005, Hawkins 2006, Heino *et al*. 2007, Mykrä *et al*. 2007, 2008). If biological variation associated with local environmental gradients (e.g., reach slope or substrate size) is masked by environmental factors that vary over large spatial scales (e.g., climatic factors and geology), then indices developed from more spatially restricted datasets may be required for site-specific assessments.

Recently derived biological indices developed for the EPA's national WSA and the WEMAP project (Stoddard *et al*. 2005, 2006; EPA 2006) presented an opportunity to evaluate this idea by comparing performance metrics (precision, bias, responsiveness, and sensitivity) of these indices with those of indices developed specifically for California (Ode *et al*. 2005, Rehn *et al*. 2005). The comparability of both site-specific and regionally aggregated biological assessments, where CA indices <WEMAP indices <WSA-West indices in geographic extent and geoclimatic heterogeneity, were evaluated. For these comparisons, assessments of an independent set of evaluation (test) sites that had not been used in developing any of the indices were conducted. To the extent that the test dataset permitted, parallel analyses for both MMI and O/E indices of benthic macroinvertebrate (BMI) assemblage condition were performed.

Methods

O/E Development

Three sets of predictive models were used to produce the O/E index values for comparison. All the O/E models were developed following a standardized process (Clarke *et al*. 2003, Hawkins *et al*. 2000a, Moss *et al*. 1987) described in the EMAP Western Streams and Rivers Statistical Summary (Stoddard *et al*. 2006). The process included: 1) sampling a set of environmentally diverse sites for BMIs, 2) specifying which of these sites would be used as reference sites, 3) applying a standard taxonomy (operational taxonomic units; OTUs) to all samples, 4) clustering of reference sites according to their similarity in BMI assemblage composition, 5) calculating and screening candidate predictor variables, and 6) calibrating linear discriminant functions models for predicting assemblage composition at new sites. All models were developed with map-level predictor variables (with the exception that field measured reach slope was used in one model) to allow more universal applicability of models (Table 1). Aside from the specific combination of predictor variables used in the models, the major difference among models was the range of environmental heterogeneity or geographic extent encompassed by the reference sites used in each model. Models were based on data from either targeted-riffle benthic samples (CA models) or a combination of targeted-

Table 1. Predictor variables used for all predictive models.

	California Models	WEMAP Models	WSA Model (no sub-models)
Sub-model 1	Watershed area Longitude Latitude Temperature	No predictors (null models)	Watershed area Longitude Day of year Minimum temperature Elevation Precipitation Slope
Sub-model 2	Longitude Preipitation Day of year Watershed area	Watershed area Longitude Elevation Precipitation	
Sub-model 3	Watershed area Temperature	No predictors (null models)	

riffle and reach-wide, multiple-habitat samples (WEMAP and WSA-West models). These two types of samples appear to be generally comparable for CA streams (Rehn *et al*. 2007). Other aspects of model development were similar (Table 2).

WSA-West model

A single western US model (WSA-West) developed during the national wadeable streams assessment (Yuan *et al*. 2008) encompassed the most heterogeneous environmental conditions and the largest geographic scope (~2,500,000 km^2; Figure 1). The WSA-West model was developed for all mountainous and xeric regions of the western United States and excluded only plains ecoregions (Figure 1; see Environmental Protection Agency 2006). To produce the WSA-West O/E index, 519 reference sites were clustered into 31 groups, and 7 predictor variables were selected to predict group membership (Table 1).

WEMAP models

The same data used to construct the WSA-West model had been previously used to develop five separate ecotype-specific submodels (Stoddard *et al*. 2006, 2008). All sampled sites (reference and non-reference) were assigned to one out of five broad ecotypes based on a k-means classification (MacQueen 1967) of long-term climatic (temperature and precipitation), geographic variables (latitude, longitude and elevation), and topographic variables (watershed area and channel slope). This pre-classification of sites was mainly designed to reduce the range of environmental heterogeneity encompassed by each model. The geographic scope of the resulting submodels ranged from ~200,000 km^2 to ~1,800,000 km^2 (Figure 2). Of the five submodels developed for the WEMAP study area (Stoddard *et al*. 2005, 2006), four submodels applied to geoclimatic conditions found in California. One model used predictor variables, whereas the other three were null models that predicted the same biota at all

Table 2. Comparison of BMI collection method, taxonomic effort levels and organism counts used both to build models and score test sites. See methods for definitions.

Indicator	Model	Field Method	Taxonomic Effort	Organism Count
O/E	WEMAP WSA 2 CA sub-models	RWB RWB TRB	Some species, but mostly genus (including Chironomidae)	300 (after removal of ambiguous individuals)
MMI	WEMAP	RWB	Some species, but mostly genus (including Chironomidae)	300
	WSA	RWB	Some species, but mostly genus (including Chironomidae)	300
	CA model (NCIBI / SCIBI)	TRB	Genus, Chironomidae to family	500

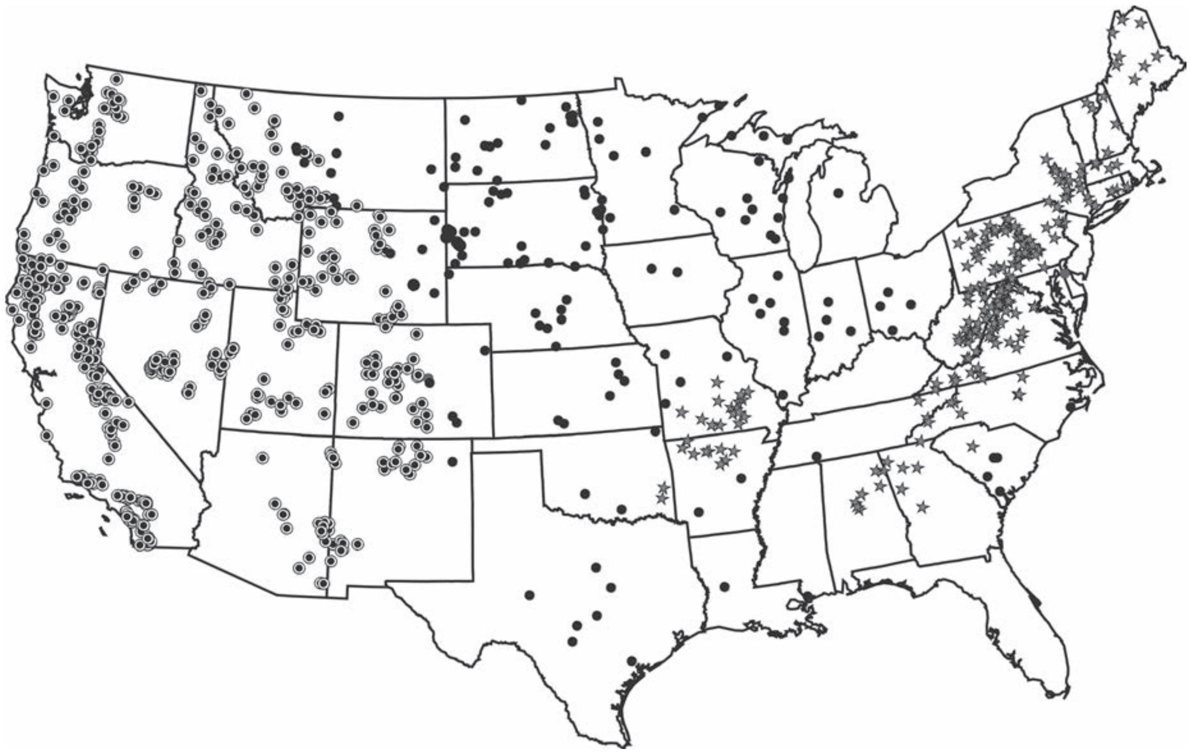


Figure 1. Location of reference sites used to create the three WSA predictive sub-models. Only the western sub-model applies to California sites. Each symbol represents a different sub-model.

sites within a geoclimatic region (Van Sickle *et al*. 2005; Table 1).

CA models

The third model set included three submodels that were developed for three types of climatic conditions in CA: cool-wet sites (mean monthly temperature (MMT) >9.9ºC and mean monthly precipitation (MMP) >895 mm), warm-dry sites (MMT >9.9ºC and MMP <895 mm), and cold-mesic sites (MMT <9.9ºC; Figure 3). The three CA submodels were calibrated from data collected at 209 reference sites, 179 of which were also used in calibrating WEMAP and WSA-West models (the other 30 sites were used as validation samples in the WEMAP and WSA projects). The spatial extent of the reference sites for these submodels was ~150,000 km^2 each (Figure 3). These three submodels also used unique combinations of predictor variables (Table 1).

MMI Development

The WSA, WEMAP, and CA MMIs were developed following similar methods as first developed by Karr (1981) and extended by others (Kerans and Karr 1994, Hughes *et al*. 1998, McCormick *et al*. 2001, Klemm *et al*. 2003): 1) assignment of a large pool of sites to either reference or test sets based on their degree of anthropogenic stress, 2) division of the site pool into calibration and validation sets, 3) using the calibration set to screen biological response metrics based on their responsiveness to important stressor gradients, their signal-to-noise ratios, and their non-redundancy with other metrics,

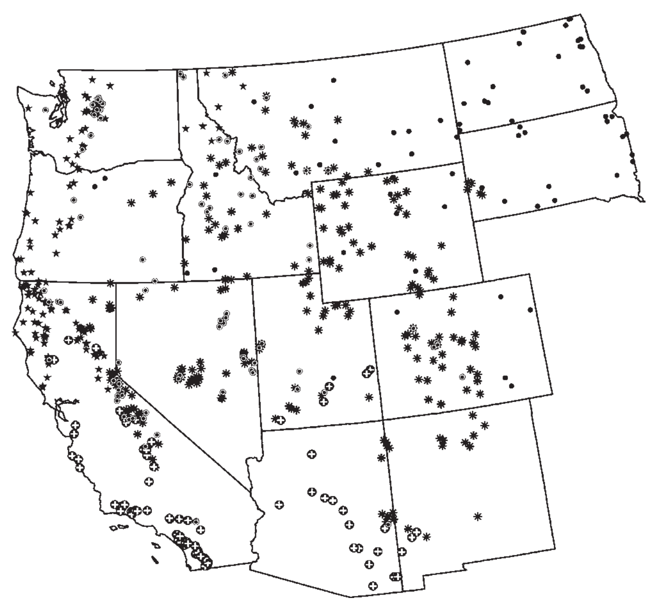


Figure 2. Location of reference sites used to create the five WEMAP predictive sub-models. Note that four of the five sub-models apply to California.

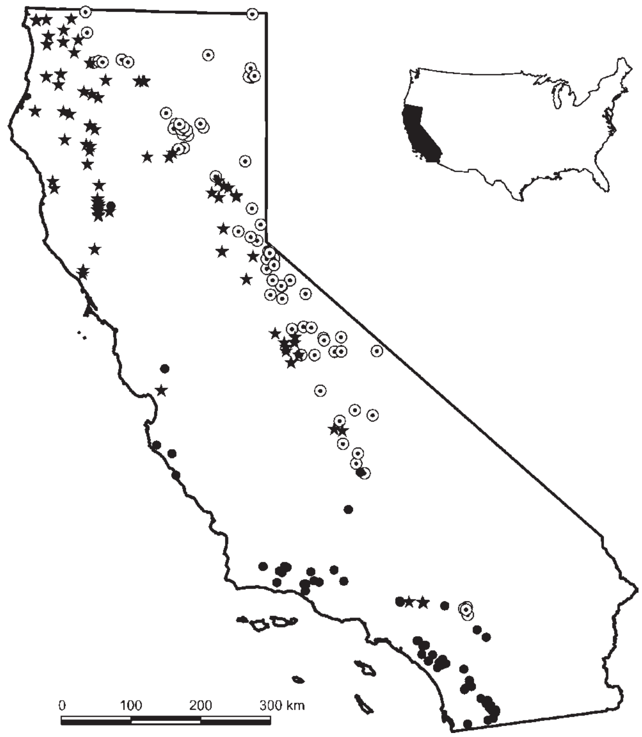


Figure 3. Location of reference sites used to create the three CA predictive sub-models. Each symbol represents a different sub-model.

4) establishing scoring ranges for selected metrics, 5) assembling a composite MMI from the component metrics, 6) establishing impairment thresholds for the index, and 7) evaluating index performance against the validation dataset (Herlihy *et al*. 2008, Stoddard *et al*. 2008).

These MMIs differed in a few important respects (Tables 2 and 3). The CA indices were based on subsamples of 500 organisms collected from targeted-riffle habitats (TRB) and identified primarily to genus level, but the WSA-West and WEMAP indices were based on subsamples of 300 organisms collected from multiple reach-wide composite habitats (RWB) with some individuals identified to species level (see text below for details on field and lab methods).

WSA-West MMIs

The EPA's WSA program developed two MMIs (xeric and western mountain ecoregions; Omernik 1987) to support its assessments of western streams using a calibration dataset of 775 sites (235 xeric and 540 mountain; Stoddard *et al*. 2008, EPA 2006). Both indices used six metrics, five of which were in common (Table 3). Scoring ranges for both WSA-West MMIs were scaled from 0 to 100 (Van Sickle and Paulsen 2008).

WEMAP MMIs

WEMAP developed three MMIs (xeric, plains and mountain ecoregions) for its analyses (Stoddard *et al*. 2005, 2006), two of which (xeric and mountain) applied to CA sites. The calibration dataset was comprised of 244 xeric and 565 mountain sites, nearly all of which (754 of 809) were used in WSA-West MMI development. As in the WSA-West index, the xeric and mountain versions of the WEMAP MMI consisted of six metrics, but shared fewer metrics in common (Table 3). Index values for both WEMAP MMIs were scaled from 0 to 100 (Stoddard *et al*. 2005).

CA MMIs

Two MMIs were developed for use in coastal California: the Southern Coastal California Index of Biotic Integrity or SCIBI (Ode *et al*. 2005) and the Northern Coastal California Index of Biotic Integrity or NCIBI (Rehn *et al*. 2005). The two CA MMIs included both the mountain and xeric aggregate ecoregions used for the WSA and WEMAP MMIs, and separate metric scoring ranges were established for the Omernik Level III (1987) ecoregions within each CA MMI development area (Figure 4). Of the 502 sites used to develop the CA MMIs, 119 were also used in WEMAP and WSA-West MMI development. The NCIBI consisted of eight metrics, whereas the SCIBI consisted of seven metrics, with four metrics in common (Table 3). The CA MMIs were also scaled from 0 to 100 (Ode *et al*. 2005, Rehn *et al*. 2005).

Test Site Data

These analyses incorporate BMI data collected for two large-scale probability surveys of CA streams. For clarity, use of the term "test sites" was restricted to refer only to these probabilistic samples of evaluation sites and not to non-reference sites used to calibrate MMIs, which are sometimes referred to as "test sites" in MMI development. For the O/E comparisons, data collected from 127 sites during the WEMAP 2000-2003 survey were used. For the MMI comparisons, data from 68 sites sampled by the California State Monitoring and Assessment Program (CMAP) between 2004 and 2006 were used. It was necessary to use different test sets for the O/E and MMI analyses because: 1) the restricted geographic boundaries of the CA MMI models limited the number of sites shared between O/E and MMI data sets, and 2) the MMI calibration datasets were

Table 3. BMI metrics comprising the multimetric indices. EPT = Ephemeroptera, Plecoptera, and Trichoptera.

Metric	California		WEMAP		WSA	
	NCIBI	SCIBI	Mountain	Xeric	Mountain	Xeric
EPT Richness	X	X	X	X	X	X
% Individuals in Top 5 Taxa			X		X	X
% Non-Insect Taxa	X	X		X		
Clinger % Taxa				X	X	X
% Intolerant Individuals	X	X				
% Non-Insect Individuals			X			X
% Tolerant Taxa		X	X			
Coleoptera Richness	X	X				
Scraper Richness					X	X
Tolerant % Taxa					X	X
% Burrower Individuals			X			
% Collector Individuals		X				
% EPT Taxa					X	
% Intolerant Taxa				X		
% Non-Gastropod Scraper Individuals	X					
% Omnivore Taxa			X			
% Predator Individuals	X					
% Shredder Taxa	X					
Diptera Richness	X					
Predator Richness		X				
Shannon Diversity				X		
Shredder Richness				X		

partially comprised of sites used for the O/E test set. The 127 sites used to evaluate predictive models were distributed throughout California (Figure 4a), whereas the 68 sites used to evaluate MMI models were restricted to coastal watersheds (Figure 4b). Most MMI test sites were concentrated in the northern half of the state (61 sites north of Monterey Bay), and the majority of these sites (40) were located within the boundaries of the NCIBI calibration sites (Figure 4b). The remaining 21 northern California sites were concentrated in the San Francisco Bay and Santa Cruz Mountains regions, which lie between the development regions of the two CA MMIs (Figure 4b). We used the NCIBI to score sites located between the NCIBI and SCIBI regions for the cross-index comparisons because this region is ecologically more similar to the North Coast than the South Coast and because reference conditions for this area were better represented in the NCIBI (Rehn *et al*. 2005). SCIBI scores were used for another 14 sites located within the region defined by the SCIBI calibration sites. Although the different geographic distributions in test sites may affect comparisons between MMIs and O/E indices, they do not affect comparisons of the performance of each type of index among the three geoclimatic scales.

Test site, field, and laboratory methods

All test sites were sampled in accordance with standard WEMAP field methods (Peck *et al*. 2006). A sampling reach was defined as 40 times the average stream width at the center of the reach, with a minimum reach length of 150 m. Two BMI samples were collected from each reach with standard 500-µm D-frame nets: 1) a RWB sample consisting of eleven 0.09-m^2 samples taken from equally spaced locations throughout the reach and 2) a TRB sample consisting of eight 0.09-m^2 samples taken from fast water habitat units within the reach (Hawkins *et al*. 2003).

All BMI samples used for the test datasets were processed at the California Department of Fish and Game's Aquatic Bioassessment Laboratory in Chico, CA. At least 500 individuals were identified to the standard taxonomic resolution targets described in Richards and Rogers (2006), i.e., those levels of taxonomic resolution that can be consistently achieved. A true, fixed 500-count random subsample was then

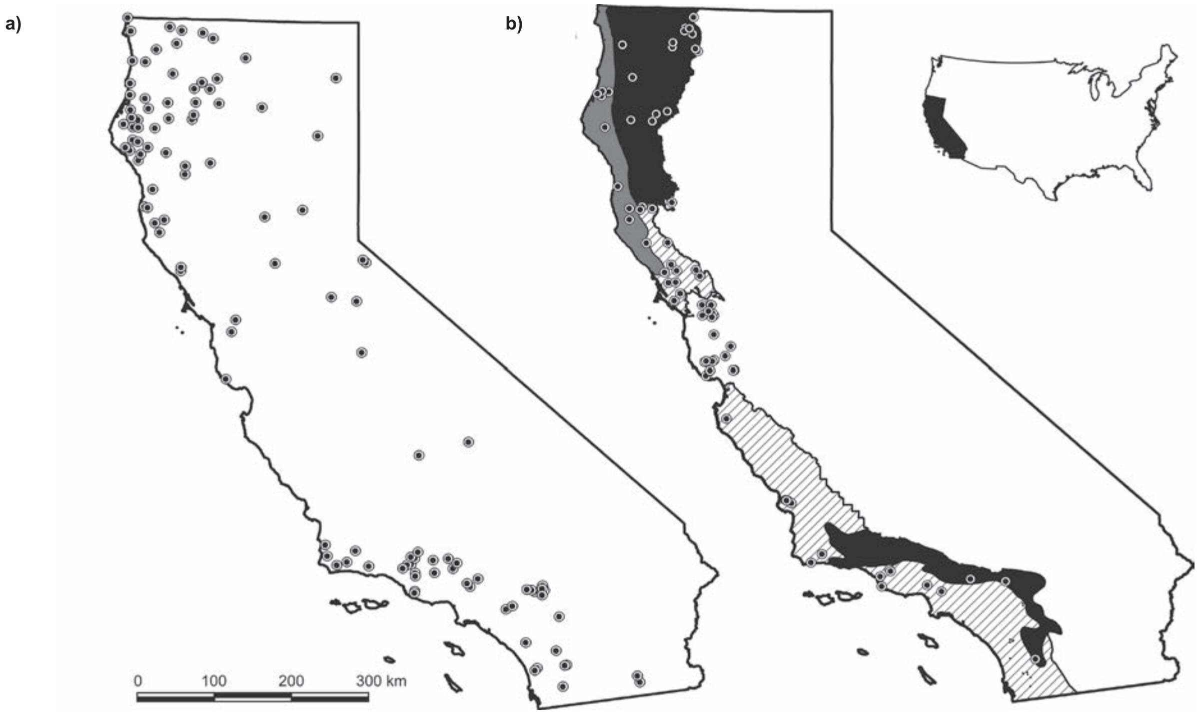


Figure 4. Location of test sites used for comparative analyses: 127 test sites used in O/E comparisons (a) and 68 test sites used in IBI comparisons (b). The three solid shaded regions correspond to mountain ecoregions used in the western and national models, whereas the three hatched regions correspond to the xeric ecoregions used in the western and national models. The two solid shaded regions in the northwest part of the state circumscribe the region for which the North Coast IBI was developed. The hatched regions and the continuously shaded region in southwestern part of the state circumscribe the region for which the South Coast IBI was developed. The inset shows the location of California in the United States.

obtained by computer resampling the sample data. Samples with between 450 and 500 individuals were retained in analyses. These raw data were then used to produce the standardized taxa lists and metrics needed for the various indices (Table 3). All analyses were based on the field methods, sample sizes and taxonomic levels used to develop each index (as indicated in Table 2).

Scoring Sites: Predictive Models

BMI taxonomic data

The raw subsample count data were further processed for use with the predictive models by: 1) converting the original identifications to the taxonomic levels used in the models (e.g., OTUs), 2) eliminating individuals that could not be assigned to an OTU (i.e., ambiguous individuals), and 3) resampling the remaining non-ambiguous individuals to 300-count samples. Samples with <300 individuals were retained in analyses.

Predictor variables

Geographic coordinates (latitude and longitude) were obtained via GPS measurements taken during sample collection. Watershed area were calculated after delineating upstream watershed boundaries for each site with automated GIS scripts or manual delineation where necessary. Long-term MMP, MMT, and MMA values for each site were estimated from GIS grids of (1961-1990) obtained from the Oregon Climate Center (http://www.ocs.orst.edu/prism). Site elevations were derived from 30-meter digital elevation models (http://ned.usgs.gov). Channel (reach) slope was measured in the field (as it was in model development).

Geographic and environmental attributes were used to assign each site to the appropriate WEMAP and CA models. Assignment of sites to the five WEMAP models was based on latitude, longitude, elevation, MMP, MMT, watershed area, and channel slope. These assignments were made prior to model building during the k-means analysis (MacQueen 1967). Assignment of test sites to the appropriate CA model was conducted after model development. This study used a simple classification and regression tree model based on long-term precipitation and air temperature to assign sites to the CA submodels.

O/E values were calculated based on just those taxa with site probabilities of capture ≥0.5 because these values result in more precise O/E values that are also usually more sensitive to stress (Hawkins *et al*. 2000a, Ostermiller and Hawkins 2004, Van Sickle *et al*. 2007) than O/E values based on all taxa in the reference calibration data set. When reporting impairment decisions for the test sites, impairment thresholds were set at two standard deviations below the mean value of reference sites for all O/E models (Table 4).

Scoring Sites: MMIs

BMI taxonomic data

Because the MMIs differed with respect to organism count and taxonomic resolution, MMI scores were calculated based on the sample counts and taxonomy used when developing each index (Table 2). MMI values were then calculated for test samples that had been collected in a standard manner to avoid confounding comparisons with inter-method variability. All sites were assigned to either the xeric or mountain aggregate ecoregions, with mountain ecoregions being further divided into Southern California Mountains, Klamath Mountains, Coast Ranges, and Southern and Central California Chaparral and Oak Woodlands for the CA MMIs (Omernik 1987). MMI values were then calculated based on the specific scoring ranges developed for each individual metric and region and rescaled these MMI values from 0 to 100. As for O/E models, impairment thresholds for all MMIs were set at two standard deviations below the mean value at reference sites (Table 4).

TRB was used as the default sample type, although RWB samples were used at six sites where TRB samples were unavailable or had low sample counts (<450 organisms). Because it was found elsewhere that RWB samples on average scored 7.8 points lower on the CA IBIs than TRB samples (Rehn *et al*. 2007), 7.8 points were added to CA IBI scores for these RWB samples. To evaluate the potential effect of using TRB samples instead of RWB samples (the method used in national and western model development; Table 2) in comparisons, an additional analysis was performed in which both the TRB and RWB data from 21 sites with all three MMIs were scored. If paired t-tests indicated significant differences between methods, RWB scores were adjusted by a correction factor corresponding to the difference between mean site scores.

Comparison of Index Scores

The CA index values were used as a benchmark for comparing the performance of the WSA-West and WEMAP indices. Comparisons were based on index precision, bias, responsiveness, and sensitivity.

O/E comparisons

Precision was measured as the standard deviation (sd) of reference site O/E values. Bias was measured as the tendency for reference site O/E values to vary systematically with one or more of four

Table 4. Standard deviation (sd) values and impairment thresholds (IT) for each predictive model (O/E) and coefficients of variation (CV) and impairment thresholds (IT) for MMIs. Note that only WEMAP sub-models 2 through 5 apply to California.

California O/E			WEMAP O/E			WSA O/E		
Sub-model	sd	IT	Sub-model	sd	IT	Model	sd	IT
1	0.13	0.74	1	0.24	0.52	West	0.20	0.59
2	0.17	0.66	2	0.15	0.70			
3	0.16	0.68	3	0.20	0.60			
			4	0.20	0.60			
			5	0.17	0.66			

California MMI			WEMAP MMI			WSA MMI		
Sub-model	CV	IT	Sub-model	CV	IT	Sub-model	CV	IT
NCIBI	0.14	52	Mountain	0.13	55	Mountain	0.26	28
SCIBI	0.19	39	Xeric	0.23	36	Xeric	0.25	34

natural gradients (reach slope, elevation, watershed area, and percent of reach with fast water habitats). The study also assessed relative bias between pairs of O/E indices using linear regression; slopes were tested for significant differences from 1, and intercepts were were tested for significant differences from 0. The consequences of these types of biases were illustrated by plotting the pair-wise differences in index scores against these natural gradients. Responsiveness was measured as the mean difference between reference and test sites in O/E values. Sensitivity was measured as the proportion of test sites assessed as impaired by the models. This measure of sensitivity is a joint function of precision, bias, and responsiveness. For these assessments, the threshold values for inferring impairment were defined as 2 sds below the reference (calibration) means (Table 4). Binomial tests (Zar 1999) were used on sites with disagreeing impairment decisions to determine if the indices were equally likely to detect impairment. This test was performed within each of the three CA submodels, as well as on all sites combined. In addition to comparison of impairment determinations based on 2 sds thresholds, two different threshold corrections for ecoregional differences were also evaluated. In the WSA, impairment thresholds were established separately for xeric and mountain ecoregions at the 5th percentile of the calibration reference population (estimated as 1.64 standard deviations below the reference mean; Herlihy *et al*. 2008). We also estimated separate thresholds for mountain and xeric regions at 2 sd below the mean for each ecoregion, an approach consistent with previous comparisons. For all relevant analyses, Bonferroni adjustments were applied for multiple comparisons when the correction was conservative. That is, the correction was not applied when we were screening natural gradients as potential drivers of bias, but was applied for hypothesis tests of index agreement (e.g., impairment decisions, responsiveness tests).

Multimetric index comparisons

MMI analyses paralleled the O/E comparisons. However, raw MMI scores were not directly comparable because the scores at calibration reference sites differed among the MMIs. Therefore, MMI scores were rescaled by dividing the raw score by the index's reference mean. These adjusted scores were then used as a "common currency" in all analyses in which scores were compared directly. Thus, the MMI scaling in these analyses was similar to the ~1.0 reference mean in O/Es. Only the comparisons of impairment decisions were based directly on the raw MMI scores.

Results

O/E Comparisons

Precision

The predictions of the WSA-West and WEMAP models were less precise (reference site O/E sd = 0.17 to 0.20) than those of the CA models (sd = 0.13 to 0.17; Table 4). Imprecision in model predictions contributed, in part, to weak relationships between the CA O/E indices and the WSA-West and WEMAP O/E indices (CA vs. WSA-West $r^2 = 0.32$, CA vs. WEMAP $r^2 = 0.35$; Figure 5). However, the stronger agreement between the less precise WSA-West and WEMAP O/E indices (WSA-West vs. WEMAP $r^2 = 0.58$) indicates that factors other than precision (e.g., bias) must also be affecting differences in agreement (Figure 5).

Bias

The WSA-West and WEMAP O/E values were biased predictors of the CA O/E values and each other, with slopes and y-intercepts significantly different ($p < 0.001$) than 1 and 0, respectively, for all comparisons (Figure 5). Differences were large, with slopes as low as 0.58 and intercepts as high as 0.36. These results showed that the nature of the bias was not constant across all sites. Instead, differences in index scores depended on the site-specific differences among models in how they either over- or under-estimated E (the expected number of predicted taxa) relative to one another. The reason that the O/E indices were biased predictors of one another occurred, at least in part, because the WSA-West and WEMAP models failed to adjust predictions of E for the effects of at least one natural gradient. This failure is illustrated by systematic variation in reference site O/E values produced by the WSA-West and WEMAP models across percent slope (WSA-West score = 0.025% slope + 0.80, p = 0.001; WEMAP score = 0.023% slope + 0.67, p = 0.002) and percent fast water habitat gradients (WSA-West score = 0.0051% fast water + 0.747, p <0.001; WEMAP score = 0.0045% fast water + 0.63, p <0.001). No such relationships were evident for CA O/E values (CA score = 0.0086% slope + 0.78, p = 0.259; CA score = 0.0016% fast water + 0.77, p = 0.205). The reason the CA O/E indices were unrelated to reach slope is probably related to the fact that, within CA, channel

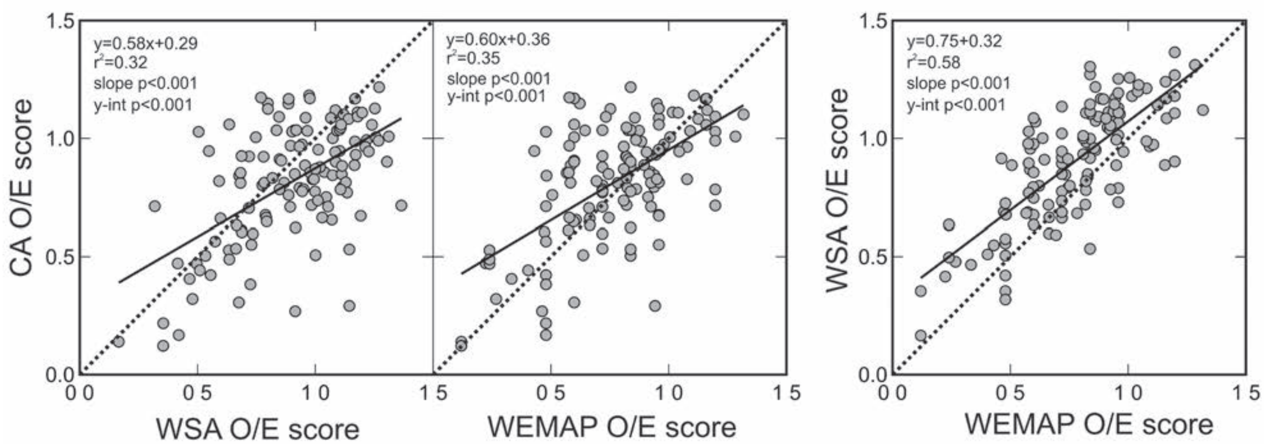


Figure 5. Regressions between O/E scores at CA test sites for all combinations of models. The dotted diagonal lines represent perfect 1:1 relationship between the models, and the thick solid lines indicate linear best-fit relationships. Significance tests are for y-intercept (y-int) = 0 and slope = 1.

slope was associated with watershed area (Area), a predictor in all three CA models (square root slope = 4.11 -0.531*log Area – 0.040*latitude across all reference sites, n = 209, $r^2 = 0.14$, model $p < 0.001$). It is therefore possible that watershed area was a surrogate predictor of reach slope within CA. Percent fast water was measured at too few sites to determine its relationship with watershed area within CA. As a consequence of the bias between the WSA-West and WEMAP model predictions, pair-wise differences between O/E values for both the WSA-West and WEMAP indices and the CA indices were significantly related to channel slope and percent fast-water habitat (Figure 6). Similar biased predictions associated with either elevation or watershed area were not observed, nor were any of these relationships observed for pair-wise differences in values between WSA-West and WEMAP (Figure 6; Table 5). Furthermore, correlation coefficients were low for all of these relationships (Table 5), indicating that very little variance in differences between the indices was explained by these natural gradients. Although not related to the four natural gradients we examined, there was a tendency for the WSA-West model to produce higher O/E scores than the WEMAP submodels, especially at lower O/E values ($p < 0.005$; Table 5; Figures 5 and 6).

Responsiveness

The WEMAP models tended to produce the lowest O/E values and the WSA-West models the highest O/E values at the test sites (Table 6). O/E values based on the CA models tended to be intermediate in magnitude. This pattern generally occurred for both mountain and xeric ecoregions, although differences were not always statistically significant. However, the magnitude of difference in mean test site O/E values between mountain and xeric test sites varied with the models used. The CA models resulted in lower average O/E values for xeric than for mountain sites (Table 6), whereas both the WEMAP and WSA-West models produced statistically similar mean O/E values at xeric and mountain test sites.

Index sensitivity and concordance among assessments

The WSA-West O/E was much less likely to lead to inferences of impairment (16 of 127 sites; Table 7) than either the WEMAP O/E (43 of 127 sites) or the CA O/E (35 of 127 sites, binomial tests, $p < 0.001$). When an ecoregion correction based on 2 sds (consistent with primary analyses) was applied, there was no effect on any impairment decision (16 out of 127 sites impaired) because the separate xeric and mountain thresholds were within 2 points on a 100 point scale of their combined threshold. However, when an ecoregion correction based on the 5th percentile threshold used for the national wadeable streams assessment (Herlihy *et al*. 2008) was applied, the number of sites determined to be impaired by the WSA-West index (27 of 127 sites) was not significantly different from the 35 impairment decisions produced by the CA O/E index (binomial test, $p = 0.081$; Table 7).

Multimetric Index Comparisons: Comparison of TRB vs. RWB for WSA and WEMAP MMIs

MMI scores derived from the TRB and RWB sampling methods were highly correlated for both

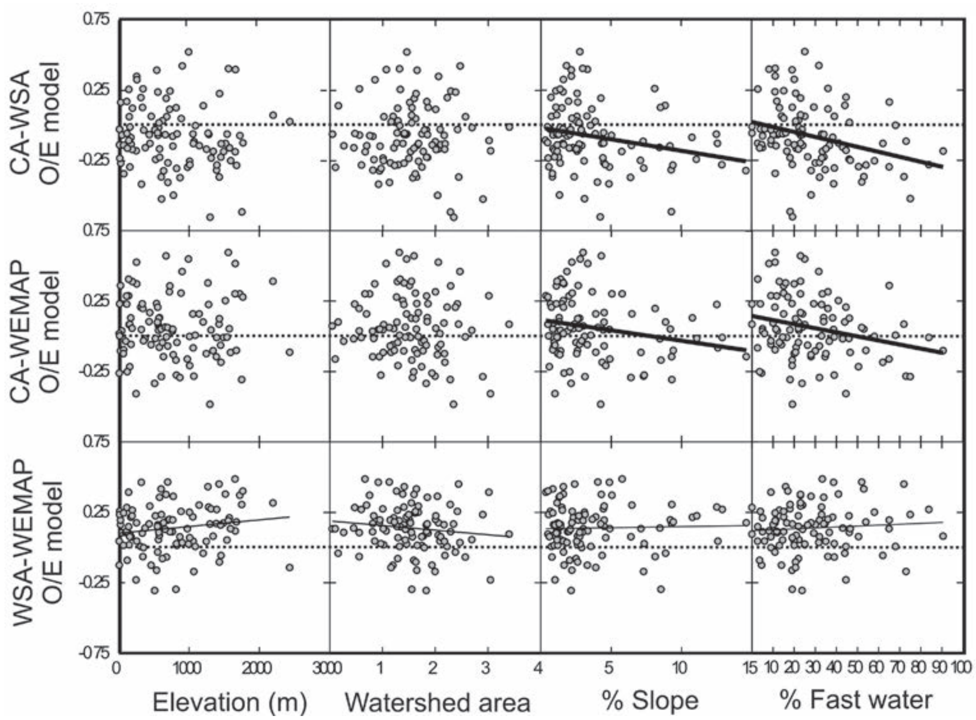


Figure 6. Scatterplots and regressions between the pair-wise differences in O/E values for three different O/E indices and four environmental gradients at CA test sites. The dashed horizontal lines represent zero difference. Thick solid lines denote regressions with r^2 and slopes significantly different from 0; thin solid lines denote those with intercepts significantly different from 0 but non-significant slope.

WSA and WEMAP indices (WSA $r^2 = 0.75$, WEMAP $r^2 = 0.73$), as has been shown elsewhere for CA MMIs (Rehn *et al*. 2007). For WEMAP MMIs, RWB samples collected in the mountain ecoregion scored on average 7.2 points lower than TRB samples (paired t test, $p < 0.001$), but samples based on the two methods collected in the xeric ecoregion produced statistically indistinguishable scores ($p = 0.65$). Mountain WSA MMI values were also lower for RWB samples (4 points, $p = 0.02$), but RWB MMI values from the xeric region were higher than TRB samples by 6 points ($p = 0.002$). For the purpose of inter-index comparisons, these scoring biases were corrected by adding 7.2 points to the MMI values for those mountain WEMAP RWB samples used in comparisons (three sites), adding 4 points to values for the mountain WSA RWB samples, and subtracting 6 points from values for xeric WSA RWB samples (three sites). However, only TRB samples were used in remaining MMI analyses.

Precision

The northern and southern CA MMIs were more precise (reference site CVs = 0.14 and 0.19) than the WSA-West mountain and xeric MMIs (CVs = 0.26, 0.25), but comparable to those of the WEMAP mountain and xeric MMIs (CVs = 0.13, 0.23; Table 4). Associations among the rescaled MMI indices (CA vs. WSA-West $r^2 = 0.70$, CA vs. WEMAP $r^2 = 0.76$, and WSA-West vs. WEMAP $r^2 = 0.75$; Figure 7) were much stronger than we observed for the O/E indices (Figure 5).

Bias

The rescaled WSA-West MMI was a biased predictor of both the CA and WEMAP MMIs, with slopes significantly different ($p < 0.001$) from 1 (Figure 7). In addition, the WEMAP MMI on average produced higher scores at test sites than the CA MMI (Table 6). The WEMAP MMI rated low-scoring sites higher than the WSA-West MMI and high-scoring sites lower than the WSA-West MMI (Figure 7). However, most of these differences in MMI values were not associated with the natural gradients we considered, except for the significant relationships between CA and WEMAP pairwise differences and both elevation and watershed area (Figure 8).

Responsiveness

On average, the rescaled CA MMIs scored test sites lower than the rescaled WEMAP MMIs, which

Table 5. Regressions (y = a + bx) for relationships shown in Figures 6 and 10 where y is the difference between the index scores of two models and x is a natural gradient variable. Asterisks indicate significant slopes, y-intercepts, or r^2 values at p = 0.05 level (significance threshold not adjusted for multiple comparisons).

Index	Natural Gradient	Model Difference	b	p-value for b	a	p-value for a	r^2
O/E (n = 101)	Elevation	CA-WSA	-0.000043	0.283	-0.043	0.259	0.01
		CA-WEMAP	0.0000042	0.918	0.059	0.132	0
		WSA-WEMAP	0.000048	0.112	0.1	<0.001*	0.03
	log Watershed Area	CA-WSA	0.0029	0.928	-0.081	0.125	0
		CA-WEMAP	-0.025	0.424	0.1	0.06	0.01
		WSA-WEMAP	-0.028	0.23	0.18	<0.001*	0.01
	Percent Slope	CA-WSA	-0.016	0.019*	-0.017	0.606	0.05*
		CA-WEMAP	-0.015	0.035*	0.12	<0.001*	0.04*
		WSA-WEMAP	0.0015	0.77	0.13	<0.001*	0
	Percent Fastwater	CA-WSA	-0.0035	0.002*	0.023	0.543	0.09*
		CA-WEMAP	-0.0029	0.012*	0.14	0.001*	0.06*
		WSA-WEMAP	0.00064	0.458	0.12	<0.001*	0.01
MMI-rescaled (n = 68)	Elevation	CA-WSA	0.000047	0.586	-0.24	<0.001*	0
		CA-WEMAP	0.00012	0.041	-0.15	<0.001*	0.06
		WSA-WEMAP	0.000073	0.415	0.086	0.028*	0.01
	log Watershed Area	CA-WSA	-0.043	0.19	-0.13	0.105	0.03
		CA-WEMAP	-0.057	0.01	0.011	0.832	0.1
		WSA-WEMAP	-0.014	0.674	0.14	0.095	0
	Percent Slope	CA-WSA	0.0024	0.832	-0.23	<0.001*	0
		CA-WEMAP	0.011	0.151	-0.14	<0.001*	0.03
		WSA-WEMAP	0.0085	0.46	0.09	0.020*	0.01
	Percent Fastwater	CA-WSA	0.0021	0.182	-0.28	<0.001*	0.03
		CA-WEMAP	-0.00071	0.518	-0.1	0.004*	0.01
		WSA-WEMAP	-0.0028	0.086	0.18	<0.001*	0.04

in turn scored test sites lower than rescaled WSA-West MMIs (Table 6). This trend generally held for both mountain and xeric ecoregions, although the WSA-West vs. WEMAP mountain contrast was not significantly different. All MMIs tended to score test sites in the xeric ecoregion lower than test sites in the mountain ecoregion, although the difference in mean values based on the WSA-West MMI was not significant (Table 6).

Index sensitivity and concordance among assessments

As with the O/E indices, impairment decisions differed considerably among the rescaled MMI indices (Table 8). The number of sites assessed as impaired was far fewer for the WSA-West and WEMAP MMIs (21 and 17 sites of 68 total sites, respectively) than the CA MMI (39 of 68 sites, binomial tests, p <0.001). This pattern occurred in both xeric and mountain ecoregions but was only significant in the xeric ecoregions (binomial tests: mountain p = 0.219, xeric p <0.001).

Summary of WEMAP and WSA-WEST indices performance relative to CA indices

Differences in index precision, bias, and responsiveness can each contribute to differences in index performance as measured by index sensitivity, the likelihood that an assessment will identify impairment. In this study, assessment differences between WEMAP or WSA-West indices and CA indices depended on the type of index examined and specific differences in index precision, bias, and responsiveness (Table 9). Although the large-scale indices tended to lead to different inferences regarding biological condition than the CA indices, the specific differences

Table 6. Results of t-test comparisons for differences in index responsiveness between sets of mountainous and xeric test sites, or between model pairs. Mean 1 and Mean 2 indicate the mean scores of the first and second members of each tested pair. All MMI scores were rescaled by dividing scores by the appropriate reference mean.

Test Dataset	Comparison Type	Test group	Indices in Test	Mean 1	Mean 2	Difference	p (*significant α =0.0167)	Test (2-tailed)
O/E	*Index Comparison*	*Both Ecoregions (n = 127)*	CA vs. WSA	0.82	0.90	0.09	<0.001*	paired t-test
			CA vs. WEMAP	0.82	0.77	0.04	0.032	
			WSA vs. WEMAP	0.90	0.77	0.13	<0.001*	
		MTN only (n = 74)	CA vs. WSA	0.87	0.93	0.06	0.023	paired t-test
			CA vs. WEMAP	0.87	0.80	0.07	0.002*	
			WSA vs. WEMAP	0.93	0.80	0.13	<0.001*	
		XER only (n = 53)	CA vs. WSA	0.75	0.87	0.12	0.005*	paired t-test
			CA vs. WEMAP	0.75	0.74	0.00	0.938	
			WSA vs. WEMAP	0.87	0.74	0.12	<0.001*	
	Ecoregion Comparison	*MTN vs. XER*	CA	0.87	0.75	0.12	0.006*	2 sample t-test
			WSA	0.93	0.87	0.06	0.156	
			WEMAP	0.80	0.74	0.05	0.248	
MMI	*Index Comparison*	*Both Ecoregions (n = 68)*	CA vs. WSA	0.65	0.88	0.23	<0.001*	paired t-test
			CA vs. WEMAP	0.65	0.77	0.12	<0.001*	
			WSA vs. WEMAP	0.88	0.77	0.11	<0.001*	
		MTN only (n = 30)	CA vs. WSA	0.80	1.00	0.20	<0.001*	paired t-test
			CA vs. WEMAP	0.80	0.88	0.07	0.009*	
			WSA vs. WEMAP	1.00	0.88	0.13	0.018	
		XER only (n = 38)	CA vs. WSA	0.53	0.78	0.24	<0.001*	paired t-test
			CA vs. WEMAP	0.53	0.69	0.15	<0.001*	
			WSA vs. WEMAP	0.78	0.69	0.09	0.006*	
	Ecoregion Comparison	*MTN vs. XER*	CA	0.80	0.53	0.27	<0.001*	2 sample t-test
			WSA	1.00	0.78	0.23	0.0219	
			WEMAP	0.88	0.69	0.19	0.001*	

Table 7. Counts of CA sites declared impaired (I) and not impaired (NI) by CA O/E estimates and corresponding WEMAP and WSA O/E estimates. WSA-Adjusted: Impairment thresholds set at 5th percentile for each ecoregion.

		CA Sub-model 1 (n = 58)		CA Sub-model 2 (n = 44)		CA Sub-model 3 (n = 25)		Total (n = 127)		All Sites
		I	NI	I	NI	I	NI	I	NI	
CA	I	13	-	16	-	6	-	35	-	35
	NI	-	45	-	28	-	19	-	92	92
WEMAP	I	10	7	11	8	4	3	25	18	43
	NI	3	38	5	20	2	16	10	74	84
WSA	I	5	1	7	2	0	1	12	4	16
	NI	8	44	9	26	6	18	23	88	111
WSA-Adjusted	I	9	4	9	4	0	1	18	9	27
	N	4	41	7	24	6	18	17	83	100

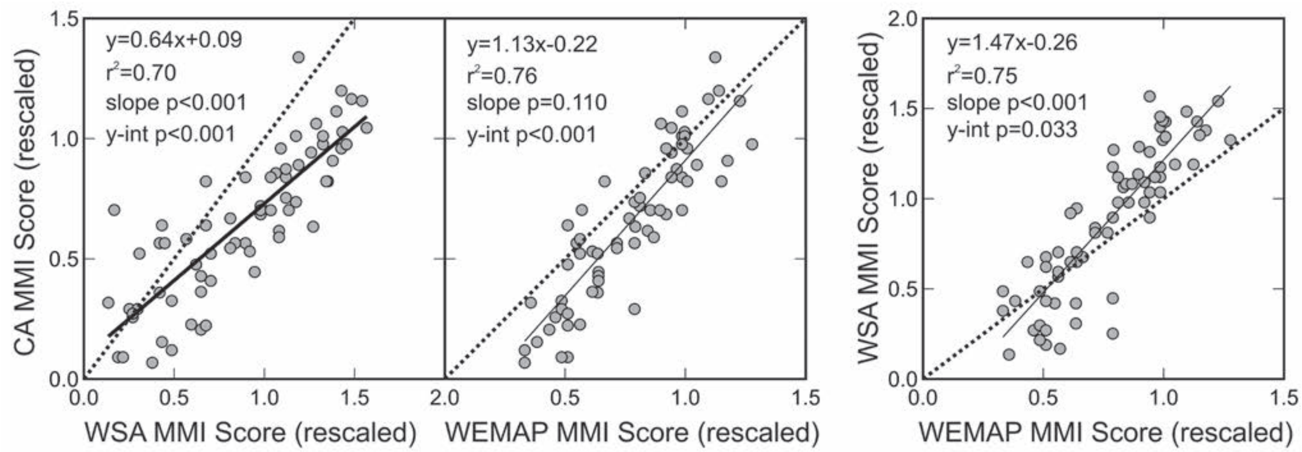


Figure 7. Regressions between rescaled scores at CA test sites between rescaled index scores for different combinations of the MMIs. The dashed diagonal lines represent perfect 1:1 relationship between the models, and the thick and thin solid lines indicate linear best-fit relationships. Significance tests are for y-intercept (y-int) = 0 and slope = 1.

among indices were variable. These differences lead to the WEMAP O/E index having similar sensitivity to the CA O/E indices, whereas the WSA-West O/E index was less sensitive. The difference between these two large-scale indices appeared to be largely associated with differences in their responsiveness. The MMI comparisons showed the opposite response in that the WEMAP MMI was slightly more sensitive than the CA MMI in mountain regions while the WSA-West MMI was less sensitive than the CA MMI in xeric regions. As we saw for the O/E comparisons, the differences between the WEMAP and WSA-West MMI sensitivities were also most clearly associated with differences in their responsiveness.

Discussion

The multiple spatial scales over which environmental gradients influence the taxonomic and functional composition of freshwater assemblages has been the focus of considerable interest in recent

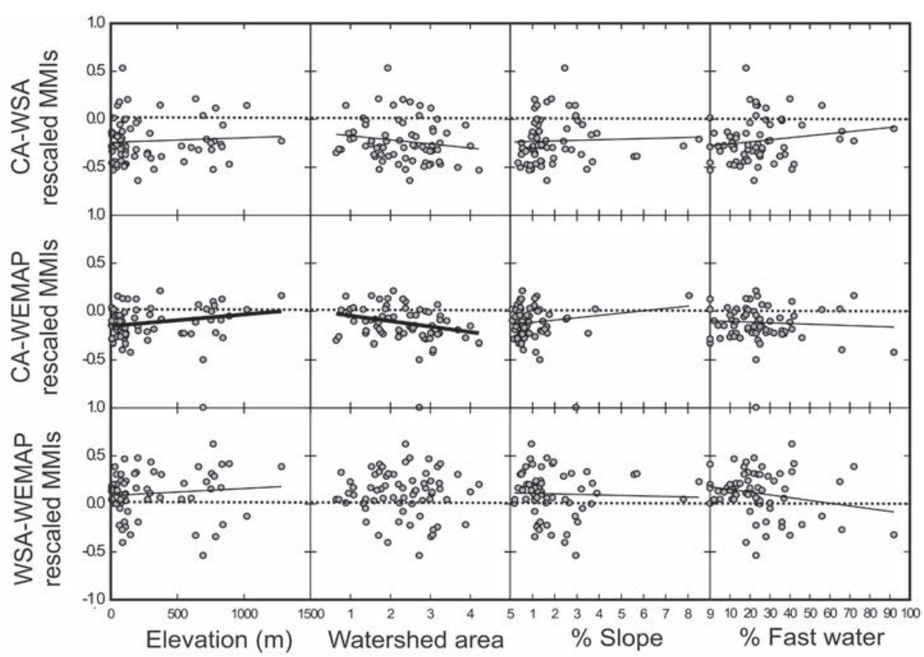


Figure 8. Scatterplots and regressions between the pair-wise differences in rescaled MMI values for the three different MMIs and four environmental gradients at CA test sites. The dashed horizontal lines represent zero difference. Thick solid lines denote regressions with r^2 and slopes significantly different from 0; thin solid lines denote those with intercepts significantly different from 0 but non-significant slope.

Table 8. Counts of CA sites declared impaired (I) and not impaired (NI) by CA MMI estimates and corresponding WEMAP and WSA MMI estimates.

		CA Mountain (n = 30)		CA Xeric (n = 38)		Total (n = 68)		All Sites
		I	NI	I	NI	I	NI	
CA	I	10	-	29	-	39	-	39
	NI	-	20	-	9	-	29	29
WEMAP	I	5	1	15	0	20	1	21
	NI	5	19	14	9	19	28	47
WSA	I	5	1	11	0	16	1	17
	NI	5	19	18	9	23	28	51

years (Poff 1997, Johnson *et al*. 2004, Johnson *et al*. 2007, Heino *et al*. 2007, Hoeinghaus *et al*. 2007, Mykrä *et al*. 2007, Mykrä *et al*. 2008). At the heart of all of these studies is a desire to clarify understanding of the factors that determine species distribution limits, one of the central goals of ecological theory (Levins 1966, Wiens 1989, Peters 1991, Brown *et al*. 1996, Guisan and Zimmermann 2000). This issue has significant implications for the utility of biotic indices because their effectiveness depends understanding how distribution patterns of individual taxa are influenced by landscape and waterway environmental heterogeneity, and how those effects are expressed at different scales of observation.

Index Comparability

O/E indices

Matching test sites with appropriate reference condition is a critical element of all bioassessments (Moss *et al*. 1987, Hughes *et al*. 1995, Stoddard *et al*. 2007). Errors in specifying the correct reference condition can lead to either under- or over-estimates of the true biological condition at individual sites. Our results show that the failure of the large-scale predictive models to account for the effects of some naturally occurring environmental factors caused substantial systematic differences among the O/E values derived from these models relative and those derived from the CA models. The fact that the most spatially extensive models (WEMAP and WSA-West models) did not adjust for the effects of local environmental heterogeneity (i.e., slope, percent fast-water habitats) on E, and hence O/E, shows that such spatially extensive models may have limited applicability for site-specific assessments and use of these assessments to generate regional assessments. There are several reasons the more spatially extensive models may have failed to account for the effects of reach slope and percent fast water on assemblage composition. First, available map-derived variables may not have been good surrogates for these variables when used at large scales. For example, watershed area is likely related to one or more factors that influence taxa presence at a site, including channel slope and amount of fast-water habits (Hynes 1970, Allan and Castillo 2007). However, watershed area might not be consistently associated with channel slope across a region the size of the western United States. In the three sets of models we examined, watershed area appeared to account for differences among sites in channel reach for only the spatially less extensive CA models. Even in those models that used direct measures of channel slope as a predictor variable (e.g., the WSA-West model), the relationship between invertebrate taxa and slope may be obscured by strong relationships between invertebrate composition and predictors that vary markedly across regions, such as temperature and precipitation. Furthermore, a predictive model based on linear relationships between biotic composition and predictor variables will fail to accurately describe any non-linear relationships and hence inaccurately predict the taxa that should occur under specific states of that variable. In contrast, over a smaller range of environmental conditions, surrogate predictors such as watershed area, temperature, or precipitation may adequately capture differences between sites in local habitat features such as channel slope and type of habitat. In general, these problems of prediction bias might be reduced in the future by both improving how well reference site networks represent all

Table 9. Summary of differences in precision, bias, responsiveness, and sensitivity of the WEMAP and WSA indices relative to CA indices. M = mountain ecoregion, X = xeric ecoregion. The term "similar" indicates no statistical difference; the terms "lower" and "higher" indicate the direction of a significant difference.

Performance Measure	O/E		MMI	
	WEMAP	WSA	WEMAP	WSA
Precision	Lower	Lower	Similar	Lower
Bias	Yes	Yes	Yes	Yes
Responsiveness	Lower	Lower	Lower	Lower
Sensitivity	Similar	Lower	Lower	Lower

streams of interest (in terms of both sample size and types of streams) and by using robust predictors such as Random Forests (Cutler *et al*. 2007) that do not assume linear relationships.

The fact that the WSA-West model strongly underestimated impairment relative to the CA model has at least two possible explanations: 1) poorer precision in the WSA-West model resulted in lower impairment thresholds and thus fewer impairment decisions, and 2) WSA underestimated the probabilities of capture of some of the taxa that contribute to the O/E calculations. The second result could have arisen if the reference sites used to predict the fauna in California streams were less rich on average than the otherwise similar California sites assessed. Vinson and Hawkins (1996) reported that invertebrate taxa richness in streams draining mountainous regions of California (Coast Range Mountains and Sierra Nevada) was higher than streams draining other mountainous regions in the western USA. Models based on a mix of reference sites from across the western United States might therefore be expected to under-predict richness at CA mountain sites. This explanation seems plausible for the WSA-West model, because average WSA-West O/E values for CA mountainous reference sites were greater than 1 on average (Sierra Nevada = 1.04, Southern Coastal Mountains = 1.11, and Klamath Mountains = 1.04). However, WEMAP reference site O/E values did not exhibit this trend. It seems prudent that we should refine models to explicitly account for the effects of biogeographic history on taxa richness. Such modeling might be accomplished through the use of categorical predictive variables that classify sites by their relevant zoogeographic region rather than general purpose ecoregions (Hawkins and Vinson 2000, Hawkins *et al*. 2000b). The contrasting result for the WEMAP model (i.e., that WEMAP model did not underestimate impairment relative to the CA model despite precision values intermediate between the CA and WSA models) is likely the consequence of the tendency of the WEMAP model to score sites lower than the WSA model.

Multimetric indices

Although, agreement among the MMI scores was considerably stronger than for the O/E indices, the relationships between scores were not consistent across the scoring range, indicating differences in responsiveness of the indices at low vs. high biotic condition sites. Also, although the WEMAP and WSA-West MMIs were derived from nearly identical datasets, there were numerous differences in the performance of the two larger MMIs, including precision, responsiveness and sensitivity. These differences reflect the different approaches used to develop the MMIs (Ode *et al*. 2005; Rehn *et al*. 2005; Stoddard *et al*. 2005, 2008).

Differences in MMI responsiveness were likely caused by one or more of the following: 1) differences in how metrics were scaled in the separate indices, 2) differences in the quality of sites used to calibrate the indices, or 3) differences in how individual metrics in each MMI respond to stress. Because there was considerable overlap in metrics among the indices, much of the difference among the MMIs in their assessments probably lies in differences in the scoring ranges of specific metrics. For example, although the number of EPT taxa is a nearly ubiquitous metric in MMIs (Karr and Chu 1999), the scoring range for this metric varies among regions. An EPT scoring range established from reference site data combined across a large spatial extent will not necessarily reflect local reference conditions. In some regions, test sites will be underscored; in others they will be overscored. We found evidence of this effect in the number of disagreements in impairment decisions made under the different MMIs. Furthermore, the WSA-West MMI did not indicate a difference in biotic condition between mountain and xeric test sites, whereas the CA and WEMAP MMI did. This finding was echoed in the way impairment decisions differed between WEMAP and WSA-West indices in xeric and mountain regions. Both WEMAP and WSA-West MMIs tended to overestimate impairment at mountain sites relative to the CA MMI, whereas the WSA-West MMI underestimated impairment at xeric sites relative to the CA MMI.

A final potential explanation is that differences in MMI performance were related to differences in the calibration sets used to derive the metric scoring ranges. Because MMIs are calibrated with both reference and test data, any difference in the biological quality of either set of calibration sites can affect a site's scoring, just as they can in O/E models (Hawkins 2006). Because of incomplete information regarding the quality of reference and test sites used to calibrate the different indices, how seriously such differences affected index performance could not be addressed at this time.

Effects of spatial scale on index performance

It has been long known that taxonomic composition is influenced by natural environmental gradients. How these relationships are expressed at different spatial scales, and hence affect biological indices, is much less clear, but is of increasing interest (Finn and Poff 2005, Heino *et al*. 2007, Cao *et al*. 2007, Mykrä *et al*. 2008). MMIs and predictive models use different methods for accounting or adjusting for natural gradients. Predictive models are explicitly designed to describe how natural environmental gradients affect the distribution of individual taxa (Wright *et al*. 1989, 2000). However, some natural gradients may be important at certain geographic scales, but cease to matter at other scales, as shown in this study and elsewhere (Mykrä *et al*. 2008).

In contrast to O/E indices, MMIs attempt to minimize the effects of natural gradients by a priori classification of reference sites into environmentally homogeneous sets of sites. In addition, metrics are selected to be insensitive to natural gradients, or by adding correction factors that adjust for scoring differences along gradients (Karr and Chu 1999). In this study, for example, scoring ranges for the EPT richness metric varied little across spatial scales within ecoregions (Ode *et al*. 2005; Rehn *et al*. 2005; Stoddard *et al*. 2005, 2008), and the CA MMI for the North Coast explicitly corrects for watershed area in affected metrics (Rehn *et al*. 2005).

In this study, the large-scale predictive models were not completely successful in adjusting for two of the gradients (percent slope and percent fastwater habitats) we examined. Likewise, the CA and WSA-West MMIs were not completely effective at controlling for an elevation gradient.

Index performance and model traits

All the biological indices in our evaluations produce scores by comparing biological expectations to observed biology. Although E in O/E is explicitly modeled (i.e., predicted), MMI expectations are derived from a set of reference sites that are grouped (by ecoregion, stream size, etc.) to maximize similarity of the biological assemblages at reference sites. Thus, both O/E and MMI are indices based on modeled expectations. Levins (1966) postulated that there is an inherent tradeoff among three desirable model traits: reality (i.e., accuracy, or lack of bias), precision, and generality (see also Guisan and Zimmermann 2000). Although these model traits are not necessarily mutually exclusive, we cannot expect the models used to predict biotic conditions to optimize each trait. In creating standardized indices applicable across a large range of geoclimatic conditions, generality was improved at the expense of both reality and precision. This tradeoff points to the need to develop more localized models for bioassessment programs, especially those that use biocriteria to infer if streams are supporting their designated aquatic life uses. However, the fact that impairment decisions can be very sensitive to the thresholds used to define impaired conditions (as seen when an ecoregion-based correction was applied to the WSA-West model for O/E comparisons), suggests that it may be possible to adjust for some of the systematic differences among the models. Larger models could be rendered more suitable for local application by calibrating impairment thresholds to local reference conditions. In practice, a local regulatory entity could recalculate the standard deviations for O/E or MMI models based only on local reference sites and use these to set locally relevant thresholds.

Concluding Remarks

The answer to the central question of whether indices developed from geoclimatically extensive data can substitute for more locally produced indices depends both on their intended use and the type of indicator. In regional condition assessments, accuracy (lack of bias) is more important than precision. That is, for low precision can be compensated by looking at large numbers of samples with the expectation that the estimated average condition will still be accurate. For the purpose of regional assessments, use of the WEMAP O/E index produced results that were generally comparable to the CA indices. In contrast, because of its strong bias, the WSA-West O/E index would probably underestimate regional impairment. Likewise, lower precision and differences in responsiveness across the scoring range make the WSA-West MMIs less desirable for regional condition assessments.

For site-specific assessments, where both accuracy and precision are important, it seems clear that locally derived indices should outperform large-scale indices for both types of index (see also Mykrä *et al*. 2008). Because most applications of bioassessment tools are site-specific, there is a clear need to continue to develop regional models that explicitly take locally important gradients into account (Heino *et al*. 2007). However, because the WEMAP MMI had similar precision and WEMAP MMI scores were highly correlated with CA MMI scores, the WEMAP MMI might provide an acceptable substitute in California (and potentially other regions in the western US) until local MMIs are developed, assuming care is taken to adjust impairment thresholds to reflect local reference conditions.

Finally, these results suggest three related applied research needs: 1) identifying the geographic or geoclimatic scale that optimizes index performance, 2) determining the factors that most strongly influence index performance and identifying the geographic scales at which they vary, and 3) identifying ways of more accurately specifying the reference condition from geoclimatically extensive sets of reference site data. It is not known much about which factors influence the optimal geographic scale for producing either predictive models or multimetric indices, but the rapidly expanding field of bioassessment would benefit greatly from the ability to predict these factors.

Literature Cited

Allan, J.D. and M.M. Castillo. 2007. Stream Ecology: Structure and Function of Running Waters. 2nd Edition. Kluwer. Dordrecht, The Netherlands.

Andrewartha, H.G. and L.C. Birch. 1954. The Distribution and Abundance of Animals. The University of Chicago Press. Chicago, IL.

Bailey, C.R., R.H. Norris and T.B. Reynoldson. 2004. Bioassessment of Freshwater Ecosystems: Using the Reference Condition Approach. Kluwer. Dordrecht, The Netherlands.

Barbour, M.T. and C.O. Yoder. 2000. The multimetric approach to bioassessment, as used in the United States. pp. 281-292 *in*: J.F. Wright, D.W. Sutcliffe and M.T. Furse (eds.), Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. Freshwater Biological Association. Ambleside, United Kingdom.

Bonada, N., N. Prat, V.H. Resh and B. Statzner. 2006. Development in aquatic insect biomonitoring: a comparative analysis of recent approaches. *Annual Review of Entomology* 51:495-523.

Brown, J.H., G.C. Stevens and D.M. Kaufman. 1996. The geographic range: size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics* 27:597-623.

Cao, Y., C.P. Hawkins, J.Olson and M. Kosterman. 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. *Journal of the North American Benthological Society* 26:566-585.

Clarke, R.T., J.F. Wright and M.T. Furse. 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecological Modelling* 160:219-233.

Cutler, D.R., T.C. Edwards, Jr., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson and J.J. Lawler. 2007. Random forests for classification in ecology. *Ecology* 88:2783-2792.

Environmental Protection Agency (EPA). 2006. Wadeable Streams Assessment: A Collaborative Survey of the Nation's Streams. EPA 841-B-06-002. EPA Yosemite. Yosemite, CA.

Finn, D.S. and N.L. Poff. 2005. Variability and convergence in benthic communities along the longitudinal gradients of four physically similar Rocky Mountain streams. *Freshwater Biology* 50:243-261.

Guisan, A. and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modeling* 135:147-186.

Hawkins, C.P. 2006. Quantifying biological integrity by taxonomic completeness: evaluation of a potential indicator for use in regional- and global-scale assessments. *Ecological Applications* 16:1251-1266.

Hawkins, C.P., R.H. Norris, J.N. Hogue and J.M. Feminella. 2000a. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10:1456-1477.

Hawkins, C.P., J. Ostermiller, M. Vinson, R.J. Stevenson and J. Olson. 2003. Stream algae, invertebrate, and environmental sampling associated with biological water quality assessments. Western Center for Monitoring and Assessment of Freshwater Ecosystems. Utah State University. Logan, UT.

Hawkins, C.P., R.H. Norris, J. Gerritsen, R.M. Hughes, S.K. Jackson, R.H. Johnson and R.J. Stevenson. 2000b. Evaluation of landscape classifications for biological assessment of freshwater ecosystems: synthesis and recommendations.

Journal of the North American Benthological Society 19:541-556.

Hawkins, C.P. and M.R. Vinson. 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessments. *Journal of the North American Benthological Society* 19:501-517.

Heino, J., H. Mykrä, J. Kotanen and T. Muotka. 2007. Ecological filters and variability in stream macroinvertebrate communities: do taxonomic and functional structure follow the same path? *Ecography* 30:217-230.

Herlihy, A.T., S. Paulsen, J. Van Sickle, J. Stoddard, C.P. Hawkins and L.L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference condition approach at a continental scale. *Journal of the North American Benthological Society* 27:860-877.

Hoeinghaus, D.J., K.O. Winemiller and J.S. Birnbaum. 2007. Local and regional determinants of stream fish assemblage structure: inferences based on taxonomic vs. functional groups. *Journal of Biogeography* 34:324-338.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. pp. 31-47 *in*: W.S. Davies and T.P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making. Lewis Publishers. Ann Arbor, MI.

Hughes, R.M., P.R. Kaufmann, A.T. Herlihy, T.M. Kincaid, L. Reynolds and D.P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. *Canadian Journal of Fisheries and Aquatic Sciences* 55:1618-1631.

Hutchinson, G.E. 1959. Homage to Santa Rosalia; or, why are there so many kinds of animals. *American Naturalist* 93:145-159.

Hynes, H.B.N. 1970. The Ecology of Running Waters. University of Toronto Press. Toronto, Canada.

Johnson, R. K., M.T. Furse, D. Hering and L. Sandin. 2007. Ecological relationships between stream communities and spatial scale: implications for designing catchment-level monitoring programs. *Freshwater Biology* 52:939-958.

Johnson, R.K, W. Goedkoop and L. Sandin. 2004. Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. *Freshwater Biology* 49:1179-1194.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6:21-27.

Karr, J.R. and E.W. Chu. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Island Press. Washington, DC.

Kerans, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications* 4:768-785.

Klemm, D.J., K.A. Blocksom, F.A. Fulk, A.T. Herlihy, R.M. Hughes, P.R. Kaufmann, D.V. Peck, J.L. Stoddard, W.T. Thoeny, M.B. Griffith and W.S. Davis. 2003. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic Highland streams. *Environmental Management* 31:656-669.

Levins, R. 1966. The strategy of model building in population ecology. *American Scientist* 54:421-431.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. pp. 281-297 *in*: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. Berkeley, CA.

McCormick, F.H., R.M. Hughes, P.R. Kaufmann, D.V. Peck, J.L. Stoddard and A.T. Herlihy. 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands Region. Transactions of the *American Fisheries Society* 130:857-877.

Moss, D., M.T. Furse, J.F. Wright and P.D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology* 17:41-52.

Mykrä, H., J. Aroviita, J. Kotanen and H. Hämäläinen. 2008. Predicting the stream macroinvertebrate fauna across regional scales: influence of geographical extent on model performance. *Journal of the North American Benthological Society* 27:705-716.

Mykrä, H., J. Heino and T. Muotka. 2007. Scale-related patterns in spatial and environmental components of stream macroinvertebrate assemblage variation. *Global Ecology and Biogeography* 16:149-159.

Ode, P.R., A.C. Rehn and J.T. May. 2005. A quantitative tool for assessing the integrity of southern coastal California streams. *Environmental Management* 35:493-504.

Omernik, J.M. 1987. Ecoregions of the conterminous United States Map (scale 1:7,500,000). *Annals of the Association of American Geographers* 77:118-125.

Ostermiller, J.D. and C.P. Hawkins. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. *Journal of the North American Benthological Society* 23:363-382.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D. Klemm, J.M. Lazorchak, F.H. McCormick, S.A. Peterson, P.L. Ringold,T. Magee and M. Cappaert. 2006. Environmental Monitoring and Assessment Program-Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA/620/R-06/003. US Environmental Protection Agency. Washington, DC.

Peters, R.H. 1991. A Critique for Ecology. Cambridge University Press. Cambridge, MA.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society* 16:391-409.

Rehn, A.C., P.R. Ode and C.P. Hawkins. 2007. Comparisons of targeted-riffle and reach-wide benthic macroinvertebrate samples: implications for data sharing in stream-condition assessments. *Journal of the North American Benthological Society* 26:332-348.

Rehn, A.C., P.R. Ode and J.T. May. 2005. Development of a benthic index of biotic integrity (B-IBI) for wadeable streams in northern coastal California and its application to regional 305(b) assessment. Report to the State Water Resources Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory. Rancho Cordova, CA.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day and D.M. Rosenberg. 1997. The reference condition: A comparison of multimetric and multivariate approaches to assess water quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* 16:833-852.

Richards, A.B. and D.C. Rogers. 2006. List of freshwater macroinvertebrate taxa from California and adjacent states including standard taxanomic effort levels. Southwest Association of Freshwater Invertebrate Taxonomists. http://www.waterboards.ca.gov/swamp/docs/safit/ste_list.pdf.

Simpson, J.C. and R.H. Norris. 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs. pp. 125-142 *in*: J.F. Wright, D.W. Sutcliffe and M.T. Furse (eds), Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. Freshwater Biological Association. Amblesite, United Kingdom.

Statzner, B., B. Bis, S. Dolédec and P. Usseglio-Polatera. 2001. Perspectives for biomonitoring at large spatial scales: a unified measure for the functional classification of invertebrate communities in European running waters. *Basic and Applied Ecology* 2:73-85.

Stoddard, J.L., A.T. Herlihy, D.V. Peck, R.M. Hughes, T.R. Whittier and E. Tarquino. 2008. A process for creating multi-metric indices for large scale aquatic surveys. *Journal of the North American Benthological Society* 27:878-891.

Stoddard, J.L., D.V. Peck, A.R. Olsen, D.P. Larsen, J. Van Sickle, C.P. Hawkins, R.M. Hughes, T.R. Whittier, G. Lomnicky, A.T. Herlihy, P.R. Kaufmann. 2006. Environmental Monitoring and Assessment Program (EMAP) Western Streams and Rivers Statistical Summary. EPA 620/R-05/006. US Environmental Protection Agency. Washington, DC.

Stoddard, J.L., D.V. Peck, S.G. Paulsen, J. Van Sickle, C.P. Hawkins, A.T. Herlihy, R.M. Hughes, P.R. Kaufmann, D.P. Larsen, G. Lomnicky, A.R. Olsen, S.A. Peterson, P.L. Ringold and T.R. Whittier. 2005. An Ecological Assessment of Western Streams and Rivers. EPA 620/R-05/005. US Environmental Protection Agency. Washington, DC.

Van Sickle, J., C.P. Hawkins, D.P. Larsen and A.T. Herlihy. 2005. A null model for the expected macroinvertebrate assemblage in streams. *Journal of the North American Benthological Society* 24:178-191.

Van Sickle, J., D.P. Larsen and C.P. Hawkins. 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. *Journal of the North American Benthological Society* 26:319-331.

Van Sickle, J. and S.G. Paulsen. 2008. Assessing the attributable risks, relative risks, and regional extents of aquatic stressors. *Journal of the North American Benthological Society* 27:920-931.

Vinson, M.R. and C.P. Hawkins. 1996. Effects of sampling area and subsampling procedure on comparisons of taxa richness among stream. *Journal of the North American Benthological Society* 15:392-399.

Wiens, J. 1989. Spatial scaling in ecology. *Functional Ecology* 3:385-397.

Wright, J.F., P.D. Armitage, M.T. Furse and D. Moss. 1989. Prediction of invertebrate communities using stream measurements. *Regulated Rivers: Research and Management* 4:147-155.

Wright, J.F., D.W. Sutcliffe and M.T. Furse. 2000 Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. Freshwater Biological Association. Ambleside, United Kingdom.

Yoder, C.O. and E.T. Rankin. 1995. The role of biological criteria in water quality monitoring, assessment and regulation. Technical Report MAS/1995-1-3. Ohio EPA. Columbus, OH.

Yuan, L.L., C.P. Hawkins and J. VanSickle. 2008. Effects of regionalization decisions on an O/E index for the national assessment. *Journal of the North American Benthological Society* 27:892-205.

Zar, J.H. 1999. Biostatistical Analysis (4th edition). Prentice Hall. Englewood Cliffs, NJ.

Acknowledgements

The authors would like to thank staff of the California Department of Fish and Game Aquatic Bioassessment Laboratory who collected the field data and identified the benthic invertebrates used in the test datasets and Alan Herlihy who provided helpful information about WEMAP and WSA-West model development. Funding for the EMAP and CMAP programs was provided by the EPA (Assistance Agreement No. CR82823801) and the California State Water Resources Control Board's (SWRCB) Non Point Source Program (Agreement Number 03-273-250-2). This study was supported by the SWRCB's Surface Water Ambient Monitoring Program. CPH was supported, in part, by USEPA Science To Achieve Results (STAR) grants R-82863701 and R-83059401 and a contract with Region 5 of the USDA Forest Service. This manuscript was greatly improved by comments from John Van Sickle, two anonymous reviewers, and Kerry Ritter.

Appendix 5

B182

SWAMP
Surface Water Ambient Monitoring Program

Final Technical Report | **2009**

Recommendations for the Development and Maintenance of a Reference Condition Management Program (RCMP) to Support Biological Assessment of California's Wadeable Streams

March 2009

California Water Boards

www.waterboards.ca.gov/swamp

B183

Recommendations for the development and maintenance of a reference condition management program (RCMP) to support biological assessment of California's wadeable streams

Report to the State Water Resources Control Board's
Surface Water Ambient Monitoring Program (SWAMP)

Peter Ode, SWAMP Bioassessment Coordinator
Aquatic Bioassessment Laboratory/
Water Pollution Control Laboratory
California Department of Fish and Game
2005 Nimbus Road
Rancho Cordova, CA 95670

Ken Schiff, Deputy Director
Southern California Coastal Water Research Project
3535 Harbor Blvd., Suite 110
Costa Mesa, CA 92626

March 2009

Technical Report 581

Table of Contents

EXECUTIVE SUMMARY

Direct measures of the ecological condition of waterbodies have received a recent surge in interest within California's water quality management and regulatory programs because biology-based assessments have several advantages over chemistry- or toxicity-based assessments. Biological assessments are more closely linked to the beneficial uses to be protected and chemistry- or toxicity-based criteria usually lack the predictive ability to infer biological condition. Ultimately, California needs to develop biology-based standards, or biocriteria, as a regulatory tool for monitoring and protecting aquatic life use.

Biological assessment tools, including biocriteria, attempt to objectively "score" the biological integrity at a given site. A crucial component to the development of assessment tools is understanding biological expectations at reference sites that consist of natural, undisturbed systems. These reference systems set the biological condition benchmarks for comparisons to the site(s) being evaluated. Two recent external reviews of the State Water Board's Surface Water Ambient Monitoring Program (SWAMP) affirmed the importance of a sound statewide reference condition program (i.e., TetraTech 2002, SPARC 2006).

In October 2007, the SWAMP bioassessment committee assembled a technical panel of statewide and national experts in bioassessment. The panel met for three days to develop a set of recommendations that the SWAMP program could use to establish and maintain a comprehensive reference condition management program (RCMP). The program accounts for biological variation caused by natural environmental gradients and balances statewide consistency with the flexibility needed to adapt to California's diverse regional settings. Furthermore, the plan allows for adaptive refinement over time.

The panel defined a general strategy for establishing the RCMP that has four components:

1. California will be divided into different geographic regions based on coarse biogeographic similarities in order to partition some of the natural variability among regions (these boundaries should be consistent with those used for the SWAMP Perennial Streams Assessment)
2. A pool of reference sites will be assembled within each region through a sequential process of identification and screening of candidate sites
3. The sites within each reference pool will be managed through iterative review of data to refine regional boundaries, ensure continued suitability of sites and ensure adequate representation of natural gradients
4. A monitoring design will be created for sampling this pool of reference sites to document the range of biological and physical condition at reference sites, and monitor for changes to this condition over time

The panel recommended identifying and screening candidate locations to create a pool of verified reference sites using either a "standard model" or an "alternative model". The standard model will cover the vast majority of the state where high quality sites are available. The alternate model will apply in those regions where an insufficient quantity of high quality sites exist and another strategy is required for selecting candidates for the reference pool. This may include regions such as the agriculturally dominated Central Valley or the intensely urbanized southern California coastal plain.

The standard model is a synthesis of widely used techniques for selection and screening candidate sites using a toolbox consisting of existing site data, GIS techniques, expert knowledge and site visits. The alternative approach consists of two general strategies: 1) modification of standard tools (e.g., lowering the GIS screening thresholds, collecting more intensive site data) and 2) use of non-standard approaches. The non-standard approaches include:

- Select best sites using existing biological indices
- Species pool approach
- Factor-ceiling approach
- Model taxon preferences for limiting environmental gradients

These different approaches are not mutually exclusive and several panel members recommended they be used in combination to provide weight-of-evidence that candidate sites are acceptable for the reference pool in these difficult locations.

The panel outlined a monitoring strategy for the RCMP, which included recommendations for sampling methods, sampling density and frequency, and the set of biological, chemical and physical attributes that should be collected at each reference site. The panel strongly recommended that the RCMP should be compatible with ongoing statewide monitoring programs such as the newly developed SWAMP Perennial Streams Assessment. For the monitoring design, the panel recommended both random and targeted sites. A probabilistic rotating panel was suggested for the random design because it provides an unbiased method for defining natural variability while still optimizing large-scale trend detection. Targeted repeated sampling designs are useful for detecting trends at specific locations; some of these sites have been sampled for years and provide a rich history that should not be lost.

To guide the SWAMP program as it implements the RCMP, the panel made a series of recommendations for prioritizing the elements of the plan. The panel recommended that the implementation begin by screening existing datasets for reference sites, followed by a combination of GIS screens and site visits to fill in gaps in regions with few reference sites.

FOREWORD

The recommendations in this document were developed by a technical panel composed of experts in bioassessment. The panel reflected a broad range of local, statewide, and national experiences with freshwater bioassessment, specifically with defining reference conditions for bioassessment and biocriteria. The panel met for three days on October 17-19, 2007 to outline the content of this document. The meeting followed a four-step process:

1) Defining the background of the problem
2) Establishing a set of guiding philosophies for the development of a reference site management plan
3) Providing general guidance by outlining an overall approach
4) Providing detailed guidance for specific technical issues

This document follows a similar format. This document captures all of the items agreed to by consensus of the group and attempts to point out diverging opinions or unresolved issues. On occasion, we expand on key concepts that were implicit to our discussions, but may not have been discussed directly. Where appropriate, we use sidebars, tables, and figures to illustrate key concepts or provide additional information. Thank you to Dr. Robert Hughes (Oregon State University) for additional document review.

Panel Members (from left to right): David Herbst (University of California at Santa Barbara, Sierra Nevada Aquatic Research Laboratory), Peter Ode (California Department of Fish and Game, Aquatic Bioassessment Laboratory), Raphael Mazor (Southern California Coastal Water Research Project), D. Phil Larsen (US EPA retired, Western Ecology Division), Andrew Rehn (California Department of Fish and Game, Aquatic Bioassessment Laboratory), Lenwood Hall (University of Maryland, Wye Research and Education Center), Terrence Fleming (US EPA Region IX, Office of Water), Charles Hawkins (Utah State University, Western Center for Monitoring and Assessment of Freshwater Ecosystems), Alan Herlihy (Oregon State University, Department of Fisheries and Wildlife), Kenneth Schiff (facilitator, Southern Coastal California Water Research Project).

CONTEXT: LINKING BIOASSESSMENT TO BIOCRITERIA[1]

Aquatic bioassessment is the applied science of interpreting the ecological condition of waterbodies directly from the organisms that inhabit them. Biocriteria are narrative or numeric standards that define whether the integrity of biological communities is impaired at a specific site. Water quality regulatory programs can receive many benefits from adopting biology-based standards as targets of their policies and management actions. The key to using biology-based methods effectively is the establishment of benchmarks that objectively define the biological expectations (or potential) of a given site. Reference conditions provide these objective benchmarks.

Why bioassessment?

The Clean Water Act (Section 101a) requires states to "restore and maintain the chemical, physical and biological integrity" of their waterbodies. For decades, most state water quality monitoring programs have focused on the chemical integrity (and to a lesser extent physical integrity) of waterbodies largely because these parameters are relatively simple to sample, relatively straightforward to measure and evaluate, and methods for developing chemical criteria are relatively standardized. While chemical/ toxicological and physical condition monitoring may provide indirect measures of ecological condition, exclusive focus on these measures is inadequate for protection of aquatic life uses, one of the primary beneficial uses of concern in water quality management. Because many chemical/ physical water quality thresholds are based on toxicity to aquatic organisms (USEPA WQS handbook, 2nd Edition 1994), these indirect measures are often surrogates for the beneficial use that is the target of protection efforts. Furthermore, biological integrity is frequently impaired by factors other than chemical contamination (e.g., hydrologic alteration, instream and riparian habitat alteration). Ultimately, ecological condition assessments provide the most appropriate assessment endpoint for protecting beneficial uses associated with aquatic life.

Why biocriteria?

Adoption of biology-based regulatory standards has the potential to provide significant enhancements to the protection of water resource integrity because biocriteria provide a regulatory mechanism for applying bioassessment's benefits to numerous water resource objectives.

The State Water Resources Control Board's Surface Water Ambient Monitoring Program (SWAMP) is supporting the biocriteria goal by developing tools for using benthic macroinvertebrates as indicators of the health of aquatic life in perennial streams. SWAMP's objective is to develop the bioassessment infrastructure (i.e., standardized methods, analytical tools, objective reference conditions, interpretive framework) that will enable water quality programs to employ biocriteria in a variety of regulatory applications.

---

[1] Much of the information summarized in this section was synthesized from several key sources: Barbour *et al*. 1996a, Karr 1995, 1997, Stoddard *et al*. 2006.

Importance of reference conditions to bioassessment and biocriteria

The development of chemical criteria for aquatic life follows a relatively straightforward process in which numerical standards are based on results from lab-based toxicity testing. For most chemical contaminants, management objectives are focused on keeping concentrations below these toxicity-derived numerical thresholds. In contrast, biological objectives are based on maintaining the integrity of an assemblage (or multiple assemblages) of organisms. The challenge in developing biocriteria is translating what is currently a narrative standard into an ecologically relevant numerical standard. Development of biological criteria, however, is complicated by the fact that the composition of stream communities varies naturally even in the absence of anthropogenic stress. Thus, biocriteria will require a fundamentally different approach to establishing the expectations for unimpaired waterbodies.

Reference conditions (based on reference sites) provide a widely accepted mechanism for defining appropriate expectations and accounting for this natural variability (Hughes *et al*. 1986, Barbour *et al*. 1996, Karr and Chu 1999, Bailey *et al*. 2004). Reference sites are sections of streams that represent the desired state of stream condition (*sensu* Meyer 1997) for a region of interest. Once suitable reference reaches have been identified, these are used to characterize the range of biotic conditions expected for minimally disturbed sites. Deviation from this range is then used as evidence that test sites are impaired. [2]

A standardized lexicon of terms used to define biological expectations (adapted from Stoddard *et al.* 2006):

Reference Condition (RC(BI)) ~ Because this term has been used for a wide range of meanings, Stoddard *et al.* (2006) argue that the term should be restricted to meaning "reference condition for biological integrity … in the absence of significant human disturbance or alteration"

Minimally Disturbed Condition (MDC) ~ stream condition in the absence of "significant" human disturbance. Assumes all streams have some anthropogenic stresses, but in most cases will approach true RC(BI)

Historical Condition (HC) ~ stream condition at a specific point in time (e.g., pre-Columbian, pre-industrial, pre-intensive agriculture, etc.)

Least Disturbed Condition (LDC) ~ the best physical, chemical and biological conditions currently available ("the best of what's left"). This definition is sufficiently flexible to establish biological expectations even in highly altered systems

Best Attainable Condition (BAC) ~ the expected ecological condition of least disturbed sites given use of best management practices for an extended period of time. This definition is helpful for communicating the potential for improving ecological condition above the currently best available conditions

Tiered aquatic life use (TALU) framework

The potential for biocriteria to improve aquatic life beneficial use protection can be greatly enhanced by a flexible framework for interpreting beneficial use attainment in a variety of settings. The current system of aquatic life use designations in California is outdated and does not adequately take advantage of advances in our ability to assess aquatic life use attainment. The USEPA and other states (notably, Maine and Ohio) have recognized this problem and have

[2] Approaches to the selection of reference sites have been discussed extensively (Hughes and Larsen 1988, Hughes 1995, Rosenberg *et al.* 1999, Stoddard *et al*. 2006). Although there has been much debate about terminology used to describe expected biological conditions, the concept is flexible and can be applied either very narrowly (e.g., the condition of waterbodies before European invasions) or more broadly (e.g., the "least disturbed" or "best available" conditions currently found in a region of interest). The strategy in this document follows terminology usage recommended by Stoddard *et al*. 2006 (see text box).

developed a “tiered” system of aquatic life use designations, which utilize the power of biological information to develop graduated levels of protection.

“Tiered aquatic life uses” (TALU), supported by numeric biocriteria, can be thought of as defining different management levels for biological condition across a quality continuum that ranges between “natural” conditions to complete loss of the natural biological community (Figure 1). In the TALU system, “tiers” represent classes of waterbodies that are grouped based on similarities in anthropogenic disturbance levels, resulting biological condition, and recovery potential (USEPA 2005). Under this flexible system, designated uses to support aquatic life can cover a broad continuum of biological conditions, with some waters being closer to the ideal of “natural” or “minimal human impact” than others. Biocriteria applied in a framework of TALU designations can help shift the regulatory focus from performance-based standards (e.g., limiting the number of chemical criteria exceedences) to impact-based standards (e.g., attainment of ecological condition targets).

Reference conditions play two distinct roles in the TALU framework

The y-axis in the TALU framework (see Figure 1) is biological condition, a scale that measures ecological integrity of a site. The upper limit of the biological condition axis is anchored by an idealized target that represents the natural state of ecological conditions, or RC(BI) in the strict sense of Stoddard *et al.* (2006).

In addition, within each tier, there is some best attainable condition (BAC, *sensu* Stoddard *et al.* 2006) for waterbody classes in these tiers.

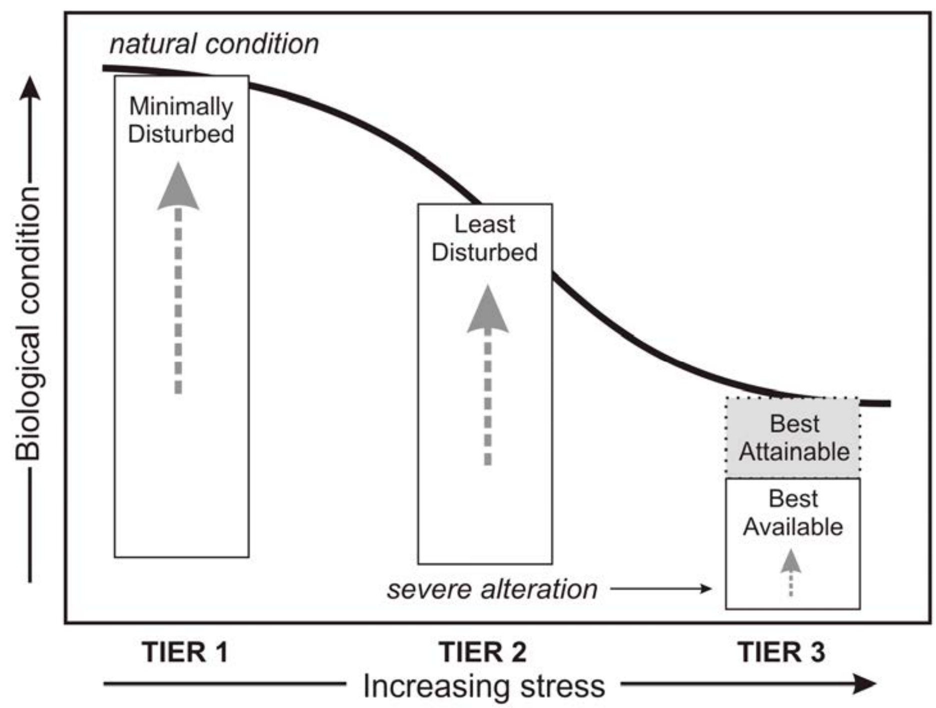


Figure 1. The biological condition gradient (BCG) used to define stream condition tiers in the TALU framework. Boxes indicate the expected range of biological condition scores at sites within each tier. Figure modified from Stoddard *et al*. 2006.

INTRODUCTION

General background

As the use of biological information in states' water quality regulatory programs has expanded across the US, these programs have followed a typical progression in which biosurveys (collection of biological samples, often as supplements to existing chemical monitoring) are followed by bioassessments (assessing ecological condition from biological data), finally progressing to full biocriteria (use of biological data to make regulatory decisions about aquatic life use condition).

As other programs proceeded along the path toward standardized interpretation of bioassessment data, they all recognized the need for grounding their programs with explicitly defined expectations for biological condition. Although criteria and procedures used to identify reference sites vary from program to program, the basic approaches used by most programs are quite similar. A partial review of water quality assessment programs in the North America (both state and federal programs), European Union (Water Framework Directive) and Australia (Water Reform Framework) revealed that many programs employed a similar GIS-based landscape-scale analysis to identify candidate watersheds, followed by site reconnaissance to evaluate reach-scale impacts (Barbour *et al*. 1996a, Whittier *et al*. 1987, Rosenberg *et al*. 1999, ANZECC and ARMCANZ 2000, Drake 2003, REFCOND 2003, Grafe 2004).

Reference sites manage natural variation

The composition of organisms at a site is a function of both natural and anthropogenic factors. These factors can be viewed as a series of "filters" that determine which taxa occur at a site (Poff and Ward 1990, Poff 1997, Statzner *et al*. 2001). For example, the pool of benthic macroinvertebrate taxa occurring within a large region like California's Sierra Nevada is a function of large scale processes (e.g., parent geology, climate and evolutionary history); the subset of taxa that occur at a given site at a given point in time is determined by a series of biotic and abiotic filters (e.g., life history traits, competition and predation, substrate composition, pH, thermal and hydrologic regimes, pollution tolerance) that further limit the occurrence of each taxon. The central challenge in bioassessment is to develop techniques that maximize the detection of signals of anthropogenic stress filters while minimizing the noise from natural filters. The identification of reference sites (that captures sources of natural variation) is a key component of most strategies for meeting this challenge (Hughes 1995, Wright and Li 2002, Bailey *et al*. 2004).

California's progress toward biocriteria implementation has followed a similar path. Since the early 1990s, bioassessment samples have been collected from more than 4000 sites by state and federal agencies alone (Figure 2). Some of these programs have been spatially extensive probability assessments of environmental condition such as the US EPA's Environmental Monitoring and Assessment Program (EMAP) and the California's Monitoring and Assessment Program (CMAP). Others are more directed studies to assess watershed-specific conditions or trends at locations of interest such as regional SWAMP monitoring, US Forest Service monitoring, and the US Geological Survey's National Water Quality Assessment Program (NAWQA). In addition, an abundance of additional sites have been sampled for National Pollutant Discharge Elimination System (NPDES) permit monitoring, and by citizen monitoring groups.

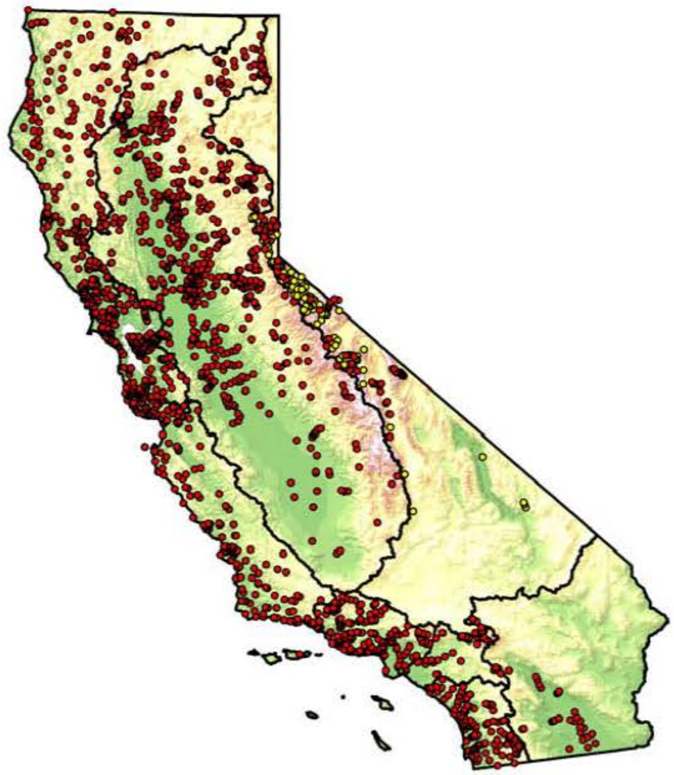


Figure 2. Approximately 3000 bioassessment sampling locations in California sampled between 1994 and 2007. Red circles represent sites processed by Aquatic Bioassessment Laboratory, yellow circles represent those processed by Sierra Nevada Aquatic Research Laboratory. More than 1000 other sites have been sampled by other state and federal agencies, permitted dischargers and citizen monitoring groups.

Because the early applications of bioassessment techniques in California were fragmented, the procedures for defining reference condition were largely *ad hoc* or project specific, with little or no attempt to apply consistent methods from project to project. Most of the reference or "control" sites used in early California bioassessment studies (e.g., point source enforcement cases, watershed specific bioassessments) were selected to define local expectations and were not selected using common criteria that would enable comparisons among projects.

Several large scale efforts to screen reference sites were undertaken in the early 2000's to support biological index development or as part of large state probability surveys: Western EMAP (2000-2003) and CMAP (2004-2007). In a concurrent effort, the USFS collaborated with scientists at Utah State University to identify over 200 reference sites on forest service lands in California between 1998 and 2000. Sites from these sampling programs were combined with other regional datasets to produce several of the main biotic indices used in California (statewide O/E models, North Coast IBI, South Coast IBI). Separate reference sites were used to develop the Eastern Sierra IBI (Herbst and Silldorf 2006).

In all of the large-scale studies between 1998 and 2007, both landscape scale and local scale factors were used for screening reference sites. Although common approaches were used to screen sites for most of these projects, little or no attempt was made to ensure consistency in screening among projects. This limits the utility of existing reference sites for statewide applications for several reasons. First, each project may use very different factors for selecting reference sites (e.g., one program may rely more on landscape scale factors while another may rely more on local scale factors). Second, some projects may use similar factors to select reference sites, but use different thresholds to screen sites (e.g., road density cutoffs or % upstream development cutoffs). Third, even when similar screening criteria are used for the same landscape or local scale factors, temporal variation in the reference site data has rarely been accounted for.

Why SWAMP needs an RCMP

The recent commitment by the SWAMP program to develop bioassessment/ biocriteria infrastructure provides us with an opportunity and impetus to standardize the reference site selection process statewide. The SWAMP program has long recognized this need, recently devoting a significant portion of its funding to developing reference condition datasets. Three recent peer reviews of SWAMP affirmed the importance of this effort:

1. In 2002, the SWAMP program funded an external review of bioassessment programs throughout California. That review was conducted by the lead author of the USEPA's bioassessment guidance document for streams and rivers.[3]

2. In 2005-06, the entire SWAMP program was peer-reviewed by an external "Scientific Planning and Review Committee" (SPARC), comprised of water quality experts from around the country.[4] The SPARC strongly recommended that SWAMP continue to develop its bioassessment program as a very high priority, specifically commenting that: a) the state board should consider revamping its entire standards program to make better use of biological endpoints (i.e., bioassessments) and b) the bioassessment program should focus particular attention on fostering consistency in its scoring indices.

3. In 2008, the USEPA (2009) conducted a Critical Elements Review of SWAMP's progress toward developing the technical elements to support biocriteria. The review stressed the fundamental importance of defining reference conditions and supported CA's reference condition strategy.

Establishing consistency in SWAMP's reference site selection process is clearly a key to effective implementation of biocriteria. However, identifying reference sites for California's perennial streams is complicated by its size (i.e., there are more than 300,000

---

[3] The external review, conducted by Dr. Michael T. Barbour and Colin Hill of Tetra Tech, Inc., produced a final report in January 2003 titled *The Status and Future of Biological Assessment for California Streams*, which may be viewed on the Internet at http://www.swrcb.ca.gov/swamp/reports.html

[4] The SPARC's final report is posted at: http://www.waterboards.ca.gov/swamp/docs/reports/sparc486_swampreview.pdf

stream kilometers), diverse ecological settings (12 Level III Omernik ecoregions are present in California, Figure 3), and anthropogenic settings (vast regions of the state are entirely converted to either agricultural or urban land uses). There are many natural gradients within each ecoregion. For example, the elevation in the Southern California Coastal Ecoregion extends from sea level to 8,000 feet encompassing cold water, high gradient mountain streams, but also includes warm water, low gradient streams in the flood plain. To complicate matters further, there are extreme natural temporal cycles of dry and wet years, which may not occur in all regions of the state during the same year. This is compounded by the episodic natural disturbance of flooding and fires. Finally, human-dominated landscapes can be so pervasive in locations such as urban southern California and the agriculturally dominated Central Valley that no undisturbed reference sites may currently exist in these regions. A statewide framework for consistent selection of reference sites must account for this complexity.

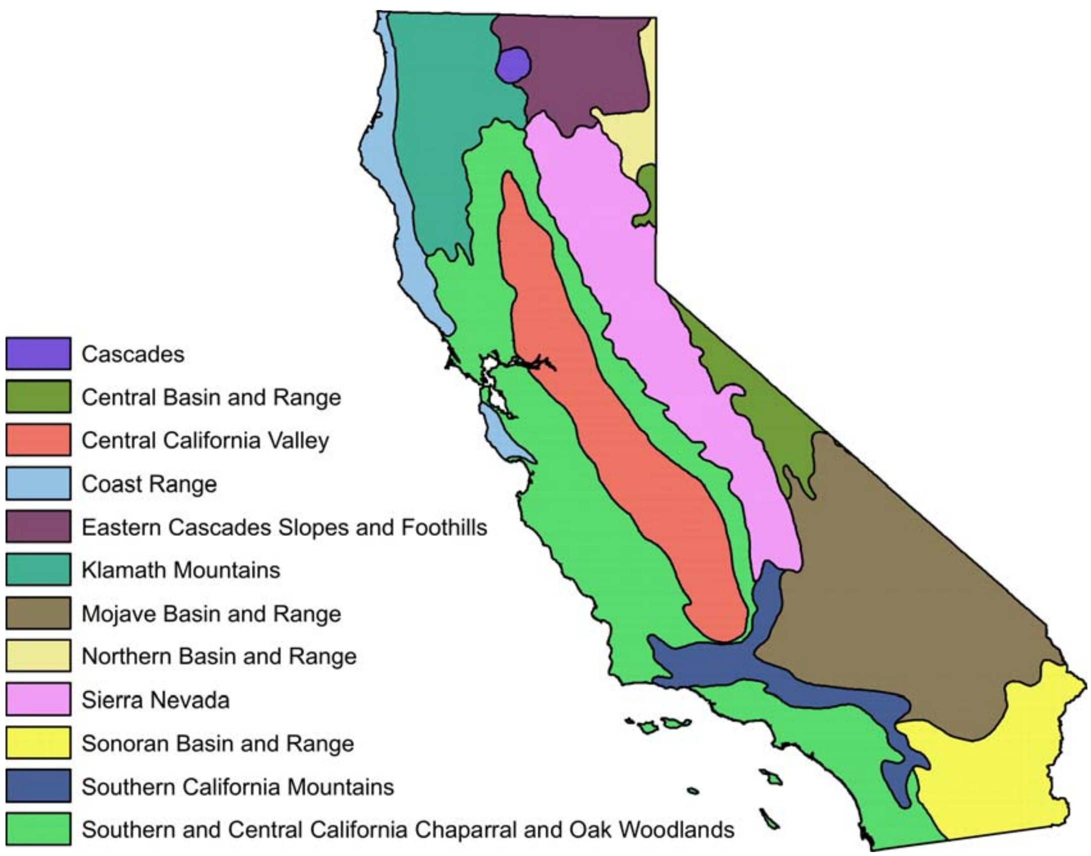


Figure 3. Boundaries of 12 Level III Omernik ecoregions present in California.

GOALS AND OBJECTIVES

This document summarizes recommendations to SWAMP for the development and maintenance of a Reference Condition Management Program (RCMP) that will support its regulatory biological assessment programs. The goal of the SWAMP RCMP is to provide an objective system for defining the expected biological and physical condition for wadeable streams and rivers in California. This system will identify pools (populations) of verified reference sites and outline procedures for sampling them to determine the range of biological expectations in these pools.

The monitoring objective

Data collected from reference sites will be used to answer a primary question: "what is the expected natural composition of lotic freshwater organisms in each of the major biogeographical regions of California"? The answer needs to be determined with sufficient rigor to serve as the basis for setting defensible numeric biocriteria. Our primary focus is on establishing expectations for benthic macroinvertebrate assemblages in perennial wadeable streams, but we expect that the approach will allow similar assessments of algal and fish assemblages as well as instream habitat condition and riparian condition.

Accounting for natural variability

An extension of the central monitoring question is: "what is the range of biotic measures (e.g., taxonomic composition, individual metrics and biological indices) in high quality sites and which natural environmental gradients (both spatial and temporal) are most strongly related to this variation." Ultimately, the goal is to identify the major sources of natural variability for all biological response measures (Figure 4). To account for these gradients, reference sites should be distributed to represent the full gradient range.

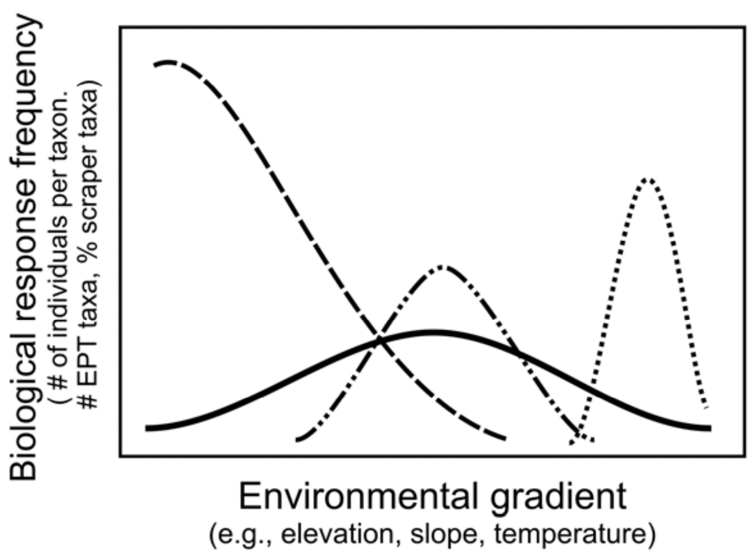


Figure 4. Hypothetical frequency distribution relationships between biological responses and environmental gradients.

GUIDING PHILOSOPHIES

In order to guide the development of the RCMP, the panel agreed upon a set of basic philosophies. These philosophical principals were used to guide their decision-making:

- **Use natural condition as the desired state whenever possible -** The panel's goal was to identify sites in natural or near-natural conditions whenever possible. However, the panel recognized that there are regions in the state where an insufficient number of sites in near-natural condition were likely to be found. The panel agreed that setting biological expectations were no less important in these regions. Therefore, the panel endeavored to identify the best attainable condition in these suboptimal regions of the state.

- **Balancing statewide consistency with regional flexibility** - The panel agreed that the reference strategy should balance a set of desirable, but sometimes naturally conflicting, traits: objectivity, consistency and flexibility. For example, a reference program that works for all of California can't be both perfectly consistent and perfectly flexible. This strategy aims to balance the competing demands of statewide consistency with the flexibility needed to adapt to unique regional conditions.

- **Reference site management is an iterative process -** The management of a reference site network is an ongoing and iterative process. The monitoring program should be responsive to new information and perspectives gained from selecting and monitoring reference sites. The general strategy should build in analysis of data to optimize selection strategies (process of selecting sites) and management design (e.g., how many sites, regional boundaries, which natural gradients to account for).

- **The RCMP should be transparent -** The technical process of determining reference conditions should be transparent to external review. As the state moves toward implementation of biocriteria, transparency and comprehension of the RCMP process will improve stakeholder confidence and provide structure for discussions about setting objective standards.

- **These recommendations are a starting point -** The panel understood that their recommendations provide a starting point for evaluating reference condition rather than an exhaustive set of operating procedures for selecting reference sites. This document is written assuming that SWAMP will develop a technical workplan that details a more refined program as the RCMP is implemented.

GENERAL GUIDANCE

The general approach for establishing the SWAMP reference site network has four components:

1. California should be divided into different geographic regions based on coarse biogeographic similarities in order to partition some of the natural variability among regions
2. A pool of reference sites should be assembled within each region through a sequential process of identification and screening of candidate sites
3. The reference pools should be managed through iterative review of data to refine regional boundaries, ensure continued suitability of sites and ensure adequate representation of natural gradients
4. A monitoring design should be created for sampling this pool of reference sites to document the range of biological and physical condition at reference sites, and monitor for changes to the condition of reference sites over time

All but the second component, site selection, apply equally to all regions of the state. The site selection process has two versions depending on the availability of high quality reference sites. We refer to the two versions in this document as: 1) the "standard model", which applies to regions with a sufficient number of reaches with relatively low levels of anthropogenic stress; and 2) the "alternate model", which applies to regions that do not have a sufficient number of high quality reaches. The vast majority of California should be able to apply the standard model.

Component I: Partitioning CA into biogeographic regions

Two general schemes are available for delineating California's ecoregions (Omernik 1995 and Bailey *et al.* 1994). We follow Omernik's divisions here because the boundary delineation decisions were generally based on a broader range of geology, climate and zoogeography than Bailey's. Omernik Level III ecoregions have been delineated for all of North America (Omernik 1995), with 12 Level III ecoregions falling in California (Figure 3).

Partitioning the state into different regions based on habitat similarities has some precedence in California bioassessment. The SWAMP Perennial Streams Assessment (PSA) has relied on a combination of Omernik ecoregions and regional board boundaries to partition the state for assessment purposes (Figure 5). Because these definitions include significant ecological gradients that contribute to natural variability in biological assemblages, and because they comprised existing assessment units, the panel agreed that these delineations were appropriate to use as initial boundaries for the reference network. However, the panel also stressed that ecoregions do not always adequately capture natural gradients that are key drivers of aquatic assemblages (insert references here, Hawkins and Norris 2000). Thus, data analyses must address the suitability of these boundaries as the program collects more data.

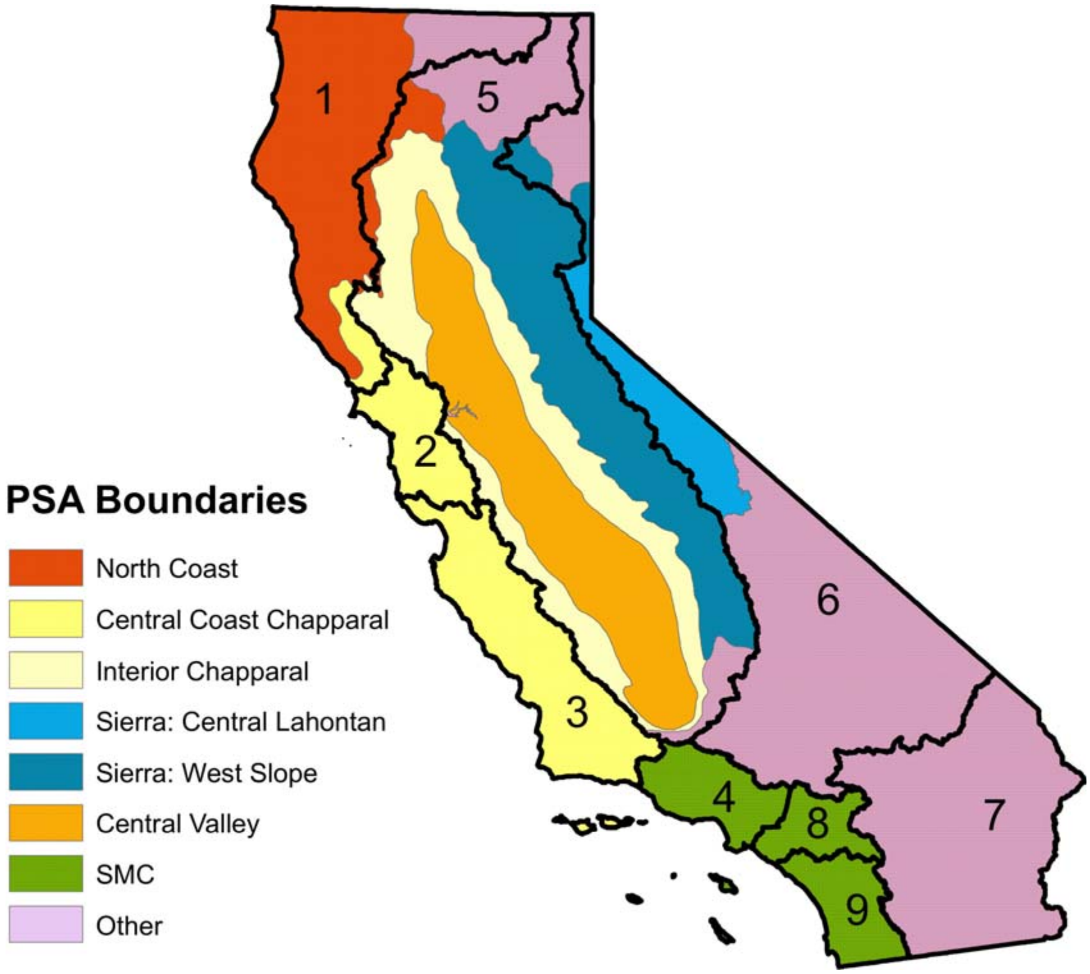


Figure 5. Boundaries used for defining the regional subunits of the SWAMP Perennial Stream Assessment (PSA) survey. SWAMP regional board boundaries one through nine are indicated by thick lines. SMC=Southern Coastal California Stormwater Monitoring Council.

Component II (a): Selecting sites: the "standard model"

The second step in the general approach is the most resource intensive and technically challenging: to develop a large pool of reference sites within each ecoregion. The ability to precisely establish biological expectations within each region is a function of the number of sites that are sampled and natural variability within each region. Therefore, the pool of reference sites should be large enough to provide a robust characterization of natural variability. Furthermore, reliance on a small number of reference sites is risky because it increases the consequences of catastrophic failure of individual sites. The size of the site pool in each region will depend on the number of major environmental gradients in each region (e.g., elevation, temperature, etc.) and the strength of influence of these gradients on biotic assemblages.

The panel recommended a sequential approach for assembling a set of candidate reference sites and screening suitable sites for the final reference pools within each region. The process includes: 1) screening data from previous site visits to identify candidate sites, 2) application of remote sensing and point-source GIS data screens of all potential stream reaches (combining landscape and local scale) to identify candidate sites, 3) use of best professional judgment/ local knowledge to add sites to the candidate pool.

Once a set of candidate sites is assembled, each candidate site should receive an on-site visit to evaluate its suitability. The exact type of data collected for evaluations during this stage will vary by region, but at a minimum should include: observations of local landuse activities, instream and channel habitat condition, riparian condition, evidence of recent natural disturbance. Some regions may require additional chemical data (water column or sediment) or toxicological data to confirm site suitability.

Sites that pass both the remote sensing and field reconnaissance screens become part of the reference pool for that region.

Component II (b): Selecting sites: the "alternate model"

The panel recognized that the standard model is not likely to work in all regions of California. The conversion of natural landscapes to agricultural and urban land uses is so extensive in some parts of the state that the entire region is devoid of waterbodies that could be used to define reference condition. Most regions of California should be able to use the standard model; the alternate model should only be used when the standard model is not feasible.

The panel defined the following criteria as triggers for acceptable use of alternate site selection strategies (both criteria must apply):

1) Insufficient high quality sites are available within one of the main regions (or a large section of one of the main regions) to adequately characterize ecological potential. Suitable stream reaches are unavailable for one or more of the following reasons:
    a) Anthropogenic landuse is a dominant factor in all watersheds within the region (or subregion)
    b) Normal flow is modified (e.g., flow diversions, dams, withdrawal or augmentation)
    c) Natural channels are altered (e.g., all or most channels converted to conveyances, irrigation supply/drains)
    d) Riparian corridors are impacted throughout the region (e.g., concretized riparian or surrounding landscape modified)

2) No comparable region exists from which to draw inference about biological expectations. That is, the areas are unique in their biological expectation so regions with few reference sites are not able to incorporate sites from another region.

This situation is not unique to California streams and many large programs have recognized the need to deal with regions with insufficient reference sites (REFCOND 2003, Stoddard *et al*. 2005, Paulsen *et al*. 2006). National guidance for developing state biocriteria programs highlighted the need for special treatment of these conditions (Barbour *et al*. 1996a,b). While the unique needs of these regions are widely recognized, the approaches for establishing ecological potential for reference-poor regions are far from standardized.

The RCMP panel outlined a general strategy for approaches to explore in reference-poor regions. The RCMP panel did not take any strong position on the relative strengths of these alternatives nor how different approaches should be combined to define expected conditions in reference-poor regions. Some of the alternative strategies included:

1. Use a modified version of the standard approach (e.g., use lower thresholds, emphasize local condition measures)
2. Alternate approaches
    a. Use existing tools to screen sites
    b. Species pool approach
    c. Factor-ceiling approach
    d. Model taxon preferences for key environmental gradients

These alternative strategies are not mutually exclusive and, when appropriate, should be used as multiple lines of evidence to reinforce an objective definition of biological expectation in regions without reference sites. In the "specific guidance" sections of this document (see Alternate Strategies for Selecting Sites) we describe these approaches and discuss strategies for applying them to California's challenging landscapes.

Component III: Managing the regional site pools

After the site pools have been assembled for each region, the RCMP requires an ongoing evaluation of data from these sites to address several key management questions. There are two major components to managing the reference pools: 1) evaluation of the regional representation of natural gradients and 2) periodic review of sites to evaluate changes to their suitability.

The ability to effectively understand natural sources of biological variation is fundamental to establishing sound biocriteria[5]. Therefore, the RCMP must directly assess the reference pools to ensure representation of regionally important natural gradients. This review should include a periodic review of the suitability of the initial regional boundaries proposed here.

The second aspect to site management is periodic review of sites in the reference pools to assess their continued suitability as reference sites. Conditions within stream reaches and in their upstream drainages can change over time (e.g., timber harvest, conversion of natural landscapes to agricultural or urban/suburban/exurban uses). Furthermore, we may discover sources of stress that were unknown when sites were initially added to the reference pools (e.g., discovery of nonpoint source discharges, mines, flow withdrawals/diversions, small-scale placer mining, etc.). Sites that fall into this category may be monitored to measure the impacts of these stressors, but they should be removed

[5] See discussion on p. 5.

from the reference site pools. In contrast, natural disturbances (e.g., forest fires, catastrophic flooding or landslides) can also alter the biological condition at sites and they should be excluded for sampling temporarily, but should remain in the reference site pool[6].

Component IV: The monitoring strategy

The panel recommended an integrated probabilistic and targeted sampling design for the RCMP. The probabilistic approach will sample a rotating subset of randomly-selected (rotating panel design) sites from within the reference pool each year to estimate average biological condition. A subset of the randomly-selected sites should be sampled annually to measure year-to-year variability at sites and improve SWAMP's ability to detect drift in reference condition within each region over time. This design provides an unbiased assessment of natural variability with enhanced trend detection.

Targeted sampling is comprised of fixed sites near locations of special interest, but this should be supplemental to the probabilistic sampling effort. Fixed sites provide additional power to detect trends, but suffer from its inability to extrapolate to other locations. However, many agencies already monitor reference sites and, provided they meet the RCMP selection criteria, these sites have the added benefit of years of historical data. As SWAMP extends its reference monitoring program through collaboration with other state and federal programs, it should retain the ability to incorporate these sites.

The panel emphasized sampling more probabilistically selected sites over targeted sites, but did not make any recommendations about relative proportion of each type. This decision should reflect the relative importance to the SWAMP program of estimating current biological expectation versus detecting changes in the reference state. Changes in the reference state may become increasingly important due to factors such as climate change.

[6] A special study of natural disturbance recovery could be especially enlightening with regard to understanding natural variation.

SPECIFIC GUIDANCE

1.0 Site Selection: Assembling the reference candidate pool

The panel recommended a sequential approach for assembling a pool of potential reference sites using a series of tools to identify candidate sites (Figure 6). The toolbox components included: 1) use of existing data from previous site visits, 2) GIS data screens of all potential stream reaches using databases of stressor data (combining landscape and local scale), 3) expert selection of site locations based on regional experience.

1.1 Use of existing sites

Previously sampled sites are an excellent source of candidate reference sites and where available in sufficient numbers, can constitute a ready-made pool of reference sites. However, previously sampled sites vary widely in the amount of information associated with them, and they fall into two categories: 1) sites with a large amount of associated environmental data that is sufficient to evaluate without additional data collection, 2) sites that require additional data collection to produce adequate evaluations. Several programs in the state have collected sufficient data to meet the first condition (e.g., EMAP, Central Valley WEMAP, CMAP, SNARL, some regional board programs), but most sampled sites fall into the second class.

The current distribution of existing candidate sites in California is illustrated in Figure 7. Sites were pre-screened from ABL and SNARL databases and sorted into one of three tiers based on the availability of different types of screening data. Under the RCMP, Tier 1 sites would pass to the pool of verified reference sites if they passed a BPJ screen (see following section), sites in other tiers would be placed in the candidate pool and be subjected to the full site screening process (Figure 6).

1.2. GIS data screens of all potential stream reaches using databases of stressor data [7]

If regions do not have sufficient existing sites to fill the final pool of fully screened reference sites (steps 1 - 3 of the general guidance), then new candidate sites should be identified through use of geographic information systems (GIS) techniques for screening remote sensing data and GIS databases of point source stressors. GIS-based searches for candidate reaches are expected to contribute the majority of sites in many regions.

Ode (2002) described a GIS based method for identifying candidate stream reaches using a series of remote sensing data filters. Under this approach, candidate watersheds are identified for a region with GIS techniques and then stream reaches within these watersheds are targeted for reconnaissance to verify reference quality characteristics. The RCMP generally follows this approach, which consists of the following steps:

[7] GIS techniques are used at two different stages of the RCMP process: 1) searching for potential new reference streams (described in this section) and 2) quantifying impacts to existing sites (described in the following section). The techniques are very similar, but differ somewhat in their application. The search phase is a relatively coarse screen of candidate watersheds while the verification phase is site specific and allows for multiple spatial scales of GIS analysis (see Figure 8).

1.2.1 Assemble GIS layers of important landuse disturbances
The list of potential impacts to stream condition is very long and includes multiple point and non-point sources of disturbance. Quantitative measures of many human or human-influenced activities are available in digital spatial (GIS) formats from various state and federal agencies (see Tables 1 and 2), but there is a very large amount of variation in the degree to which datasets are accurate, current, and consistent across wide geographical ranges.

1.2.2 Determine appropriate reporting units (areas of analysis) and create necessary GIS layers~ Current GIS applications for locating least disturbed waterbodies in a region (see ATtILA text box) calculate summary stressor metrics (e.g., % urban landuse, road density) for each reporting unit (typically watersheds) in the region of interest. Candidate stream sites are then selected from within these watershed areas. It is recommended that the RCMP use a modified version of watershed polygons developed by the national NHD+ program.[8]

1.2.3 Use ATtILA extension to calculate stressor metrics using remote sensing and point source datasets (see ATtILA text box)~ ATtILA produces summary output in a spreadsheet containing multiple stressor metrics for each candidate watershed (i.e., % agricultural landuse, % impervious surface, # of mines, # road crossings/stream km).

1.2.4 Analyze distribution of stressor metrics and select appropriate thresholds
Screening thresholds for GIS stressor metrics can be set using a variety of approaches: 1) visual inspection of frequency histograms for natural breaks in distributions, 2) statistical criteria[9] (e.g., eliminate watersheds with road densities greater than 1.5 standard deviations above the mean for all watersheds in the region, or eliminate all but the lowest 25th percentile of all road densities), 3) established (i.e., literature based) impact thresholds. At this stage in the screening process, the RCMP panel recommended the use of fairly liberal screening thresholds since GIS data are often inexact and impacted sites can be screened during later stages of the site verification process.

1.2.5 Eliminate watersheds that fail GIS screens
Because of the large number of stressor variables that are quantified in this step, there will be a large number of metrics to evaluate. The panel discussed two options for how to combine the information from these different screens:

---

[8] With funding from the SWAMP program, CSU Chico's Geographic Information Center (GIC) has developed a method for creating nested watersheds from the native polygons available from the NHD+ program. The NHD+ polygons are limited in their utility as reporting units because they are non-overlapping. Thus, 2nd order watershed boundaries in NHD+ do not include their tributary 1st order basins. The GIC's modification creates new watershed polygons that are aggregates of all upstream polygons (e.g., 4th order watersheds contain all upstream 3rd, 2nd and 1st order polygons).

[9] Effective of statistical properties of distributions to define thresholds depends on a normal distribution of scores. Some distributions (e.g. highly skewed or bimodal) may be better interpreted by looking for natural breaks or using literature based criteria.

a) Screens could be applied as a series of filters, with failure in any metric resulting in elimination of the watershed from the candidate pool.

b) Alternately, a multi-metric index of stressors could be used to create a composite score for each candidate site and low scoring watersheds would be removed from the candidate pool.

The panel recommended the use of a hybrid approach, in which the multi-metric scoring would be used to screen watersheds, but "kill-switches" would be employed to eliminate watersheds that exceeded high impact thresholds for particular stressors (e.g., eliminate watersheds with > 10% urban landuse).

As an additional consideration, the panel recommended that the RCMP explore quantitative methods for deciding which impacts to use for selection. For example, some stressors may have a greater effect than others and, thus, should be weighted more heavily than relatively benign influences. A corollary would apply to data sets with different levels of confidence. For example, information about mine locations may be available, but not about which are actively contributing contaminants to streams.

ATtILA extension for GIS Landscape Analysis

http://www.epa.gov/esd/land-sci/attila/intro.htm

To quantify landuse activities occurring upstream of sites, the Ebert and Wade (2004) developed a user friendly interface that accepts a range of GIS data layers and produces summary statistics for areas defined by the user. The extension, Analytical Tools Interface for Landscape Analysis (ATtILA), is a plugin to ESRI's ArcView® (version 3.x) GIS software (ESRI Products) and takes advantage of ESRI's Spatial Analyst extension to run the spatial calculations.

- The ATtILA extension calculates the percentages of various landuse activities occurring in specified areas (urban; forested; agricultural-row crops; agricultural- orchards/vineyards; agricultural-total), other correlated measures of human activity (population density; road length; road density; road crossings/stream mile; percent impervious surface), and estimated nitrogen and phosphorus loadings.
- ATtILA can use polygons of any spatial extent as reporting units (e.g., entire upstream basin, local buffers)
- In 2007, the SWAMP program provided funds for a project to adapt the ATtILA extension to meet the GIS needs the RCMP process. Specific enhancements being developed include the ability to add custom stressor coverages, summarize point source data, and facilitate rapid adjustment of stressor thresholds for screening candidate sites. The project will be coordinated with the implementation of the RCMP
- It is expected that the capabilities of the modified ATtILA extension will expand as the RCMP process develops over time.

1.2.6 Identify candidate stream reaches within candidate watersheds[10]

After eliminating watersheds using GIS screens, the remaining watersheds represent potential candidates for the reference pool. These areas may be able to be further refined to further isolate candidate stream reaches (see Figure 8).

[10] An alternative strategy is to select candidate stream segments directly using analytical tools designed to work with the NHD+ datasets. Under this approach, confluence points would be the the reporting unit and NHD+ tools would summarize all upstream landuses. Errors in the current version of NHD+ (primarily problems with flowline connectivity) currently limit the effectiveness of this approach, but it may become more useful as NHD+ improves. The RCMP should remain open to both approaches and revisit this issue as new versions of NHD are released.

1.3 Use of local knowledge to add sites to the candidate pool

Although existing data and GIS searches will contribute the majority of sites to the candidate pool, a few sites may be added to the candidate pool on the basis of local knowledge. Local knowledge can sometimes help in identifying candidate sites because GIS datasets are imperfect and GIS screens may pass over good sites because of inaccurate or outdated disturbance information. These sites, however, should be critically evaluated because subpar sites based on local knowledge will dilute the quality of the reference pool. More rigorous evaluation of these sites should include examination of existing data.

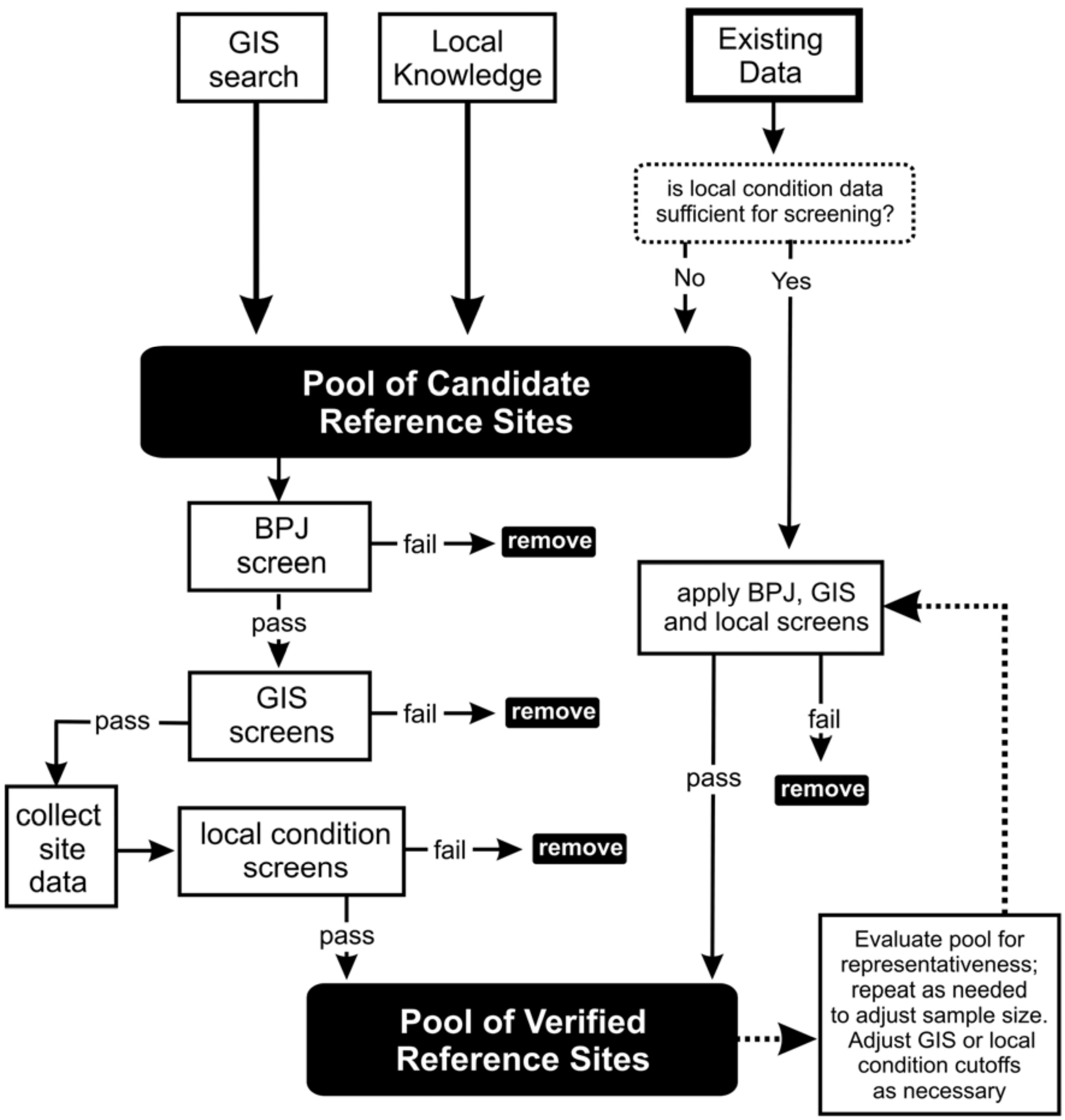


Figure 6. Schematic of the standard reference site selection and verification process.

Table 1. Potential GIS data coverages for nonpoint sources.

NON POINT-SOURCE COVERAGES			
Information Type	**Data Source(s)**	**Notes**	**Coverage**
Landuse/Landcover	National Landcover Dataset (NLCD), MRLC	1992, 2001 satellite imagery, allows for 9-yr landcover change assessments	Statewide
Impervious Surface	NLCD, Others	Quality varies regionally	NLCD statewide, others patchy
Road Density	USFS, TIGER		Statewide, but patchy
Timber Harvest	CDF, THPs		
Vegetative Change/ Vegetative Change Cause (LCMMP)	USFS/CDF		Not Statewide
Population Density	Census Blocks, CDF	Produced in conjunction with decadal population censuses; censuses can be combined to estimate population change	Statewide
Grazing	Cattlemen's Association		Not Statewide
Fire History	CDF, USFS		Best for FS lands

Table 2. Potential GIS data coverages for point sources.

POINT-SOURCE COVERAGES			
Information Type	**Data Source(s)**	**Notes**	**Coverage**
Mining	USGS	Possibly outdated	Statewide
NPDES	EPA	Prone to inaccuracies	Statewide
303(d) listed streams	SWRCB	Every three years	Statewide
Water Diversions/ Extractions	USGS, NHD+	Possibly outdated	Statewide
Dams	CalWater	Doesn't include overflow info	Statewide
Stormwater Inputs	NHD+, Counties	Uneven coverages	Patchy
POTW	EPA	Prone to inaccuracies	Statewide
Landslide Datasets	CalTrans		Statewide

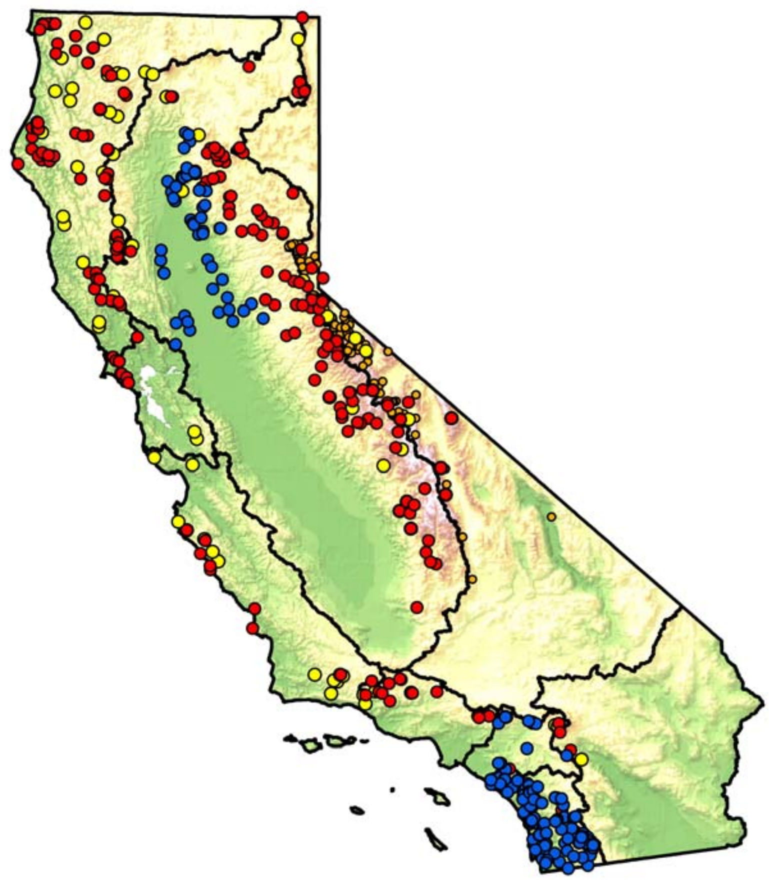


Figure 7. Partial set of bioassessment sites available for initial screens assigned to one of three tiers. Tier 1 sites (yellow circles) are EMAP and CMAP sites that passed a full suite of screens based on the most complete data for evaluation. Chemical and habitat thresholds were based on Stoddard *et al*. (2005) and landuse thresholds were based on Ode *et al*. (2005) and Rehn *et al*. (2005). Tier 2 sites (red circles) are USFS and Regional Water Board sites that have passed a less stringent screening process, but might very well be reference and need additional data before they either passed into Tier 1 or eliminated from the candidate pool. Tier 2 sites were screened based on land use, less extensive physical habitat data and limited or no chemical data. Tier 3 sites (blue circles) are cases in the Sacramento Valley, Sierra Nevada foothills and southern coastal California that probably need an alternative reference screening process (e.g., the factor ceiling approach). SNARL sites (orange circles in Eastern Sierra Nevada) used different screening thresholds, but are likely equivalent to Tier 1 sites.

2.0 Site Selection: Screening the candidate pool

Once a large set of sites is selected for the candidate pool, sites in the pool undergo a series of screening steps to either validate sites as appropriate reference sites or eliminate them from the pool. The major screening tools are: 1) expert opinion (BPJ), 2) landscape screens (GIS), and 3) local condition screens.

2.1 BPJ screens

While BPJ can play a role in identification supplementing the pool of candidate sites, it plays a bigger role in eliminating candidate sites. Sites should be eliminated on the basis of BPJ knowledge that there are known problems that aren't accounted for in GIS datasets. For example, GIS datasets may miss recent development, known pollutant spills, or nonpoint sources. This step should include coordination with local watershed groups, landowner groups and other stakeholders to eliminate inappropriate sites. The rationale for rejection should be documented.

2.2 Landscape scale screens (GIS)

Just as GIS techniques are essential for adding sites to the candidate pool (Figure 6), they also play a crucial role in reference site verification. The datasets and techniques used in this step are essentially the same as those used in searching for candidate watersheds/ stream segments, but the application of the tools differs somewhat. Whereas the GIS analyses were applied at a fairly coarse spatial scale in Section 1.2, GIS tools can be applied at multiple spatial scales during the screening stage.

The first step in the second GIS stage is to convert candidate watershed areas into specific sampling sites by selecting a common point on the stream segments in each watershed (e.g., the downstream confluence point), making them equivalent to other sites in the candidate pool (as in Figure 8a).

The chief benefit to the two-stage application of GIS techniques is that it gives us the opportunity to identify multiple sampling locations within reference watersheds.
While sites would normally be screened using stream confluence points as the candidate site locations, site locations could be moved to other points in the watersheds to identify additional reference sites within good watersheds or to avoid portions of the watershed with undesirable sources of human disturbance (Figure 8b).[11]

Using watershed delineation tools and local site buffering tools currently available for use with GIS software, polygons should be created to represent different spatial scales upstream of each site (e.g., the entire watershed draining to the site, the upstream area within a 5 km radius of the site, the area within a 200m buffer on either side of the stream within 1km upstream). Once created, these areas can be used as reporting units for

[11] Although the two stage application of GIS techniques gives us greater flexibility to identify multiple candidate stream reaches within each candidate watershed, an alternative strategy would be to eliminate the coarse search for watershed described in Section 1.2 and go straight to the more refined screening analysis indicated in Figure 8a.

ATtILA analyses. Metrics calculated for the different spatial scales can be screened as in Section 1.2.5.

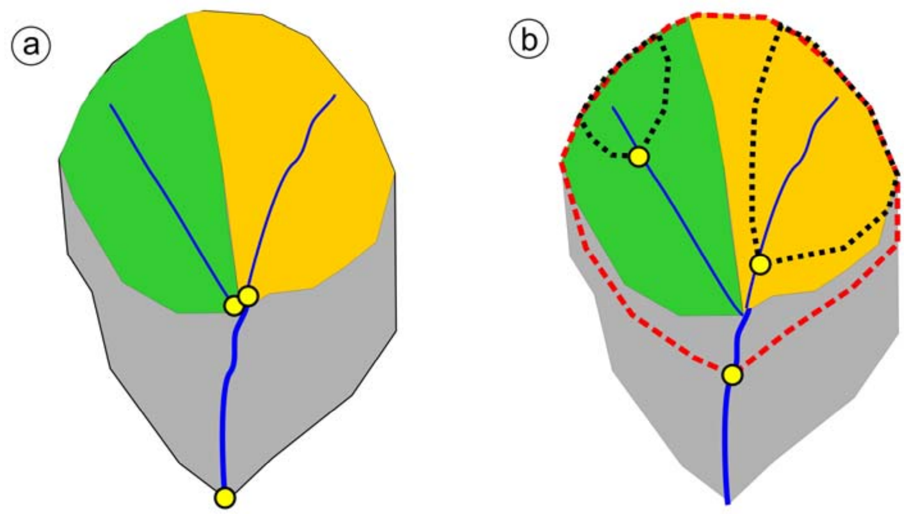


Figure 8. Illustration of alternative applications of the second stages of GIS analysis in the RCMP using a hypothetical second order watershed containing two first order watersheds: a) normal site locations represented by yellow circles, b) alternate site locations and their watershed boundaries (represented by dotted lines).

2.3 Local Condition Screens

Sites that have passed BPJ and GIS screens are then subjected to an evaluation of site scale stressors. Some of the local scale information can be obtained from aerial photography of sites, but the majority of this information will come from site visits and in some cases collection of water quality data.

2.3.1 Site scale data: Aerial photography

Aerial photography provides a unique view of potential site scale stressors. Digital orthophoto quadrangles (DOQs) are available for the entire state of California (DFG). Google Earth is another source of digital satellite imagery. DOQs and other sources of aerial photographic images can provide excellent information about local stressors not available through other sources, but are subject to the same timeframe limitations as other digital sources.

2.3.2 Site scale data: Site visits

The panel strongly recommended site visits as a crucial component of reference site verification. Once candidate list have been narrowed down to sites that meet BPJ, GIS and DOQ screens, land ownership should be determined for each site and owners contacted to obtain access permission and or sampling permits as needed. Site owners can also be contacted at this point to determine if there are any reasons for rejecting sites.

Field visits should be used to collect both qualitative (e.g., presence of obvious disturbances) and quantitative data (e.g., % intact riparian zone). Quantitative measures should focus on data that can be collected and analyzed cost-effectively.

2.3.3 Qualitative data

Visual assessments of site suitability should include a minimum set of observations:

- Upstream impoundments, or evidence of water withdrawal or diversion
- Evidence that the site is non-perennial
- Evidence of recent fire, flooding or landslides
- Local grazing impacts
- Presence of significant anthropogenic use (e.g., campgrounds, etc.)

2.3.4 Quantitative data

At a minimum, site visits should include characterization of physical habitat using the SWAMP Physical Habitat Procedures (Ode 2007) and conventional water chemistry. Physical habitat characteristics should include measures of both instream and riparian condition. SWAMP habitat procedures may be supplemented with riparian condition measures collected with the California Rapid Assessment Method (CRAM) for riverine wetlands. Water chemistry analyses should include the following analytes: chloride, turbidity, pH, total nitrogen, total phosphorus, conductivity, and alkalinity. Some chemical analytes may not be needed in all regions. For example, sulfate (a good indicator of mining activity) is not likely to be informative in xeric regions. One recommendation was to create a checklist of activities by region. Another option is to supplement with sediment and/or water column toxicity. While these tests may be expensive, they are less expensive than a screen for a long list of toxic constituents.

2.3.5 Combining site data for screening decisions

As with GIS screens (Section 1.2.5), there are many ways to combine site data to make determinations. The panel again recommended use of a hybrid approach in which site scale data is combined to calculate a multi-metric site condition score. The use of kill switches was also recommended for excessively high or low scores for individual habitat or chemistry.

3.0 Alternate strategies for selecting reference sites

While most regions of California can follow the standard approach for selecting reference sites, there are at least two large regions in California that lack sufficient high quality sites. The first is the Central Valley where natural landscapes have been almost entirely converted to agricultural and urban land uses. Most natural stream reaches in this region have been channelized or otherwise modified to support irrigation and flood control. The second is in coastal southern California (elevations below 1200 ft – upper elevations can follow standard model) where conversion to urban and suburban land uses has led to the channelization of most stream reaches. Recent studies in these regions demonstrate that at least some waterbodies in highly modified regions can support fairly rich BMI assemblages, even under considerable alteration and agricultural development (Griffith *et* *al*. 2003, deVlaming *et al*. 2004, deVlaming *et al*. 2005, Ode *et al*. 2005). Thus, there is enough range in biotic condition to differentiate degrees of impairment in these regions.

The panel recognized the unique limitations of these regions and recommended that a separate set of approaches be developed for them. Despite the differences in methodology, the goal of the alternate strategy is the same as the standard approach: to characterize the best attainable biological condition in these regions. This section outlines a set of approaches that the RCMP could follow. These fall into two general categories:

- Use a modified version of the standard approach
- Explore non-standard approaches

3.1 Modified use of standard approach

The first option is to use the set of techniques described for the standard approach, but to modify the way the techniques are applied. Modifications fall into two general types: 1) much greater emphasis on reach scale screening data, 2) use of less stringent criteria for rejecting sites.

One of the panel's philosophies is that potential reference sites in highly modified regions need a much larger amount of supporting data to verify their status than in less modified regions. In both the Central Valley and southern coastal California lowlands, streams exist in a landscape matrix with a universally high level of unnatural land uses. Furthermore, many streams have extensive flow manipulation, including water diversion, re-introduction, and inter-basin transfers that render watershed based tools irrelevant. For both these reasons, watershed based stressor analyses are less informative screening tools. Accordingly, much greater reliance should be placed on data collected from direct site visits than on remote sensing data. The panel recommended increased emphasis on riparian condition, instream habitat condition, and water column chemistry. In some cases, additional data (e.g., sediment and or water column toxicity) will be necessary to verify sites.

Selective relaxation of screening thresholds may also be an effective means of identifying the best available sites in a region. For example, acceptable road densities are likely to be much higher in southern coastal California than in other regions of the state. Likewise, acceptable local agricultural landuse percentages and acceptable levels of fine sediments are likely to be higher in the Central Valley than in less modified regions. While less stringent thresholds may help identify some of the best sites in highly modified regions, the use of kill switches is an essential safeguard against accepting unacceptably low thresholds. Specific cutoffs such as >10% local impervious surface, or toxin concentrations greater than the standards set by the California Toxic Rule may be more appropriate in these heavily modified landscapes.

A version of this modified standard approach was applied to search for reference sites in the Central Valley (Ode *et al*. 2005). Remote sensing data (e.g., landuse percentages) and other GIS datasets (e.g., pesticide application rates) was used as a coarse screening tool, but this data was de-emphasized in favor of riparian condition and instream habitat scores. This study identified approximately 20 potential reference creeks in the Sacramento Valley (see Figure 7), but these still need to be screened for water chemistry and toxicity before they are acceptable.

3.2 Non-standard approaches

Although modified use of the standard techniques can go a long way toward providing the data needed to adequately characterize biological expectations in these areas, it is unlikely to resolve the entire problem of identifying a sufficient number of candidate reference sites. The panel recommended the exploration of several different alternative, non-standard techniques:

- Select best sites using existing biological indices
- Species pool approach
- Factor-ceiling approach
- Model taxon preferences for limiting environmental gradients

All of the non-standard strategies suffer to a greater or lesser degree from circularity since the establishment of a biological reference site is being established with biological data. However, the extreme lack of reference sites in these regions requires us to consider accepting some circularity while adding additional steps to guard against the risks of circularity. The best way to guard against these risks is to use independent datasets to select the biotic response metrics.[12]

3.2.1 Use of existing indices to select sites with high quality biology

A straightforward alternate approach is to use existing biological assessment tools from the same region to identify sites that could be used to establish biological expectation in problem regions. [13] High scoring sites would be assumed to represent the "least disturbed" sites in the region. The method assumes that BMI assemblages in the target region have similar responses to anthropogenic stress as the region(s) for which the indices were created. Issues with circularity are mitigated by the fact that the scoring tools were derived objectively using independent datasets.

A variation on this approach is possible in regions where only a few reference sites can be identified (either using the standard methods or the modified standard described above). Under this variation, a model (either MMI or O/E) would be created using a small number of reference sites. Then new sites with similar BMI assemblages would be added to the reference pool and the model recalculated. This recursive approach results in more explanatory power because it is based on a larger number of reference sites, but it is inherently circular because the new sites are not chosen based on independent information.

---

[12] Note also that some have argued that the circularity concern is less of a problem in highly modified systems than more pristine systems because relationships between metrics and stressors are simpler (Karr and Chu 1999).

[13] Examples of existing biological assessment tools include the Southern California IBI (Ode *et al*. 2005), northern California IBI (Rehn *et al*. 2005) and the California RIVPACS models (Hawkins unpublished).

3.2.2 Species pool

Another option is the species pool approach, which uses the total faunal diversity of a region (i.e., central valley or southern California coastal urban lowlands) to establish a biological condition axis. The process involves assembling a pool of all BMI taxa ever collected from the region, then using taxonomic richness as the measure of biological integrity at test sites. The inventory could be compiled from existing data sets, historical records (i.e., museums or other voucher collections), or directed field surveys. This technique assumes that richness is a good measure of condition, that there hasn't been extensive extinction of native fauna and that the constituent species in the pool are all potential colonists of any test stream.

The utility of this approach could be enhanced in at least two ways. The number of richness metrics could be increased by breaking richness out by taxonomic groups (midges, worms, mayfies, etc.), isolating the different information content in these groups. Further, the species pool could be modeled to associate expected taxa with key environmental gradients (i.e., substrate composition, elevation, etc.) and the proportion of taxa present at reference sites could be a potential target for attainment of reference state. If this approach were taken, then the species pool concept should be tested first in a region where identifying reference sites are not problematic as proof of concept.

3.2.3 Factor-ceiling approach

Carter and Fend (2005) developed a technique for defining a range of biotic expectation that takes into account the decrease in biotic condition caused by physical modification along an axis of increasing urbanization. In their example, a simple statistical technique (partitioned least squares regression, OLS) was used to identify the highest biotic scores along an urbanization gradient. Upper values define the range of expected biotic conditions for the region. Since a full urbanization gradient was used to take into account decreasing biotic potential with increasing urbanization, the resulting range of expected conditions is a conservative estimate of biotic potential for the region. While this approach could be used in both the Central Valley and southern coastal California lowlands, the method would work especially well in the Central Valley because the agricultural impact gradient is not as strongly confounded by elevation or other longitudinal gradients as the urban ones studied by Carter and Fend (2005).

The first step is to identify key measures of physical modification (hydrologic modification, channel modification, streambed modification) and to combine these into a multifactor axis of agricultural modification (i.e., the primary axis in a PCA of these stressors). The second step would be to identify appropriate metrics for detecting biotic impairment in valley streams.

3.2.4 Modeling taxon preferences for limiting environmental gradients

The final alternate strategy involves modeling taxon preferences for key environmental gradients, or limiting environmental differences (LED) and then using these relationships to select the most appropriate sites for setting biological benchmarks. Different habitat features (e.g., climate, channel morphology, water chemistry, substrate characteristics) can be thought of as acting as "filters" that select for particular species traits (Poff 1997).

This conceptual framework provides a way of accounting for the influence of both natural and anthropogenic factors on species distributions. Chessman and others (Chessman 1995, 2006, Chessman and Royal 2004, Chessman *et al*. 2008) recently developed a technique for using the tolerance or preference of individual taxa for key environmental filters (e.g., water temperature range, substrate composition, flow regime) to predict the assemblage of taxa that could be expected to occur at any test site under minimal human stress. Deviation from that expectation is used to infer degradation just as it is in predictive models (e.g., RIVPACS).

This is a promising approach; even the primitive assignment of taxa to simple preference classes used by Chessman and Royal (2004) resulted in stronger associations between their water quality assessments and independent measures of human disturbance than did the Australian predictive models developed from reference sites. They achieved similar results when applying the technique to fish assemblages (Chessman *et al*. 2007).

To adapt this to California's heavily modified regions, there is a need to develop models of the environmental affinities of Central Valley and southern coastal California lowland BMI taxa. It will likely take several years to collect enough samples to characterize individual BMI responses across key environmental gradients, but some of this data has already been collected and could be worked with now.

3.3 Combining approaches

The alternatives described in this section are not mutually exclusive; the RCMP could use more than one in each region. It is possible that not all approaches will work equally well in all regions and, as a result, different alternatives might be used in different regions. The panel was silent on which approaches, or which combinations of approaches should be prioritized.

The panel cautioned that using these non-standard approaches would require significant effort. Since these non-standard approaches have been used sparingly elsewhere, and essentially not at all in California, pilot studies looking into their applicability was recommended. The first step in the panel's recommendation was to evaluate existing datasets to determine if historical data exists for implementing any of these approaches. As mentioned in section 3.2.2, these approaches should be tested in a location where reference sites exist. Developing any non-standard approach needs to be ground-truthed before widespread use of the tool should be applied. Once this proof-of-concept occurs, then targeted data collection in one of the reference-poor regions can be initiated.

MANAGING THE REFERENCE POOLS

Accounting for natural variation

Classification of streams according to natural gradients can help partition natural sources of variation in biological assemblages and thereby improve our ability to detect deviation from reference condition (see Hughes 1995 for a review of the history of stream classifications). The RCMP needs to ensure that the regional reference site pools are representative of the most important regional gradients. The best way to test the representation of these gradients is through ordination of BMI datasets to determine which natural gradients explain most BMI variation in each region. Assessment of natural variation should include a periodic review of the suitability of the initial regional boundaries. The initial boundaries may either expand or contract and regions may need to be subdivided or merged as we gain more detailed information about the drivers of natural biological variation in each region.

However, since most regions do not have many reference sites to begin with, these analyses will have to take place iteratively as the program builds up a sufficient number of sites in each region. As initial guide, the panel recommended that the RCMP attempt to distribute sites to represent the following natural gradients:

- Stream size (stream order, discharge volume, etc.)
- Geology (with special attention to gradients in calcareous composition)
- Climate (temperature and precipitation)
- Elevation
- Reach slope (an important driver of stream morphology and substrate composition)
- Conductivity and natural nutrient gradients (associated with alkalinity)

The second component to site management is periodic review of sites in the reference pools to assess their continued suitability as reference sites. Conditions within stream reaches and in their upstream drainages can change over time (e.g., timber harvest, conversion of natural landscapes to agricultural or urban/suburban/exurban uses). Furthermore, we may discover sources of stress that were unknown when sites were initially added to the reference pools (e.g., discovery of point source discharges, mines, flow withdrawals/diversions, small-scale placer mining, etc.). Sites that fall into this category may be monitored to measure the impacts of these stressors, but they should be removed from the reference pools.

Dealing with natural disturbance

Natural disturbances such as forest fires, catastrophic floods and landslides can have a significant impact on biological assemblages and physical habitat conditions. As such, they can contribute considerable noise to reference distributions, thereby reducing the precision of biological assessment tools based on these distributions.

There are several competing philosophies for how to handle sites with recent natural disturbances. For example, Idaho's program flagged sites affected by natural disturbance to assess in parallel with other reference sites (Grafe 2004). In contrast, Oregon explicitly included these sites with other reference sites, as a means of incorporating natural disturbance as a component of natural variability (Drake 2003). The RCMP will keep these sites in the reference pools, but will not sample them after the disturbance. The appropriate time to avoid sampling disturbed reference sites is not currently known and should be the subject of targeted research or special study.[14]

[14] The San Diego Regional Water Quality Control Board has funded a multi-year project with the ABL to track biological assemblage recovery in reference and test sites following two large scale forest fires events in 2003 and 2007.

MONITORING STRATEGY

Monitoring Design

The primary question to be answered from the monitoring of the RCMP is "what is the expected natural composition of lotic freshwater organisms in each of the major biogeographical regions of California"? In order to answer this question, the panel agreed it is most important to gather information from a large number of sites in order to capture the full range of natural variability within a region. To collect this information in a spatially balanced and unbiased fashion, the panel advocated a probabilistic sampling design. Probabilistic designs were used in the REMAP, WEMAP, CMAP and PSA surveys in order to get unbiased estimates of stream condition and the approach for this design would be similar. In this case, the regional reference pool would represent the sample frame where sites would be selected at random for sampling. As in the PSA, these randomly drawn sites could be stratified to ensure the spatial distribution across natural gradients such as stream order, elevation, slope, geology, precipitation, or other factors.

An important secondary component to answering the monitoring question is to assess how the range of natural conditions changes over time. Certainly year-to-year variability can alter the distribution and abundance of organisms based on climatic conditions (i.e., wet vs. dry year, warm vs. cold year, etc.). Revisiting sites is the most powerful way to gather this type of temporal information. Two designs lend themselves to answering this question. The first would be to revisit a subset of the probabilistic sites. The panel favored this type of design, termed "rotating panel", because it provides both temporal and spatial variance terms. Urquhart and Kincaid (1999) and Larsen *et al*. (2004) describe the rotating panel strategy in more detail. However, a large number of potential reference sites are already being monitored on a regular basis. Provided these sites can pass the large- and local-scale screening criteria, the panel recommended sampling these sites as a cost-effective method to gain trends information at specific locations of interest. The main drawback to the targeted design, however, is the lack of ability to extrapolate to other reference locations.

Indicators and methods

Once the reference site pools are established, they can be sampled to meet the needs of a variety of programs. However, the panel agreed that a base program should monitor those indicators that are currently being used for SWAMP's statewide assessments (see PSA text box). These indicators include BMIs, physical habitat quality and basic water quality measurements. In some instances, enhancement of the indicators in certain regions or at certain sites may be needed to

Indicators sampled for the SWAMP Perennial Stream Assessment (PSA)

Biological
- BMIs
- Algae (diatoms, soft algae)
- CRAM riverine wetland methods

Physical Habitat
- SWAMP instream and riparian condition (derived from EMAP field protocols)

Chemical
- Nutrients (SRP, NO_2, NO_3,TP, TN, Si)
- Major ions (Cl^-, $S0_4$)
- SSC, turbidity
- pH
- Hardness, alkalinity, conductance

address local concerns. Region-specific enhancements were deemed acceptable as long as the base program is not handicapped to implement the enhancements. For example, additional biological indicators such as fish have been used by others (Hughes *et al*. 2005; Brown and Moyle 2005). Field and laboratory methods and quality assurance measures should also be consistent with SWAMP.

Number of reference sites

The appropriate number of sites to sample in each region depends on the extent of variation related to natural gradients, which is currently unknown for most regions. The panel therefore could not provide specific guidance on sample size. Instead the panel made two recommendations:

1. The RCMP should sample approximately 50 sites in each region to support assessments of natural variability. Intensification of sampling in initial years was recommended to establish the reference baseline, with potentially reduced intensity in later years.

2. The RCMP should conduct power analysis to determine the optimal sample size for assessing confidence in the statistical parameters of the distribution of biological metrics (Figure 9). For example, an assessment of variance at reference sites within a region can be calculated based on existing data (although not all regions have enough sites to support this at present). The inflection point of this power curve represents an efficient sample size where additional sites provide little improvement in confidence, yet fewer sites might dramatically broaden the confidence limits.

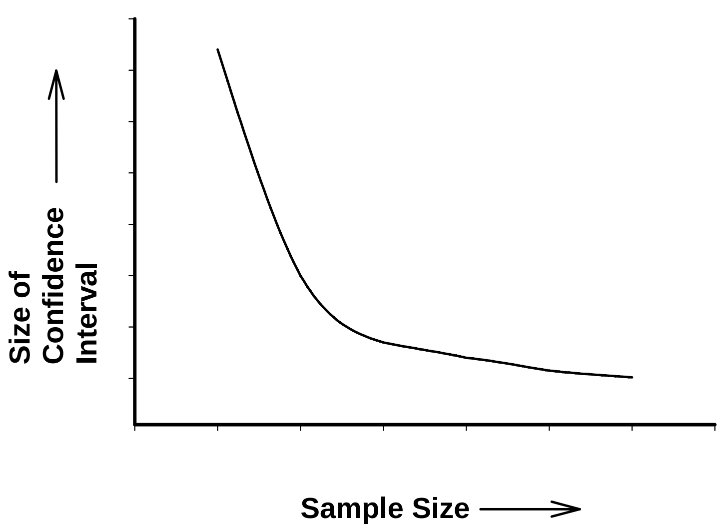


Figure 9. Example power curve defining sample sizes relative to site variability.

Sampling frequency
Sampling frequency affects trend detection. The optimal sampling frequency for trend detection is a function of sampling design. Trend detection as part of the probabilistic design is a function of number of sites (spatial variability), sampling frequency (temporal variability), amount of change to be detected, and other factors. The panel recommended a subset of probabilistic sites be sampled once within the appropriate index period for the region (should be consistent with the index period used for the SWAMP PSA). The recommended index period should capture a time frame where benthic macroinvertebrate communities are sufficiently stable to produce repeatable results, but prior to stress from late season flow reductions. Revisiting a subset of probabilistic sites each year will provide an estimate of interannual variability, thus improving large-scale trend detection. The proportion of revisited sites was not addressed specifically by the panel, but could be optimized using power analysis.

The panel agreed that targeted sites were an efficient way to assess long-term trend detection. Sampling frequency at targeted sites is a function of variability in the biological metrics, the amount of time required to detect a trend, and the amount of detectable change. The panel recommended that the RCMP use power analysis to establish the optimal sampling frequency (Figure 10). Once again, this could possibly be accomplished using data from existing sites that have been sampled for a number of years.

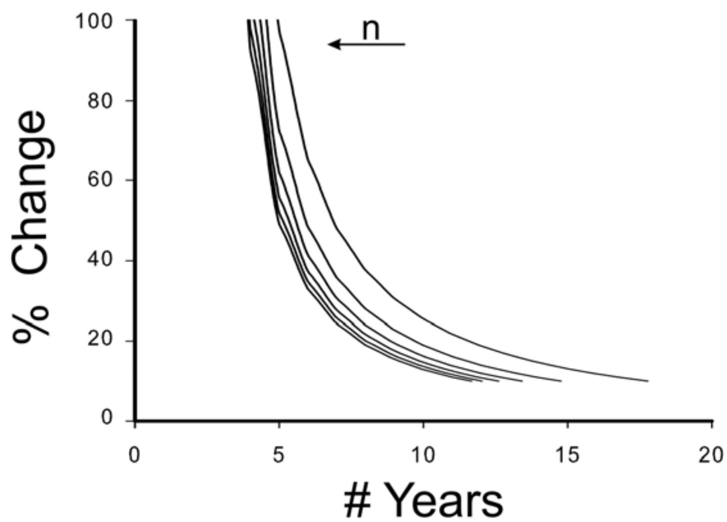


Figure 10. Theoretical power curves describing the relationship between the number of samples collected and the magnitude of detectable change at fixed sites. Individual curves represent different numbers of samples per year, with higher numbers toward the left of the figure.

ADDITIONAL RECOMMENDATIONS

Funding

Defensible bioassessment techniques and biocriteria require a reference condition program that can document both spatial and temporal variation. While the panel did not recommend a minimum level of funding, they advised that funding will need to be long term and stable. Several cost-effective strategies are available, but options discussed included trade-offs between probabilistic and targeted sites, optimizing sample size using power analysis (see previous section on sample size and frequency), and finding additional partners to help support the RCMP (see section below on collaboration). Regardless, SWAMP should prioritize some sampling effort every year to document annual variation in reference condition.

Inter-regional consistency

The RCMP should continue to focus on the issue of fostering consistency among the various regions of the state. Statewide assessments and comparisons among regions require a common currency for interpreting statewide assessments, and for inter-regional comparisons. However, this goal is complicated by the need for regional specific reference selection criteria. While the panel did not deliberate extensively on this topic, it recognized the importance of the issue and provided some initial guidance to help focus the thinking of the program. The main advice from the panel was that the objective of inter-regional consistency can probably not be resolved by the reference site selection process itself, but rather must be dealt with through data analysis and interpretation.

Development and application of assessment tools can be based on either regional reference pools or combined statewide reference pools. Regionalized assessment tools provide sensitivity to local environmental gradients and are more likely to pick up sites that deviate from the regional expectation. In contrast, statewide assessment tools would judge all of the state's sites on the same basis, but may reduce responsiveness to locally important gradients. Furthermore, we may have to accept that the performance of statewide analytical tools may vary regionally depending on the quality of the respective regional reference pools.

An example of an analytical solution is a hybrid approach in which both the regional and statewide indices are built and both tools are used to score test sites. Where both tools agree, there is relative certainty in the assessment of that site (i.e., both tools indicate reference-like or both indicate impacted). Where the tools disagree, a greater degree of relative uncertainty exists and additional information may be required to help interpret the status of that site (i.e., other indicators, additional sampling).

Collaborations/Coordination

Consistent with its policy to coordinate with other state and federal water quality monitoring efforts, SWAMP should seek opportunities to build partnerships with other state and federal agencies. Many of these entities have current reference programs (e.g., USFS, EPA, USGS), while others would benefit from joining an established reference program (e.g., Non-point Source Monitoring, State Parks, Irrigated Lands Program, Agricultural Coalitions, National Park Service, etc.). In addition SWAMP should explore ways to combine its bioassessment RCMP with other program components that would benefit from reference condition (e.g., CRAM, wetland monitoring, nutrient and sediment criteria monitoring).

The panel recommended exploration of ways to improve the types and quality of data used in GIS analyses. For example, the program could seek opportunities to coordinate with other state/federal/university efforts to enhance base layers like the NHD+ and stressor layers for quantification of grazing, timber harvest, pesticide application, etc. Further, the RCMP would should explore research efforts designed to improve prediction of specific stressor impacts and efforts to develop models that can be used to assess impact components that are not easily summarized by the ATtILA model. For example a model predicting sediment load (AnnAGNPS sedimentation model, USDA 2000) was applied by the University of Nevada, Reno. Other needs include estimating mining impacts, pesticide impacts and a means for summarizing the intensity of water manipulation within candidate areas.

Involving stakeholders in the process

It is often desirable to select sampling locations that occur on publicly owned land or land with easy access. Since it is important to sample streams from a truly representative set of sites within an area, it is often necessary to sample from reaches running through privately owned land. Reasonable efforts should be taken to obtain permission from landowners before rejecting candidate sites. This stage is very important and the quality of the final data set (and the ability to make inferences about reference conditions in the region of interest) will depend on the ability to obtain a representative set. The degree to which this stage is important varies regionally since some areas have more private ownership than others (e.g., western Sierra Nevada has many more publicly-owned lands than the interior chaparral).

Building effective relationships with local stakeholders (regional boards, watershed groups, landowner group, tribal groups, etc.) is clearly a critical part of making this reference site selection methodology work, especially in regions with a large degree of private ownership. To this end, implementation of this RCMP should include efforts to promote transparency in methods, encourage feedback and participation and explore opportunities to improve access to important privately held reference sites.

CONSIDERATIONS FOR OTHER FLOWING WATERS

The following section is not intended to be an exhaustive review of issues for defining reference conditions for these waterbodies, but a summary of the panel's preliminary guidance regarding issues that are likely to be important in these systems.

Large Rivers/ Non-wadeable streams

Large rivers are likely to require non-standard approaches to defining biological expectations because there are relatively few non-wadeable streams/rivers in the state and most receive the cumulative impacts of all human activities in their watersheds. Furthermore, several panelists suggested that standard chemical and physical habitat screening was unlikely to work in these systems. Screening criteria should include quantification of hydromodification, distance downstream from dams or other stressors.

Several of the alternative strategies could apply to these systems. Another alternative would be to target sampling at points along river just before they experience significant increases in sources of anthropogenic stress (e.g., where rivers in the western Sierra Nevada descend into the Central Valley).

Non-perennial streams

Non-perennial streams tend to have more variable biological assemblages than perennial streams. The standard approach should work for most of these systems statewide, but special attention should be given to classification of non-perennial streams by their degree of "intermittent-ness" in both space and time. The panel suggested that the RCMP should take advantage of current statewide vegetative mapping efforts to explore the potential for classifying non-perennial streams.

LITERATURE CITED

ANZECC and ARMCANZ. 2000. Australian and New Zealand guidelines for fresh and marine water quality. Volume 2. Aquatic ecosystems- rationale and background information. Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Australian Government, Department of the Environment and Heritage, Canberra, Australia.

Arnold, C.L.J. and C.J. Gibbons. 1996. Impervious surface coverage: The emergence of a key environmental indicator. *Journal of the American Planning Association* 62:243–258.

Bailey, R.G., P.E. Avers, T. King, and W.H. McNab, (eds.). 1994. Ecoregions and subregions of the United States (1:7,500,000 scale map, with supplementary table of map unit descriptions, compiled and edited by W.H. McNab and R.G. Bailey). USDA Forest Service. Washington, DC.

Bailey, R.C., M.G. Kennedy, M.Z. Dervish, and R.M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. *Freshwater Biology* 39:765-774.

Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. Bioassessment of Freshwater Ecosystems: using the Reference Condition Approach. Kluwer Academic Publishers. Norwell, MA.

Barbour, M.T., J.B. Stribling, J. Gerritsen, and J.R. Karr. 1996a. Biological Criteria: Technical guidance for streams and small rivers. Revised Edition. EPA 822-B-96-001. USEPA Office of Water. Washington, DC.

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastian. 1996b. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15:185-211.

Barbour, M.T. and C. Hill. 2003. The status and future of biological assessment for California streams. Report to the California State Water Resources Control Board. Tetra Tech, Inc. http://www.swrcb.ca.gov/swamp/reports.html

Brown, L.R. and P.B. Moyle. 2005. Native fishes of the Sacramento-San Joaquin Drainage, California: a history of decline. Pages 75-98 in: J.N. Rinne, R.M. Hughes, and B. Calamusso (eds.) Historical Changes in Large River Fish

Assemblages of the Americas. American Fisheries Society, Symposium 45. Bethesda, MD.

Carter, J.L and S.V. Fend. 2005. Setting limits: the development and use of factor-ceiling distributions for an urban assessment using benthic macroinvertebrates. *American Fisheries Society Symposium* 47:179-191.

Chessman, B.C. 1995. Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family-level identification, and a biotic index. *Australian Journal of Ecology* 20:122-129.

Chessman, B.C. 2006. Prediction of riverine fish assemblages through the concept of environmental filters. *Marine and Freshwater Research* 57:601-609.

Chessman, B.C. and M.J. Royal. 2004. Bioassessment without reference sites: use of environmental filters to predict natural assemblages of river macroinvertebrates. *Journal of the North American Benthological Society* 23:599-615.

Chessman, B.C., M. Muschal, and M.J. Royal. 2008. Comparing apples with apples: use of limiting environmental differences to match reference and stressor-exposure sites for bioassessment of streams. *River Research and Applications* 24:103-117.

deVlaming, V., D. Markiewicz, K.L. Goding, T. Kimball, and R. Holmes. 2004. Macroinvertebrate assemblages in agriculture- and effluent-dominated waterways of the lower Sacramento River watershed. Report to the Central Valley Water Quality Control Board. UC Davis Aquatic Toxicology Laboratory. Davis, CA.

deVlaming, V., D. Markiewicz, K.L. Goding, A. Morrill, and J. Rowan. 2005. Macroinvertebrate assemblages of the San Joaquin River watershed. Report to the Central Valley Water Quality Control Board. UC Davis Aquatic Toxicology Laboratory. Davis, CA.

Drake, D. 2003. Selecting Reference Condition Sites: An Approach for Biological Criteria and Watershed Management. Oregon Department of Environmental Quality, Watershed Assessment Division. Portland, OR.

Grafe, C.S. 2004. Selection of Reference Condition for Small Streams in Idaho: A Systematic Approach. Department of Environmental Quality, Surface Water Quality Program. Boise, ID.

Griffith, M.B., P. Husby, R.K. Hall, P.R. Kaufmann, and H.H. Brian. 2003. Analysis of macroinvertebrate assemblages in relation to environmental gradients among lotic habitats of California's Central Valley. *Envirnomental Monitoring and Assessment* 28:281-309.

Hall, L.W. and W.D.Killen. 2001. Characterization of benthic communities and physical habitat in an agricultural and urban stream in California's Central Valley. University of Maryland-Agricultural Experiment Station. Queenstown, MD.

Hawkins, C.P. and R.H. Norris (eds.). 2000. Landscape classifications: aquatic biota and bioassessments. Journal of the North American Benthological Society 19.

Herbst, D.B. and E.L. Silldorff. 2006. Development of an Index of Biological Integrity (IBI) for Stream Assessments in the Lahontan Region of California. DRAFT Report to the Lahontan Regional Water Quality Control Board.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-47 in: W.S. Davies and T.P. Simon (eds.), Biological assessment and criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers. Ann Arbor, MI.

Hughes, R.M. and D.P. Larsen. 1988. Ecoregions: an approach to surface water protection. *Journal of the Water Pollution Control Federation* 60:486-493.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference sites: A method for assessing stream potentials. *Environmental Management* 10:629-635.

Hughes R.M., J.N. Rinne, and B.Calamusso. 2005. Historical changes in large river fish assemblages of the Americas: A synthesis. Pages 603-612 in: R.M. Hughes, J.N. Rinne, and B.Calamusso(eds.), Historical Changes in Large River Fish Assemblages of the Americas. American Fisheries Society, Symposium 45. Bethesda, MD.

Karr, JR. 1995. Protecting aquatic ecosystems: clean water is not enough. Chap. 2. Pages 7-13 in: W.S. Davis and T.P. Thomas (eds.), Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. CRC Press, Inc., Boca Raton.

Karr, JR. 1997. Measuring biological integrity. Essay 14A. Pages 483-485 in: G.K. Meffe and G.R. Carroll (eds.), Principles of Conservation Biology, 2nd edition. Sinauer, Sunderland, Massachusetts.

Karr, J.R. and E.W. Chu. 1999. Restoring Life in Running Waters - Better Biological Monitoring. Island Press. Washington, DC.

Larsen, D.P., P.R. Kaufmann, T.M. Kincaid, and N.S. Urquhart. 2004. Detecting persistent change in the habitat of salmon-bearing streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences 61:283-291.

Meyer, J.L. 1997. Stream health: incorporating the human dimension to advance stream ecology. Journal of the North American Benthological Society 16:439-447.

Ode, P.R. 2002. A quantitative framework for reference site selection: case study from the Sierra Nevada Foothills ecoregion. Sacramento River Watershed Program. California Department of Fish and Game, Water Pollution Control Laboratory. Rancho Cordova, CA.

Ode, P.R. 2007. Standard operating procedures for collecting macroinvertebrate samples and associated physical and chemical data for ambient bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 001. Sacramento, CA.

Ode, P.R., D. P. Pickard, J. P. Slusark, and A.C. Rehn. 2005. Adaptation of a bioassessment reference site selection methodology to creeks and sloughs of California's Sacramento Valley and alternative strategies for applying bioassessment in the valley. Report to the Central Valley Regional Water Quality Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory, Rancho Cordova, CA.

Omernik, J.M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49-62 in: W.S. Davies and T.P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making. Lewis Publishers, Ann Arbor, Michigan.

Paulsen, S.G., J.L. Stoddard, S. Holdsworth, A. Mayio, and E. Tarquino. 2006. Wadeable streams assessment: A collaborative survey of the nation's streams. EPA 841-B-06-002. USEPA Office of Research and Development/Office of Water. Washington, DC.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16: 391-409.

Poff, N.L. and J.V. Ward. 1990. Physical habitat template of lotic systems: recovery in the context of historical pattern of spatiotemporal heterogeneity. Environmental Management 14: 629-645.

Reference Condition Working Group 2.3 (REFCOND). 2003. Rivers and Lakes – Typology, Reference Conditions and Classification Systems. Guidance Document No. 10. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Luxembourg.

Rehn, A.C., P.R. Ode, and J.T. May. 2005. Development of a benthic index of biotic integrity (B-IBI) for wadeable streams in northern coastal California and its application to regional 305(b) assessment. Report to the State Water Resources

Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory. Rancho Cordova, CA.

Rosenberg D.M., T.B. Reynoldson, and V.H. Resh. 1999. Fraser River Action Plan: Establishing Reference Conditions for Benthic Invertebrate Monitoring in the Fraser River Catchment. Environment Canada No. DOE-FRAP 1998-32. British Columbia, Canada.

Statzner, B., A.G. Hildrew, and V.H. Resh. 2001. Species traits and environmental constraints: entomological research and the history of ecological theory. *Annual Review of Entomology* 46: 291-316.

Stoddard, J.L., D.V. Peck, S.G. Paulsen, J. Van Sickle, C.P. Hawkins, A.T. Herlihy, R.M. Hughes, P.R. Kaufmann, D.P. Larsen, G. Lomniki, A.R. Olsen, S.A. Peterson, P.L. Ringold, and T.R. Whittier. 2005. An Ecological Assessment of Western Streams and Rivers. EPA 620/R-05/005. US Environmental Protection Agency, Washington, D.C.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R. Johnson, and R. Norris. 2006. Setting expectations for ecological condition of streams: the concept of reference conditions. *Ecological Applications* 16:1267-1276.

Urquhart, N.S. and T.M. Kincaid. 1999. Designs for detecting trend from repeated surveys of ecological resources. *Journal of Agricultural, Biological, and Environmental Statistics* 4:404-414.

US Environmental Protection Agency (USEPA). 2005. Use of Biological Information to Better Define Designated Aquatic Life Uses in State and Tribal Water Quality Standards: Tiered Aquatic Life Uses. EPA-822-R-05-001. USEPA Office of Water, Washington, DC.

USEPA. 2009. Evaluation of the California State Water Resource Control Board's Bioassessment Program. Prepared by C. Yoder and R. Plotnikoff, Center for Applied Bioassessment and Biocriteria, Midwest Biodiversity Institute, Columbus, OH, for USEPA Office of Science and Technology. Washington, DC.

Wright, K.K. and J.L. Li. 2002. From continua to patches: examining stream community structure over large environmental gradients. Canadian Journal of Fisheries and Aquatic Sciences 59: 1404-1417.

Waite, I.R., A.T. Herlihy, D.P. Larsen, N.S. Urquhart, and D.J. Klemm. 2004. The effect of macroinvertebrate taxonomic resolution in large landscape bioassessments: an example from the Mid-Atlantic Highlands, U.S.A. Freshwater Biology 49: 474-489.

Whittier, T.R., D.P. Larsen, R.M. Hughes, C.M. Rohm, A.L. Gallant, and J.M. Omernik. 1987. The Ohio Stream Regionalization Project: A compendium of results. EPA/600/3-87-025. Corvallis, OR.

Yoder, C.O. and E.T. Rankin. 1995. Biological criteria program and implementation in Ohio. Pages 109-144 in: Davis W.S. and T.P. Simon (eds), Biological Assessment and Criteria: Tools for Water Resource Planning. CRC Press. Boca Raton, FL.

Appendix 6

MEMORANDUM

Tetra Tech, Inc.
400 Red Brook Blvd., Suite 200
Owings Mills, MD 21117-6102
phone 410-356-8993
fax 410-356-9005

DATE: 31 July 2009
TO: Phil Markle
FROM: Jerry Diamond, Ph.D.

SUBJECT: Reference conditions and bioassessments in southern California streams

All bioassessment methods depend on having appropriate reference conditions with which to base an assessment; i.e., bioassessment data for a given site cannot be accurately interpreted by themselves—interpretation or assessment of the site data is done within the context of the biology that can be expected to occur naturally, given the type of habitat present, the type of aquatic system, and the physiographic region (i.e., ecoregion) of the country (Stoddard et al., 2006). Identifying appropriate reference conditions for certain types of aquatic systems, habitats, and ecoregions can be problematic because of wide-scale human land use changes such as hydrological modification (e.g., dams, levees, concrete channelization), urbanization (e.g., increased runoff, removal of riparian vegetation, bank protection structures), and agricultural/ livestock effects (e.g., water removal for irrigation, removal of riparian vegetation).

Southern California (Los Angeles, San Diego and surrounding counties) is an area that has experienced intense land use changes over the past 50 years, particularly in terms of urbanization and its many environmental consequences (e.g., changes in the natural hydrology, changes in stream geomorphology, etc.). In particular, low gradient as well as low elevation streams in this region have been especially prone to land use effects. This situation has resulted in high uncertainty regarding appropriate reference conditions for low gradient and low elevation streams in this region.

This observation was identified in a Technical Report I and others at Tetra Tech prepared for the Los Angeles Regional Water Quality Control Board (Tetra Tech, 2005; 2006). In that report we evaluated stream biological condition with respect to a generalized human disturbance gradient in the region, as part of an EPA-funded project to evaluate the possibility of developing tiered aquatic life uses (TALU) for southern California coastal streams. Relying on SWAMP and other data for the region, we attempted to use the recently developed southern California IBI (SoCal IBI, Ode et al., 2005) to define certain attributes of the Biological Condition Gradient for the region, which could then be used to develop TALU (Davies and Jackson, 2006). We observed that the BCG should be different (i.e., expectations lower) for low versus high elevation streams in that project and that low elevation streams lacked a clear reference condition in this region.

Working with a Technical Advisory Committee (TAC) on this project (consisting of regional experts from California Fish & Game, State Water Resources Control Board, other Regional Boards, EPA Region 9, and universities), we identified a lack of appropriate reference sites for low elevation/low gradient streams as a critical data gap in moving forward with TALU. A fairly extensive search of existing biological data in the region by Tetra Tech and the TAC indicated that suitable reference sites at lower elevations and/or for lower stream gradients were not available with which to benchmark a biological condition gradient.

Subsequent to the above project, I have been working with the Southern California Coastal Water Research Project (SCCWRP) and the LA Regional Board in facilitating two workshops on TALU for the region. In the most recent stakeholder workshop (held June 2008), there was focused discussion on the issue of appropriate reference conditions, in which there was agreement that low gradient (rather than low elevation) was perhaps the most critical factor distinguishing stream biology in the region and that reference condition for low gradient streams (many but not all of which occur at low elevation) is a critical data gap (Schiff and Diamond, 2009). In fact, in the "road map" of projects developed from this workshop, defining reference condition for streams in this region was identified as one of the top priority needs.

Given the difficulty in identifying appropriate reference conditions for low gradient coastal streams in southern California, it is perhaps premature to set regulatory requirements based on biology observed at these types of sites. The TALU framework, as well as the regional stakeholder workshops (e.g., Schiff and Diamond, 2009) recognize that different hydrologic, geomorphic, and other habitat-related factors will dictate the biological characteristics that can be expected in a given stream. The type of aquatic life uses one can reasonably expect from a low gradient or modified stream in southern California, for example, are not the same as from a high gradient or natural stream, as our previous work has demonstrated. What is the expected biological condition for low gradient or modified streams in southern California is a question that needs more attention and, as noted by all stakeholders at the June 2008 workshop, incorporation of information using other assemblages (e.g., algae) in addition to macroinvertebrates.

Literature Cited

Davies, S. and S. Jackson. 2006. The Biological Condition Gradient: A descriptive model for interpreting change in aquatic ecosystems. Ecological Applications: 16: 1251–1266.

Ode, P.R., A.C. Rehn and J.T. May. 2005. A quantitative tool for assessing the integrity of southern California coastal streams. *Environmental Management* 35:493-504.

Stoddard, J., D. Larsen, C. Hawkins, R. Johnson, and R. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications, 16: 1267–1276.

Tetra Tech. 2005. Evaluation of Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams. Draft Summary Report. Prepared for: EPA Region 9 and California Los Angeles Regional Water Quality Board. Tetra Tech, Inc. Owings Mills, MD.

Tetra Tech. 2006. Revised Analyses of Biological Data to Evaluate Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams. Prepared for: EPA Region 9 and California Los Angeles Regional Water Quality Board. Tetra Tech, Inc. Owings Mills, MD.

Appendix 7

B233

IDENTIFYING BARRIERS TO TIERED AQUATIC LIFE USES (TALU) IN SOUTHERN CALIFORNIA

Ken Schiff
and
Jerry Diamond

Southern California Coastal Water Research Project

Technical Report 590 - June 2009

B234

Identifying Barriers to Tiered Aquatic Life Uses (TALU) in Southern California

Ken Schiff[1] and Jerry Diamond[2]

[1]Southern California Coastal Water Research Project
3535 Harbor Blvd., Suite 110, Costa Mesa, CA 92626

[2]Tetra Tech, Inc.
400 Red Brook Blvd., Suite 200, Owings Mills, MD 21117

June 2009

Technical Report 590

PREFACE

The goal of this document is to explore the use of a new environmental management tool in southern California known as Tiered Aquatic Life Use or TALU. TALU focuses on the traditionally difficult regulatory problem of maintaining balanced biological communities. The existing California State regulatory framework only lists broad, categorical biological expectations such as warmwater (WARM) or coldwater (COLD) habitat. TALU has the potential to refine the biological expectations within each of these broad categories based on a variety of factors including physical habitat, hydrology, or level of habitat alteration. More detailed expectations tailored to the specific habitat could dramatically improve environmental managers' ability to assess biological impairment and set appropriate benchmarks for improvement.

The goal of this document was to create a workplan for implementing TALU in southern California. We compiled existing information about TALU and, by working with local stakeholders, identified some of the largest technical and potential policy barriers for implementation. This was not an easy task since southern California stakeholder opinions, sensitivities, and personal agendas can dramatically differ. TALU is a powerful tool that can be utilized as a positive step towards conservation and restoration or, alternatively, abused as a means of limiting regulatory oversight. Ultimately, this report lists 13 projects that should be undertaken to help resolve these barriers and develop scientifically defensible tiered aquatic life uses, and integrate these tiered uses into the existing water quality standards program to the betterment of the environment.

This document does not focus on the many non-technical factors that will be fundamental for TALU to be a successful management tool. These factors, which can be political and procedural, are built into the State and Federal regulatory policy development process. Many times, divisive policy issues are a function of perception rather than fact. It is the aim of this document to ensure that the all of the facts are available to evaluate the viability of TALU as a meaningful regulatory tool.

ACKNOWLEDGEMENTS

Attendees at the TALU Workshops held November 27, 2007 (1) and June 19, 2008 (2) at the Southern California Coastal Water Research Project.

Jeff Armstrong (Orange County Sanitation District) [1]
Zora Baharians (City of Los Angeles) [1]
Polly Barrowman (Heal the Bay) [1]
Shirley Birosik (Los Angeles Regional Water Quality Control Board) [1]
Angela Bonfiglio (Ventura County Watershed Protection District) [1]
Richard Boone (Orange County Resources and Development Management Division) [1]
Lilian Busse (San Diego Regional Water Quality Control Board) [1, 2]
Seth Carr (City of Los Angeles) [1]
Arlene Chun (Riverside County Flood Control District) [1]
Chris Crompton (Orange County Resources and Development Management Division) [1]
Erika DeHollan (Los Angeles County Sanitation District) [1]
Jerry Diamond (TetraTech, Inc.) [1, 2]
Sabrina Drill (University of California Extension) [1, 2]
Jennifer Dulay (California Conservation Corps) [1]
Melenee Emanuel (State Water Resources Control Board) [1]
Terry Fleming (US EPA Region IX) [1]
Germenew Amenu (Los Angeles County Department of Public Works) [1]
Stuart Goong (Orange County Resources and Development Management Division) [1]
Gerald Greene (City of Downey) [1]
Lenwood Hall (University of Maryland) [2]
Jim Harrington (California Department of Fish and Game) [1]
Ann Heil (Los Angeles County Sanitation District) [1, 2]
Evan Hornig (US EPA Office of Water) [1]
Scott Johnson (ABC Laboratories, Inc.) [1, 2]
Michael Lyons (Los Angeles Regional Water Quality Control Board) [1, 2]
Phil Markle (Los Angeles County Sanitation District) [1]
Rafael Mazor (Southern California Coastal Water Research Project) [1, 2]
Kathleen McGowan (Geosyntec Consultants) [1]
Gerry McGowen (City of Los Angeles) [1]
Sofia Mohaghegh (City of Los Angeles) [1]
Jenny Newman (Los Angeles Regional Water Quality Control Board) [1]
Peter Ode (California Department of Fish and Game) [1, 2]
Craig Pernot (ABC Laboratories, Inc.) [1]
Renee Purdy (Los Angeles Regional Water Quality Control Board) [1, 2]
Rik Rasmussen (State Water Resources Control Board) [1, 2]
Emily Reyes (State Water Resources Control Board) [1]
Deborah Smith (Los Angeles Regional Water Quality Control Board) [1]
Ken Schiff (Southern California Coastal Water Research Project) [1, 2]
Naeem Siddiqui (California Department of Fish and Game) [1]
Joyce Sisson (Heal the Bay) [1]
Eric Stein (Southern California Coastal Water Research Project) [1]
Julie Stephenson (Geosyntec Consultants) [1]
Thomas Suk (Lohantan Regional Water Quality Control Board) [1]
Jack Topel (Santa Monica Bay Restoration Commission) [1]
Rebecca Vega-Nascimiento (Los Angeles Regional Water Quality Control Board) [1]
Penny Weiand (City of Los Angeles) [1]
Vera Williams (State Water Resources Control Board) [1]
Joanna Wisniewska (County of San Diego Department of Environmental Health) [1]
Clayton Yoshida (Los Angeles Department of Water and Power) [1]
Vada Yoon (Flow Science Inc.) [1]

Participants TALU workshop June 19, 2008, Southern California Coastal Water Research Project, Costa Mesa, CA.

LIST OF ACRONYMS

ACOE	Army Corps of Engineers
BCG	Biological condition gradient
COLD	coldwater habitat
CSUSM	California State University San Marcos
DWR	Department of Water Resources
EMAP	Environmental monitoring and assessment program
EPA	Environmental Protection Agency
EWH	exceptional warmwater habitat
GSG	Generalized stressor gradient
IBI	Index of biological integrity
MWH	modified warmwater habitat
NOAA	National Oceanic and Atmospheric Administration
NPDES	National Pollutant Discharge Elimination System
PSA	Perennial Stream Assessment
SANDAG	San Diego Association of Governments
SCAG	Southern California Association of Governments
SCCWRP	Southern California Coastal Water Research Project
SFEI	San Francisco Estuary Institute
SMC	Stormwater Monitoring Coalition
SNARL	Sierra Nevada Research Laboratory
SWAMP	Surface water ambient monitoring program
TALU	Tiered aquatic life use
USFS	United Sates Forest Service
WARM	warmwater habitat (California)
WWH	warmwater habitat (Ohio)

TABLE OF CONTENTS

BACKGROUND

What are Tiered Aquatic Life Uses (TALU)?

All states, including California, have designated uses (known as beneficial uses in state terminology) that protect aquatic life. California has several different beneficial uses relevant to protecting aquatic life including warmwater and coldwater habitat, and protection of different life stages such as fish migration and spawning.[1] Most ecosystem managers recognize that the more specific the designated use definition, the clearer it is to describe attainment goals and ensure maintenance and protection of the designated use. EPA also acknowledged this fact and, in response, developed a framework for states to develop Tiered Aquatic Life Uses (TALU).

TALU recognizes different management goals for waterbodies within a given waterbody class and these goals are defined based on detailed information on biological condition and stressor intensity. An example of TALU would be the three tiers of warmwater use defined by the Ohio EPA (OEPA, 2008): exceptional warmwater habitat (EWH), warmwater habitat (WWH), and modified warmwater habitat (MWH). All of these tiers are part of a designated use for warmwater habitat, but each of these tiers is associated with different biological expectations based on detailed knowledge of these systems. EWH has a higher expectation of biological condition (i.e., the types of flora and fauna that should be present represent higher water quality and higher habitat quality) than WWH, which in turn, has a higher biological expectation than MWH.

It is important to recognize that tiered uses are defined based on fundamental differences in structural or hydrological condition, not the current biological or water quality condition. Instead, biological expectations for each tiered use are based on knowledge of what biota is capable of occurring in a waterbody given the fundamental structural or hydrological template that exists. In this way, environmental managers utilize TALU to achieve effective stewardship of beneficial uses by: (1) identifying high quality waterbodies and preventing the gradual degradation of these waterbodies; and (2) identifying restoration benchmarks for degraded biological condition in waterbodies given their structural and hydrologic condition.

Southern California is a tremendously valuable location for examining the application of TALU because of its wide array of biological habitats, extensive structural and hydrologic modification, and regulatory agencies' desire to regulate on biological as well as chemical condition. Streams, coastal lagoons, and bays support sensitive aquatic species, diverse wildlife, and unique habitats. As a result, southern California needs a more refined way of defining Aquatic Life Uses. For example, coastal perennial streams in southern California can range widely in terms of the degree of urbanization, hydrologic regime, and habitat alteration. The TALU framework could be a powerful tool to refine the WARM designated beneficial use and to better reflect attainable aquatic life goals for different stream conditions.

---

[1] Categorical aquatic life beneficial uses that are designated for waterbodies in California include: Warm Freshwater Habitat; Cold Freshwater Habitat; Inland Saline Water Habitat; Estuarine Habitat; Wetland Habitat; Marine Habitat; Rare, Threatened, or Endangered Species; Migration of Aquatic Organisms; and Spawning, Reproduction, and/or Early Development.

Initial Steps of the TALU Process in Southern California

There has been some exploration of TALU concepts in southern California. These initial steps have included a pilot study (Tetra Tech, 2005; 2006) and a subsequent public workshop. Between 2005 and 2007, the pilot study gathered a group of experts to discuss the technical underpinnings of a TALU framework for southern California coastal streams. No new data were collected as part of this effort, but relevant available biological data were compiled to conceptualize the two primary components of TALU: (1) the biological condition gradient (BCG); and (2) the generalized stressor gradient (GSG).

The BCG describes how ten general ecological attributes of aquatic ecosystems change in response to increasing levels of stressors. These attributes include several common aspects of community structure (e.g., pollution sensitive species, endemic long-lived species) organism condition, ecosystem function, and biological attributes related to stream connectivity and the larger watershed scale. The gradient can be considered analogous to a field-based dose-response curve where dose (x-axis) = increasing levels of stressors and response (y-axis) = biological condition (Figure 1). The BCG is divided into six levels of biological condition along the stressor-response curve, ranging from observable biological conditions found at no or low levels of stressors (Level 1) to those found at high levels of stressors (Level 6).

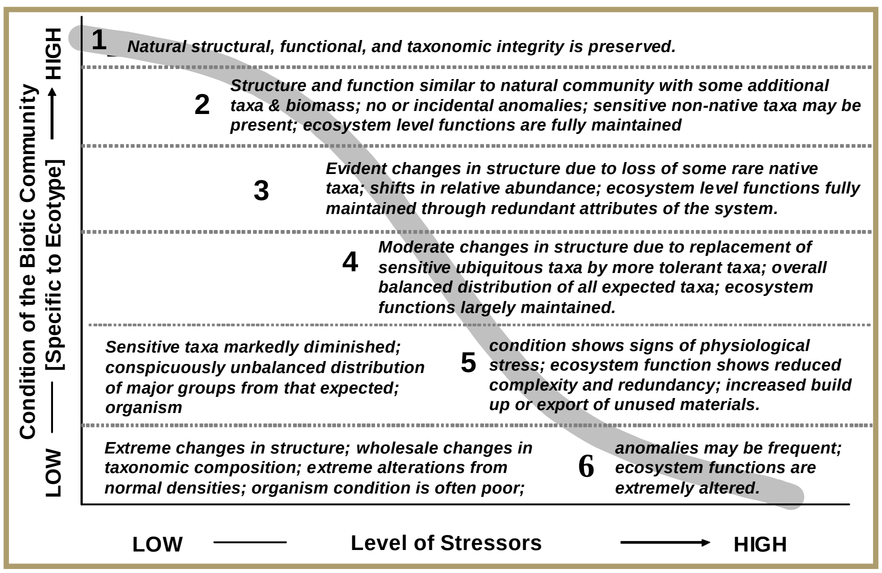


Figure 1. Conceptual model of the Biological Condition Gradient.

The GSG describes the stressor gradient present in the region of interest. Stressors are physical, chemical, or biological factors that adversely affect aquatic biota. Stressors can occur at different scales including instream, within the riparian area and floodplain, or within the watershed. Understanding the linkages between stressors and the response of aquatic biota will help determine existing and potential biological conditions of the aquatic biota. Multiple stressors are usually present and the GSG on the x-axis seeks to represent the cumulative influence of stressors, much as the y-axis generalizes biological condition.

The primary outcome of the pilot study was that TALU could be created in the unique stream environments of southern California. Although much work was left to be accomplished, a BCG and GSG were conceptualized, as well as potential tiered use definitions for perennial streams in the region. The BCG was based largely on the existing Southern California Index of Biological Integrity (IBI; Ode *et al.* 2005) and its associated biological metrics, while the GSG was based primarily on physical habitat measurements and watershed scale disturbance metrics. Relationships were identified between types of coastal perennial streams in southern California, observed aquatic life condition, and preliminary tiered aquatic life uses, along with their corresponding biological expectations.

Several uncertainties were also identified during the pilot study regarding the BCG, GSG, and biological expectations for different tiers. Examples of key uncertainties included defining truly natural conditions in areas where little natural condition remains. Identifying unimpaired sites is vitally important for setting the upper range (i.e. Level 1) of the BCG. Another key uncertainty was the efficacy of additional indicators such as fish or amphibians. One additional uncertainty was optimizing metrics for quantitatively expressing the GSG.

In November, 2007, the Los Angeles Regional Water Quality Control Board sponsored a stakeholder workshop on TALU. The goal of the workshop was two-fold: (1) communicate the findings of the pilot study; and (2) garner input from stakeholders on the viability of TALU as a management tool. Presentations by the US EPA Office of Water and Region IX, the Los Angeles Regional Water Board, and Tetra Tech (US EPA's technical contractor) laid out the rationale, approach, and goals of TALU. The participants were educated about the TALU framework with insight provided by the results of the Southern California pilot study.

The primary outcome of the stakeholder workshop was an earnest interest in TALU. Break-out discussions identified a multitude of issues that were classified into four general areas: (1) determining reference conditions, best attainable conditions, and levels within the BCG; (2) defining stressor gradient metrics; (3) protecting high quality sites and encouraging restoration of degraded sites; and (4) clarifying the regulatory process for developing TALUs.

Identifying Barriers

In June 2008, a second workshop was held to further investigate the specific barriers to implementing TALU in southern California. The workshop was comprised of 12 invited participants representing a cross-section of stakeholders including regulatory, regulated, scientific, and non-governmental sectors (please see Acknowledgements). The group focused on a single goal: design a workplan to overcome the barriers associated with TALU development. Ultimately, the workplan will provide guidance to regulatory and regulated stakeholders that outline the steps necessary to develop TALU in a way that is scientifically defensible and feasible for management. There were three chief considerations asked of participants:

- What are the primary data gaps or information needs?
- How do we combine data gaps into unique project designs?
- What are the factors for prioritizing projects to fill data gaps?

In an effort to constrain the scope of the workplan, the workshop participants immediately decided to limit the scope to perennial wadeable streams in the southern California region.

The workshop ideas and concerns fell into one of three areas including biological, stressor, and implementation related data gaps. The biological-related data gaps included identifying appropriate indicators, adequate representation of reference conditions and range of impact (for defining the BCG scale), capturing natural temporal variation (seasonal/interannual), and specific biological responses to changes in flow (hydromodification).

The stressor related data gaps included improving the understanding and quantification of the human disturbance gradient (to build the GSG), improving the information for quantifying and defining stressor gradients at both the local and watershed scales (e.g., physical habitat and GIS/land use, respectively), and identifying site specific factors that influence stressor impact on aquatic life (e.g., best management practices).

The implementation related issues included identifying appropriate habitat breaks for TALU application, development of appropriate criteria, setting tiers, determining values for nonbiological indicators (i.e. water quality objectives) for the tiers, and integrating TALU with other state or federal regulatory programs.

There were several factors the workshop participants utilized for prioritizing project concepts. These included availability of data/information for compilation as opposed to new data collection, estimated cost, time for completion, and perceived importance in providing defensibility of TALU structure. Ultimately, 14 projects were derived for the workplan based on these criteria.

Table 1. Summary of data gaps or information needs identified at the June 19, 2008 technical meeting regarding the advancement of Tiered Aquatic Life Uses (TALU) in southern California coastal perennial streams and proposed projects that address these gaps.

DATA GAP	PROPOSED PROJECTS
Biology-related	
• The BCG needs to include more than one type of indicator, so that expected responses to human development are accurately evaluated	• Project #1: Develop algal indicators of biological condition for perennial coastal California streams; • Project #2: Develop riparian vegetation and habitat indicators suitable for BCG development
• Natural condition needs to be defined for each stream classifications to determine Level 1 for the BCG	• Project #3: Define minimally impacted (natural) biological condition for coastal perennial streams and determine appropriate stream classification factors
• Temporal variability needs to be captured in the BCG	• Project # 4: Determine seasonal and interannual variability for relevant biological indicators and identify appropriate ranges of indicators for BCG development
• Representation of biological sites needs to be broad and complete enough to ensure accurate BCG development	• Project #5: Characterize range of available biological indicator information and identify gaps in BCG
• Biological expectations for hydrologically modified streams need to be defined	• Project #6: Determine appropriate BCG for different degrees of hydrologic modification
Stressor-related	
• Need to evaluate if recent changes in physical habitat sampling methods provide useful information for quantifying the GSG	• Project #7: Evaluate and develop a refined set of physical habitat measures that help develop the GSG
• Better base maps are needed for quantifying stressor information	• Project #8: Develop refined base maps of stressor information
• Need to better define and integrate landscape and reach scale stressors to quantify the human disturbance gradient	• Project #9: Research and evaluate different indices of human disturbance as GSG surrogates
• Need to understand why individual outlier sites have unpredictably good or bad biological condition	• Project #10: Examine BMP effects on biological condition
Implementation-related	
• Need to translate science to policy when setting stream classifications and tiered uses	• Project #11: Determine appropriate implementation criteria for identifying stream classes and tiered uses
• Consider biocriteria as a means to evaluate whether tiered uses are being achieved	• Project #12: Integrate BCG development and TALU with potential biocriteria
• Examine how other water quality objectives should be tiered along with biological uses (e.g., DO, temperature)?	• Project #13: Determine potential tiered water quality objectives
• Need to link TALU with other regulatory programs (TMDL, 401/404, stormwater) • State-wide implementation vs. region-specific approaches need to be evaluated	• Project #14: Link TALU with other regulatory programs

SPECIFIC PROJECTS

Project 1:	**Develop algal indicators of biological condition for perennial coastal California streams**
Issue:	Previous BCG development efforts were based primarily on macroinvertebrate data and assessment tools. However, macroinvertebrate data and assessment tools alone may not be sufficiently sensitive and robust to characterize perennial coastal California streams. Several examples exist including low gradient streams. Therefore, BCG development should include more than one type of indicator so that expected responses to human disturbance are accurately evaluated. Algae often respond differently to stressors, particularly nutrients, than macroinvertebrates. Therefore, inclusion of algal indicators will provide a more comprehensive BCG.
Tasks:	1. Compile existing algal data for southern California. 2. Segregate algal data and related assessment tools into various habitat types, including consideration of elevation, stream gradient, and degree of channelization. 3. Identify whether sufficient algal data is available for reference sites in southern California to develop an algal indicator. If not, identify sites and collect data as needed. 4. Correlate algal data and related assessment tools with physical or chemical stressors, land use, etc. Other stream systems can provide insight into these relationships. 5. Determine if algal data show sufficient sensitivity to stressors in southern California to serve as useful indicators of human impacts. 6. If algal indicators are sufficiently sensitive to act as useful indicators of biological condition in perennial southern California streams, select an indicator, or suite of indicators, to develop the BCG for algae. This process should be reviewed using an expert panel to verify BCG attributes for algae. 7. Set detection, precision, and accuracy estimates for the algal index developed.
Product:	Identification of algal indicators and expected changes with increasing stress. Detailed description of BCG for algal indicators.
Information Available:	Algal bioassessment methods and data collection are currently underway as part of SWAMP program. Some data is available through Western EMAP. A South Coast periphyton IBI is currently under development at SCCWRP. Additional sampling could be conducted to fill in gaps or verify correlations, as needed.
Estimated Cost:	$ 100,000 to 500,000, depending upon whether sufficient data are available
Schedule:	Two to three years, depending on availability of data
Potential Collaborators:	SCCWRP, EMAP, SWAMP, SNARL, CSUSM

Project 2:	**Develop riparian vegetation and habitat indicators suitable for BCG development**
Issue:	During the BCG Pilot Study for southern California coastal streams, the Technical Advisory Committee clearly recognized that riparian vegetation/habitat is a useful indicator of biological condition. However, use of riparian vegetation/habitat as an indicator of biological condition must be approached cautiously, as lack of vegetation/habitat can also be considered part of the stressor gradient. Preliminary work using the California Rapid Assessment Method (CRAM) was used as a placeholder absent any other standardized riparian quantification method. However, more work is needed to refine the usefulness of riparian vegetation and habitat indicators in TALU development, including identifying reference conditions and determining whether quantifiable metrics can be developed that characterize the condition gradient in response to stressor intensity.
Tasks:	1. Examine current status of CRAM to see if quantitative metrics of disturbance have been assessed. 2. If not, collate existing CRAM information along with metrics of stress or disturbance level. 3. Determine appropriate riparian/waterbody classifications (habitats) for which individual natural conditions will be defined. These could include high elevations streams, low elevation/high gradient streams, and low elevation/low gradient streams. 4. Identify specific changes in riparian indicators with stressor intensity, characterizing natural conditions as well as conditions under various levels of stress. During this process, develop a means to consider lack of vegetation due to hydrologic modification as a stressor. Identify BCG thresholds for riparian condition using CRAM. 5. Assess whether CRAM serves as an appropriate and sufficiently sensitive metric for riparian vegetation/habitat in southern California perennial coastal streams. If CRAM does not appear to be a good metric, assess whether other metrics should be used instead.
Product:	Identification of riparian indicators and expected condition gradient with increasing stress. Detailed BCG for riparian indicators.
Information Available:	Current on-going work on CRAM, including the State’s Wetland Monitoring Program; 404/401 monitoring for restoration/mitigation projects. SWAMP/Perennial Stream Assessment monitoring.
Estimated Cost:	$100,000 to 500,000, depending upon whether sufficient data are available
Schedule:	Two to three years, depending upon availability of data
Potential Collaborators:	SCCWRP, SFEI, CA Coastal Conservancy, US ACOE, Southern CA Wetland Recovery Project

Project 3:	**Define minimally impacted (natural) biological condition for coastal perennial streams and determine appropriate stream classification factors.**
Issue:	BCG development depends on having Level 1 (natural condition) defined, even if it is not represented in the region at present. The Pilot Study suggested that high elevation streams were a different class from low elevation streams, but this may not be the case and the exact elevation cutoff is unknown. The separation of stream classifications is driven largely by ecotonal gradients of physical factors and biological assemblages in the absence of stressors, i.e. a comparison of reference conditions. Identifying different classes of streams is critical because this is what determines ultimate biological expectations (i.e., low elevation or low gradient stream biological assemblages may never look like those of a high elevation or high gradient stream, even with outstanding habitat and water quality).
Tasks:	1. Compile biological indicator data, water quality data, pertinent classification metadata (elevation, gradient, geology, etc.), and stressor data. 2. Identify sites and data that are believed to represent natural conditions (Level 1) using the stressor data. If unstressed sites are unavailable, then alternative approaches can be evaluated including using sites outside of the Southern California Bight, historical information, museum archives, etc. 3. Evaluate the degree to which biological expectations differ between different coastal streams in southern California and determine classes. This is typically accomplished using multivariate statistical techniques. 4. Verify stream class determination and Level 1 attribute conditions using expert opinion.
Product:	Database of macrobenthos, other biological indicators, and pertinent physical and stressor information. Statistical analysis of biological assemblages sufficient to delineate stream classes. List and range of data for biological metrics, physical, and stressor information that characterizes Level 1 of the BCG for different classes of streams in the region.
Information Available:	Macroinvertebrate data are available from a wide range of sources including SWAMP, EMAP, SMC, NPDES monitoring, amongst others. Sufficient data may also be available for other indicators such as algae, riparian condition, and fish. (See projects 1 and 2.) SWAMP is also creating a Reference Condition Management Plan that will directly address this issue in future years.
Estimated Cost:	$150,000 - $250,000
Schedule:	One to two years.
Potential Collaborators:	SWAMP, SMC, USFS, EMAP

Project 4:	**Determine Seasonal and Interannual Variability for Relevant Biological Indicators and Identify Appropriate Ranges of Indicators for BCG Development**
Issue:	A comprehensive and accurate BCG depends, in part, on understanding and incorporating natural variability in the biological condition of the indicators. All biological indicators have some variability between seasons and between years resulting from differences in hydrological or climate regime, or innate differences in population recruitment or mortality rates. To a large extent, this type of variability has not been evaluated, creating an information gap in terms of uncertainty in biological indicator thresholds for different levels of the BCG.
Tasks:	1. Compile biological indicator data for individual sites over time. Preferably, each site will have multiple seasons and/or multiple years of record. 2. Characterize and quantify the variability of biological data, including individual metrics and composite metrics for various indicators. 3. Identify multi-year variability for given index periods and evaluate the need for a single index period in BCG development for a given indicator. Quantify appropriate ranges for individual indicators under natural conditions (Level 1 of the BCG) as well as for various stress levels.
Product:	Time-series data for specific biological indicators and sites, and statistics for seasonal and inter-annual variability based on different classes of streams. Identification of appropriate ranges of indicators to be used in setting Level 1 of the BCG.
Information Available:	Multi-year site data for macrobenthic assemblages are collected largely by NPDES permittees, although the data for reference sites may be limited. EMAP has revisited a subset of sites. The USFS has revisited some sites, but many are not in the southern California region.
Estimated Cost:	$100,000-$200,000 if data are available
Schedule:	One year
Potential Collaborators:	SWAMP, EMAP, USFS, NPDES permittees

Project 5:	**Characterize range of available biological indicator information and identify gaps in biological condition gradient**
Issue:	BCG development depends on having a complete understanding of how various biological indicators change with increasing stressor intensity. While the character of natural conditions and extremely stressed conditions is often known with some precision, changes in biological condition with intermediate levels of stress are not often as well characterized, yet this information is crucial to having a useful BCG for TALU development. Without sufficiently represented gradients of biological condition, inappropriate thresholds for BCG levels may be established. Therefore, it is critical that datasets of appropriate indicators cover the entire range of biological conditions in response to stressors. If gaps are present in the data (i.e., not enough intermediate-stressed sites), additional sampling will be needed.
Tasks:	1. Compile data sets for biological indicators, physical habitat, and stressor data. This may coordinate well with Projects 1-3. 2. Characterize the distribution of data for biological indicators and determine potential breaks or groups that may define thresholds for BCG levels, based on response of the data to stressors. Identify areas of the distribution in which there are relatively few sites represented or parts of the distribution in which there are sharp changes in indicator condition. 3. Determine if locations of missing data represent areas where thresholds will be placed. These areas of the gradient would be the prioritized data gaps for additional sampling.
Product:	Compiled data set of biological, physical habitat, and stressor information. Graphs and tables describing the distributions of each indicator. Prioritized list of data gaps requiring additional sampling.
Information Available:	For a focus on macroinvertebrates, spatially distributed data sets are preferred such as SWAMP, EMAP, PSA, SMC, USFS and others.
Estimated Cost:	$50,000 to $150,000; perhaps >$500,000 if additional sampling is included.
Schedule:	One year for data compilation and analysis
Potential Collaborators:	SWAMP, EMAP, PSA, SMC, USFS and others

Project 6:	**Determine appropriate BCG for different degrees of hydrologic modification**
Issue:	Hydromodification is one of many potential stressors. However, the pervasiveness of hydrologic modification in southern California and the significant degree to which it can impact biota makes it a particularly important stressor. Since hydrologic modification represents a stressor condition that is difficult to reverse in the short- to medium-term, this may be one basis upon which TALU is considered for southern California coastal streams (i.e., for low gradient/low elevation streams, assign tiers based on degree of hydromodification such as full channelization, concrete sides with soft botton, and unchannelized). Therefore, understanding how biological expectations change with hydrologic modification is an essential step towards refining the BCG and developing TALU in the region.
Tasks:	1. Compile biological, physical habitat, stressor condition, and water quality data as well as hydromodification attributes from existing data. This can include various biological indicators (benthic macroinvertebrates, algae, riparian vegetation, fish, amphibians, etc.) and could be done in coordination with Projects 7, 8, and 9. Develop metrics of hydrologic modification that can be scaled from natural (no modification) to extreme modification. 2. Develop a relationship between biological metrics or IBI and hydromodification metrics. 3. Verify relationships and identify a refined and comprehensive BCG that takes these relationships into account, using an expert review panel. The expert panel should help derive decision rules for weighting different data and determining BCG level based on various biological datasets (i.e., macroinvertebrates, algae, riparian vegetation, fish, amphibian, etc.).
Product:	A refined BCG based on level of hydrologic modification. Proposed tiered aquatic life uses based on varying levels of hydrologic modification.
Information Available:	SCCWRP, Counties of Ventura and Los Angeles, and the SMC are currently working on hydrologic modification projects related to erosion. For a focus on macroinvertebrates, spatially distributed data sets are preferred such as SWAMP, EMAP, PSA, SMC, USFS and others
Estimated Cost:	$50,000 to $150,000; perhaps >$500,000 if additional sampling is included.
Schedule:	Two to three years. One and one half years for data compilation and the remainder for developing the BCG
Potential Collaborators:	SWAMP, EMAP, PSA, SMC, USFS and others

Project 7:	**Evaluate and develop a refined set of physical habitat measures that help develop the GSG.**
Issue:	Physical habitat quality should be an important factor in determining biological condition expectations. Until recently, most physical habitat sampling followed protocols that were semi-quantitative and subject to large sampler-to-sampler variance. The Pilot Study showed that these highly variable, semi-quantitative physical habitat measurements were insufficiently robust for developing a predictable GSG. More quantitative, less variable, physical habitat protocols have recently been developed and are now being implemented throughout the region. These new protocols may be more useful in developing the GSG since they are more quantitative, but no one has examined their results critically for this type of TALU application.
Tasks:	1. Compile physical habitat data for sites using the new protocols along with biological data, as available. 2. Characterize the statistical distribution of various physical habitat measures. It may be useful to examine multi-metric indices of physical habitat condition. It may also be useful to differentiate the data by stream classification and degree of hydromodification. 3. Determine relationships between physical habitat metrics and biological measures. Recommend the physical habitat metrics that best predict biological responses. 4. Pilot test recommended metrics at a range of sites to evaluate the utility of the proposed physical habitat metrics.
Product:	Series of correlation plots or matrices of physical habitat metrics and biological responses. Recommend validated physical habitat metrics for use in developing the GSG.
Information Available:	EMAP has the most quantitative physical habitat measurements. SWAMP and the Perennial Stream Assessment have developed new methods for physical habitat that are derived from the EMAP protocols. The SMC will be using the SWAMP protocols in the upcoming years and the data generated could serve as the validation data set.
Estimated Cost:	$200,000 - $500,000, not including additional data collection
Schedule:	Two to three years
Potential Collaborators:	EMAP, SWAMP, PSA, SMC

Project 8:	Develop refined base maps of stressor information
Issue:	Development of a reliable GSG is dependent upon having accurate stressor information. Moreover, this information will help define the tiers for TALU implementation. Currently, insufficient stressor information exists with which to draw relationships with existing biological indicators. For example, macroinvertebrate data are available for many sites in the region, but associated stressor information is not complete. This stressor information comes in many varieties, but can be broken into two types: watershed scale and reach scale. Watershed stressors focus on large-scale cumulative impacts such as upstream land use. Reach stressors focus on local impacts such as physical habitat, flow, or water quality.
Tasks:	1. Compile data on watershed scale stressors. This may include, but is not limited to, land use, imperviousness, flow augmentation or diversions as well as associated structures (i.e., dams, reservoirs, etc.), and point source discharges. 2. Compile data on reach scale stressors. This may include, but is not limited to, stream bed material (i.e., fully channelized, concrete-lined with soft bottom, unchannelized), nonpoint source inputs, road crossings and associated structures (i.e., bridges, culverts, Arizona crossings). 3. Place all of this information into a GIS platform for use in future projects. Use the GIS to create maps of the stressor distributions. 4. Evaluate maps to ensure they are using the most up-to-date information and identify sites needing follow-up reconnaissance to ensure desired accuracy.
Product:	GIS layers and base maps of watershed and reach scale stressors.
Information Available:	Much of the watershed scale stressor information is currently available and compiled. Less information has been compiled for reach scale stressors.
Estimated Cost:	$250,000 to $500,000
Schedule:	One to three years, depending on number of stressors and scale.
Potential Collaborators:	DWR, SCAG/SANDAG, most public works and flood control agencies, NOAA.

Project 9:	**Research and evaluate different indices of human disturbance as GSG surrogates**
Issue:	There are myriad of biological stressors, which often have cumulative impacts on southern California streams. Successful TALU delineations depend on having a clear understanding of these stressors and their gradations (i.e., the GSG). Through the process of defining GSG attributes, stakeholders can determine which stressors are controllable (and therefore, not an appropriate aspect of tiered uses) and which are not readily controllable (and might make for good attributes to use in defining tiers). Previously, only landscape scale stressors were evaluated. However, these large-scale stressor evaluations were incomplete and virtually no reach-scale stressors appeared adequate for describing biological response in the biological indices examined to date (i.e., macroinvertebrates). The goal of this project is to improve the GSG for developing TALU.
Tasks:	1. Compile the existing knowledge of stressor indices from the literature, particularly those used in other water programs. 2. Use the existing knowledge from task 1 to create metrics to characterize stressors. This may include multi-metric approaches. 3. Evaluate the biological responses along each stressor metric gradient to identify the best (most predictive) approach. Conduct this process with several types of biological responses to determine the most sensitive biological response to stress. 4. Verify the pros and cons of potential stressor metrics and select preferred approach using an expert review panel. 5. Create a GIS map of stressor metrics for perennial streams region wide.
Information Available:	There are a number of stressor metrics recently developed and published in the literature. Land cover data are readily available, but should be checked for currency and accuracy (see Project 8). Hydrologic as well as physiochemical data are available from several sites and time periods. Where data do not exist, a targeted sampling program may be required.
Product:	Literature review of existing approaches to stressor identification. Series of correlation plots or matrices of stressor metrics and biological responses. Recommended GSG options for use in developing TALU.
Estimated Cost:	$200,000 - $400,000
Schedule:	Two to three years
Potential Collaborators:	SWAMP, NPDES permittees, USGS, DWR

Project 10:	**Examine BMP effects on biological condition**
Issue:	Condition assessments from the Pilot Study indicated that some sites had relatively “good” biological condition considering the level of stressors such as surrounding land use. Similarly, some sites had relatively “poor” biological condition despite an apparent lack of significant stressor sources. The initial assumption has been that unique, site-specific circumstances help dictate the outlier conditions of these sites. To determine whether site-specific circumstances are the cause of the outlier conditions, sites that are uncharacteristically “good” or “bad” should be examined to determine if this is a result of specific practices, such as BMPs or the presence of industrial discharges. This analysis can help determine whether the indicators are appropriate, and potentially identify the key physical and/or hydrologic factors that can help improve degraded sites.
Tasks:	1. Using the compiled data set from Projects 6, 8, and 9, look for anomalous sites that do not fit the BCG/GSG relationship. 2. Conduct site reconnaissance to determine site-specific factors, including BMPs or specific discharges, if any. 3. Based on BMPs or other factors that yielded better than expected biological condition, recommend approaches that may help improve other lower quality sites (e.g., BCG Level 5 or 6). An alternative is to work with agencies that are preparing to install BMPs to test BMP effectiveness. 4. Recommend procedures for handling outlier or anomalous sites within a TALU framework.
Product:	Report with maps showing outlier sites and evaluation of factors causing site-specific condition. Create a list of BMPs that will improve biological condition at these sites. Guidelines for dealing with outlier sites in TALU implementation where site-specific factors need to be accounted for.
Information Available:	SWAMP and the Perennial Stream Assessment have a large number of sites that can contain outliers for investigation. SCCWRP has just completed an assessment of BMPs for habitat restoration.
Estimated Cost:	$100,000 to $200,000, more if sampling or BMP construction is required.
Schedule:	One to two years
Potential Collaborators:	SWAMP, EMAP, PSA

Project 11:	**Determine appropriate implementation criteria for identifying stream classes and tiered uses**
Issue:	BCG and GSG-related projects will determine appropriate classes of perennial streams in Southern California, within which more specific aquatic life uses can be defined. To implement this classification, there needs to be objective science-based criteria for distinguishing classes so that water quality standards can clearly identify to which class a given segment belongs. However, there are policy implications for how stream classifications are attributed. It is this intersection of science and policy that requires thoughtful implementation to ensure equity, effectiveness, and cost efficiency. Several questions need to be answered such as, if classification is based on elevation (or gradient), what is the specific cutoff for high vs. low elevation streams (or high vs. low gradient); are there exceptions to this classification; and how is this classification scheme best applied to ensure efficient implementation of TALU? Similarly, TALU tier thresholds are derived from application of scientific information, but these thresholds need to be re-evaluated once they are applied to actual stream reaches to ensure the biological expectations are appropriate.
Tasks:	1. Compile, summarize, and analyze statistically the database from Projects 3, 4, 6, 8, and 9 will be to identify stream classes that should be considered for separate TALU "regions". This will be done in a pilot watershed. 2. Conduct GIS analysis and create a map of stream classification assignments and proposed tiered uses in the pilot watershed. 3. Evaluate the stream assignments to confirm appropriate classes and tiered uses within each class using a task force of scientists, regulatory and regulated agency staff, as well as nongovernmental organizations. While the goal is not to agree on every stream reach assignment, this project will help to define a framework for conducting the public process in the remainder of the region.
Product:	Framework document detailing the criteria and process for assigning stream classifications and tiered uses.
Information Available:	Results of Projects 3, 4, 6, 8, and 9
Estimated Cost:	$75,000 – 150,000
Schedule:	One year
Potential Collaborators:	Regulatory agencies and regulated stakeholders

Project 12:	**Integrate BCG and TALU development with potential biocriteria**
Issue:	Formulation of tiered aquatic life uses will be most useful if there are appropriate criteria available to ensure protection of waterbodies within each tier. Currently, no biocriteria have been established as regulatory water quality standards for southern California streams although the Southern California IBI for macroinvertebrates has been suggested. On-going algae work, including that proposed in Project 1, could provide information with which to develop biocriteria for algae, if algae criteria can be developed that serve as good indicators of biological condition. If appropriate biocriteria can be formulated, they could be used as measurement benchmarks with which to evaluate impairments and restoration progress as well as document protection of different aquatic life uses.
Tasks:	1. Establish a task force consisting of regulatory, regulated, and nongovernmental agencies to provide a context for biocriteria interpretation. This group may best be served by using a regulatory agency as the lead. 2. Create a framework that maps the relationship between beneficial uses in basin plans, biocriteria, use attainability analysis, and antidegradation policies. Data compiled and used as part of this workplan should help immensely. 3. Write a consensus-based white paper outlining the regulatory model that can be used as the basis for integrated policy development.
Product:	White paper outlining the regulatory model that can be used as the basis for integrated policy development
Estimated Cost:	$75,000-$150,000
Schedule:	One to two years
Potential Collaborators:	Regulatory agencies and regulated stakeholders

Project 13:	Determine potential tiered water quality objectives
Issue:	In developing tiered aquatic life uses, it may be appropriate to modify water quality objectives to reflect what is necessary to obtain and maintain aquatic life uses for that tier. For example, if a high quality tiered aquatic life use is identified (and supported by both BCG and available biological condition data), it may be critical to have more stringent water quality objectives for certain parameters, such as oxygen, temperature, sediment, and possibly certain chemical pollutants, than are necessary for more standard aquatic life uses. Likewise, if a tiered use is identified for highly modified waterbodies, it may be desirable to modify objectives in cases where a less stringent objective may be adequately protective. Tiered or modified water quality objectives may not be appropriate for certain types of parameters. While there have been some evaluations of this issue at the national level, no guidance has been developed. If and how objectives are modified in concert with TALU will have a direct bearing on how TALU is implemented.
Tasks:	1. Convene a workshop consisting of regulatory agencies, resource agencies, and invited scientists to discuss appropriate actions in tasks 2-3 below. 2. Evaluate what EPA and others have considered, and list the pros and cons of different strategies for dealing with tiered water quality objectives. 3. Identify a preliminary list of parameters for possible tiering, as well as a list of parameters for which tiered objectives would be inappropriate. 4. Identify a pilot study to test the feasibility of tiered water quality objectives. Where possible, actual data for parameters should be examined from segments representing all tiers.
Product:	Topical Workshop. Position paper recommending results of evaluation and parameters potentially subject to tiering, if any. Design for Pilot Study.
Estimated Cost:	$50,000 to $75,000
Schedule:	Six months to one year
Potential Collaborators:	Regulatory and regulated entities.

Project 14:	**Link TALU with other regulatory programs**
Issue:	Local, State, and Federal regulatory programs do not operate in isolation from one another. Water quality standards, biocriteria, total maximum daily loads (TMDLs), NPDES permitting, 401/404 certification for streambed alteration are just a few examples. Optimizing the interplay between regulatory programs and regulatory agencies will help reduce redundancy and increase effectiveness of the regulatory framework. This will be particularly important in determining if TALU should be initiated at the local, regional, or statewide level.
Tasks:	This project will require two tasks. First, a policy committee should be gathered to help evaluate optimal implementation strategies. This policy committee should contain representatives from regulatory, regulated, and environmental advocacy organizations. Regulatory program representation should include RWQCB, SWRCB, and EPA. Second, the committee should draft an implementation workplan to coordinate efforts.
Product:	Implementation strategy workplan.
Information Available:	There are other examples that can serve as a model for this Committee including the State's Sediment Quality Objectives.
Estimated Cost:	$100,000 to $200,000
Schedule:	Two years
Potential Collaborators:	Regulatory and regulated entities

PROJECT INTEGRATION AND SYNTHESIS

The projects outlined in the previous section are designed to address major data gaps in our understanding of biological responses to stressors in southern California perennial streams and how the stressor axis of the BCG should be constructed and applied. These projects are necessary to formulate a scientifically defensible framework upon which tiered aquatic life uses can be developed and implemented. To make the most efficient use of available resources, certain projects should be completed or at least largely completed prior to others. Ideally, regulators and stakeholders would cooperatively lay out the TALU development framework in order to make the process efficient, effective, and transparent. To that end, we see projects being conducted in four phases, understanding that there will be (and should be) some overlap in the timing of different phases so that the process is as efficient as possible.

In the first phase, basic information is needed regarding biological responses to stressors, characterizing the stressor gradient, and the types of data available for BCG analyses. Therefore, Project #3 (natural condition definition and appropriate classification) and Project #5 (characterize range of biological condition data available) should be initial priorities. Unless these projects are addressed, subsequent BCG or GSG-related projects may be flawed or incomplete. Simultaneously, Project #7 (improve physical habitat measures to develop the GSG), Project #8 (improved base maps for stressors), and Project #9 (evaluate indices of human disturbance) should also be first phase projects of high priority. Results of Projects 7, 8, and 9 will be instrumental in developing a sound GSG axis with which subsequent BCG development can occur. The outcome of the first phase of projects will be:

- A better understanding of how natural condition should be described biologically
- Available data or information to characterize Level 1 of the BCG (at least for macroinvertebrates)
- Degree to which the full range of biological condition is represented using available site data for the southern California
- Preferred ways to characterize the stressor gradient and data refinements needed to define and quantify the GSG
- Refinements to physical habitat metrics and results that will feed into the GSG characterization and provide useful information for other programs and applications
- More informative base maps to allow better characterization of the range of stressor intensity represented using current biological sites

A second phase of projects would build on the ones noted above, refining the BCG further using other assemblage data (algae, Project #1, and riparian vegetation, Project #2). The inclusion of algae and riparian vegetation condition attributes is considered key to making the BCG more robust and scientifically defensible. The inclusion of these assemblages, as well as macroinvertebrates (and fish or other vertebrates to the extent possible), will ensure that a broader range of effects of stressors are included in the BCG and properly interpreted. The timing of these projects would also allow completion of current algal and CRAM data collection efforts, which will be instrumental in completing Projects 1 and 2. Results of Phase 2 would be a more comprehensive BCG that can now be refined in Phase 3 using expert consensus and site-specific information.

The third phase of projects would further refine and ultimately complete previous work in the form of more complete, robust BCG characterization (Project #6), and consideration of ways that may be effective in restoring certain tiers of aquatic life uses in some cases (Project #10, evaluate effects of BMPs and other site-specific factors on biological condition). The analysis of more site-specific biological-stressor relationships (Project #10) is neither necessary, nor desirable when formulating the BCG for a region (Phases 1 and 2) but is useful once a regional BCG is developed and the beginnings of implementation are being considered. Site-specific relationships can also be helpful in validating the BCG and determining the types of stream conditions that may be highest priority for restoration efforts.

The fourth and final phase of projects addresses TALU implementation issues (Projects 11, 12, and 13). In order to develop appropriate implementation criteria for stream classification, tiered uses, biocriteria, and appropriateness of tiered water quality objectives, a well-characterized and accepted BCG (including a robust GSG) is critical. The science provided in the first 3 phases will help guide appropriate implementation strategies. While biocriteria can be developed without TALU, implementation of biocriteria in the context of TALU is likely to have greater environmental benefits, be easier for regulatory agencies to implement in the long run, and be more defensible to stakeholders. Phase 4 projects could start as Phase 2 projects are being completed, once better information becomes available to characterize the BCG and GSG. However, Phase 4 implementation projects are not likely to be completed until after BCG development is complete (Phase 3).

While approximate costs are provided in the project descriptions, the estimates are by no means rigorous and there are many opportunities for cost savings by leveraging among projects and outside studies. For example, there are at least eight projects that rely on compiled databases of biological condition, hydrology, physical habitat, and stressor information. Obviously, this needs only to be done once and, even then, portions will be done in individual project development (i.e., stressor specific information, Project 8). Another example would be the formation of expert panels and task force committees. Virtually every project would benefit from the use of independent, multi-sector review as a means for oversight, validation, and transparency. These committees are crucial to success, but a new committee is not needed for every study. One committee could take on the challenge of several projects, especially if the projects are similar in nature such as those described within each of the implementation phases. Finally, the potential collaborators for these projects were repeated over and over again. An integrated approach with multiple agencies attacking these data gaps will increase the cost leveraging necessary to overcome the hurdles to achieving TALU.

REFERENCES

Ode, P.R., A.C. Rehn and J.T. May. 2005. A quantitative tool for assessing the integrity of southern California coastal streams. *Environmental Management* 35:493-504.

OEPA. 2008. State of Ohio Water Quality Standards. Chapter 3745-1 of the Administrative Code. Effective April 23, 2008. Ohio Environmental Protection Agency, Division of Surface Water, Standards and Technical Support Section. Columbus, OH.

Tetra Tech. 2005. Evaluation of Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams. Draft Summary Report. Prepared for: EPA Region 9 and California Los Angeles Regional Water Quality Board. Tetra Tech, Inc. Owings Mills, MD.

Tetra Tech. 2006. Revised Analyses of Biological Data to Evaluate Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams. Prepared for: EPA Region 9 and California Los Angeles Regional Water Quality Board. Tetra Tech, Inc. Owings Mills, MD.

Appendix 8

B263

Revised Analyses of Biological Data to Evaluate Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams

Prepared For:
Susan Jackson (EPA HECD)
Terry Fleming (EPA Region 9)

Prepared By:
Tetra Tech, Inc.
400 Red Brook Blvd
Suite 200
Owings Mills, MD 21117

December 8, 2006

Data Report: Revised Analyses of Biological Data to Evaluate Tiered Aquatic Life Uses (TALU) for Southern California Coastal Streams

Introduction

Under a previous work assignment with EPA Region 9 and the Los Angeles Regional Water Quality Control Board, Tetra Tech used available biological and habitat quality data (provided primarily by EMAP), as well as information provided by local and regional experts, to develop a preliminary Biological Condition Gradient (BCG), which is a framework that characterizes changes in biological condition going from undisturbed (reference) to very impaired conditions (Davies and Jackson, 2006). The range of potential impaired conditions encountered in the region constitutes the Generalized Stressor Gradient (GSG), which is a framework that characterizes changes in stressor attributes going from undisturbed to very impaired conditions (Davies and Jackson, 2006). In order to develop a defensible framework for tiered aquatic life uses (TALU), streams in the region need to be categorized with respect to their biological expectations considering the types of classes that either occur naturally or that are distinguishable based on what are major habitat alterations due to anthropogenic factors.

Since the initial work was completed by Tetra Tech, several other sources of macroinvertebrate and habitat data became available, primarily through California's Statewide Assessment and Monitoring Program (SWAMP) as well as other sources. These data provided substantially more information on the low elevation, urbanized streams in the region (e.g., in and around Los Angeles and San Diego), a major data gap identified by Tetra Tech in the previous work. As a result, we were able to more confidently identify the range of biological conditions currently observed in streams affected to varying degrees by anthropogenic alterations. Through these analyses, the revised results presented in this report should provide more confidence in terms of how streams might be classified in the region, and ultimately, potential tiered aquatic life use definitions.

Tetra Tech previously incorporated several suggestions from Technical Advisory Committee (TAC) members in the region regarding the types of attributes that should be considered in developing the BCG and the GSG for the region. As noted previously, certain attributes identified in EPA's national BCG framework were either modified or removed for the southern California region because they are either not relevant to this region or were better incorporated as part of the generalized stressor gradient (GSG). Key biological characteristics that were included in the BCG are: (1) Southern California IBI and component metrics developed by Department of Fish and Game (DFG) for macroinvertebrates; (2) fish assemblage information obtained from Drs. Jonathan Baskin, Thomas Haglund, and Camm Swift; (3) and algae diatom information obtained from EPA's Rapid Bioassessment Protocols and Western EMAP sources.

This revised report updates the macroinvertebrate attribute information for the BCG based on the new data evaluated. Presented here is a conceptual BCG that is intended to serve as a precursor to a final, fully calibrated BCG that could be used in the TALU framework or in Use Attainability Analyses (UAA). Other biological information was not updated in this exercise. We would note that new periphyton information being collected in the region by the Southern California Coastal Water Research Project (SCCWRP) and by Tetra Tech could be very useful in further refining the BCG in the future. We would also note that the TAC felt that the BCG attribute long-lived or regionally endemic species may be especially useful in terms of discriminating biological condition over the stressor gradient in this region. This attribute is characterized mostly in terms of vertebrate species information (number or types of fish, amphibian and reptile species) since these species are relatively long-lived and/or endemic to a particular drainage or watershed in this region. The TAC agreed that better information concerning these types of species would be very beneficial in refining the BCG and perhaps aquatic life uses as well.

Data Sources

Additional macroinvertebrate data used in these analyses were obtained from California Department of Fish and Game (Pete Ode) and from EPA Region 9 (Terry Fleming). Data for approximately 1700 benthic macroinvertebrate samples and physical habitat assessments were compiled, along with geographical coordinates at over 300 sites in southern California between 1998 and 2005. Biological data included data for the seven different metrics, which comprise the Southern California IBI (SoCal IBI), as well as the IBI score for each sample (Table 1). Habitat assessments were based on the Rapid Bioassessment Protocols (Barbour et al. 1999) and included data scores for the 10 different parameters on a 0-20 scale (0 poor, 20 optimal) as well as the total habitat score for each site (Table 1).

Table 1. Biological metrics and physical habitat parameters used in analyses.

Biological Metrics	**Physical Habitat Parameters**
EPT taxa	Epifaunal substrate
Intolerant taxa percent	Sediment deposition
Predator taxa	Embeddedness
Coleoptera taxa	Riffle frequency
Non-insect percent	Channel alteration
Tolerant taxa percent	Channel flow
Collector percent	Bank vegetative protection
	Bank stability
	Velocity/ depth regime
	Riparian zone width
Southern California IBI	**Total Habitat Score**

In addition to instream physical habitat measures, the stressor gradient was characterized by landscape influences on sampling locations. For each location, 5 km radius circles were delineated and land use/land cover (LULC) percentages (MRLC 1992) were calculated within these circles to represent general landscape activities in the vicinity of the sample sites. These LULC percentages were used to calculate a Landscape Development Intensity (LDI) index (Brown and Vivas 2006) that weights each land use type base on the energy that each uses. Potential LDI index scores range from 1 to 10 with 1 representing natural systems and 10 representing the most intense urban land uses. Agricultural land uses have LDI coefficients between 2 (low intensity pasture) and 7 (high intensity feed lots, dairy farms, etc.). Urban land uses have LDI coefficients that range between 7 (low density residential) and 10 (central business district). This LDI index is used as another indicator of the stressor gradient as it serves as a surrogate for chemical and hydrologic impacts, which may not be included in instream physical habitat measures. LDI has been used by Florida in its biological assessment program (Fore 2004) and is particularly useful for distinguishing an urbanized gradient.

Preliminary Stream Classification

Natural variations in streams of this region can be attributed generally to differences in elevation. Through basic knowledge of the study area, as well as inspection of aerial photographs, it was determined that an elevation of 1200 feet appeared to be a relatively reliable threshold for distinguishing between higher and lower gradient stream systems. Using this elevation threshold, four types of site classes were identified with which BCG attributes were evaluated:

1) natural high elevation foothills (>1200 ft),
2) natural low elevation (<1200 ft),
3) low elevation partially altered channel or riparian zone,
4) low elevation concrete-lined channel.

Sites were grouped into one of these categories based on visual inspection of aerial photographs of each site and its surrounding area. These four stream classes cover the range of stressor and biological conditions observed in the Southern California Bight region. In addition, these four classes were clearly distinguishable from each other visually and were thought to be distinct ecologically as well.

Stressor Measures in Relation to Stream Classes

Median habitat scores were related to natural and anthropogenic influences as represented by the four site classes (Figure 1). Habitat scores were also related to LDI index scores, demonstrating a relationship between habitat quality and overall landscape stress (Figure 2).

Macroinvertebrate Data

BCG attributes that were refined based on the updated macroinvertebrate data included attributes 3, 4, and 5. Other BCG attributes remained unchanged from the previous version developed by Tetra Tech because there were no new data or other information that would help further refine other BCG attributes. In conducting these analyses, we compiled relevant macroinvertebrate metric data for each attribute by stream class as defined in previous work and as noted above. One of the key questions examined in this exercise is whether the initial classifications used previously continue to be scientifically defensible given the more extensive biological data made available.

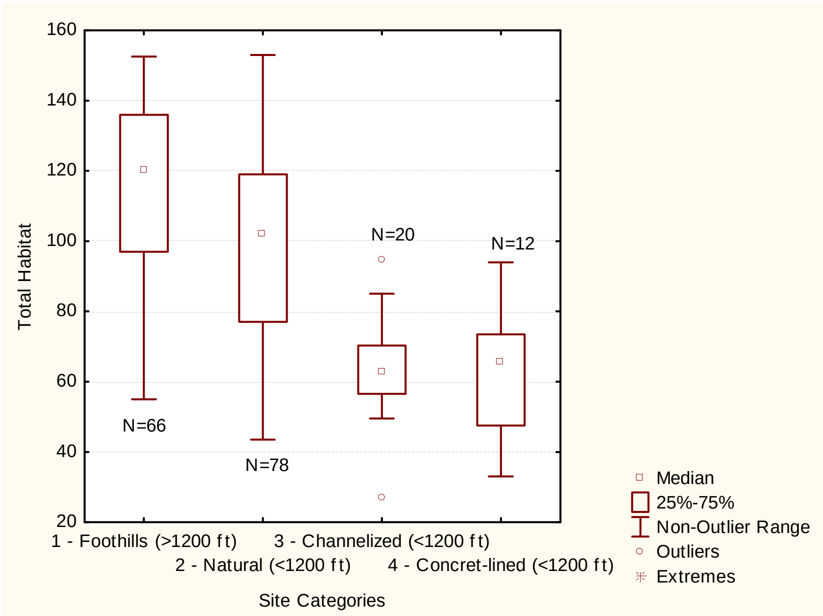


Figure 1. Total habitat scores organized among four site categories

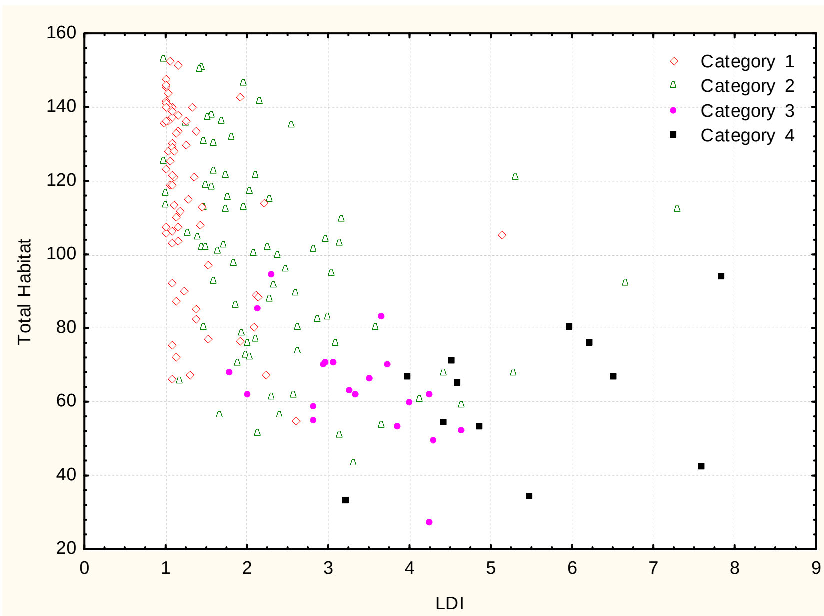


Figure 2. Total habitat scores versus LDI index scores organized among four site categories.

Scatterplots of the SoCal IBI and biological metrics versus habitat assessment score and LDI index score were used to examine relationships between habitat condition and overall landscape stress on macroinvertebrate assemblages. In addition, biological data were categorized according to the four site classes to illustrate variability within site classes in terms of response to stress. Non-parametric Kruskal-Wallis tests (at an alpha = 0.05) were used to statistically evaluate differences in results among the four site categories.

Results

Southern California IBI scores ranged from 0 to approximately 90 and about 60 percent of the sites were impaired according to the classifications developed by Ode et al. (2005) (i.e., IBI scores less than 40) (Figure 3). For the two selected metric distributions (Figure 3), about 8 percent of the sites had no EPT taxa and approximately 40 percent of the sites had percent non-insect less than 10%. Sites located above 1200 ft elevation generally had higher IBI and sensitive metric scores than those found below 1200 ft (Figures 4 and 5). Approximately 30 percent of the sites above 1200 ft were impaired, while 80 percent of the sites below 1200 ft were impaired and half of these were rated as very poor. For the non-insect percent metric, about 50 percent of the sites above 1200 ft had non-insect percents less than 20%, while about 70 percent of the sites below this elevation had values less than 20%. EPT taxa values had relatively similar distributions among the two elevation categories.

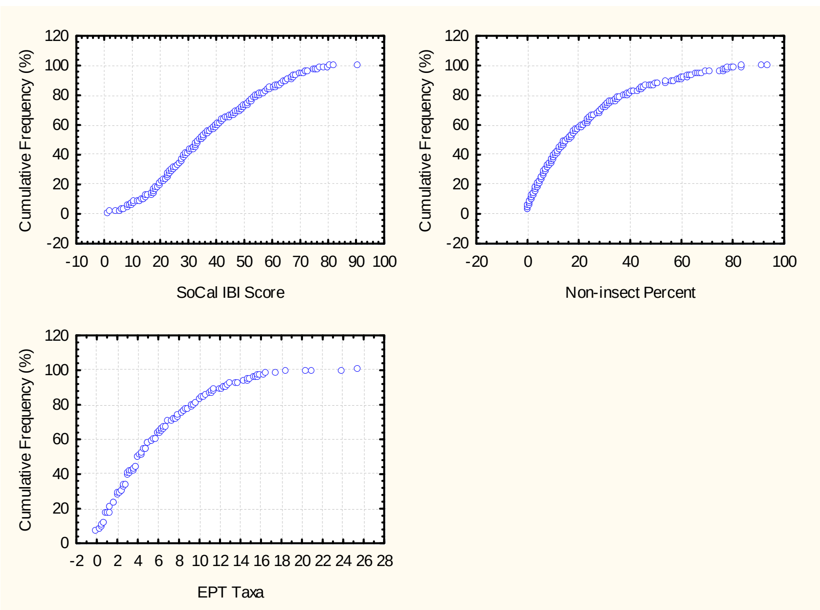


Figure 3. Cumulative frequency distribution plots for the SoCal IBI and two example metrics, intolerant percent and EPT taxa.

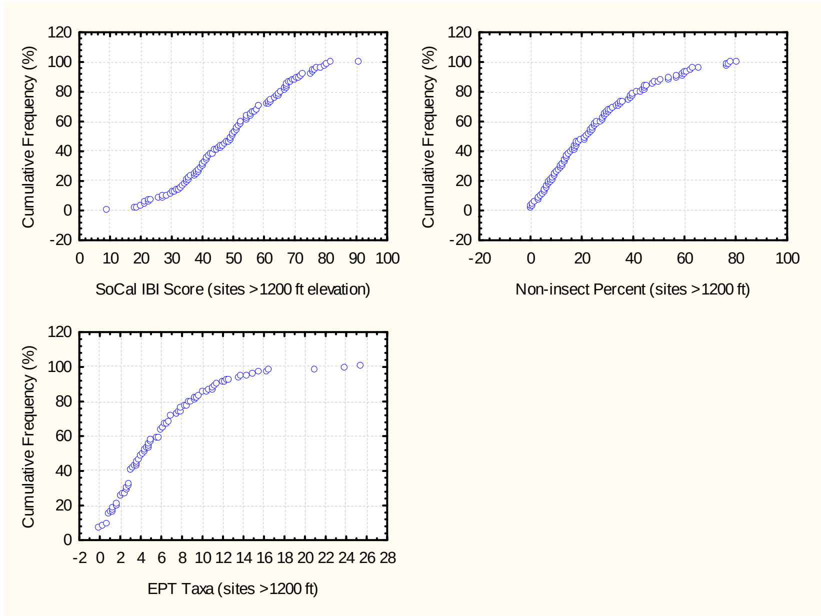


Figure 4. Cumulative frequency distribution plots for the SoCal IBI and two example metrics for sites located at elevations greater than 1200 feet.

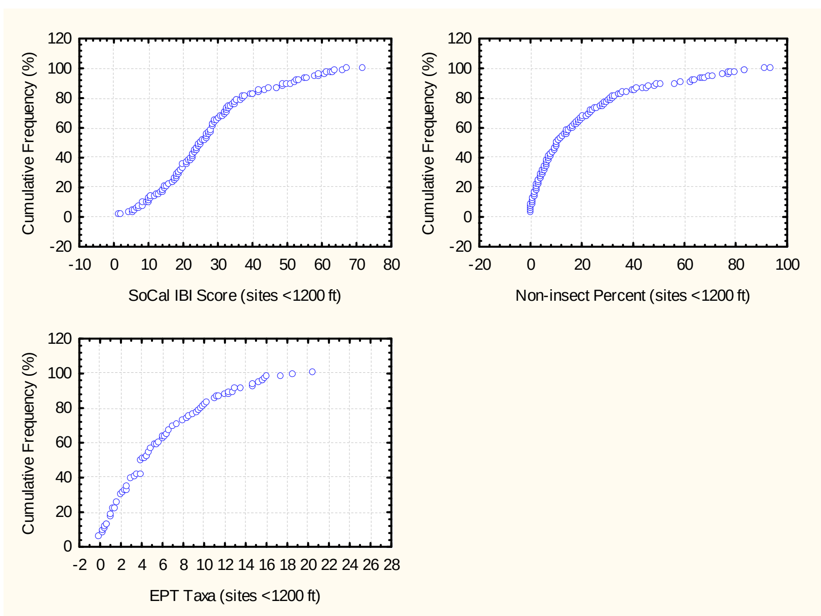


Figure 5. Cumulative frequency distribution plots for the SoCal IBI and two example metrics for sites located at elevations lower than 1200 feet.

Southern California IBI

As shown in Figure 6 the SoCal IBI scores were higher in natural channel sites (both >1200 ft and <1200 ft) than at human-altered sites (both partially altered and concrete lined categories). A non-parametric Kruskal-Wallis test confirmed that the two natural categories (1 and 2) were significantly different (p<0.05) from one another and each was significantly different from both of the human-altered site classes (3 and 4). SoCal IBI scores, however, were not significantly different between the two altered site classes. The following summarizes relationships regarding three of the BCG attributes that were subject to change based on the additional data in this analysis, and site classification. The three BCG attributes examined were: (1) sensitive ubiquitous taxa, (2) taxa of intermediate tolerance, and (3) tolerant taxa.

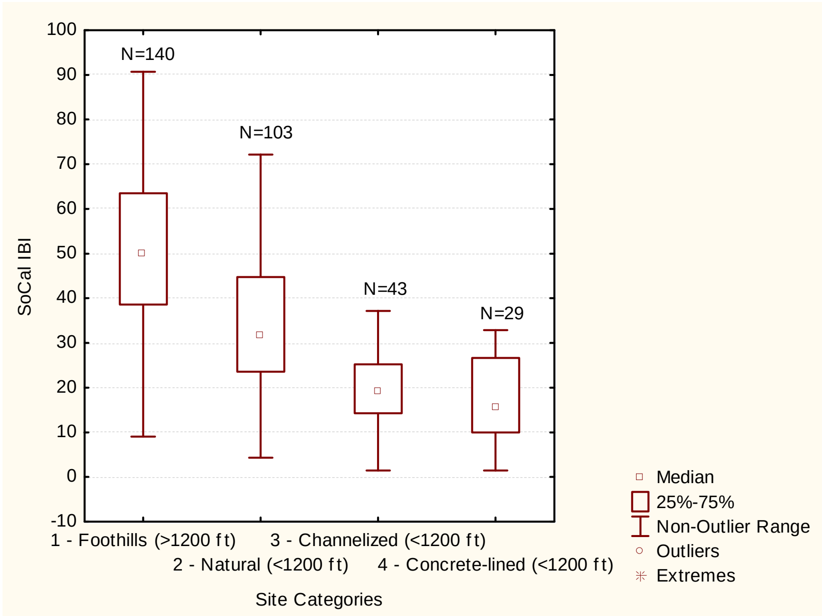


Figure 6. Southern California IBI scores in relation to the four site class categories used in this evaluation.

Attribute 3: Sensitive Ubiquitous Taxa

The Southern California IBI developed by DFG and others (Ode et al. 2005) includes four metrics that represent sensitive ubiquitous macroinvertebrate taxa: intolerant percent, number of EPT taxa, Coleoptera taxa, and number of predator species. All four sensitive ubiquitous taxa metrics showed similar patterns in response to the four site class categories (Figure 7). For intolerant percent, EPT taxa, and Coleoptera taxa, the two impacted classes of sites did not appear to be different from one another. For all four metrics, values for the two classes of natural sites were noticeably different from the two impacted classes. Additionally, the foothills class (i.e., category 1) was substantially different than the other natural site class (<1200 ft). A Kruskal-Wallis test on all the four metrics showed that all groups were significantly different (p<0.05) from one another, except the two altered classes which were statistically the same. Although predator taxa values among the two impacted classes (3 and 4) appeared different (Figure 4), the Kruskal-Wallis test indicated that this difference was not significant at an alpha level = 0.05.

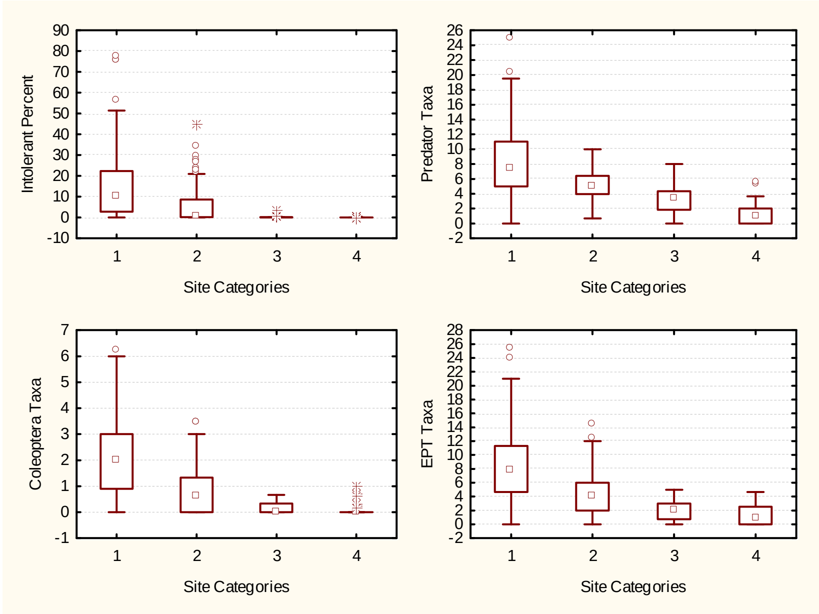


Figure 7. Intolerant percent, predator taxa, Coleoptera taxa, and EPT taxa reported in benthic samples as a function of four site class categories (see text for description of site categories).

Attribute 4: Taxa of Intermediate Tolerance

The SoCal IBI does not have a metric that includes only intermediate tolerant taxa. However the TAC recognized certain taxa that they considered to be representative of this attribute. These taxa included the caddisfly *Hydropsyche*, the mayfly *Baetis*, and elmid beetles. Dominance of these taxa is thought to signify fair – poor biological condition in this region. However, Figures 8 and 9 suggest otherwise – *Baetis* and *Hydropsyche* percent were lowest in site categories 3 and 4 (altered channels) and declined in response to increasing landscape disturbance as represented by LDI scores. Kruskal-Wallis tests confirmed these differences. Percent *Baetis* was significantly different ($p<0.05$) between category 1 and the two altered classes and category 2 was significantly different from category 4. The two natural stream class categories, as well as categories 2 and 3, and 3 and 4, were statistically the same. For percent *Hydropsyche*, the two natural categories were statistically the same, as were the two altered categories; otherwise, all categories were significantly different from one another.

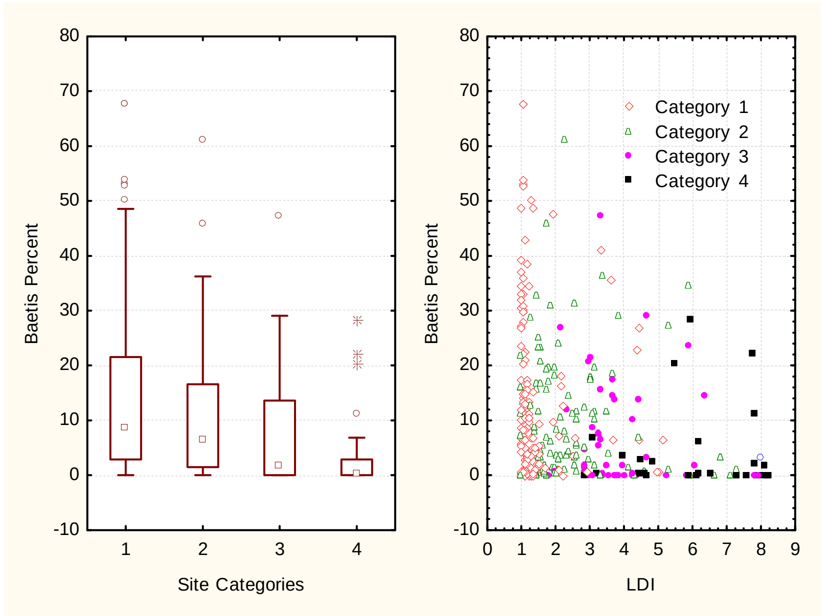


Figure 8. Baetis percent among four site categories and plotted versus LDI scores

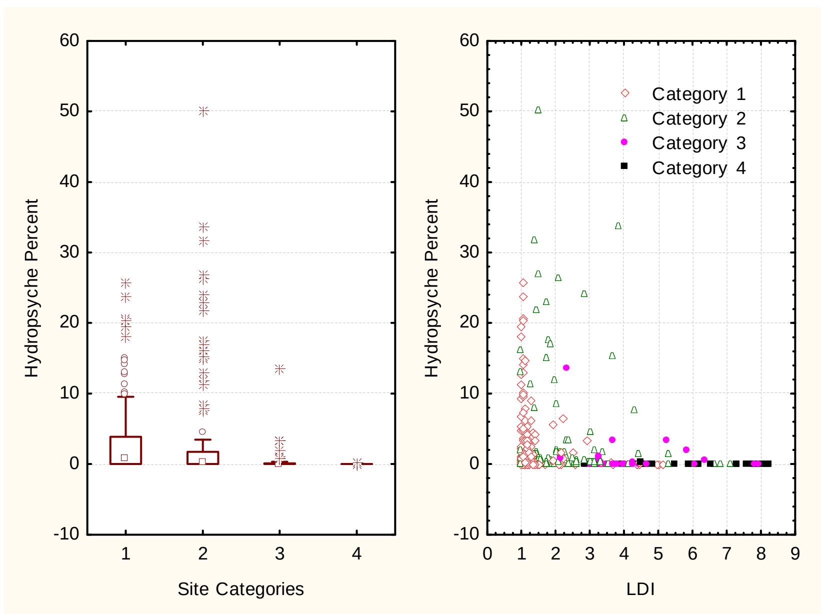


Figure 9. Hydropsyche percent among four site categories and plotted versus LDI scores

Attribute 5: Tolerant Taxa

The SoCal IBI includes three metrics that are indicative of tolerant taxa: percent collectors, number of non-insect taxa, and percent tolerant taxa. The percent collector metric showed a gradual increase from natural foothill (>1200 ft) streams (Category 1) to the concrete lined channels (Category 4) (Figure 10). Non-insect and percent tolerant metric scores were actually higher at the partially-altered sites than at the concrete lined sites. In fact, for these two metrics, concrete-lined channels appeared to be similar to both types of natural stream classes (Categories 1 and 2). A Kruskal-Wallis test on the non-insect taxa metric values indicated that categories 1 and 2 (natural sites) were not significantly different ($p>0.05$) from category 4 (concrete-lined channels); all other categories were significantly different from one another. For the percent tolerant metric, categories two and three were statistically the same as category 4, while category 1 was significantly different from all the other categories. A Kruskal-Wallis test on the percent collector metric indicated that all categories were significantly different from one another except categories 2 and 3 (low elevation natural channel and channelized), which were statistically the same.

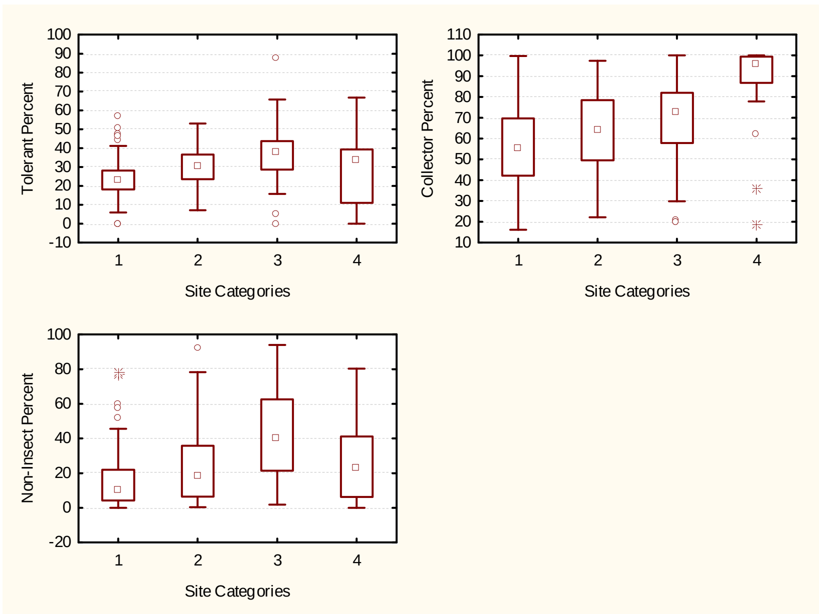


Figure 10. Percent tolerant taxa, percent collector taxa, and percent non-insect taxa reported in benthic samples as a function of four site class categories (see text for description of site categories).

Refinement of the BCG

Based on our method of site classification, we could not distinguish biologically, partially altered channels from concrete-lined channels for the majority of metrics, as well as the SoCal IBI; i.e., the concrete-lined channels can apparently achieve biological condition levels similar to those observed in partially altered low elevation streams. As we observed in the previous work, higher elevation streams have a higher biological expectation than lower elevation streams in the region, independent of the degree of channel alteration. In addition, the types of taxa often observed in the higher elevation cooler streams is different than those observed in the warmer lower elevation streams. This is borne out by the fishery information as well. While the exact elevation threshold to be used to separate low from high elevation stream classes is somewhat flexible (we used 1200 feet elevation), there are scientific data to support distinguishing higher elevation streams from lower elevation streams in terms of biological expectations. Use

Attainability Analyses might be necessary in some cases to clarify whether a borderline stream segment belongs to the lower or higher elevation stream class.

Figures 11 and 12 show relationships between the SoCal IBI, its component metrics, and increasing stress, as measured by either stream habitat quality score or the LDI index. The SoCal IBI and metrics were generally responsive to habitat degradation (Figure 11) and overall landscape alteration (Figure 12). Particular metrics that appeared most related to both habitat and LDI scores are percent tolerant taxa, predator taxa, and EPT taxa.

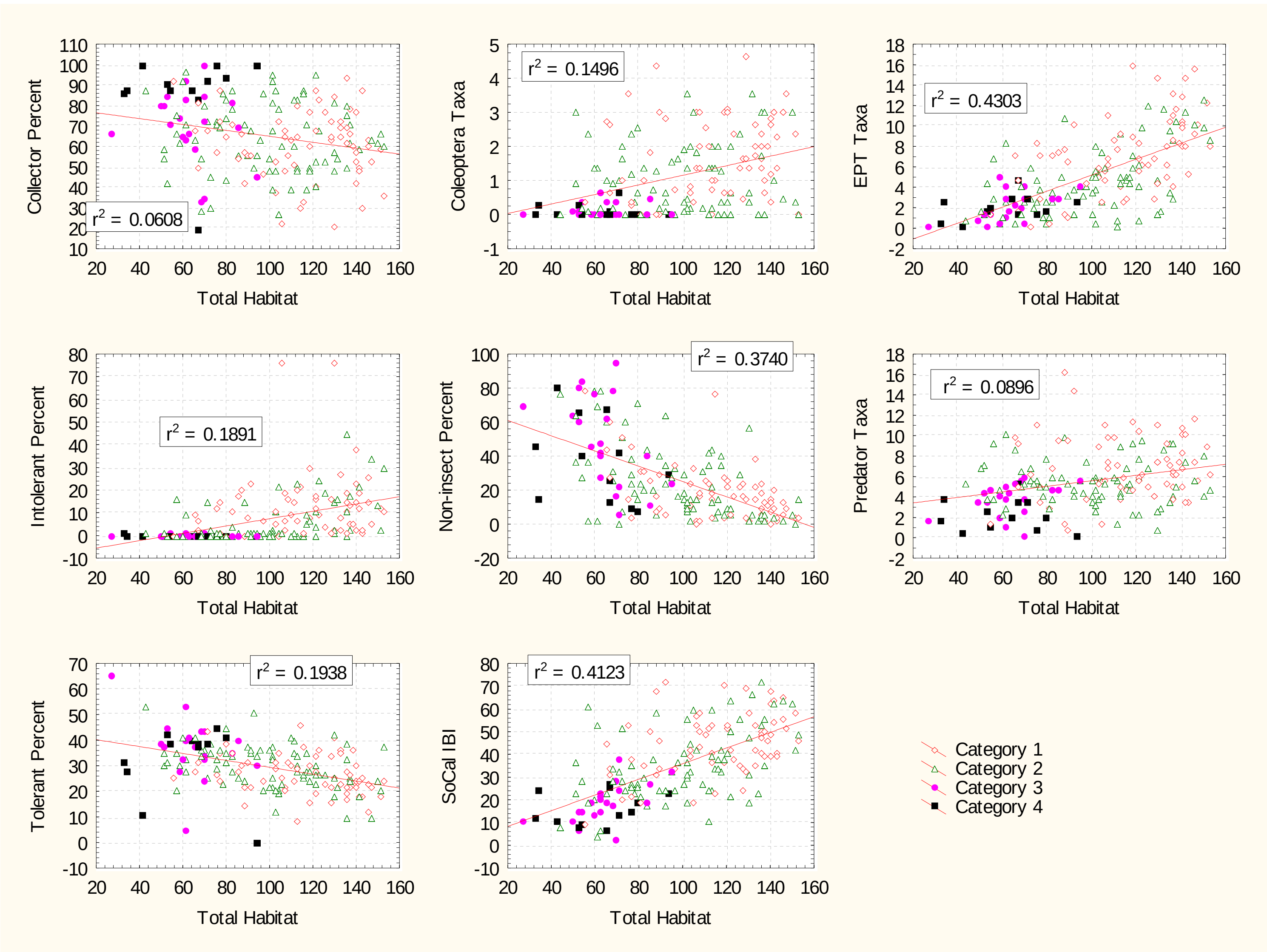


Figure 11. SoCal IBI and associated metrics versus total habitat scores organized among four site categories. All correlations were significant ($p<0.05$).

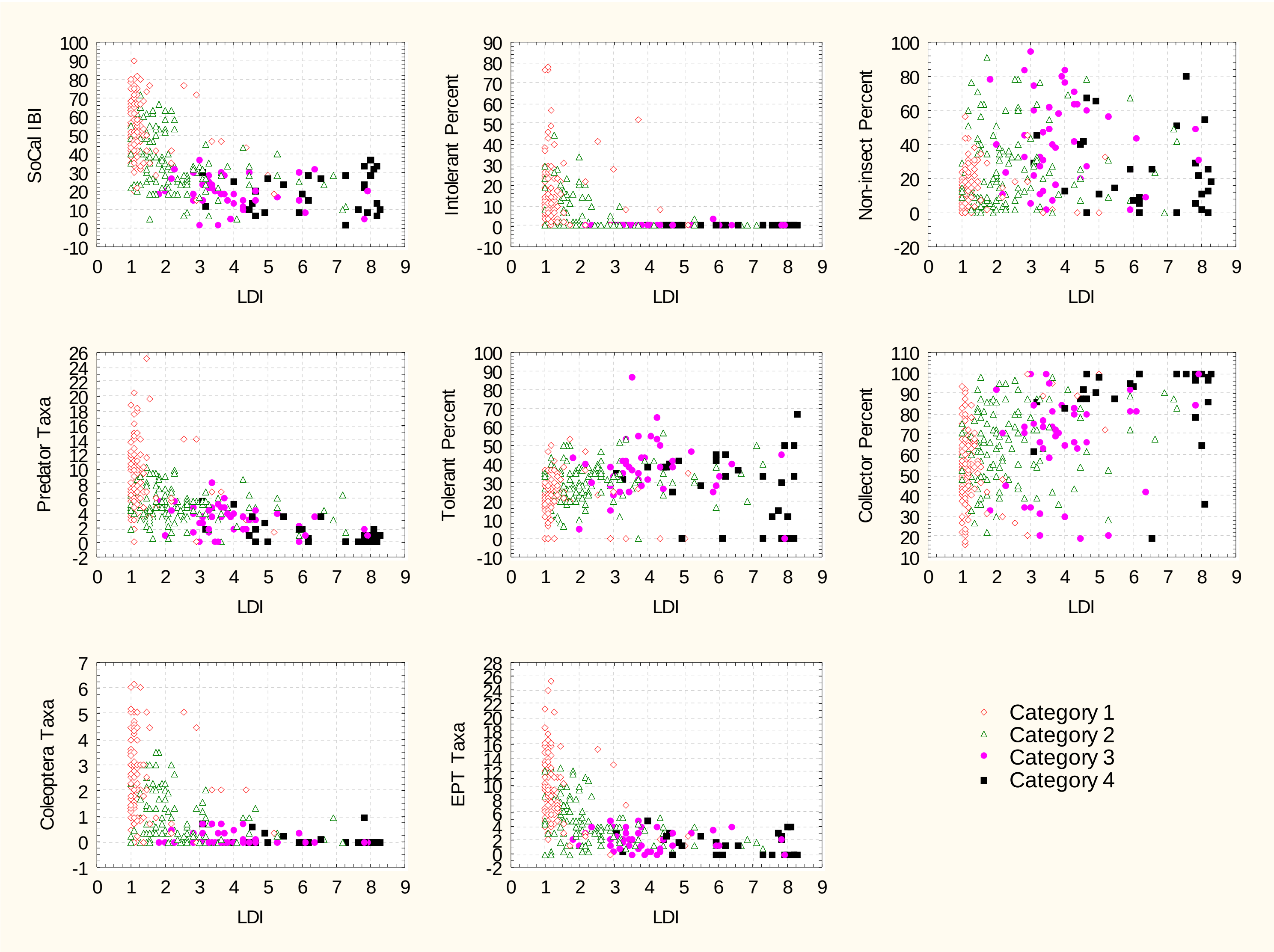


Figure 12. SoCal IBI and associated metrics versus LDI index scores organized among the four site class categories.

Tables 2 and 3 present the revised BCG incorporating the findings observed in the present analyses. In higher elevation streams, some sites appeared to be fairly pristine, as judged by a completely naturally vegetated land cover for many miles around the site. The macroinvertebrate assemblage at these sites showed all the signs of being minimally disturbed (i.e., true reference sites *sensu* Stoddard et al., 2006) and the TAC acknowledged this as well. Therefore, there is the possibility that the natural condition (i.e., BCG level 1) is known and quantifiable for Attributes III and V, and perhaps other attributes, for higher elevation streams in southern California (Table 2). Definitions for Attributes III and V in terms of macroinvertebrate indicators were updated based on the current analyses. Attribute IV (intermediate tolerant taxa) was not updated and it is not clear whether this attribute is relevant to southern California streams. Taxa that are thought to be intermediately tolerant (e.g., *Baetis*, *Hydropsyche*), did not display the expected trend with increasing stress, as measured by either habitat quality or LDI. Other studies have found that taxa of intermediate tolerance are found in roughly similar proportions across BCG tiers 2-5, representing a wide variety of conditions (Gerritsen and Leppo, 2004; Gerritsen and Jessup, 2006). Perhaps other faunal or algal indicators are more discriminating in terms of this attribute.

For lower elevation streams, it is not clear whether truly natural, unimpaired sites still exist in the southern California bight. However, at least a few low elevation sites displayed IBI and metric values approaching the highest scores found anywhere in the region. This may, of course, be a natural outcome of how the IBI was developed. As a placeholder, BCG level 1 (native condition) was defined for Attributes III and V for macroinvertebrates based on a compilation of the best metric scores observed for all low elevation sites combined (total of 175 sites; Table 3). Again, intermediately tolerant taxa (Attribute IV) may not be an informative attribute in terms of macroinvertebrates for this region. Number of Coeloptera (beetle) taxa is thought to be another indicator of sensitive ubiquitous taxa (Figure 7); however, the total number of taxa observed in the dataset (6 taxa) is few, making it difficult to discern fine differences with stressor level. Therefore, this metric was removed from the BCG table pending more information.

Among lower elevation streams, there are currently some differences in biological condition between natural and human-altered streams. However, while available habitat quality data suggests several factors that are different between the two types of streams (e.g., substrate heterogeneity and stability, channel sinuosity and complexity, riparian condition quality), it is unclear what is potentially attainable in the human-altered streams in the region (i.e., a least disturbed condition). When low elevation streams are examined with respect to increasing stress (as measured by either the habitat quality index or the LDI index), we can distinguish two separate classes corresponding to relatively natural channels and those that are altered hydrologically on the basis of certain metrics such as percent collectors. However, there appear to be more similarities than differences in terms of biological expectations between these two classes (Figure 10). Using the BCG framework, the best achievable condition (not necessarily best attainable) for altered low elevation streams in the region corresponds to a BCG level of 4, an LDI index score of approximately 4, and a SoCal IBI score of approximately 37 (Figure 12). The best achievable score for a given site, based on this dataset for the more natural channel low elevation streams appears to correspond to a BCG level of 2, an LDI index score of 2, and a SoCal IBI score of 72. No one site appeared to meet all of the indicator criteria identified under BCG level 1 for low elevation streams.

Table 2. Biological Condition Gradient Matrix: California Bight (High Elevation; >1200 ft)

	Biological Condition Gradient					
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
Ecological Attributes	Native structural, functional and taxonomic integrity is preserved; ecosystem function is preserved within the range of natural variability	**Minimal** changes in structure due to loss of some rare native taxa; shifts in relative abundance of taxa but Sensitive-ubiquitous taxa are a dominant component; ecosystem functions are fully maintained through redundant attributes of the system;	**Some** changes in structure due to loss of sensitive or rare native taxa; shifts in relative abundance of taxa but Sensitive-ubiquitous taxa are common and abundant; ecosystem functions are fully maintained through redundant attributes of the system	**Major** changes in structure due to replacement of some Sensitive-ubiquitous taxa by more tolerant taxa,; overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes	Sensitive taxa are nearly absent; conspicuously unbalanced distribution of major groups from that expected; organism condition shows signs of physiological stress; system function shows reduced complexity and redundancy; increased build-up or export of unused materials	Extreme changes in structure; wholesale changes in taxonomic composition; extreme alterations from normal densities and distributions; organism condition is often poor; ecosystem functions are severely altered
I Historically documented, long-lived or regionally endemic taxa Relies on fish and other vertebrates; May need to break out by basin*	As predicted for natural occurrence except for global extinctions (e.g., unarmored 3-spine stickleback, Pacific Treefrog, California newt, or garter snakes present); steelhead and lampreys in foothills.	As predicted for natural occurrence except for global extinctions; 3-spine stickleback present in lowland;	Some may be absent due to global extinction or local extirpation; 3-spine stickleback rare or extirpated	Some may be absent due to global, regional or local extirpation	Usually absent; stickleback very rare or absent.	Absent

* LA Basin may have historically more endemic fish species than either San Gabriel, Malibu, San Diego drainages. Also need to distinguish upland from lowland sites. Trout more upland; sticklebacks and sculpins lowland. Most long-lived species extinct in region; may be similarity between long-lived or endemics and sensitive-rare species.

Table 2. Biological Condition Gradient Matrix: California Bight (High Elevation; >1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
II Sensitive-rare taxa (currently rare)*	As predicted for natural occurrence, with at most minor changes from natural densities Sculpin (Cottus asper) (Ventura); lamprey adults in upland streams) red-legged frogs present; 3 spine armored stickleback	Virtually all are maintained with some changes in densities	Some loss, with replacement by functionally equivalent Sensitive-ubiquitous taxa	May be markedly diminished	Absent	Absent
III Sensitive-ubiquitous taxa [% intolerant individual EPT]	As predicted for natural occurrence, with at most minor changes from natural densities Partially armored Stickleback common; speckled dace species present in upland streams. Trout present in higher elevation streams. > 40% Intolerant; > 22 EPT taxa; > 20 Predator taxa	Present and abundant; > 16 EPT taxa > 14 predator taxa; > 30% intolerants Diatoms main form of periphyton; Achnanthes oblongella, ventralis; Cymbella amphioxys, gracilis, Amphora inariensis	Common and abundant; ≥10 EPT; ≥ 11 predator; >20% intolerants	Present but some replacement by functionally equivalent taxa of greater tolerance. ≤10 EPT, ≤ 11 predator, < 20% intolerants	Frequently absent or markedly diminished; less sensitive EPT (e.g., Baetidae) may be present but not more sensitive taxa. < 7 EPT; < 6 predator; < 4% intolerants	Absent ≤4 EPT taxa; <2% intolerant; <3 predator taxa

Table 2. Biological Condition Gradient Matrix: California Bight (High Elevation; >1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	
IV Taxa of intermediate tolerance	As predicted for natural occurrence, with at most minor changes from natural densities Native sucker present Western toad Common stickleback	As naturally present with slight increases in abundance	Often evident increases in abundance Diatom species include: Achnanthes biasolettiana, Cymbella sinuata, Denticula tennis, Fragilaria construens, Navicula capitata.	Common and often abundant; relative abundance may be greater than Sensitive-ubiquitous taxa	Often exhibit excessive dominance	May occur in extremely high OR extremely low densities; richness of all taxa is low
V Tolerant taxa [non-insect taxa %tolerant taxa Collectors]	As naturally occur, with at most minor changes from natural densities Arroyo chub present <10% tolerant; <5% Non Insect taxa; >40% Intolerant <30% collectors	As naturally present with slight increases in abundance; <45% collectors; >30% intolerants; <10% non-insects; coleopteran taxa present; <15% tolerant taxa Arroyo chub present	May be increases in abundance of functionally diverse tolerant taxa; <50% collectors; >20% intolerants; <15% non-insects; <25% tolerant Arroyo chub present	May be common but do not exhibit significant dominance; few coleopteran taxa; >15% non-insects; >25% tolerant taxa, >50% collectors; <20% intolerant taxa Diatom indicators include: Nitzchia palea, Navicula atomus, minima, Fragilaria capucina, Cymbella affinis, Stephanodiscus. Attached green algae more prolific – Cladophora, Stigeoclonium, Oedogonium – as well as blue-greens such as Oscillatoria, Ababena Arroyo chub present	Often occur in high densities and are dominant; high percentage of collectors and non-insect taxa; few predator or EPT taxa >60% collectors; >30% tolerant taxa; >20% non-insect taxa; <10% intolerant taxa Arroyo chub less abundant	Comprise ≥ one-third of the assemblage; often extreme departures from normal densities (high or low); no coleoptera, sensitive EPT taxa, and few predator taxa. Mostly collector taxa and often high proportion of non-insect taxa >75% collectors; >40% non-insect taxa; >40% tolerant taxa; <2% intolerant taxa Arroyo chub scarce

Table 2. Biological Condition Gradient Matrix: California Bight (High Elevation; >1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	
VI Non-native or intentionally introduced taxa Include riparian vegetation	Non-native taxa not present	Non-native taxa may be present, but in few numbers and very few species represented	Introduced non-native taxa may be more common in some assemblages (e.g. fish, amphibians, or macrophytes).	Non-native taxa fairly numerous but may not dominate assemblage	Some assemblages (e.g., fish, amphibians, or macrophytes) are dominated by non-native taxa (e.g., brown trout, Cottus asperus in upland)	Often dominant; may be the only representative of some assemblages (e.g., plants, fish, amphibians).
VII Organism Condition (especially of long-lived organisms) More data needed*[*]	Any anomalies are consistent with naturally occurring incidence and characteristics	Any anomalies are consistent with naturally occurring incidence and characteristics	Anomalies are infrequent	Incidence of anomalies may be slightly higher than expected	Biomass may be reduced; anomalies increasingly common	Long-lived taxa may be absent; Biomass reduced; anomalies common and serious; minimal reproduction except for extremely tolerant groups

[*] Percent fish anomalies (DELTS) higher in more stressed systems in the Central Valley (USGS report); should be useful attribute for LA region but unclear whether there are sufficient data available.

Table 2. Biological Condition Gradient Matrix: California Bight (High Elevation; >1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	
VIII Ecosystem Functions	All are maintained within the natural range of variability. Algal as well as plant source of energy.	All are maintained within the natural range of variability	Virtually all are maintained through functionally redundant system attributes; minimal increase in export except at high storm flows	Virtually all are maintained through functionally redundant system attributes though there is evidence of loss of efficiency (e.g., increased export or decreased import)	There is apparent loss of some ecosystem functions manifested as increased export or decreased import of some resources,. Shift to almost entirely algal production: % Collector-filterers dominate the macroinvertebrate assemblage indicative of filamentous algae and DOC as the major energy sources.	Most functions show extensive and persistent disruption
* For southern California streams, may work in opposite direction? Limited connectance naturally, at least in uplands; greater connectance is artificially derived – leads to increase in exotics and decrease in natives.						

Table 3. Biological Condition Gradient Matrix: California Bight (Low Elevation; <1200 ft)

	Biological Condition Gradient					
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
Ecological Attributes	Native structural, functional and taxonomic integrity is preserved; ecosystem function is preserved within the range of natural variability	**Minimal** changes in structure due to loss of some rare native taxa; shifts in relative abundance of taxa but Sensitive-ubiquitous taxa are a dominant component; ecosystem functions are fully maintained through redundant attributes of the system;	**Some** changes in structure due to loss of sensitive or rare native taxa; shifts in relative abundance of taxa but Sensitive-ubiquitous taxa are common and abundant; ecosystem functions are fully maintained through redundant attributes of the system	**Major** changes in structure due to replacement of some Sensitive-ubiquitous taxa by more tolerant taxa,; overall balanced distribution of all expected major groups; ecosystem functions largely maintained through redundant attributes	Sensitive taxa are nearly absent; conspicuously unbalanced distribution of major groups from that expected; organism condition shows signs of physiological stress; system function shows reduced complexity and redundancy; increased build-up or export of unused materials	Extreme changes in structure; wholesale changes in taxonomic composition; extreme alterations from normal densities and distributions; organism condition is often poor; ecosystem functions are severely altered
I Historically documented, long-lived or regionally endemic taxa Relies on fish and other vertebrates; May need to break out by basin*	As predicted for natural occurrence except for global extinctions (e.g., unarmored 3-spine stickleback, Pacific Treefrog, California newt, or garter snakes present); steelhead and goby in coastal reaches, stickleback and sculpin in lowlands	As predicted for natural occurrence except for global extinctions; 3-spine stickleback present in lowland	Some may be absent due to global extinction or local extirpation; 3-spine stickleback rare or extirpated	Some may be absent due to global, regional or local extirpation	Usually absent; stickleback very rare or absent.	Absent

* LA Basin may have historically more endemic fish species than either San Gabriel, Malibu, San Diego drainages. Also need to distinguish upland from lowland sites. Trout more upland; sticklebacks and sculpins lowland. Most long-lived species extinct in region; may be similarity between long-lived or endemics and sensitive-rare species.

Table 3. Biological Condition Gradient Matrix: California Bight (Low Elevation; <1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
II Sensitive-rare taxa (currently rare)*	As predicted for natural occurrence, with at most minor changes from natural densities Sculpin (Cottus asper) (Ventura) lamprey ammocoetes in lowland streams; red-legged frogs present; Speckled dace – lowlands; 3 spine armored stickleback	Virtually all are maintained with some changes in densities	Some loss, with replacement by functionally equivalent Sensitive-ubiquitous taxa	May be markedly diminished	Absent	Absent
III Sensitive-ubiquitous taxa [% intolerant, prediator taxa]	As predicted for natural occurrence, with at most minor changes from natural densities Partially armored Stickleback common; speckled dace species present in upland streams. Trout present in higher elevation streams. > 40% Intolerant; > 12 EPT taxa; > 14 Predator taxa	Present and abundant; > 10 EPT taxa, > 10 predator taxa; > 20% intolerants Diatoms main form of periphyton; Achnanthes oblongella, ventralis; Cymbella amphioxys, gracilis, Amphora inariensis	Common and abundant; ≥ 8 EPT; ≥ 6 predator; > 10% intolerants	Present but some replacement by functionally equivalent taxa of greater tolerance. ≤8 EPT, ≤ 6 predator, < 10% intolerants	Frequently absent or markedly diminished; less sensitive EPT (e.g., Baetidae) may be present but not more sensitive taxa. < 3 EPT; < 4 predator; < 5 % intolerants	Absent ≤1 EPT taxa; ≤ 1% intolerant; ≤ 2 predator taxa

Table 3. Biological Condition Gradient Matrix: California Bight (Low Elevation; <1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
IV Taxa of intermediate tolerance	As predicted for natural occurrence, with at most minor changes from natural densities Native sucker present Western toad Common stickleback	As naturally present with slight increases in abundance	Often evident increases in abundance Diatom species include: Achnanthes biasolettiana, Cymbella sinuata, Denticula tennis, Fragilaria construens, Navicula capitata.	Common and often abundant; relative abundance may be greater than Sensitive-ubiquitous taxa	Often exhibit excessive dominance	May occur in extremely high OR extremely low densities; richness of all taxa is low
V ***Tolerant taxa*** [non-insect taxa %tolerant taxa Collectors]	As naturally occur, with at most minor changes from natural densities Arroyo chub present <15% tolerant; <5% Non Insect taxa; <40% collectors	As naturally present with slight increases in abundance; <50% collectors; >30% intolerants; < 8% non-insects; coleopteran taxa present; <20% tolerant taxa Arroyo chub present	May be increases in abundance of functionally diverse tolerant taxa; <60% collectors; <12% non-insects; <25% tolerant Arroyo chub present	May be common but do not exhibit significant dominance; few coleopteran taxa; >12% non-insects; >20% tolerant taxa, >60% collectors Diatom indicators include: Nitzchia palea, Navicula atomus, minima, Fragilaria capucina, Cymbella affinis, Stephanodiscus. Attached green algae more prolific – Cladophora, Stigeoclonium, Oedogonium – as well as blue-greens such as Oscillatoria, Ababena Arroyo chub present	Often occur in high densities and are dominant; high percentage of collectors and non-insect taxa; few predator or EPT taxa >75% collectors; >33% tolerant taxa; >20% non-insect taxa; Arroyo chub less abundant	Comprise ≥ one-third of the assemblage; often extreme departures from normal densities (high or low); no coleoptera, sensitive EPT taxa, and few predator taxa. Mostly collector taxa and often high proportion of non-insect taxa >90% collectors; >45% non-insect taxa; >40% tolerant taxa; Arroyo chub scarce

Table 3. Biological Condition Gradient Matrix: California Bight (Low Elevation; <1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
VI Non-native or intentionally introduced taxa Include riparian vegetation	Non-native taxa not present	Non-native taxa may be present, but in few numbers and very few species represented	Introduced non-native taxa may be more common in some assemblages (e.g. fish, amphibians, or macrophytes).	Non-native taxa fairly numerous but may not dominate assemblage	Some assemblages (e.g., fish, amphibians, or macrophytes) are dominated by non-native taxa (e.g., bluegill, bass, African clawed frog, carp in lowland streams).	Often dominant; may be the only representative of some assemblages (e.g., plants, fish, amphibians).
VII Organism Condition (especially of long-lived organisms) More data needed*[*]	Any anomalies are consistent with naturally occurring incidence and characteristics	Any anomalies are consistent with naturally occurring incidence and characteristics	Anomalies are infrequent	Incidence of anomalies may be slightly higher than expected	Biomass may be reduced; anomalies increasingly common	Long-lived taxa may be absent; Biomass reduced; anomalies common and serious; minimal reproduction except for extremely tolerant groups

[*] Percent fish anomalies (DELTS) higher in more stressed systems in the Central Valley (USGS report); should be useful attribute for LA region but unclear whether there are sufficient data available.

Table 3. Biological Condition Gradient Matrix: California Bight (Low Elevation; <1200 ft)

Biological Condition Gradient						
	1 **Natural or native condition** Historical reference condition in many cases	2 **Very Good** Minimal changes in the structure of the biotic community and minimal changes in ecosystem function Least disturbed conditions – current reference condition	3 **Good** Evident changes in structure of the biotic community and minimal changes in ecosystem function	4 **Fair** Serious changes in structure of the biotic community and minimal changes in ecosystem function	5 **Poor** Severe changes in structure of the biotic community and moderate changes in ecosystem function	6 **Very Poor** Radical changes in structure of the biotic community and major loss of ecosystem function
VIII Ecosystem Functions	All are maintained within the natural range of variability. Algal as well as plant source of energy.	All are maintained within the natural range of variability	Virtually all are maintained through functionally redundant system attributes; minimal increase in export except at high storm flows	Virtually all are maintained through functionally redundant system attributes though there is evidence of loss of efficiency (e.g., increased export or decreased import)	There is apparent loss of some ecosystem functions manifested as increased export or decreased import of some resources,. Shift to almost entirely algal production: % Collector-filterers dominate the macroinvertebrate assemblage indicative of filamentous algae and DOC as the major energy sources.	Most functions show extensive and persistent disruption
* For southern California streams, may work in opposite direction? Limited connectance naturally, at least in uplands; greater connectance is artificially derived – leads to increase in exotics and decrease in natives.						

Generalized Stressor Gradient (GSG)

The GSG attributes and characteristics developed for this project were based on qualitative information compiled from various regional references, from TAC members, and from knowledge developed as part of the arid west GSG (see Table 4). Southern California streams differ from most other arid west systems in the degree of natural flashiness in undisturbed reaches, the amount of channel braiding that occurs naturally, and the numbers of exotic species that profoundly affect the distribution of endemic biota. Therefore, departures from “natural” or minimally impaired systems (Level 1) are characterized in terms of the degree of departure from the natural hydrograph, the degree of channel and flood plain alteration, and the degree and types of exotic species present. Similar to results from other regions of the country, it is generally thought that Level 1, or completely natural streams, are unlikely to exist in southern California, except perhaps in remote foothill areas. Furthermore, because the hydrology is naturally variable in this region, it may be difficult to quantitatively characterize Level 1 in any case. The TAC suggested several changes to the national GSG framework to make it more relevant to southern California streams. These include:

- Habitat should be divided into two attributes: instream habitat and riparian habitat. The former includes substrate condition, channel morphology, and the presence of barriers or channel alterations such as culverts. Riparian habitat includes riparian vegetation condition (including native or lack of native species) and lateral connectivity with floodplain. Tetra Tech obtained and included metrics from the California Rapid Assessment Method (CRAM) for wetlands that pertain to riparian condition as well as hydrology.
- Water Quality should be divided into two attributes: conventional and naturally-occurring pollutants and anthropogenic toxics. The TAC agreed that tiered uses will not allow for water quality degradation. However, natural water quality characteristics could be a stressor.

Table 4. Stressor Condition Gradient Matrix: California Bight

	Stressor Condition Levels			
Attribute	**1**	**3**	**4-5**	**6**
Flow	Natural hydrograph; includes periodic seasonal floods and very low flows (dry conditions in some cases); dry season flow from natural sources; rising water has unrestricted access to floodplain; Most of channel characterized by equilibrium conditions.	Moderately changed hydrograph; more consistent flows seasonally through treated wastewater inputs and other sources; some irrigation withdrawals or groundwater removal for other human purposes; noticeable change in flashiness; lateral excursion of rising waters partially restricted by unnatural features; Some aggradation or degradation present but not severe.	Significantly changed hydrograph; both managed and natural flow factors present; stormwater runoff dramatically increases flows temporarily; lateral excursion of rising waters partially restricted by unnatural features; Most of channel actively degrading or aggrading.	Severely changed hydrograph; flow human-controlled; peaking flows, “rafting flows”, or water diversions common; stream is all treated wastewater effluent flow;; diversions such that stream is dry periodically; stream flow result of dam releases; rising waters completely contained within artificial banks; Channel has completely artificial hydrogeology and equilibrium.
Instream Habitat	Natural substrate and channel sinuosity; Braided channels common in lowlands; natural cover available for fish and other aquatic life.	Substrate somewhat modified (often tending to be smaller in size); channel morphology may be slightly modified.	Natural bottom but concrete sides or altered bottom. Substrate size typically fine. Culverts or instream structures present – clear effects on channel morphology	Severely changed channel morphology; channelized; concrete sides and bottom; substrate radically altered.
Riparian Habitat	lateral connection between stream and riparian corridor; native riparian vegetation predominates; underwater willow roots or other riparian plants serve as habitat for aquatic life; 75-100% of stream has riparian buffer; average buffer width ≥ 100m; intact soils.	some exotic-invasive riparian vegetation; connection with flood plain/riparian corridor mostly intact; 50-75% of stream has riparian buffer; average buffer width 60-99m intact or moderately disrupted soils.	25-50% of stream has riparian buffer; average buffer width 30-60m; moderate-extensive soil disruption.	exotic vegetation only if any at all; no connection to flood plain; < 25% of stream has riparian buffer; average buffer width < 30m; barren ground or highly compacted soils.
Conventional Water Quality parameters and naturally occurring chemicals	DO generally near saturation in upland streams – generally > 5 mg/L in lowland streams; temperature cool in upland streams – generally < 30 °C in lowland streams in the summer.	DO and temperature may be slightly altered but still satisfactory for native aquatic life.	Altered DO and/or temperature regimes; elevated concentrations of metals or other constituents naturally	DO and/or temperature radically altered – temperature often > 30° C in summer; conductivity, salinity, or dissolved solids generally much higher than typical for supporting aquatic life; metals or other chemicals naturally high and known to be toxic to aquatic life

Table 4. Continued

	Stressor Condition Levels			
Attribute	**1**	**3**	**4-5**	**6**
Anthropogenic Toxics	No anthropogenic toxics	Infrequent pollutant exceedences of standards; generally non-toxic conditions	Occasional exceedences of WQ objective(s); Stormwater runoff may decrease water quality in certain segments or over short time periods.	Toxics exceed water quality objectives; multiple toxic chemicals co-occur or multiple exceedences of a WQ objective
Watershed Condition	All natural land cover; natural longitudinal connectivity and connectivity with ground water; Contiguous natural riparian buffer between segments.	Mostly natural land cover – some human developed areas; longitudinal connectivity mostly in tact – some fragmentation of habitat or barriers	Mostly human land uses, Urban intensity moderate (30-50 out of 100); longitudinal connectivity fragmented, interrupted; agricultural uses may be relatively predominant	Nearly all human land uses; urban intensity > 50/100; connectivity severely altered; agricultural land uses dominant
Invasive Species	Exotics or introduced species absent. Riparian vegetation as naturally occurs.	A few non-invasive exotics may be present (e.g., crayfish, fathead minnow), including riparian plant species; but generally has little effect on native species or riparian habitat.	Some non-invasive exotics combined with one or two aggressive exotic species (e.g., brown trout; Tamerisk; Arrando).	Invasive, predatory, or aggressive exotic species common (e.g., bass, bluegill, African clawed frog, bull frog). Clear evidence of extirpation of native species due to exotic species. Highly altered riparian habitat due to invasive species present.

- Energy source attribute has questionable relevance to southern California streams. The TAC suggested deleting this attribute pending further discussions.
- Watershed condition attribute was added. This includes land uses and longitudinal and vertical connectivity issues. The urban intensity index, which Tetra Tech calculated for several sites in the Region is one descriptor that is useful here. The CRAM connectivity metric is also relevant here.
- Invasive species attribute was added. This includes riparian plants as well as fauna.

Urbanization, Hydrology and SoCal IBI

There also appears to be some separation in the GSG based on flow regimes and hydrology; streams with more constant flows year-round (e.g., effluent dominated streams) appear to have a higher likelihood of harboring exotic species. Highly urbanized areas are often subject to much greater wet weather runoff than normal resulting in much higher peak flows and a very altered hydrograph.

Plotting the SoCal IBI against the LDI, there are sites that appear to be better than most within its class of urban intensity (see labels in Figure 13). One possibility is that while potential urban sources are present (e.g., residential housing is relatively dense, many roads), the actual level of stressors is less because of the way road runoff and other human-derived stressors are routed. Another possibility is that the stream has certain features that help protect it from urban-related stressors (e.g., riparian vegetation). A third possibility is that sites with lower IBI scores for a given LDI are affected by non-urban stressors as well (e.g., agriculture derived stressors) and are therefore, subject to more stressors than those sites with better IBI scores. Future efforts should plan to compile what is known about these sites so that we can identify factors that mitigate urban effects and better define the GSG.

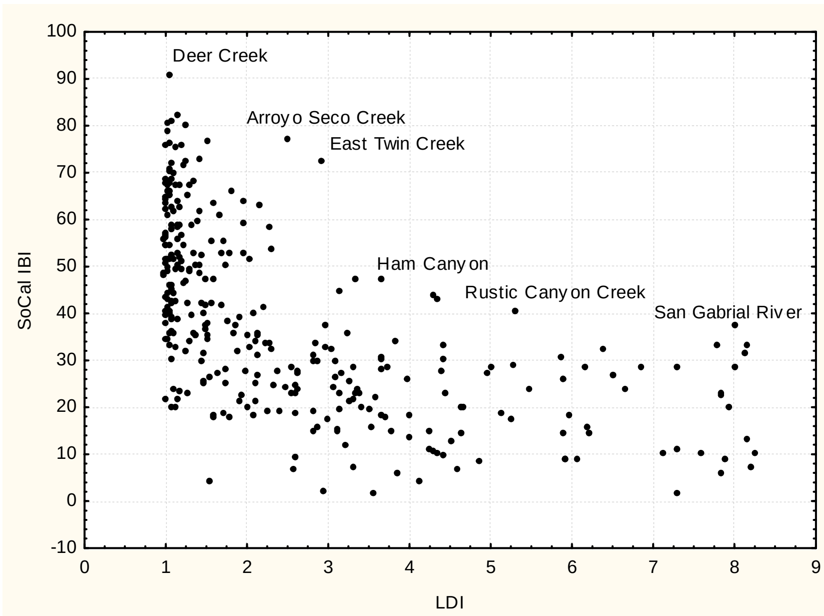


Figure 13. Plot of the southern California macroinvertebrate (SoCal) IBI in relation to the LDI. Higher IBI scores indicate better biological condition. Higher LDI values indicate greater landscape disturbance and probable urban stressors.

The results presented here can be described as a conceptual development of a Southern California BCG based on the existing SoCal IBI and its associated biological metrics. Although the conceptual BCG presented here is a promising step, a fully calibrated BCG is necessary in order for the biological and stressor data to be used in tiered aquatic life uses, as well as for use attainability analyses.

It is recommended that a workshop be organized to initiate development of a calibrated BCG. Individuals involved in the workshop should have extensive knowledge on the type of biological assemblage being investigated and should understand its responses in pristine to severely stressed conditions. Generally, these workshops last two to three days, depending on participants' familiarity with TALU and BCG concepts. The strong relationships of these biological measures with stress (as described by habitat quality and the LDI index), as well as the variation in biology among the two natural and two altered site classes, suggest that generating a calibrated BCG would be possible using the currently available data. To do this, macroinvertebrate data (and to the extent feasible, other types of biological data) need to be explored in more detail to identify specific taxa that are common, as well as sensitive, to the stressors found in the Southern California Bight region. Additionally, the knowledge of local experts must be used in order to reduce uncertainty associated with ambiguous or incomplete data. It may also be necessary to assemble a more comprehensive GSG based on a larger assemblage of data types (i.e., stressors).

Literature

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. Second edition. U.S. Environmental Protection Agency; Office of Water; Washington, D.C. EPA 481-B-99-002.

Brown, M. T., N. Parker, and A. Foley. 1998. Spatial Modeling of Landscape Development Intensity and Water Quality in the St. Marks River Watershed. DEP Contract #GW138. Center for Wetlands, University of Florida, Gainesville.

Brown, M.T. and M.B. Vivas. 2005. Landscape development intensity index. Environmental monitoring and assessment 101:289-309.

Davies, S.P. and S.K. Jackson. 2006. The biological condition gradients: A descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16(4):1251-1266.

Fore, L.S. 2004. Development and Testing of Biomonitoring Tools for Macroinvertebrates in Florida Streams. Prepared for: Russel Frydenborg and Ellen McCarron, Florida Department of Environmental Protection, 2600 Blair Stone Rd., Tallahassee, FL 32399-2400.

Gerritsen, J. and B. Jessup. 2006. Draft: Identification of the Biological Condition Gradient for Non-Calcareous Stream of Pennsylvania. Prepared for: Susan Jackson, U.S. EPA Office of Science and Technology; Maggie Passmore, U.S. EPA, Region 3; Rod Kime, Pennsylvania Department of Environmental Protection by Tetra Tech, Inc., 400 Red Brook Blvd, Ste 200, Owings Mills, MD 21117.

Gerritsen, J. and E.W. Leppo. 2004. Draft: Tiered Aquatic Life Use Development for New Jersey. Prepared for: William Swietlik, USEPA OST; James Kurtenbach, U.S. EPA, Region 2; Kevin Berry, New Jersey Department of Environmental Protection by Tetra Tech, Inc., 400 Red Brook Blvd, Ste 200, Owings Mills, MD 21117.

Ode, P.R., A.C. Rehn, and J.T. May. 2005. A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams. Environmental Management 35(4): 493-504.

Richter, B.D., J.V. Baumgartner, J. Powell, and D.P. Braun. 2000. A method for assessing hydrologic alteration within ecosystems. Conservation Biology 10(4):1163-1174.

Stoddard, J., D. Larsen, C. Hawkins, R. Johnson, and R. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications, 16: 1267–1276.

Appendix 9

Bioassessment tools in novel habitats: An evaluation of indices and sampling methods in low-gradient streams in California

Raphael D. Mazor, Kenneth C. Schiff, Kerry J. Ritter, Andy Rehn[1] and Peter Ode[1]

Abstract

Biomonitoring programs are often required to assess streams for which assessment tools have not been developed. For example, low-gradient streams (slope ≤1%) comprise 20 to 30% of all stream miles in California and are of particular interest to watershed managers, yet most sampling methods and bioassessment indices in the State were developed in high-gradient systems. This study evaluated the performance of three sampling methods: targeted riffle composite (TRC), reachwide benthos (RWB), and the margin-center-margin modification of RWB (MCM); and two indices: the Southern California Index of Biotic Integrity (SCIBI) and the ratio of observed to expected taxa (O/E) in low-gradient streams in California for application in this habitat type. Performance was evaluated in terms of efficacy (i.e., ability to collect enough individuals for index calculation), comparability (i.e., similarity of assemblages and index scores), sensitivity (i.e., responsiveness to disturbance), and precision (i.e., ability to detect small differences in index scores). The sampling methods varied in the degree to which they targeted macroinvertebrate-rich microhabitats, such as riffles and vegetated margins, which may be naturally scarce in low-gradient streams. The RWB method failed to collect sufficient individuals (i.e., ≥450) to calculate the SCIBI in 28 of 45 samples, and often collected fewer than 100 individuals, suggesting it is inappropriate for low-gradient streams in California. Failures for the other methods were less common (TRC:16 samples; MCM:11 samples). Within-site precision, measured as the minimum detectable difference (MDD), was poor but similar across methods for the SCIBI (ranging from 19 to 22). RWB had the lowest MDD for O/E scores (0.20 vs. 0.24 and 0.28 for MCM and TRC, respectively). Mantel correlations showed that assemblages were more similar within sites among methods than within methods among sites, suggesting that the sampling methods were collecting similar assemblages of organisms. Statistically significant disagreements among methods were not detected, although O/E scores were higher for RWB samples than TRC. Index scores suggested impairment at all sites in the study. Although index scores did not respond strongly to several measurements of disturbance in the watershed, % agriculture showed a significant, negative relationship with O/E scores.

Introduction

Large-scale biomonitoring programs are often confronted with the need to assess habitat types for which assessment tools have not been developed. This problem is severe in large heterogeneous regions like California (Carter and Resh 2005). Developing and maintaining unique assessment tools for multiple habitat types may be prohibitively expensive and may impede comparison of data from different regions. Therefore, assessing the applicability of tools in diverse habitat types is a critical need for large biomonitoring programs.

In southern California, biomonitoring programs use tools like the SCIBI (Ode *et al*. 2005), which were developed using reference sites that were predominantly in high-gradient (i.e., >1% slope) streams. However, low-gradient streams are a major feature in alluvial plains of this region (Carter and

[1] *California Department of Fish and Game, Aquatic Bioassessment Laboratory, Water Pollution Control Laboratory, Rancho Cordova, CA*

Resh 2005). According to the National Hydrography Dataset Plus (NHD+; USEPA and USGS 2005) approximately 20 to 30% of all stream miles in California have slopes below 1%. Because these habitats are subject to numerous impacts and alterations (SMCBWG 2007), several biomonitoring efforts in California specifically target low-gradient streams, even though the applicability of assessment tools created and validated in high-gradient streams has not been tested.

Low-gradient streams differ from high-gradient streams in many respects (Montgomery and Buffington 1997). For example, bed substrate is typically composed of fines and sands, rather than cobbles, boulders, or bedrock. In California and other semiarid climates, low-gradient channels are often complex, with ambiguous and dynamic bank structure. Frequent floods create new channels and cause streams to abandon old ones (Carter and Resh 2005). For bioassessment programs, an important distinction between high- and low-gradient streams is the scarcity of riffles and other microhabitats that are typically targeted by macroinvertebrate sampling protocols (e.g., Harrington 1999).

In this study, application of three sampling methods and two bioassessment indices for use in low-gradient streams in California were evaluated. Sampling methods were assessed for efficacy (i.e., the ability to collect sufficient numbers of benthic macroinvertebrates), comparability (i.e., community similarity and agreement among assessment indices), sensitivity (i.e., responsiveness of the indices to watershed disturbance), and precision of the assessment indices (i.e., power of assessments to detect differences among sites).

Methods

Study Areas

Twenty-one low-gradient sites were sampled in several regions across California (Table 1; Figure 1). Most sites were in heavily altered rivers, although a few were in protected watersheds. Slopes were estimated from the NHD+ (USEPA and USGS 2005), or from digital elevation models (at Jack Slough, Wadsworth Canal, and the Santa Ana River, which lacked associated data in the NHD+). All sites were on reaches defined in the NHD+ as having slopes below 1%.

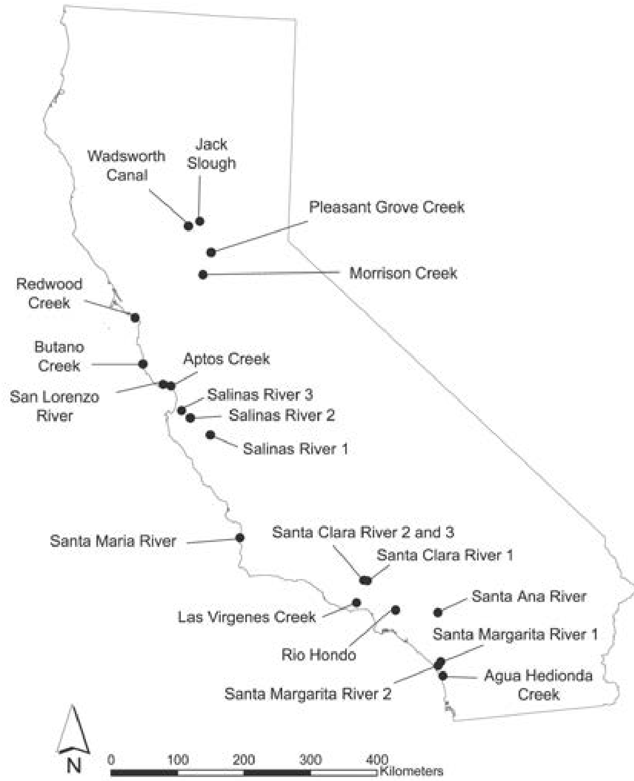


Figure 1. Location of study sites.

Sampling

At each site, TRC, RWB, and MCM sampling methods were used to collect benthic macroinvertebrates. The three sampling methods differ in the degree to which they target the richest microhabitats (e.g., riffles or vegetated margins). TRC and RWB are similar to methods used in the nationwide Environmental Monitoring and Assessment Program (EMAP; Peck *et al*. 2006), and both methods are currently used in California's bioassessment programs (Ode 2007). MCM is intended to capture marginal habitats not sampled by RWB, and has been adopted for use in low-gradient streams in California (Ode and van Buuren 2008). Samples were displaced upstream or downstream by 1 m when necessary to avoid interference among different methods. At 12 sites, triplicate samples were collected for each method (Table 1).

For the TRC method, 11 equidistant transects were established along the 150-m reach, and 3 1-ft^2 areas of streambed were sampled at three randomly selected transects. At each transect, field crews targeted the richest microhabitats and sampled a total of 9 ft^2 of streambed in three riffles. This method is

Table 1. Low-gradient sites included in the study. S = assessed using Southern California Index of Biotic Integrity; X = not assessed using an index of biotic integrity; WS = watershed; Local = within 500 m of sampling point; Ndel = ambiguous watersheds which could not be delineated; Ndet = ambiguous stream network for which stream order could not be determined; and * = triplicate samples collected.

Site	Watershed	County	Watershed Size (km^2)	Stream Order	% Developed		% Agricultural		% Open space	
					WS	Local	WS	Local	WS	Local
Within Central and Southern California										
Central Coast										
S	Aptos Creek*	Santa Cruz	200	3	18	92	0	0	82	8
S	Salinas River 1	Monterey	10940	6	14	71	0	1	86	28
S	Salinas River 2*	Monterey	10666	7	5	28	7	61	88	11
S	Salinas River 3	Monterey	9141	7	5	13	4	27	90	60
S	San Lorenzo River	Santa Cruz	378	4	5	7	6	56	88	37
S	Santa Maria River*	Santa Barbara	1844	6	4	4	6	0	91	96
South Coast										
S	Agua Hedionda Creek*	San Diego	80	3	76	77	0	0	24	23
S	Las Virgenes Creek*	Los Angeles	63	3	19	29	0	0	81	71
S	Rio Hondo*	Los Angeles	325	3	70	83	0	0	30	17
S	Santa Ana River	Riverside	1965	6	25	78	1	0	74	22
S	Santa Clara River 1	Los Angeles	817	4	14	68	0	0	86	32
S	Santa Clara River 2	Los Angeles	1107	5	16	76	0	1	84	23
S	Santa Clara River 3	Los Angeles	1107	5	16	75	0	5	84	20
S	Santa Margarita River 1*	San Diego	1856	6	13	48	3	0	84	52
S	Santa Margarita River 2*	San Diego	1888	6	14	24	3	0	83	76
Outside Central and Southern California										
Bay Area										
X	Butano Creek	San Mateo	234	3	11	34	0	0	89	66
X	Redwood Creek*	Marin	44	2	4	10	2	24	94	67
Central Valley										
X	Jack Slough	Yuba	Ndel	3		7		91		2
X	Morrison Creek*	Sacramento	114	3	40	100	4	0	56	0
X	Pleasant Grove Creek	Placer	40	3	69	34	3	16	28	50
X	Wadsworth Canal	Sutter	Ndel	Ndet		12		87		1

similar to the targeted riffle composite method used by EMAP, which sampled a total of 8 ft^2 of streambed from four to eight riffles (Peck *et al*. 2006). A second difference was the fixed reach length of 150 m, in contrast to EMAP, which had a variable reach length set at 40 times the wetted width.

In contrast to TRC, which allowed the field crew to sample the richest microhabitats within transects, the RWB method used systematically distributed sampling locations. For RWB, eleven equidistant transects were established along the 150-m reach, and one sample was collected with a D-frame kick-n*et al*ong each transect at 25, 50, or 75% of the stream width (with the position changing at each transect). A total of 11 ft^2 of streambed was sampled. This method is similar to the Reach-Wide Benthos method used by EMAP, except that EMAP used variable reach length set to 40 times the wetted width (Peck *et al*. 2006).

The MCM method was identical to RWB with minor modification. Instead of collecting samples at 25, 50 and 75% of stream width, samples were collected at 0, 50, and 100%. Unlike RWB, MCM samples were collected from the margins, which in low-gradient streams often contain the richest, most stable microhabitats (e.g., vegetated margins). As with RWB, 11 ft^2 of streambed were sampled.

Benthic macroinvertebrates were sorted and identified to the Standard Taxonomic Effort Level 1 (i.e., most taxa to genus, with Chironomidae left at family) established by the Southwestern Association of Freshwater Invertebrate Taxonomists (Richards and Rogers 2006). When possible, at least 500 individuals were identified in each sample.

Data Analysis

For each sample, bioassessment metrics and indices were calculated and analyzed to evaluate the efficacy, comparability, sensitivity, and precision of the three sampling methods.

Calculation of indices and metrics

The SCIBI was calculated for 15 sites located on coastal drainages from Santa Cruz to San Diego Counties. No IBIs were calculated for the two sites in the San Francisco Bay Area and the four sites in the Central Valley because IBIs for these regions were not available at the time of the study. Furthermore, small sample sizes in these regions and unknown comparability of IBIs for different regions would limit the utility of including these sites. In order to calculate the SCIBI, benthic macroinvertebrate data were processed according to the index requirements. For example, samples containing more than 500 individuals were randomly subsampled with replacement to obtain 500 individuals per sample.

Calculation of O/E scores

Observed-over-expected scores were calculated for all sites using a predictive model developed for the state of California (Charles P. Hawkins pers. com.; Western Center for Monitoring and Assessment. Accessed online March 30, 2007: http://129.123.10.240/wmcportal/DesktopDefault.aspx). These scores are the ratio of observed to expected taxa, and are based on only those taxa with a probability of occurrence ≥50%. The original identifications were converted to operational taxonomic unit (OTU) names used in the models, and ambiguous taxa (i.e., those that could not be assigned to an OTU and those that could not be adequately identified, such as early instars), as well as all Chironomidae larvae, were eliminated. The resulting sample counts were reduced to 300, if more than 300 individuals remained after removal of ambiguous taxa. Sites were assigned to the appropriate submodel based on climate (i.e., low mean annual precipitation, and high mean monthly temperature), which were used to predict expected taxa occurrence (E) using longitude, percent sedimentary geology in the watershed, and log mean annual precipitation. Climatic data were obtained from the Oregon Climate Center (accessed online March 30, 2007: http://www.ocs.orst.edu/prism), and geologic data were obtained from a generalized geological map of the United States (accessed online March 30, 2007: http://pubs.usgs.gov/atlas/geologic). Details of these predictive models can be found in Ode *et al*. 2008.

The two Central Valley sites were located in streams with ambiguous watersheds, and therefore required that percent sedimentary geology be estimated, rather than calculated by geographic information systems (GIS). For this study, percent sedimentary geology was estimated at 100%. Using different percent sedimentary geology values (i.e., 0, 20, 40, 60, and 80%) had negligible effect on O/E scores; coefficient of variation for scores within each sample at the two Central Valley sites was <2%, (data not shown), perhaps as a result of the low numbers of observed taxa at these sites.

Evaluation of sampling methods and indices

Efficacy

To assess the efficacy of the sampling methods, the percentage of samples was calculated for each method that collected at least 450 individuals (within 10% of the minimum number for calculating the SCIBI) or at least 270 individuals (within 10% of the minimum number for calculating O/E, counting only unambiguous taxa). In bioassessment applications, smaller samples would be rejected and represent wasted resources. In order to minimize the effects of pseudoreplication, the percentage of samples containing an adequate number of individuals was calculated for each site; then, this percentage was averaged across all 21 sites. This rate estimated the likelihood of collecting adequate samples from the population of sites in the study. McNemar's test was used to test differences between methods (paired within sites) for statistical significance (Zar 1999, Stokes *et al*. 2000). Because McNemar's test requires binary data, within-site rates were rounded to 1 or 0 at replicated sites. A Bonferroni correction was used to account for multiple tests across methods (i.e., $\alpha = 0.05/3 = 0.017$).

Comparability

To see if the different sampling methods collected similar types of organisms, community structure between sampling methods was compared using a Mantel test (Mantel 1967). Mantel tests provide a measure of correlation (Mantel's R) between two sampling methods. Sorensen distance was used as a dissimilarity measure. For sites where multiple samples were collected, mean distances were used; that is, matrices comprised mean or observed distances between pairs of sites, not samples. All samples were included in this analysis, regardless of the number of individuals collected. Significance was tested against correlation values for 999 runs with randomized data.

A Bonferroni correction was used to account for multiple tests across methods (i.e., $\alpha = 0.05/3 = 0.017$). PC-ORD [Version 5.12] was used to run Mantel tests (MJM Software Design, Glendeden Beach, OR).

To determine the relative influence of sampling method on assessment indices, a variance components analysis was used to determine how much of the variability was explained by differences among sites, sampling methods, and their interaction. Restricted maximum likelihood (REML) was used to calculate variance components because of the unbalanced design. SAS was used for all calculations (using PROC VARCOMP method=REML, SAS Institute Inc. 2004). Unlike the mean square method of estimating variance components, REML ensures that all components are greater than or equal to zero (Larsen *et al*. 2001). Because sites were a fixed factor and not a random factor, the variance component attributable to site must be considered a finite, or pseudo variance (Courbois and Urquhart 2004). Only sites where all three sampling methods were represented (after excluding samples containing inadequate numbers of organisms) were used in this analysis.

To assess agreement among the sampling methods, mean SCIBI and O/E values were calculated and regressed for each pair of methods. Slopes were tested against 1 and intercepts to 0 ($\alpha = 0.05$); Theil's test for consistency and agreement, which is based on differences between sampling methods, was used as an additional test of comparability (Theil 1958). Pairwise differences between mean SCIBI and O/E scores were regressed against log watershed area and stream order to see if these gradients contributed to the observed disagreements. A Bonferroni correction was not used for either analysis in order to increase the ability to detect disagreements. Bias was not explicitly assessed because none of the methods could be assumed to represent a true value. Only samples with adequate numbers of individuals were used in this analysis.

Sensitivity

The sensitivity of the assessment indices to watershed alteration was assessed by correlating mean SCIBI and O/E scores against land cover metrics, including percent open, developed, and agricultural land within the watershed for all sites with unambiguous watersheds (Table 1). This analysis assumed that the biology of the streams respond to these watershed alterations. Open water was excluded from all calculations. Land cover data was obtained from the National Land Cover Database (USGS 2003). Relationships were assessed by calculating the Spearman rank correlation, which is robust to non-normal distributions and extreme values in land cover metrics (Zar 1999). Only samples with the minimum number of individuals for each index were used in this analysis. Data from each sampling method were analyzed independently. A Bonferonni correction was used to account for multiple comparisons ($\alpha = 0.05/6 = 0.008$) across two indices and three land cover classes within each method.

Precision

Precision was evaluated by calculating the MDD of each sampling method for SCIBI and O/E scores (Zar 1999, Fore *et al*. 2001). The MDD was calculated using the mean squared error from a nested ANOVA (replicates within site) as an estimate for average within-site variance. Only data from site and method combinations with replication (after exclusion of samples lacking adequate numbers of individuals) were used to estimate variability. These estimated variabilities were applied to a two-sample *t*-test ($\alpha = 0.05$, $\beta = 0.10$) with three replicates in each sample. Additionally, the coefficient of variation (CV) of the indices for each method, averaged across sites, was calculated.

Results

One hundred thirty-five samples were collected at 21 sites throughout the state; 15 of these sites were located along the southern and central California coast. All three methods were used at each site, and 196 taxa were identified. For all sampling methods, SCIBI and O/E scores were low at most sites (Figure 2). For example, mean SCIBI scores were well under 39 (the impairment threshold) at all but one site (Aptos Creek). Observed-over-expected scores indicated impairment in nearly every sample, as scores were below the impairment threshold of 0.66 in all but three samples.

Efficacy

Efficacy was low for all methods, and many samples contained fewer than the required number of individuals. Ideally, each sample should have contained at least 500 individuals. However, only 46 of 135 samples met this target; 34 of the remaining 89 samples had at least 450 individuals, the minimum required for calculation of the SCIBI. For the 55 samples with fewer than 450 individuals, IBIs may

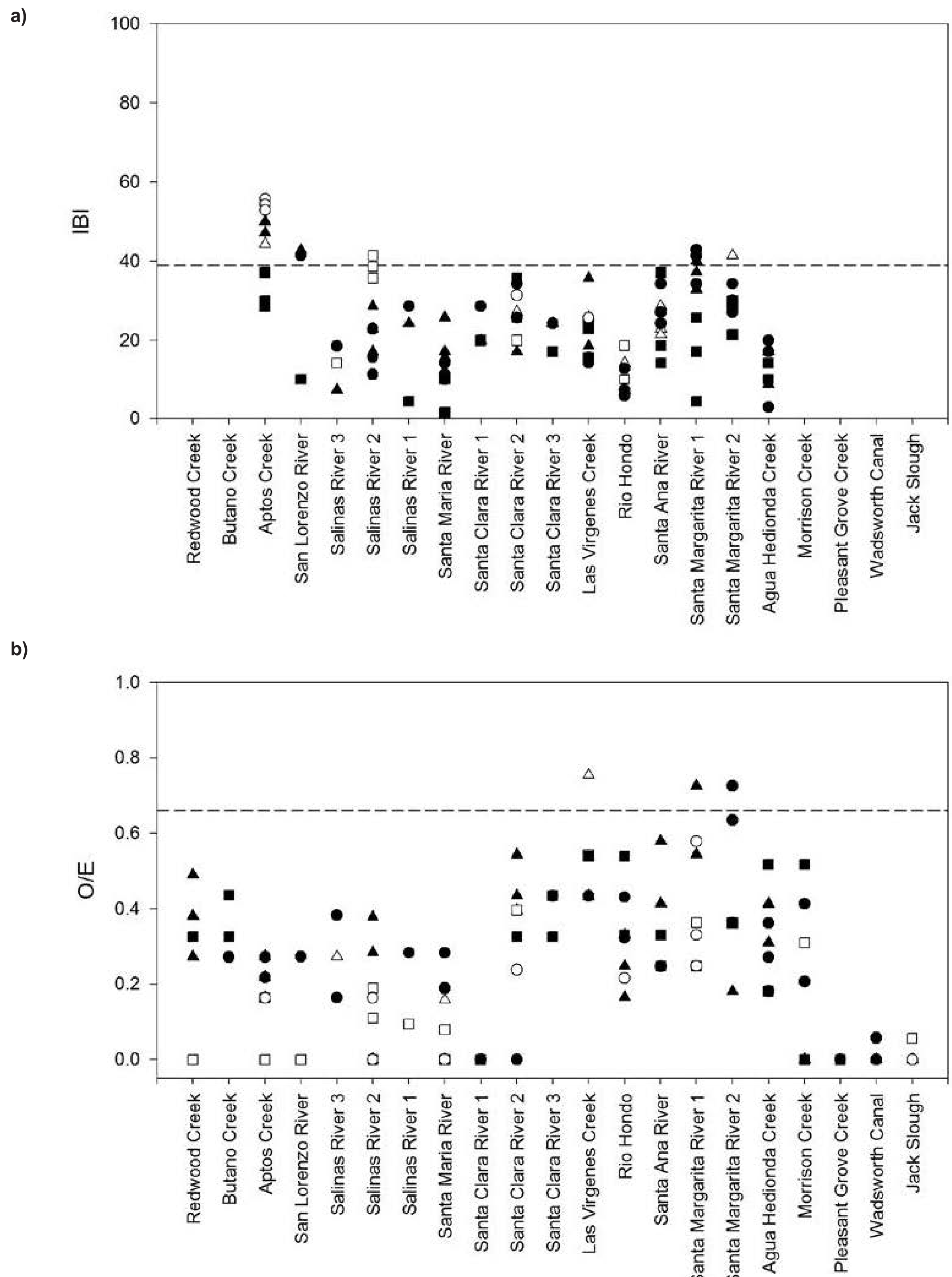


Figure 2. Southern California Index of Biotic Integrity (SCIBI; a) and Observed/Expected (O/E; b) scores by site and method. Each point represents an individual sample. Triangles represent MCM samples. Squares represent RWB samples. Circles represent TRC samples. Black symbols are samples containing sufficient individuals for index calculation, and white symbols are samples containing insufficient individuals for index calculation. Dashed lines represent the threshold for identifying impairment with each index (i.e., 39 for the SCIBI, and 0.66 for the O/E).

B299

not be valid. Furthermore, 55 samples had fewer than 270 unambiguously identified individuals, meaning that O/E scores may not be valid for these samples.

Several samples had extremely low counts (e.g., four individuals; Table 2). Most of these samples were collected by the RWB sampling method. Nearly half (21 out of 45) of RWB samples had fewer than 450 individuals. In contrast, only 2 MCM samples and 6 TRC samples had fewer than 450 individuals. The adjusted efficacy rate, a site-adjusted estimate of sampling efficacy, for the MCM method (54%) was twice that of RWB (27%). The adjusted efficacy rate for TRC (46%) was nearly as high as that of the MCM method. However, these differences fell short of statistical significance after Bonferroni corrections were applied (i.e., $p > 0.017$). The rates were slightly higher for samples with at least 270 individuals at 67, 32, and 67% for MCM, RWB, and TRC, respectively, and these differences were statistically significant (McNemar's test $p = 0.0039$).

Comparability

Sampling methods comparability was good in terms of both multivariate community structure and index scores. Mantel's test showed significant correlations among benthic macroinvertebrate communities collected by all three sampling methods (Table 3). However, the RWB method had weaker correlations with both TRC (0.40) and MCM (0.45), compared to the higher correlation observed between TRC and MCM (0.69). In all cases, the correlations were significant ($p < 0.002$).

Variance components analysis showed that the methods were highly comparable and that site accounted for nearly all of the explained variance in both indices. The analysis of SCIBI scores included 7 sites and 26 samples; the analysis of O/E scores included 10 sites and 52 samples. Site accounted for

Table 3. Mantel correlations between sampling methods. Asterisk denotes statistical significance ($p < 0.017$).

Method 1	Method 2	Mantel's R	P
RWB	MCM	0.45	0.001*
RWB	TRC	0.40	0.002*
MCM	TRC	0.69	0.001*

100% of the explained variance in SCIBI scores and 95% in O/E scores. Method and interaction between site and method explained none or negligible components of the variance in these indices (0 to 5%).

Significant disagreements between pairs of sampling methods were not observed for either index (Table 4; Figure 3). Slopes for all three comparisons were not significantly different from 1, and no intercepts were significantly different from 0. Consistency among SCIBI scores was best (i.e., slope closest to 1) between the MCM and TRC methods (slope = 0.96) and worst for the MCM and RWB methods (slope = 0.62). In contrast, consistency among O/E scores was best between the MCM and RWB methods (slope = 0.97) and worst for the RWB and TRC methods (slope = 0.72). Theil's test confirmed the lack of significant disagreements among IBI and O/E scores between pairs of methods. No differences between sampling methods were significantly related to log watershed area or stream order (regression slope and intercept $p > 0.05$).

Sensitivity

Sensitivity of both indices to gradients in land cover was poor, although to some extent the relationships were affected by sampling method, specific cover type, and geographic scale (Table 5; Figure 4). For example, O/E scores were strongly and negatively correlated with agricultural land cover in the

Table 2. Samples, sites, and efficacy by method. Adjusted Rate = site-adjusted estimate of efficacy rate.

Method	Total		≥ 450 Organisms			≥ 270 Organisms		
	Samples	Sites	Samples		Adjusted Rate	Samples		Adjusted Rate
MCM	45	21	34	76%	54%	32	71%	67%
RWB	45	21	17	38%	27%	14	31%	32%
TRC	45	21	29	64%	46%	30	67%	67%

Table 4. Regressions of mean IBI and O/E scores for each method. Slopes were tested against 1 and intercepts were tested against 0. Methods 1 and 2 plotted on x and y axis, respectively, in Figure 3. SE = Standard error.

Index	Method 1 (x)	Method 2 (y)	n	r^2	Slope	SE	p	Intercept	SE	p
SCIBI	MCM	TRC	14	0.77	0.96	0.15	0.803	2.52	3.96	0.537
	MCM	RWB	7	0.45	0.62	0.25	0.194	6.31	5.53	0.305
	MH	TRC	7	0.74	1.18	0.28	0.540	-0.30	5.63	0.959
O/E	MCM	TRC	14	0.78	0.86	0.13	0.284	0.02	0.04	0.633
	MCM	RWB	8	0.90	0.97	0.13	0.816	0.02	0.04	0.653
	RWB	TRC	8	0.71	0.72	0.19	0.185	0.06	0.06	0.401

watershed (Spearman's ρ ranged from -0.46 to -0.89 across sampling methods). However, most relationships between index scores and land cover metrics were not statistically significant (i.e., p <0.008). Only the relationship between O/E scores from RWB samples were significantly correlated with agricultural land use in the watershed (ρ = -0.89, p = 0.003). Although the direction of correlation often met expectations (e.g., % open space in the watershed vs. SCIBI; Figure 4c), a few showed no clear relationship (e.g., % developed land in the watershed vs. O/E; Figure 4d).

Precision

Sampling method affected the precision of both the SCIBI and O/E scores (Table 6). For example, the RWB sampling method had the largest MDD for the SCIBI: 22 vs. 19 for the other two methods. However, RWB had the lowest MDD when O/E

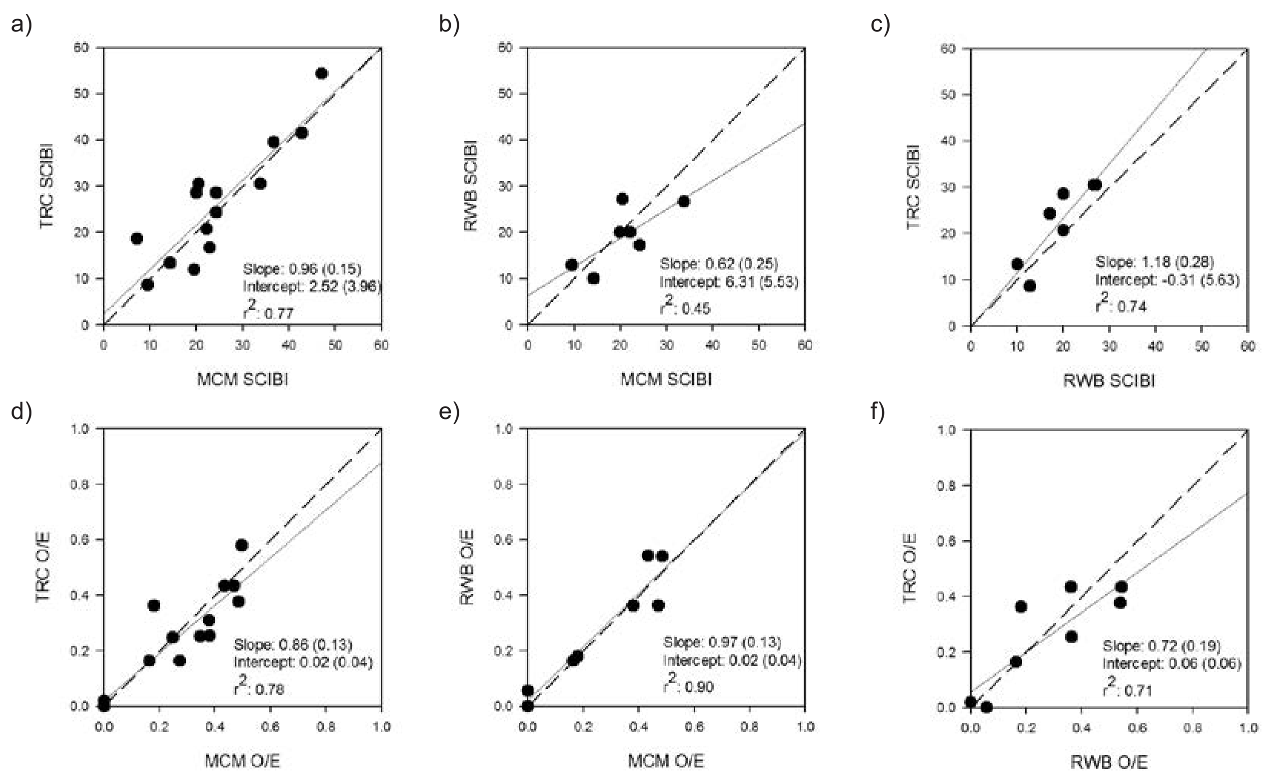


Figure 3. Agreement between the sampling methods for Southern California Index of Biotic Integrity (SCIBI; a – c) and Observed/Expected (O/E; d – f) scores Each point represents the mean index score at a site. Solid lines represent linear regressions, and dashed lines represent perfect 1:1 relationships. Numbers in parentheses are standard errors. Slopes were tested against 1, and intercepts were tested against 0.

Table 5. Spearman rank correlations (ρ) between bioassessment indices and landscape metrics. * = statistical significance (p <0.008).

Index	Land Cover	Method	Watershed			1 km radius		
			n	ρ	p	n	ρ	P
SCIBI	% Developed	MCM	15	-0.08	0.783	15	0.11	0.685
		RWB	7	-0.75	0.054	7	-0.59	0.159
		TRC	14	-0.32	0.914	14	0.20	0.487
	% Open	MCM	15	-0.04	0.892	15	0.09	0.742
		RWB	7	0.62	0.139	7	0.67	0.102
		TRC	14	-0.04	0.890	14	-0.08	0.782
	% Agricultural	MCM	15	0.06	0.842	15	-0.11	0.689
		RWB	7	0.12	0.799	7	0.22	0.628
		TRC	14	0.00	0.991	14	-0.02	0.954
O/E	% Developed	MCM	15	0.14	0.640	15	0.35	0.202
		RWB	8	-0.28	0.509	8	-0.07	0.866
		TRC	17	0.23	0.370	17	0.31	0.222
	% Open	MCM	15	-0.05	0.857	15	0.01	0.980
		RWB	8	0.40	0.333	8	0.17	0.693
		TRC	17	-0.24	0.355	17	0.02	0.948
	% Agricultural	MCM	15	-0.67	0.009	15	-0.24	0.388
		RWB	8	-0.89	0.003*	8	-0.15	0.719
		TRC	17	-0.46	0.064	17	-0.31	0.220

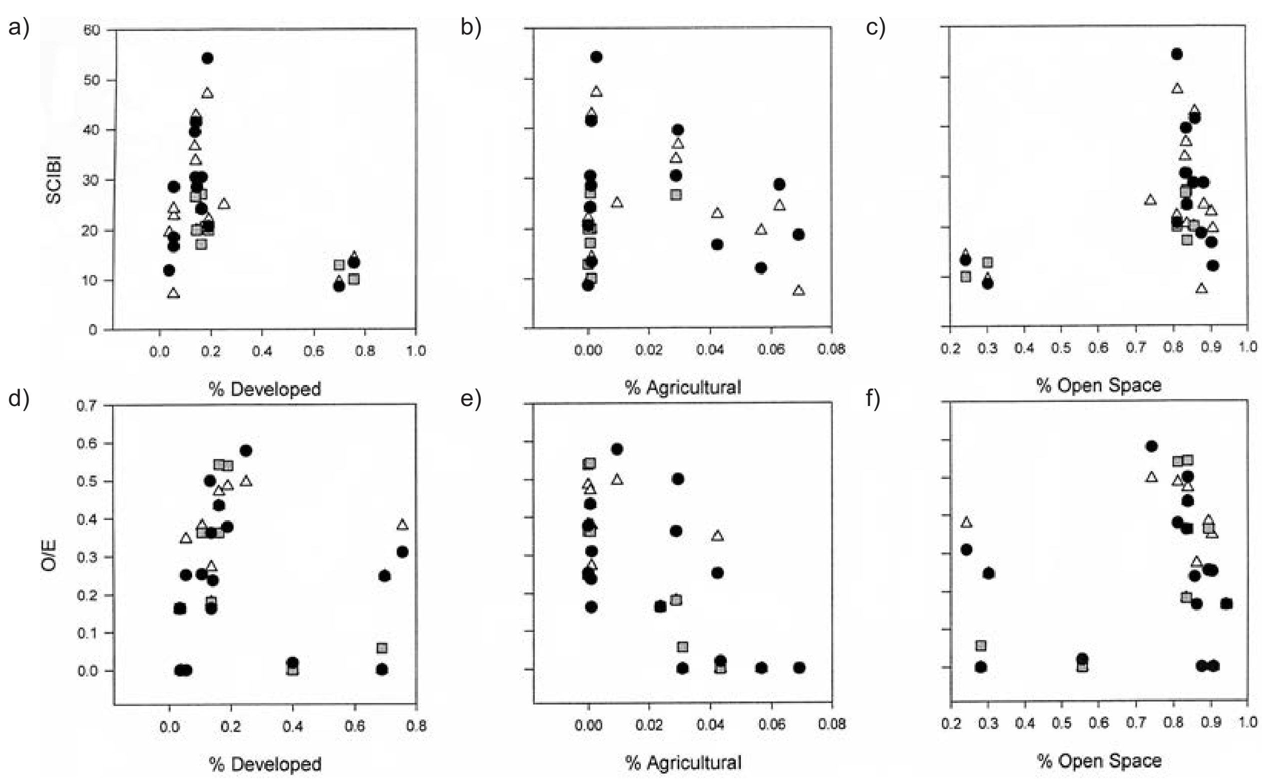


Figure 4. Index scores versus land cover metrics. Each point represents the mean of all samples collected by one method at each site. White triangles represent MCM samples. Gray squares represent RWB samples. Black circles represent TRC samples.

Table 6. Within-site variability (expressed as mean square error, MSE) and minimum detectable difference (from a two-sample, 2-tailed t-test with n = 30, α = 0.05, and β = 0.1) for each of the sampling methods. d.f.: degrees of freedom. SS: sum of squares. MSE: mean square error. MDD: mean detectable difference.

Index	Method		d.f.	SS	MSE	F	p	MDD
SCIBI	TRC	Sites	7	2507	358	12.5	>0.0001	19
		Residuals	15	430	29			
	RWB	Sites	3	403	134	3.7	0.0701	22
		Residuals	7	254	36			
	MCM	Sites	8	1745	218	8.0	0.0002	19
		Residuals	16	437	27			
O/E	TRC	Sites	8	0.625	0.078	12.7	>0.0001	0.28
		Residuals	13	0.074	0.006			
	RWB	Sites	3	0.115	0.038	14.5	0.0037	0 20
		Residuals	6	0.016	0.003			
	MCM	Sites	9	0.860	0.096	20.9	>0.0001	0.24
		Residuals	17	0.078	0.005			

scores were used: 0.20 vs. 0.28 for TRC and 0.24 for MCM. Coefficients of variation showed similar trends in variability among methods when SCIBI scores were used, (ranging from 22 to 27%), and lower CVs for RWB when O/E scores were used: 12 vs. 20% for MCM and 45% for TRC.

The low number of samples containing adequate numbers of individuals meant that estimates of within-site variance were sometimes based on very small samples. For example, only four sites in the region using the SCIBI had multiple samples with sufficient numbers of organisms collected by the RWB method. This problem was less severe for estimates based on O/E scores because fewer individuals per sample are required for index calculation, and because sites in the Central Valley and San Francisco Bay area could be included in the estimates.

Discussion

Low-gradient streams are distinct from other streams in many aspects, such as substrate material, bed morphology, and the distribution of microhabitats (Montgomery and Buffington 1997). As a consequence of these differences, traditional bioassessment approaches in California that were developed in high-gradient streams with diverse microhabitats have limited applications in low-gradient reaches. The sampling methods evaluated in this study differed in the extent to which they targeted the richest microhabitats (such as riffles, or vegetated margins). For example, the TRC method allows field crews to select the richest microhabitats specifically. In contrast, the RWB method may systematically undersample or miss these habitats entirely, as the richest areas in low-gradient streams are typically found at the margins (Montgomery and Buffington 1997). The MCM method, a modification of the RWB method, was designed so that these margins could be targeted.

Caution should be used when applying sampling methods or assessment tools that were calibrated for specific habitat types (e.g., high-gradient streams) to new habitats (e.g., low-gradient streams). The present study's evaluation of assessment tools unveiled a number of shortcomings that weaken application of these tools in low-gradient streams, including the inability to collect adequate numbers of organisms, poor sensitivity of assessments, and low precision of the sampling methods. Significant disagreements among the methods were not detected, although power was low because of the low number of samples. The inability of the RWB sampling method to collect an adequate number of individuals in nearly half of all samples makes it unsuitable for low-gradient streams, even though this method is widely used by bioassessment programs in California (Ode 2007) and across the USA (Peck *et al*. 2006). Although biomonitoring programs must assess a diverse range of habitat types with available tools, the present study indicates that these programs may be well served by evaluating tools in novel habitats where monitoring activities occur.

Variance components analysis of assessment indices showed that differences among sites explained more of the variance in index scores than differences among sampling methods, suggesting that similar types of benthic macroinvertebrates are collected by the different methods. However, analysis of disagreements among the methods indicated that some samples collected by RWB were distinct from those collected by TRC, and samples collected by MCM were intermediate between the other two. For example, samples collected by TRC had lower O/E scores than samples collected by MCM, which in turn were lower than those collected by RWB. However, differences among these methods did not reach statistical significance.

Other studies comparing single, targeted habitat sampling methods (e.g., TRC) to multi-habitat sampling methods (e.g., RWB) have shown similar results. For example, MDDs reported in other studies (or calculated from reported variabilities) were comparable to those reported here, although generally larger (Rehn *et al*. 2007, Blocksom *et al*. 2008). However, these studies found that multi-habitat sampling reduced variability in multimetric indices, whereas the present study found that variability was lower for the single habitat method (i.e., TRC; Table 7). As in Rehn *et al*. (2007), the present study found that TRC samples had higher O/E scores than RWB samples, but that the strength of disagreement was inconsistent in the largest watersheds.

The generally weak response of the indices to land cover metrics suggests that the SCIBI and O/E may not be sensitive to variability in watershed-scale disturbance in low-gradient streams. This conclusion is tempered by small sample sizes that limited power, and sensitivity to reach-scale degradation was not explored in this study for lack of data. Several studies have shown the strong impact of reach-scale factors on benthic macroinvertebrates, which may exceed the influence of watershed-scale stressors (e.g., Hickey and Doran 2004, Sandin and Johnson 2004). Furthermore, most of the watersheds in the study were highly altered, particularly those in the region of the SCIBI, and portions of the disturbance gradient to which these indices are more sensitive may not have been adequately sampled. Several studies have found that biota responds to disturbance gradients ≤10% development in a watershed, but responses above this gradient are muted (e.g., Hatt *et al*. 2004, Walsh *et al*. 2007). Agricultural land cover, which was low in most watersheds (<10%), showed strong responses with the indices, suggesting that the study was able to capture portions of this gradient to which both the SCIBI and O/E were sensitive.

The low numbers of organisms collected from the low-gradient streams in the study may reflect the naturally low population densities of benthic macroinvertebrates in these reaches. The River Continuum Concept hypothesizes that higher order streams with larger watersheds have a lower energy base because of reduced allochthonous input and depressed autochthonous productivity (Vannote *et al*. 1980). This lower energy base would be expected to support reduced biomass. However, observation of the sites in this study suggests that the lack of stable microhabitats (e.g., riffles and vegetated margins) may account for the reduced numbers of macroinvertebrates, as few species are adapted to the shifting sandy substrate found in most low-gradient streams in California. A well known, but extreme, example of the impact of shifting sandy substrates on maintaining low densities of benthic macroinvertebrates are the migrating submerged dunes in the lower

Table 7. Minimum detectable differences in multimetric indices. Southern California Index of Biotic Integrity (SCIBI); Northern California Index of Biotic Integrity (NICIBI); Virginia Stream Condition Index (VSCI); Macroinvertebrate Biotic Integrity Index (MBII); California O/E Index (O/E); and NT = not tested.

Index type	Method	Present study	Rehn *et al*. 2007	Blocksom *et al*. 2008	
Multimetric index	Single-habitat	19.2 (SCIBI)	19.7 (SCIBI + NCIBI)	19.88 (VSCI)	29.79 (MBII)
	Multi-habitat	22.6 (SCIBI)	15.5 (SCIBI + NCIBI)	17.37 (VSCI)	17.91 (MBII)
Predictive model	Single-habitat	0.28 (O/E)	0.22 (O/E)	NT	NT
	Multi-habitat	0.20 (O/E)	0.19 (O/E)	NT	NT

Amazon River (Sioli 1975, Lewis, Jr. *et al*. 2006). Although very high productivity of Chironomidae and other benthic macroinvertebrates has been observed in low-gradient sandy rivers of the southeastern United States, this productivity was attributed to snags and other stable microhabitats, more than to the shifting sandy substrate (Benke 1998). Thus, the vast majority of the macroinvertebrate activity in a large reach of river was found in small areas containing snags (Wallace and Benke 1984). Snag microhabitats are arguably less common in streams of the arid Southwest, which lack dense riparian forests to contribute snag-forming woody debris and may be less likely to be sampled using a systematic sampling method like RWB.

Bioassessment programs are often required to make do with available tools to fulfill regulatory mandates, yet they lack resources to evaluate the tools for applications in all habitats of concern. Although all sampling methods in this study suffered from poor efficiency in collecting organisms, the MCM method greatly improved efficacy and reduced the frequency of rejected samples. Furthermore, the lack of significant disagreements and inconsistencies suggests that the MCM method produced results that were comparable to the other methods already in use in California, which may facilitate integration of historical data sets (Cao *et al*. 2005, Rehn *et al*. 2007). Therefore, the present study supports the use of MCM in low-gradient streams in California as a substitute for the currently preferred RWB method. Overall, bioassessment programs can improve data quality and avoid unnecessary expenses by explicitly evaluating assessment tools when assessing novel habitat types.

Literature Cited

Benke, A.C. 1998. Production dynamics of riverine chironomids: Extremely high biomass turnover rates of primary consumers. *Ecology* 79:899-910.

Blocksom, K.A., B.C. Autrey, M. Passmore and L. Reynolds. 2008. A comparison of single and multiple habitat protocols for collecting macroinvertebrates in wadeable streams. *Journal of the American Water Resources Association* 44:577-593.

Cao, Y., C.P. Hawkins and A.W. Storey. 2005. A method for measuring the comparability of different sampling methods used in biological surveys: implications for data integration and synthesis. *Freshwater Biology* 50:1105-1115.

Carter, J.L. and V.H. Resh. 2005. Pacific Coast Rivers of the Coterminous United States. pp. 541-590 *in*: A.C. Benke and C.E. Cushing (eds.), Rivers of North America. Elsevier Academic Press. Boston, MA.

Courbois, J.-Y.P. and N.S. Urquhart. 2004. Comparison of survey estimates of the finite population variance. *Journal of Agricultural, Biological, and Environmental Statistics* 9:236-251.

Fore, L.S., K. Paulsen and K. O'Laughlin. 2001. Assessing the performance of volunteers in monitoring streams. *Freshwater Biology* 46:109-123.

Harrington, J. M. 1999. California Stream Bioassessment Procedures. California Department of Fish and Game, Aquatic Bioassessment Laboratory, Water Pollution Control Laboratory. Rancho Cordova, CA.

Hatt, B.E., T.D. Fletcher, C.J. Walsh and S.L. Taylor. 2004. The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams. *Environmental Management* 34:112-124.

Hickey, M.B.C. and B. Doran. 2004. A review of the efficiency of buffer strips for the maintenance and enhancement of riparian ecosystems. *Water Quality Research Journal* 39:311-317.

Larsen D.P., T.M. Kincaid., S.E. Jacobs and N.S. Urquhart. 2001. Designs for evaluating local and regional scale trends. *Bioscience* 51:1069-1078.

Lewis, Jr., W.M., S.K. Hamilton and J.F. Saunders, III. 2005. Rivers of northern South America. pp. 219-256 *in*: C.E. Cushing, K.W. Cummins and G.W. Minshall (eds.), River and Stream Ecosystems of the World. University of California Press. Berkeley, CA.

Mantel, N. 1967. The detection of disease clustering and generalized regression approach. *Cancer Research* 27:209-220.

Montgomery, D. and J. Buffington. 1997. Channel-reach morphology in mountain drainage basins. *Geological Society of America Bulletin* 109:596-611.

Ode, P.R. 2007. Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Available from http://mpsl.mlml.calstate.edu/phab_sopr6.pdf

Ode, P.R. and B.H. van Buuren. 2008. Amendment to SWAMP Interim Guidance on Quality Assurance for SWAMP Bioassessments. Surface Water Ambient Monitoring Program. Sacramento, CA.

Ode, P.R., A.C. Rehn and J.T. May. 2005. A quantitative tool for assessing the integrity of southern coastal California streams. *Environmental Management* 35:493-504.

Ode, P.R, C.P. Hawkins and R.D. Mazor. 2008. Comparability of biological assessments derived from predictive models and multimetric indices of increasing geographic scope. *Journal of the North American Benthological Society* 27:967-985.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Kemm, J.M. Lazorchek, F.H. McCormick, S.A. Peterson, S.A. Ringold, T. Magee and M. Cappaert. 2006. Environmental Monitoring and Assessment Program—Surface Waters Western Pilot study: Field operations manual for wadeable streams. EPA/620/R-06/003. United States Environmental Protection Agency, Office of Research and Development. Corvallis, OR.

Rehn, A.C., P.R. Ode and C.P. Hawkins. 2007. Comparison of targeted-riffle and reach-wide benthic macroinvertebrate samples: implications for data sharing in stream-condition assessments. *Journal of the North American Benthological Society* 26:332-348.

Richards, A.B. and D.C. Rogers. 2006. List of Freshwater Macroinvertebrate Taxa from California and Adjacent States including Standard Taxonomic Effort Levels. Southwest Association of Freshwater Invertebrate Taxonomists. Accessed online March 1, 2008: http://www.swrcb.ca.gov/swamp/docs/safit/ste_list.pdf

Sandin, L. and R.K. Johnson. 2004. Local, landscape and regional factors structuring benthic macroinvertebrate assemblages in Swedish streams. *Landscape Ecology* 19:504-514.

SAS Institute, Inc. 2004. SAS OnlineDoc 9.1.3. Cary, NC.

Sioli, H. 1975. Tropical river: The Amazon. pp. 461-488 *in*: B.A. Whitton (ed.), River Ecology. University of California Press. Berkeley, CA.

Stokes, M.E., C.S. Davis and G.G. Koch. 2000. Categorical Data Analysis Using the SAS System. 2nd edition. SAS Institute, Inc. Cary, NC.

Stormwater Monitoring Coalition Bioassessment Working Group (SMCBWG). 2007. Regional Monitoring of Southern California's Coastal Watersheds. Technical Report 539. Southern California Coastal Water Research Project. Costa Mesa, CA.

Theil, H. 1958. Economic Forecasting and Policy. North Holland Publishing Co. Amsterdam, The Netherlands.

United States Environmental Protection Agency and United States Geological Survey (USEPA and USGS). 2005. National Hydrography Dataset Plus. Edition 1.0. Available from: http://www.horizon-systems.com/nhdplus/

United States Geological Survey (USGS). 2003. National Land Cover Database. Edition 1.0. http://www.lsc.usgs.gov.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell and C.E. Cushing. 1980. The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Science* 37:130-137.

Wallace, J.B. and A.C. Benke. 1984. Quantification of wood habitat in subtropical Coastal Plain streams. *Canadian Journal of Fisheries and Aquatic Sciences* 41:1643-1652.

Walsh, C.J., K.A. Waller, J. Gehling and R. MacNally. 2007. Riverine invertebrate assemblages are degraded more by catchment urbanisation than by riparian deforestation. *Freshwater Biology* 52:574-587.

Zar, J. 1999. Biostatistical Analysis. 4th edition. Prentice Hall. Upper Saddle River, NJ.

Acknowledgements

The authors thank the Aquatic Bioassessment Laboratory of the California Department of Fish and Game for laboratory processing and identification; Aquatic Bioassay and Consulting, Weston Solutions, Inc., and California Regional Water Quality Control Boards for field sampling; and Chuck Hawkins for assistance with predictive models. This project was partially supported by the Stormwater Monitoring Coalition of Southern California.

Appendix 10

B309

Surface Water Ambient Monitoring Program

Final Technical Report	2009

Evaluation of the California State Water Resource Control Board's Bioassessment Program

March 15, 2009

California Water Boards

www.waterboards.ca.gov/swamp

B310

Evaluation of the California State Water Resource Control Board's Bioassessment Program

March 15, 2009

Final Report to U.S. EPA-OST and Region IX

Chris O. Yoder, Research Director
Center for Applied Bioassessment and Biocriteria
Midwest Biodiversity Institute
P.O. Box 21561
Columbus, OH 43221-0561

and

Rob Plotnikoff, Senior Aquatic Scientist
Tetratech, Inc.
Center for Ecological Sciences
400 Red Brook Blvd.
Suite 200
Owings Mills, MD 21117

EXECUTIVE SUMMARY

The State of California's bioassessment, monitoring and assessment (M&A), and water quality standards (WQS) programs were reviewed in January 2008 using the U.S. EPA's Critical Technical Elements and Programmatic Review process (Barbour and Yoder 2008; Quasney and Yoder 2008), which evaluates key components of these state programs and existing and planned capacities. The review process results in technical, policy, and management recommendations for building, refining and maintaining functional and effective bioassessment and M&A tools that support the full spectrum of WQS and management programs. This review was conducted by a two person review team with national expertise at evaluating, building, and implementing state and tribal programs.

Bioassessment, the use of resident aquatic biota as direct indicators of the biological integrity of water bodies, is a powerful tool for water resource regulatory programs. The need for state water quality agencies to develop and maintain robust bioassessment programs is underscored by the National Research Council's critical review of state TMDL, M&A, and WQS programs (National Research Council 2001). The NRC's review makes clear that all states need better biological endpoints, adequate M&A, and tiered aquatic life uses (TALU) in order to develop and refine appropriate and effective WQS that result in more accurate and appropriate management outcomes including TMDLs.

While the federal Clean Water Act (CWA) has long required states to protect and restore the chemical, physical, **and** biological integrity of the nation's waters, California has only recently begun to consider the developments that it will need in order to implement modernized WQS that lead to more effective water quality management programs. Because of the prior investment in the development of bioassessment tools since the mid-1990s, California is now positioned to initiate the process of integrating bioassessments into their WQS and monitoring and assessment programs via the development and implementation of narrative and numeric biocriteria.

Key Findings of the Review:

1. California's bioassessment program has made great strides in recent years due primarily to investments made by the Dept. of Fish and Game's Aquatic Bioassessment Laboratory (DFG-ABL) and the Water Board's Surface Water Ambient Monitoring Program (SWAMP). With continued management support, SWAMP is capable of building, maintaining and refining the technical tools that the Water Boards will need to incorporate biological criteria and assessments into their water quality programs.

2. As determined by the U.S. EPA Critical Technical Elements methodology California's bioassessment program is currently at an above average level of rigor (Level 3; 88.3%) and is being used in statewide 305(b) assessments, the 303(d) listing/delisting process, and in support of specific regulatory needs in selected Regions. Continued investment

and active management support will be needed to achieve a fully functional (CE Level 4) program that will provide more comprehensive support for the suite of regulatory needs and in all Regions.

3. California's bioassessment program is currently capable of addressing wadeable perennial streams. Additional investment and technical development will be needed to address other waterbody types including large non-wadeable rivers, non-perennial streams, lakes, and wetlands.

4. SWAMP has invested a significant amount of financial resources to develop the current bioassessment infrastructure. However, full implementation of California's bioassessment program is constrained by the fact that most of the program is conducted by contractors. This review affirms the findings of prior peer reviews that the Water Board needs its own in-house bioassessment coordinator and staff. This will enhance the integration of monitoring and assessment results in al facets of water quality management.

5. While the DFG-ABL, SWAMP, and their contractors are building a solid technical foundation for a robust freshwater bioassessment program, they can only provide the technical tools for developing biological endpoints and Tiered Aquatic Life Uses (TALUs). The State and Regional Water Boards will need additional biologists and planning staff to develop, refine and implement narrative/ numeric biocriteria and TALUs in support of all applicable regulatory programs and at the same spatial scale at which they are being applied.

Management Recommendations:

1. The Water Boards should revise the structure and content of the beneficial uses and criteria related to aquatic life uses to more accurately reflect the natural attributes of the diversity of watersheds through the state. This is consistent with recommendations from the NRC (2001) and the SPARC (2006).

2. The Water Boards should integrate biological assessment tools into their water quality programs (WQS, NPDES, and TMDLs). This represents a fundamental paradigm shift that will require strong management understanding and support.

3. The State Water Board should develop statewide narrative biocriteria which incorporate numeric biological endpoints to interpret the narrative objectives for aquatic life use protection as soon as possible.

4. The Water Boards should require key program units (e.g., WQS, NPDES, TMDL) to incorporate biological assessments into their programs and program evaluation. Adopting biological criteria within a framework of TALUs would enhance its implementation in these programs.

5. The Water Boards should assign staff and provide training to programs incorporating biological assessments. This includes support for statewide efforts and ongoing efforts at the Regional Boards.

6. The State Water Board should create and maintain a specialist position for a state-wide bioassessment policy coordinator. This is consistent with the recommendation made in a prior external peer review of SWAMP's bioassessment program (Barbour and Hill 2003).

Technical Recommendations: [NOTE: the following recommendations are based in part on the Critical Elements evaluation conducted during the January 2008 program review and are based on elevating the technical rigor of the statewide and regional board programs to level 4.]

1. SWAMP should continue to support the technical infrastructure development strategy outlined in its FY06-07 and FY07-08 bioassessment work plans.

2. SWAMP should establish reference conditions for the objective interpretation of biological data and implement the reference condition management plan. This investment will pay dividends to all water quality programs using biological assessments. This would also serve the development of chemical/physical endpoint s and indicators as part of a program of integrated bioassessment.

3. SWAMP should develop additional indicators of ecological condition to supplement the benthic macroinvertebrate indicators currently in use. The consistent addition of a second assemblage in the bioassessment process is needed to elevate the program to level 4. Options for this include an algal assemblage indicator (currently in development by SWAMP), a wetland indicator (CRAM, also in development), and fish assemblage indicators (currently in development by USGS). SWAMP should continue to support efforts to determine which supplemental indicators are best suited to California's needs and for specific waterbody ecotypes (perennial wadeable streams, non-perennial streams, non-wadeable large rivers, wetlands).

4. SWAMP should continue to support development and maintenance of the biological component of the database. This provides the essential framework for statewide integration of biological and physical habitat data. Two priorities are tools to calculate biological metrics for water resource managers and tools to convey results to the public.

5. SWAMP should develop a QA/QC oversight program for the collection of ambient biological data. This would set the standard for SWAMP comparability for other Water Board programs and provide guidance to other agencies wishing to become SWAMP compatible. Adopting biocriteria and TALUs in the WQS would contribute to the compulsory standardization of the use of biological assessment data throughout the state and between the regions.

6. SWAMP should continue to support the statewide perennial stream assessment. This addresses the need to assess the condition of a major class of surface waters in California and provides a solid framework for integrating stream monitoring with other programs in the state.

TABLE OF CONTENTS

Section – Heading/Subheadings **Page**

INTRODUCTION

U.S. EPA has supported the development of state and tribal bioassessment programs via the production of methods documents, case studies, regional workshops, and evaluations of individual state and tribal programs since 1990. Since 2000, EPA has fostered a more detailed and “hands on” developmental and implementation process for incorporating tiered aquatic life uses (TALUs) and numeric biocriteria in state and tribal water quality programs. The successful development and implementation of biocriteria and TALUs is directly dependent on the rigor, comprehensiveness, and integration of monitoring and assessment (M&A) with state water quality standards (WQS) and water quality management programs. This framework can also provide measures to evaluate the effectiveness of major water quality management programs such as NPDES permitting, TMDLs, nonpoint source management, stormwater management, and watershed planning.

On January 23-24, 2008 the U.S. EPA sponsored an evaluation of the Water Board’s biological assessment program. The purpose was to evaluate both the State’s technical program elements and its regulatory structure in order to make recommendations that will enhance CA’s ability to make informed decisions about the ecological condition and management of California’s rivers and streams. The scope of the review included a range of topics about the surface water monitoring and assessment program, the structure of the existing WQS, the development of bioassessment tools to delineate impaired waters and determine stressor effects, and the use of biological data to support Water Board programs including NPDES permitting, non point source management, stormwater management, and TMDLs.

The evaluation process consisted of direct interactions with state program management and staff to evaluate the status of their bioassessment, M&A, and WQS programs and to describe how each is used to support water quality management. The following include the principal reports and products of the EPA TALU development and implementation process since 1998.

1) *Important Concepts and Elements of a State Watershed Monitoring and Assessment Program (Yoder 1998):* This document was developed as a state oriented document following the Intergovernmental Task Force on Monitoring Water Quality and the U.S. EPA environmental indicators initiatives of the 1990s. It outlines the essential concepts and elements of what is referred to as an “adequate” state monitoring and assessment program. The term adequate was chosen to represent a cost-effective, yet comprehensive approach to monitoring that assures the use and development of chemical, physical, and biological indicators collected and arrayed in a strategic manner that results in supporting water quality management decisions at all relevant scales.

2) *Use of Biological Information to Better define Designated Aquatic Life Uses in State and Tribal Water Quality Standards: Tiered Aquatic Life Uses (August 2005):* This document serves as a detailed presentation of methods for developing and implementing TALUs in state WQS. It consists of detailed descriptions of the baseline elements of TALU – the

Biological Condition Gradient, elements and milestones for the incorporation of TALUs in WQS.

3) *Critical Elements Technical Elements of a Bioassessment Program (November 2007; updated September 2008):* The rigor of a state's program is evaluated in order to determine the capacity to assess ecological condition and diagnose impairment. This evaluation consists of thirteen technical elements associated with design, methods, and interpretation features of a bioassessment program that are rated jointly with the state program management and staff. The cumulative rating of the elements provides a level of rigor (ranging from level 1, the lowest level of rigor, to level 4, the highest and best suited for full program support) of the overall bioassessment program. The capacity to accurately address a suite of management questions and issues is dependent upon the level of rigor. A critical technical elements evaluation of the California bioassessment program was completed using the standardized checklist and scoring methodology (Barbour and Yoder 2007, 2008).

Part 1. Use of Bioassessment in State Water Board Programs

1. Monitoring and Assessment Program

The Surface Water Ambient Monitoring Program (SWAMP) is a statewide effort designed to monitor and assess the conditions of surface waters throughout the state of California. SWAMP was proposed in 2000 (SWRCB 2000) in response to a legislative directive to integrate existing water quality monitoring activities of the State Water Resources Control Board and the nine Regional Water Quality Control Boards (Regional Water Boards), and to coordinate with other monitoring programs. The needs of an emerging TMDL process, inconsistencies between regional boards, and information needs for regulatory decision-making were some of the principal drivers.

SWAMP has fostered the development of biological assessments because they provide a direct and quantitative measure of aquatic life use protection. A major review of the program was conducted in 2003 (Barbour and Hill 2003). In 2005, the SWAMP Scientific Planning and Review Committee (SPARC 2006) recommended "The State Board should adjust water quality management approaches to take advantage of the more direct measures SWAMP is developing of aquatic life condition through bioassessment monitoring".

Tools for assessing biological assemblages in perennial wadeable streams are currently the most well-developed of the biological monitoring tools; this is largely the result of investments made by the Department of Fish and Game's Aquatic Bioassessment Laboratory (DFG-ABL) and SWAMP since the mid 1990s. The State has made significant progress with the use of benthic macroinvertebrates in stream bioassessments, but has recently begun to develop and implement an Algae Plan (SCCWRP 2008), and is also evaluating the utility of the California Rapid Assessment Methodology (CRAM) as a tool for assessing riparian wetland habitat. SWAMP is also considering the utility of fish bioassessments in California. It is also recognized that there are additional freshwater ecotypes and strata that need to be addressed to meet the goal of providing full water quality management program support (Table 1)

Ecotype	Habitat	Algae	Invertebrates	Fish
Ephemeral	Y			
Intermittent	Y	?	?	?
Perennial	pHab CRAM	% cover Biomass Algal IBI	IBI or O/E	?
Rivers	pHab CRAM	Y	Y	Y
Lakes/Reservoirs	pHab CRAM	Y	Y	Y
Bay/Estuaries	CRAM	Y	BRI	Y
Coast/Ocean		Y	So Cal BRI	So Cal Fish Index

Table 1. Summary of biological indicator development efforts in California by major aquatic ecotypes.

Additional investment will be needed to develop and maintain a program that is capable of addressing other waterbody types (e.g., large non-wadeable rivers, non-perennial streams, lakes, wetlands). Indicator work done on perennial streams may applicable to other waterbody types. For instance studies are underway to investigate the use of macroinvertebrate assemblages and periphyton to assess intermittent and ephemeral streams. The CRAM wetland methodologies can be applied to intermittent and ephemeral streams but also to lakes and estuaries. California has also participated in national and regional bioassessment projects such as U.S. EPA-EMAP and REMAP surveys, the National Wadeable Streams Assessment, the National Lakes Assessment, and the National Rivers and Streams Assessment each of which lends experience with these other waterbody types.

California has begun moving from conducting simple biosurveys (i.e., the collection of limited sets of biological samples) to more spatially robust bioassessments of ecological condition. This has occurred within selected Regional Boards and these can serve as a template for all Water Boards. The next challenge will be the development of biological criteria to better inform and guide water quality management decision-making. While SWAMP and the selected Region Board programs have contributed the technical rigor required by this process, it will require considerations that apply within specific regions of the state. Hence it needs to be a coordinated effort with consistent participation and integration between the state and regional water boards.

2. Role of Bioassessment in Listing Decisions

Waterbody listings are presently based on exceedences of water quality criteria. The State Board's listing policy (SWRCB, 2004) provides detailed guidance on the interpretation of chemical and toxicity data. Listing and de-listing decisions are based on the frequency with which numeric water quality criteria are exceeded as defined in the listing policy and interpreted through the use of a binomial probability distribution. Assessment of physical and biological data is more difficult because there are no numeric criteria. Listings are therefore based on the interpretation of narrative criteria.

A water body may be listed if there is significant degradation in biological populations and/or assemblages as compared to reference site(s), but only if it is associated with a pollutant. The analysis of biological communities must rely on measurements conducted using published protocols from at least two stations and requires that comparisons to reference site conditions shall be made during similar seasonal and/or hydrological periods.

Regional Boards using biological information in the listing process are required to: 1) identify appropriate reference sites and document methods for the selection of reference sites, 2) document the sampling methods, index period and quality assurance/quality control procedures for the habitats being sampled, and 3) compare bioassessment data to conditions at reference sites and evaluate physical and other water quality data to support any assessment conclusions. The listing policy encourages the use of indices of biological integrity (such as the IBIs developed by SWAMP).

A significant number of waterbodies have been listed in the past for sediment, excess algae, hydromodification, and water diversions using best professional judgment to interpret narrative standards in the Basin Plans. The lack of a quantitative biological endpoint or numeric biocriteria for attainment of aquatic life can create challenges for managers.

3. Water Quality Standards

Water quality standards (WQS) provide the objectives for both developing the requirements for and judging the effectiveness of pollution controls and management programs. The California WQS are comprised of beneficial uses, numeric and narrative criteria (objectives) to protect those beneficial uses, and an antidegradation policy.

<u>Beneficial Uses</u>. At present there are 6 defined "beneficial uses" that apply to the protection of aquatic life use in fresh water across the state. These are cold fresh water habitat (COLD), warm fresh water habitat (WARM), spawning (SPAWN) and migration (MIGR) of aquatic species, habitat for wildlife in general (WILD) and for rare, threatened, and endangered species (RARE), and the preservation of biological habitats of special significance (BIOL). These uses are applied to specific watersheds through Regional Water Quality Control Plans (Basin Plans) that

are developed, administrated and enforced by the Regional Water Boards. Two Regional Boards have wetland habitat (WET) as a defined beneficial use and the State Board is considering application of the wetland use as part of a hydromodification policy.

These aquatic life use designations define the general types of organisms, assemblages and habitats that are being protected. For instance, COLD use designation protects “uses of water that support cold water ecosystems including, but not limited to, preservation and enhancement of aquatic habitats, vegetation, fish, or wildlife, including invertebrates”. Aquatic life use support assessment is challenging in California because expectations for the aquatic life use support will naturally vary across the state. The current generic aquatic life use designations simply do not account for the natural variability in rivers and streams across the broad biogeographic regions of the state.

The SPARC recommended that the Water Boards use the National Research Council (2001) recommended framework to revise and refine the designated uses, the supporting protective criteria, and the attainment assessment procedures to more fully reflect the diversity of watersheds and their respective/desired attainable human and aquatic life uses. U.S. EPA (2005) largely followed the NRC (2001) recommendations in their methodological guidance for developing and implementing a TALU approach to WQS and monitoring and assessment. That framework and the technical developments to date are the basis for this review.

<u>Numeric and Narrative Objectives</u>. There are relatively few numeric objectives for the protection of aquatic life. The California Toxics Rule contains numeric water quality objectives for 22 chemicals. The Basin Plans have limited objectives for additional toxics. Narrative objectives in the Basin Plans related to the protection of aquatic life use are generally expressed in the form of “no toxics in toxic amounts”, “no significant degradation”, or “no significant deviation from reference”. State and Regional Board staff engaged in assessments have little guidance on how to interpret these narrative objectives. A TALU framework and numeric biocriteria would greatly clarify these endpoints.

The biological information being generated through SWAMP can be used to establish biological expectations for different waterbodies across the state. This is a first step in the establishment of biological criteria. Such information and data may also be used to support the development or refinement of other water quality objectives (e.g. temperature, dissolved oxygen or nutrient criteria) or program applications (e.g., 401 certifications) across the state. The SWAMP Reference Condition Management Plan (Ode and Schiff, 2008) lays out a strategy for establishing biological expectations.

<u>Antidegradation</u>. The state’s antidegradation policy is incorporated in the Basin Plans by reference. Biological information is not typically used in antidegradation analyses, but it has the potential to enhance their application. Biological assessment could be used as a direct measure of instream aquatic life use and to provide a trigger for antidegradation analyses when such assessments indicate that there is degradation of water quality. The biological assessment tools developed already provide a method to measure condition incrementally thus enhancing

its utility for detecting incremental changes that may not reflect a violation of standards. This capacity will enhance its usefulness in new antidegradation applications.

4. Use of Bioassessment in other Board Programs

Monitoring and assessment activities should be designed to provide information and tools to support multiple programmatic activities with the same data and information. As biological assessments provide a direct measure of aquatic life use they can help program managers prioritize management actions to protect and restore beneficial uses. They can also be used as outcome measures to evaluate the effectiveness of various programs (e.g., NPS, NPDES, and TMDLs) to protect and restore beneficial uses.

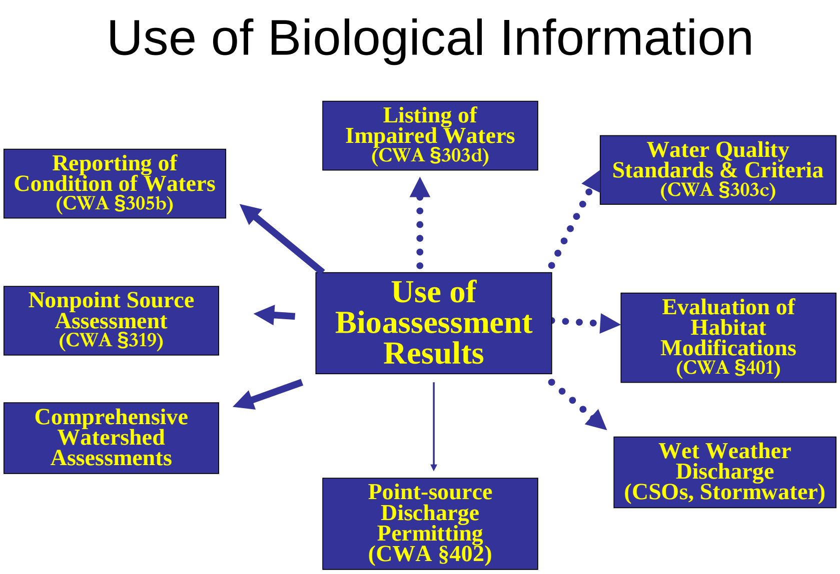


Figure 1. Efforts to develop strong monitoring and assessment programs lead to support for multiple water quality management programs.

<u>NPDES</u>. The use of biological data in NPDES permits and WDRs in California dates back to the early 1990's. Bioassessments have been used mostly in "upstream-downstream" designs to assess the impact of point source dischargers such as POTWs, but are also increasingly being used in stormwater permits. In Southern California alone, 323 bioassessment samples are collected by stormwater agencies each year as part of their MS4 permit requirements. The State Board's SWAMP program is developing draft permit language to assist Water Board staff that wish to incorporate freshwater bioassessment into permit requirements and/or other Water Board programs or projects. The boilerplate language will include guidance on field and lab methods, index periods for sampling, and the required QA and data submittal procedures. Interpretation of bioassessment results have largely been relative to reference site or locally derived IBIs, where available.

<u>NPS</u>. Bioassessments have been used in a number of nonpoint source projects to assess the effectiveness of actions on water quality (instream biota). The State Board's nonpoint source

program has helped fund monitoring of perennial streams to identify the extent of the states streams that are impacted by nonpoint source pollution and to identify the stressors that are impacting streams (Ode, 2007). Funds have also been used to support development of stressor identification tools (Rehn, 2006) and improve understanding of associations between biological assemblages and key stressors associated with NPS activities (e.g., agricultural and urban land uses).

TMDLs. Bioassessments have been used primarily as targets for TMDL monitoring in California rather than as direct biological endpoints. The endpoints in most TMDLs are primarily water quality endpoints rather than biological endpoints. However, the translation of bioassessment results to relevant TMDL endpoints is a next step in increasing the programmatic uses of bioassessment information in California. It will also enhance the comprehensiveness, relevancy, and applicability of TMDLs by focusing on the most limiting factors beyond the expected impact of individual pollutants by also highlighting their associated interactions and co-occurring stressors such as habitat and land use.

Part 2. Critical Elements Evaluation

The critical technical elements of bioassessment programs are described and divided into four general levels of rigor supported by a sliding scale of resolution and development (Barbour and Yoder 2007, 2008). A level 4 program is the most rigorous and the most capable of fully addressing the myriad of management issues regarding aquatic resources that are commonly faced by states and tribes. The remaining three levels of bioassessment rigor may be appropriate for some, but not all of the water quality management program support needs of state programs. Delineating the extent and severity of aquatic life impairments and diagnosing categorical and parameter-specific stressors are the primary tasks for a TALU based approach to monitoring and assessment that is intended to support multiple water quality management programs (Yoder and Barbour 2009).

A critical elements (CE) evaluation was conducted by proceeding through the CE checklist in accordance with the methodology in Barbour and Yoder (2007). The statewide program yielded a raw score of 53 out of the maximum possible score of 60 which equates to a mid-level 3 program; the two Regional board programs that were also evaluated were borderline level 3. The results for each element are discussed below (See Appendix 2 for checklist):

1. Index Period. *An index period is a consistent seasonal time frame for sampling the assemblage that is a cost-effective alternative to sampling on a year-round basis to account for seasonal variations. Ideally, the optimal index period corresponds to recruitment cycles of the organisms (based on reproduction, emergence, growth, and migration patterns). Sampling during an index period minimizes between-year variability*.

The statewide program adheres to a standardized index period (April to October) that slides from north to south to reflect differences in temperature. In southern California the index period is from April to October for the multimetric index (April to June for the O/E models); in northern California the index period is generally from August to September. Most Regional Boards adhere to this but there is some accommodation to support program needs. A CE score of 4.0 out of 4.5 was given to the California program.

2. Spatial Coverage. *Available resources and the desired outcome of the sampling design are key determinants in achieving adequate coverage.*

The “universe” of monitoring and assessment needs in California is spatially extensive and diverse. The nine regional boards incorporate a wide diversity of hydrological, landscape, and natural regional strata. No single design can meet all the State of California’s monitoring objectives.

SWAMP is using a probabilistic sampling design to obtain unbiased estimates of the biological conditions of perennial streams across the state and to track trends in biological conditions over time. The design of the SWAMP Perennial Stream Assessment (PSA) survey is cost effective because the entire resource need not be sampled – only a representative sample of streams. Another advantage of the probability-based design is that it allows the coordination/integration of other probability-based designs. In California the perennial stream survey is being coordinated with national stream assessments, regional watershed assessments being performed by Regional Boards Southern California (i.e., RB4, RB8 and RB9) and includes significant contributions from the regulated community including the Stormwater Monitoring Coalition in southern California and the Regional Monitoring Program in the San Francisco Bay area. The principal spatial designs include a statewide probabilistic network consisting of approximately 100 sampling sites per year, stratified by 6 ecological regions.

Many Regional Boards use targeted monitoring designs. These might involve watershed scale designs that include a resolution at an 8-11 digit HUC spatial scale to meet their specific needs. The designs vary from upstream/downstream sampling to bottom-of-watershed monitoring designs to more distributed networks. Some Regional Boards are using a rotating watershed approach, with a goal of sampling all watersheds in a region within a fixed time period (5 years is a common goal). The actual numbers of targeted sites are dependent on regional funding levels and annual monitoring priorities. Measurements include core chemical, physical, and biological parameters per the statewide SWAMP methodology with supplemental parameters added based on region-specific needs. The results from the statewide SWAMP perennial stream surveys provide context for local sampling.

The combined score of 4.0 reflects the practical integration of the statewide (which includes a combination of probability and targeted sites) and partial integration of Regional Board programs into the overall State Board effort.

3. Natural Classification. *In developing a bioassessment program, USEPA recommends classifying waterbodies more specifically than simply by waterbody type (e.g., river, lake, etc.), because it is highly unlikely that the biological condition of any given waterbody type is uniform throughout any anthropogenically-defined boundary. The classification of waterbodies is useful* *in partitioning natural variability and distinguishing it from variability resulting from human-induced changes. Classification of waterbodies can be based on a combination of characteristics, i.e., watershed drainage size, ecological regions, elevation, temperature, and other physical features of the landscape and/or waterbody for each waterbody type (e.g., large rivers, wadeable streams, headwater streams). The number of sites sampled and the availability of candidate reference sites within each class may limit the number of classifications*.

The challenge for the SWAMP program is to develop a program accounts for biological variation caused by natural environmental gradient and balances statewide consistency with the flexibility to adapt to California's diverse regional settings. In the present scheme, California will be divided into different geographic regions based on coarse biogeographic similarities in order to partition some of the natural variability among regions. These boundaries are consistent with those being used in the SWAMP perennial stream survey. Within the biogeographic classification, additional factors such as watershed size, elevation, and precipitation may be used to define biological expectations.

The CE score of 3.5 will be elevated to 5.0 with the developments that are already underway.

4. Criteria for Reference Sites. *A reference site should be natural or minimally disturbed while maintaining essential attributes. When reference sites are used to establish reference conditions, the State needs to document how it selects reference sites (by what criteria) and how it uses them to define regional reference conditions. Factors to be considered in selecting reference sites include human population density and distribution, road density, and the proportion of mining, logging, agriculture, urbanization, grazing, or other land uses. Candidate reference sites are evaluated for these factors to determine the degree of human modification that has occurred. Sites are eliminated if they have undergone direct human modification.*

The SWAMP strategy for selecting and sampling reference sites is documented in its Reference Condition Management Plan (RCMP, Ode and Schiff, 2008 In Prep)". The SWAMP RCMP program has proposed a general strategy for identifying reference sites. California will be classified into broad biogeographic regions. A pool of reference sites will be assembled within each region through a sequential process of identification and screening of candidate sites. This pool of reference sites will be managed through an iterative review of data to refine regional boundaries, ensure continued stability of sites and ensure adequate representation of natural gradients. Finally a monitoring design will be created for sampling this pool of reference sites to document the range of biological and physical condition at reference sites, and to monitor changes to this condition over time.

Screening of candidate sites will be done primarily through a combination of evaluation of existing data, GIS techniques, expert knowledge and site visits. It is recognized that high quality reference sites may not exist in certain areas of the state such as the agriculturally dominated Central Valley or the intensely urbanized southern California coastal plain. An alternate model for site selection will be used in these cases.

The score of 5.0 for the statewide program reflects the high degree of development of reference site selection criteria and procedures. These criteria and procedures are likely to be refined as the RCMP is implemented.

5. Reference Conditions. *The issue of reference conditions is critical to the interpretation of biological data. Generally, USEPA recommends the use of a regional reference condition based on an aggregate of sites that allows for broader application in State water resource programs than site-specific conditions. There must be a sufficient number of reference sites to capture regional stratification and the range of natural variations in biological assemblages due to geology, climate, and other natural physicochemical differences. Ideally, reference conditions represent the highest biological conditions found in waterbodies undisturbed by anthropogenic stressors. Recognizing that pristine habitats are rare or non-existent, resource managers must decide on an acceptable level of disturbance to represent an attainable or existing reference condition. Reference condition can be derived from reference sites, an empirical model of expectations that may include knowledge of historical conditions, or a model extrapolated from ecological principles. Usually, data from sites that represent best attainable conditions (i.e., least disturbed) of a waterbody are used.*

The SWAMP plan for development of reference conditions is embodied in the RCMP (Ode and Schiff 2008). Currently, reference condition is being determined from a still growing network of 300+ "least impacted" reference sites (1998-present). The reference site plan envisages sampling at 50-75 sites/year. The design includes ecoregional stratification and representation of the full range of regionally important natural gradients (e.g., elevation, precipitation, etc.). Development of regional reference condition is in progress – not yet completed for all regions. The goal is to have 50 sites per region.

The CE score of 3.5 should improve to 4.0 with the addition of regional reference sites that are being established as part of the ongoing improvements above.

6. Taxonomic Resolution. *An assemblage is defined as an association of interacting populations of organisms in a given waterbody. Although a single assemblage may be sufficient to make an attainment determination, USEPA recommends the use of at least two to enhance confidence in the assessment findings (USEPA 1996) because each assemblage serves a different function in the aquatic community and may be susceptible to stress in varying manners and degrees. Taxonomic identification of each assemblage to genus or species level provides reliable information about sensitivity, tolerance, and ecological/environmental relationships. Genus/species identifications improve assessments using richness values or metrics as key endpoints. Identification to family level requires less expertise to perform and usually speeds up the assessment process.*

For macroinvertebrate taxonomic identifications, the SWAMP program has recommended resolution to genus/species for development datasets; scoring tools are usually calibrated to work with genus level identifications. To ensure consistency and rigor in taxonomic data, SWAMP provides primary support for the Southwest Association of Freshwater Invertebrate Taxonomists (SAFIT), which establishes and maintains taxonomic standards.

SWAMP should also support the activities of an algal taxonomic workgroup, similar to SAFIT, to develop a standard algal taxonomic effort as recommended by Fetscher and McLaughlin (2008). SWAMP is currently leveraging the efforts and expertise of its partners to develop algal indices in southern California and the central coast. In both these efforts, soft algae and diatoms are currently analyzed to the lowest practicable taxonomy (usually genus/species), but recommendations for level of taxonomy for general assessment purposes are pending the results of the index development process.

The CE score of 4.5 reflects the full development of the macroinvertebrate assemblage and the in progress development of a periphyton indicator. Reaching the CE score of 5.0 is contingent on the full development and use of a second assemblage.

7. Sample Collection. *Standardization of field methods is necessary to establish the validity and reliability of biological data used in an assessment. Thorough training of investigators, coupled with rigorous certification processes, enhances the ability to provide a consistent unit of effort. Strong oversight of activities and leadership of apprentice professionals are critical. Standardization is especially important when information will be used in later trend analysis. The development of standard operating procedures (SOPs) for field and laboratory methods must include an effective quality assurance (QA) program with QC checks.*

The SWAMP program has developed a statewide protocol for macroinvertebrate sampling and physical habitat characterization that is derived from the EPA's national EMAP protocols (Ode 2007b). The SWAMP bioassessment group will work closely with the SWAMP QA Officer to develop comprehensive Quality Assurance Oversight Plan for quality assurance and quality control of bioassessment data. This guidance will cover personnel qualifications, training and field audit procedures, procedures for documenting sources of field and lab (including taxonomic data) error, procedures for chain of custody documentation, requirements for measurement precision, health and safety warnings, cautions (actions that would result in instrument damage or compromised samples), and interferences (consequences of not following the standard operating procedure). As most of the SWAMP sampling is performed by the California Department of Fish and Game, the procedures for quality assurance and quality control are currently addressed in the Quality Assurance Project Plan.

The SWAMP program is currently sampling periphyton using procedures developed for the EMAP program. However methods for field and laboratory protocols for algal sampling, identification and quantification used by various agencies have not been standardized across the state. The recently drafted SWAMP Algae Plan (Fetscher and McLaughlin, 2008) details key considerations for algae-based bioassessments, including the need to standardize sample collection and taxonomic methods across the state.

The CE score of 5.0 reflects the full development of the macroinvertebrate and partial development of the periphyton assemblage methodologies for the statewide and regional programs.

8. Sample Processing. *A systematic treatment of samples is needed to ensure the greatest extent of accuracy and precision. A strong QA/QC program is desired to ensure that (1) sample sorting procedures are being followed and no organisms are missed in the sample, and (2) the taxonomy is consistent and accurate.*

The CE score of 5.0 out of 5.0 for the statewide program reflects the full development of sample processing procedures for macroinvertebrates (Ode, 2008). The State also has a plan to develop standard statewide sample processing methods for periphyton (Fetscher and McLaughlin, 2008).

9. Data Management. *A reliable, efficient and quality assured database management system is fundamental to a program's ability to use monitoring information effectively to solve environmental problems. A proper system for aggregating data and performing the necessary quality control checks is essential. Furthermore, integration of assessment information from multiple assemblages (fish, macroinvertebrate, algae, etc) can contribute important diagnostic information. Data management includes not only proper stewardship of raw data elements but also proper computation of biological metrics and biocriteria threshold information. A strong geographic information system (GIS) linked to a well-designed relational database moves programs toward a more comprehensive watershed perspective in interpreting monitoring data and improves the ability of biological data to meet the increasing information demands of State and federal programs, responsible parties, and the public.*

The SWAMP 2.5 database is a relational database that encompasses all SWAMP monitoring data and links to a large distributed network of state and federal monitoring data (CEDEN). New bioassessment modules for entering, storing and reporting bioassessment data are nearly complete. Future work includes the development of tools to facilitate QA/QC procedures, summarize physical habitat data, and calculate bioassessment metric and IBI calculations. The CE score of 4.5 for the statewide program can be improved to 5.0 once the current data management system includes all reporting fields and calculation routines.

10. Ecological Attributes. *Ecological attributes are those aspects of an aquatic assemblage or community that correspond to the structure and function of that assemblage or community for a given condition. EPA has suggested 10 primary ecological attributes that form a continuum of responses to human disturbance (USEPA 2005). Ten primary ecological attributes have been identified as the basis for evaluating the BCG (USEPA 2005; Davies and Jackson 2006). The first six attributes relate to taxonomic identity, composition, and tolerances. They are 1) historically documented, sensitive, long-lived, or regionally endemic taxa, 2) sensitive rare taxa, 3) sensitive ubiquitous taxa, 4) taxa of intermediate tolerance, 5) tolerant taxa, and 6) non-native taxa that tend to displace endemic taxa. The seventh attribute is organism condition, which provides* *information on individual health. The remaining three attributes are functional integrity, ecosystem connectance, and spatial and temporal extent of stressors.*

The SWAMP program has developed several regional macroinvertebrate MMIs that use ecological attribute metrics in their calibration. SWAMP will continue to refine ecological attribute characterizations as it completes/ revises future MMIs. The State is also in the early stages of developing periphyton indices for coastal stream and has developed a plan for the use of periphyton in stream assessments (Fetscher and McLaughlin 2008).

The CE score of 4.0 out of 4.5 should increase with the development of the macroinvertebrate MMI and O/E model for all bioregions and the addition of a second assemblage.

11. Biological Endpoints & Thresholds. *State bioassessment programs should implement index development and threshold selection. Numerous methods are available for analyzing biological indicator data to assess attainment status, including both univariate and multivariate analysis techniques. Thresholds are the benchmarks from which the biological condition needed to support designated uses are described Selecting this threshold is perhaps the most critical aspect in reporting and documenting attainment status.*

Multimetric indices for macroinvertebrate data have been developed for perennial streams in the North Coast (Rehn and Ode, 2005), for perennial streams in Southern California (Ode et al., 2005) and for perennial streams in the Sierra Nevada (Herbst and Silldorff 2008, Rehn 2007). The State is also using a set of three predictive models based on the River Invertebrate Prediction and Classification System (RIVPACs), which compares the list of taxa observed at a site (O) to the list of taxa expected (E) to occur at a given site in the absence of human disturbance. The statewide California RIVPACs models (C. Hawkins unpublished) incorporates geographic coordinates (latitude and longitude), watershed area, average precipitation, average temperature and percent sedimentary geology into its predictions.

The SWAMP program uses statistical criteria to generate impairment thresholds. In the case of the northern and southern coastal IBIs, thresholds separating impaired from non-impaired were set at 2 standard deviations below a mean reference score. For the RIVPACS scores categorization to into "Good", "Poor" and "Very Poor used thresholds of 1.5 and 3 standard deviations below an O/E score of 1.0 (the score expected under no impairment).

The State Board is funding projects in Southern California and the Central Coast to develop periphyton indices. The products from these two studies are expected in 2009. The State is currently testing the use of the California Rapid Assessment Methodology (CRAM) for assessment of riparian habitat. As with the macroinvertebrate scores, it is likely that threshold values for these indices will be derived statistically from reference populations.

The CE score of 3.5 out of 4 will improve with the full development of the macroinvertebrate MMI and O/E models, a second assemblage, and the derivation of appropriately detailed numeric biocriteria

12. Diagnostic Capability. *The diagnostic capacity of bioassessment data and results is dependent on the development of patterns and response signatures from a database that includes a variety of stressors and the full gradient of human disturbance and biological response. This increases the value of biological data beyond the determination of status (attainment/non-attainment) to include inferences and decisions about causal associations and elimination of candidate causes in a stressor identification process. The development and use of a diagnostic capability is only possible within programs that have specifically developed methods and for which precision and accuracy issues have been addressed.*

The SWAMP and the NPS program have made some tentative steps in this direction. With funding from the NPS program the perennial stream survey (formerly known as CMAP) was modified to investigate associations between bioassessment scores and land use using associative techniques such as relative risk assessment. The NPS program also funded research to associate benthic invertebrate assemblages with land use (e.g., agricultural, forested and urban land uses). SWAMP has also funded the development of stressor specific tolerance values for benthic macroinvertebrates. The SWAMP bioassessment program receives a score of 2.5 out of a possible 4.0. For perspective, this score is similar to that of other states that have been reviewed.

13. Professional Review and Documentation. *Subjecting documented methods and assessment reports to a rigorous peer review is ultimately the best way to ensure an agency's credible data and scientific underpinnings. Inherent in a review is that it is conducted in an objective and independent manner (outside the agency and with no vested interest in the outcome) by technical and policy experts able to provide valid critique and suggestions, and recommendations for improvement and refinement are taken in good faith. Validation of standard operating procedures for all aspects of the assessment and monitoring program by outside experts is an initial step in establishing confidence in the resulting data. Programs that do not address and implement critical recommendations fail to benefit from an independent endorsement of their procedures and assessments.*

The SWAMP has a solid peer-review process for evaluating individual technical studies and reports. The overall SWAMP program underwent a technical review in 2005 (SPARC, 2006). There was a review of the bioassessment program in 2003 (Barbour and Hill, 2003) and this critical elements review also serves as peer review. The program receives a CE score of 4.5 out of 4.5.

Summary

The SWAMP bioassessment program is presently operating a high quality program at the state level and in selected regions. The information that we gathered and reviewed shows that the program operates at level 3 and is appropriate for 305(b) assessments and to support 303(d) listings. Ongoing development activities will eventually result in a level 4 program capable of being used more rigorously in regulatory decisions in perhaps 4-5 years.

Improvements that are planned or already underway will directly affect 9 elements and increase the CE score by 5-7 points resulting in a level 4 program for both the statewide and regional programs. Achieving a L4 program is contingent on the (1) full development and use of a second assemblage, (2) developing more detailed diagnostic capabilities, (3) improving data management and (4) developing the capacity of the other regional boards and linking regional monitoring to statewide efforts. This will take time to accomplish, perhaps 4-5 years depending on the rate of progress, resources devoted to the developmental effort, etc. Making these improvements should lead to an improved delineation of condition along the BCG and an improved diagnostic capability via an increased capacity to detect biological responses to specific types of stressors, provided that adequate and concurrent data about relevant stressors are also collected and analyzed.

The consistent addition of a second assemblage in the bioassessment process is needed to elevate the program to level 4. Three commonly used bioassessment assemblages (benthic macroinvertebrates, algae and fish) all provide unique perspective on the biological condition of a stream and its watershed. To be clear we advocate the use of *a minimum* of two assemblages in a given stream or river, but recommend that all three be available to choose from as each is applicable. The decision about which assemblage(s) to use in a particular situation should be made from all perspectives in addition to the obvious logistical and resource related perspectives. SWAMP has made strong progress toward developing algal indicators as a second indicator, but options for a third indicator are still under consideration. The use of fish indicators in CA is complicated by the State's limited fish fauna and may be not be a cost-effective indicator, but this should be explored further because fish can provide information about larger scale ecological condition (e.g., watershed connectivity, loss of spawning and other habitats, impacts of introduced species, etc.) that other assemblages cannot. Alternately, riverine wetland tools (e.g., CRAM) currently being explored by SWAMP may provide a means of partially filling the need for larger scale context.

We recommend that a follow-up CE review be conducted when these decisions are being made and upon the implementation of the improvements that are more immediately attainable. We would recommend in this case that new assemblages be developed and applied alongside macroinvertebrates based on the resource and management issues at hand.

The integration of the bioassessment results with chemical/physical data and other stressor information that is already included in the SWAMP will lead to a better understanding how human disturbance influences measurable biological response and lead to better support for all water quality management programs. Case examples of how this can be accomplished are available in the EPA TALU document (U.S. EPA 2005). Finally, these improvements will enable California to more fully develop a TALU (Tiered Aquatic Life Use) framework that will improve its current WQS and enhance the utility of aquatic life designated use classes for regulatory and other management applications.

Part 3. Moving from Bioassessment to Biocriteria

California's bioassessment program has made great strides in recent years due primarily to investments made by the Dept. of Fish and Game's Aquatic Bioassessment Laboratory (DFG-ABL) and the State Water Board's Surface Water Ambient Monitoring Program (SWAMP). California's bioassessment program is currently at a fairly high level (Level 3) and is being used within the recommended scope of that level to support 305(b) assessments and 303(d) listing. Continued investment and active management support will be needed to achieve a fully functional (Level 4) program that will support other regulatory needs and at relevant spatial scales of implementation.

It is clear from the extensive and well organized documentation that was provided before and during the review that California's scientists have a solid conceptual understanding of the steps required to reach the end goal of numerical biological criteria in the state's WQS in order to provide support for all relevant water quality management programs. It is equally clear that there is inadequate management support and commitment to achieve timely implementation of biocriteria in California.

While the DFG-ABL, SWAMP and their contractors are building a solid technical foundation for a robust freshwater bioassessment program, they can only provide the technical tools for developing biological endpoints and Tiered Aquatic Life Uses (TALUs). The State and Regional Water Boards need additional biologists and planning staff to develop, refine and implement TALUs as envisioned by the USEPA. This review affirms the findings of past peer reviews that the State Water Board needs its own in-house bioassessment coordinator and staff.

Managers at the State Water Board should be aware that the SWAMP program is, with continued management support, capable of building, maintaining and refining the technical tools that the Water Boards will need to incorporate biology into their water quality programs. Implementation of these tools including the development of TALUs, biocriteria and biological endpoints for TMDLs will be a fundamental paradigm shift that will require the detailed involvement of qualified biologists and planning staff. The Water Board's SWAMP program cannot be expected to fulfill those planning and implementation roles. Following U.S. EPA directives and the examples set by many other states, managers at the Water Boards should seek to provide the resources that are necessary to implement the technical bioassessment tools being developed by SWAMP.

As a first step, and consistent with the prior external peer review of SWAMP's bioassessment program (*see* Barbour and Hill 2003), the State Water Board should strive to create and maintain a specialist position for a state-wide bioassessment policy coordinator. The State Water Board needs a high-level in-house bioassessment policy coordinator to shepherd the implementation of the technical tools currently being developed by SWAMP into regulatory framework that is biocriteria and TALU. As the program develops, the State Board should create a team of staff that will work with the coordinator to integrate bioassessment/ biocriteria into the State's water regulatory programs.

Bioassessment toBiocriteria

SWAMP

Standardized biological protocols

↓

Classify water bodies into similar groups or classes

↓

Identify reference sites in each class

↓

Conduct bioassessments at reference sites in each class

↓

Develop Assessment Tool

STANDARDS

↓

Develop Biocriteria for each Aquatic Life Use

↓

Apply Biocriteria to all Water Bodies

Figure 2. Schematic framework for moving from bioassessment to biocriteria n California.

1. Refine Beneficial Uses. Use refinement is a broad term that encompasses any activity undertaken by a state to review and revise the designated uses applied to its waters. A state may refine its designated uses by revising the language defining what it intends to protect with this particular designated use or by revising a designated use by adopting more refined or specific designated uses in its place.

As recommended by the NRC (2001) and the SPARC (2006), the Water Board should consider refining beneficial uses relating to aquatic life use support. Generic beneficial use designations such as cold water habitat (COLD) simply do not account for the natural variability in rivers and streams across broad biogeographic regions. Cold water habitat in the North Coast is clearly different than cold water habitat in southern California. The State Board should develop a structure for examining the existing structure of designated uses to determine what parts, if any, will need to be changed or refined. This should be consistent with the principles and structure of the Biological Condition Gradient (BCG; U.S. EPA 2005; Davies and Jackson 2006).

The State Board should consider subclassifications of waterbodies in their use refinement process. Subclassifications based on similarities in the natural conditions of the waters could be established from major flowing water classes (such as large rivers, perennial stream, intermittent streams and ephemeral streams) or ecoregions (areas of biogeographic similarity) or a combination of these.

The State Board should support regional efforts to develop tiers of aquatic life uses and expand these efforts statewide. Tiers are subdivisions within subclasses of water based on similarities in the history of anthropogenic disturbance, the resulting biological condition, and the recovery potential within a tier (Figure x). Tiering of uses based on potential for recovery would also provide a framework for use attainability analyses. We advocate that UAAs be developed carefully and from the perspective of achieving the highest potential for each waterbody. It is tempting to plunge into a UAA process prematurely as a way to resolve impaired waters listings in the short-term, but we recommend that this be reserved for a time when the biocriteria and TALUs are sufficiently developed.

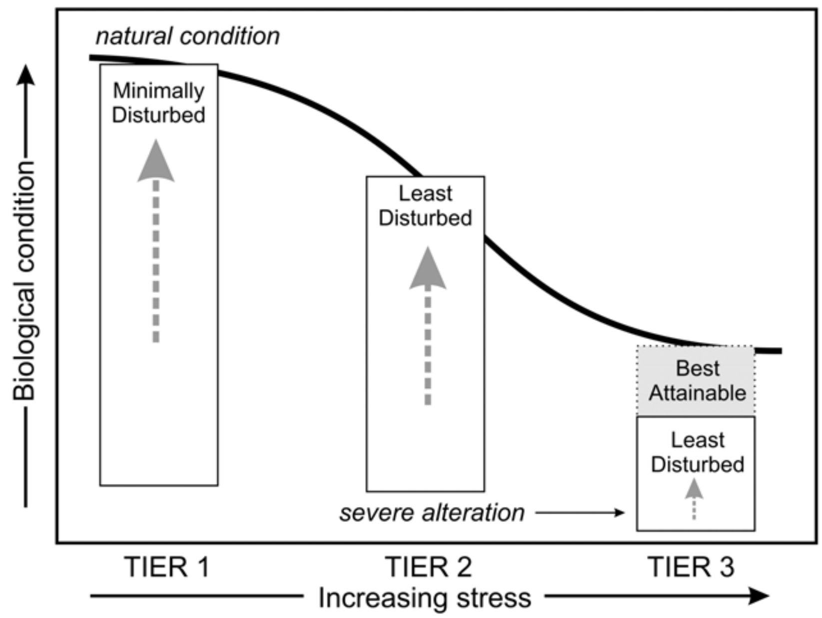


Figure 3. The biological condition gradient (BCG) used to define stream condition tiers in the TALU framework. Boxes indicate the expected range of biological condition scores at sites within each tier. Figure modified from Stoddard et al. 2006.

2. Develop Biological Objectives. The Water Board should develop statewide narrative biological objectives (biocriteria) to protect beneficial uses in Basin Plans that are associated with aquatic life use support. This should not preclude efforts by Regional Water Boards to develop biocriteria. However, many Regional Water Boards lack tools for interpreting existing narrative objectives in their Basin Plans. Currently, bioassessment data are used by Regional Water Boards in an "informal" manner where the assessments are used to support attainment decisions, but they lack any formal linkage to a designated aquatic life use. This lack of formal regulatory linkage to beneficial uses will limit the fuller use and true potential of bioassessment as a regulatory tool.

Biocriteria should be developed at both the State Board and Regional Board levels. However, development of numeric biocriteria will need to proceed in a series of phases. A key first step is the development of a statewide narrative objective that would set a common framework for the development and application of bioassessment tools to beneficial use protection. The interim step of developing statewide narrative biocriteria following the model set forth by Oregon Department of Environmental Quality (ODEQ) is likely to be an effective first step in California.

ODEQ's Statewide Narrative Criterion

Waters of the State shall be of sufficient quality to support aquatic species without detrimental changes in the resident biological communities.

Without detrimental changes in the resident biological communities means no loss of ecological integrity when compared to natural conditions at an appropriate reference site or region.

Ecological integrity means the summation of chemical, physical, and biological integrity capable of supporting and maintaining a balanced, integrated adaptive community of organisms having a species composition, diversity, and functional organization

Numeric biological criteria could then be achieved with the addition of defining language that pertains to the subclassification of different types of streams and rivers, ecotype specificity, biogeographical regions, and the level of protection afforded by tiered uses. It may be possible to use the predicted taxa list generated by the RIVPACs model to help identify highest attainable use for perennial streams across the state.

3. Develop Implementation Plan for Narrative Criteria: Biological criteria may appear to be more complicated to implement than traditional water quality criteria, but mostly because they achieve a congruence with natural factors that chemical criteria can not. A plan should be written which describes the technical components of the biocriteria (i.e., how to interpret biological data) as well as the policy components of the biocriteria (i.e., how they are to be used in programs. Technical tools and training will be necessary for staff, permittees and the general public. Policies will need to be developed regarding use of biocriteria in 305(b) assessments, 303(d) listings, NPDES monitoring, compliance and enforcement and in TMDLs. The State Water Board needs a high-level in-house bioassessment policy coordinator in order to shepherd the implementation of the technical tools currently being developed by SWAMP.

Part 4. Summary Conclusions

The State Board monitoring and assessment program is presently operating a high quality bioassessment program at the state level and in selected regions. The information that we gathered and reviewed shows that the statewide program operates at level 3 (the two regional board programs are borderline L3), and that the ongoing development activities will eventually result in a level 4 program in perhaps 4 - 5 years. These developmental tasks are one and the same as those that are necessary for developing biocriteria within a TALU framework. Hence this developmental process should deliver the technical capacity to support full TALU implementation.

SWAMP has and is making very effective use of their current resources to develop bioassessment tools which will support water quality programs (Figure 1). This means that SWAMP is positioned to provide data and information for more than general status assessments as required by Sections 303d and 305b, but to all Water Board programs including NPDES, NPS and TMDLs. These programs rely on monitoring and assessment information to provide an accurate and complete delineation of waterbody impairments and their associated causes. SWAMP data can also provide measures of the overall environmental outcomes produced by Water Board programs.

It is clear from examples in other States (Rankin 2003; U.S. EPA 2005) that a TALU based program will be a direct benefit to the California WQS, TMDL, NPDES permitting, and other water quality management programs (Figure 4). A TALU based approach would result in more refined aquatic life use designations that are appropriate to various water body types throughout the state. It would also lead to more specific biological objectives that are tailored to protect aquatic life in these waterbodies.

These biological objectives could be used to support listing and delisting decisions made by the Regional Boards. The tiered objectives can be used by Water Board programs to establish incremental goals for improvement for impaired waters. The objectives can also be used by the Water Boards to identify the high quality waters in the state and serve as backstops to ensure that these high quality waters are not degraded

The SWAMP program is developing a white paper to outline the technical infrastructure elements and identify current and future research needs to support bioassessments in California. A second white paper is being developed to identify the programmatic and policy issues that are necessary to move from bioassessments to biocriteria and TALU. These would provide the framework for developing TALU in California. Both Maine and Ohio provide case histories that describe the evolution of each program’s WQS and monitoring and assessment program to the attainment of level 4. These case studies are included in the EPA TALU document (U.S. EPA 2005; Appendices A and B). In addition, states that are involved in detailed developmental projects (e.g., Minnesota) can also provide a measure of comparability via their experiences.

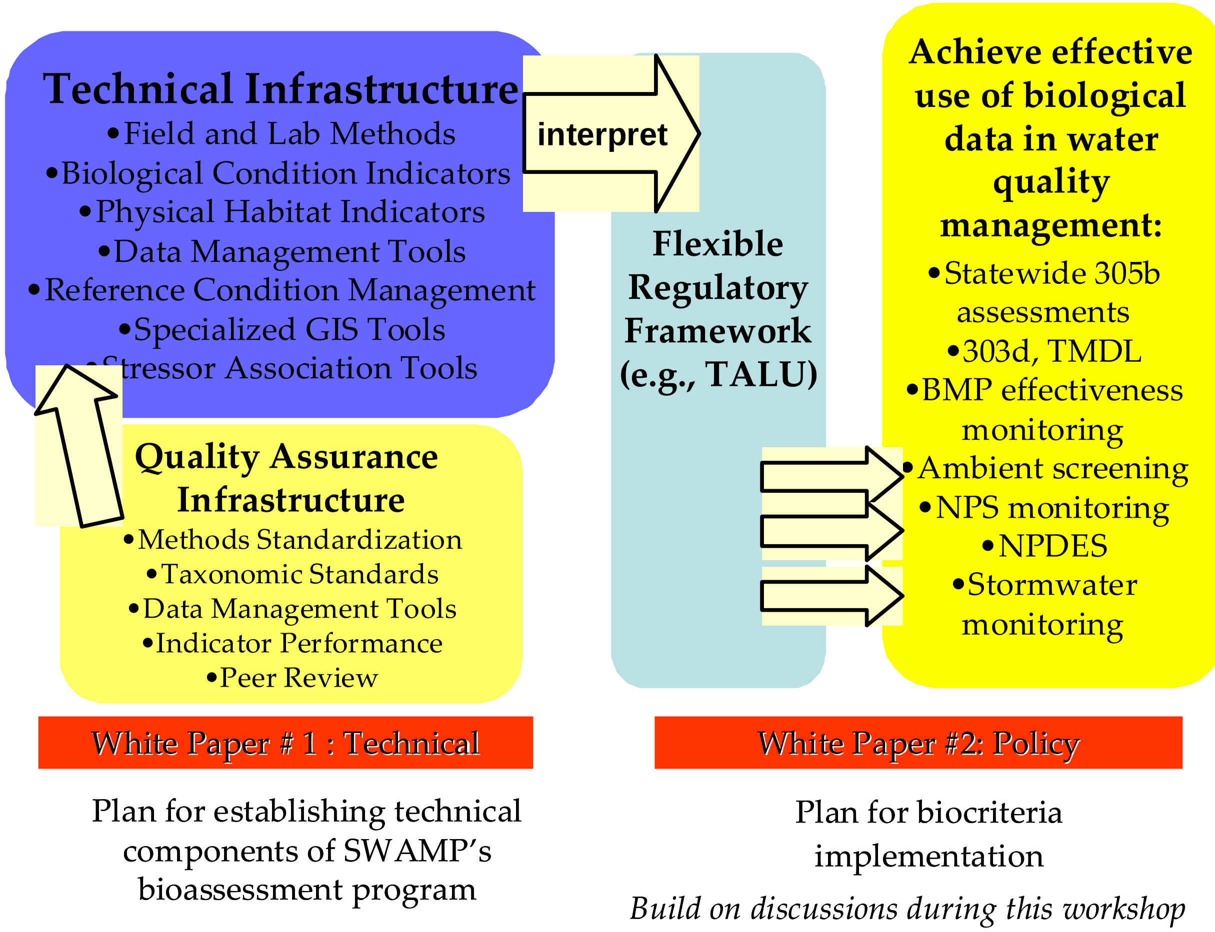


Figure 4. Process being used by the State Water Board to develop the technical infrastructure needed to use biological assessments and biocriteria within a TALU framework to provide full water quality management program support.

It is recognized here that this evaluation is a first step towards identifying the specific actions and needs of the California program to attain a level 4 program and achieve the support role for all management programs that is envisioned by the TALU process (U.S. EPA 2005; Barbour and Yoder 2007). Chapter 5 of the EPA TALU document describes the general milestones that a state program can use to gauge their own progress. This is now amplified in the 2008 update of the Critical Elements document using an active state development process as a working example (Barbour and Yoder 2008). The State Board should consider using the framework outlined in Figure 5 below to determine their existing position. This would accomplish an "inventory" of the existing program and determine what components are "TALU ready" and which areas are in need of further development and in which priority. Once this is done, a specific plan and timeline can be developed. At this time, we would estimate at least 5+ years to accomplish the tasks associated with full TALU development, but some aspects could be done more quickly if given a higher priority.

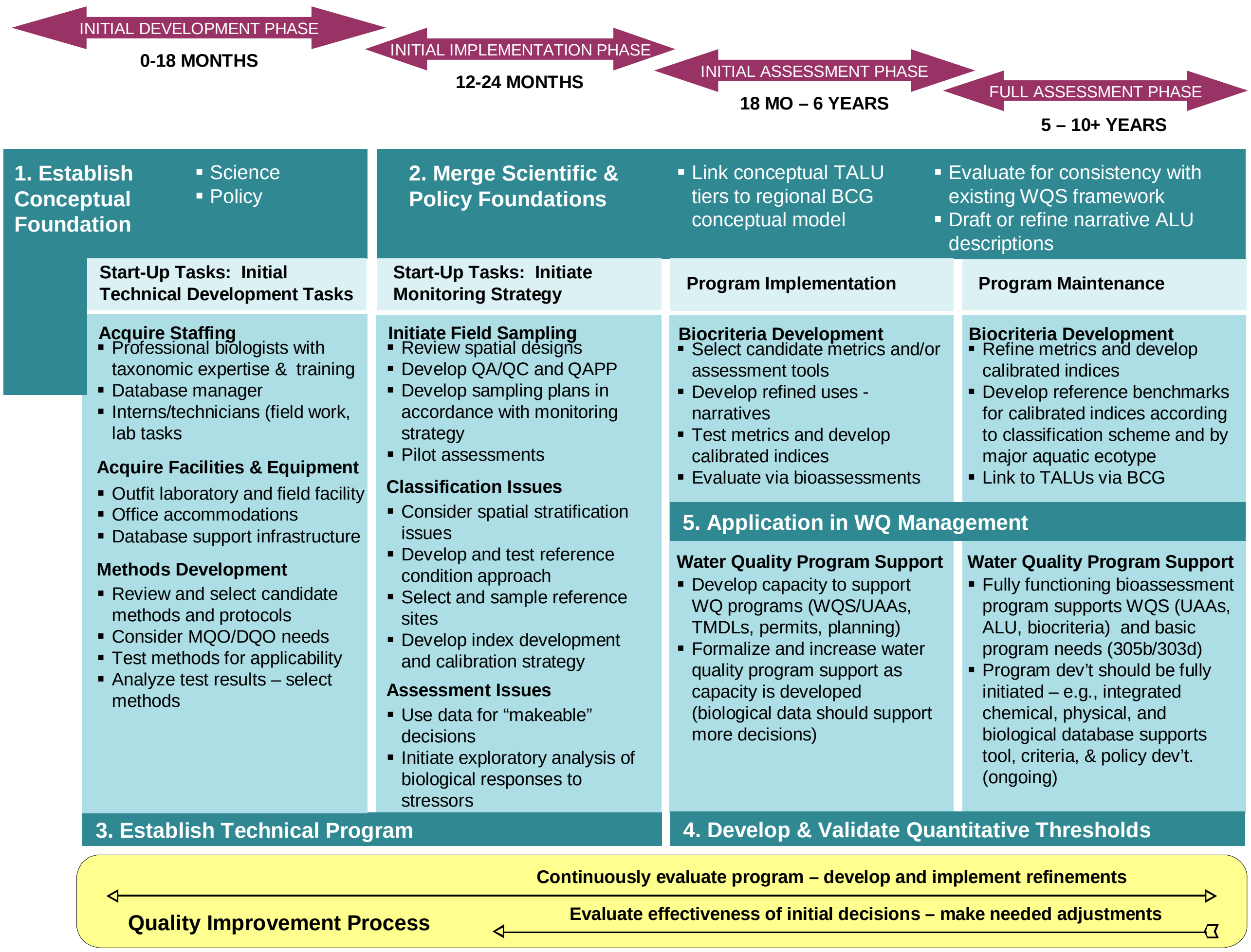


Figure 5. Hypothetical timeline for moving from bioassessment to biocriteria (U.S. EPA 2005).

The review of the California WRCB monitoring and assessment and WQS programs and the Critical Technical Elements results can be used to identify the specific technical and programmatic aspects that are in need of further development, refinement, and/or additional resources to accomplish full TALU program development. This review process is an essential component of the implementation process as generalized in *"Use of Biological Information to Better define Designated Aquatic Life Uses in State and Tribal Water Quality Standards: Tiered Aquatic Life Uses (August 2005)"* (U.S. EPA 2005). This includes general guidance and case examples for developing and implementing a TALU-based approach to monitoring and assessment and water quality standards (WQS) by States and Tribes. It contains a hypothetical timeline (Figure 5) that describes a sequence of steps including the development of a baseline bioassessment program (already in place via SWAMP), initial support for baseline management programs (partially in place and in selected Water Boards), development of narrative and numeric biocriteria (concept in place), increasingly sophisticated support for all relevant water quality management programs (yet to be accomplished), and long term maintenance of the program (the result of full TALU program development and implementation). The ultimate goal is the adoption of numeric biocriteria and Tiered Aquatic Life Uses (TALU) in the California WQS.

This template provides a framework within which the State can first determine where their program is along the timeline in Figure 5.

We expect that California will be positioned “somewhere” along the TALU timeline once a detailed exercise is undertaken to inventory the existing program. The “position” along the timeline is determined by first conducting a baseline review of the state programs and its technical elements, which is represented by this memorandum. The development of a full TALU program could take several years if a State or Tribe is starting from “scratch”. However, it is likely that States and Tribes already operate at least a basic program (i.e., Level 2; Yoder and Barbour 2009) and will likely determine that the time for implementing a more refined program consistent with Level 4 is considerably less than the 10 years depicted in Figure 5. Based on the information garnered by this baseline review we expect that the development of the bioassessment program via SWAMP and select Region Boards will show California to be further along this timeline than most states given the Level 3 status of the current program. We do recommend that this be done considering the unique roles of the statewide SWAMP program and the Regional Board programs in TALU implementation.

We recommend that the next step for California is to use this process to determine “where” the program currently stands and what tasks are yet to be accomplished to reach the above stated program goals. This process is a prerequisite to producing a detailed plan for the eventual development and adoption of TALU based narrative and numeric biocriteria in the California WQS, supported by a Level 4 program. The example in the latest draft of the Critical Technical Elements (Barbour and Yoder 2008) represents a working example of how California can use the results of the baseline program review and CE process to develop a “blueprint” for making orderly improvements and attaining full TALU status. This will include a mix of technical and policy development tasks.

REFERENCES:

Barbour, M.T., and C. O. Yoder. 2008. Critical Technical Elements of a Bioassessment Program. USEPA, Office of Water, Washington, DC. (September 2008 Draft). 75 pp.

Barbour, M.T., and C. O. Yoder. 2007. Critical Technical Elements of a Bioassessment Program. USEPA, Office of Water, Washington, DC. (November 2007 Draft). 71 pp.

Barbour, M.T. and C. Hill. 2003. The Status and Future of Biological Assessment for California Streams. California State Water Resources Control Board, Division of Water Quality, Final Report. 52 pp. + appendices.

Davies, S.P. and S.K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16 (4): 1251-1266 + appendices.

Fetscher, A.E. and K. McLaughlin. 2008. Incorporating Bioassessments Using Freshwater Algae into California's Surface Water Ambient Monitoring Program. Prepared for the State Water Resources Control Board by Southern California Coastal Water Research Project Tech Report 563.

Herbst, D.B. and E.L. Silldorff. 2006. Development of an index of biological integrity (IBI) for stream assessment in the Lahontan region of California. Draft report to the Lahontan Regional Water Quality Control Board.

National Research Council. 2001. Assessing the TMDL Approach to Water Quality Management. Committee to Assess the Scientific Basis of the Total Maximum Daily Load Approach to Water Pollution Reduction, Water Science and Technology Board, National Research Council. ISBN: 0-309-07579-3, 122 pp. Download at: www.nap.edu/catalog/10146.html

Ode, P.R., A.C. Rehn and J.T. May. 2005. A quantitative tool for assessing the integrity of southern coastal California streams. Environmental Management 35(4): 493-504.

Ode, P.R. 2007a. Ecological condition assessment of California's perennial wadeable streams (2000 through 2006). Report to the State Water Resources Control Board's Non-Point Source Program.

Ode, P.R.. 2007b. Standard Operating Procedures for Collecting Benthic Macroinvertebrate Samples and Associated Physical and Chemical Data for Ambient Bioassessments in California. California State Water Resources Control Board, Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 001.

Ode, P.R. and K. Schiff. In Prep. Recommendations for the Development and Maintenance of a Reference Condition Management Program to Support Biological Assessment of California's Wadeable Streams. Prepared for the State Water Resources Control Board's Surface Water Ambient Monitoring Program (SWAMP).

Quasney, C.T. and C.O. Yoder. 2008. Bioassessment Technical Evaluation: A method for assessing quality and level of scientific rigor of an agency's biological monitoring and criteria. EPA/XXX/R-XX/XXX. U.S. EPA, Office of Water. 70 pp.

Rankin, E.T. 2003. Comparison of Biological-based and Water Chemistry-based Aquatic Life Attainment/Impairment Measures under a Tiered Aquatic Life Use System. Center for Applied Bioassessment and Biocriteria Fact Sheet 3-CABB-03. Midwest Biodiversity Institute, Columbus, OH. 28 pp.

Rehn, A.C. 2006. Development of stressor-specific tolerance values for western benthic macroinvertebrates. Draft Report prepared for California State Water Resources Control Board, Surface Water Ambient Monitoring Program (SWAMP) Bioassessment

State Water Resources Control Board. 2007. 2006 CWA Section 303(d) List of Water Quality Limited Segments Requiring TMDLs.

State Water Resources Control Board. 2007 SWAMP FY06-07 Bioassessment Workplan. (January 2007).

State Water Resources Control Board. 2006. Water Quality Assessment of the Condition of California Coastal Waters and Wadeable Streams. Clean Water Act Section 305b Report 2006. (October 2006).

State Water Resources Control Board. 2005. Comprehensive Monitoring and Assessment Strategy to Protect and Restore California's Water Quality. Water Boards, State Water Resources Control Board and Regional Water Quality Control Boards. Surface Water Ambient Monitoring Program (SWAMP). 62 pp. + appendices.

State Water Resources Control Board. 2005. Comprehensive Monitoring and Assessment Strategy to Protect and Restore California's Water Quality, Surface Water Ambient Monitoring Program (SWAMP). State Water Resources Control Board and Regional Water Quality Control Boards, Sacramento, CA. 62 pp. + appendices (October 2005).

State Water Resources Control Board. 2004. Water Quality Control Policy for Developing California's Clean Water Act Section 303(d) List. California EPA, Water Boards. State Water Resources Control Board and Regional Water Quality Control Boards. 28 pp. + appendices.

State Water Resources Control Board. 2003. The Status and Future of Biological Assessment for California Streams. Final Report to the California State Water Resources Control Board, Division of Water Quality. (January 2003)

State Water Resources Control Board. 2002. Quality Assurance Management Plan for the State of California's Surface Water Ambient Monitoring Program (SWAMP). Prepared for California State Water Resources Control Board, Division of Water Quality, Sacramento, CA. First version (December 2002).

State Water Resources Control Board. 2000. Legislative report: Summary of Information from Surface Water Monitoring Programs. Appendix to the Plan for Implementing a Comprehensive Program for Monitoring Ambient Surface and Ground Water Quality. (January 2000).

SPARC. 2006. Review of California's Surface Water Ambient Monitoring Program (SWAMP). Technical Report 486, Southern California Coastal Water Research Project. 25 pp.

U.S. EPA. 2005. Use of biological information to better define designated aquatic life uses in state and tribal water quality standards. Office of Water, Washington, DC. EPA 822-R-05-001. 188 pp.

Yoder, C.O. and M.T. Barbour. 2009. Critical technical elements of state bioassessment programs: a process to evaluate program rigor and comparability. Environ. Mon. Assess. DOI 10.1007/s10661-008-0671-1.

Yoder, C.O. 1998. Important concepts and elements of an adequate state watershed monitoring and assessment program, pp. 615-628. *in* Proceedings of the NWQMC National

Appendix Table 1. List of Participants

Emilie Reyes, (SWAMP Chief - State Water Board)	ereyes@waterboards.ca.gov	(916) 341-5556
Dawit Tadesse (SWAMP Unit – State Water Board)	dtadesse@waterboards.ca.gov	(916) 341-5486
Bruce Fujimoto (Stormwater – State Water Board)	bfujimoto@waterboards.ca.gov	(916) 341-5523
Ken Harris (TMDL Program – State Water Board)	kharris@waterboards.ca.gov	(916) 341-5500
Toni Russell (SWAMP Unit – State Water Board)	trussel@waterboards.ca.gov	(916) 322-2578
Joanne Cox (TMDL Coordinator – State Water Board)	jcox@waterboards.ca.gov	(916) 341-5552
Steve Fagundes (Nonpoint Source – State Water Board)	sfagundes@waterboards.ca.gov	(916) 341-5487
Melenee Emanuel (Nonpoint Source – State Water Board)	memanuel@waterboards.ca.gov	(916) 341-5271
George Nichol (SWAMP Unit – State Water Board)	gnichol@waterboards.ca.gov	(916) 341-5504
Rik Rasmussen (Fw Stds./Planning – State Water Board)	rrasmussen@waterboards.ca.gov	(916) 341-5549
Dominic Gregario (Oceans Stds. – State Water Board)	dgregario@waterboards.ca.gov	(916) 341-5488
Val Connor (Regulatory Section – State Water Board)	vconnor@waterboards.ca.gov	(916) 341-5573
Tom Suk (Regional Water Board-Region 6, Lahontan)	tsuk@waterboards.ca.gov	(530) 542-5419
Dave Gibson (Water Board-Region 9, San Diego)	dgibson@waterboards.ca.gov	(858) 467-4387
Pete Ode (CA DFG-ABL)	pode@ospr.dfg.ca.gov	(916) 358-0316
Terry Fleming (U.S. EPA Region IX)	Fleming.Terrence@epa.gov	(415) 972-3462
Rob Plotnikoff, Tetratech (EPA TALU Team)	robert.plotnikoff@tetratech.com	(206) 728-9655
Chris Yoder, MBI (EPA TALU Team)	yoder@rrohio.com	(614) 403-9592

Appendix Table 2. A checklist for evaluating the degree of development for each technical element of a bioassessment program and associated comments on the elements for the California WRCB bioassessment program (both SWAMP and applicable Regional Boards). The point scale for each element ranges from lowest to highest resolution (na – not applicable).

Element 1	(Lowest) 1.5	2.0 2.5	3.0 3.5	4.0 4.5 (Highest)	Comments
Index Period	Collection times are variable throughout the year, and sampling is performed without regard to seasonal influences.	An index period is conceptually recognized, but sampling may take place outside of this period for convenience or to match existing programs; sampling outside of the index is not adjusted for seasonal influences.	A well-documented seasonal index period(s) is calibrated with data for reference conditions, but sampling may take place outside of this period for convenience or to match existing programs; sampling outside of the index is adjusted for seasonal influences. Index periods are selected based on known ecology to minimize natural variability, maximize gear efficiency, and maximize the information gained about the assemblage.	Same as Level 3, but administrative needs and index periods fully reconciled. Scientific basis of temporal sampling influences management decision framework.	April-October seasonal index period that "slides" from south to north: SoCal – April to early June; NoCal – August to September; most regional boards adhere to this, but some do not to accommodate program support needs.
					Points **Statewide: 4.5** **Regional: 4.0**

Element 2	(Lowest) 1.5	2.0 2.5	3.0 3.5	4.0 4.5 (Highest)	Comments
Spatial Coverage	An individual site is used for assessment of watershed condition; simple upstream/ downstream and fixed station designs prevail; assessments at local scale.	Multiple sites are used for watershed assessment; spatial coverage only for questions of general status or locally specific problem areas; synoptic (non-random) design at coarse scale (e.g., 8-digit HUC common); spatial extrapolation is based on "rules of thumb"; may be supplemented by simple upstream/downstream assessments.	Spatial network suitable for status assessments; statewide spatial design using rotating basins with single purpose design at coarse scale (e.g., 8 digit HUC); may be supplemented by occasional intensive surveys.	Comprehensive spatial network suitable for reliable watershed assessments in support of multiple water quality management programs at more detailed scale (e.g., 11-14 digit HUC); statewide rotating basin approach or similar scheme to complete statewide monitoring in a specified period of time; multiple spatial designs appropriate for multiple issues.	Statewide probability design (WEMAP) and "pour point" integrator sites at 8 digit HUC scale; Regional boards employ watershed scale intensive survey designs at HUC 11 scale (currently in 4 of 9 regions).
					Points **Statewide: 3.5** **Regional: 4.0** **Combined: 4.0**

Appendix Table 2. (continued)

Element 3	(Lowest) 2.0 / 2.5	3.0 / 3.5	4.0 / 4.5	5.0 (Highest)	Comments
Natural Classification	No partitioning of natural variability in aquatic ecosystems. Minimal classification limited to individual watersheds or basins with generalized stratification on a regional basis; does not incorporate differences in stream characteristics such as size, gradient.	Classification recognizes one stratum, usually a geographical or other similar organization such as fishery based cold or warmwater, and is applied statewide; lacks other intra-regional strata such as watershed size, gradient, elevation, temperature, etc.	Classification is based on a combination of landscape features and physical habitat structure (inter-regional); achieves highest level of classification possible by considering all relevant intra-regional strata and subcategories of specific stream types.	Fully partitioned and stratified classification scheme based on a true regional approach that transcends jurisdictional (i.e., State) boundaries to strengthen inter-regional classification and recognizes zoogeographical aspects of assemblages.	Classification includes intra-regional factors such as watershed size, elevation, and other stratifying factors; not yet developed for all bioregions. **Points** **Statewide: 3.5** **Regional: na**

Element 4	(Lowest) 2.0 / 2.5	3.0 / 3.5	4.0 / 4.5	5.0 (Highest)	Comments
Criteria for Reference Sites	No criteria, except informal BPJ selection of control sites. May be little documentation and supporting rationale.	Based on "best biology", i.e., BPJ on what the best biology is in the best waterbody; minimal non-biological data used.	Non-biological criteria supported by narrative descriptors only; combine BPJ with narrative description of land use and site characteristics; may use chemical and physical data thresholds as primary filters.	Quantitative descriptors used to support non-biological criteria; characteristics of sites are such that the best biological organization expected to be supported; chemical and physical characteristics of sites used only as secondary and tertiary filters to avoid circularity in other criteria.	A quantitative procedure for screening reference sites is used; **Points** **Statewide: 5.0** **Regional: na**

Appendix Table 2. (continued)

Element 5	(Lowest) 1.0	1.5 2.0	2.5 3.0	3.5 4.0 (Highest)	Comments
Reference Conditions	No reference condition; presence and absence of key taxa or best professional judgment. rather than established reference conditions may constitute the basis for assessment.	Reference condition based on biology of a 'best' site or waterbody; a site-specific control or paired watershed approach may be used for assessment; regional reference sites lacking.	Reference conditions based on site-specific data, but are used in watershed scale assessments; regional reference sites are conceptually recognized, but are too few in number and/or spatial density to support the deviation of biocriteria.	Applicable regional reference conditions are established within the applicable waterbody ecotypes and aquatic resource classes; consist of multiple sites that either represent reference or are along the BCG in such a manner to allow extrapolation of expected conditions for assessing and monitoring within waterbody ecotype. Re-sampling of reference sites done systematically over a period of years.	Development of regional reference condition is in progress – not yet completed for all regions. **Points** **Statewide: 3.5** **Regional: na**

Element 6	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Taxonomic Resolution	Gross observation of biota; single assemblage only; very low taxonomic resolution (e.g., order/family level for macro-invertebrates.; family for fish by non-biologists).	Single assemblage (usually macroinvertebrates); low taxonomic resolution (e.g., family level) by experienced biologists.	Single assemblage with high taxonomic resolution (e.g., "lowest practical" i.e., genus/species); if multiple assemblages, others are lower resolution or infrequently used.	Two or more assemblages with high taxonomic resolution (e.g., "lowest practical" i.e., genus/species); capacity to use each assemblage concurrently is maintained; practitioners are certified in accordance with available offerings (e.g., NABS, state credible data provisions).	Statewide program employs lowest practicable taxonomy (usually genus/species); SoCal employs genus level; second assemblage (periphyton) is under development; fish may be used regionally. **Points** **Statewide: 4.5** **Regional: 4.5**

Appendix Table 2. (continued)

Element 7	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Sample Collection	Approach is cursory and relies on operator skill and BPJ, producing highly variable and less comparable results; Training limited to that which is conducted annually for non-biologists who compose the majority of the sampling crew. Documentation of methods more as an overview.	Textbook methods are used rather than in-house development of detail of SOPs to specify methods; a QA/QC document may have been prepared; training consists of short courses (1-2 days) and is provided for new staff and periodically for all staff.	Methods are evaluated and refined (if needed) for State purposes; detailed and well documented; SOPs are updated periodically and supported by in-house testing and development; a formal QA/QC program is in place with field replication taken; rigorous training is for all professional staff, regardless of skill mix to raise skill levels and enhance interaction and consistency.	Same as Level 3, but methods cover multiple assemblages.	Sample collection methods are fully developed for two assemblages (macroinvertebrates, periphyton); fish methods also exist in other agencies. **Points** **Statewide: 5.0** **Regional: 5.0**

Element 8	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Sample Processing	Biological samples are processed in the field using visual guides; sorting and identification are dependent on operator skill and effort.	Organisms are identified and enumerated primarily in the field prohibiting ample QC but done by trained staff; for fish cursory examination of presence and absence only; no in-house development of SOPs.	Laboratory processing of all samples (except for fish); A formal QA/QC program is in place; rigorous training is provided; vouchering of organisms done for ID verification.	Same as Level 3, but is applicable to multiple assemblages; subsampling level tested. Notations made on fish as to diseased, erosion, lesion, tumors.	Sample processing fully developed for statewide program and for two assemblages; some regional programs are not as well developed. **Points** **Statewide: 5.0** **Regional: 4.0**

Appendix Table 2. (continued)

Element 9	(Lowest) 2.0	2.5 3.0	3.5 4.0	4.5 5.0 (Highest)	Comments
Data Management	Sampling event data organized in a series of spreadsheets e.g., (by year, by data-type, etc); QC cursory and mostly for transcription errors.	Separate quasi-databases for physical-chemical and biological data (Excel, Access, dBase, etc) with separate GIS shape files of monitoring stations; data-handling methods manuals available; QC for data entry, value ranges, and site locations.	True relational database containing biological and sampled site info (Oracle, etc); fully documented and implemented data QAPP.	Relational database of bioassessment data (including indices and biocriteria) with real-time connection to spatial data coverage showing monitored sites in relation to other relevant spatial data layers (population density; impervious surfaces; vegetation coverage, low-flight photos, nutrient concentrations, ecoregion, etc); fully documented and implemented data QAPP; data available from multiple assemblages to enable integrated analysis.	**Points** **Statewide: 4.5** **Regional: 3.0**

Element 10	(Lowest) 1.5	2.0 2.5	3.0 3.5	4.0 4.5 (Highest)	Comments
Ecological Attributes	Linkage to the BCG or adherence to the basic ecological attributes as a foundation is lacking; simple measures of presence/absence.	Only inferences can be made for a few of the comparatively simple ecological attributes, e.g., sensitive/tolerant taxa of a ubiquitous nature; single dimension measures used.	Ecological attributes used as a foundation for bioassessment, but may not be fully developed, or may be lacking. BCG incorporated into conceptual underpinnings.	The ecological attributes of the BCG form the conceptual foundation; level of rigor represents or extends to all underpinnings of the ecological attributes.	Statewide O/E model and 3 regional MMIs have been developed; periphyton index under development. **Points** **Statewide: 4.0** **Regional: na**

Appendix Table 2. (continued)

Element 11	(Lowest) 1.0	1.5 2.0	2.5 3.0	3.5 4.0 (Highest)	Comments
Biological Endpoints and Thresholds	Assessment may be based only on presence or absence of targeted or key species; (Some citizen monitoring groups use this level); attainment thresholds not specified; this approach may be sufficient for Coarse problem identification. Coarse method (low signal) and detects only high and low values.	A biological index or endpoint is established for specific water bodies, but is likely not calibrated to waterbody classes or statewide application; index is probably relevant only to a single assemblage; presence/absence based on all taxa; BPJ thresholds based on single dimension attributes. Limited to pass/fail determinations of attainment status that does not reflect incremental measurement along the BCG.	A biological index, or model, has been developed and calibrated for use throughout the State or region for the various classes of a given waterbody type; the index is relevant to a single assemblage; attainment thresholds are based on discriminant model or distribution of candidate reference sites, or some means of quantifying reference condition. Can distinguish 3-4 increments along the BCG; supports narrative evaluations based on multimetric or multivariate analysis that are relevant to the BCG.	Biological index(es), or model(s) for multiple assemblages is (are) developed and calibrated for use throughout the State or region and corresponds to the BCG; integrated assessments using the multiple assemblages are possible, thus improving both the assessment and diagnostic aspects of the process; multiple parameters for evaluation, based on integrated data calibrated to regional reference condition. Able to detect status (integrated signal) on a continuous scale along the BCG; power to detect at least 5-6 categories of condition.	O/E model is statewide; MMI developed for selected regions; periphyton in development. **Points** **Statewide: 3.5** **Regional: na**

Element 12	(Lowest) 1.0	1.5 2.0	2.5 3.0	3.5 4.0 (Highest)	Comments
Diagnostic Capability	Diagnostic capability lacking.	Coarse indications of response via assemblage attributes at gross level, i.e., general indicator groups (e.g., EPT taxa); Supporting analysis across spatial and temporal scales limited.	More detailed development of indicator guilds and other aggregations to distinguish and support causal associations; usually involves refined taxonomy (at least genus level); supported by analysis of larger datasets and/or extensive case studies; patterns repeatable across different sources; developed for a single assemblage only.	Response patterns are most fully developed and supported by organized and extensive research and case studies across spatial and temporal scales; results are actively used in biological assessment and in assigning associated causes and sources for program support purposes; involves refined taxonomy; accomplished for two assemblage groups.	Baseline research to support diagnosis has not been completed; baseline database is being developed – need to assure full range of stress:response in statewide and regional datasets. **Points** **Statewide: 2.5** **Regional: na**

Appendix Table 2. (continued)

Element 13	(Lowest) 1.5	2.0 2.5	3.0 3.5	4.0 4.5 (Highest)	Comments
Professional Review and Documentation	Review limited to editorial aspects.	Internal scientific review only, Outside review for objectivity left for higher levels.	Outside review of documentation and reports conducted. However, selection of peer review can be subjective.	Formal process for technical review to include multiple reference and documented system for reconciliation of comments and issues. Process results in methods and reporting improvements. Can include peer-reviewed journal publications.	A formal process is in place and is used; methods and protocols are in the process of being published in journals. **Points** **Statewide: 4.5** **Regional: na**

Statewide CE Score = 53 (Regional = 50.5)
Statewide CE % = 88.3% (Regional = 84.2%)
Statewide Level = L3 [85-95%] (Regional Level = L2 [70-85%])

Appendix Table 3. Summary of the critical technical elements evaluation for the California WRCB statewide bioassessment program conducted January 23-25, 2008.

Element	Comment
Element 1: Index Period Maximum score = 4.5 Statewide = 4.5 Regional = 4.0	The statewide program adheres to a standardized index period that slides from north to south. The regional board score will improve to 4.5 once the standard permit boilerplate language developed by the Lahontan Region is standardized statewide.
Element 2: Spatial Coverage Maximum score = 4.5 Statewide = 3.5 Regional = 4.0 Combined = 4.0	The current score of 3.5 for the statewide program reflects the statewide probabilistic design and "pour point" design for integrator sites. Regional boards apply watershed scale designs that include a resolution at an 8-11 digit HUC spatial scale and other designs such as upstream/downstream sampling. The regional board score of 4.0 reflects the watershed design and rotating subbasin approach applied by some, but not all boards. The combined score of 4.0 reflects the practical integration of the statewide and regional board programs as a reflection of the overall WRCB effort. Attaining a score of 4.5 will be realized when the watershed design is applied by all of the regional boards.
Element 3: Natural Classification Maximum score = 5.0 Statewide = 3.5 Regional = na	The CE score of 3.5 will be elevated to 5.0 with the developments that are already underway including the inclusion of other bioregions (the na score for the regional boards reflects the relevancy of this element to a statewide setting).
Element 4: Criteria for Reference Sites Maximum score = 5.0 Statewide = 5.0 Regional = na	The score of 5.0 for the statewide program reflects the high degree of development of reference site selection criteria and procedures (the na score for the regional boards reflects the relevancy of this element to a statewide setting).
Element 5: Reference Conditions Maximum score = 4.0 Statewide = 3.5 Regional = na	The CE score of 3.5 should improve to 4.0 with the addition of regional reference sites that are being established as part of the ongoing improvements described for elements 3 and 4 (the na score for the regional boards reflects the relevancy of this element to a statewide setting).

Appendix Table 3. (continued).

Element	Comment
Element 6: Taxonomic Resolution Maximum score = 5.0 Statewide = 4.5 Regional = 4.5	The CE score of 4.5 reflects the full development of the macroinvertebrate assemblage and the in progress development of a periphyton indicator. Fish may be applicable in certain regions pending developments by USGS. Reaching the CE score of 5.0 is contingent on the full development and use of a second assemblage.
Element 7: Sample Collection Maximum score = 5.0 Statewide = 5.0 Regional = 5.0	The CE score of 5.0 reflects the full development of the macroinvertebrate and periphyton assemblage methodologies for the statewide and regional programs. Fish methods also exist n other agencies.
Element 8: Sample Processing Maximum score = 5.0 Statewide = 5.0 Regional = 4.0	The CE score of 5.0 for the statewide program reflects the full development of the macroinvertebrate and periphyton assemblage sample processing methods. The regional boards have the capacity to apply the macroinvertebrate assemblage. Reaching the CE score of 5.0 is contingent on the full use of a second assemblage by the regional boards.
Element 9: Data Management Maximum score = 5.0 Statewide = 4.5 Regional = 3.0	The CE score of 4.5 for the statewide program can be improved to 5.0 once the current data management system includes all reporting fields and calculation routines. The regional board score should likewise improve when their data is routinely uploaded to the statewide data management system.
Element 10: Ecological Attributes Maximum score = 4.5 Statewide = 4.0 Regional = na	The CE score of 4.0 should increase with the development of the macroinvertebrate MMI and O/E model for all bioregions and the addition of a second assemblage (the na score for the regional boards reflects the relevancy of this element to a statewide setting).
Element 11: Biological Endpoints & Thresholds Maximum score = 4.0 Statewide = 3.5 Regional = na	The CE score of 3.5 will improve with the full development of the macroinvertebrate MMI and O/E models, a second assemblage, and the derivation of appropriately detailed numeric biocriteria (the na score for the regional boards reflects the relevancy of this element to a statewide setting).

Appendix Table 3. (continued).

Element	Comment
Element 12: Diagnostic Capability Maximum score = 4.0 Statewide = 2.5 Regional = na	The comparatively low CE score of 2.5 is a common characteristic of bioassessment programs that are in development and/or which have singularly been focused on status assessments. Improving the score for this element will occur as a result of addressing preceding elements 2, 3, 6, 10, and 11 and gaining a familiarity with how diagnostic capacity is developed; a familiarity with the concepts involved is encouraging. This will require some dedication to exploratory analyses in which the response of the biological assemblages is evaluated along the stressor axis of the BCG.
Element 13: Professional Review Maximum score = 4.5 Statewide = 4.5 Regional = na	The CE score of 4.5 reflects a thorough and complete peer review process. Statewide methods and procedures are in the process of being published in refereed journals.

Appendix 11

B354

2015 Report on the SMC Regional Stream Survey

Special study on engineered channels

Program update

Preliminary results from new indicators

Applications of SMC data

Southern California Stormwater Monitoring Coalition
Regional Watershed Monitoring Program

Contents

Cover photo: Los Angeles River at the confluence of Calabasas and Bell Creek.

This report should be cited as:
Southern California Stormwater Monitoring Coalition. 2017. 2015 Report on the Stormwater Monitoring Coalition Regional Stream Survey. SCCWRP Technical Report 963. Southern California Coastal Water Research Project. Costa Mesa, CA.

Program update

Overview of a redesigned survey

In 2015, the SMC initiated the first year of its redesigned stream bioassessment survey, sampling 102 sites and implementing several major changes to address information gaps identified in the initial five-year survey, including:

- Inclusion of nonperennial streams in the survey. Whereas nonperennial streams were previously excluded from sampling, we now attempt to include them among the 55 “condition” sites (i.e., sites selected in a probabilistic way to represent the typical condition of streams in the region) where bioassessment occurs. By shifting the sampling period earlier in the season (starting as early as March), intermittent streams that dry up before May are more likely to be represented in the survey.
- Improved trend detection through site revisits. A total of 47 “trend” sites that were sampled in the first cycle of the survey were revisited in 2015. With a sufficient number of revisits, the survey will be able to determine the extent of stream-miles that are improving or degrading over time, and identify factors that are associated with these trends.
- A change in analytes and indicators measured at each site. In order to focus on new priorities and concerns, SMC participants sampled a number of new indicators (highlighted elsewhere in this report), such as hydromodification impact potential, aquatic invasive vertebrate occurrences, hydrologic state, cellular bioassays, and non-target analysis of chemicals of emerging concern. Assessment of sediment contamination, although part of the updated survey workplan, was deferred so that a pilot study in limited areas could be completed in 2016.

What is the Stormwater Monitoring Coalition (SMC)?

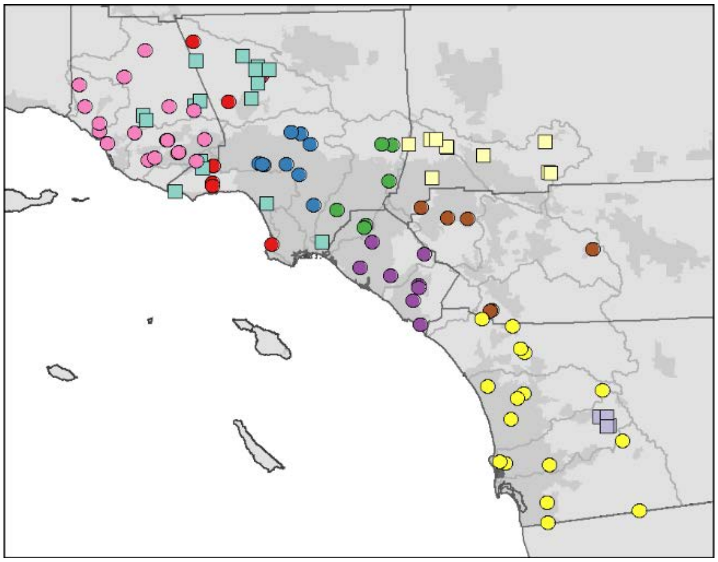
The SMC is a coalition of multiple state, federal, and local agencies that works collaboratively to improve the management of stormwater in Southern California. SMC members include regulatory agencies, flood control districts, and research agencies: County of Los Angeles Department of Public Works, County of Orange Public Works, County of San Diego Department of Public Works, Riverside County Flood Control and Water Conservation District, San Bernardino County Flood Control District, Ventura County Watershed Protection District, City of Long Beach Public Works Department, City of Los Angeles Department of Public Works, California Regional Water Quality Control Board—Santa Ana Region, Los Angeles Region, and San Diego Region, State Water Resources Control Boards, California Department of Transportation, Southern California Coastal Water Research Project (SCCWRP). In addition, the SMC collaborates with the U.S. Environmental Protection Agency Office of Research and Development. For more information, visit the SMC webpage at www.socalsmc.org.

The SMC has conducted a probabilistic survey of streams in the South Coast region since 2009. The goals of this survey are to provide the technical foundation for scientifically sound management of stormwater by answering three questions:

1. What is the biological condition of streams in the South Coast region?
2. What stressors are associated with streams in poor condition?
3. Are the conditions of streams changing over time?

The first five-year cycle of survey took place between 2009 and 2013. The results of the first cycle are summarized in a report available on SCCWRP’s website. The survey continues with a new cycle that spans from 2015 to 2019, evolving to address new questions. This report summarizes the current status of the survey and describes major developments and accomplishments that occurred in 2015. A comprehensive report will be released after completion of the fifth year of the current cycle.

Changes in cost from the first cycle were minimized, as certain indicators (i.e., toxicity, metals, and pyrethroids in the water column) were dropped based on recommendations by the SMC workgroup. Priority indicators that were retained, such as benthic macroinvertebrates, algae, riparian wetlands (i.e., CRAM), physical habitat, nutrients, and major ions, were sampled at every site.



Sampling effort in 2015 by agency.

Stormwater agencies	Condition (# sites)	Trend (# sites)	Total (# sites)
Ventura County	10	8	18
Los Angeles County	5	2	7
Los Angeles WMP	3	6	9
San Gabriel RMP	2	4	6
Orange County	5	3	8
Riverside County	3	3	6
San Diego WMAs	12	4	16
Water boards			
RB4	9	7	16
RB8	4	6	10
RB9	2	4	6
Total	55	47	102

San Diego Watershed Management Areas (black text) nested within SMC watersheds (brown text). Local jurisdictions take the lead in monitoring each WMA and setting management priorities, contributing to and making use of the SMC's regional survey.

New watershed-based permits enhance interactions with multiple agencies in San Diego County

Marking a major transition in the implementation of the SMC survey in San Diego County, smaller municipalities (including the cities of Oceanside, Encinitas, San Diego, and Imperial Beach) are now working directly alongside SMC member agencies to collect data for the survey, as opposed to working indirectly through San Diego County Public Works as a lead agency. This transition is intended to increase interaction between stormwater co-permittees and the San Diego Regional Water Quality Control Board, while also making the survey more useful to local managers. These municipalities contribute to the survey through coalitions focused within Watershed Management Areas (WMAs). The WMAs have the effect of spreading responsibility among the individual municipalities to fulfill the permit obligations. As a result, more municipalities are engaged with the regional monitoring program in supporting their management and regulatory needs.

The formation of WMAs began when the San Diego Regional Water Board consolidated municipal stormwater permits into a single regional stormwater permit. Whereas previously, all monitoring in San Diego County was coordinated through a single agency (i.e., the County of San Diego), each WMA coalition is now tasked with collecting data and identifying management priorities for its own WMAs. Survey data are used to develop a Water Quality Improvement Plan, or WQIP (see article on the San Juan WQIP below) for each WMA, with stakeholders responsible for identifying priority issues and associated stressors that each coalition should address. For watersheds that cross county borders (e.g., Santa Margarita), the WMAs facilitate cooperation among municipalities in the different jurisdictions.

Not only do the WMAs help the partners outside the SMC with the survey, but they also carry forward the SMC's vision of collaborative monitoring to the local level. Through minor adjustments to the SMC's sampling plan (e.g., allocating trend sites by watershed rather than by land use), combined with enhanced dialogue between permittees and the Regional Board, the new partners were able to acquire data for their own needs, as well as contribute to the regional assessment goals of the SMC survey.

B359

What are the biological conditions in engineered channels?

The SMC survey helps managers evaluate biological conditions in engineered channels and understand the potential policy implications.

Key Points

- Engineered channels surveyed to date are, generally speaking, in worse ecological health than natural channels based on biological indicators based on benthic macroinvertebrate and algae assemblages. These preliminary results suggest that tradeoffs between ecological health and flood protection may be unavoidable.
- While engineered channels invariably have poor scores for the California Stream Condition Index (CSCI) based on benthic macroinvertebrates, algal indices occasionally indicated better biological conditions—sometimes similar to reference condition. This wide range in index scores suggest that some engineered channels support more ecosystem functions than others.
- Within engineered channels, design and construction characteristics (e.g., armoring material or presence of low-flow features) did not influence index scores or other measures of ecological condition
- Within engineered channels, algal indices may reflect water quality conditions better than the macroinvertebrate index. For example, lower specific conductivity was associated with higher diatom index scores, but not CSCI scores. However, both types of indices have some capacity to respond to stressor gradients in these systems.
- Targeted sampling (particularly from hardened channels with good water quality, or engineered channels with high bioassessment index scores) and experimental studies may clarify the factors that support better ecological conditions.
- Survey data can provide a context for evaluating the biological condition of streams in engineered channels, thereby helping managers recognize factors, such as water quality or stream temperature, that may lead to better conditions.

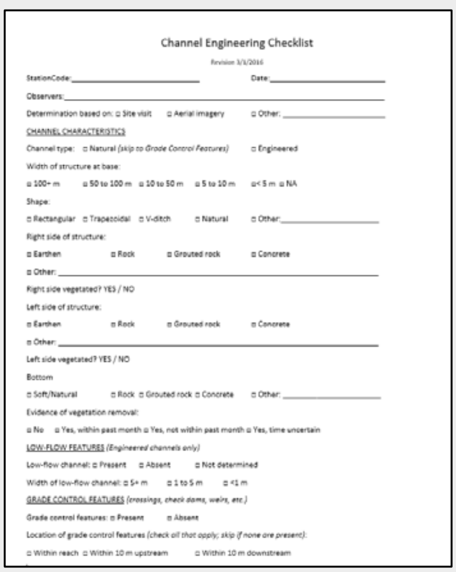
Engineered channels are common features in urban stormwater systems, protecting surrounding neighborhoods from floods that could damage property or endanger lives. However, this service often comes at a cost, as engineered channels do not provide the same quality habitat that natural stream channels provide. Additionally, engineered channels may reduce groundwater recharge, or degrade water quality through alterations of biochemical processes. Consequently, engineered channels often fail to support designated beneficial uses related to ecosystem health, such as those related to aquatic life or wildlife. Faced with these tradeoffs between competing uses, stormwater agencies and regulators encounter questions from stakeholders, such as what range of ecological conditions are possible in engineered channels? And what factors can be managed to support better conditions? The SMC stream survey provides a rich source of data to answer these questions. By developing methods to characterize engineered channels, analyzing bioassessment scores in different channel types, and exploring responses to water chemistry gradients, the SMC stream survey offers preliminary answers to these questions.

Bioassessment indices, such as the California Stream Condition Index (CSCI, based on benthic macroinvertebrate communities) and the Southern California algal Indices of Biotic Integrity (IBIs), are the key indicators used by the State and Regional Water Boards to assess attainment of aquatic life beneficial uses in streams. These indices will have a central role in the implementation of the State's bio-integrity policies; it is therefore necessary that stormwater managers understand how these indices work in engineered channels. Aquatic organisms have diverse life history traits with sensitivities to a wide range of stressors. As a result, bioassessment indices provide a holistic measure of the combined impacts of poor water quality, habitat alteration, hydrologic modification, and other disturbances. This integration allows assessment of cumulative and diverse impacts on ecosystem health. Three indices are sampled in the SMC program: the CSCI, a diatom index, and a soft algal index; each of these three indices provide an independent measure of a stream's ability to support aquatic life.

To assess the range of biological conditions in engineered channels, the SMC took advantage of the extensive bioassessment data collected by the survey since its inception in 2009. In prior years, the SMC collected benthic macroinvertebrates and algae samples at hundreds of sampling reaches across Southern California, many of which were in engineered channels. In order to make the use of these data, the SMC bioassessment workgroup developed a simple procedure for characterizing and classifying the different types of channels found in the region (Sidebar 1). The protocol was designed for rapid application in the field or in the office (if aerial imagery or other data are available). This ease of use meant that the SMC could generate a large data set from recent and older data that would support robust analyses on the features of engineered channels associated with variability in bioassessment scores. Elements of this protocol have been adopted by the Surface Water Ambient Monitoring Program (SWAMP), and resource managers throughout the state are looking to the SMC to provide guidance on how to evaluate engineered channels in their regions. These data will also be used in mapping and modeling efforts to determine locations of engineered channels in the landscape.

Characterizing engineered channels

Modification of stream channels takes many forms, exhibiting a variety of designs and constructions. To characterize the diversity of engineered channels, the SMC developed simple forms to record key features, like shape, material, size, and presence of low-flow channels. These forms were filled out during site visits for the 2015 sampling season, as well as for sites visited in earlier years (relying on aerial imagery, photographs, data from earlier field visits, and other sources of information). Elements of the SMC's approach for characterizing engineered channels have been incorporated into SWAMP's standard bioassessment protocols.



Channel Engineering Checklist

Revision 3/1/2016

StationCode:____________ Date:____________

Observers:____________

Determination based on: ☐ Site visit ☐ Aerial imagery ☐ Other: ____________

CHANNEL CHARACTERISTICS

Channel type: ☐ Natural *(skip to Grade Control Features)* ☐ Engineered

Width of structure at base:

☐ 100+ m ☐ 50 to 100 m ☐ 10 to 50 m ☐ 5 to 10 m ☐ < 5 m ☐ NA

Shape:

☐ Rectangular ☐ Trapezoidal ☐ V-ditch ☐ Natural ☐ Other:____________

Right side of structure:

☐ Earthen ☐ Rock ☐ Grouted rock ☐ Concrete

☐ Other: ____________

Right side vegetated? YES / NO

Left side of structure:

☐ Earthen ☐ Rock ☐ Grouted rock ☐ Concrete

☐ Other: ____________

Left side vegetated? YES / NO

Bottom

☐ Soft/Natural ☐ Rock ☐ Grouted rock ☐ Concrete ☐ Other: ____________

Evidence of vegetation removal:

☐ No ☐ Yes, within past month ☐ Yes, not within past month ☐ Yes, time uncertain

LOW-FLOW FEATURES *(Engineered channels only)*

Low-flow channel: ☐ Present ☐ Absent ☐ Not determined

Width of low-flow channel: ☐ 5+ m ☐ 1 to 5 m ☐ <1 m

GRADE CONTROL FEATURES *(crossings, check dams, weirs, etc.)*

Grade control features: ☐ Present ☐ Absent

Location of grade control features *(check all that apply; skip if none are present)*:

☐ Within reach ☐ Within 10 m upstream ☐ Within 10 m downstream

Forms developed by the SMC to characterize engineered channels are simple enough to complete within minutes during field visits, or from the office if aerial imagery and other data are available.

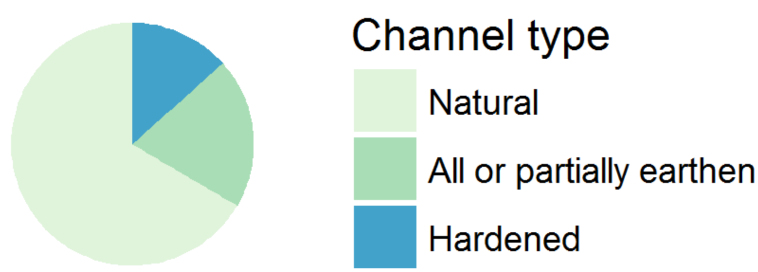


Figure 1. Proportion of stream types observed in the study

Armed with this protocol, the SMC bioassessment workgroup evaluated 724 unique bioassessment sites, with about 20% of these evaluations being made in the field. About two-thirds of the sites were natural, lacking any evident armoring, artificial structures (apart from road or bridge crossings), or straightening (Figure 1). Ninety-seven sites were entirely hardened, with concrete or grouted rock banks and a hardened bottom. The remaining 145 sites retained some earthen elements—typically a natural bottom, with earthen or partially armored banks. Because CSCI and algae IBI scores had already been calculated for these sites from the previous survey cycle, and because water chemistry and habitat quality measurements were also available, the data set was a good starting point for analyzing biological conditions in engineered channels.

Engineered channels are largely in poor condition, but some are in better condition than others

Nearly all engineered channels were in poor health, as measured by both the CSCI and the algal IBIs (Figure 2). Although a wide range of invertebrate CSCI scores was evident in engineered channels (inter-quartile range: 0.44 to 0.66), they rarely exceeded 0.79 (the threshold used in previous SMC reports to identify healthy streams similar to reference conditions). None of the entirely hardened channels met this benchmark, and only 14% of the earthen or partially hardened engineered channels did so. In contrast, 63% of the natural channels in the analysis met the healthy stream benchmark. Aquatic insect communities appear to be strongly affected by partial or complete channel hardening (see Sidebar 2).

Algal indices, however, provided different insights into stream condition. While the diatom and soft algae IBIs, like the CSCI, showed that engineered channels were generally in worse condition than natural channels, high algal IBI scores indicative of healthy (i.e., similar to reference) conditions were not uncommon. In fact, 43% of hardened channels had diatom IBI scores above the reference threshold, and 20% had high soft algae IBI scores. Whereas the CSCI indicated almost exclusively poor conditions in engineered channels, algal indicators suggest that engineered channels can support healthy streams under conducive conditions (such as good water quality).

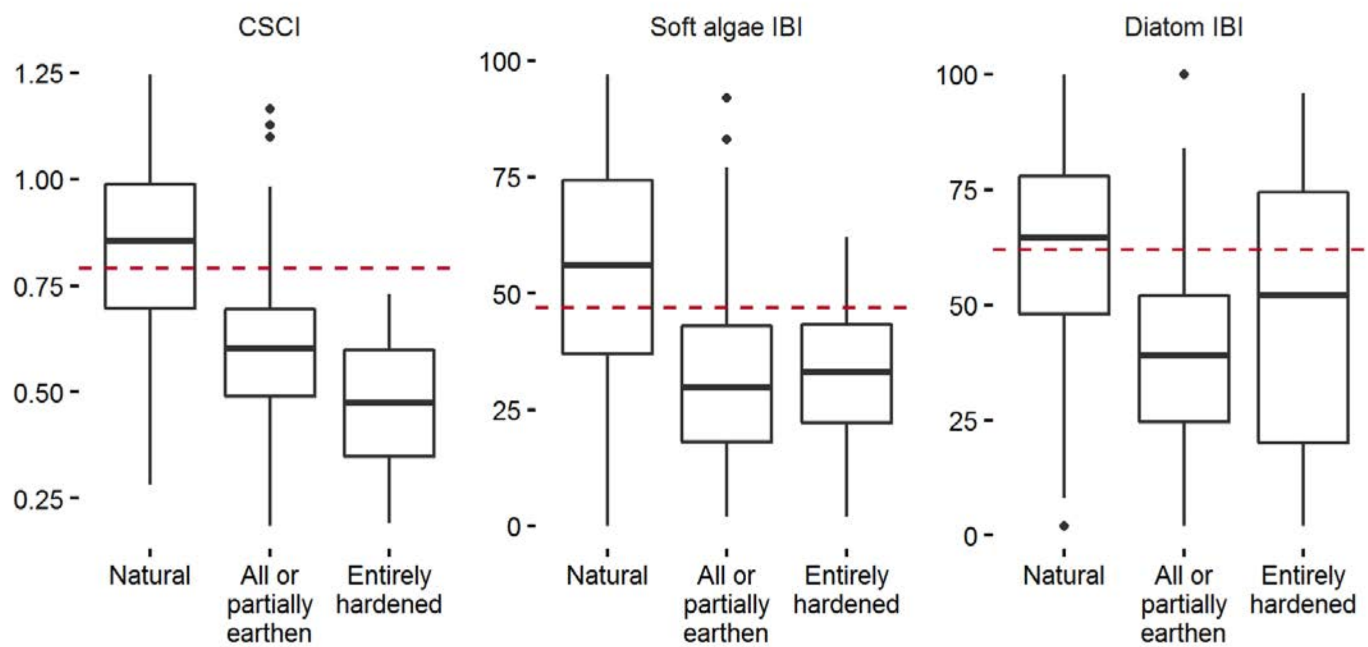


Figure 2. Bioassessment scores were typically higher in natural channels than in engineered channels. However, high scores for the algal indices were sometimes observed in engineered channels, occasionally exceeding the threshold for identifying sites in reference condition (red dashed line).

What kind of organisms are found in engineered channels?

Despite the poor in-stream ecological condition noted in this study, engineered channels do, in fact, support aquatic life, as well as terrestrial wildlife that depend on streams and rivers. Because of their accessibility and proximity to populated areas, engineered channels are frequently enjoyed for their wildlife-viewing opportunities, particularly for waterfowl and wading birds that forage in shallow areas. Although fish and amphibians are sometimes observed as well, these are almost exclusively non-native species, such as carp (*Cyprinus carpio*), tilapia (Cichlidae), bullhead (*Ameiurus*), and mosquito fish (*Gambusia affinis*).

Great blue herons and black-necked stilts forage on the concrete banks of the San Gabriel River.

Photo courtesy of Kerry Matz

***Dasyhelea*, a fly in the family of biting midges (*Ceratopogonidae*), are particularly common in hardened channels.**

The benthic macroinvertebrates found in engineered channels are only a small subset of the diversity of species found in the natural channels, typically with life history adaptations that provide resilience to disturbance (for example, rapid life-cycles with multiple generations per year, or tolerance to temperature extremes). A few invertebrate species show a particular affinity for engineered channels: Biting midges (*Dasyhelea*), soldier flies (*Euparyphus*), minnow mayflies (*Fallceon*), snails (*Physa*), worms (Oligochaeta), flatworms (Turbellaria), and seed shrimp (Ostracoda) were all more common than expected within hardened channels. Species that require complex substrates, such as those that burrow in the substrate (e.g., midges in the subfamily Tanypodinae) were less common than might be expected in a natural channel. Most sensitive and moderately tolerant species (e.g., net-spinning caddisflies, like *Hydropsyche*) were entirely eliminated. The abundance of tolerant species, and rarity of sensitive species, is reflected in the lower CSCI scores observed in engineered channels.

As with macroinvertebrates, algal assemblages within engineered channels contained subsets of species found in natural channels. Many planktonic diatoms, such as species in the *Scenedesmus* genus, were common, as well as the green filamentous algae *Cladophera glomerata*, found at nearly all concrete channels. These species are sometimes a concern. For example, *C. glomerata* form large, unsightly mats that trap debris, smother streambeds, and create odor problems.

The green alga *Cladophora glomerata* often proliferates in engineered channels, particularly if nutrient inputs are high and shading has been reduced.

What factors support higher bioassessment scores in engineered channels?

Why do some engineered channels score better than others? And why are high scores more common for algal IBIs than for the invertebrate CSCI? Design features (such as construction material or presence of a low-flow channel) had no discernible impact on either CSCI or algal IBI scores, so perhaps water quality or other habitat features were more important. That is, relatively high scores in engineered channels may indicate better water or habitat quality than lower scores.

Analyses of the data provide some support for this hypothesis. The diatom index responded to a range of water quality conditions, even within concrete channels (Figure 3, Table 1). For example, the diatom IBI declined with increasing specific conductivity in all channel types, whereas scores for the soft algae index and the CSCI exhibited responses within natural or partially earthen channels. The hypothesis that the constraints within engineered channels overwhelm the ability of bioassessment indicators to respond to stress is not well supported for diatoms.

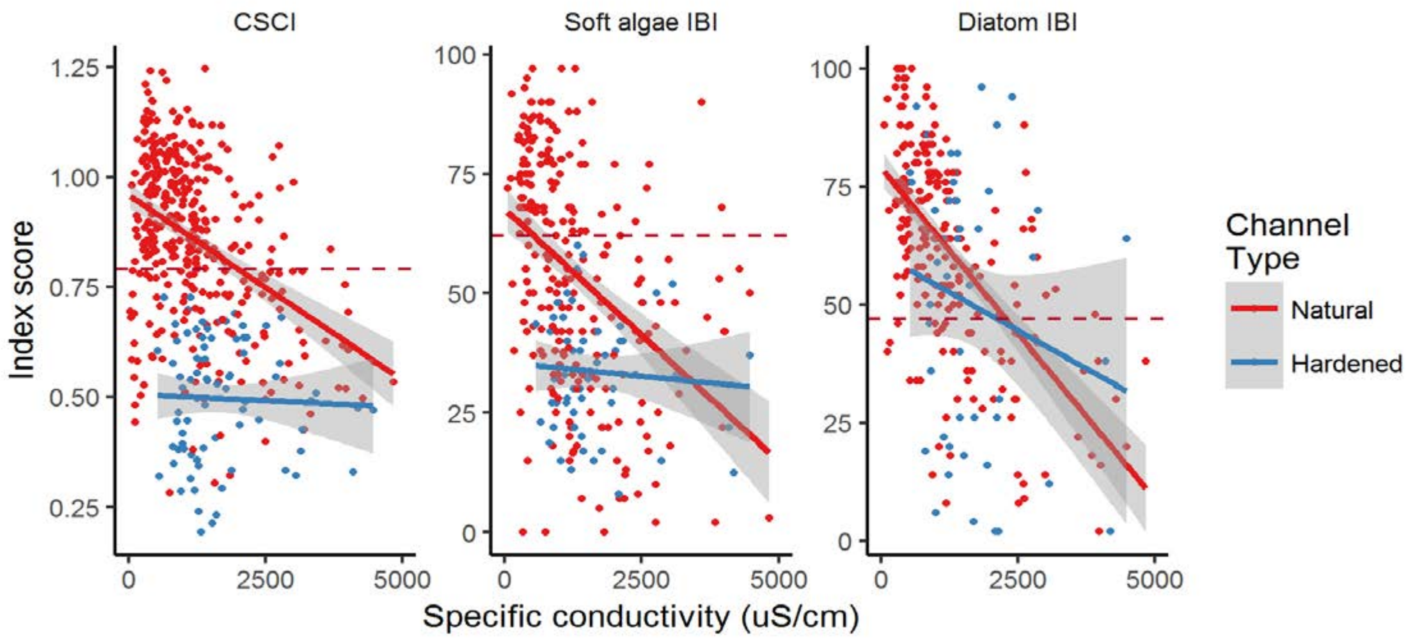


Figure 3. Specific conductivity versus bioassessment index scores in hardened and natural channels. The red dashed line is the threshold for sites in reference condition. For clarity, earthen and partially hardened channels are excluded from this plot.

Factors related to habitat showed a similar pattern of responses. For example, high levels of sands and fines in the streambed were associated with lower scores for all indices, but the relationships within hardened channels were strongest for the diatom index (Figure 4). Although the CSCI did not respond to many water chemistry and physical habitat gradients within hardened channels, shading and temperature appears to be important for this index, with higher scores observed in hardened channels where shading was high (Figure 5). Stream-side vegetation, which is often removed for flood control purposes, may provide the conditions that improve CSCI scores. However, shading had the opposite relationship with diatom IBI scores, and no relationship with soft algae IBI scores.

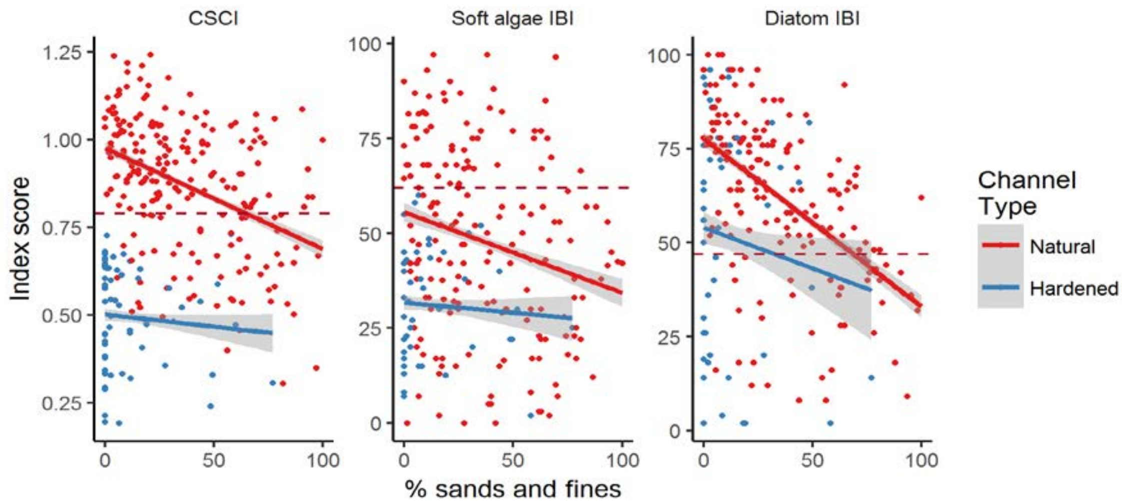


Figure 4. Percent sands and fines in the streambed versus bioassessment index scores in hardened and natural channels. The red dashed line is the threshold for sites in reference condition. For clarity, earthen and partially hardened channels are excluded from this plot.

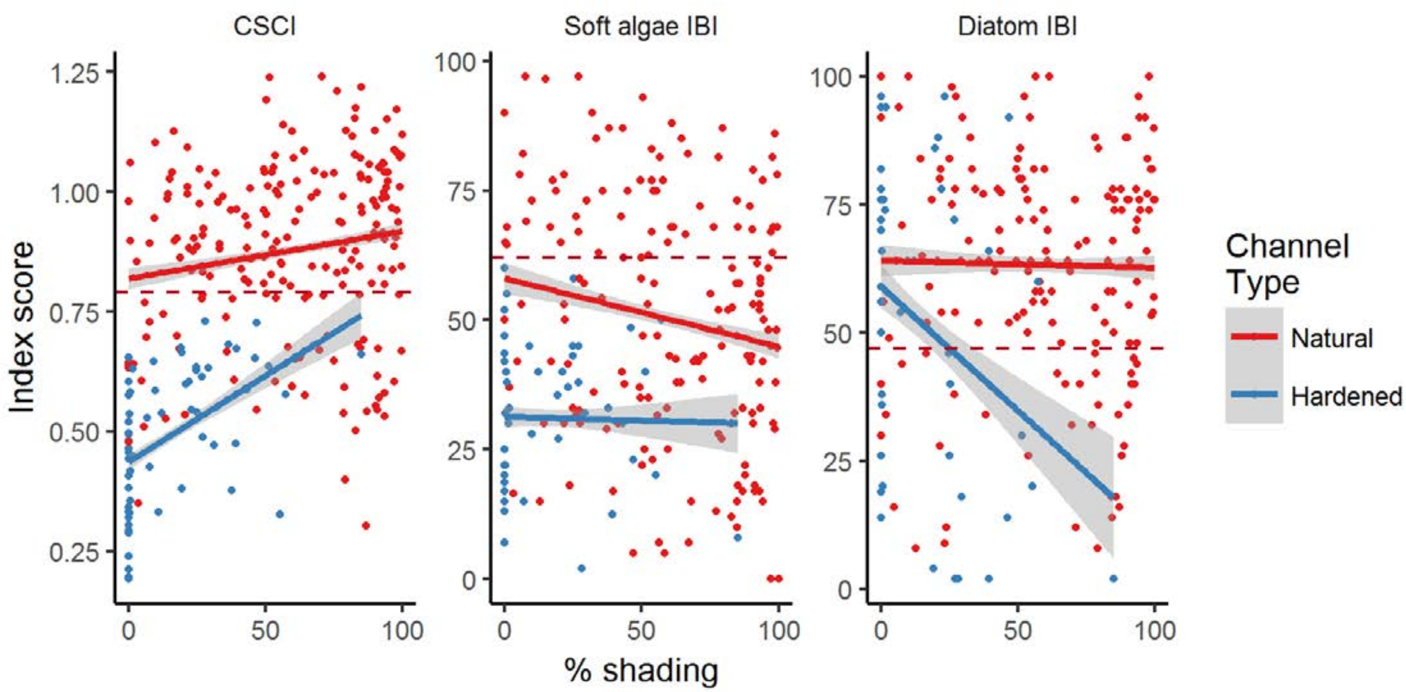


Figure 5. Shading versus index scores in natural and hardened channels. The red dashed line is the threshold for sites in reference condition. For clarity, earthen and partially hardened channels are excluded from this plot.

B365

Table 1. Correlations between water quality and habitat variables and index scores in different channel types. ρ: Spearman rank correlation coefficient. Coefficients indicating stronger relationships (ρ > 0.3) are highlighted in blue. *: p-value < 0.05.

	CSCI			Diatom IBI			Soft Algae IBI		
	Natural	Partial	Hardened	Natural	Partial	Hardened	Natural	Partial	Hardened
	ρ	ρ	ρ	ρ	ρ	ρ	ρ	ρ	ρ
Water quality									
Alkalinity	-0.30 *	0.02	0.21	-0.49 *	0.18	-0.39	-0.31 *	-0.16	-0.16
Chloride	-0.66 *	-0.52 *	-0.05	-0.73 *	-0.19	-0.30	-0.44 *	-0.34	-0.10
Total Nitrogen	-0.44 *	-0.38 *	-0.42	-0.52 *	-0.42 *	0.23	-0.51 *	-0.43 *	-0.38
pH	0.25 *	-0.16	0.08	0.15	-0.36 *	0.04	-0.05	-0.20	0.22
Temperature	-0.36 *	-0.21	-0.30	-0.42 *	-0.23	-0.13	-0.11	-0.38 *	0.11
Specific conductivity	-0.58 *	-0.51 *	-0.05	-0.66 *	-0.16	-0.25	-0.46 *	-0.29	-0.11
Physical habitat									
% fast-water	0.45 *	0.31	0.24	0.53 *	-0.11	-0.44 *	0.09	-0.18	-0.37
% thick algae cover	-0.31 *	-0.35	-0.16	-0.45 *	0.01	0.55 *	-0.03	-0.60 *	0.09
% sands and fines	-0.51 *	-0.64 *	-0.23	-0.64 *	-0.36 *	0.15	-0.19	-0.56 *	0.17
Flow diversity	0.29 *	0.36 *	0.43 *	0.31 *	-0.06	-0.27	0.15	0.23	0.01
Habitat diversity	-0.01	0.14	0.18	-0.28 *	0.14	0.40 *	-0.07	0.24	0.34
Substrate diversity	0.09	0.30	-0.18	0.09	0.25	0.20	0.23 *	0.45 *	0.20
Riparian disturbance	-0.36 *	-0.23	-0.25	-0.22 *	-0.31	0.22	-0.33 *	-0.38 *	0.24
Shading	0.13	0.43 *	0.63 *	-0.03	0.23	-0.10	-0.13	0.41 *	0.40 *
Riparian vegetation	-0.17	0.21	0.47 *	-0.20	0.20	-0.06	0.03	0.14	0.28

Little Dalton Wash: An example of a high-scoring engineered channel

Figure 6. Little Dalton Wash.

Index	Score	Percentile of reference	Percentile of hardened channels
CSCI	0.73	3	92
Diatom IBI	92	84	92
Soft algae IBI	23	0	15

Table 2. Index scores at Little Dalton Wash compared to reference sites and to other hardened channels. Percentiles calculated through a normal approximation.

Perhaps the most valuable insight provided by the SMC's study of engineered channels is that it helps managers identify examples of high-scoring sites, providing a target for managing streams in poorer condition. One such site is Little Dalton Wash, part of the San Gabriel River watershed in Azusa (Figure 6). Although the CSCI score of 0.73 was somewhat lower than the threshold of 0.79 for identifying sites in reference condition, it was higher than nearly all other hardened channels in the data set. Moreover, the diatom IBI score of 92 was well above the threshold of 62, although the soft algae IBI score was low (23). When compared to other hardened channels in the SMC survey, the unusually high scores at Little Dalton Wash are evident (Table 2).

The field conditions at Little Dalton Wash are not different from other hardened channels in any obvious way. The sampled reach is in a rectangular concrete box that lacks low-flow features. Located in the midst of a heavily developed area, it receives drainage from a 27-km^2 watershed that is more than one-third urbanized. However, comparison with survey data from other hardened channels suggest a few possibilities. Several water quality analytes, as well as physical habitat metrics, were better at Little Dalton Wash than at lower-scoring hardened channels, including chloride, total nitrogen, temperature, and specific conductivity (Figure 7). The diversity of flow microhabitats (e.g., riffles and glides) was high as well. These factors may explain the higher scores observed at this site.

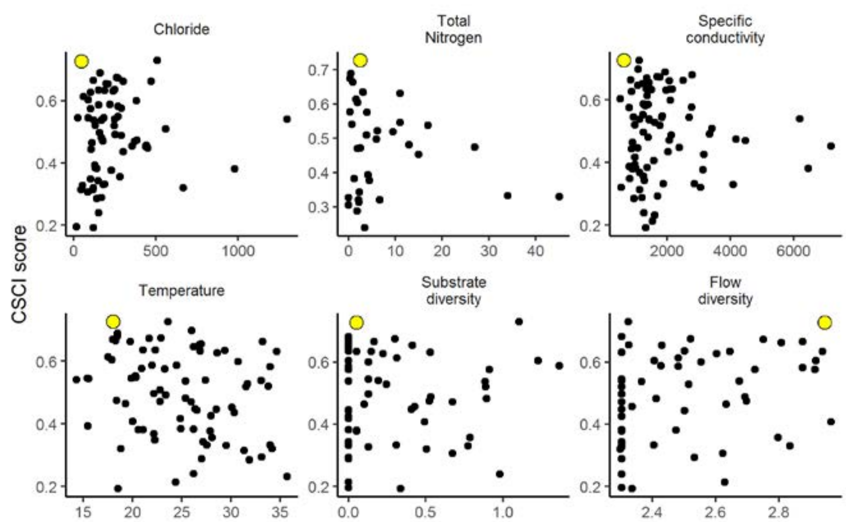


Figure 7. CSCI scores versus water quality and habitat metrics in hardened channels. The large yellow symbol represents Little Dalton Wash.

Conclusions

These preliminary results suggest that, although ecological health is clearly degraded in hardened channels, higher bioassessment index scores could be supported in certain reaches if water quality and in-stream habitat conditions are good. The ranges of observed index scores provide a starting point for managers, regulators, and stakeholders to discuss which types of actions are needed to achieve the desired level of health in modified channels, or in downstream receiving waters.

The SMC survey has cleared up a few major questions about engineered channels. It demonstrated that, although conditions in engineered channels are generally poor, some channels support better conditions than others. Additionally, this analysis underscores the value of a multi-indicator approach to ecological health assessment, as each assemblage adds to a more well-rounded view of the condition of engineered channels. Additional data may help further identify the factors that lead to better ecosystem health in engineered channels, including targeted sampling of concrete channels with good water quality, monitoring after the removal of concrete features from a channel (see Sidebar 3), and tracking water quality improvements following the implementation of best management practices that remove pollutants. Although this opportunistic analysis of available SMC survey data suggests that an engineered channel may not be able to support aquatic life as well as natural streams can, and tradeoffs between flood protection and ecological condition may be unavoidable, it shows that a range of conditions is possible, and that better conditions may be possible through management of water quality and habitat.

Restoration of engineered channels

Restoring natural features in engineered channels is sometimes proposed as a way to improve ecological conditions, as well as create amenities like improved flood control and enhanced recreational opportunities. In the County of San Diego, concrete walls and bank armoring were removed from a 1.2-mile segment of Forester Creek in 2006 at a cost of $36 million, returning the channel to a more naturalistic, vegetated form. Some water quality impairments improved following restoration (e.g., pH), while others did not (e.g., total dissolved solids). Bioassessment scores (measured with the Southern California IBI, which preceded the CSCI) increased from 25 to 40 points, although too few samples have been collected to see if this difference is statistically significant.

Left: Forester Creek upstream of the restoration site. Right: The restored portion of Forester Creek.

The Los Angeles River provides a much larger-scale example. The revitalization master plan for the Los Angeles River calls for the removal of concrete walls from up to 32 miles of the river, wherever it is safe and feasible to do so. This project may be one of the largest urban river restoration efforts undertaken in the country. With a cost that will exceed $1 billion, the impact on the river's ability to support aquatic life are not clear. Fortunately, the SMC stream survey provides abundant data, both from the Los Angeles River itself, as well as from comparable hardened and restored rivers, to provide benchmarks that enable the success of this effort to be evaluated.

Updates on new indicators

Cell bioassays evaluate the potential for harm from chemicals of emerging concern

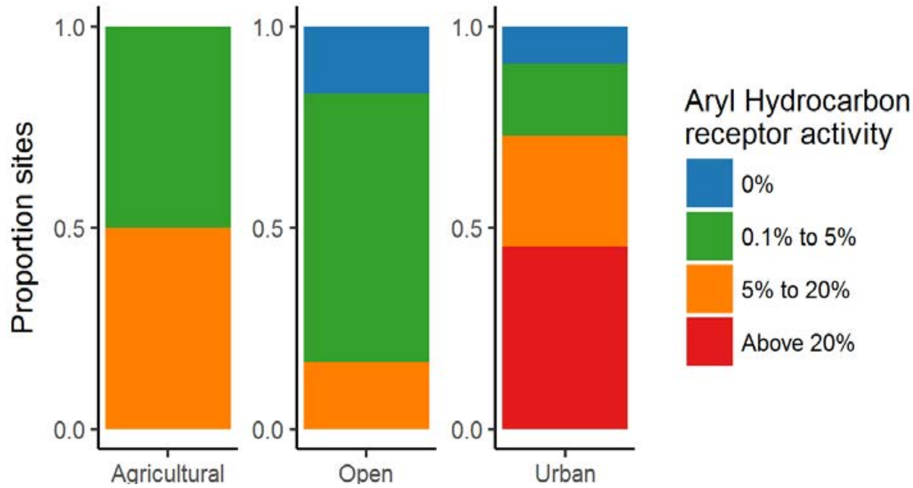


Figure 1. Aryl hydrocarbon receptor (AhR) responses were measured at sites representing different land uses within the SMC stream survey area.

Chemicals of emerging concern (CECs) have the potential to degrade ecological condition and harm human health through endocrine disruption and other physiological pathways. Comprising over 10,000 distinct chemical compounds, CECs come from pharmaceuticals, personal care products, and other sources. Many of them are biologically active, with the potential to disrupt hormonal pathways of organisms. With hundreds of new compounds being added to commercial markets every year, most without disclosure of their composition, measuring the extent and impact of CECs in the environment through traditional (i.e., single-compound) approaches is unrealistic.

The SMC survey tested an alternative approach that promises to be more effective and less expensive than traditional methods. First, samples are used in bioassays to detect cellular-level responses, followed by a non-target (i.e., multiple-compound) analysis to identify the compounds that could cause the observed response. This screening approach provides new information about potential risks of contaminant exposure to humans, aquatic life, and wildlife. For example, estrogen receptor (ER) assays can help detect the presence of hormone-mimicking chemicals that affect growth, development, and reproduction. The SMC screened 31 samples collected in 2015—one of the first applications of this new technology to a stream biomonitoring program.

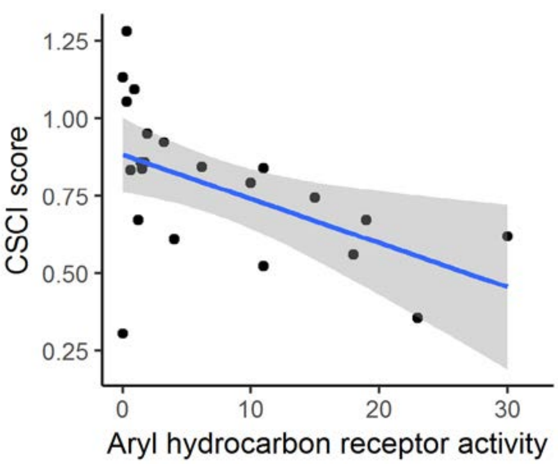


Figure 2. Aryl hydrocarbon receptor response versus CSCI scores

Responses for receptors of steroid hormones, such as glucocorticoid and estrogen, were rare, affecting only 2 and 8 sites respectively. In contrast, aryl hydrocarbon receptor (AhR) responses were widespread, affecting 28 of the 31 sites; furthermore, AhR responses were stronger at urban sites than at undeveloped sites (Figure 1). The AhR receptor is thought to play a role in mediating environmental toxicity and immune response, as well as supporting normal vascular development. Dioxins and other pollutants are known to provoke AhR responses.

Bioassay responses may explain why some sites are in poor biological condition. For example, AhR activity was negatively correlated with CSCI scores ($r = -0.84$, Figure 2), suggesting that contaminants known to cause AhR responses (e.g., polycyclic aromatic hydrocarbons, commonly associated with runoff from asphalt or combustion) may alter benthic macroinvertebrate assemblages (Figure 2). Follow-up targeted chemical analyses found sunscreen ingredients at sites with ER activity and flame retardants at sites with AhR activity, although concentrations were generally too low to explain the observed responses, meaning that other, unmeasured compounds are responsible. Field blanks were clean, meaning that contamination of the samples was not a likely cause of the response. Non-targeted analyses to identify these unknown chemicals are underway.

Assessing the ability of streams in southern California to support aquatic vertebrates


Figure 1. California tree frog (*Pseudacris cadaverina*), one of the more common native species of vertebrate found in Southern California streams.

Although the initial SMC stream survey provided a great quantity of data about stream condition based on benthic macroinvertebrates and algae, a lingering question remained about what our findings meant for higher trophic levels, such as fish, amphibians, and other vertebrates (Figure 1). Although a thorough investigation of this question is beyond the scope of the current regional monitoring program, the SMC found a way to get some answers, and at remarkably low costs.

In 2015, SMC field crews received training in identifying common aquatic vertebrates in the region, and began reporting observations of species they encountered during normal bioassessment sampling (that is, no additional time was spent trying to observe vertebrates). This effort began as a collaborative venture initiated by the SMC, the US Geological Survey (USGS), the US Fish and Wildlife Service (USFWS), and SWAMP, all of whom were hoping to improve their understanding of the spatial distributions of both native and non-native vertebrates in the region. The survey provided a concrete example of how the core regional monitoring program can be used to opportunistically collect data to answer important management questions. The resources necessary to successfully complete the survey were relatively trivial for several reasons: the SMC field teams were already visiting the sites; the teams already included biologists easily trained to identify stream vertebrates; and the sampling design was based on a time-saving casual observation approach, instead of a more traditional rigorous search at each site.

Despite the low costs, this survey provided a great deal of new and valuable data on vertebrates in the region's streams, with observations attempted at a total of 95 sites (Figure 2). Vertebrates were seen at 46% of the sites, and surprisingly, the distributions of native frogs were fairly widespread across urban, agricultural and open land use types. These native amphibians were unexpectedly tolerant to the presence of non-native fish, frogs or crayfish. In contrast, native fish species were only observed at five mountainous sites. Mosquito fish (*Gambusia affinis*) were observed at 21 sites and were the most common non-native fish species, likely as a result of deliberate introduction as a vector-control measure.

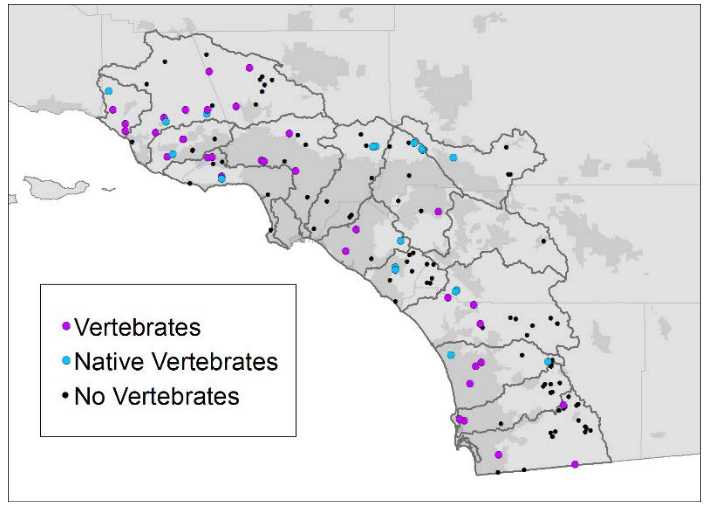


Figure 2. Location of vertebrate observations conducted in 2015.

Of the sites located on agricultural land, 68% supported vertebrates, although the many of these were non-natives (47%). In contrast, 49% of open land sites supported vertebrates, but only 17% of these were non-natives. It is important to note that these numbers likely underestimate the actual distribution of vertebrates because the field crews did not conduct exhaustive surveys of each site.

The addition of vertebrate observations to the survey yielded detailed information regarding the distributions of vertebrates throughout the southern California region, despite the limited amount of resources and training required to successfully implement it. Although more intensive efforts may have detected more species (especially nocturnal or cryptic species), opportunistic sampling was sufficient to improve our understanding of the ability of Southern California streams to support wildlife. Future work for this program might focus on the environmental and habitat factors that contribute to the presence or absence of vertebrates on agricultural, urban and open land use types; investigation of the relationships among vertebrates and other biological condition indicators including the CSCI, CRAM and Southern California algal IBIs; and improving our understanding of the spatial distribution of these important taxa by combining the SMC vertebrate dataset with those from iNaturalist, regional fish surveys, the USFWS, the USGS, and the California Department of Fish and Wildlife.

Applications of survey data

A water-quality improvement plan supported by survey data

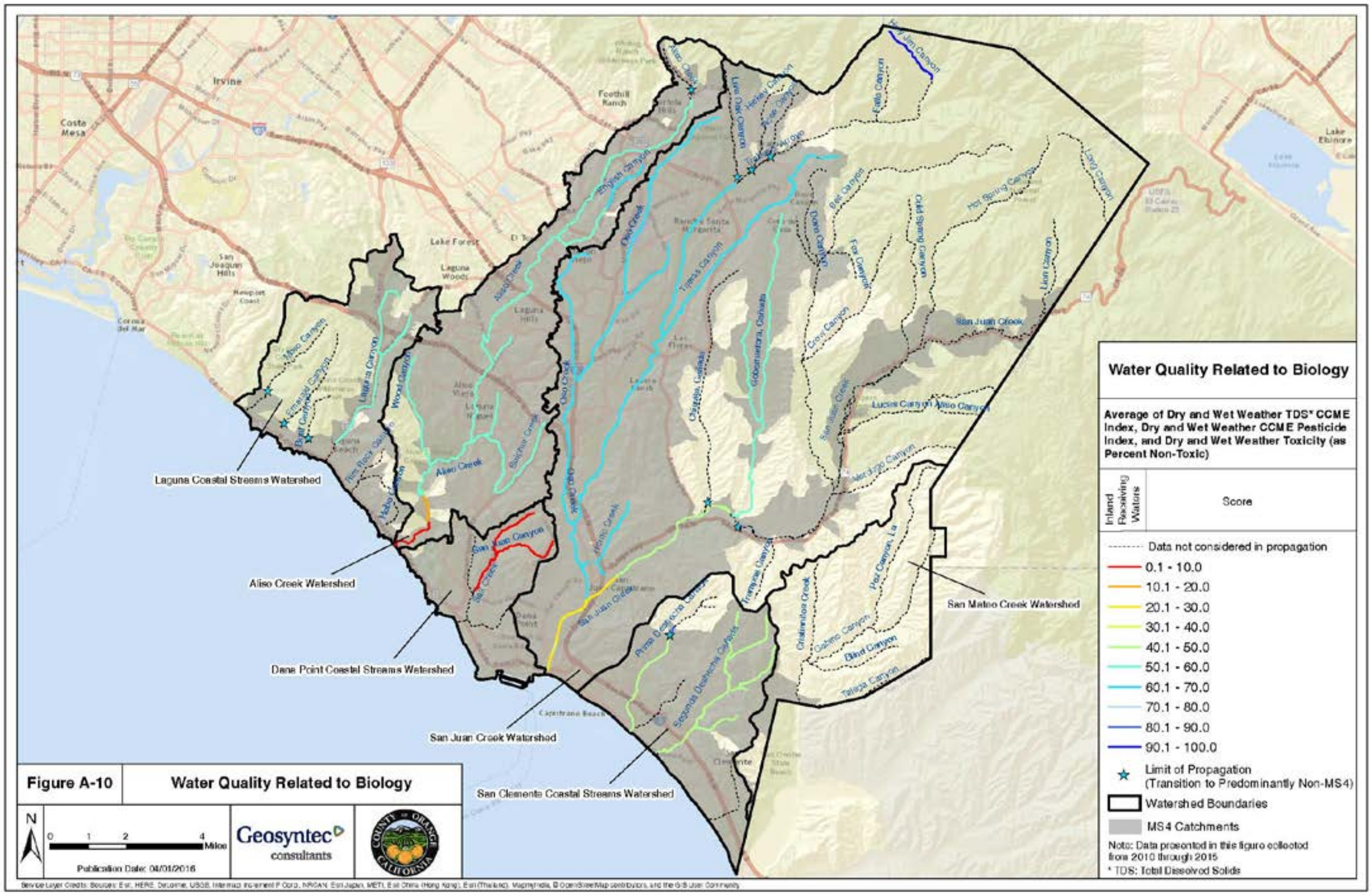


Figure 1. Excerpt from the San Juan Water Quality Improvement Plan shows how the County of Orange used SMC bioassessment and water quality data to prioritize problem areas in the watershed. Red, orange, and yellow stream segments have low-scoring bioassessment sites, in conjunction with measures of poor water quality. A separate analysis identifies stream reaches where low-scoring bioassessment occur in conjunction with geomorphic alteration.

A key objective of the SMC stream survey is to provide participants with data that helps them manage watersheds. One recent notable example is Orange County's Water Quality Improvement Plan (WQIP) for the San Juan Hydrologic Unit, which prioritizes problems in the watershed based on SMC data, emphasizing biological indicators like benthic macroinvertebrates and algae. The goal of the WQIP is to 1) determine high-priority water-quality problems; 2) identify goals, strategies, and schedules to address them; and 3) propose an approach to monitor and assess progress. In all three elements, the SMC survey provides the foundation and the framework for implementing these goals.

The WQIP identified three priority problems, and two of them were determined through bioassessment data: geomorphic alteration, and unnatural flow regimes. These problems were identified by the association of stressors related to these problems (e.g., hydromodification and habitat degradation), and their relationships with poor bioassessment index scores (Figure 1). Best management practices to mitigate these stressors will be identified, and their success will be partly determined in terms of improvements in biological condition. The monitoring and assessment component of the WQIP is currently under preparation, and it is expected that biological monitoring through the SMC stream survey will play a crucial role in this component.

Regional flow targets to support biological integrity

The SMC stream survey data provides a strong foundation to explore the problems affecting streams in the region, such as hydrologic alteration, which previous surveys suggested is a major factor affecting biological condition. Hydrologic alteration results from water diversions, inter-basin transfers, and increased imperviousness that alter the natural flow regime in a stream. Taking advantage of a new ensemble-modeling approach to estimate current and historic flows at ungauged streams, hydrologic alteration was estimated at 572 bioassessment sites, most of which are part of the SMC stream survey. The ensemble was built by calibrating simple rainfall-runoff models at 26 stream gauges in Southern California, then assigning one model to ungauged sites with similar catchment properties. Biological responses (e.g., California Stream Condition Index [CSCI] scores) were modeled against metrics reflecting hydrologic alteration, thresholds of biological response were established for multiple flow metrics, and metrics were combined into an overall index of hydrologic alteration with scores ranging from 0 (no alteration) to 14 (all metrics severely altered).

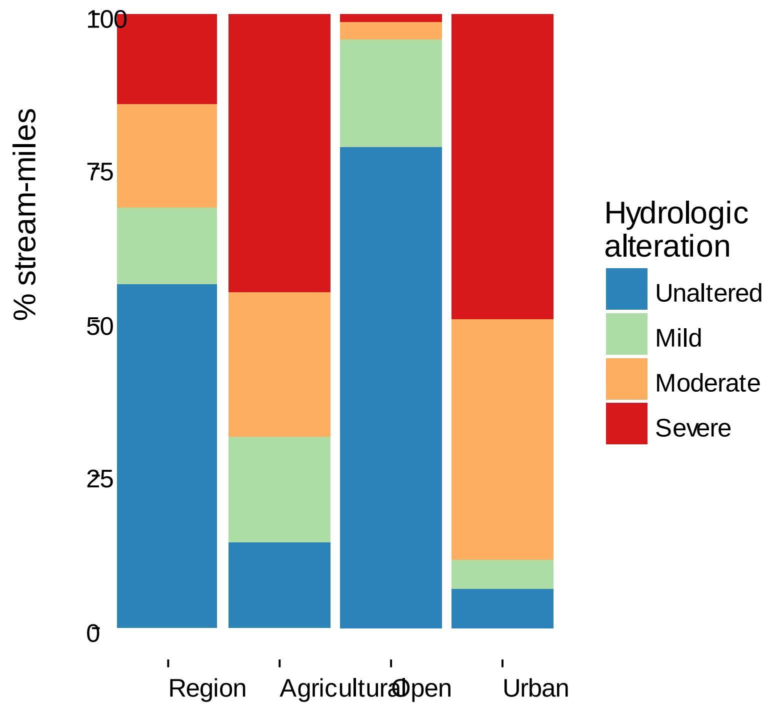


Figure 1. Extent of hydrologically altered streams in the region, as well as within three land-use classes.

Because this index was applied to survey data, it allowed the first-ever estimate of the extent of hydrologic alteration in the region. Approximately 34% of stream-miles in Southern California were estimated to be moderately or severely hydrologically altered, and alteration was more pervasive in urban (91% stream-miles altered) and agricultural (80%) than undeveloped (11%) streams (Figure 1).

The index also allowed rapid setting of management priorities and causal assessment screenings (Table 1, Figure 2). Among the biologically healthy sites (i.e., CSCI scores > 0.79), hydrologically unaltered sites (52% of total stream-miles) were prioritized for protection, and hydrologically altered sites (4%) were prioritized for monitoring. Among the biologically degraded sites, 30% were hydrologically altered, and prioritized for evaluation of flow management (such as increased stormwater detention or groundwater infiltration). Evaluation of other stressors was prioritized at the remaining 14% of stream-miles.

Table 1: Management action priorities based on measures of biological condition and hydrologic alteration.

	Unhealthy biology	Healthy biology
Altered hydrology	Evaluate flow management: 30%	Monitor: 4%
Unaltered hydrology	Evaluate other stressors: 14%	Protect: 52%

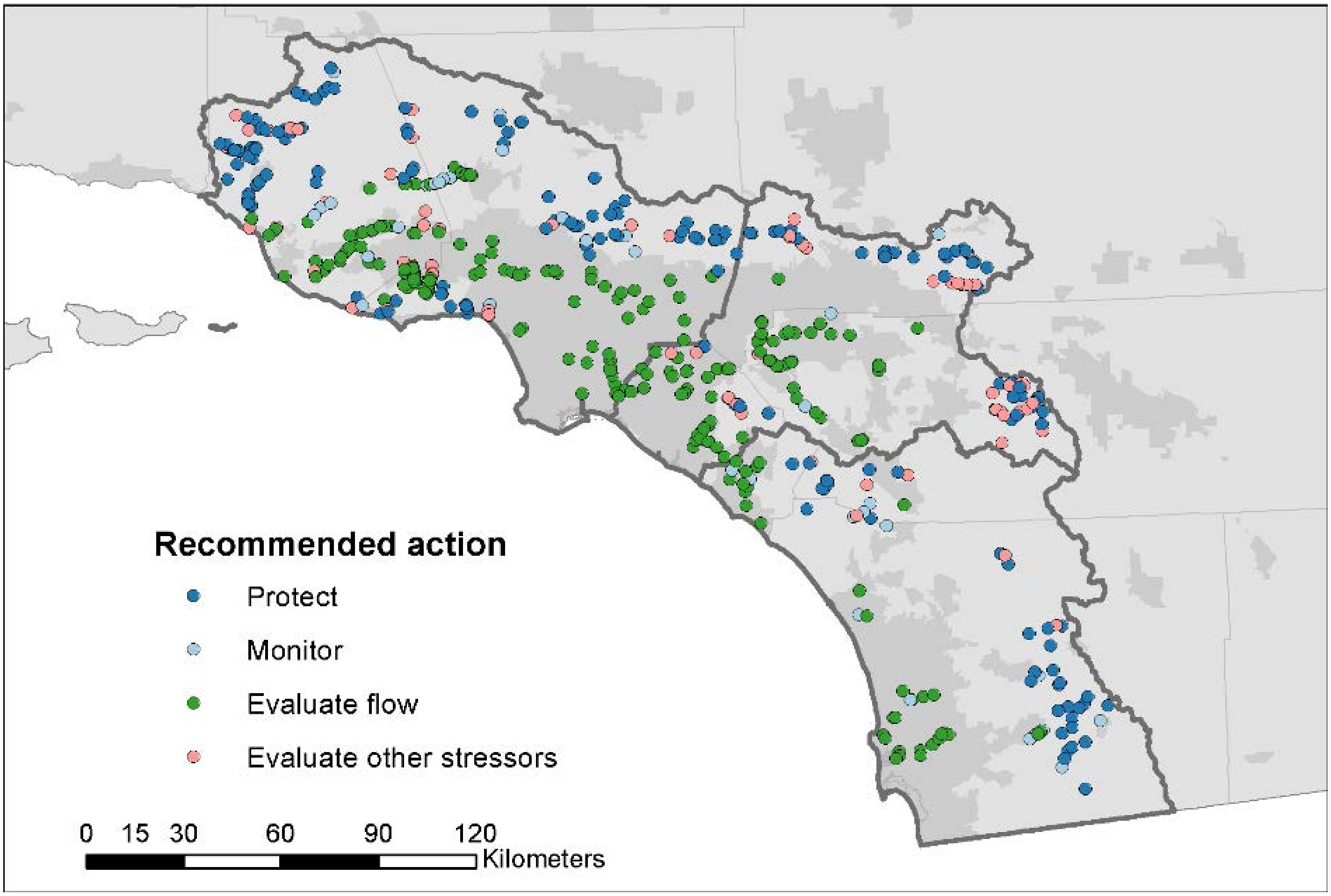


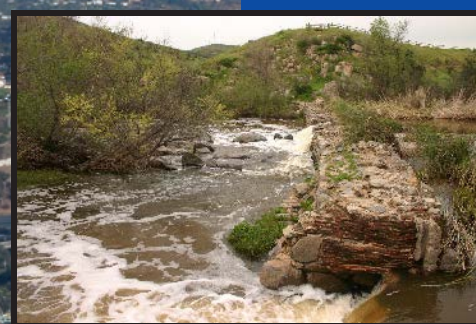
Figure 2. Management priorities for streams in the SMC region, based on estimates of hydrologic alteration and biological condition.

Regionally derived, biologically based targets for flow allow watershed managers to rapidly prioritize activities and conduct screenings for causal assessments at many sites across large spatial scales. Furthermore, regional tools pave the way for incorporation of hydrologic management in policies and watershed planning designed to support or enhance biological integrity in streams. Development of regional tools should be a priority in regions where hydrologic alteration is pervasive or expected to increase in response to climate change.

Appendix 12

B375

Application of Regional Flow-ecology to Inform Management Decision in the San Diego River Watershed



Eric D. Stein
Ashmita Sengupta
Raphael Mazor
Kenny McCune

Southern California Coastal Water Research Project

SCCWRP Technical Report 948

B376

Application of Regional Flow-ecology to Inform Management Decision in the San Diego River Watershed

Eric D. Stein, Ashmita Sengupta, Raphael D. Mazor, and Kenny McCune

Southern California Coastal Water Research Project, Costa Mesa, CA

October 2016

Technical Report 948

ACKNOWLEDGEMENTS

We thank the stakeholder workgroup for their participation and dedication to completing this demonstration project. The proof of concept and lessons learned would not have been possible without their engagement and support. We also thank Drs. Derek Booth and Brian Bledsoe for their review of draft manuscripts. Their critical reviews greatly improved the quality of the final product. This work was funded by the State Water Resources Control Board under Grant Agreement #12-430-550 for "Development of Numeric Flow Criteria to Support Freshwater Bio-Objectives, Hydromodification Management, and Nutrient Numeric Endpoints."

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Changes to instream flow are known to be one of the major factors that affect the health of biological communities. Regulatory, monitoring, and management programs are increasingly using biological community composition, particularly benthic invertebrates, as one measure of instream conditions, stormwater project performance, or regulatory compliance. Understanding the relationship between changes in flow and changes in benthic invertebrate communities is, therefore, critical to informing decisions about ecosystem vulnerability, causes of stream and watershed degradation, and priorities for future watershed management.

Taking advantage of large, robust regional monitoring data sets and recently completed regional watershed models, the Southern California Coastal Water Research Project (SCCWRP) has developed a set of "flow–ecology" relationships for southern California that relate changes in specific flow metrics to changes in benthic invertebrate indices that have been shown to be indicative of stream health. These relationships are based on the Ecological Limits of Hydrologic Alteration (ELOHA) framework, which uses a variety of hydrologic and biologic tools to determine and implement environmental flows at the regional scale. Results of the ELOHA analysis can inform management decisions, such as release rates from dams, reservoirs or basins; diversion volumes for irrigation or water re-use, or flows associated with stream restoration.

The goal of this project is to demonstrate how regionally derived flow–ecology relationships can be implemented at a watershed scale to inform management decisions. Regional relationships allow us to describe general patterns of response in biological communities to changes in hydrology. Local case studies are critical to determine how these relationships can be applied to site-specific decisions, and to identify areas where the regional relationships may need to be refined to better support local application.

Our case study focused on the San Diego River Watershed in southern California, where the potential effects of urban growth and water/runoff management on stream flow and biological condition are currently being considered. We worked with a group of local watershed stakeholders to identify three questions that that would both inform local management decisions (along with other planning considerations) and demonstrate the utility of the regional flow–ecology relationships. Close coordination with the stakeholder group enhanced the relevancy of the analysis and helps to determine how the technical approach to establishing targets may be applied in other areas. The case study focused on the following management questions:

1. How will future land use changes affect flow conditions and impact biological endpoints in the San Diego River watershed? This involves a comparison of the current hydrologic conditions to modeled conditions based on San Diego County's 2050 land use projection. Future scenarios did not include any assumptions about best management practices, low impact development or hydromodification, which would be expected to reduce potential effects of future hydrologic alteration.

2. How can we use our understanding of current and expected future hydrologic conditions along with the regional flow–ecology relationships to prioritize regions of the watershed where flow management may be most critical to maintain or improve future stream health?

3. What are the biological implications of two future management decisions that will affect in-stream flow conditions:

4. What would be the effects of reduced discharge from Santee Lakes Reservoir due to increased capture and storage to meet demand for reclaimed water?

5. What would be the effect of disconnecting imperviousness and implementing stormwater capture strategies in a currently developed portion of the watershed?

These local management questions were addressed using regional flow-ecology relationships that relate changes in stream health to changes in hydrology. Stream health was assessed using the California Stream Condition Index (CSCI), a statewide index of benthic macroinvertebrates community composition. Hydrologic alteration was assessed based on the following hydrologic metrics, which were shown to have strong statistical and ecological relationships with the CSCI (Table ES-1; See Mazor et al. in review). Metrics were also selected to ensure representation of different components of the flow regime (e.g. duration, magnitude, etc.) and different climate conditions (e.g. wet vs. dry vs. average years).

Table ES-1. Priority hydrologic metrics used in the regional flow-ecology relationships. Metrics are grouped by the hydrograph component they represent. Metric effects on biology were typically strongest during either average, wet, or dry rainfall years, or all years combined (overall).

Hydrograph Component	Metric	Metric Definition	Critical precipitation condition
Duration	NoDisturb (days)	median annual longest number of consecutive days that flow is between the low and high flow threshold	Average
	HighDur (days/event)	median annual longest number of consecutive days that flow was greater than the high flow threshold	Wet
Magnitude	MaxMonthQ (m3/s)	Maximum mean monthly streamflow	Wet
	Q99 (m3/s)	streamflow exceeded 99% of the time	Wet
Variability	RBI (unitless)	Richards-Baker index of stream flashiness	Dry
	QmaxIDR (m3/s)	interdecile range of flow	Overall
Frequency	HighNum (events/year)	median annual number of events that flow was greater than high flow threshold	Dry

Effect of future land use change

Under current land use conditions, 44% of the catchments in the watershed were considered hydrologically altered based on the metrics shown in Table ES-1. There is a broad spatial gradient of hydrologic degradation in the watershed, with the most hydrologically intact areas in the upper watershed, moderately altered catchments in the middle watershed, and the most hydrologically altered catchments in the lower watershed (Figure ES-1). Hydrologic alteration is largely correlated with total impervious cover, with hydrologic alteration generally becoming measurable as the impervious cover reaches and exceeds 5%. Given this pattern, hydrologic conditions are expected to degrade under San Diego County's projected 2050 land use for the watershed (Figure ES-1). The majority of new impacts are expected to occur in the upper watershed where current open space may convert to low-density residential land use and exceed the 5% impervious cover level. Based on the regional flow-ecology relationships, we expect that future hydrologic changes will also manifest as declines in benthic invertebrate communities, reflecting an overall impairment of biological conditions. Efforts to reduce *effective impervious cover* through low impact development or hydromodification control (which act to disconnect total imperviousness from streams) would be expected to reduce future impacts.

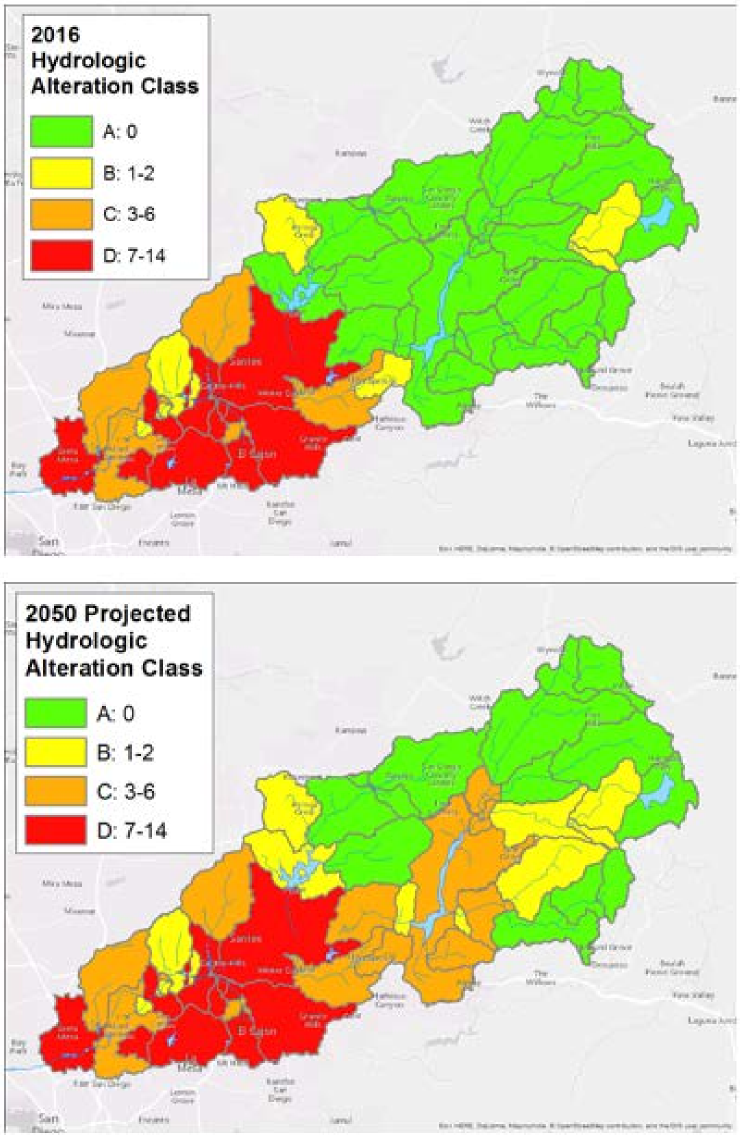


Figure ES-1. Hydrologic alteration under current (top) and 2050 projected (bottom) land use.

Prioritization of areas for various management actions

We prioritized areas of the watershed for various management actions using a combination of hydrologic alteration (see Figure ES-1) and biological condition based on existing bioassessment data (using the CSCI). The majority of upper watershed sites were considered intact and thus a high priority for preservation or protection (Figure ES-2). Two sites in the middle watershed had altered hydrology, but healthy biological communities. This suggests that the communities are either resilient or have not yet responded to the hydrologic alteration. Therefore, these sites should be monitored for potential future degradation. The lower watershed largely expressed both poor biological condition and altered hydrology. For sites in the lower watershed where both hydrology and biology were in altered, we examined available data on water quality and channel condition to better understand the relative contribution of flow alteration vs. other stressors to reduced biological health. This analysis allowed us to provide preliminary management recommendations that can be prioritized for each location (Figure ES-2). We estimate that flow alteration alone is the principle factor affecting biology at only 3 of the 13 biologically degraded sites in the lower watershed. At all other sites, flow management should be coupled with habitat or water quality remediation in order to improve biological conditions.

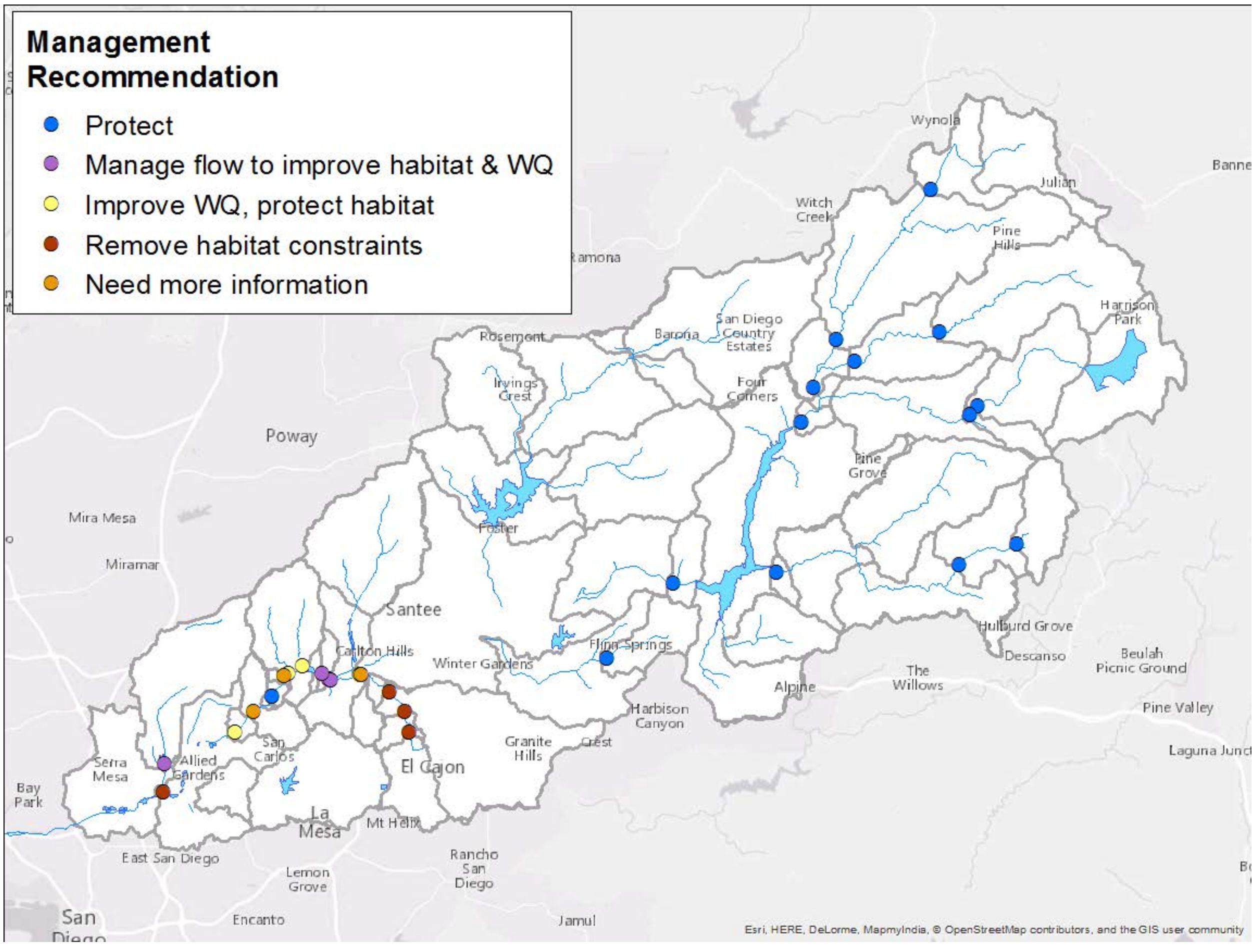


Figure ES-2. Recommended management actions for all sites based on a combination of hydrologic and biological condition. Recommendations are based on both flow-ecology information and available data on habitat and water quality obtained through the local regional monitoring program. Only sites with existing bioassessment data are included.

Evaluation of management scenarios

We demonstrated application of the flow–ecology tools to evaluate both a reservoir management scenario and an urban runoff management scenario. The reservoir management scenario involves eliminating discharge of treated wastewater into the Santee Lakes Reservoir and redirecting it for reuse to help meet increased demands for recycled water. This would reduce reservoir outflow and change hydrology of the downstream Sycamore Creek. Our analysis showed that modifying reservoir management would reduce several flow metrics closer to reference conditions; however, they will probably not fully return to reference condition due to the ongoing contribution of urban runoff. Overall, certain components of the hydrograph (usually under high flow conditions) in Sycamore Creek will improve, but is likely to remain in degraded hydrologic and biological condition, even if discharges from Santee Lakes Reservoir are eliminated following the proposed management scenario.

We modeled two urban runoff management scenarios: 1) disconnecting impervious areas from discharging to streams (i.e. reducing impervious cover), and 2) implementing stormwater retention facilities that can capture 85^{th} percentile of a 24-hour rain event. Disconnecting imperviousness decreases the extent of hydrologic alteration in the downstream reaches. However, flow metrics do not return to levels associated with healthy biological communities until the total imperviousness is at or below 5%. Analysis shows that for most metrics, there is a 66% likelihood of meeting flow targets at 5% total impervious cover and an 80% likelihood of meeting flow targets at 2% total impervious cover (i.e. with stormwater control measures installed). The sensitivity of the creek to relatively low levels of impervious cover is consistent with past studies from southern California. In contrast, retention of the 85% storm event (as is currently required through the local stormwater permit) resulted in flow metrics that met all target values.

Utility of the ELOHA approach for establishing flow-ecology relationships

A major objective of this case study was to evaluate the ability to apply the flow–ecology relationships derived from the regional ELOHA analysis to inform local watershed-scale decisions. Our results illustrate that several of the stated advantages of the ELOHA approach do aid in such watershed-scale application. The ability to apply regionally derived flow targets to inform local decisions is a major advantage of the ELOHA approach. This eliminates the need to develop local flow–ecology relationships for every stream of interest, as is the case in more traditional instream flow methods. The tools developed through the regional analysis provided readily transferable tools for local stakeholders to produce measures of hydrologic change for any location of interest and to explore how those values would change under different land-use or management scenarios. This had the dual benefit of allowing for robust analysis and providing a vehicle for stakeholder engagement in setting management priorities related to instream flow. A potential downside of the ELOHA approach is that the regionally established flow targets may not fully address all concerns or considerations at a specific project location in the same manner as a site-specific analysis would. Ultimate policy decisions about how streams are managed must balance many competing needs. This case study shows how regional flow-ecology relationships can help inform these decisions.

Lessons learned for future implementation of regional flow-ecology relationships

Future efforts can build on the experiences from this case study and continue to refine an iterative process of developing flow targets that are scientifically defensible, practical (i.e., can lead to management actions), and consistent with local stakeholder needs. Key lessons learned from this effort include:

1. Include a broad set of engaged stakeholders, including regulatory agencies, municipalities, water agencies, non-governmental organizations, and researchers. This ensures a broad perspective in the deliberations and increases the likelihood of developing balanced recommendations.

2. Invest in educating the stakeholders early in the process on the underlying science and the rationale behind how regional flow targets were developed. This promotes engagement and fosters creative solutions to the complex challenges of flow management.

3. Invest the time to compile high quality local data sources and show how local data can be used in the evaluation process. Identify the areas were future data collection can most improve outputs of

the flow–ecology analysis (e.g., local rainfall data, more refined land use, water quality data). This can inform future monitoring.

4. Develop documentation that clearly illustrates how the products of the flow–ecology analysis can be used in the context of existing regulatory or management programs.

The San Diego River implementation case study also produced several technical recommendations that can improve our ability to apply flow-ecology relationships to manage southern California streams:

1. Several flow metrics, particularly those associated with flow duration, may require modification for use in streams where the natural condition is intermittent or ephemeral. Application of regionally derived flow thresholds to specific streams that may have been naturally intermittent can lead to erroneous results.

2. Metrics associated with flow durations should be calculated on a single threshold value based on reference conditions. Estimating hydrologic change based on a moving threshold estimated separately for current and reference conditions may produce erroneous results.

3. Need to improve the representation of the drainage system to provide a more accurate hydrologic foundation for analysis. This would ultimately include improved mapping of discharges, diversions, stormwater control facilities, low impact development (LID), etc. for incorporation into modeling scenarios and effects.

4. Consider expanding the analysis to include additional elements in future case studies
    a. Include other stream or water body types
    b. Include other indicators (e.g. algae)
    c. Explore how consistent/transferable findings are from one watershed to another
    d. Explore application in watersheds that cross jurisdictional boundaries

INTRODUCTION

Flow regime has been shown to affect a broad suite of ecological processes and biological communities (Bunn and Arthington 2002, Naiman et al. 2002, Poff et al. 1997, Poff and Zimmerman 2010, Novak et al. 2015). Many studies have demonstrated that alterations of flow regime can be associated with changes in macroinvertebrate assemblages, which are used as key bioindicators for many regulatory and management programs globally (Pringle et al. 2000, Miller et al. 2007, DeGasperi et al. 2009, Poff & Zimmerman 2010). Although a basic understanding of the relationship between flow alteration and ecological response exists (Poff et al. 2010), few studies have demonstrated how to develop regulatory or management objectives (or targets) based on these relationships. Establishing quantitative and predictive relationships between change in flow and change in biological community composition is a critical step in using bioassessment indicators to establish measures of project performance or regulatory compliance.

Various approaches have been used to develop relationships between flow characteristics and biological response. Examples include use of habitat suitability models that relate flow change to requisite habitats for target taxa (e.g., MesoHABSIM, Parasiewicz 2009; and PHABSIM, Beecher et al. 2010); establishment of functional flow regimes to support species of management concern (McClain et al. 2014, Yarnell et al. 2015); and use of statistical ranges of sustainability based on unaltered hydrographs (Richter et al. 2011). Concepts from several of these approaches have been organized into the Ecological Limits of Hydrologic Alteration (ELOHA) framework (Poff et al. 2010). The ELOHA framework uses a variety of hydrologic and biologic tools to determine and implement environmental flows at the regional scale. Results of the ELOHA analysis can inform management decisions, such as release rates from dams, reservoirs or basins, diversion volumes for irrigation or water re-use, or flows associated with stream restoration. Because the ELOHA framework provides a way to assess the effect of flow alteration on the condition of biological communities (vs. individual taxa) on a regional basis, it is a useful approach for setting targets across a wide range of geographies and stream types where comprehensive detailed site-specific investigations are not practical. The ELOHA framework includes elements of stream classification, estimation of flow alternation (termed “delta H”) and development of flow ecology relationships based on the relationship between delta H and changes in the biological community (“delta B”).

There have been several recent applications of the ELOHA framework to develop flow targets for benthic invertebrates, fish, mussels, amphibians, and aquatic and riparian vegetation. Buchanan et al. (2013) completed the ELOHA approach in the mid-Atlantic region of the U.S. and was able to show clear relationships between changes in a subset of six flow metrics and six benthic invertebrate endpoints. This allowed the authors to recommend specific metrics that could be used for monitoring and assessment. McManamay et al (2013) applied ELOHA through a case study in North Carolina to assess the effect of a stream restoration on fish and riparian communities. Although the ELOHA framework worked well at documenting effects of the restoration projects, confounding factors (e.g., associations between delta H and water chemistry alteration) produced equivocal relationships between flow alteration and response of the fish community. The Nature Conservancy has developed ecosystem flow recommendations for the Susquehanna River Basin (DePhilip and Moberg 2010) and the upper Ohio River Basin (DePhilip and Moberg 2013) that provide seasonally differentiated targets for different stream classes and multiple biological endpoints (e.g., fish, mussels, amphibians, vegetation). Solans and Jalon (2016) used a series of flow alteration-ecological response curves to develop environmental flow standards for the Ebro River Basin in the Iberian Peninsula. Most recently, Mazor et al. (in review) capitalized on extensive regional biomonitoring data and a set of regional hydrologic models developed by Sengupta et al. (in review) to develop flow-ecology relationships for southern California based on benthic macroinvertebrate communities as a measure of stream health.

Previous studies have demonstrated the utility of the ELOHA framework for establishing flow targets and thresholds using relationships between changes in flow and changes in biological condition. Broad scale application of ecologically derived flow targets (or thresholds) can be informed by case studies that demonstrate how flow-ecology relationships can be used to inform actual management decisions. In addition to the study by McManamay et al (2013), the main place where flow-targets have been implemented to inform management actions is in the Juanita Creek Watershed in Washington State, USA (King County 2012). The Juanita Creek study evaluated the effectiveness of seven potential stormwater mitigation scenarios at achieving biologically relevant flow targets using a calibrated Hydrological Simulation Program-Fortran (HSPF) model; a single scenario was identified which would accomplish the stated watershed goals. To our knowledge, none of the previous cases studies attempted to apply regionally-derived flow-ecology relationships (such as those developed for southern California) to inform decisions at the watershed scale. Additional case studies that demonstrate this application can provide a template for future applications of flow-ecology based targets, and allow for consideration of lessons learned to refine these future applications. Such case studies are also important because they provide an opportunity to work with local watershed stakeholders to identify management needs and apply ecohydrology analyses to inform decisions in a way that balances consideration of ecological endpoints with other needs (e.g., water supply management, new infrastructure and development, flood control).

The goal of this project is to demonstrate how the regionally derived flow–ecology relationships developed by Mazor et al. (in review) can be implemented at a watershed scale to guide management targets/decisions. Regional relationships allow us to describe general patterns of response in biological communities to changes in hydrology. Local case studies are critical to determine how these relationships can be applied to site-specific decisions, and to identify areas where the regional relationships may need to be refined to better support local application.

METHODS

Study area

We conducted the demonstration in the San Diego River watershed, in San Diego County, California, where the potential effects of urban growth and water/runoff management on stream flow and biological condition are currently being considered (Figure 1). At 440 square miles (1,140 square km), it is among the largest watersheds in San Diego County and also has the highest population (~475,000), containing portions of five cities and several unincorporated communities. Important hydrologic resources in the watershed include five water storage reservoirs, a large groundwater aquifer, extensive riparian habitat, and coastal wetlands. Approximately 58% of the San Diego River watershed is currently undeveloped. The majority of this undeveloped land is in the upper, eastern portion of the watershed, while the lower reaches are more highly urbanized. The San Diego River watershed is a valuable case study because it includes a range of stream types, including reference (as defined by Ode et al. 2016) and highly impacted reaches; it is affected by several types of hydrologic alteration, including urban runoff, flood control, and reservoir management; it is relatively data-rich, benefiting from years of ambient and targeted monitoring programs (e.g., Mazor 2015); and there is an active and engaged watershed workgroup that is willing to participate in the demonstration project.

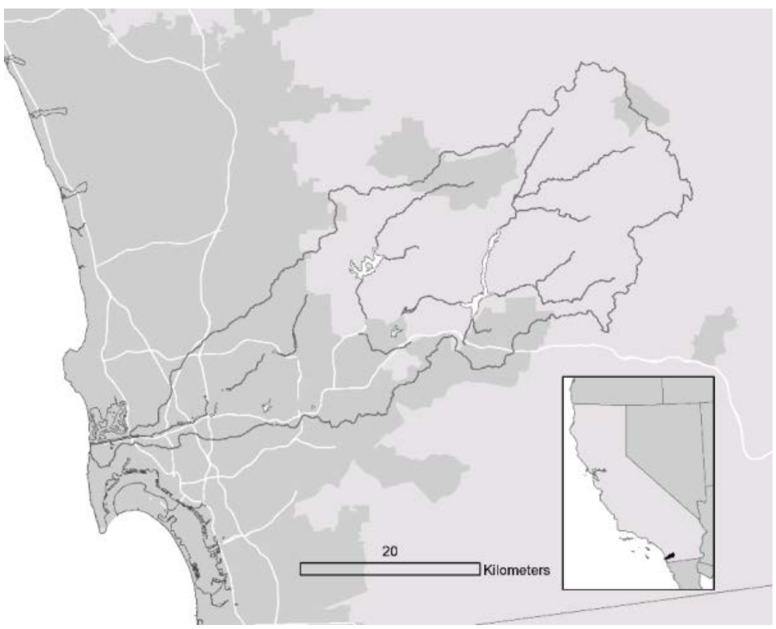


Figure 1. San Diego River Watershed

Stakeholder Process

Active stakeholder participation is integral to a successful demonstration case study because the stakeholders must identify the issues and interpret the utility of the recommendations resulting from the analysis. Stakeholders for the San Diego River case study included local municipalities, water districts, a land conservancy, a non-governmental organization, water quality regulatory agencies, the U.S. Forest Service as the upper watershed landowner and a local consulting firm (Table 1).

The stakeholder workgroup met monthly over an eight-month period and was facilitated by technical staff from the Southern California Coastal Water Research Project (SCCWRP), who had recently completed a regional ELOHA analysis (Mazor et al. in review). The workgroup was engaged in all aspects of the project including detailed scoping, assisting in modeling and analysis, and interpretation and refinement of findings. This intimate participation was key to developing products that would be acceptable for incorporation into future management decisions. A list of workgroup participants and topics for each workgroup meeting are provided in Appendix B.

Table 1. Stakeholders who participated in the San Diego River case study

- City of San Diego
- U.S. Forest Service
- Helix Water District
- Padre Dam Municipal Water District
- San Diego County
- Southern California Coastal Water Research Project
- San Diego River Conservancy
- The San Diego River Park Foundation
- San Diego Regional Water Quality Control Board
- San Diego State University
- AMEC Environmental

The stakeholder workgroup identified three questions that would both demonstrate the utility of the regional flow–ecology relationships and inform local management decisions.

1. How will future land use changes affect flow conditions and impact biological endpoints in the San Diego River watershed? This involves a comparison of the current hydrologic conditions to those that would be expected under a 2050 land use scenario.

2. How can flow–ecology relationships be used to prioritize regions of the watershed into various flow management classes that can inform future planning decisions?

3. What are the biological implications of two future management decisions that will affect in-stream flow conditions?
    a. reduced discharge from Santee Lakes Reservoir due to increased capture and storage to meet demand for reclaimed water
    b. disconnecting imperviousness, and implementing stormwater capture strategies in a currently developed portion of the watershed

Regional ELOHA (flow-ecology) analysis

The local management questions were addressed using regional flow-ecology relationships conducted for southern California that relates changes in stream health to changes in hydrology. Stream health was assessed using the California Stream Condition Index (CSCI), a statewide index of benthic macroinvertebrates community composition. Hydrologic alteration was assessed based on a series of hydrologic metrics, which were shown to have strong statistical and ecological relationships with the CSCI (Mazor et al. in review). Metrics were also selected to ensure representation of different components of the flow regime (e.g. duration, magnitude, etc.) and different climate conditions (e.g. wet vs. dry vs. average years). Because we lack measured flow data for both current and historic conditions at most bioassessment sites, both were estimated using watershed models.

Regional benthic macroinvertebrate data were obtained from the southern California regional bioassessment program (Figure 2, Mazor 2015). A total of 799 wadeable stream sites were sampled between 2008 and 2014 using a probabilistic sample design. Sites were randomly distributed across the entire stream network using a spatially balanced generalized random-tessellation design that ensured representation across all natural and anthropogenic gradients in the region (Stevens and Olsen 2004).

Benthic macroinvertebrates were collected using protocols described by Ode (2007). At each transect established for physical habitat sampling, a sample was collected using a D-frame kicknet at 25, 50, or 75% of the stream width. A total of 11 ft2 (~1.0 m2) of streambed was sampled. This method was identical to the Reach-Wide Benthos method used by EMAP (Peck et al. 2006). However, in low-gradient streams (i.e., gradient <1%), sampling locations were adjusted to 0, 50, and 100% of the stream width, because traditional sampling methods fail to capture sufficient organisms for bioassessment indices in these types of streams (Mazor et al. 2010). Benthic macroinvertebrates were collected and preserved in 70% ethanol, and sent to one of five labs for identification. At all labs, a target number of at least 600 organisms were removed from each sample and identified to the highest taxonomic resolution that could be consistently achieved (i.e., SAFIT Level 2 in Richards and Rogers 2006); in general, most taxa were identified to species and Chironomidae were identified to genus.

Figure 2. Locations of bioassessment sites used to support the regional flow-ecology analysis

Benthic macroinvertebrate data was used to calculate the California Stream Condition Index (CSCI; Mazor et al. 2016). The CSCI is a predictive index that compares observed taxa and metrics to values expected under reference conditions based on site-specific landscape-scale environmental variables, such as watershed area, geology, and climate. It includes two components: a ratio of observed-to-expected taxa (O/E) and a predictive multi-metric index (MMI) made up of 6 metrics related to ecological structure and function of the benthic macroinvertebrate assemblage. Because the CSCI and all of its components are based on site-specific reference expectations, scores are minimally influenced by major natural gradients. Therefore, CSCI scores, by definition, compare existing to reference conditions and can be used as a measure of biological alteration (delta B) under anthropogenic stress. CSCI scores and all components were classified as indicating "intact" or "altered" condition, using the normal approximation of the 10^{th} percentile of CSCI reference calibration scores as a threshold (Mazor et al. 2016). For the CSCI, this equates to a score of 0.79 (where 1 is the reference expectation) as the threshold between biologically intact and altered.

Hydrologic alteration was modeled at 584 of the 799 bioassessment sites using HEC-HMS (ACOE 2000). The remaining 215 bioassessment sites were dropped from the analysis because the rainfall data at those locations was insufficient or did not meet quality control criteria for use in model development. Past studies have assessed hydrologic alteration based on empirical observations, often using a space for time substitution (i.e. comparing distinct hydrologically intact vs. altered locations instead of comparing hydrologic change over time). Modeling provides a mechanism to estimate hydrologic alteration at any location where biological data is available, thereby allowing larger data sets to be included in flow-ecology analysis (DeGasperi et al. 2009). Given the size of the southern California data set (584 sites), there was a need to balance the desire to model hydrologic alteration with the practical considerations of needing a tool that could be readily applied to a high number of sites (Sengupta et al. in review). HEC-HMS provides the ability to produce a continuous time series of estimated flow through parameterization of relatively small number of variables in the model (HEC-HMS manual version 4.1, Xuefeng and Steinman 2009).

A set of 26 HEC-HMS models was developed as part of the regional flow-ecology analysis to represent the range of watershed conditions present in the region. Therefore, one of the 26 models can be applied to produce a daily flow time series for every bioassessment site based on basin properties draining to that site. This obviates the need to develop a unique model for every site. Inputs used to develop and parameterize the models are grouped in three categories (Table 2): 1) watershed-specific data (e.g., area, and imperviousness), 2) site-specific data (e.g., observed flow, precipitation) and 3) model-specific parameters (e.g., initial loss, number of reservoirs).

Table 2. Parameters used to develop HEC-HMS models for application to the regional bioassessment sampling sites. Parameters in bold were adjusted during simulation of natural conditions at each site.

	HEC-HMS Method	Parameters
Watershed Specific		Area
		Imperviousness
		Time of concentration
Site Specific		Observed flow
		Observed precipitation
Model Specific	Simple Canopy	Maximum Storage (in)
		Initial Storage (%)
	Simple Surface	Maximum Storage (in)
		Initial Storage (%)
	Deficit and Constant (Loss)	Initial Deficit (in)
		Maximum Deficit (in)
		Constant Rate (in/hr)
	Clark Unit Hydrograph (Transform)	**Time of Concentration (hr)**
		Storage Coefficient (hr)
	Linear Reservoir (Baseflow)	Ground Water (GW) 1 Initial Discharge (cfs)
		GW 1 Storage Coefficient (hr)
		# of GW 1 Reservoirs
		GW 2 Initial Discharge (cfs)
		GW 2 Storage Coefficient (in)
		# of GW 2 Reservoirs

Each model was sequentially calibrated for four criteria: visual hydrograph match, Nash-Sutcliffe efficiency (NSE), percent low flow days, and Richard-Baker Index of flashiness. These calibration endpoints were selected based on relevance for supporting the instream biological communities (Konrad and Booth 2005, Morley and Karr 2002). Calibrating to all four measures produced models tuned to simulate flow conditions relevant for supporting in-stream biological communities. Models were calibrated for a 3-year period and were then validated for temporal and spatial performance. For temporal validation, the calibrated models were run for years outside of the calibration period and matched with the observed flow data. In all cases, model performance (as measured by NSE) during the validation period was within 15% of performance during the calibration period.

To evaluate spatial performance, we applied statistical 'jackknifing' to all calibrated gages. In this analysis, each modeled gage is treated as an 'ungaged' site, and the remaining 25 models are used to predict flows at that site. The models were fitted to the 'ungaged' site by inputting watershed-specific data and model-specific parameters, but without changes to the model-specific parameters. These simulations were run for the 3-year calibration period. Approximately 75% of the sites had an acceptable NSE value higher than 0.5 (Moriasi et al. 2007). A final validation was performed by comparing modeled output to measured flow at 16 bioassessment sites with nearby flow gages (but not included in the model development). At 11 of the sites, the R^2 values averaged 0.61; the range varied from 0.20 to 0.95. Further details on the model validation for accuracy and bias are found in Sengupta et al. (in review).

One of the 26 validated models was assigned to each of 584 bioassessment sites with adequate rainfall data in the southern California region based on similarity of watershed characteristics that were associated with observed hydrology. The assignment was done with a model-selection tool built by 1) classifying the models into 8 clusters based on observed flow metrics; 2) creating a random forest model to predict cluster membership based on watershed characteristics (i.e., elevation maximum and range, mean annual temperature, watershed area, mean catchment-wide summer precipitation, and soil erodibility factor); and 3) calculating proximity values (i.e., the frequency that a site and a model are predicted to be in the same cluster) between novel sites and each of the 26 models. For each bioassessment site, the model with the highest proximity value was selected for further analysis. Details about the development of the model-selection tool, and its performance, are provided in Sengupta et al. (in review).

The watershed models were used to produce an hourly time series of flow for a period of 23 years (1990 - 2013) for the 584 bioassessment sites. A subset of 6 years was selected for each site to calculate specific flow metrics. The six years were chosen to include two wet, two dry, and two average rainfall years based on long-term climate records. The six years were also selected based on the availability of high quality, complete rainfall records (i.e. no missing values or apparent anomalies). A challenge of the ELOHA approach is the need to compare current hydrologic conditions to reference in order to estimate hydrologic change (delta H). Because we seldom have data on historical flows, we rely on modeling to estimate reference conditions. Hourly hydrographs were estimated for both current and reference conditions at each site following Sengupta et al. (in review). Hourly hydrographs were aggregated to daily discharge, and a suite of flow metrics that represent different aspects of flow were calculated for both current and reference conditions (Table 3) Metrics were calculated for wet, dry, and average precipitation conditions, as well as for all 6 years combined. Metric-precipitation combinations that validated poorly (i.e., $r^2 < 0.25$) with observed flow were excluded from further analysis. This resulted in a total of 116 metric–precipitation combinations for analysis. For each metric–precipitation combination, hydrologic alteration (delta H) was characterized as differences between simulated current and reference conditions. “Reference condition” was estimated by adjusting model parameters to reflect undeveloped watershed conditions. Delta H for magnitude metrics was normalized by reference condition or 0.0283 cms (whichever was larger) to account for the effect of catchment size on discharge magnitude. Details on the hydrological analysis and modeling approach can be found in Appendix A.

Table 3. Flow metrics sorted by metric type and period of evaluation. O=overall, W=wet years, A= average years, D=dry years. Unless otherwise noted, metrics are from Konrad et al. (2008), Konrad, personal communication, Colwell (1974), or Bledsoe (personal communication).

Metric			Description		O	W	A	D
Duration								
	LowDur	days/event	Median annual longest number of consecutive days that flow was less than or equal to the low flow threshold		•		•	•
	HighDur	days/event	Median annual longest number of consecutive days that flow was greater than the high flow threshold			•		•
	NoDisturb	days/year	Median annual longest number of consecutive days that flow between the low and high flow threshold		•	•	•	•
	Hydroperiod	proportion	Fraction of period of analysis with flows		•		•	•
	Per_LowFlow	proportion	Percent of time with flow below 0.0283 cms		•	•	•	•
Frequency								
	HighNum	events/year	Median annual number of events that flow was greater than high flow threshold, an event is a continuous period when daily flow exceeds the threshold		•	•	•	•
	FracYearsNoFlow	proportion	Fraction of years with at least one no-flow day		•			
	MedianNoFlowDays	days/year	Median annual number of no-flow days		•	•	•	•
Magnitude								
	MaxMonthQ	cms	Maximum mean monthly streamflow		•	•	•	•
	MinMonthQ	cms	Minimum mean monthly streamflow		•	•	•	•
	Q01	cms	1st percentile of daily streamflow		•	•	•	•
	Q05	cms	5th percentile of daily streamflow		•	•	•	•
	Q10	cms	10th percentile of daily streamflow		•	•	•	•
	Q25	cms	25th precentile of daily streamflow		•	•	•	•
	Q50	cms	50th percentile of daily streamflow		•	•	•	•
	Q75	cms	75th precentile of daily streamflow		•	•	•	•
	Q90	cms	90th precentile of daily streamflow		•	•	•	•
	Q95	cms	95th precentile of daily streamflow		•	•	•	•
	Q99	cms	99th percentile of daily streamflow		•	•	•	•
	Qmax	cms	Median annual maximum daily streamflow		•	•	•	•
	Qmean	cms	Mean streamflow for the period of analysis		•	•	•	•
	QmeanMEDIAN	cms	Median annual mean streamflow		•	•	•	•
	Qmed	cms	Median daily streamflow		•	•	•	•
	Qmin	cms	Median annual minimum daily streamflow		•	•	•	•
Timing								
	C_C	ratio	Colwell's constancy (C) a measure of flow uniformity.		•	•	•	•
	C_CP	ratio	Colwell's maximized constancy (C/P). Likelihood that flow is constant through the year		•	•	•	•

Metric			**Duration**					
	C_M	ratio	Colwell's contingency (M). Repeatability of seasonal patterns.		•	•	•	•
	C_MP	ratio	Colwell's maximized contingency (M/P). Likelihood that the pattern of high and low flow events is repeated across years.		•	•	•	•
	C_P	ratio	Colwell's predictability (P=C+M). Likelihood of being able to predict high and low flow events		•	•	•	•
	MinMonth	month	Month of minimum mean monthly streamflow				•	•
	MaxMonth	month	Month of maximum mean monthly streamflow			•		
Variability								
	RBI	Unitless	Richard Baker Index (flashiness)		•	•		•
	SFR	proportion	90th percentile of percent daily change in streamflow on days when streamflow is receding (storm-flow recession)				•	
	QminIDR	cms	Interdecile range of annual minima		•			
	QmeanIDR	cms	Interdecile range of annual means		•			
	QmaxIDR	cms	Interdecile range of annual maxima		•			

Hydrologic thresholds that result in biological response were evaluated for each flow metric-precipitation condition combination, based on nine biological response variables (i.e. the CSCI and its component metrics). Hydrologic metrics were evaluated for overall climatic conditions, as well as for wet, dry, or average precipitation years. The 116 metric-precipitation condition combinations were used to predict each of the nine biological response variables in boosted regression tree models using the gbm package in R (Ridgeway 2015, R Core Team 2016), and the importance of each predictor was ranked (Friedman 2001). Ranks were averaged across all models, and the best ranked precipitation condition within each metric was selected for further analysis. Ecologically derived targets were then established for each flow metric. Further detail about modeling biological responses to hydrologic alteration can be found in Mazor et al. (in review).

In order to set targets for hydrologic metrics based on biological response, we developed logistic regression models of the probability of healthy biological condition as a function of different levels of hydrologic alteration. Targets were set at the level of hydrologic alteration where the probability of healthy biological condition was 50% of the probability at hydrologically unaltered sites. It is important to note that these targets do not represent reference conditions. Increasing and decreasing gradients of hydrologic alteration were analyzed independently against each biological response variable. Across all biological response variables, the most conservative target was selected for further analysis. Logistic regression models were created using the glm function in R with a binomial error distribution and a logit link function (R Core Team, 2016). Metrics were scored 0 if they met targets, 1 if they failed targets, and 2 if they failed targets by more than twice the target value (Figure 3).

Metric	100% above UT	Upper Threshold (UT)	Within Threshold	Lower Threshold (LT)	100% below LT
Flow (Q min cfs)	0.04	0.02		-0.02	-0.04
Value	1.2	0.03	0.01	-0.025	-0.6
Score	2	1	0	-1	-2

Figure 3. Example scale for assigning hydrologic alteration scores

An objective of the regional flow-ecology analysis was to identify a subset of priority flow metrics that can be used to inform management actions. Metrics were prioritized based on the following criteria (Mazor et al. in review):

- Differentiate hydrologic condition at reference sites vs. altered sites
- Have the strongest relationship to biological condition based on boosted regression tree analysis and can produce a hypothesized ecological response
- Can be modeled under both current and reference conditions with a high level of confidence
- Are amenable to management actions and are expected to respond in predictable ways to deliberate changes in flow conditions
- Have minimal redundancy with other metrics; the goal is to select metrics that represent different components of the hydrograph (e.g. magnitude vs. duration)

Based on these criteria and the logistic regression analysis described above, Mazor et al. (in review) identified seven priority flow metrics and associated thresholds of biological response (Table 4). The importance of the seven priority flow metrics varied by climatic condition, with some metrics only being important during certain precipitation conditions (Table 4). Using a subset of metrics has the advantage of allowing management actions to focus on controlling a reasonable set of flow properties that will have the greatest biological effects, as opposed to trying to manage for all 116 metric-precipitation combinations.

Table 4. Priority hydrologic metrics and associated thresholds used in the regional flow-ecology relationships. Metrics are grouped by the hydrograph component they represent. Thresholds are expressed as the change in metric value (delta H) associated with poor biological condition (CSCI <0.79). Metric effects on biology were typically strongest during either average, wet, or dry rainfall years, or all years combined (overall). NT= no threshold established.

Hydrograph Component	Metric	Metric Definition	Critical precipitation condition	Decreasing Threshold	Increasing Threshold
Duration	NoDisturb (days)	median annual longest number of consecutive days that flow is between the low and high flow threshold	Average	-64	NT
	HighDur (days/event)	median annual longest number of consecutive days that flow was greater than the high flow threshold	Wet	-3	24
Magnitude	MaxMonthQ (m3/s)	Maximum mean monthly streamflow	Wet	NT	1.5
	Q99 (m3/s)	streamflow exceeded 99% of the time	Wet	-0.01	32
Variability	RBI (unitless)	Richards-Baker index of stream flashiness	Dry	NT	0.25
	QmaxIDR (m3/s)	Interdecile range of flow	Overall	-5	2.5
Frequency	HighNum (events/year)	median annual number of events that flow was greater than high flow threshold	Dry	NT	3

Application of regional flow-ecology (ELOHA) relationships to guide watershed management actions

Current and future watershed hydrologic condition was evaluated for 52 distinct catchments defined by major stream nodes (Figure 4). For each catchment, we simulated current and reference hydrology using the most appropriate of the regional HEC-HMS models. Hydrologic alteration (delta H) was calculated for the seven priority flow metrics shown in Table 4, and each metric was scored based on its distance above or below the established threshold (Figure 3). To provide an easy way to convey general hydrologic condition, an overall composite hydrologic condition score was developed by adding the absolute values of the score for each individual metric. The hydrologic condition score ranged from 0-14 because each of the metrics can receive a score between zero and two depending on how far the score is from the threshold (see Table 3). This approach assumes that each metric is of equal importance and that positive or negative changes in metric values have comparable effects. Scores were binned into four categories as shown in Table 5 and each of the 52 catchments was assigned an A – D designation, representing its overall hydrologic condition.

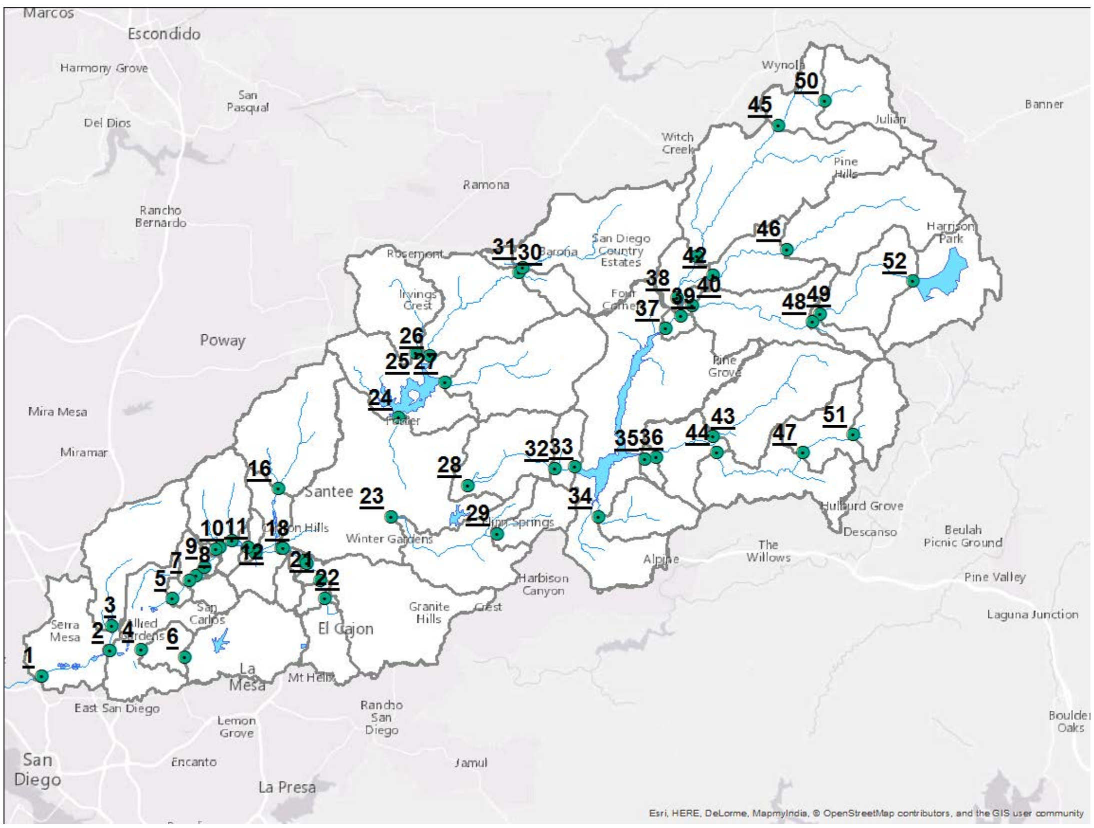


Figure 4. Individual catchments used for watershed analysis

Table 5. Definition of hydrologic condition score (0-14) based on how far each of the priority metrics is from its threshold value. See Figure 3 for explanation of scoring.

Overall Hydrologic Condition Score	Ranges of Metrics Above or Below Threshold
A	0
B	1-2
C	3-6
D	7-14

Flow management classes were assigned to each of the 29 locations where previous bioassessments had been completed. Sites were assigned to one of four classes based on their biological and hydrological status. Biological status was inferred using CSCI scores: Sites with scores greater than 0.79 were designated as biologically intact, and sites with lower scores were designated as biologically altered (Mazor et al. 2016). Hydrological status was assigned using the composite hydrologic condition score described above. Classes A and B were considered hydrologically unaltered when assigning sites to different management classes.

Hydrologically unaltered and biologically unaltered sites were put into a “protection” class; the good conditions at these sites should be protected from further designations. Hydrologically altered and biologically unaltered sites were put into a “monitoring” class; these sites may be resilient to stressors related to hydrologic alteration, but factors related to this apparent resiliency should be monitored to ensure that they continue to support biological health. Hydrologically altered and biologically altered sites were put into a “flow management” class; these sites should undergo a causal assessment to determine if flow management is likely to improve biological condition or if other constraints (e.g., channelization) may limit the ability of a stream to respond to improved flows. Hydrologically unaltered and biologically altered sites were put into an “other management” class; these sites should also undergo causal assessments, but other management options should be prioritized over flow management, such as habitat or water quality improvements (Table 6). Potential additional causes of biological alteration were evaluated for all locations where the CSCI was less than 0.79 based on additional stressor data such as water chemistry and physical habitat assessments that are routinely collected as part of the regional ambient monitoring programs (Mazor 2015).

Table 6. Management categories defined based on combination of hydrologic and biologic alteration

	Poor hydrologic condition	Good hydrologic condition
Poor biology (CSCI < 0.79)	***Flow Management***: Evaluate hydrologic alteration among other stressors. Determine relative importance of flow management for improving biological condition, relative to other stressors.	***Other Management/Causal Assessment***: Evaluate other stressors to determine cause of poor biology. Evaluation of flow management not recommended.
Good biology (CSCI > 0.79)	***Monitor***: Communities may be resilient to flow alteration. Continue to monitor for factors that may reduce resilience.	***Protect:*** Intact area. Target for preservation. Explore factors that may contribute to resilience or vulnerability.

Following the watershed mapping, the stakeholders prioritized management questions and scenarios for setting flow targets aimed at protecting (or recovering) instream biological health (as measured by CSCI). The scenarios retained for detailed analysis were selected based on consensus of the workgroup and represented a range of different management situations (e.g. reservoir operation, effluent recycling, and stormwater management). The most appropriate model was selected for each priority scenario using the model selection tool (described above) and was used to simulate both current hydrology and future hydrology based on the proposed management action. Future conditions largely consisted of changes in reservoir discharge, runoff capture, or reduction in impervious cover (i.e. low impact development). The subset of seven priority flow metrics based on the regional flow ecology analysis was calculated for each scenario (see Table 4). The projected delta H for each scenario (and each alternative within a scenario) was evaluated relative to the flow–ecology relationships and thresholds developed by the regional analysis. To aid in management interpretation of the results of the scenario analysis, the regional thresholds, which are expressed as *change in the metric value* were converted to the actual target values specific for the situation of the case study. The results of this analysis were used to develop flow management recommendations for each scenario. Ultimately, these flow recommendations should be considered in concert with other management needs for the watershed.

RESULTS AND DISCUSSION

Effect of future land-use change on hydrologic condition

To address the question, "***how will future land-use changes affect flow conditions and impact biological endpoints in the San Diego River watershed?"*** we compared the current overall hydrologic condition to the expected future condition based on 2050 SanGIS land-use projections, assuming no installation of stormwater control device or low impact development features.

Under current conditions, 17 of the 52 catchments (33%) scored in the worst two categories of hydrologic alteration, while 35 of 52 (67%) scored in the least hydrologically altered category (Table 7). There appears to be a spatial gradient of hydrologic condition in the watershed, with the most hydrologically intact areas are in the upper watershed, where much of the land is in public ownership and/or there is currently little urban development. Catchments in the poorest hydrologic condition are concentrated in the lower watershed where most of the current development exists. These areas are also downstream of all the reservoirs in the watershed (Figure 5).

Table 7. Distribution of hydrologic alteration scores under current conditions ("A" is least altered, "D" is most altered).

Category	# of catchments	Proportion of catchments
A	25	48%
B	10	19%
C	6	12%
D	11	21%

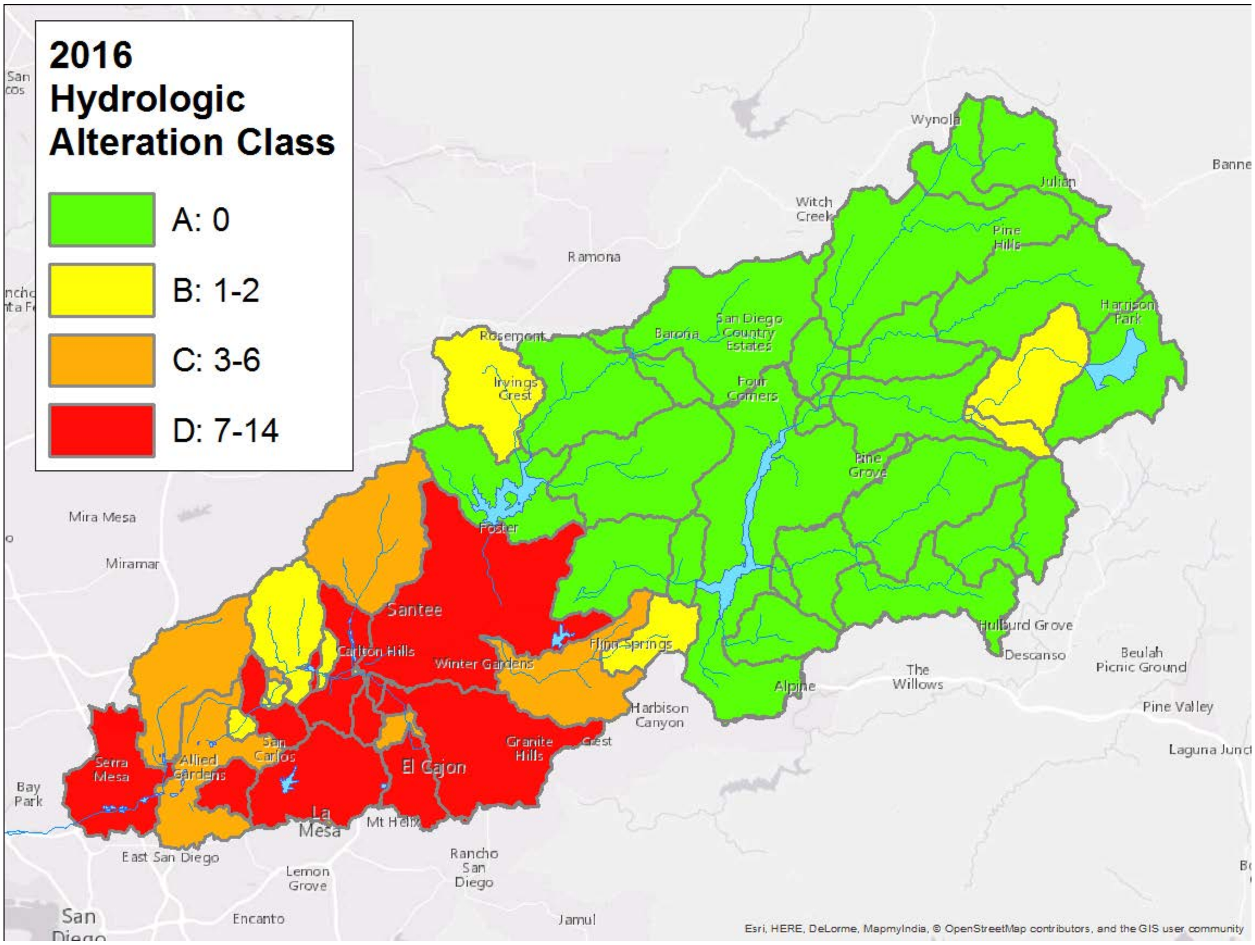


Figure 5. Hydrologic condition of each of the 52 catchments in the watershed. Numbers indicate the number of metrics that failed to meet the designated threshold.

We evaluated all 35 flow metrics in order to provide additional information about the type of hydrologic alteration occurring in each catchment (Figure 6). Catchments that are hydrologically unaltered (Classes A and B in Figure 6) generally “failed” less than 10% of the overall set of 35 metrics. This suggests that the targeted set of metrics used in Figure 5 (based on our screening filters described above) is representative of overall hydrologic condition. The most commonly exceeded metrics range across nearly all categories: duration metrics (e.g. high duration), magnitude metrics (e.g. Q95), frequency metrics (e.g. HighNum), and variability metrics (e.g. RBI). This suggests that when hydrologic alteration occurs, it tends to affect most aspects of runoff hydrographs rather than preferentially influencing certain hydrologic elements.

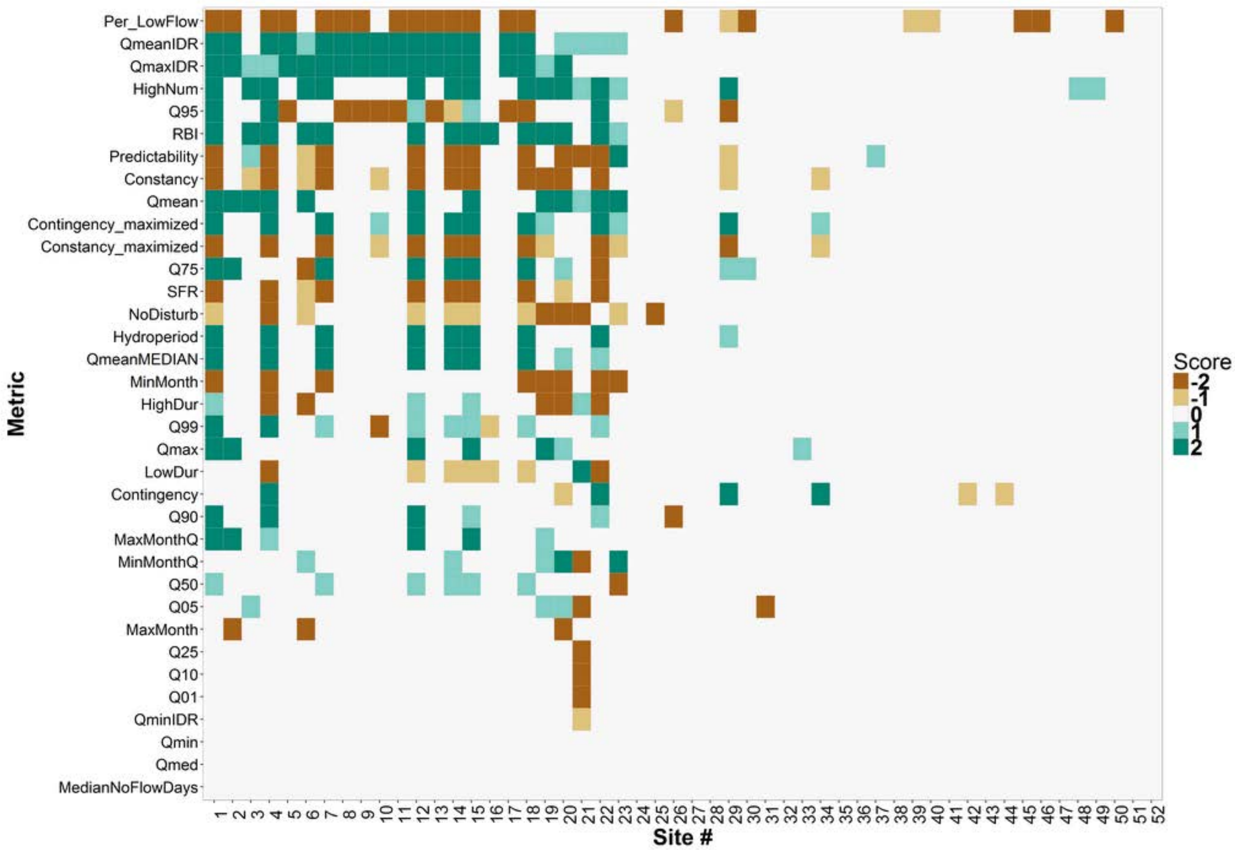


Figure 6. Heatmap showing hydrologic metric scores for all catchments and all metrics. Catchment numbers/positions on the x-axis are based on the catchment positions shown in Figure 4.

Hydrologic condition was generally related to catchment imperviousness (Figure 7). In most cases, severe hydrologic alteration was associated with total impervious cover greater than 5%. In all cases, hydrologically unaltered catchments (Classes A and B) had less than 5% total impervious cover, often only 1-2%. This is consistent with past studies that have shown hydrologic and geomorphic responses associated with modest increases in total impervious cover (Hawley and Bledsoe 2011, Vietz et al. 2016).

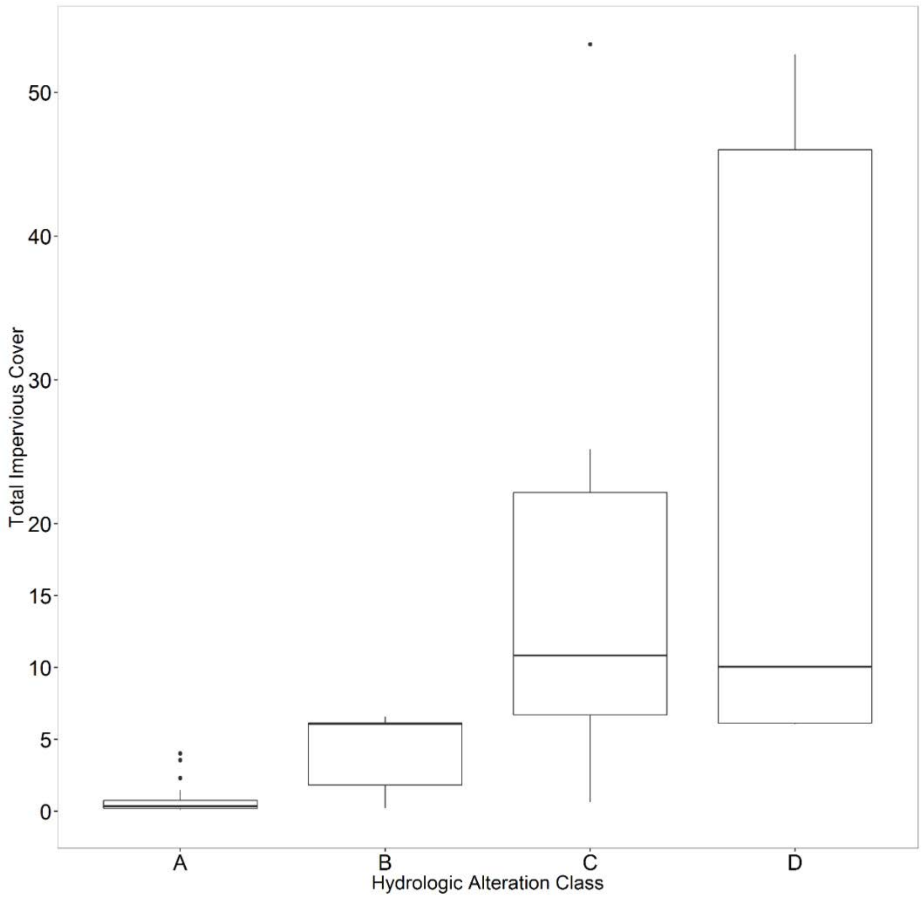


Figure 7. Relationship of hydrologic condition class and percent total impervious cover in the contributing catchment.

Under 2050 land use projections, hydrologic conditions of the watershed are expected to degrade, mainly in the middle portion of the watershed (Figure 8). Mid watershed catchments, around existing reservoirs, are expected to degrade the most in association with future land use changes, with several catchments going from Class A to Class C. Little change is expected in the upper watershed since many of the catchments in the upper watershed are hydrologically unaltered, in public ownership and hydrologic conditions are expected to remain unaltered into the future. Most of the lower watershed is already in poor hydrologic condition and is expected to remain that way in 2050, unless substantial hydrological management and/or remediation measures are implemented. It is important to note that future conditions were modeled using the same precipitation values as the current and historical scenarios since reliable downscaled future precipitation values are not available. Furthermore, the future conditions assumed no stormwater control devices, low impact development or hydromodification management, since we have no information on where/how these will be installed in the future. Therefore, the results of the 2050 analysis should be considered a worst-case scenario.

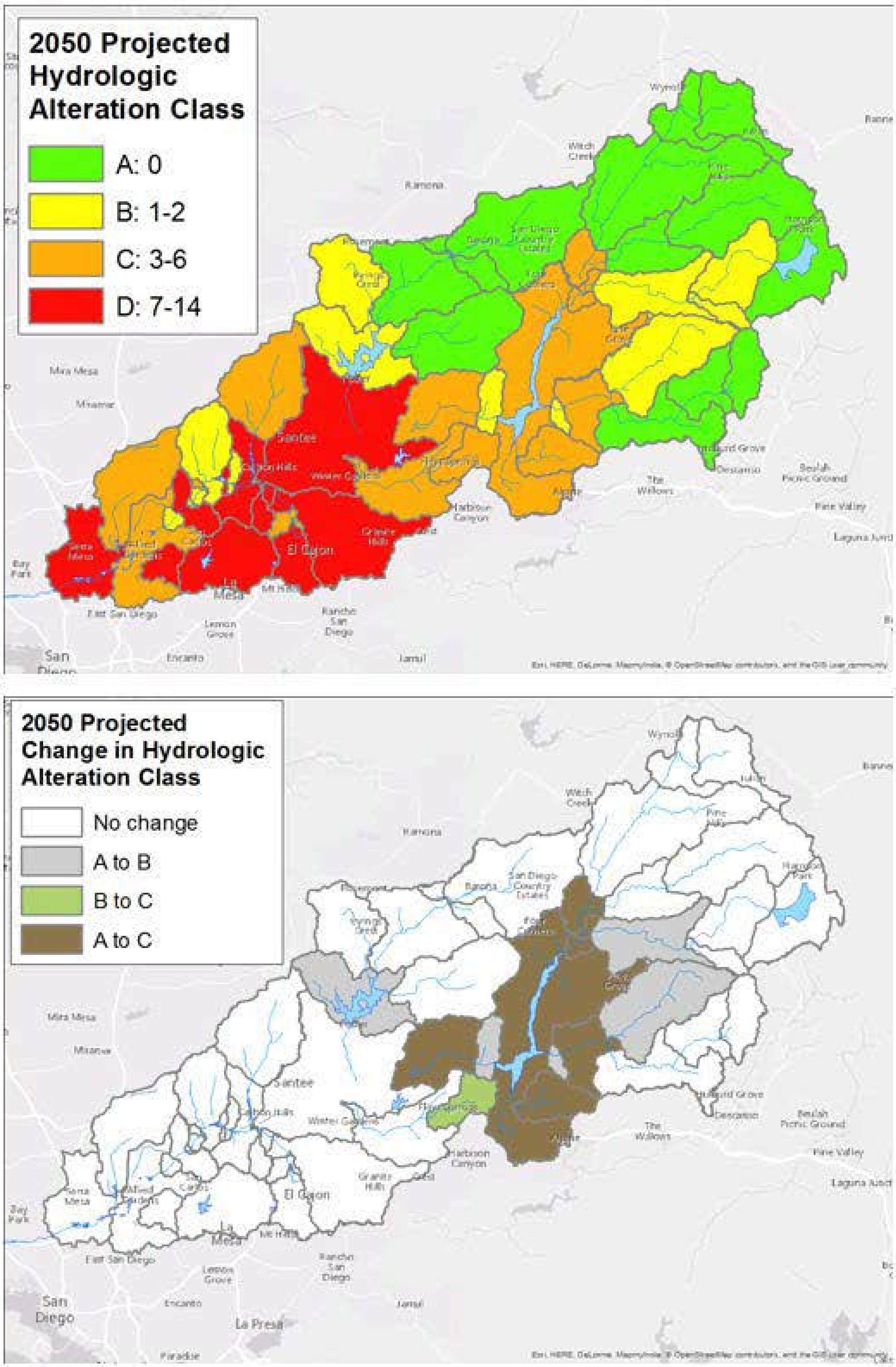


Figure 8. Overall hydrologic condition under 2050 projected land use (top) and change in hydrological condition between 2015 and 2050 (bottom). Categories are defined as in Figure 5.

Future land use changes were associated with sufficient hydrologic alteration to affect all seven metrics that contribute to the overall hydrologic rating. Of the seven metrics, QmaxIDR (a measure of flow variability), Q99 (a measure of high flow magnitude), and HighNum (a measure of the frequency of high flow events) were affected in the greatest number of catchments, and therefore most responsible for the predicted changes in overall hydrologic condition. Changes in these hydrologic metrics are associated with changes in biological condition; this suggests that future hydrologic changes are likely to result in declines in the condition of instream biological communities.

Prioritization of areas for various management actions

To address the question, "***How can flow–ecology relationships be used to prioritize regions of the watershed into various flow management classes that can inform future planning decisions***?" we compared the overall hydrologic condition scores to the CSCI scores at the 29 locations in the watershed where bioassessment has previously occurred.

The majority of upper watershed sites were considered intact, with unaltered hydrology, and therefore a high priority for protection (Figure 9). Candidate areas for flow management were focused in the lower portion of the watershed where both hydrology and biological condition were altered.

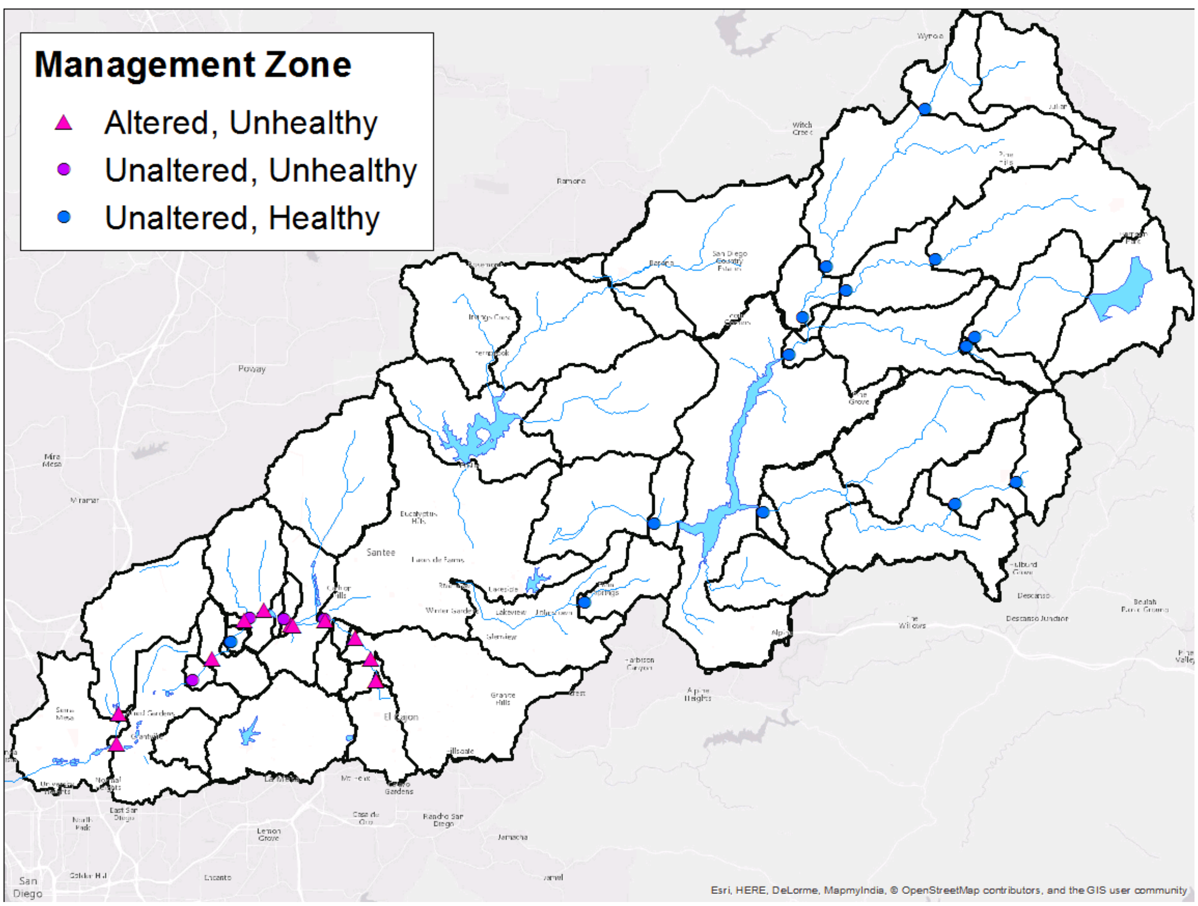


Figure 9. Management categories for bioassessment sites based on combinations of hydrologic and biologic alteration. Only three of the four possible management categories were present in the San Diego River watershed. There were no sites with altered hydrology and healthy biological communities.

Considering both the flow management zones and information available on water quality, habitat, and channel condition from ambient survey data allows us to provide specific management recommendations that can be prioritized for each location (Figure 10). We estimate that flow alteration is the primary factor affecting biology at only 3 of the 13 biologically degraded sites in the lower watershed. At all other sites, flow management should be coupled with habitat or water quality remediation in order to improve biological conditions. The lower watershed was largely in poor biological condition with altered hydrology, making flow management a good option to consider for improving watershed health. However, many of the sites in this category had highly developed floodplains or concrete-lined channels, and all lower watershed sites had poor water quality, as indicated by low scores on the diatom (D18) or soft algae (S2) indices of biotic integrity (Table 8). Therefore, flow management should always be considered in conjunction with other forms of management that address water-quality impacts and alterations to physical habitat. Flow management alone is most likely to improve biological health at sites where habitat is in poor condition, but the channel is unlined and the immediate floodplain is undeveloped. At such sites, the stream form has good capacity to respond to changes in flow, creating the microhabitat structure that supports diverse benthic macroinvertebrate assemblages. In contrast, flow management alone is unlikely to improve sites with armored banks, or where floodplain development limits the capacity of the stream form to respond. In these cases, flow management should be considered in conjunction with habitat restoration efforts that remove these constraints. At lower watershed sites with relatively good condition habitat, other stressors, such as poor water quality, may be responsible for poor biological condition; at these sites, flow management may improve water quality, but care should be taken to maintain good habitat that can support healthy instream biological communities. Finally, in one instance, two sites in close proximity were assigned to different management classes based on different models used to estimate hydrologic alternation. In this instance, we assumed the two sites were in similar condition and assigned the more conservative management class.

Table 8. Relationship of biologically unhealthy sites to water quality and physical habitat stressors.

Recommendation	Sites	Habitat quality	Bank armoring	Response capacity
Manage flows to improve habitat and water quality	3	Poor	Earthen	Moderate or good. Limited development in floodplain.
	13 and 14	Poor	None	Moderate or good. Limited development in floodplain.
Improve water quality	5, 11, 12, and 15	Good	None	Moderate or good. Limited development in floodplain. No channel armoring.
Remove habitat constraints	20, 21, 22	Poor	Concrete	Limited. Stream cut off from floodplain.
	2	Poor	Earthen	Limited. Stream cut off from floodplain.

.

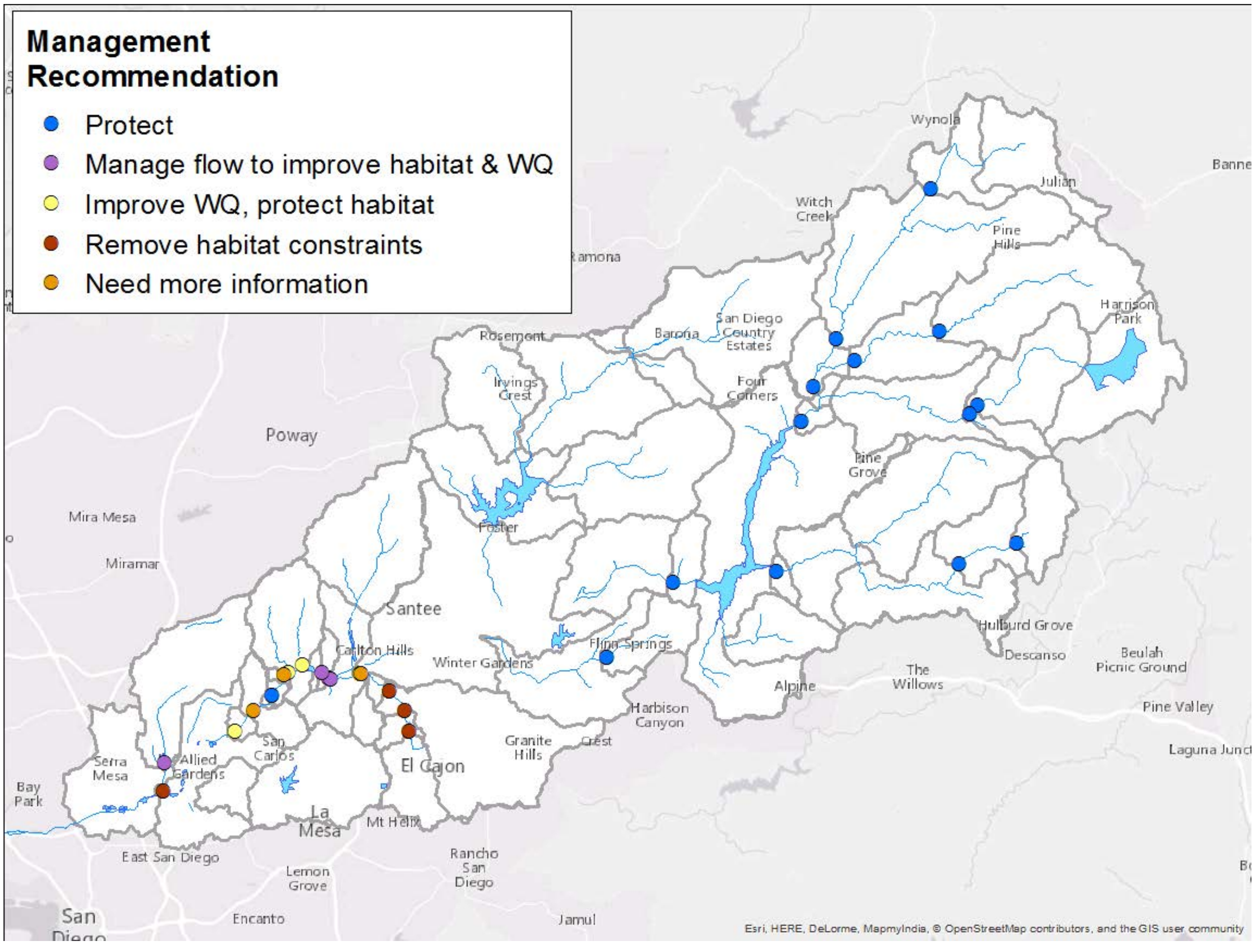


Figure 10. Recommended management actions for all sites where bioassessment has occurred. Recommendations are based on both flow-ecology information and available data on habitat and water quality obtained through the local regional monitoring program.

Evaluation of management scenarios

The stakeholder workgroup prioritized two future management scenarios for evaluation. Each of them represents potential actions that will affect in-stream flow conditions, and in turn may affect biological condition.

1. lower discharge from Santee Lakes Reservoir due to increased capture and storage to meet demand for reclaimed water
2. disconnecting imperviousness, and implementing stormwater capture strategies in a currently developed portion of the watershed

Results from each of the scenarios are described below:

Scenario 1. Lower discharge from Santee Lakes Reservoir

The Santee Lakes Reservoir receives treated wastewater from Padre Dam Municipal Water District's Ray Stoyer Water Recycling Facility (WRF). The lake releases the treated effluent to Sycamore Creek (which also receives water from a small rain-fed discharge from the lake). Future management scenario involves eliminating discharge of treated wastewater into the lakes and diverting it for reuse to help meet increased demands for recycled water. This will be associated with a proportional decrease in discharge from Santee Lakes Reservoir to Sycamore Creek (because there is less need to create capacity in the lakes); the rain-fed discharge will continue to be released to the creek (Table 9).

Table 9. Inflow into Santee Lakes Reservoir due to wastewater effluent and rainfall runoff. Values are total monthly discharge into the reservoir.

	Average Effluent Flow (Mgal)	Rain-Fed Discharge (Mgal)
January	43.00	2.30
February	33.08	2.73
March	37.60	1.76
April	22.65	1.56
May	12.88	1.31
June	4.91	0.00
July	3.13	2.27
August	2.88	0.00
September	11.25	1.51
October	17.09	0.71
November	28.92	2.79
December	42.24	4.17

Simulations of future scenarios using HEC-HMS indicate that the flow regime will continue to have natural variability, with lower magnitude of flows under the future management scenario relative to current conditions (Figure 11).

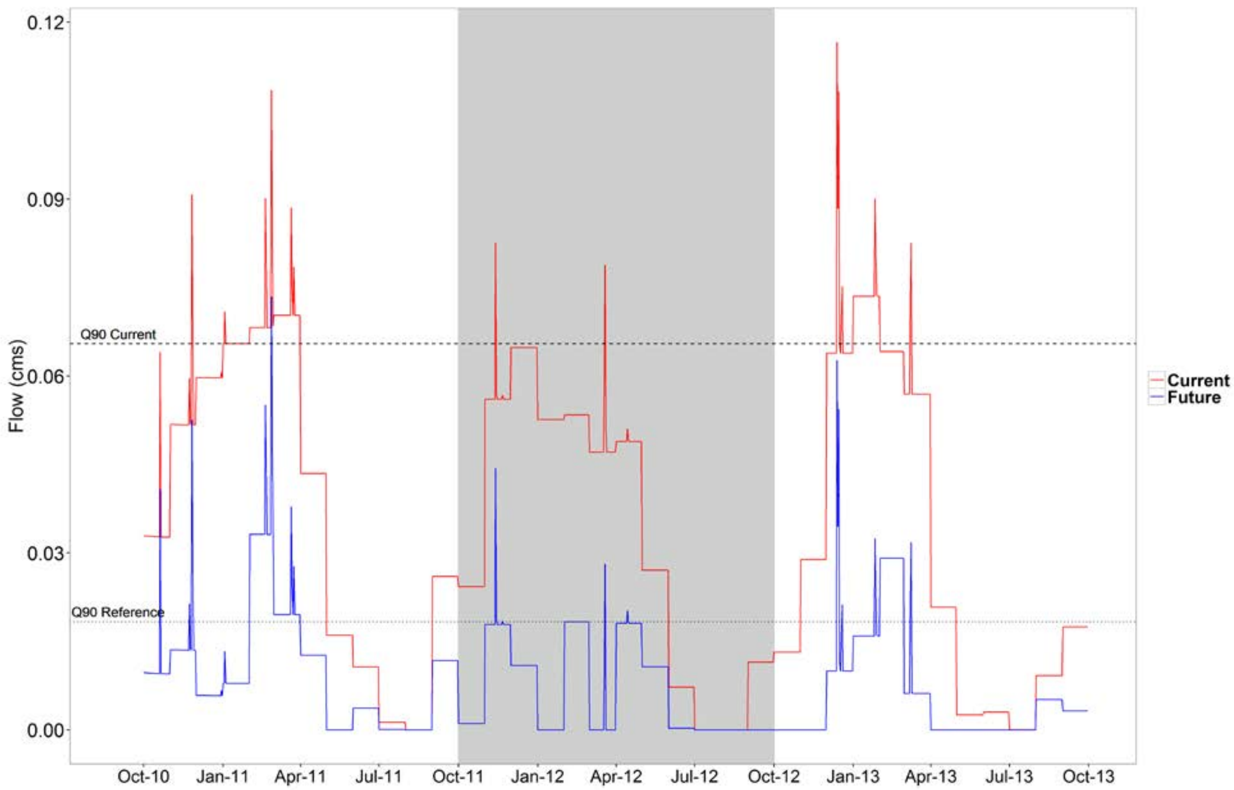


Figure 11: Modeled daily discharge under current and future scenarios at Sycamore Creek. The blue line represents the current scenario (which includes effluent discharge), and the orange line represents a future scenario where effluent is reused and not discharged into the creek.

Current conditions at Sycamore Creek are altered mainly in terms of the duration of high flow conditions (e.g. HighDur and NoDisturb). This reflects discharge from Santee Reservoir that elevates downstream high flow conditions. The balance of the priority flow metrics are currently meeting targets (Table 10). Under future scenarios, many high flow metrics are expected to improve in response to the removal of discharges from the reservoir. In contrast, the remaining metrics will remain at or slightly below the targets associated with healthy biological conditions. Failure to achieve these targets under future conditions likely reflects the effects of ongoing urban runoff, which will not be affected by changes in the reservoir operation. Overall the hydrologic condition in Sycamore Creek will improve under high flow conditions, but is likely to remain in degraded hydrologic and biological condition, even if discharges from Santee Lakes are eliminated following the proposed management scenario.

Providing clear objectives can aid in future desires to manage runoff and reservoir discharge in a manner that promotes healthy downstream biological communities. To assist in future management decisions, we developed the following specific management statements for the Santee Reservoir/Sycamore Creek scenario:

- <u>NoDisturb</u> = Maintain an average low flow between 0 and 0.02 cms (0.7 cfs) for a minimum of 119 days during the dry season

- HighDur = Maintain flow greater than 0.02 cms (0.7 cfs) for between 25 and 52 days per year
- MaxMonthQ = Maintain mean monthly flows below 0.1 cms (3.5 cfs).
- Q99 = Storm flows (or high flow events) should be between 0.03 cms (1 cfs) and 1.1 cms (39 cfs)
- HighNum = Ensure less than 4 high flow events per year with a flow greater than 0.02 cms (0.7 cfs)

Variability metrics do not lend themselves to directed management actions; therefore, we have not provided objectives for RBI or QmaxIDR. Instead these flow metrics should be used to evaluate the effectiveness of actions taken in response to the other metrics.

Table 10. Current and expected future hydrologic metric values in Sycamore Creek (SC) downstream of Santee Reservoir. The table presents site-specific targets that have been calculated based on the regional threshold values. Green cells represent conditions where flow targets would be met; yellow cells represent conditions where flow would be the same as the target value. NT =no target assigned.

Metric	Unit	Value		Target	
		Current	Future	Lower	Higher
NoDisturb	days	31	122	119	NT
HighDur	days/event	212	28	25.1	52.2
MaxMonthQ	cms	0.1	0.0	NT	0.1
Q99	cms	0.1	0.0	0.0	1.1
HighNum	events/year	1	4	NT	4
RBI	unitless	0.0	0.9	NT	0.3

Scenario 2. Impact of disconnecting imperviousness and implementing stormwater retention facilities in an urbanized catchment

Alvarado Creek catchment is located in the downstream portion of the San Diego River watershed. At an area of 14 sq. mi., and 50% total imperviousness cover, it is a heavily urbanized and hydrologically altered reach. We tested **two** scenarios in this sub-catchment: 1) effect of disconnecting imperviousness (modeled as a decrease in total imperviousness in the catchment), and 2) implementing stormwater retention facilities that can capture 85th percentile of a 24-hour rain event.

Disconnecting imperviousness decreases the extent of hydrologic alteration in the creek. However, flow metrics do not drop below levels associated with healthy biological communities until the total imperviousness is at or below 5% (Table 12). Analysis shows that for most metrics, there is a 50%

likelihood of meeting flow targets at 10% impervious cover, 66% likelihood at 5% impervious cover and finally an 80% likelihood of meeting flow targets at 2% impervious cover. Above 10% impervious cover, the likelihood of achieving flow targets declines by 15%. This is consistent with previous results that 5% impervious cover appears to be an important level or maintaining biologically protective levels of flow.

For the 85th percentile of a 24-hour storm event, based on a precipitation isohyetal developed for San Diego River watershed, any storm event with less than or equal to 0.75 inches (1.9 cm) is assumed to be 100 percent captured by the retention structures, resulting in no runoff (Table 11).

Providing clear objectives can aid in future desires to manage runoff and reservoir discharge in a manner that promotes healthy downstream biological communities. To assist in future management decisions, we developed the following specific management statements for the Alvarado Creek scenario.

- NoDisturb = Maintain an average low flow between 0 cms and 0.01 cms (0.4 cfs) for a minimum of 119 days during the dry season
- HighDur = Maintain flow greater than 0.01 cms (0.4 cfs) for between 27 and 56 days per year
- MaxMonthQ = Maintain mean monthly flows below 0.66 cms.(23 cfs)
- Q99 = Storm flows (or high flow events) should be between 0.2 (7 cfs) and 0.66 cms (23 cfs)
- HighNum = Ensure less than 4 high flow events per year with a flow greater than 0.01 cms (4 cfs)

As stated above, variability metrics do not lend themselves to directed management actions. Instead they should be used to evaluate the effectiveness of actions informed by the other metrics.

Table 11. Response of key metrics to changes in total impervious cover and 85% runoff capture. The table presents site-specific targets that have been calculated based on the regional threshold values. Green cells represent conditions where flow targets would be met. NT = no target assigned

Metric	Unit	Imperviousness					Capture	Target	
		2%	5%	10%	25%	50%	85th % storm	Lower	Higher
NoDisturb	days	32	32	32	32	31.5	32	119	NT
HighDur	days/event	35.5	34	32.5	24	9	8	27	56
MaxMonthQ	cms	0.31	0.35	0.41	0.59	0.88	0.53	NT	0.66
Q99	cms	0.19	0.45	0.89	2.04	4.04	2.64	0.2	0.67
HighNum	events/year	23.5	22.5	23.5	24	24	24	NT	4
RBI	unitless	0.22	0.47	0.75	1.15	1.4	1.39	NT	0.23

IMPLICATIONS AND RECOMMENDATIONS

The goal of this project was to demonstrate how regional flow-ecology relationships can be used to inform instream environmental flow properties necessary to meet ecological benchmarks as defined by measures of benthic macroinvertebrate community composition and structure. These target flows can be used to help establish goals for use in hydromodification management, nutrient numeric endpoints, and freshwater bioobjectives. They can also be used to develop performance targets for management actions, BMPs, etc. This case study allowed us to develop a framework for implementing regionally derived flow–ecology relationships to inform local management decisions. The stakeholder-focused process allowed us to identify technical and practical benefits and challenges associated with the approach that can inform future implementation efforts.

Utility of the regional flow-ecology approach based on the ELOHA framework

A major objective of this case study was to evaluate the ability to apply the flow–ecology relationships derived from the regional analysis to inform local watershed-scale decisions. Our results illustrate that several of the stated advantages of the ELOHA approach aid in such watershed-scale application. The ability to apply regionally derived flow thresholds to inform local decisions is a major advantage of the ELOHA approach. This eliminates the need to develop local flow–ecology relationships for every stream of interest, as is the case in more traditional instream flow methods (Beecher et al. 2010, McClain et al. 2014). The tools developed through the regional analysis provided readily transferable tools for local stakeholders to produce measures of hydrologic change (i.e., delta H) for any location of interest and to explore how those values would change under different land use or management scenarios. This had the dual benefit of allowing for robust analysis and providing a vehicle for stakeholder engagement in setting management priorities related to instream flow, an important cornerstone of the ELOHA approach.

Use of the predictive CSCI index in our regional flow-ecology analysis took advantage of the available bioassessment data and provided an easy way to provide measures of biological change (delta B), which has been a challenge for past ELOHA applications (e.g., McManamay et al. 2013). Developing the regional flow–ecology relationships and applying them at the local scale would not have been possible without the regional bioassessment data and the existence of the predictive scoring tool (Mazor et al. 2016). Large regional data sets provide sufficient sample size to develop statistically meaningful flow-ecology relationships in spite of the inherent “noise” in the data associated with other co-occurring factors that interact with flow to affect biological community condition (Solans and Jalon 2016). The predictive scoring tool is a measure of biological condition relative to expected reference conditions and thus provides a readily available measure of biological change (delta B) at every site. The availability of similar data and tools should be a major consideration for other efforts interested in developing similar regional approaches.

Other important elements of the ELOHA approach are the inclusion of a broad suite of hydrologic metrics that relate to ecologically relevant biological metrics through hypothesized flow–ecology relationships. Our seven priority flow metrics included two measures of magnitude, two of duration, two measures of variability, and one of frequency. This combination ensures that all elements of the hydrograph will be addressed through flow management. The selected metrics have hypothesized relationships that affect macroinvertebrate communities, allowing us to communicate their ecological relevance to managers and local stakeholders (Table 12). They are also amenable to management and minimize redundancy between metrics (Table 13). Interestingly, our metrics are similar to those identified by DeGasperi et al. (2009) who found that decreases in macroinvertebrate indices in urbanizing watersheds in the Puget Sound area of Washington were associated with changes to the number and duration of high and low flow events, and flow flashiness. It is important to note, however, that hypothesized relationships for both this study and other similar studies were derived through statistical analysis of regional bioassessment data sets. Additional mechanistic studies will be important to validate these relationships and confirm their ecological relevancy. As such studies are completed, they can be used to refine flow management targets based on improved understanding of the flow–ecology relationships.

Table 12. Hypothetical biological responses to alterations in six selected flow metrics

NoDisturb:

- Decrease: Times between spates and droughts are too short to support the expected abundance and diversity of long-lived taxa (e.g., semivoltine insects). Flood-dependent reproducers (e.g., cottonwoods) have fewer opportunities to establish. Good recolonists (drifters, strong fliers, exiters) will flourish.
- Increase: Long-lived taxa are able to out-compete taxa that reproduce quickly or recolonize.

HighDur:

- Decrease: Reduced time with floodplain access, reducing floodplain subsidies to fish and inverts, and diminishing time for riparian seedlings to establish.
- Increase: Desiccation resistance is less useful. More opportunities for aerial colonization (good fliers)

HighNum:

- Decrease: Fewer flushing flows. Allows more clogging of substrate and encroachment of macrophytes. Reduction of spawning gravels for fish. Deposition will fill pools. Greater accumulation of algae may lead to increased grazing.
- Increase: More scouring flows. More incision and bank erosion, leading to mortality of riparian vegetation. Direct mortality of long-lived organisms may eliminate semivoltine taxa.

Q99 and MaxMonthQ:

- Decrease: Reduces size of flushing flows, allowing more clogging of substrate and encroachment of macrophytes. Reduction of spawning gravels for fish. Deposition will fill pools. Greater accumulation of algae may lead to increased grazing. More desiccation-resistant taxa. More predation, and more predation-resistant (armored, or quick reproducers) taxa.
- Increase: Greater scour, leading to incision and bank erosion. Riparian vegetation mortality will increase, both through bank failure and lowering of the water table. Greater flushing of leaf litter will lead to a decline in shredders.

QmaxIDR:

- Decrease: Greater similarity between high and low flows will result in more stable channel morphology, with less bank erosion, leading to a reduction of large woody debris entering the stream. Access to the floodplain will be reduced, limiting growth of fish and amphibians that take advantage of this resource.
- Increase: Increased differences between high and low flows may destabilize channels, leading to greater bank erosion or incision, affecting the growth or survival of riparian vegetation. The consequent loss of riparian vegetation may decrease shading and leaf-litter input to the stream, shifting the trophic structure from an allochthonous system to an autochthonous one.

RBI

- Decrease: Reduced flashiness decreases the frequency of mortality events, allowing the proliferation of long-lived semivoltine taxa.
- Increase: Increased flashiness favors short-lived, multi-voltine taxa and good dispersers that can recover quickly after frequent flooding events.

Table 13. Description and management implications of priority flow metrics

- NoDisturb (days), is the median annual longest number of consecutive days that flow is between the low (Q10) and the high flow (Q90) threshold. Disturbance changes the bed shear stress and effects sediment transport. While an increase in the number of no-disturbance days does not have a high negative impact on the stream health, a decrease in the number of days is significant. Under urbanization scenarios we usually see a decrease in the number of no-disturbance days.

- HighDur, is the median annual longest number of days the flows were greater than upper threshold (Q90). This metric only has a lower threshold and a corresponding lower target. In terms of management, as long as the metric value is higher than the lower target, the stream is not failing the metric. Both the duration metrics require several years of data.

- MaxMonthQ (cms) is the maximum mean of the monthly flows. The MaxMonthQ has an upper threshold and associated target but no lower target. The management goal is to ensure that the metric values are below the upper target value. In cases of urbanization, we see a rapid increase in the MaxMonthQ.

- Q99 (cms) is a high flow threshold, or the top 1% of the flow and has upper and lower bound targets in cms. The management goal is to maintain the metric values within this range. In cases of urbanization, we see a rapid increase in the Q99 values.

- RBI describes the oscillation in flows (or discharge) relative to the total flows (Baker et al 2004). This flashiness metric usually increases with urbanization which impacts the runoff patterns. However, the flashiness might decrease in case there are dams or steady controlled releases from reservoirs which dampen the natural flashiness of the hydrograph. The metric has an upper target, which implies that an extreme flashy stream is unhealthy for the biological communities, and the management goals should focus on keeping the RBI scores below the upper target value.

- Qmax IDR measures variability as the difference between the high flow threshold (Q90) and low flow threshold (Q10) divided by the 50th percentile flow (Q50). A higher value implies increasing variability, which is typically the case in streams without hydrologic regulation.

- HighNum is the frequency metric which estimates the number of events where the flow is higher than Q90 threshold. This metric has an upper target which implies that the management should focus on maintaining high flow events to a number less than the upper target.

We did not stratify streams in the San Diego case study, as is suggested for the general ELOHA approach. The San Diego watershed includes three stream classes from the statewide classification (Pyne et al. in press), with 60% of the streams being in one class and the remaining 40% being equally divided between two other classes. However, our analysis did not result in substantial differences in the local flow–ecology relationships as result of stream class. Instead climate (wet, dry, average rainfall) was a more important predictor. Therefore, we classified relationships by climatic period vs. stream type.

Challenges of the ELOHA approach

The main challenges associated with local implementation of the regional flow-ecology relationships relate to availability of high-quality input data, applicability of some metrics to site-specific simulation of reference conditions, and limitations on the interpretation of the output relative to other considerations and potentially confounding factors. Quality of rainfall data was one of the most critical factors affecting confidence in the regional flow–ecology relationships (Sengupta et al. in review). Similarly, in the San Diego River watershed the uneven availability of high-quality hourly rainfall data that encompassed all climatic conditions affected our ability to apply the hydrologic models equally across the entire watershed. In some cases, we had to drop data from the nearest gages due to gaps or obvious errors and substitute with less proximate gages, but that provided better or more complete rainfall data. This spatial offset introduced some additional uncertainty that must be accounted for in interpreting the model output.

Application of the regional flow-ecology relationships to local management scenarios revealed several complications associated with the formulations of certain metrics commonly used in applications of the ELOHA framework (e.g., Solans and Jalon 2016). The first complication involves many duration metrics that are calculated based on frequency or duration of flows above or below a benchmark derived from a long-term flow record. For example, the HighNum metric is calculated as the number of flow events over the 90^{th} percentile of daily flow. This formulation may not be suitable for evaluating hydrologic change, because the benchmark may shift along with other parts of the hydrograph, thereby obscuring hydrologic impacts. Figure 12 shows the current and reference hydrographs of a site that has experienced dramatically increased flows. If the current flows are compared to a benchmark derived from the historic hydrograph, it is clear that the site experiences one very extended high flow event every year; in contrast, the historic flows experienced several, short-duration high-flow events each year. However, if the current flows are compared to a benchmark derived from the current hydrograph (as is commonly done), the site appears to experience only a few short high-flow events each year. Thus, the hydrologic alterations from historic conditions are obscured when shifting benchmarks are used to calculate certain metrics. This problem is not easily apparent in regional analyses due to the large sample size, and is most clear when applied to a specific site, as in the present study. We recommend that future analysis use a constant, unshifting benchmark based on historical conditions when estimating thresholds for duration metrics based on thresholds of high- or low-flow events.

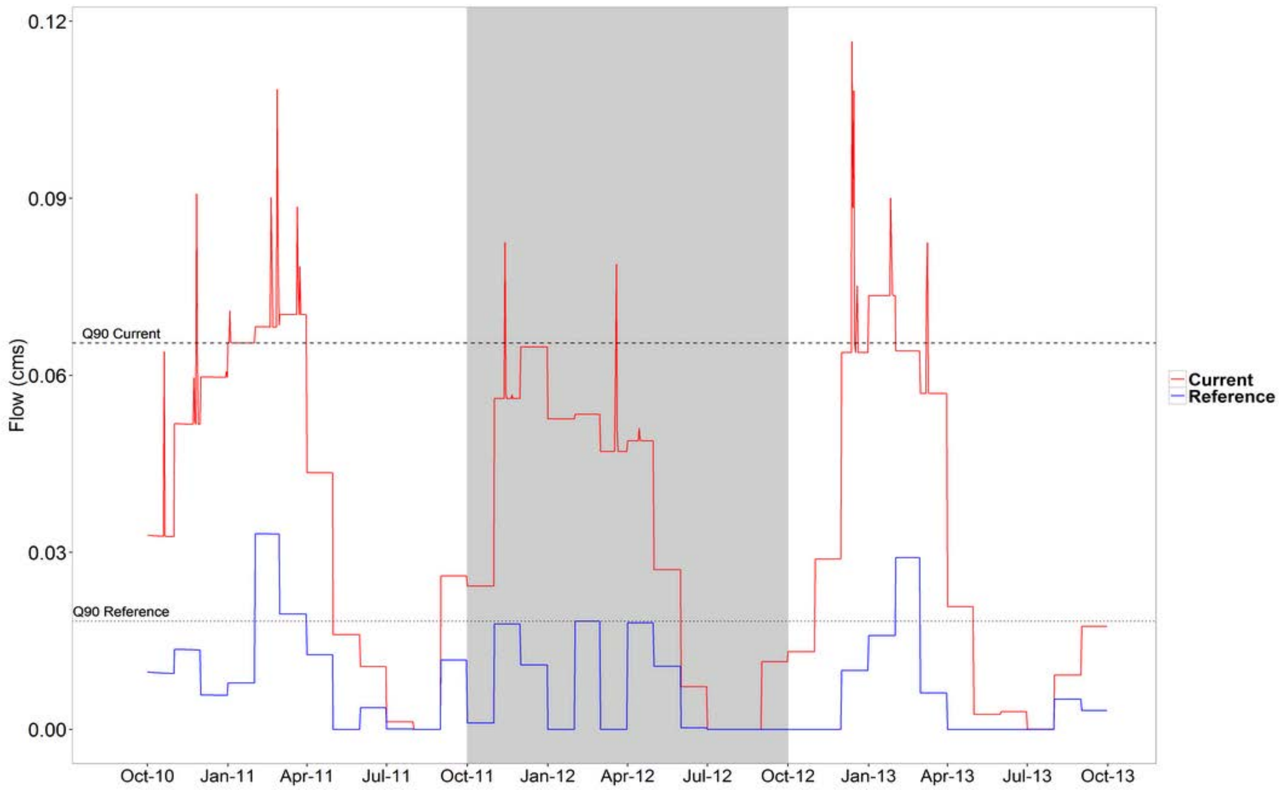


Figure 12. Comparison of current and reference flow for a sample bioassessment site showing the effect of the use of different thresholds. Conclusions about changes in duration of high flow events would vary dramatically if only a single threshold based on reference is issued vs. different thresholds were used for current and reference conditions.

The second issue associated with metric calculation relates to anomalous results that may occur when reference conditions are expected to represent intermittent streams with long periods of zero-flow days. This may result in reference flows for many of the magnitude metrics being extremely low (or zero), making it virtually impossible for management scenarios to achieve targets for certain metrics. This computational issue is confounded by the real challenge that it may not be possible to reduce runoff back to natural conditions, even with full implementation of stormwater runoff controls (DeGasperi et al. 2009). New or modified metrics may need to be developed to accommodate establishing flow management targets appropriate for naturally intermittent or ephemeral streams.

The use of HEC-HMS to produce the delta H values was a tradeoff between ease of use and model precision. HEC-HMS is arguably not optimal for evaluating BMP and other non-point source runoff management measures. We chose this model to develop the regional flow-ecology relationships because of its simplicity, availability, ability to perform long term continuous simulations of streamflow. Its status as an industry standard model developed by the US Army Corps of Engineers makes it practical for application to the hundreds of catchments evaluated during the regional analysis. Similarly, its familiarity and accessibility make it ideal for involving local stakeholders in the analysis and decision-making process. However, other lumped parameter hydrologic models that are also widely used to perform continuous simulations, such as HSPF or SWMM, may be more appropriate. SWMM is more robust in terms of modeling storm sewers and various stormwater control measures including low impact development practices. At the expense of more complexity and model parameters, HSPF includes additional details on soil moisture and subsurface processes that can enhance modeling of baseflow and groundwater behavior. These features would likely provide more precise estimates of how future management interventions could affect runoff and, consequently, stream flow metrics. We did not investigate whether/how use of an alternative or more sophisticated model would affect the output of our scenario analyses, but this should be investigated in the future.

Our reliance on developing flow targets based on the response of a single community assumes that the macroinvertebrate community reflects overall ecological condition. Although this is not a totally unreasonable assumption, we recognize that different components of the stream ecosystem may be affected differently by changes in various components of the hydrograph. Other ELOHA efforts have attempted to address this issue by developing flow–ecology relationships for multiple communities (e.g. fish, vegetation, mussels) and recommending targets around protection of each (DePhilip and Moberg 2013). This approach is more robust, but complicates development of management measures that can address all biological endpoints. Ultimately, such an approach is likely less parsimonious for regulatory applications.

Spatial and temporal factors must also be considered when applying flow–ecology relationships. Our analysis focused on catchment-scale responses. However, benthic invertebrates may also respond to local scale factors such as duration of wetting of bars and localized velocity zones (Kath et al. 2016, Kennedy et al. 2016). Hydrologic change at the local (small) scale may be ecologically important but is likely not affected by managing for the flow metrics we identified, and may be difficult to address through any regionally derived flow management framework. Although our regional flow criteria were developed in consideration of wet, dry and average climatic cycles, they likely do not account for longer term climate patterns and extreme episodic events that may be important for establishing and maintaining resilient instream habitats. This deficiency was highlighted by McManamay et al. (2013), who found that results of ELOHA analysis cannot necessarily be used in a predictive manner because biological communities may respond to other factors not included in the flow–ecology analysis, such as changes in substrate associated with infrequent events, such as catastrophic floods or fires. Moreover, they note that temporal resolution of most case studies does not coincide with the temporal period of data underlying ELOHA relationships. For example, streams may respond to episodic events and patterns operating on decadal time scales. We currently lack flow metrics that capture these interannual and longer term hydrologic patterns. Finally, as we noted in our analysis confounding factors such as changes in water chemistry typically co-occur with hydrologic changes and may contribute to biological community health in ways not captured by flow management.

These issues reinforce the concept that flow-ecology relationships should be used as one line of evidence in coordination with other factors/considerations when establishing stream management prescriptions and targets. In particular, many watersheds are subject to complex regulatory and management systems that involve combinations of new and retrofit facilities aimed at reducing runoff and retaining flows for infiltration and reuse. The regional flow targets established by Mazor et al. (in review) and applied in this case study can be an important consideration in designing and implementing integrated watershed management plans aimed at meeting both short and lon objectives.

Framework for development of local flow targets

We found the case study process to be productive because it provided a framework for considering hydrologic management in the context of watershed planning. It also provided the first opportunity in the region to develop quantitative flow targets that could be used to inform actionable management decisions. The regional flow–ecology relationships provided flexibility in establishing targets based on desired levels of confidence that those targets would be associated with healthy biological communities. Regionally derived targets took advantage of the robust regional monitoring data set and a broad set of hydrologic conditions. This improved relevance to local conditions was an important consideration for the watershed stakeholders. Given the utility of the process, we used the case study to develop a stepwise process that can serve as a framework for future implementation in other watersheds. This stepwise process is based on an adaptation of the ELOHA framework (Figure 13).

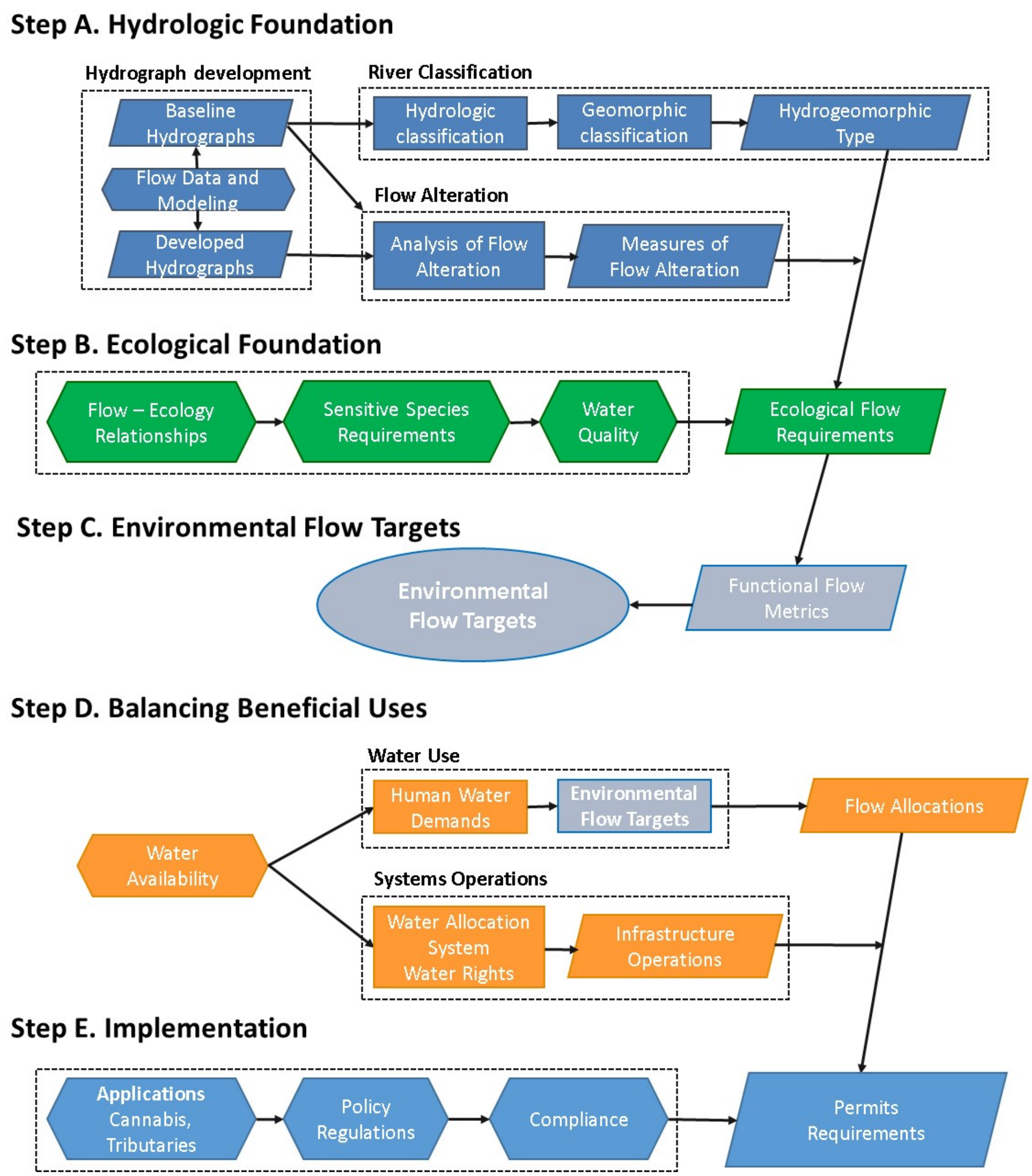


Figure 13. Process for development and implementation of instream flow targets, modified from the ELOHA framework (Poff et al. 2010).

Based on the framework in Figure 13, we identified the following steps that can be followed if other groups wish to pursue similar efforts to develop flow management recommendations:

Step 1: Determine what hydrologic class the stream of interest is in

Step 2: Identify management needs, regulatory objectives, or other targets

Step 3: Compile local data

- Contemporary and proposed future land use
- Information on contemporary and proposed water capture, storage, diversion, discharge, and other water management
- Local rainfall data at hourly time intervals (data must be checked to ensure sufficient quality and duration, at least ten years that encompass wet, dry, average rainfall conditions)

Step 4: Divide watershed in subbasins for analysis based on hydrology and management needs

Step 5: Select appropriate model(s) for catchments of interest using regional model selection tool

Step 6: Model both contemporary and natural hydrology for each catchment

Step 7: Calculate delta H metrics for each reach/node

Step 8: Select priority metrics and targets based on the following:

- Recommendations from the regional ELOHA analysis
- Relevance to local management needs
- Ability to influence through management measures

Step 9: Determine temporal factors associated with the targets

- Seasonality
- Persistence/duration
- Frequency (e.g. always, every X years)

Step 10: Evaluate various management scenarios relative to targets identified in Step 8

Step 11: Explore potential related or confounding factors (e.g. water quality, substrate)

Step 12: Develop recommended actions to achieve flow targets

- Relate actions to specific hydrologic modifications, e.g. diversion rates

Step 13: Relate flow metrics and targets to monitoring design, locations, and indicators

Step 14: Determine adaptive management actions that will be triggered if targets are not met

Informing management decisions

Stakeholder participation was critical in identifying scenarios and interpreting how the results of the analysis can be used to inform management action. Stakeholders identified the following desired applications for flow targets, which helped define our analysis:

- Identify priority management sites based on biological and hydrologic condition
- Use results to inform BMP/LID selection
- Identify areas were flow management has potential to improve CSCI scores
- Explore implication of future management of reservoirs for multiple benefits, e.g. water quality and water supply

These desired uses shaped our ultimate products. For example, we developed the overall composite index of hydrologic alteration in direct response to stakeholder desire to holistically assess the watershed for areas most vulnerable to future hydrologic alteration. Not surprisingly, the degree of hydrologic modification was correlated with impervious cover. We found that hydrologic alteration generally occurred in catchments with greater than 5% total impervious cover, which is similar to other studies that have shown that channel degradation to due to hydromodification occurs at relatively low levels of imperviousness (Hawley et al. 2012, Vietz et al. 2016). Similarly, the map of hydrologic management categories was identified as one of the most useful products for planning purposes because it allows stakeholders to prioritize areas for protection and for flow management.

We were able to demonstrate the utility of applying the flow-ecology relationships to inform management for both point source and non-point source management scenarios. For both the reservoir management scenario and the urban runoff management scenario, we were able to determine a range at which hydrologic management may facilitate recovery of impacted biological communities.

Lessons learned for future implementation

Future efforts can built on the experiences from this case study and continue to refine an iterative process of developing flow targets that are scientifically defensible, practical (i.e., can lead to management actions), and consistent with local stakeholder needs. Key lessons learned from this effort include:

1. Include a broad set of engaged stakeholders, including regulatory agencies, municipalities, water agencies, non-governmental organizations, and researchers. This ensures a broad perspective in the deliberations and increases the likelihood of developing balanced recommendations.
2. Invest in educating the stakeholders early in the process on the underlying science and the rationale behind how regional flow targets were developed. This promotes engagement and fosters creative solutions to the complex challenges of flow management.
3. Invest the time to compile high quality local data sources and show how local data can be used in the evaluation process. Identify the areas were future data collection can most improve outputs of the flow–ecology analysis (e.g., local rainfall data, more refined land use, water quality data). This can inform future monitoring.
4. Develop documentation that clearly illustrates how the products of the flow–ecology analysis can be used in the context of existing regulatory or management programs.

The San Diego River implementation case study also produced several technical recommendations that can improve our ability to apply flow-ecology relationships to manage southern California streams:

1. Several flow metrics, particularly those associated with flow duration, may require modification for use in streams where the natural condition is intermittent or ephemeral.

Natural intermittency poses fewer issues when developing regional flow-ecology relationships based on hundreds of sites. However, application of the resultant thresholds to specific streams that may have been naturally intermittent can lead to erroneous results.

2. Metrics associated with flow durations should be calculated on a single threshold value based on reference conditions. Estimating change in flow durations based on a moving threshold estimated separately for current and reference conditions may produce erroneous results.

3. Need to improve the representation of the drainage system to provide a more accurate hydrologic foundation for analysis. This would ultimately include improved mapping of discharges, diversions, stormwater control facilities, LID, etc. for incorporation into modeling scenarios and effects.

4. Consider expanding the analysis to include additional elements in future case studies

    - Include other stream or water body types
    - Include other indicators (e.g. algae)
    - Explore how consistent/transferable findings are from one watershed to another
    - Explore application in watersheds that cross jurisdictional boundaries

The original authors of the ELOHA framework promote the idea that flow targets derived by statistical analysis are a starting point. Targets should be iteratively refined using additional monitoring data, professional judgement and consideration of all complementary and competing factors necessary to develop flow standards that can address often divergent interests. The San Diego River case study provides an illustration of how watershed stakeholders are critical partners in the process. Resultant flow standards provide a starting point for developing agreed upon, adaptive flow management programs that can protect intact waterbodies and restore those that are currently impacted.

LITERATURE CITED

Baker, D.B., Richards, R.P, Loftus, T.T, and Kramer, J.W. 2004. A new flashiness index: Characteristics and applications to Midwestern Rivers and streams. JAWRA Journal of the American Water Resources Association, 40: 503–522. doi:10.1111/j.1752-

Beecher, H.A., B.A. Caldwell, S.B. DeMond, D. Seiler, S.N. Boessow. 2010. An Empirical Assessment of PHABSIM Using Long-Term Monitoring of Coho Salmon Smolt Production in Bingham Creek,Washington, North American Journal of Fisheries Management, 30: 6, 1529 — 1543: DOI: 10.1577/M10-020.1

Buchanan,C, H.L.N. Moltz, H.C. Haywood, J.B. Palmer, A.N. Griggs. 2013. A test of The Ecological Limits of Hydrologic Alteration (ELOHA) method for determining environmental flows in the Potomac River basin, U.S.A. Freshwater Biology (2013) 58, 2632–2647

Bunn, S.E. and A.H. Arthington. 2002, Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity: Environmental Management, v. 30, no. 4, p. 492–507.

Colwell R.K. 1974. Predictability, constancy, and contingency of periodic phenomena. Ecology 55(5): 1148-53.

DeGasperi C.L., H.B. Berge, K.R. Whiting, J.J. Burkey, J.L. Cassin and R.R. Fuerstenberg. 2009. Linking Hydrologic Alteration to Biological Impairment in Urbanizing Streams of the Puget Lowland, Washington, USA. Journal of the American Water Resources Association (JAWRA) 45(2):512-533. DOI: 10.1111/j.1752-1688.2009.00306.x

DePhilip, M. and T. Moberg. 2010. Ecosystem flow recommendations for the Susquehanna River Basin. The Nature Conservancy. Harrisburg, PA.

DePhilip, M. and T. Moberg. 2013. Ecosystem flow recommendations for the Upper Ohio River basin in western Pennsylvania. The Nature Conservancy. Harrisburg, PA.

Elith J., J.R. Leathwick and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802–813.

Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29: 1189–1232.

Hawley, R.J. and B.P. Bledsoe. 2011. How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study. Journal of Hydrology 405:69–82

Hawley, R.J., B.P. Bledsoe, E.D. Stein and B.E. Haines. 2012. Channel Evolution Model of Response to Urbanization in Southern California. *Journal of the American Water Resources Association* 48(4):722—744.

Horvitz, D.G. and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47:663-685.

Kath J., E. Harrison, B.J. Kefford, L. Moore, P.J. Wood, R.B. Schäfer, F. Dyer. 2016. Looking beneath the surface: using hydrogeology and traits to explain flow variability effects on stream macroinvertebrates. Ecohydrology (online). DOI: 10.1002/eco.1741

Kennedy T.A., J.D. Muehlbauer, C.B. Yackulic, D.A. Lytle, S.W. Miller, K.L. Dibble, E.W. Kortenhoeven, A.N. Metcalfe, C.V. Baxter. 2016. Flow Management for Hydropower Extirpates Aquatic Insects, Undermining River Food Webs. BioScience (online). Doi:10.1093/biosci/biw059

Kincaid, T.M. and A.R. Olsen. 2013. Spsurvey: Spatial survey design and analysysis. R package version 2.6. R Foundation for Statistical Computing. Vienna, Austria.

King County. 2012. Stormwater Retrofit Analysis for Juanita Creek Basin in the Lake Washington Watershed. Ecology Grant: G0800618. Prepared by J Burkey, M Wilgus H Berge. King County Department of Natural Resources and Parks. Water and Land Resources Division.Seattle, Washington.

Konrad C.P. and D.B. Booth. 2005. Hydrologic Changes in Urban Streams and Their Ecological Significance. American Fisheries Society Symposium 47:157–177

Konrad C,P,, A,M,D. Brasher, J.T. May. 2008. Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States. Freshwater Biology, 53:1983–1998.

Mazor, R.D., J. May, A. Sengupta, K. McCune, E.D. Stein. (In review). Setting hydrologic targets to support biological integrity. Freshwater Biology

Mazor, R.D. 2015. Bioassessment of Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey. Southern California Coastal Water Research Project Technical Report #844

Mazor, R.D., A.C. Rehn, P.R. Ode, M. Engeln, K.C. Schiff, E.D. Stein, D.J. Gillett, D.B. Herbst, C.P. Hawkins. 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshwater Science. 35(1):249–271. DOI: 10.1086/684130

Mazor, R.D., K. Schiff, K. Ritter, A. Rehn, P. Ode. 2010. Bioassessment tools in novel habitats: An evaluation of indices and sampling methods in low-gradient streams in California. Environmental Monitoring and Assessment 167:91-104.

McClain, M.E., A.L. Subalusky, E.P. Anderson, S.B. Dessu, A.M. Melesse, P.M. Ndomba, J.O.D. Mtamba, R.A. Tamatamah, C. Mligo. 2014. Comparing flow regime, channel hydraulics and biological communities to infer flow–ecology relationships in the Mara River of Kenya and Tanzania. Hydrological Sciences Journal, 59 (3–4), 801–819.

McManamay, R.A., D.J. Orth, C.A. Dolloff, D.C. Mathews. 2013. Application of the ELOHA Framework to Regulated Rivers in the Upper Tennessee River Basin: A Case Study. Environmental Management 51:1210–1235: DOI 10.1007/s00267-013-0055-3

Moriasi, D.N, J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, T.L. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE 50(3):885-900

Morley, S.A., J.R. Karr. 2002. Assessing and Restoring the Health of Urban Streams in the Puget Sound Basin. Conservation Biology 16(6):1498–1509

Naiman, R.J., S.E. Bunn, C. Nilsson, G.E. Petts, G. Pinnay, L.C. Thompson. 2002. Legitimizing fluvial ecosystems as users of water—an overview: Environmental Management, v. 30, no. 4, p. 455–467.

Novak, R., J.G. Kennen, R.W. Abele, C.F. Baschon, D.M. Carlisle, L. Dlugolecki., J.E. Flotermersch, P. Ford, J. Fowler, R. Galer, L.P. Gordon, S.N. Hansen, B. Herbold, T.E. Johnson, J.M. Johnston, C.P. Konrad, B. Leamond, P.W. Seelbach. 2015. Draft: EPA-USGS Technical Report: Protecting Aquatic Life from Effects of Hydrologic Alteration: U.S. Geological Survey Scientific Investigations Report 2015–5160, U.S. Environmental Protection Agency EPA Report

822-P-15-002, XX p., http://pubs.usgs.gov/sir/2015/5160/ and http://www2.epa.gov/wqc/aquatic life-ambient-water quality-criteria

Ode PR. 2007. Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Surface Water Ambient Monitoring Program. Sacramento, CA.

Parasiewicz, P. 2009. Habitat Time Series Analysis to Define Flow Augmentation Strategy for the Quinebaug River, Connecticut and Massachusetts, USA. River Research and Applications. Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rra.1066

Peck, D.V.,A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchak, F.H. McCormick, S.A. Peterson, P.L. Ringold, T. Magee, and M. Cappaert. 2006. Environmental Monitoring and Assessment Program -Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA/620/R-06/003. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C

Peterson, S.A., P. L. Ringold, T Magee, M Cappaert. 2006. Environmental Monitoring and Assessment Program - Surface Waters Western Pilot Study: field operations manual for wadeable streams. EPA/620/R-06/003. US Environmental Protection Agency, Office of Research and Development. Corvallis, OR.

Poff, N.L. and J.K.H. Zimmerman. 2010. Ecological responses to altered flow regimes—A literature review to inform the science and management of environmental flows: Freshwater Biology, v. 55, no. 1, p. 194–205.

Poff, N.L., B. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, C. Apse, E. Kendy, A.T. Warner, R. Tharme, B.P. Bledsoe, D. Merritt, R.B. Jacobson, M. Freeman, K. Rogers, J. Henriksen, J. Olden, J. O'Keeffe, M. Acreman. 2010. The ecological limits of hydrologic alteration: A framework for developing regional environmental flow standards. Freshwater Biology 55, 147–170.

Poff, N.L., J.D. Allan, M.B. Bain, J.R. Karr, K.L. Prestegaard, B.D. Richter, R.E. Sparks, J.C. Stromberg. 1997. The natural flow regime—a paradigm for river conservation and restoration: Bioscience, v. 47, no. 11, p. 769–784.

Pyne, M.I., D.M. Carlisle, C.P. Konrad, E.D. Stein. (in press) Classification of California streams using combined deductive and inductive approaches: setting the foundation for analysis of hydrologic alteration. Ecohydrology (in press).

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Version 3.2.4. Vienna, Austria.

Richards, A.B. and D.C. Rogers. 2006. List of Freshwater Macroinvertebrate Taxa from California and Adjacent States Including Standard Taxonomic Effort Levels. Southwest Association of Freshwater Invertebrate Taxonomists. Chico, CA.

Richter, B.D., M.M. Davis, C. Apse, C. Konrad. 2011. A presumptive standard for environmental flow protection. River Research and Applications 28(8):1312-1321. DOI: 0.1002/rra.1511

Ridgeway, G. 2015. gbm: Generalized Boosted Regression Models. R package version 2.1.1. https://CRAN.R-project.org/package=gbm

San Diego Geographic Information Source. Digital land use layers. http://www.sangis.org/index.html

Scharffenberg, W. A., and M.J. Fleming. Hydrologic Modeling System HEC-HMS: User's Manual. US Army Corps of Engineers, Hydrologic Engineering Center, 2006.

Sengupta A., E.D. Stein, K. McCune, R.D. Mazor, B.P. Bledsoe, S. Adams, C.P. Konrad. (In review). From Gaged to Ungaged: Predicting Flow and Hydrologic Alterations at a Regional Scale for Application to Regional Flow-ecology analyses. Freshwater Biology

Solans, M.A. and D.G. Jalon. 2016. Basic tools for setting environmental flows at the regional scale: application of the ELOHA framework in a Mediterranean river basin. Ecohydrology. DOI: 10.1002/eco.1745

Stevens, D.L. and A.R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262-278.

U.S. Army Corps of Engineers (ACOE). 2000. Hydrologic Modeling System HEC-HMS U.S. Army Corps of Engineers R&D Technical Reference Manual.

Vietz, G.J., M.J. Sammonds, C.J. Walsh, T.D. Fletcher, I.D. Rutherford, M.J. Stewardson. 2016. Ecologically relevant geomorphic attributes of streams are impaired by even low levels of watershed effective imperviousness. Geomorphology 206:67-78

Xuefeng, C. and Alan Steinman, A. 2009. Event and continuous hydrologic modeling with HEC-HMS. Journal of Irrigation and Drainage Engineering135.1 (2009): 119-124.

Yarnell, S.M., G.E. Petts, J.C. Schmidt, A.A. Whipple, E.E. Beller, C.N. Dahm, P. Goodwin, J.H. Viers. 2015. Functional Flows in Modified Riverscapes: Hydrographs, Habitats and Opportunities. BioScience: doi: 10.1093/biosci/biv102

APPENDIX A – DETAILED PROCEDURES FOR HYDROLOGIC ANALYSIS

Directions to run HEC-HMS Modeling packages developed for flow ecology analysis

To be able to run these modules

- Basic idea of catchments, watersheds, and delineated areas
- Moderate skills in R programming (scripts provided)
- Basic understanding of watershed modeling

Software needed:

- Streamstats (online, no download necessary)

  http://streamstatsags.cr.usgs.gov/v3_beta/

- R and Rstudio (installation needed for both, install R before installing R studio)

  https://cran.cnr.berkeley.edu/

  https://www.rstudio.com/products/rstudio2/

- HEC-HMS (install)

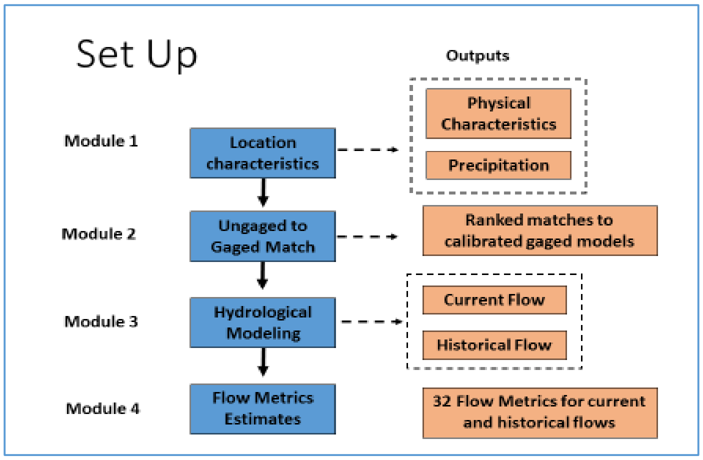
  http://www.hec.usace.army.mil/software/hec-hms/downloads.asp

Notes for running R scripts:

- setwd("../Desktop/") sets up each script to automatically read files in the folder Modeling Workshop, as long it is located on your desktop
- Mac users will need to change the ".." in "../Desktop/" to "~" (tilde)
- Each script must be opened from within R-Studio in order to correctly use setwd("../Desktop/")
- If you get an error that says you cannot change the working directory, then close the script in R-Studio, close R-Studio, re-open R-Studio, then open the script from within R-Studio

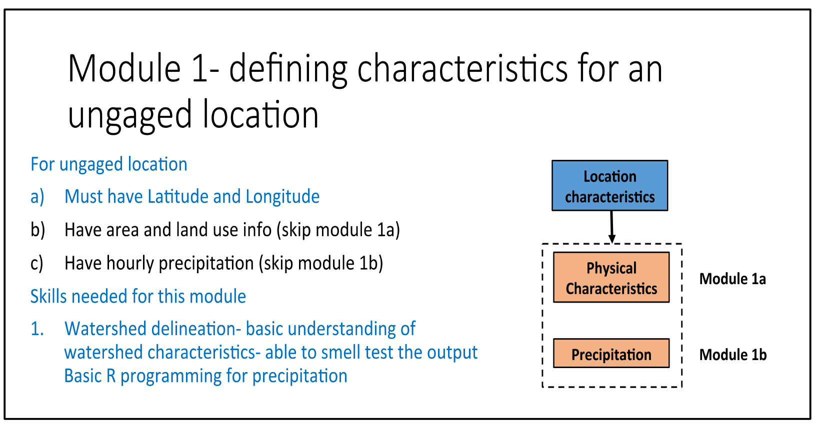
Introduction

The modeling tool has four modules. The modules should be run sequentially to get the flow metrics. Described below are four modules and their outputs.



Module 1.

Module 1 allows the users to delineate the watershed area for an ungaged site, and estimate hourly precipitation (1990-2013).



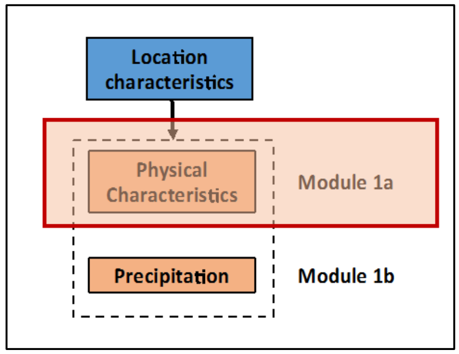
Limitations:

1. For some locations, the Streamstats outputs a square delineated catchment area. Always check the visual output, in case it looks incorrect, move the location slightly to obtain the watershed characteristics.

2. The precipitation raw data from the gages is limited to 1990-2013. The script is enabled for any period, to produce output for periods outside of 1990-2013, hourly gaged precipitation data is required.

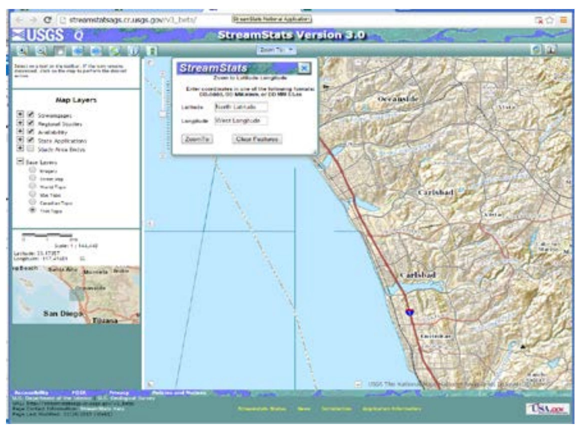
Instructions to run the modules

Delineate the subcatchment/watershed area for ungaged location

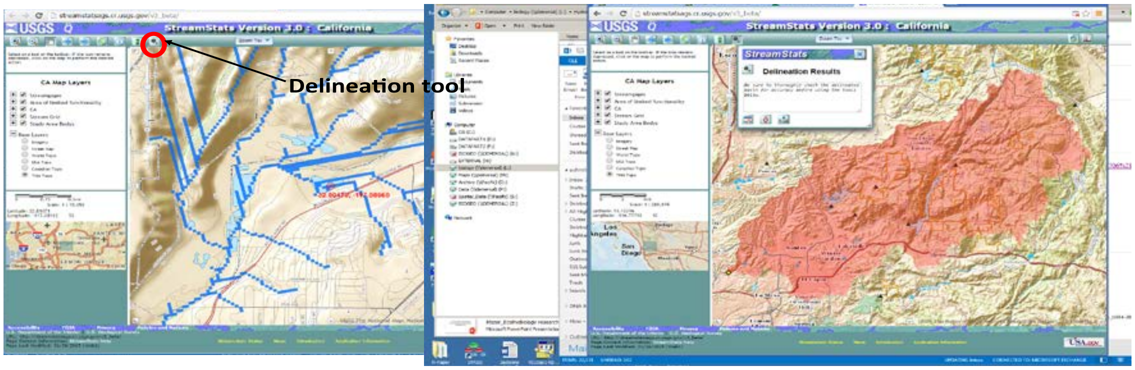


We will use Streamstats to delineate area and land use. URL provided below.

1. http://streamstatsags.cr.usgs.gov/v3_beta/
2. Press the 'zoom to' button highlighted in figure 4, and enter the latitude and longitude.

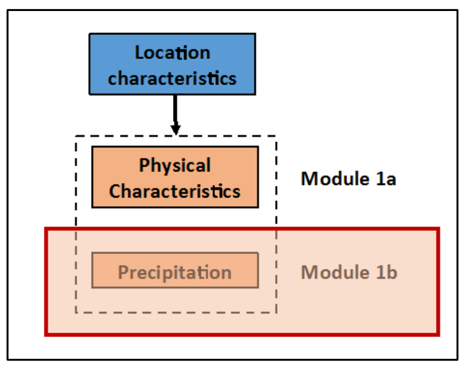


3. Zoom till you see the delineation tool (highlighted below). Select first tab in the pop out window (in figure 5b).



4. Select area and imperviousness, and compute

Module 1b. Estimating hourly precipitation



Input files: Modeling Workshop\Inverse Distance\Data*Precip_(YEAR).csv*, *AssessmentSiteCoord.csv*, *PrecipStationCoord.csv*

Output file: Modeling Workshop\Inverse Distance\Data*Assess_(YEAR).csv*

1. We use R Studio to estimate hourly precipitation.
2. To predict flows at the gages, we need hourly precipitation data
3. Daily data is available on PRISM website
4. For better flow predictions, we estimated hourly flow using precipitation data from >200 sites
5. You can use the script for 1990-2013 (and will require raw precipitation data outside this range)
6. Open the file *AssessmentSiteCoord.csv*, delete the current data in the spreadsheet and enter the data for your site ID, latitude and longitude in the appropriate columns

7. From within RStudio, open Modeling Workshop\Inverse Distance*InvDist_calc_02_SelectYears.R*
8. Run the line install.packages(…), then add a # at the beginning of the line
9. Specify year range on line 23 (default is 1990:2013)
10. Run script by clicking “Source” button at the top right of the script window
11. Look at *Assess_(YEAR).csv* in your working directory for output

Module 2: Matching ungaged sites to gaged sites

Calculates the proximity of the ungaged site to the calibration gages (models).

Input file: Modeling Workshop\Site Assignment*test.csv*

Output file: Modeling Workshop\Site Assignment*top.model.csv*

How to assign an ungauged site to a flow model:

1. Delineate watersheds, and 5-km watershed clips
2. Calculate predictors:

Variable name (CASE SENSITIVE!)	Description	Source file to use
StationCode	Unique site identifier	User
New_Lat	Latitude in decimal-degrees North. Not required for predictions, but useful for plotting.	User
New_Long	Longitude in decimal-degrees West (should be negative). Not required for predictions, but useful for plotting.	User
Imperv_percent	Mean % imperviousness in the catchment (0-100).	[StreamStats]
URBAN_2000_WS	NLCD urban land use in the catchment (0-100). For NLCD 2000, these codes count towards urban:	NLCD2000 or NLCD2006
KFCT_AVE	Mean soil erodibility in the catchment.	[RAFI]
Ag_2000_WS	NLCD agricultural land use in the catchment (0-100). For NLCD 2000, these codes count towards urban:	NLCD2000 or NLCD2006
CODE_21_2000_WS	NLCDE Code 21 (highly managed vegetation) in the catchment (0-100). For NLCD 2000, only code 21 counts.	NLCD2000 or NLCD2006
Ag_2000_5k	NLCD agricultural land use in the 5-km clip of the catchment (0-100). For NLCD 2000, these codes count towards urban:	NLCD2000 or NLCD2006
RoadDens_5K	Road density (km/km2) in the 5-km clip of the catchment. Dirt roads do not count.	[Rafi]

1. From within R Studio, open Modeling Workshop\Site Assignment*assigning_testsites_020116.R*
2. Currently *test.csv* has dummy site information; use this as a template for your site info
3. Please see handout for GIS information required for this module
4. *top.model.csv* is the output file with top matched gage info

Note: Within a year or two automated tools from the State (SWAMP) will calculate the variables, but for now, use GIS to estimate them.

Module 3: Running hydrological model (HEC-HMS) to predict hourly flow

We run the model for current and reference or historical conditions

Output files: Modeling Workshop\Hourly To Daily Flow Conversion\Hourly Flow*SDR_AssessmentSites_Hourly_Current.csv* and *SDR_AssessmentSites_Hourly_Reference.csv*

First we will estimate current flows

1. Navigate to Modeling Workshop\HEC HMS Models\Current Conditions
2. Copy the model folder that is the top matched to your site
3. Save a copy in a new folder
4. Click on HEC-HMS icon on your desktop
5. Click on open folder tab, navigate to your new folder
6. Open file with the .hms extension
7. We need to change 5 parameters- area, imperviousness, time of concentration, storage coefficient and precipitation
8. From the "Compute" tab, right click on "Run 1" then click "Compute"
9. From the "Results" tab, double click "Run 1", double click "Subbasin-1" then click on "Time-Series Table"
10. A window will appear shortly, from which you will copy all the data in the "Total Flow" column
11. Open the file Modeling Workshop\Hourly To Daily Flow Conversion\Hourly Flow*SDR_AssessmentSites_Hourly_Current.csv*, paste the "Total Flow" data in a new column on the right, starting on row 2
12. Put the site name in the first row of this column
13. Remove all the other columns, EXCEPT for your new column, the "Date_Time" column and one additional column of flow data (the metric calculation requires data from at least 2 sites)

Historical flows

1. Navigate to Modeling Workshop\HEC HMS Models\Reference Conditions

2. Copy the model folder that is the top matched to your site
3. Save a copy in a new folder
4. Click on HEC-HMS icon on your desktop
5. Click on open folder tab, navigate to your new folder
6. Open file with the .hms extension
7. We need to change 4 parameters- area, time of concentration, storage coefficient and precipitation
8. From the “Compute” tab, right click on “Run 1” then click “Compute”
9. From the “Results” tab, double click “Run 1”, double click “Subbasin-1” then click on “Time-Series Table”
10. A window will appear shortly, from which you will copy all the data in the “Total Flow” column
11. Open the file Modeling Workshop\Hourly To Daily Flow Conversion\Hourly Flow*SDR_AssessmentSites_Hourly_Reference.csv*, paste the “Total Flow” data in a new column on the right, starting on row 2
12. Put the site name in the first row of this new column
13. Remove all the other columns, EXCEPT for your new column, the “Date_Time” column and one additional column of flow data (the metric calculation requires data from at least 2 sites)

Module 4. Flow metrics are estimated on daily flow

Convert hourly flow output from HEC-HMS to daily flow

Input files: Modeling Workshop\Hourly To Daily Flow Conversion\Hourly Flow*SDR_AssessmentSites_Hourly_Current.csv* and *SDR_AssessmentSites_Hourly_Reference.csv*

Output files: Modeling Workshop\Hourly To Daily Flow Conversion\Daily Flow*SDR_Assessment_Daily_Current.csv* and *SDR_Assessment_Daily_Reference.csv*

1. From within R Studio, open Metric Workshop\Hourly To Daily Flow Conversion*HourlytoDailyFlow_Current.R*
2. Run install.packages(…) line, then add # to the beginning of the line
3. Run script by clicking “Source” button on the top right of the script window
4. Converted daily flow data will be in Modeling Workshop\Hourly To Daily Flow Conversion\Daily Flow*SDR_Assessment_Daily_Current.csv*
5. Repeat using Metric Workshop\Hourly To Daily Flow Conversion*HourlytoDailyFlow_Reference.R* to produce daily flow data for reference condition hourly flow data

Calculate Metrics for Daily Flow Data

Input files: Modeling Workshop\Hourly To Daily Flow Conversion\Daily Flow*SDR_AssessmentSites_Daily_Current.csv* and *SDR_AssessmentSites_Daily_Reference.csv*

Output files: Modeling Workshop\Metric Calculation \Results\ *Sdr_Current_Metrics.csv* and *Sdr_Reference_Metrics.csv*

1. From within R Studio, open Modeling Workshop\Metric Calculation*KonradMetrics_Current.R*
2. Click “Source” button in upper right corner of script window
3. Metric results will be in Modeling Workshop\Metric Calculation\Results*Sdr_Current_Metrics.csv*
4. Repeat using Modeling Workshop\Metric Calculation*KonradMetrics_Reference.R* to get metric results for reference condition flow data

Description of Metrics in QSUM (typically Median Annual Values)

Qmean [M3/S] - mean streamflow for the period of analysis

QmeanMEDIAN [M3/S] - median annual mean streamflow

QmeanIDR - (90th percentile of annual mean streamflow - 10th percentile of annual mean streamflow)/50th percentile of median annual mean streamflow

Qmed [M3/S] - median daily streamflow

Qmax [M3/S] - median annual maximum daily streamflow

QmaxIDR - (90th percentile of annual maximum streamflow - 10th percentile of annual maximum streamflow)/50th percentile of annual maximum streamflow

HighNum [events/year] - median annual number of events that flow was greater than high flow threshold, an event is a continuous period when daily flow exceeds the threshold

HighDur [days/event] - median annual longest number of consecutive days that flow was greater than the high flow threshold

Qmin [M3/S] - median annual minimum daily streamflow

QminIDR - (90th percentile of annual maximum streamflow - 10th percentile of annual maximum streamflow)/50th percentile of annual maximum streamflow

LowNum [events/year] - median annual number of events that flow was less than or equal to the low flow threshold, an event is a continuous period when daily flow was less than or equal to the threshold

LowDur [days/event]- median annual longest number of consecutive days that flow was less than or equal to the low flow threshold

NoDisturb [days] - median annual longest number of consecutive days that flow between the low and high flow threshold

Hydroperiod [0.01 = 1% of period of analysis] - fraction of period of analysis with flows

FracYearsNoFlow [0.01 = 1% of years] - - fraction of years with at least one no-flow day

MedianNoFlowDays [days/year]- median annual number of no-flow days

PDC50 [0.01=1% change in streamflow] - the median percent daily change in streamflow, no flow days are not included (0.01 = 1%)

SFR [-0.01=-1% change in streamflow]- the 90th percentile of percent daily change in streamflow on days when streamflow is receding (a measure of storm-flow recession)

BFR [-0.01=-1% change in streamflow] - the 50th percentile of percent daily change in streamflow on days when streamflow is receding (a measure of base-flow recession)

MaxMonth [1- Jan, 12-Dec] - month of maximum mean monthly streamflow

MaxMonthQ [M3/S] - maximum mean monthly streamflow

MinMonth [1- Jan, 12-Dec] - month of minimum mean monthly streamflow

MinMonthQ [M3/S] - minimum mean monthly streamflow

Q01, Q05, Q10, ...,Q99 [M3/S] - streamflow exceeded 1%, 5%, 10%, ..., 99% of the time

BugID	ModelMatch	Area	Imperviousness	Lat	Long
907S00577	SantaMaria_11028500	11.64	0.48	33.07609	-116.676
SMC04426	SanMateo_11046300	17.80	0.24	33.00697	-116.67
907S03210	Jamul_11014000	38.43	0.37	33.00313	-116.729
907S01418	SanMateo_11046300	24.87	0.19	32.99246	-116.719
SMC04682	SantaYsabel_11025500	21.04	0.27	32.97115	-116.648
907S46499	SanMateo_11046300	101.73	0.26	32.96269	-116.749
907S03786	SantaYsabel_11025500	11.29	0.26	32.89422	-116.658
SMC32718	Jamul_11014000	190.98	0.59	32.88455	-116.822
SMC11430	Mission_11119750	4.48	6.56	32.84835	-116.86
SMC02006	Poway_11023340	23.29	43.62	32.83115	-116.985
SMC09174	LosAngeles_11092450	346.73	6.08	32.83967	-117.002
SMC08150	Poway_11023340	367.06	6.13	32.83731	-117.02
SMC04134	Jamul_11014000	377.26	6.09	32.82874	-117.052
907P2PBxx	SanLuisRey_11042000	428.14	10.04	32.7675	-117.159
907S05514	SanMateo_11046300	66.73	0.29	32.97974	-116.742
907S01610	SanMateo_11046300	23.07	0.25	32.96676	-116.653
907S01434	SantaYsabel_11025500	5.26	0.10	32.90428	-116.626
907S02774	Poway_11023340	4.45	52.96	32.81182	-116.973
SMC10198	LosAngeles_11092450	5.60	53.24	32.82182	-116.976
907SDFRC2	LosAngeles_11092450	346.67	6.07	32.83945	-117.001
SMC04054	SanLuisRey_11042000	367.06	6.13	32.83697	-117.019
SMC19552	SanLuisRey_11042000	367.90	6.17	32.83965	-117.024
SMC07126	Mission_11119750	368.31	6.18	32.84359	-117.035
SMC12246	Mission_11119750	376.56	6.10	32.83982	-117.043
907SDSDR9	Mission_11119750	376.80	6.09	32.83894	-117.045
907SSDR11	Mission_11119750	380.87	6.12	32.82119	-117.063
SMC03110	Mission_11119750	381.67	6.17	32.81106	-117.073
SMC01990	Poway_11023340	12.19	25.15	32.79577	-117.113
SMC09286	Poway_11023340	405.87	8.53	32.78188	-117.114

APPENDIX B – STAKEHOLDER WORKGROUP AND SCHEDULE OF WORKGROUP MEETINGS

The demonstration project workgroup met six times between November 2015 and June 2016 (Table B1). All meetings were held in the San Diego River Watershed

Table B1. Workgroup participants

NAME	ORGANIZATION
Daron Pedroja	State Water Board
Gary Strawn	San Diego Water Board
Shannon Quiquley	San Diego River Park Foundation
Dustin Harrison	San Diego River Conservancy
Tracy Cline	San Diego County
Joanna Wisniewska	San Diego County
Eric Stein	SCCWRP
Raphael Mazor	SCCWRP
Ashmita Sengupta	SCCWRP
Alicia Kinoshita	San Diego State University
Trent Biggs	San Diego State University
Natalie Mladenov	San Diego State University
Charles Morloch	San Diego County
Rob Northcote	Padre Dam Municipal Water District
Arne Sandvik	Padre Dam Municipal Water District
Brian Olney	Helix Water District
Emily Blunt	U.S. Forest Service
Goldy Herbon	City of San Diego
Jeff Pasek	City of San Diego
Vicki Kalkirtz	City of San Diego
Andre Sonsken	City of San Diego
Jim Harry	City of San Diego
Anita Eng	City of San Diego
Doug Thomson	City of San Diego
James Dodd	City of San Diego
Maris Guerro	Army Corps of Engineers
John Rudolph	AMEC Environmental

The dates and goals of each meeting are listed below:

Meeting #1: November 18th, 2015
Meeting Goals:

- Provide an overview of the watershed demonstration project
- Discuss and agree upon portion of the watershed to focus on
- Agree on general roles and contributions of partners
- Develop general schedule for next set of meetings

Meeting #2: January 20th, 2016
Meeting Goals:

- Discuss work plan for priority actions/products from first meeting
- Agree on schedule for obtaining necessary data for analysis
- Compile list of primary contacts for participation in analysis

Meeting #3: February 17th, 2016
Meeting Goals:

- Technology transfer- using models to predict flows, and flow metrics at ungaged locations
- Discuss the process, and usability
- Discussion on final products

Meeting # 4: March 16th, 2016
Meeting Goals:

- Address outstanding issues on the hydrologic modeling tools
- Agree on management scenarios being evaluated

Meeting #5: April 20th, 2016
Meeting Goals:

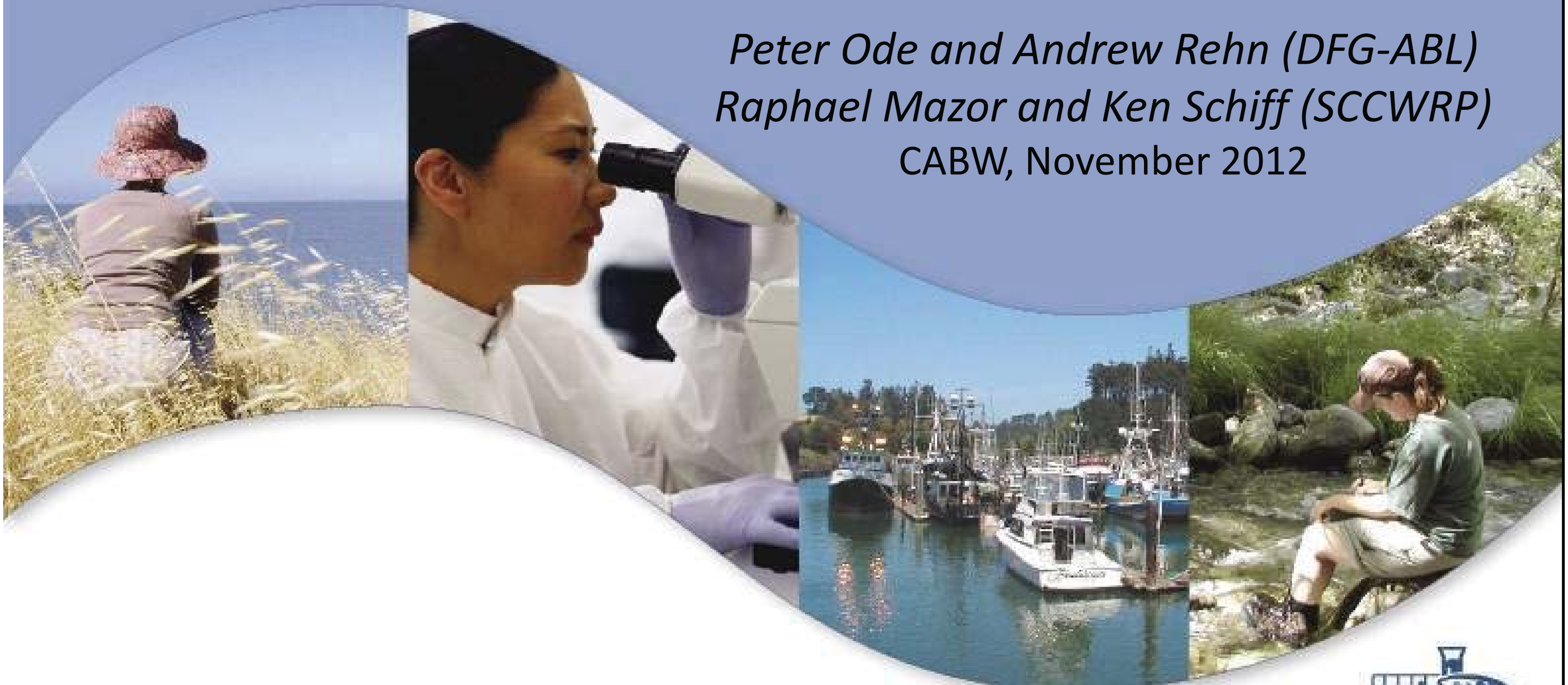
- Review products and outline for final demo project report

Meeting #6: June 15th, 2016
Meeting Goals:

- Review draft demo project report

Appendix 13

Building the Technical Foundation for Biological Objectives



Peter Ode and Andrew Rehn (DFG-ABL)
Raphael Mazor and Ken Schiff (SCCWRP)

CABW, November 2012

CALIFORNIA DEPARTMENT OF FISH&GAME

Aquatic Bioassessment Laboratory

SCCWRP Established 1969

CALIFORNIA Water Boards

SWAMP Surface Water Ambient Monitoring Program

- **Technical Foundation** (Peter Ode/ Rafi Mazor – DFG, SCCWRP)
- **Regulatory Framework** (Karen Larsen, State Water Board)
- **Causal Analysis** (David Gillett, SCCWRP)
- **Stakeholder Process** (Brock Bernstein)
- **Open Discussion**

- **Measuring Stressor Distributions** (Andy Rehn, DFG)
- **Tools for Assessing Stream/Wetland Condition** (Eric Stein, SCCWRP)
- **SWAMP's Lab SOP for BMIs** (Melinda Woodard, QA Team)

SWAMP Surface Water Ambient Monitoring Program

Technical Foundation

Part I – Laying the groundwork (20)

Part II – Creating the scoring tools (40)

Part III – Supporting Implementation (20)

SWAMP
Surface Water
Ambient Monitoring
Program

Technical Team

*Andy Rehn, *DFG-ABL*
*Raphael Mazor, *SCCWRP +DFG-ABL*
Larry Brown, *USGS*
Jason May, *USGS*
David Herbst, *SNARL*
Peter Ode, *DFG-WPCL/ABL*

Ken Schiff, *SCCWRP*
David Gillett, *SCCWRP*
Eric Stein, *SCCWRP*
Betty Fetscher, *SCCWRP*
Kevin Lunde, *SF Water Board*

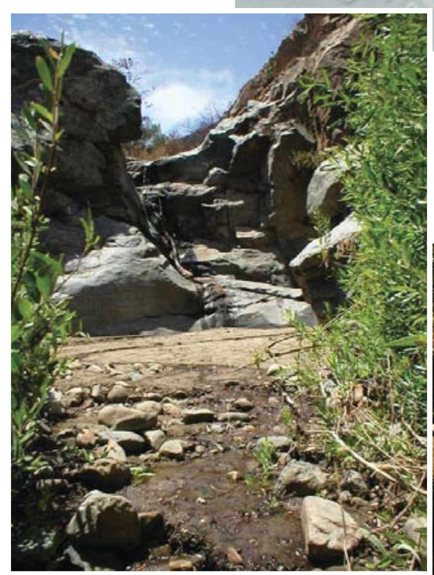
B444

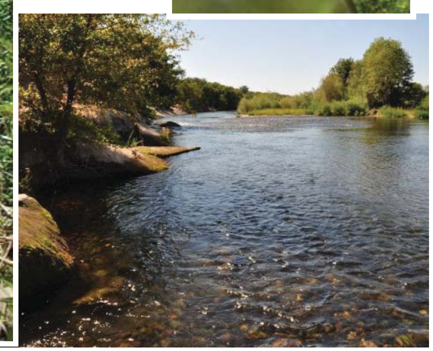
Why Develop Ecological Indicators?

- Global paradigm shift toward ecological indicators
- Provide direct evidence about resources we are trying to protect
- More relevant measures of impacts and BMP effectiveness
- Links resource protection across multiple agencies by focus on ultimate policy goals

SWAMP
Surface Water Ambient Monitoring Program

CA’s Ecological Indicators





Multiple Indicators – BMIs, algae, (fish), riparian vegetation

Multiple waterbody types – large rivers, non-perennial streams, lakes, wetlands

Start with invertebrates and perennial streams

invertebrates:

the backbone of bioassessment

- ***Abundant***
- ***Diverse***
- ***Informative***
- ***Adorable***


How do we convert a list of species into a condition score?
NABS (www.benthos.org)

Standardized Bioassessment Infrastructure Elements

Surface Water Ambient Monitoring Program (SWAMP)

Field Methods

Data Management + Reporting

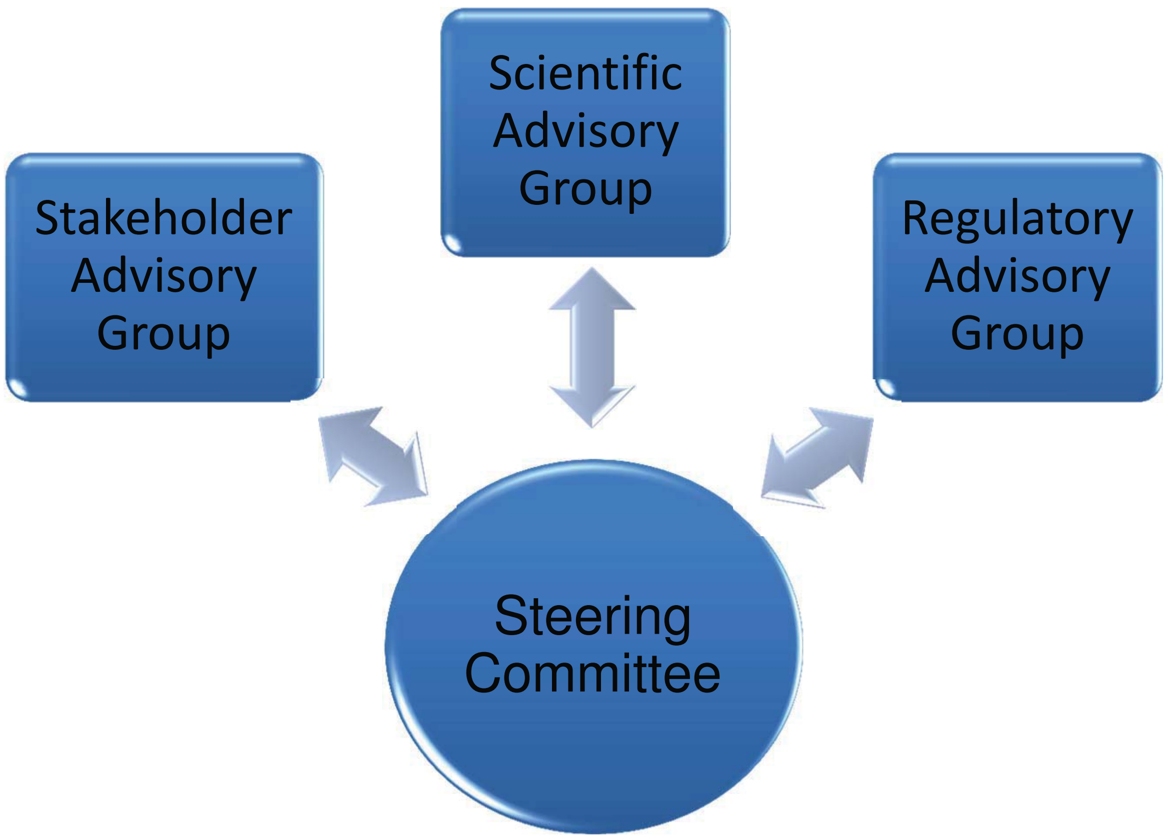
Lab Methods

Quality Assurance Documentation

SWAMP
Surface Water Ambient Monitoring Program

Biological Objectives Workgroups

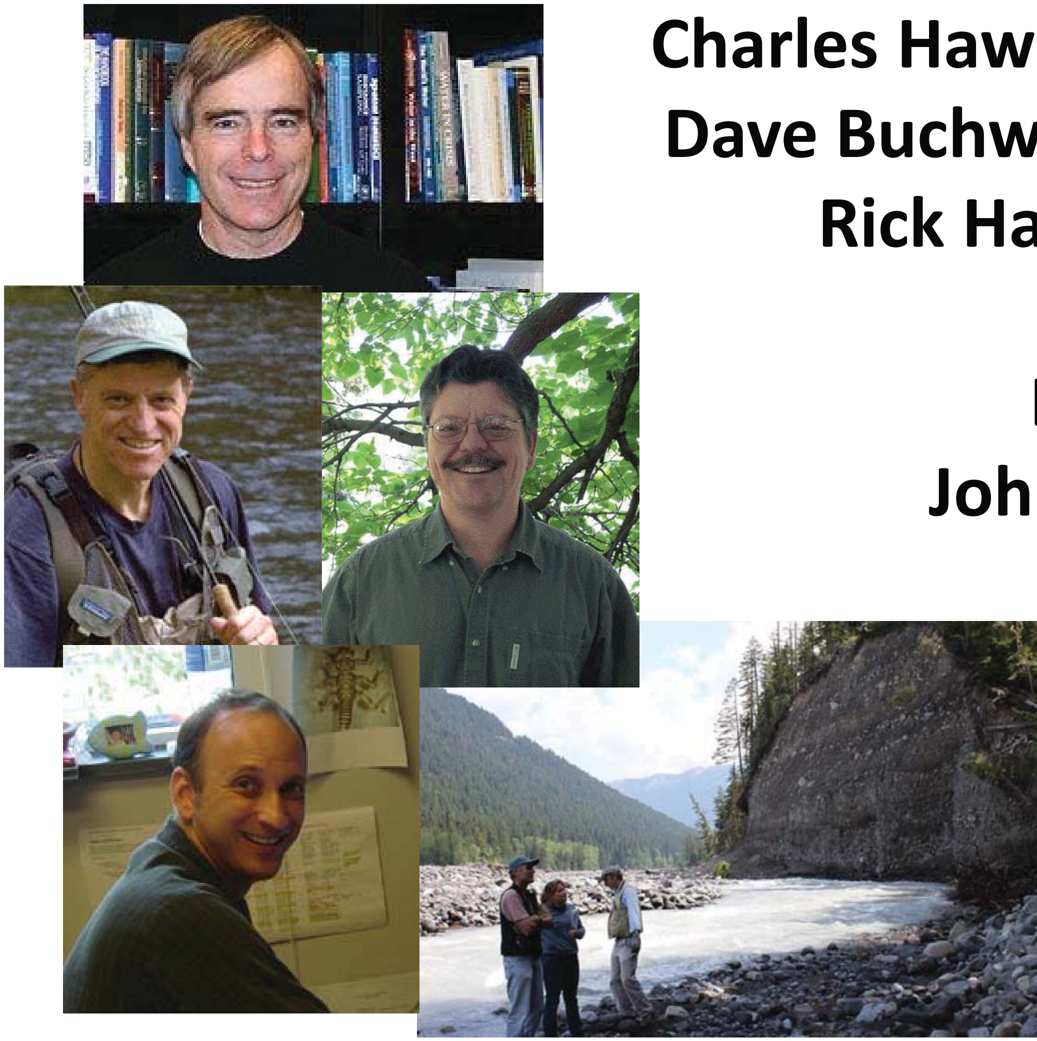
> 20 meetings, excellent feedback



SWAMP
Surface Water Ambient Monitoring Program

Scientific Advisory Panel

Charles Hawkins, *Utah State University*
Dave Buchwalter, *North Carolina State*
Rick Hafele, *Oregon DEQ (retired)*
Chris Konrad, *USGS*
LeRoy Poff, *Colorado State*
John VanSickle*, *EPA (retired)*
Lester Yuan*, *EPA*



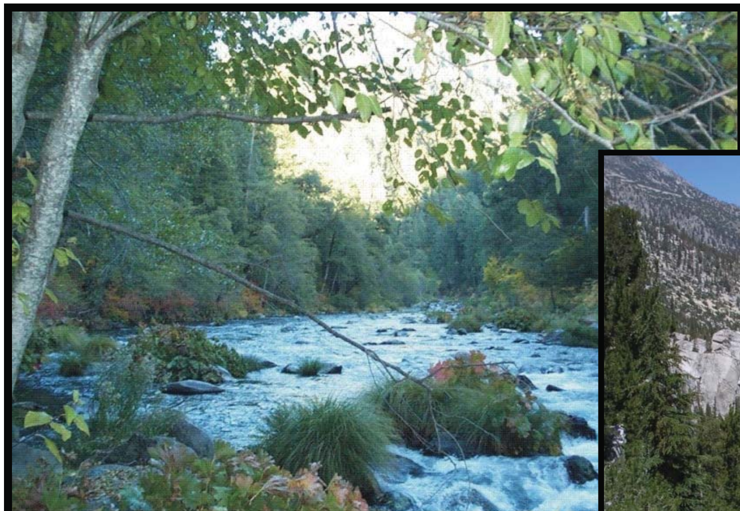
*not pictured

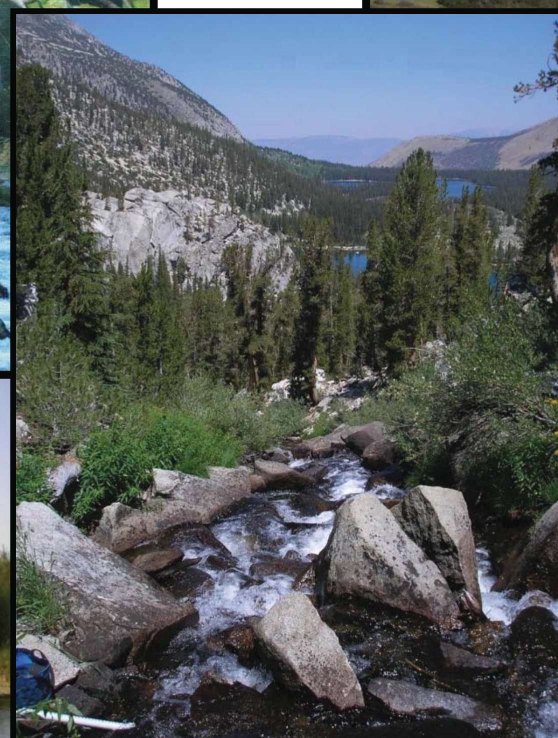
SWAMP Surface Water Ambient Monitoring Program

Scoring Tools Depend on Reference Sites

(sites with low levels of disturbance)

"What should the biology look like at a test site?"



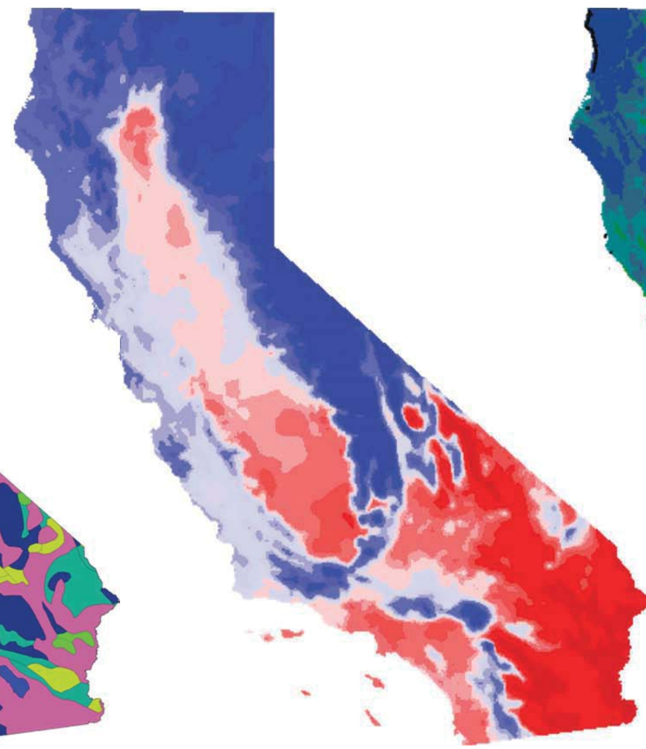


Technical Challenges:

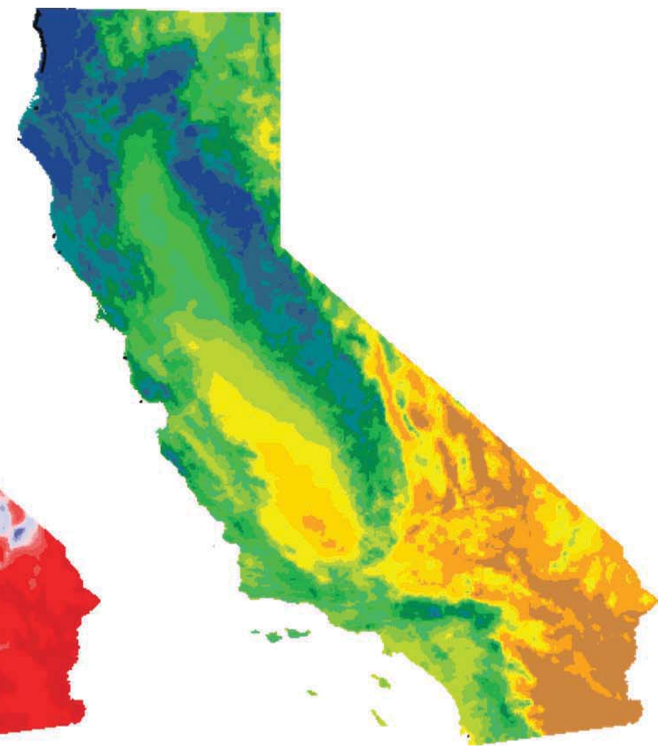
Strong natural gradients result in natural variation in biological expectations

Geology

Temperature

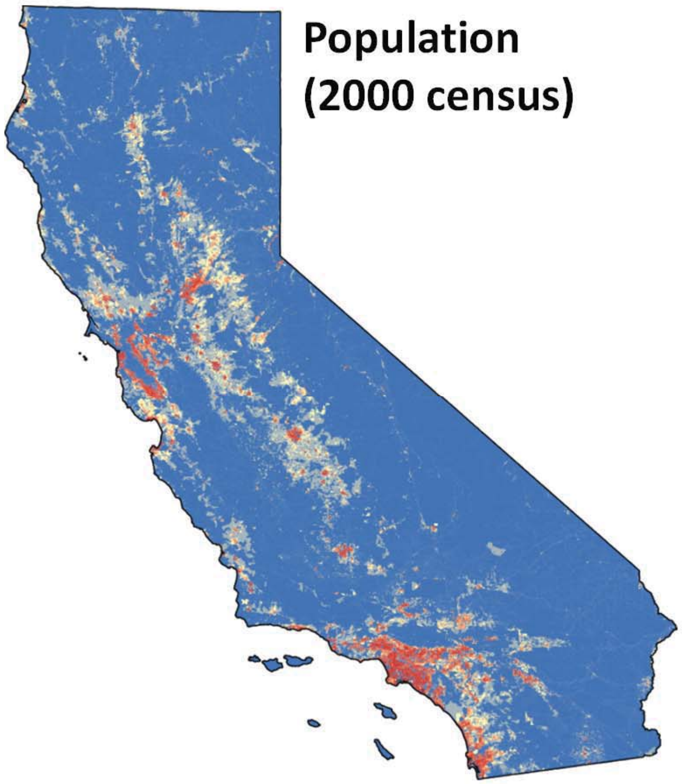


Precipitation



Technical Challenges:

Intense development can create regional gaps



Population
(2000 census)

Agricultural Areas
(2001 NLCD)

Reference Sites for Biocriteria

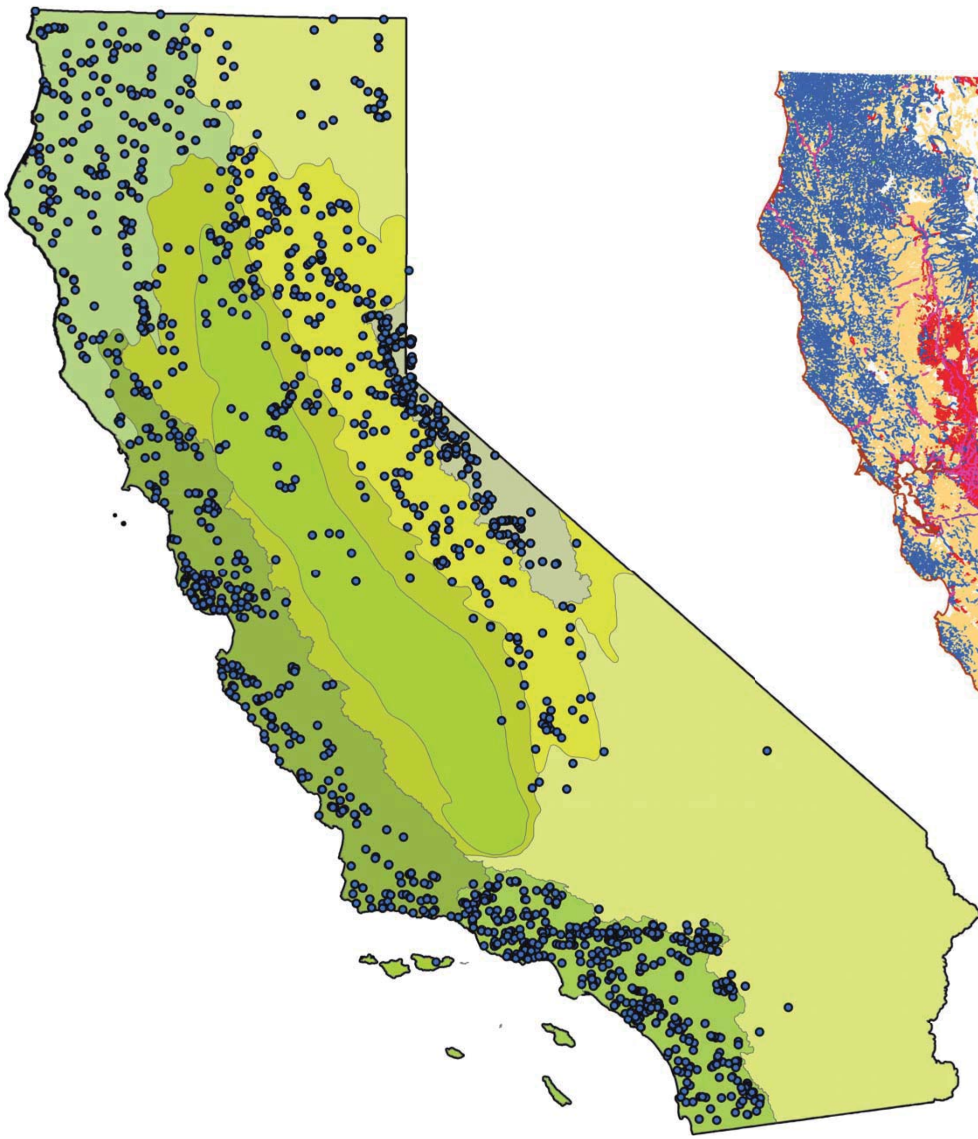
Selecting for site quality and representativeness

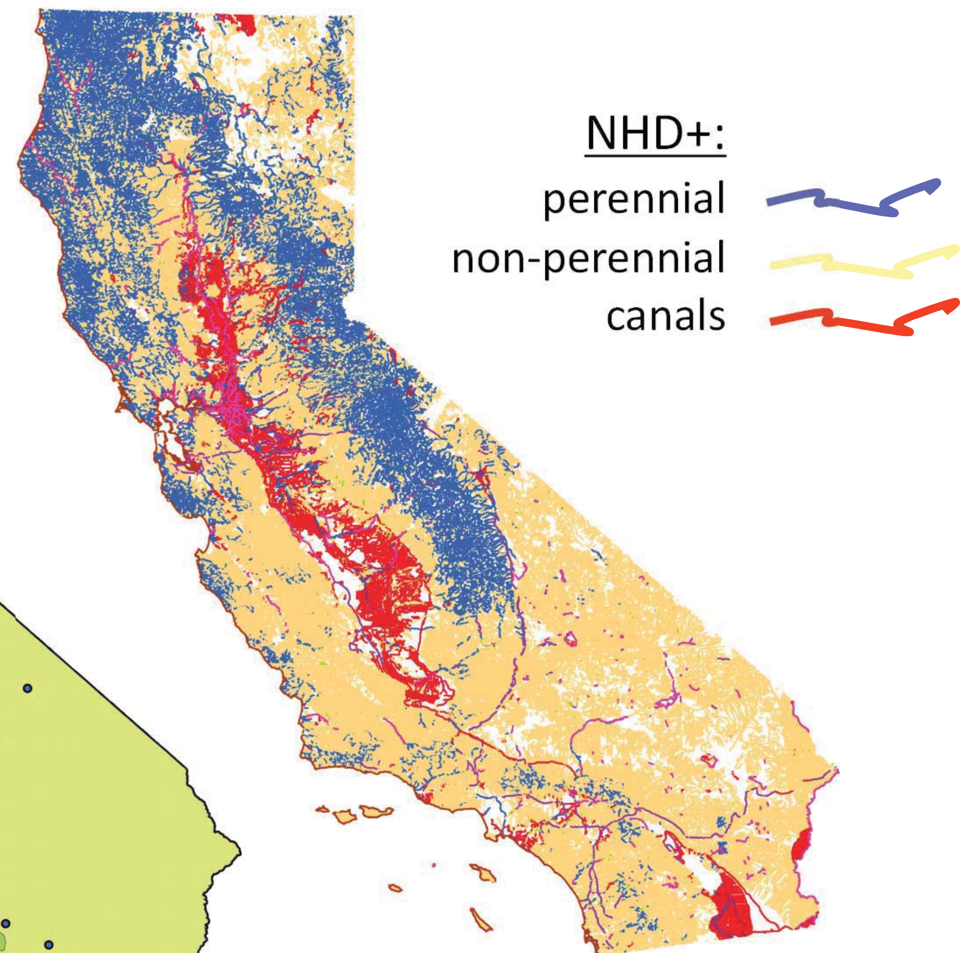
Challenge: Very few (if any) pristine streams exist; site selection process has to maximize representativeness while minimizing amount of disturbance at reference sites

Performance Objectives:

1. Reference pool represents the majority of CA streams
2. Biological "quality" is maintained at reference sites

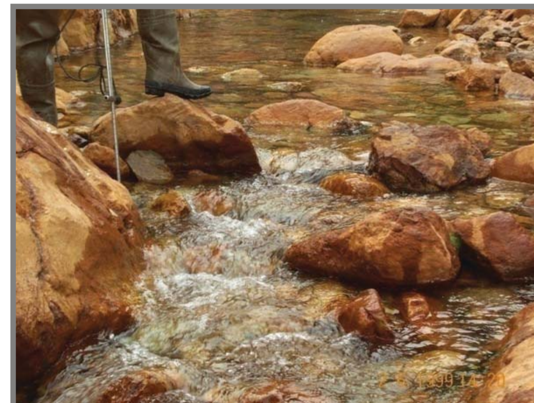
Assemble Data from > 2400 sites

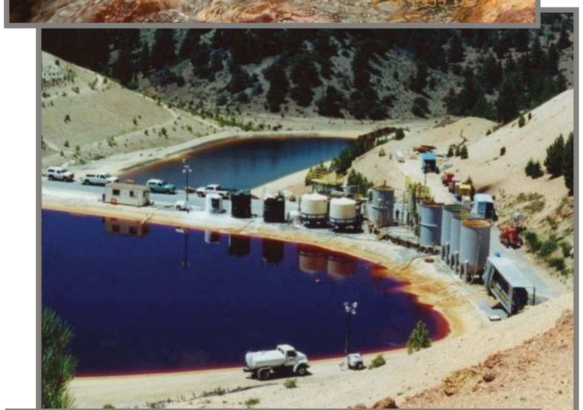


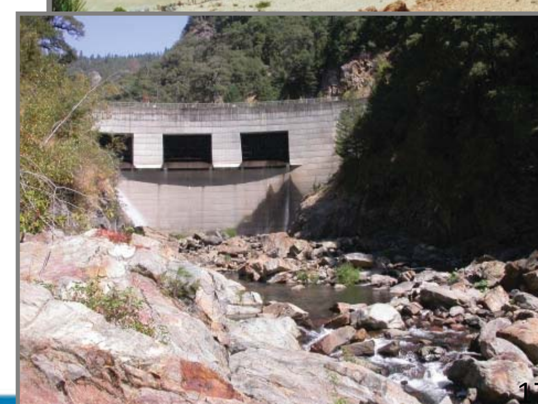


Reference sites have few sources of human stress

- **Infrastructure**: roads, railroads
- **Population**
- **Hydromodification**
  - manmade channels, canals, pipelines
- **Landuse**
  - Ag/Urban development
  - Timber Harvest, Grazing
- Fire history, dams, mines
- 303d list, known discharges
- Invasive invertebrates, plants
- Instream and riparian habitat
- Water chemistry





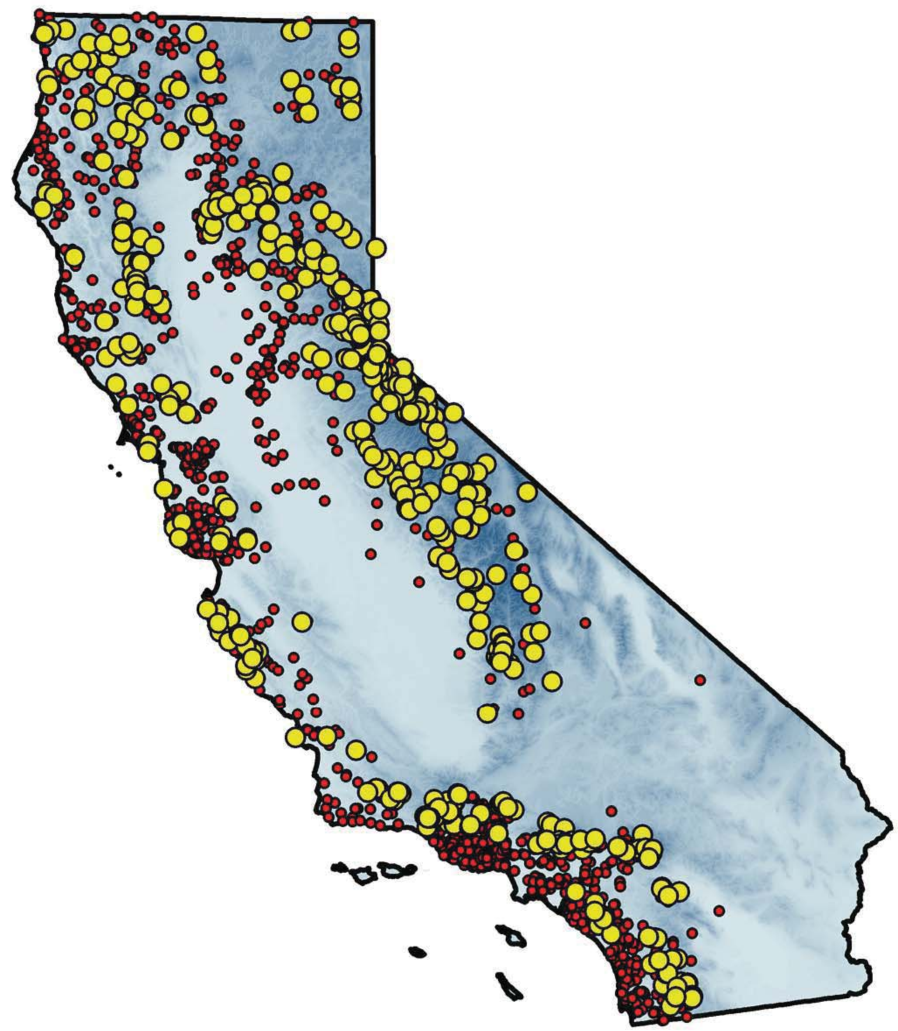


Thresholds are comparable or stricter than other CA indices and include many more criteria

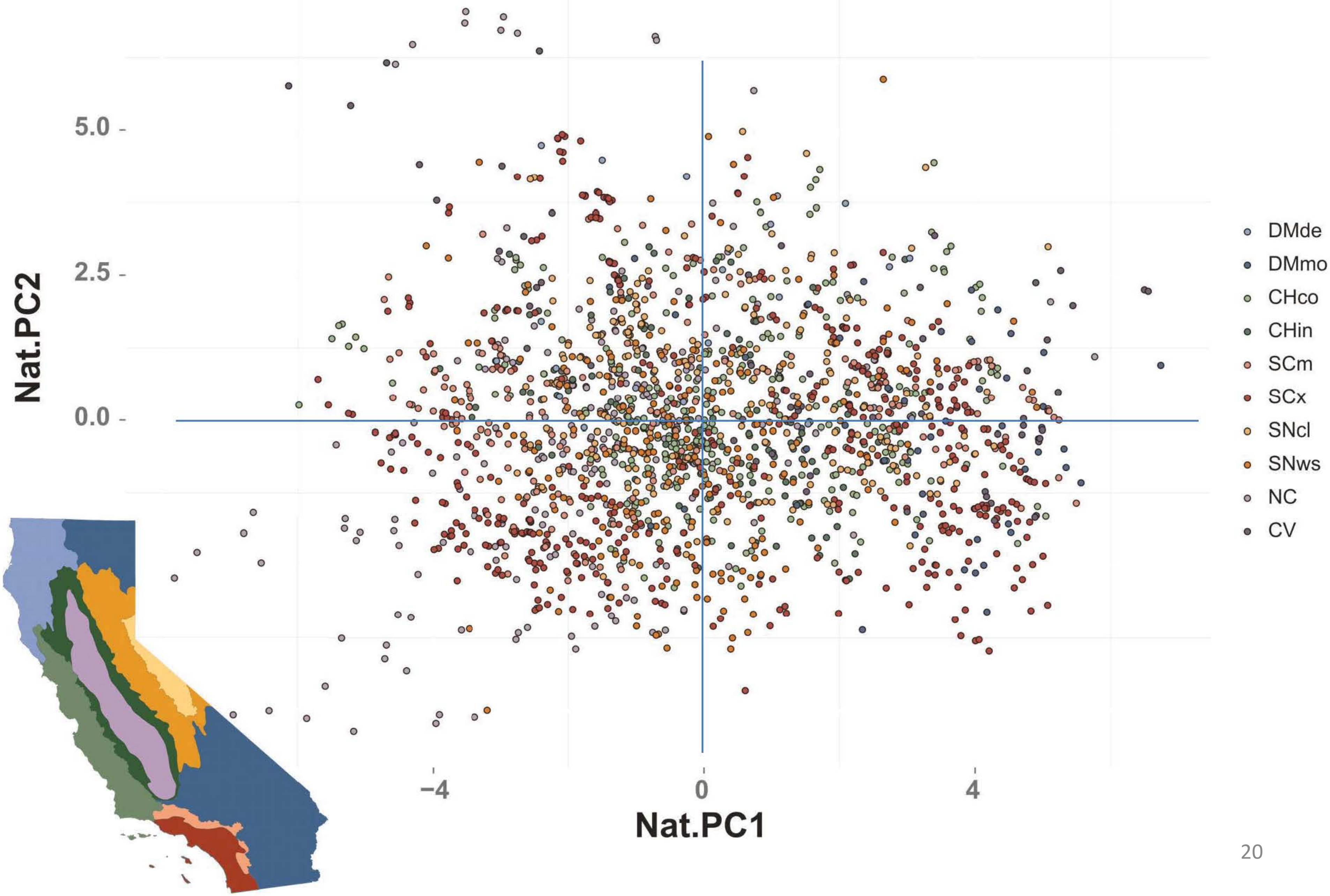
Metric	Bio-Objectives	South Coast IBI	North Coast IBI
Local Disturbance (W1_Hall)	1.5	-	-
% Agricultural	3,3,10	5	5
% Urban	3,3,10	3	3
% Ag + Urban	5,5,10		
% Code 21	7,7,10	in urban	in urban
Road Dens (km/km^2)	1.5	2.0	1.5/ 2.0
Paved Road X-ings (#/ws)	5/10/50		
Nearest Dams	>10 km	-	-
Active Producing Mines	0 (5k)	-	-
% Canals & Pipelines	10	-	-
Gravel Mine Density	0.1 (r5k)		
Conductivity	<2000 uS, + <99%, >1%		
BPJ Screen	X	X	X

Very good geographic coverage

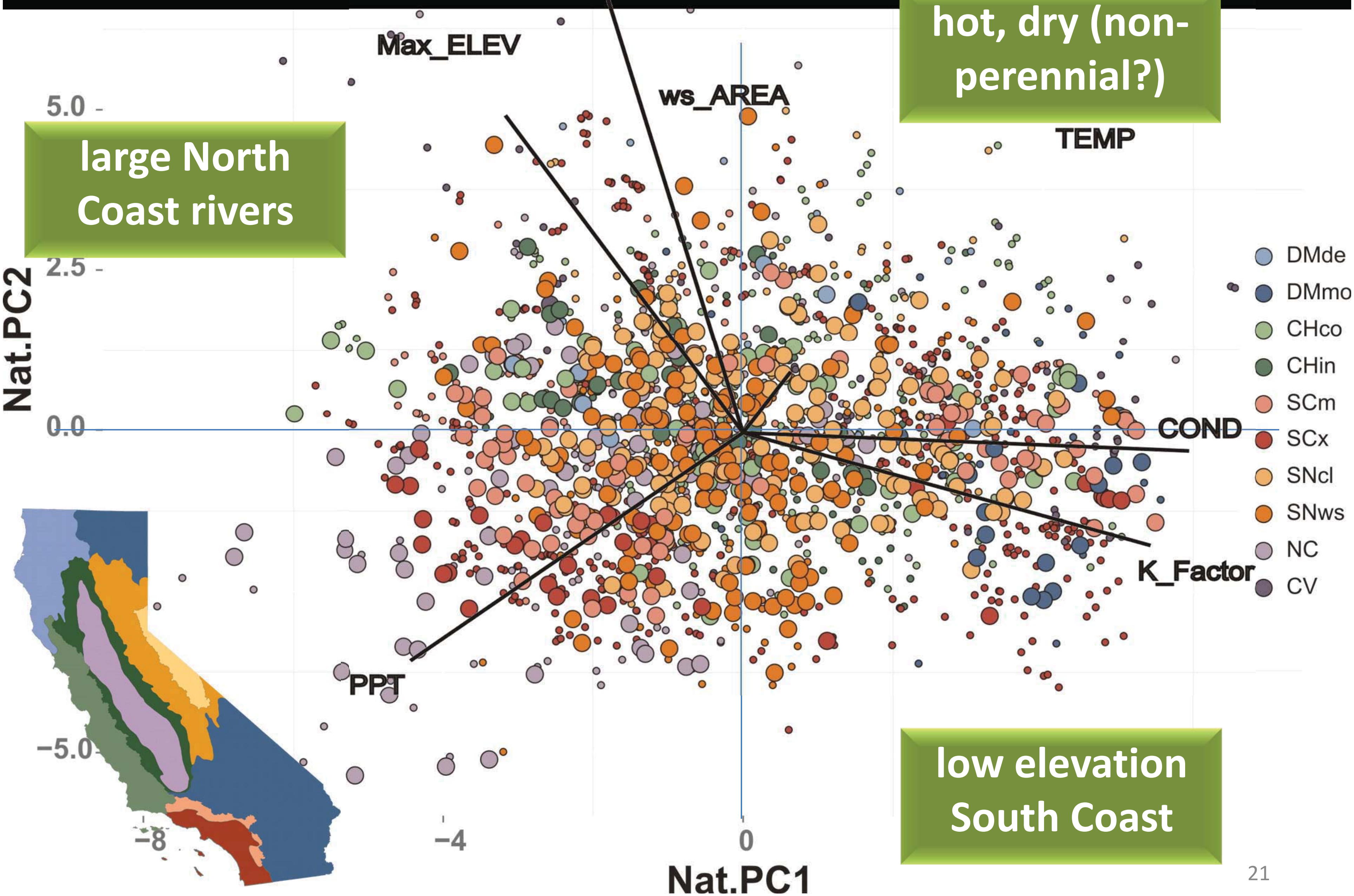
REGION	n
North Coast	75
Central Valley	1
Coastal Chaparral	57
Interior Chaparral	33
South Coast Mountains	85
South Coast Xeric	34
Western Sierra	131
Central Lahontan	114
Deserts + Modoc	27
TOTAL	**586**



Multivariate view of natural diversity



Strong environmental representativeness



B461


photo courtesy John Sandberg

B463

Part III – Supporting Implementation
(technical support for policy decisions)

- Setting Impairment Thresholds
- Ensuring statewide consistency
- Applicability: Objective approaches for setting limits to the tools
- Summary and What's Next

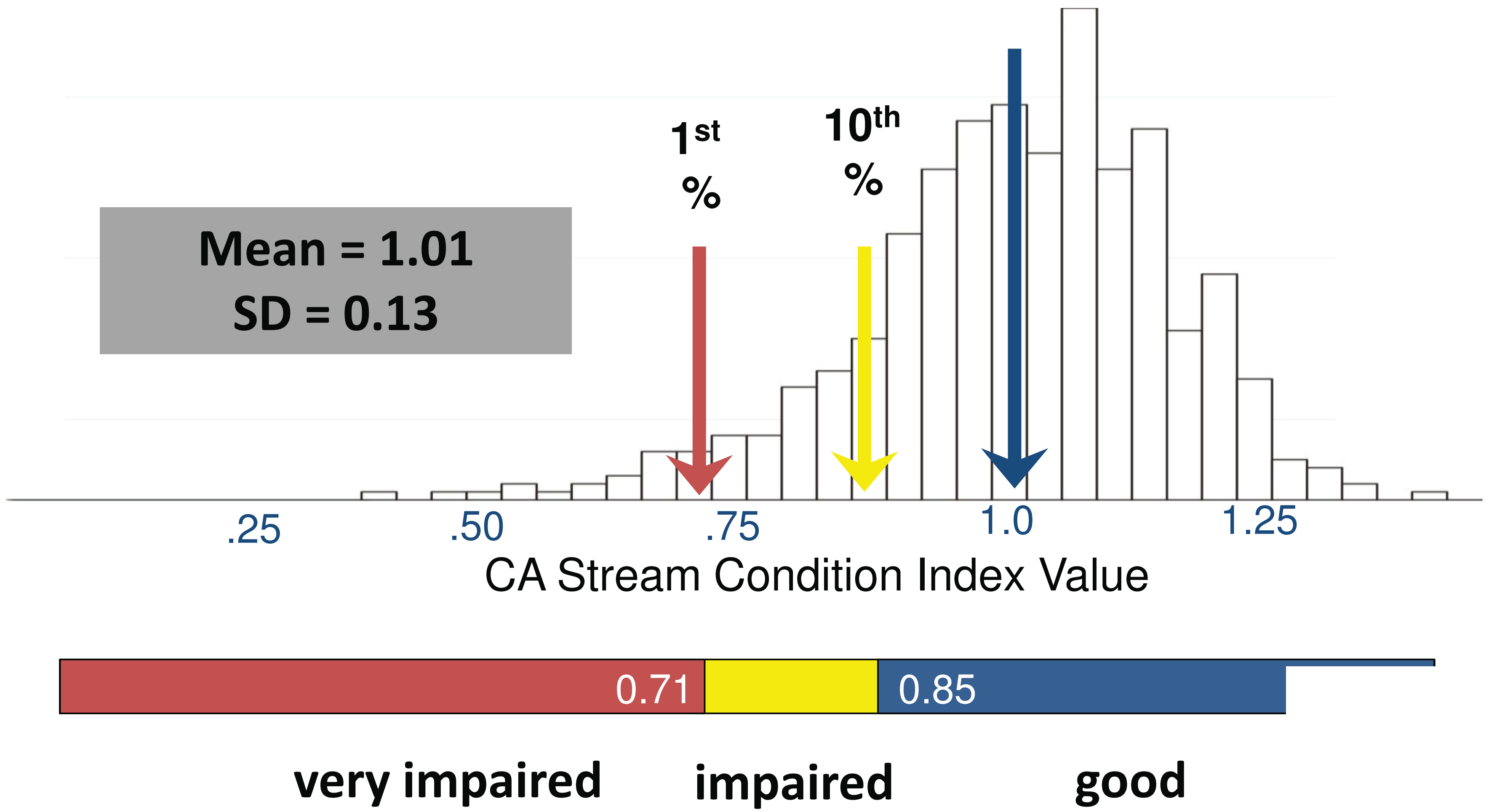
SWAMP
Surface Water
Ambient Monitoring
Program

Desirable Qualities of Regulatory Thresholds

- **Objective**
- **Balance false positives and false negatives** – should be protective of resource, but not over-sensitive
- **Incorporate uncertainty of site score**

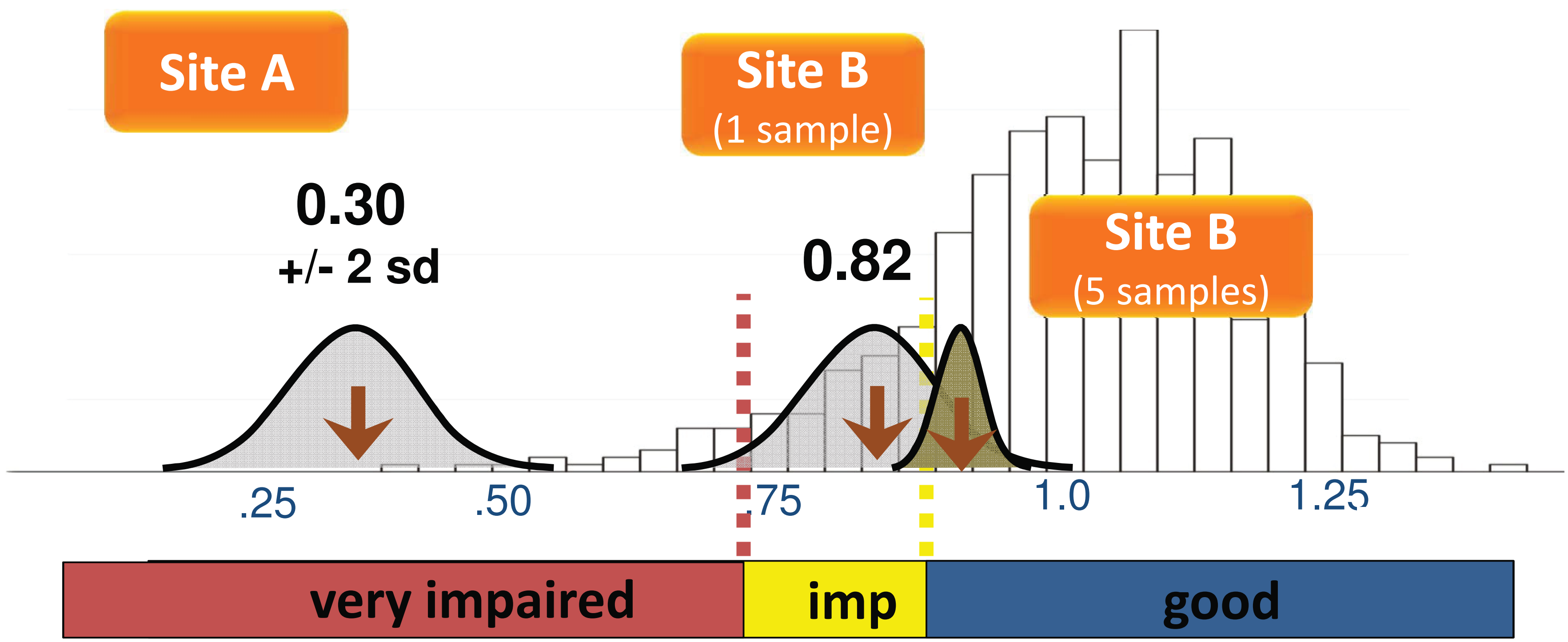
SWAMP
Surface Water Ambient Monitoring Program

Distribution based thresholds:



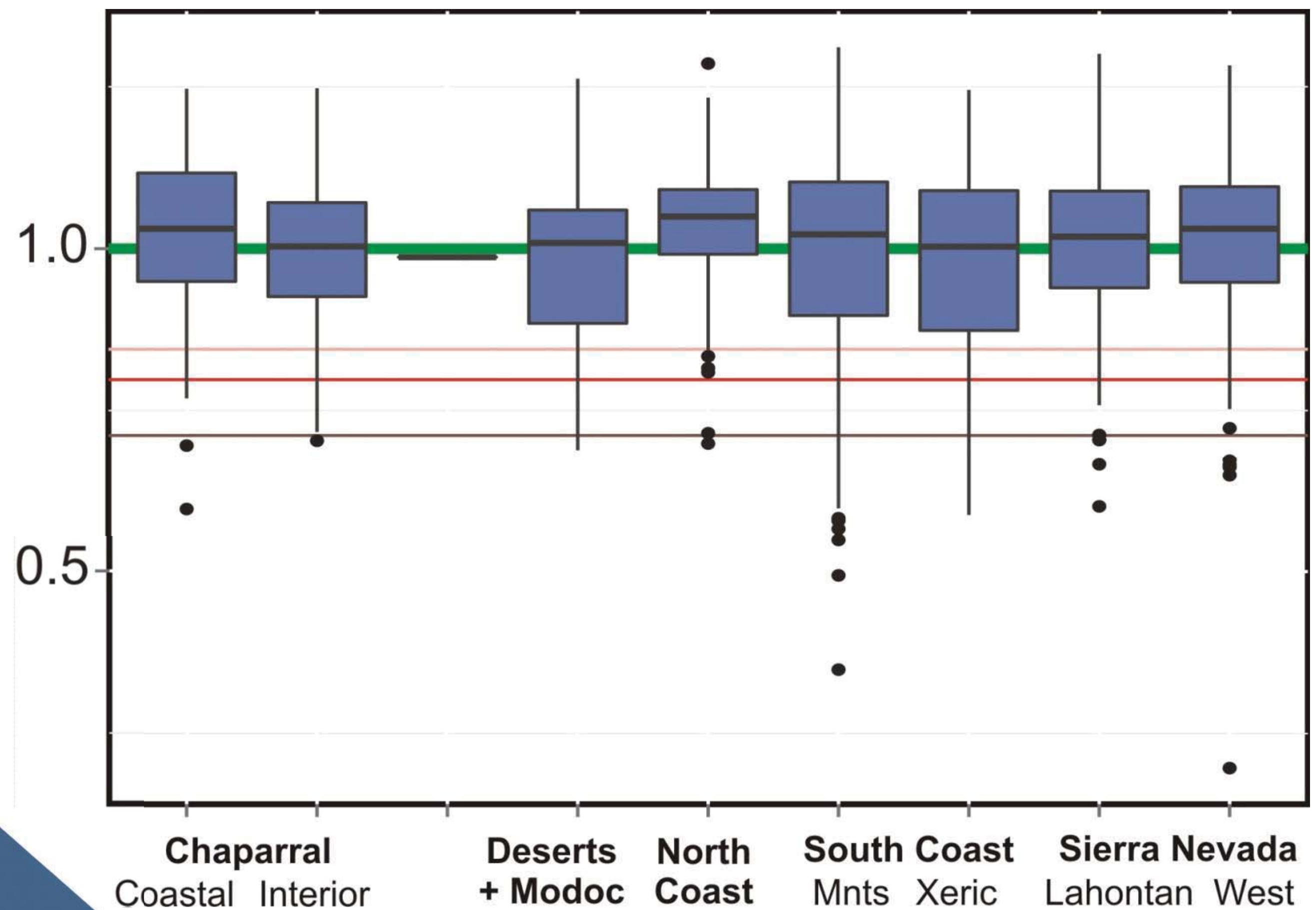
Incorporating Test Site Uncertainty

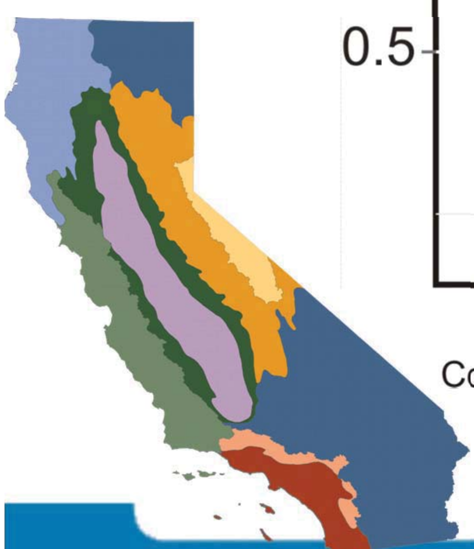
Use within-site error rate to account for uncertainty around test site score



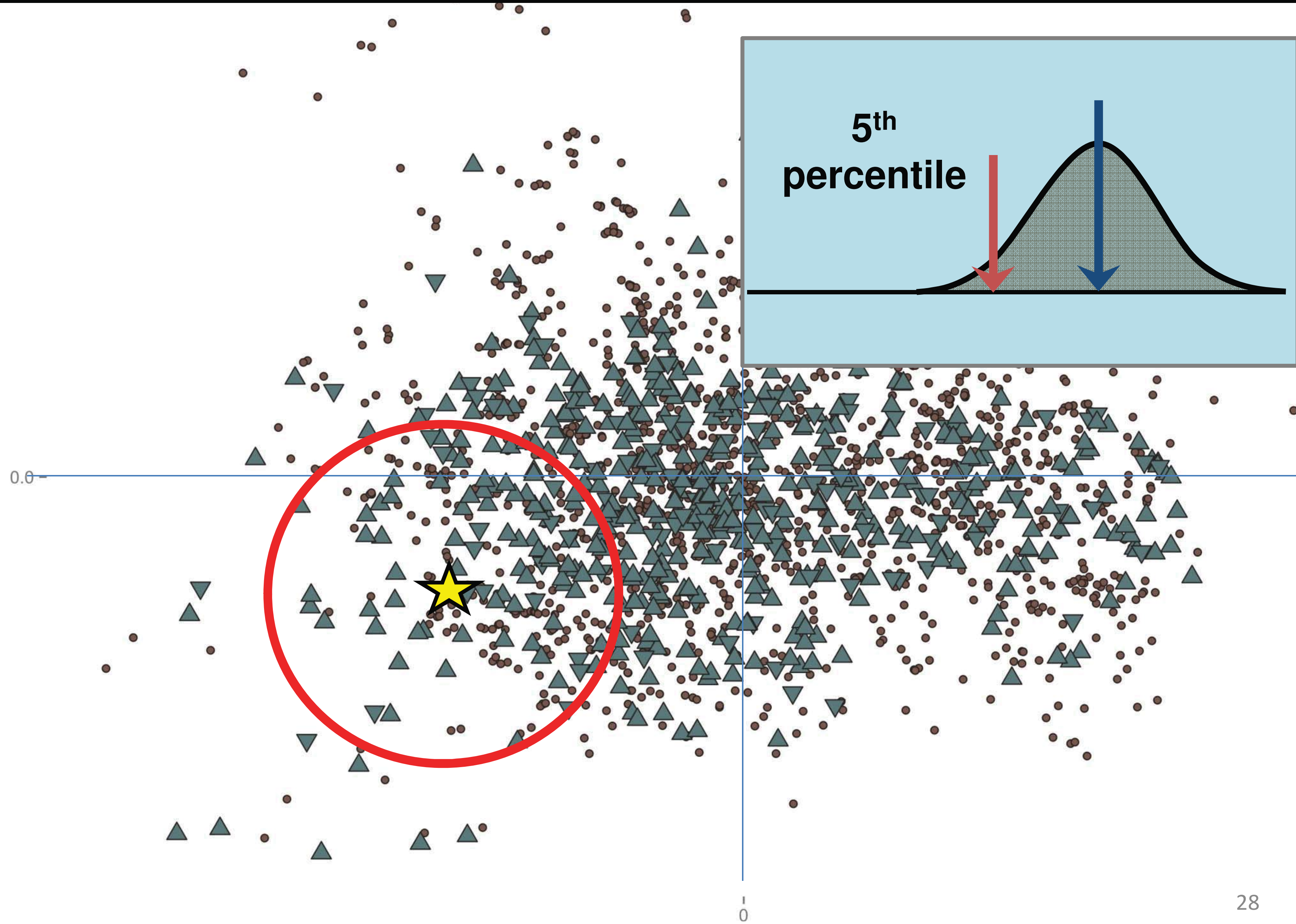
more certainty with multiple samples

Ensuring Regionally Consistent Thresholds





Enhancing threshold consistency



B468

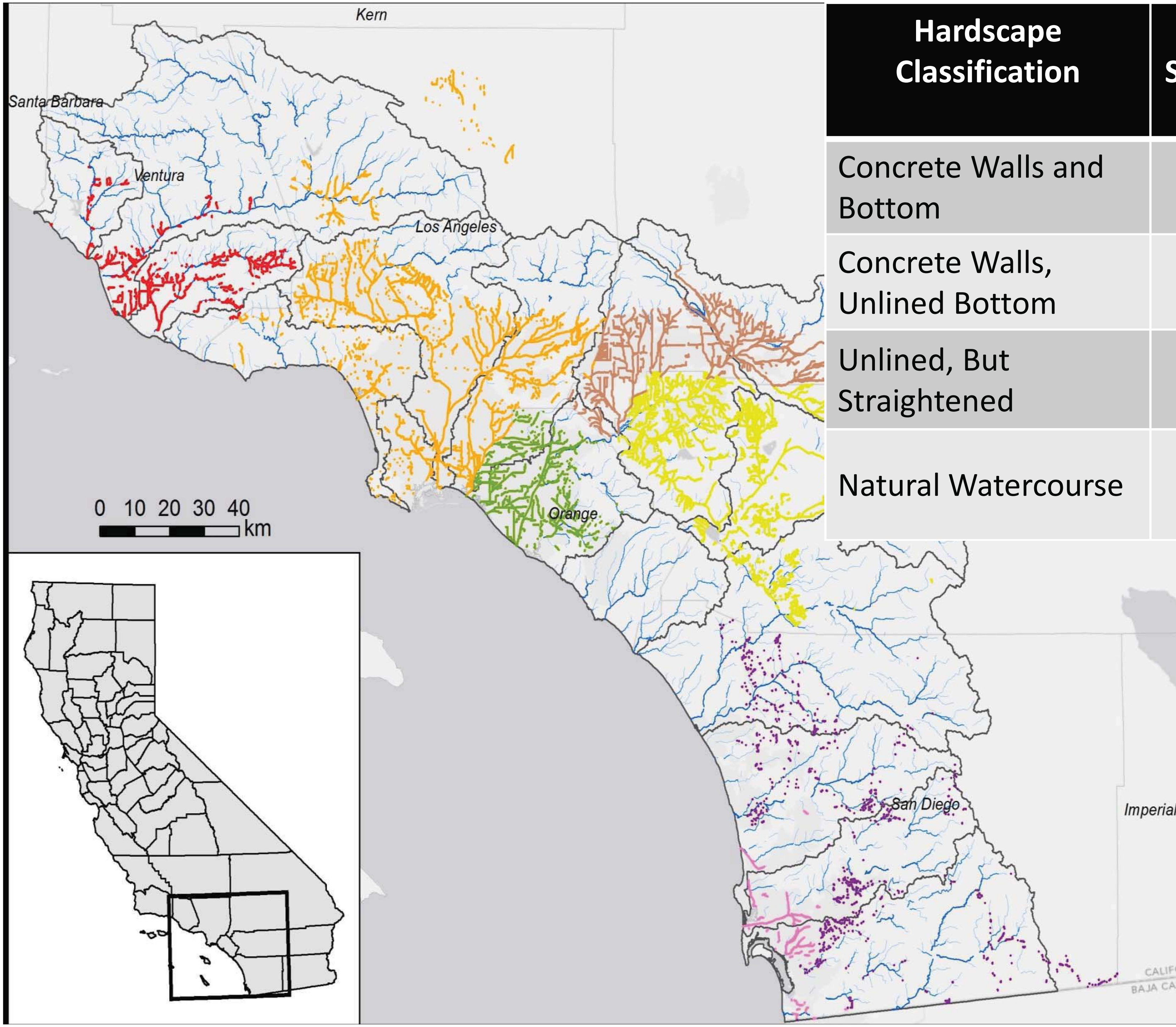
Where can we apply the CSCI?

- **Categorical** = exception classes in policy
  - Excepted regions (e.g., Central Valley)
  - Excepted waterbody types (e.g., modified channels)

- **Quantitative Approaches**

SWAMP
Surface Water Ambient Monitoring Program

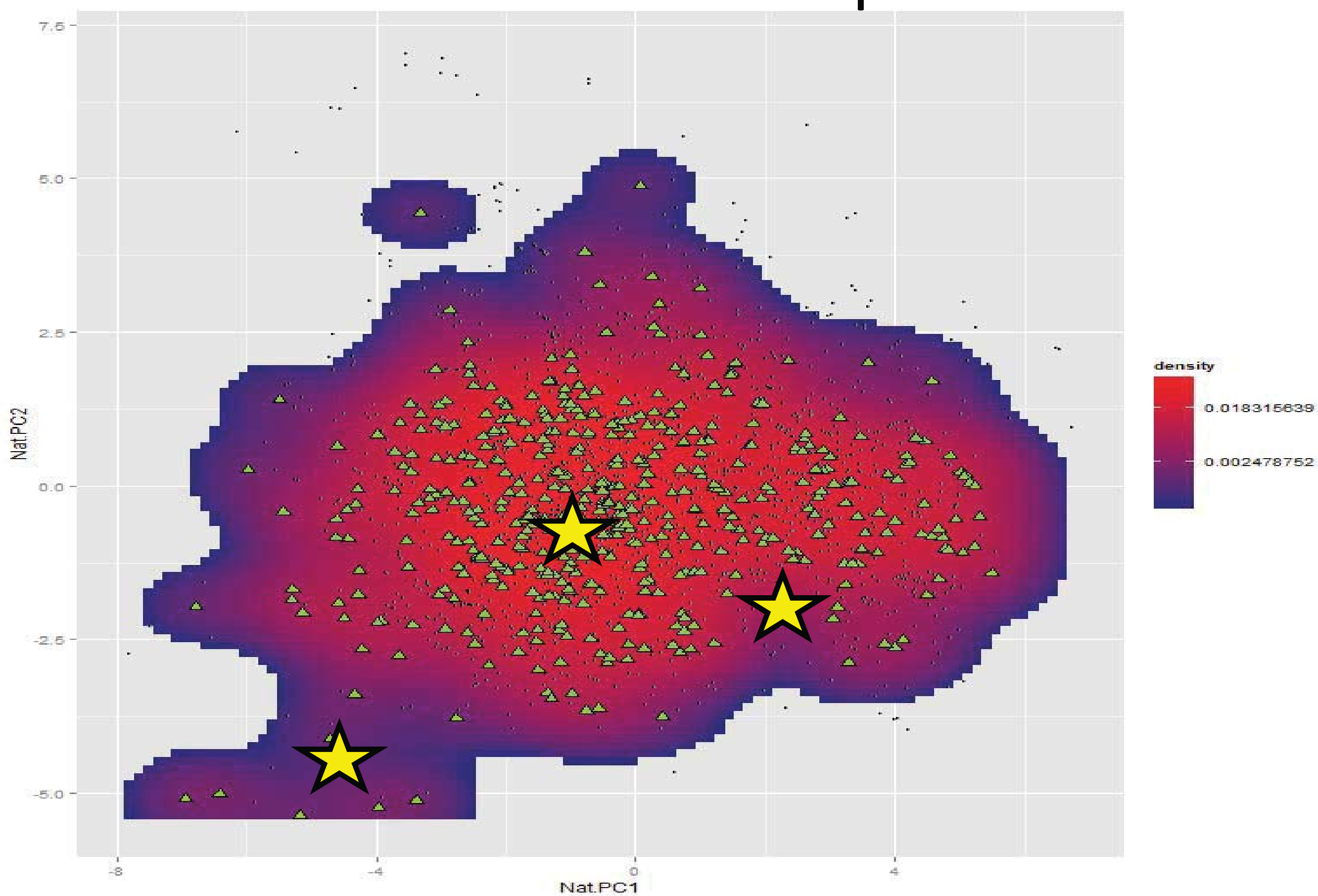
B470



Hardscape Classification	% of So Cal Stream Miles
Concrete Walls and Bottom	3.2%
Concrete Walls, Unlined Bottom	3.6%
Unlined, But Straightened	16.6%
Natural Watercourse	76.3%

Quantitative Approaches:

"is a test site within the experience of the model in environmental space?"



Could be used to establish exceptions for truly unique environmental settings

B472

Applicability of the CSCI in exception class settings

- We can still use the CSCI as a ruler, but we won't regulate based a reference-based threshold
- Could use "best attainable" approach instead of "reference" to set expectation, or use to compare among sites

B473

Automation and Documentation

STANDARD METHODS … available on SWAMP website

AUTOMATE calculations

- Package GIS layers
- Make standard calculation and reporting tools available via **CEDEN**

Document, document, document

- Journal articles
- Website 101 and FAQ
- Website appendices

Questions?

Appendix 14

Science Advisory Panel Response

18 October 2012

Prioritization of Issues

- Scoring Tool
- Determining Scoring Tool Coverage
- Thresholds
- Causal Assessment
- Regulatory Guidance

Scoring Tool

- The Panel supports the use of the hybrid scoring tool
  - Includes both species-specific and biological community responses
  - Science Team will need to work on developing a simple explanation of this tool
- Some additional model evaluation would build confidence
  - Independent validation data sets
  - Simulation of impairment data to test responsiveness
  - What are the actual taxa or metrics that are driving scoring tool disagreement
- The Team will need to automate the calculation of this more complex scoring tool

Determining Scoring Tool Coverage

- A multi-variate approach is preferred
  - Include core dimensions of natural variability that biology responds to
- Guidance should be developed for stakeholders who assert their stream is not covered by this tool
- Even for sites outside the experience of the models, assessment options are still available
  - i.e., upstream-downstream

Thresholds

- Select thresholds based on distributions of reference condition
  - Need to assure some ecological meaningfulness
  - This approach can be used for developing categories of impact
- Test site uncertainty should be included
- Incorporating multiple samples at the test site is preferable: two options
  - Binomial approach (frequency of exceedence)
  - Mean site condition
- Ensure condition is assessed consistently at all sites

Causal Assessment

- Causal Assessment is important for progress in bio-objectives development
  - Panel recognizes that CADDIS is an imperfect tool and needs refinement
- CA needs to take advantage of its large data set to streamline causal assessment
  - This unique opportunity should reduce future costs
- CA needs to improve comparator site selection
  - Incorporate comparators outside the watershed
- CA needs to improve diagnostic tools
  - Regional response models (i.e., Relative risk)
  - Species specific response models
  - Laboratory based species sensitivity distributions

Regulatory Guidance

- CA's working definition of "perennial" and "wadeable" seem appropriate
  - This definition is the foundation of the scoring tools
- Inference of segment-scale biological condition from a single site should be done cautiously
  - Additional samples at multiple locations may be needed

B483

B484

Appendix 15

WORK PLAN

Predicting Biological Integrity of Streams Across a Gradient of Development in California Landscapes

Raphael Mazor, Martha Sutula, Eric Stein (SCCWRP)
Andy Rehn and Pete Ode (CDFW)

Introduction

The California State Water Resources Control Board (State Water Board) is developing a combined Biostimulatory (nutrient) and Biointegrity policy for wadeable streams, hereto referred to as the Biostimulatory-Biointegrity Project. The scientific approach supporting this project is grounded in biological assessments of the health of benthic macroinvertebrate and algal communities. The State is supporting the use of standardized bioassessment indices to quantify the biological integrity and support of aquatic life uses in wadeable streams. The benthic macroinvertebrate index (i.e., the California Stream Condition Index, or CSCI) has previously been developed (Mazor et al. 2016). An algal stream condition index (ASCI) is currently under development, with a provisional ASCI expected fall 2017 (see ASCI work plan).

As natural landscapes are converted to support urban or agricultural uses, the underlying hydrologic, physical, and biogeochemical factors within the stream and its catchment that support healthy stream communities are altered, potentially harming aquatic life. Developed landscapes are associated with an increase of many stressors in streams, such as elevated contaminant and nutrient concentrations, altered flow regimes, sedimentation, and habitat degradation (e.g., Waite et al. 2012). In some streams, direct channel modifications (e.g., bank armoring) may also limit opportunities to sustain high-quality ecological conditions for aquatic life. In these highly developed settings, the large number of linked stressors may prevent a stream from supporting its beneficial uses or attaining high scores on indices of biological condition. Often, these stressors are difficult to mitigate or remove under the traditional mechanisms available to the Water Boards. In these circumstances, the range of CSCI and/or ASCI scores may be constrained, but targeted restoration could improve conditions. Key technical questions underpinning the range of options and prioritization of management actions for wadeable streams along the continuum from undeveloped to highly developed landscapes found within California are: For which streams is biological integrity constrained by development in the catchment? How can they be identified and mapped? What are the ranges of biological conditions these developed landscapes can support?

The State Water Board is seeking to protect biointegrity in streams, including streams where integrity is constrained by development. Identifying landscapes where development has a high likelihood of limiting biointegrity is an important first step to identifying effective management options. This creates a technical need for 1) a simple, reproducible, and easy-to-understand methodology for identifying landscapes where development has a high likelihood of limiting biointegrity and 2) predicting expectations for CSCI and (when available) ASCI indices screening tool or starting point for discussions on appropriate management strategies. These analyses create a technical foundation for the State Water Board and the Regional Boards to protect biological integrity in streams by informing appropriate biological condition expectations or by prioritizing sites long and short term restoration activities in these landscapes.

Geographic information systems (GIS) are commonly used to quantify landscape development within stream catchments. Estimating landscape alteration in catchments has traditionally been time-consuming for large-scale programs that monitor hundreds of sites annually, but recent tools (i.e., STREAMCAT, Hill et al. 2015) have made it possible to rapidly estimate landscape alteration in all streams in California represented by the National Hydrography Dataset Plus (NHD Plus) stream network. STREAMCAT therefore presents an opportunity to model the influence of landscape alterations on stream bioassessment scores on a large scale, and to apply predictions of these models to any stream represented in NHD Plus. These models have the potential to predict a range of likely scores in a stream given a degree of landscape alteration, setting the stage for policy discussions about the level of support that these streams in developed landscapes provide to beneficial uses.

Study Objective, Conceptual Approach to Model Landscape Influences on Stream Bioassessment Index Scores

The purpose of this study is to explore constraints on bioassessment index scores in streams across a continuum of landscape development, using a predictive, GIS approach. Key graphics from this analysis will be used to support discussions between the Water Board and its Regulatory and Stakeholder Advisory Groups on policy options to prioritize and improve the management of streams in developed landscapes.

The GIS approach involves developing models that predict a range of bioassessment index scores based on measures of landscape development. The product of these models is a map of likely CSCI (and when available, ASCI) scores for each segment. The intent of such a map is to identify watersheds where discussions of policy options for undeveloped versus developed landscapes could be productive. This map is intended to be used as a screening tool or starting point for discussions; it is not intended to be a one-off, definitive assessment that is used to set expectations for developed landscapes without further field level investigations of stressors and causal factors.

This approach relies on the following definition of “developed landscapes”:

> *Landscapes where development is likely to limit bioassessment index scores.*

Development of a GIS model and application to predict likely bioassessment index scores in developed landscapes require three types of decisions:

1. *Developed Land Uses*. Developed landscapes can be characterized by variables in the STREAMCAT dataset related to human alterations, such as urban and agricultural land-use types in the National Land Cover Dataset, land cover imperviousness, etc. (Table 1). Other variables could be included or excluded, but must be limited to variables included in or easily added to STREAMCAT.
2. *Likelihood*. The likelihood of achieving the desired biological condition can be calculated by statistical models, but determining if a likelihood is low enough to be considered "unlikely" is a value-based (i.e., non-technical) decision.
3. *Desired biological condition:* The management objective, as defined by bioassessment index scores, here to referred to as "assessment endpoints" (Biostimulatory-Biointegrity Project Science Plan).

Decisions on which developed landscape variables to include must occur during model development, while discussion of values appropriate to set the likelihood and desired CSCI and ASCI assessment endpoints are model application questions, all of which will ultimately be made by the Water Board. In order to foster discussion and provide the regulatory (RG) and stakeholder advisory groups (SAG) an opportunity to provide feedback on these three decisions, the Technical Team will iteratively engage the RG and the SAG in the model development and model application phases to provide ample opportunity for this feedback to occur.

Scope of Work:

The study has three tasks:

1) Develop models to predict a range of CSCI and ASCI scores based on measures of landscape development from the STREAMCAT dataset;
2) Apply the models to the entire NHD Plus stream network represented in the STREAMCAT dataset, classify stream segments based on likelihood of achieving target scores, and create maps illustrating these classifications, in order to engage Water Board staff and advisory groups on decisions on likelihood and CSCI and ASCI assessment endpoints; and
3) Produce a technical memo with key graphics and model output.

Task 1. Develop models to predict a range of CSCI and ASCI scores based on measures of landscape development derived from the STREAMCAT dataset

A dataset representing CSCI scores across a range of site conditions in California will be aggregated. Index scores from each site will be snapped to the corresponding stream segment in NHD Plus. STREAMCAT data characterizing landscape alteration variables (e.g., percent urban land cover, percent cropland, catchment imperviousness; Table 1) will be associated with each bioassessment site. Appropriate statistical models (e.g., quantile random forest) will be calibrated to associate measures of landscape development with bioassessment scores. A models will also be developed for ASCI, though decisions on which land use variables and likelihood values to use will focus on CSCI only, since a provisional ASCI index is anticipated late stage (Fall 2017).

Water Board staff will make provisional decisions on land use variables to include. The initial proposal will be based on consultation with the RG. The proposed land use variables and rationale will be presented to the SAG for feedback.

Deliverable:

1.1 Draft models and related graphics to predict bioassessment scores, in iterative stages of feedback.

1.2 Descriptive summaries of models, including evaluations of model performance, and list of landscape development variables in STREAMCAT selected for use in the models.

Table 1. List of STREAMCAT variables that can be evaluated in landscape modeling exercise. Most of these variables are calculated at multiple spatial scales.

Potential variables	Description
CanalDens	Density of NHDPlus line features classified as canal, ditch, or pipeline (km/ square km)
DamDens	Density of georeferenced dams (dams/ square km)
DamNrmStor	Volume all reservoirs (NORM_STORA in NID) per unit area (cubic meters/square km)
HUDen2010	Mean housing unit density (housing units/square km)
MineDens	Density of mines sites and within 100-m buffer of NHD stream lines (mines/square km)
PctAg2006Slp10	% area classified as ag land cover (NLCD 2006 classes 81-82) occurring on slopes ≥ 10%
PctAg2006Slp20	% area classified as ag land cover (NLCD 2006 classes 81-82) occurring on slopes ≥ 20%
PctCrop2006	% area classified as crop land use (NLCD 2006 class 82)
PctHay2006	% area classified as hay land use (NLCD 2006 class 81)
PctImp2006	Mean imperviousness of anthropogenic surfaces
PctUrbHi2006	% area classified as developed, high-intensity land use (NLCD 2006 class 24)
PctUrbLo2006	% area classified as developed, low-intensity land use (NLCD 2006 class 22)
PctUrbMd2006	% area classified as developed, medium-intensity land use (NLCD 2006 class 23)
PctUrbOp2006	% area classified as developed, open space land use (NLCD 2006 class 21)
PopDen2010	Mean populating density (people/square km)
RdCrs	Density of roads-stream intersections (2010 Census Tiger Lines-NHD stream lines) (crossings/square km)
RdDens	Density of roads (2010 Census Tiger Lines) (km/square km)

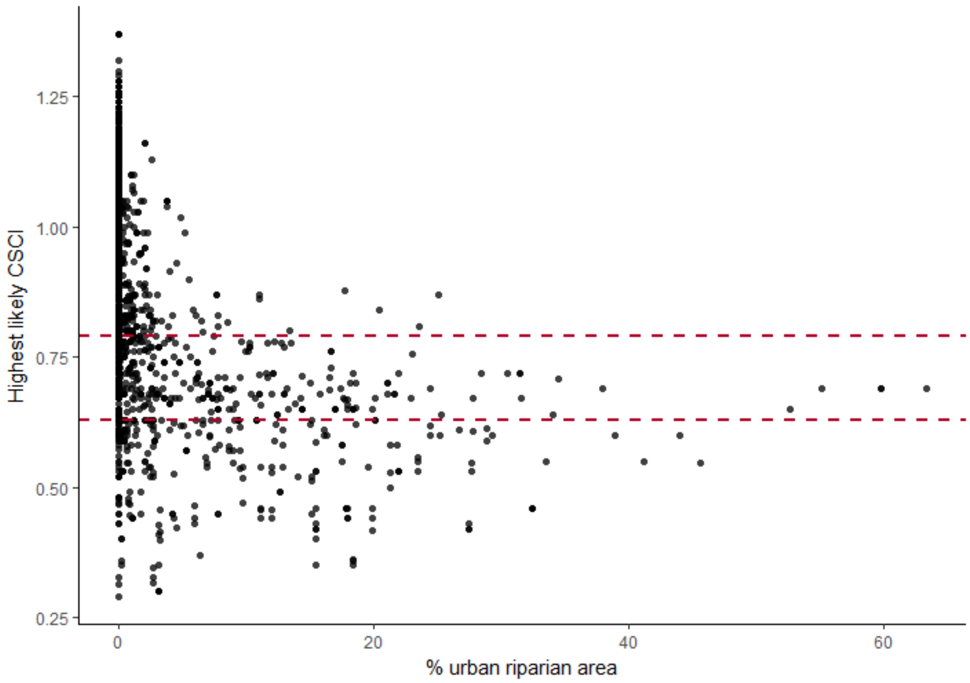


Figure 1. An example of the highest likely CSCI scores predicted by a quantile random forest model relating developed land use variables to biological integrity. The x-axis is percent of high density urban land cover within a 100-m buffer around the NHD stream lines (one of the variables included in example model). Dots above the top red line represent sites that are unconstrained by development (in this example, >10% chance of CSCI scores > 0.79). Dots between the two red lines are moderately constrained by development (<10% chance of CSCI scores > 0.79). Dots below the bottom red line are highly constrained by development (<10% chance of CSCI scores > 0.63). In this example, the 90th percentile of predicted scores represents the highest likely CSCI score.

Task 2. Apply the models to engage Water Board staff and advisory groups on discussions of sensitivity of model output to choice of likelihood and assessment endpoint

The purpose of this task is to help State Water Board staff and advisory groups understand how choice probabilities used to define modeling likelihood and desired assessment endpoint affects mapped categories of streams. The GIS mapping methodology will be applied to entire NHD Plus network of streams in California included in the STREAMCAT database. For selected regions or watersheds, the influence of key decision-points (e.g., minimum thresholds for acceptable bioassessment index scores, or minimum acceptable likelihood for attainment of these thresholds) will be illustrated by showing how the decisions described above influence the percentage and spatial extent of the stream network within the developed category. For example, the Water Board may define constrained channels as those with less than a 10% chance to achieve a CSCI score above 0.63 (e.g., dots below the bottom dashed line in Figure 1); maps will then be generated across the state to highlight which streams are designated as constrained under this definition, thereby helping stakeholders see the implications of this classification for their watersheds (e.g., segments shown as red lines in Figure 2).

Deliverable: 2.1 Interactive maps and oral presentations with maps, graphics, summary tables of the stream drainage network showing model outputs, e.g., maximum score likely to be attained in each stream segment (Figures 2 and 3) as a function of choice of likelihood and assessment endpoint value.

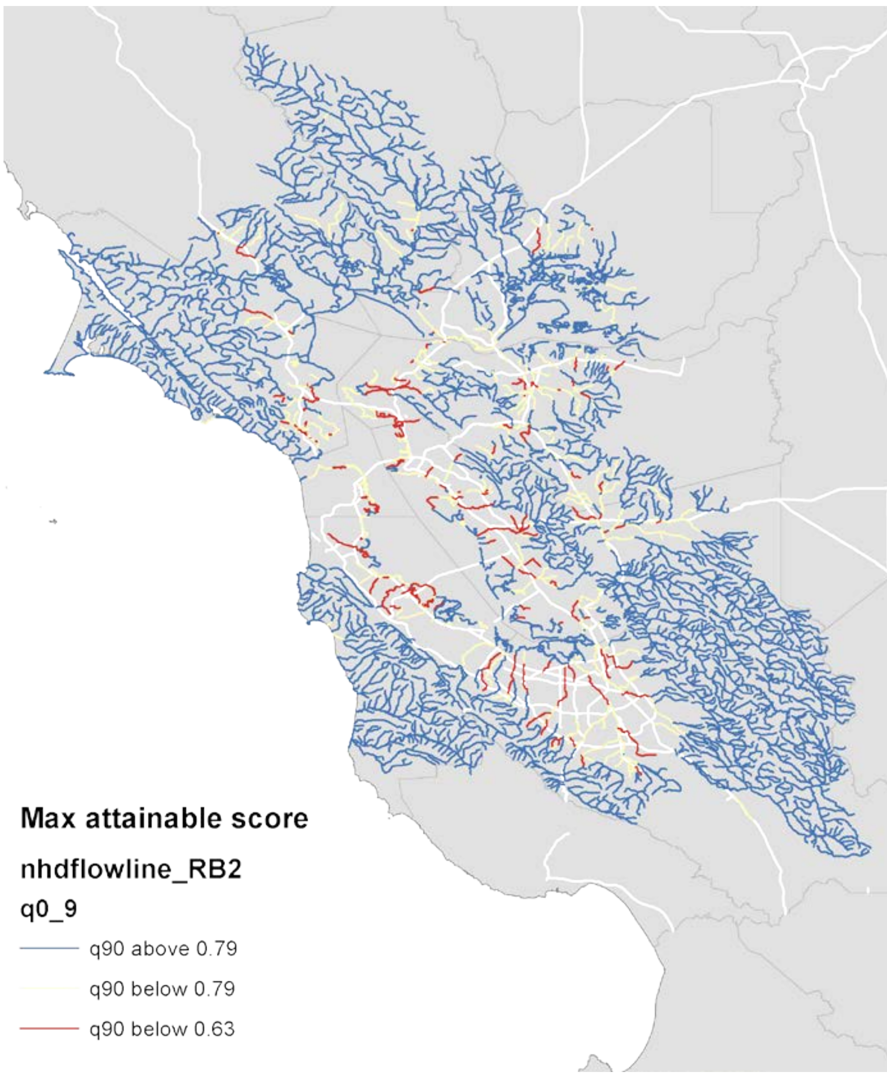


Figure 2. Map showing the classification of stream segments in the San Francisco Bay Area based on landscape development. In this map, attainability is defined as the 90th percentile of model predictions for each segment. Blue segments: Unconstrained by development, as described in Figure 1. Yellow segments: Moderately constrained by development. Red segments: Highly constrained by development.

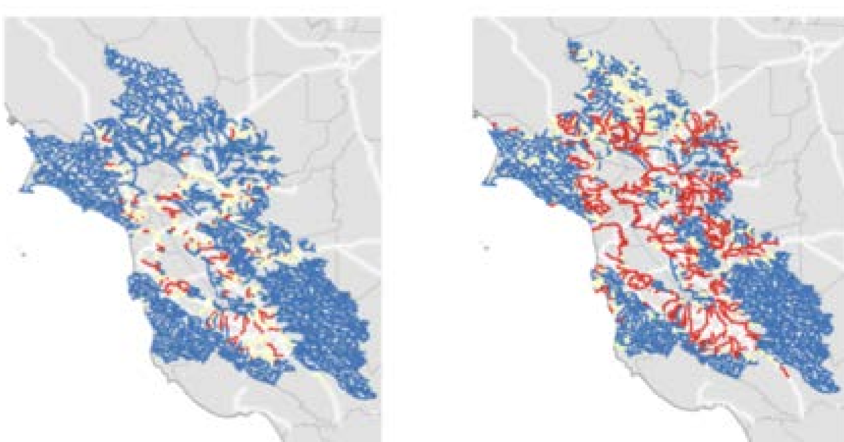


Figure 3: Example of maps generated with two different probabilities to define

likelihood. The map on the left was generated with a 10% probability to define likely scores, whereas the map on the right was generated with a 50% probability.

Task 3. Produce a Technical Memo with Key Graphics and Model Output

Based on feedback from group discussions and Water Board direction from Task 2, a reduced set of interactive maps and graphics can be generated to support this discussion. The purpose of this task is produce a technical memo with this reduced set of key graphics and model output in a format that can be easily shared and used to support discussions among Water Board staff and its advisory groups on policy options for channels in developed versus undeveloped landscapes. The maps and graphics will include ASCI scores and a linkage to Biological Condition Gradient calibration (see BCG workplan), and versions of maps and graphics to demonstrate policy options under consideration, as requested by Water Board staff.

Deliverables: 3.1) technical memo summarizing methodology and results of task 1 and 2, 3.2) model output that can be viewed in an interactive mode (e.g. Google Earth kmz file), 3.3) presentation to RG and SAG of illustrating policy options under consideration, upon request of Water Board staff.

Schedule of Interim Milestones and Deliverables

Task	Description	Estimated Date
1.1	Draft models and related graphics to predict bioassessment scores, in iterative stages of feedback	May 2017 (CSCI) September 2017 (ASCI)
1.2	Descriptive summaries of models, including evaluations of model performance, and list of landscape development variables in STREAMCAT selected for use in the models.	May 2017 and iteratively thereafter
2.1	Interactive maps and oral presentations with maps, graphics, summary tables as a function of choice of likelihood and assessment endpoint value.	May 2017 and iteratively thereafter
3.1	Draft and final technical memo summarizing methodology and results	September 2017 December 2017
3.2	Interactive model output (e.g. google earth .kmz file)	September 2017 December 2017
3.3	Presentation to RG and SAG of illustrating policy options under consideration	Upon request by Water Board staff

Citations

Hill, R.A., M.H. Weber, S.G. Leibowitz, A.R. Olsen, and D.J. Thornbrugh. 2015. The Stream-Catchment (StreamCat) dataset: A database of watershed metrics for the conterminous United States. Journal of the American Water Resources Association 52: 120-128.

Mazor, R.D., A.C. Rehn, P.R. Ode, M. Engeln, K.C. Schiff, E.D. Stein, D.J. Gillett, D.B. Herbst, and C.P. Hawkins. 2016. Bioassessment in complex environments: Designing an index for consistent meaning in different settings. Freshwater Science 35(1): 249-271

Waite, I.R. J.G. Kennen, J.T. May, L.R. Brown, T.F. Cuffney, K.A. Jones, and J.L. Orlando. 2012. Comparison of stream invertebrate response models for bioassessment metrics. Journal of the American Water Resources Association 48: 570-583.

Appendix 16

B495

JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 48, No. 3 AMERICAN WATER RESOURCES ASSOCIATION June 2012

COMPARISON OF STREAM INVERTEBRATE RESPONSE MODELS FOR BIOASSESSMENT METRICS[1]

Ian R. Waite, Jonathan G. Kennen, Jason T. May, Larry R. Brown, Thomas F. Cuffney, Kimberly A. Jones, and James L. Orlando[2]

ABSTRACT: We aggregated invertebrate data from various sources to assemble data for modeling in two ecoregions in Oregon and one in California. Our goal was to compare the performance of models developed using multiple linear regression (MLR) techniques with models developed using three relatively new techniques: classification and regression trees (CART), random forest (RF), and boosted regression trees (BRT). We used tolerance of taxa based on richness (RICHTOL) and ratio of observed to expected taxa (O/E) as response variables and land use/land cover as explanatory variables. Responses were generally linear; therefore, there was little improvement to the MLR models when compared to models using CART and RF. In general, the four modeling techniques (MLR, CART, RF, and BRT) consistently selected the same primary explanatory variables for each region. However, results from the BRT models showed significant improvement over the MLR models for each region; increases in R^2 from 0.09 to 0.20. The O/E metric that was derived from models specifically calibrated for Oregon consistently had lower R^2 values than RICHTOL for the two regions tested. Modeled O/E R^2 values were between 0.06 and 0.10 lower for each of the four modeling methods applied in the Willamette Valley and were between 0.19 and 0.36 points lower for the Blue Mountains. As a result, BRT models may indeed represent a good alternative to MLR for modeling species distribution relative to environmental variables.

(KEY TERMS: modeling; macroinvertebrates; watershed disturbance; land use; prediction; statistical assessment.)

Waite, Ian R., Jonathan G. Kennen, Jason T. May, Larry R. Brown, Thomas F. Cuffney, Kimberly A. Jones, and James L. Orlando, 2012. Comparison of Stream Invertebrate Response Models for Bioassessment Metrics. *Journal of the American Water Resources Association* (JAWRA) 48(3): 570-583. DOI: 10.1111/j.1752-1688.2011.00632.x

INTRODUCTION

Modeling has increased markedly in the past decade in all areas of ecology, and major advances have been made in conceptual models and statistical techniques (Leathwick *et al.*, 2005; Austin, 2007; Cabecinha *et al.*, 2007; Turak *et al.*, 2011), which, in turn, help practitioners derive response models that better support the needs of bioassessment programs. A

[1]Paper No. JAWRA-11-0093-P of the *Journal of the American Water Resources Association* (JAWRA). Received July 26, 2011; accepted December 2, 2011. **Discussions are open until six months from print publication**.

[2]Respectively, Biologist (Waite), U.S. Geological Survey, Oregon Water Science Center, 2130 SW 5th Avenue, Portland, Oregon 97201; Biologist (Kennen), U.S. Geological Survey, New Jersey Water Science Center, Trenton, New Jersey 08628; Biologist (May and Brown) and Hydrologist (Orlando), U.S. Geological Survey, California Water Science Center, Sacramento, California 95819; Biologist (Cuffney), U.S. Geological Survey, North Carolina Water Science Center, Raleigh, North Carolina 27607; and Physical Scientist (Jones), U.S. Geological Survey, Utah Water Science Center, West Valley, Utah 84119 (E-Mail/Waite: iwaite@usgs.gov).

fundamental goal of bioassessment in stream ecology is a better understanding of the effects of human land use on stream biota and the processes at various scales that cause these effects. However, streams are complex spatial and temporal habitat mosaics that are directly and indirectly influenced by a combination of natural geology, climate, and human disturbance (Stanford *et al.*, 2005). Stream ecologists are trying to understand the spatial scales and processes associated with human and natural disturbances that are affecting the biota. Models provide a useful framework for testing hypotheses, determining potential direct and indirect linkages, and directing where further research is needed. The expansion and application of multivariate models in stream ecology are helping to address these issues and hopefully will lead to a broader understanding of ecological and anthropogenic pathways and responses (Oberdorff *et al.*, 2001; Cabecinha *et al.*, 2007; Turak *et al.*, 2011; Waite *et al.*, 2010).

Much of the research documenting the effects of land-use change on stream biota indicates that as the total watershed area in agricultural and/or urban land use increases, individual biological metrics and multimetric indices (MMIs) (such as an Index of Biotic Integrity, IBI) that reflect compositional changes in sensitive species generally decrease (Paul and Meyer, 2001; Allan, 2004; Van Sickle *et al.*, 2004; Cuffney *et al.*, 2005; Ode *et al.*, 2008; Waite *et al.*, 2010). Though some researchers have found a threshold response (i.e., a nonlinear or step function) of individual or multimetric biological indices to land-use indicators (e.g., Davis and Simon, 1995; Wang *et al.*, 2001; Walsh *et al.*, 2005; Hilderbrand *et al.*, 2010; King and Baker, 2010) much of the literature indicates that the response more often is a simple monotonic response with no initial resistance (Booth, 2005; Cuffney *et al.*, 2005, 2010; Kennen *et al.*, 2005; Morgan and Cushman, 2005; Roy *et al.*, 2005; Stanford *et al.*, 2005; Waite *et al.*, 2008, 2010). The debate about possible threshold responses continues not only because of the interest in determining, from a management perspective, where a threshold might occur along a land-use gradient, but also because of the effect thresholds and the resultant nonlinear responses have on the application of various modeling techniques. If biological responses to landscape measures are indeed complex and nonlinear, then newer modeling techniques such as classification regression trees (CART), random forest (RF) and boosted regression trees (BRT), multilevel hierarchical modeling, structural equation models, or artificial neural networks may be necessary to model these responses (Grace, 2006). However, if various biological responses to human disturbance are commonly simple and linear, then they should be more easily modeled via standard regression techniques, which are typically easier to develop and interpret.

There are three commonly used bioassessment variable types including individual biological metrics (e.g., Ephemeroptera, Plecoptera, and Trichoptera richness or EPT), combining individual metrics into a multimetric index (e.g., IBI) and development of the observed/expected ratio metric (O/E). Each method has its advantages and disadvantages, yet sometimes they can give different results in differ environmental settings (Herbst and Silldorff, 2006; Chessman *et al.*, 2010; Hawkins *et al.*, 2010). It is possible that individual metrics may be more stressor gradient specific and multimetric indices better at more general disturbance gradients, however, detailed comparison of these three methods is beyond the scope of this paper. We focus on two common individual biological metrics, the general tolerance of invertebrates to a multitude of stressors including sediment, temperature, dissolved oxygen, hydrological and habitat changes, nutrients, and contaminants following Barbour *et al.* (1999) and the ratio of the observed/expected taxa based on the RIVSPAC method (River Invertebrate Prediction and Classification System) (Clarke, 2000; Moss, 2000). The number of tolerant taxa is expected to increase while the O/E value is expected to decrease as the amount of disturbance to the stream increases.

Using the same dataset used in this paper, Waite *et al.* (2010) developed macroinvertebrate response models for three regions in the western United States (U.S.) and the best multiple linear regression (MLR) models based on Akaike Information Criterion (AIC) and R^2 from each individual region required only two or three explanatory variables to model macroinvertebrate metrics to explain 41-74% of the variation. In each region, their best model contained some measure of urban and/or agricultural land use, yet often the model was improved by including a natural explanatory variable such as mean annual precipitation or mean watershed slope (for the MLR equations, see Waite *et al.*, 2010). Two macroinvertebrate metrics, the richness of tolerant macroinvertebrates (RICHTOL) and some form of EPT richness, were common response variables in models developed among the three regions (Waite *et al.*, 2010). Models were developed for the same two invertebrate metrics even though the geographic regions they modeled reflect distinct differences in precipitation, geology, elevation, slope, population density, and land use. L. R. Brown, J. T. May, A. C. Rehn, P. R. Ode, I. R. Waite, and J. K. Kennen (personal communication) were also able to develop strong models using linear modeling techniques (MLR), they modeled an invertebrate index of biotic integrity (BIBI) across a gradient of urbanized streams in southern California and were able to explain approximately 48% of the variation based on MLR models including classification accuracy of 69 and 87% for impaired and unimpaired sites, respectively.

One important question that researchers are working to answer is whether the use of newer, more complex modeling techniques such as CART and regression trees improves our ability to predict biological metrics and potentially provide new insights into response patterns and mechanistic pathways. Generalized linear models (GLMs) and generalized additive models (GAMs) were introduced in the 1980s and 1990s as improved methods over MLR for data with non-normally distributed errors (e.g., presence–absence and count data) or nonlinear relations and usually outperform single regression trees (Elith *et al.*, 2008). Regression trees are one type of technique within the commonly used CART or decision tree family (e.g., Breiman *et al.*, 1984; De'ath and Fabricius, 2000; Prasad *et al.*, 2006). Trees attempt to explain variation in one categorical (classification) or continuous (regression) response variable by one or more explanatory variables, the resultant output being a dendogram or tree with varying numbers of branches or nodes. These techniques have a few properties that are highly desirable for ecological data analysis: (1) they can handle numeric, categorical, and censored response variables, (2) they are not affected by explanatory variables that follow nonnormal distributions (i.e., skewed, Poisson, or bimodal), and (3) they can model complex interactions simply (De'ath, 2007). Maloney *et al.* (2009) found that CART models of watershed disturbance on BIBI values provided results that were intuitive and easy to interpret but they did not classify sites any better than logistic regression models; however, RF models showed minor improvements in performance over the other models. De'ath (2007) and Elith *et al.* (2008) show that BRTs outperform GLMs and GAMs in variable selection, predictive ability (higher R^2 and lower error), and can handle sharp discontinuities in data that are difficult for the other methods. Aertsena *et al.* (2010) also showed that BRT outperformed most modeling techniques (i.e., MLR, GLM, GAM, and CART), with the exception of artificial neural networks.

Over the past decade the estimate of O/E has become a common measure of biological condition for use in bioassessments (e.g., Hawkins, 2006; Carlisle *et al.*, 2008). The expected taxa for a site are commonly estimated by models (e.g., RIVPACS) (Clarke, 2000; Moss, 2000) of reference sites; this value is then compared to the actual taxa collected at a site. Models based on this approach have been developed in many international regions (e.g., Europe, New Zealand, and Australia) (Davies, 2000; Clarke and Murphy, 2006) and for separate regions within the U.S., including many states (Hubler, 2008). Recently, Hawkins *et al.* (2010) compared the response of three types of O/E models with five versions of MMIs for macroinvertebrates and found that in general, the O/E models were better able to distinguish managed or disturbed sites from reference sites than the MMIs. Due to these results and to its overall national and international popularity, we wanted to evaluate how models developed using O/E as the response variable would compare to models developed using single metrics, such as RICHTOL.

Our goal in this paper is to compare the overall performance (i.e., model fit, or R^2) of models developed using standard MLR techniques with more complex models developed using newer alternative techniques such as CART, RF, and boosted regression for the common macroinvertebrate metrics RICHTOL and O/E as the response variables. Also, we believe that the development of watershed disturbance predictive models such as those presented herein will build upon previous research to help the potential derivation of more complex models to better understand disturbance pathways in the landscape and ultimately the biocomplexity of aquatic systems.

METHODS

Data Aggregation and Landscape Analysis

For this comparative analysis we used the datasets (U.S. Geological Survey, U.S. Environmental Protection Agency, Oregon Department of Environmental Quality, and California Department of Fish and Game) previously aggregated for three regions in the western U.S. by Waite *et al.* (2010). A brief summary of the methods follows. Sites were evaluated based on the following criteria: invertebrate data sampled with comparable methods; upstream watershed area of between 13 and 259 km^2; and watersheds could not be nested (i.e., no spatial autocorrelation). Sites meeting these conservative criteria resulted in three study regions: Coastal Southern California ($n = 55$), the Blue Mountains ecoregion of eastern Oregon ($n = 148$), and the Willamette Valley ecoregion in north-central Oregon ($n = 96$) (Figure 1).

For consistency, watersheds were re-delineated for the selected sampling sites within the three study regions using USGS 7.5 min quadrangle digital raster graphics as base layers. The digital raster graphics were displayed on-screen along with National Hydrography Dataset (NHD) high resolution stream lines for each region (U.S. Geological Survey, 2007).

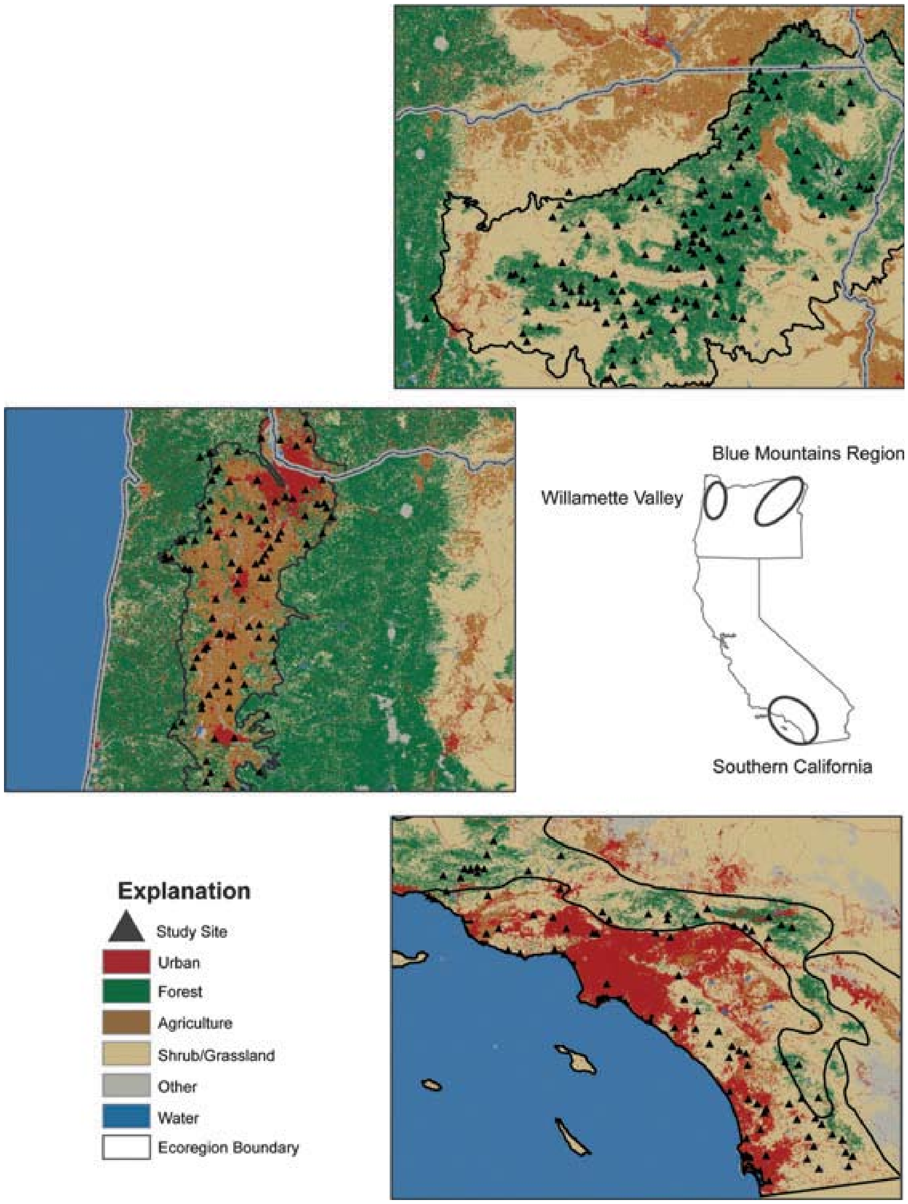


FIGURE 1. Map Showing Land Use and Land Cover for the Three Modeling Regions: Blue Mountains and Willamette Valley, Oregon, and Southern California.

Watershed boundaries were digitized on-screen at a scale of 1:10,000 or larger. Adjacent watershed polygons were edge matched to eliminate all overlaps and gaps. All work was conducted using ArcGIS, ArcMap 9.2 (Environmental Systems Research Institute, Redlands, CA; Table A1) GIS software.

Riparian buffer zone polygons were created within each watershed, extending 2 km upstream from the outlet of each watershed along the main stem and all tributaries and 90 m on either side of the stream centerlines. The buffers were created by selecting the appropriate NHD stream lines within each watershed and creating routes along each main stem and tributary flow path. The routes were then clipped to a distance of 2 km from the basin outlet and buffered. All abbreviations for riparian based explanatory variables begin with the letters "Rip"; otherwise, variables are watershed based (Table 1).

Spatial datasets representing landscape metrics of watershed disturbance were created for each watershed and riparian zone buffer from available national and regional datasets (Table A1) and included elevation, slope, land cover (1992 and 2001), population density, road networks, soil infiltration capacity, hydrography, pollution point sources, dams, and precipitation. Land-use summaries were based on either 1992 or 2001 spatial data (as described in Vogelmann *et al.*, 2001; Homer *et al.*,

TABLE 1. Description, Variable Code and Definition of Explanatory (landscape) and Predictor (invertebrate metrics) Variables Used for Response Model Development.

Explanatory Variables: Landscape

Description	**Variable Code**	**Definition**
Watershed Scale Variables		
Percent urban land use	Urban	Percent watershed area in urban land use (NLCD 2000 categories 21, 22, 23, and 24)
Percent agricultural land use	Ag	Percent watershed area in agricultural land use (NLCD 2000 category 82)
Sum of percent Ag + Urban	Ag + Urb	Sum of percent watershed area in urban (NLCD 2000 categories 21, 22, 23, and 24) and agricultural (NLCD 82) land use
Percent forest	Forest	Percent watershed area in forest land use (NLCD 2000 categories 41, 42, 43)
Percent pasture	Pasture	Percent watershed area in pasture land use (NLCD 2000 category 81)
Percent shrub/scrub	Shrub	Percent watershed area in shrubland, shrub/scrub (NLCD 2000 category 52)
Road density	RdDens	Road density in watershed = Road length (km)/watershed area (km^2)
Mean population density	PopDen	Watershed mean population density based on 2000 census (persons/km^2)
Minimum elevation	Min-Elev	Elevation (m) at stream site, pour point of watershed
Mean slope percent	Slope	Mean percent watershed slope
Manmade stream density	MmStreams	Manmade stream density in watershed = manmade stream length (km)/ watershed area (km^2)
Mean annual precipitation	MnAnnPrecip	Mean annual precipitation (cm)
Soil infiltration rate	Soil_Mod-Infil	Hydrologic soil group B, moderate infiltration rate (min. infiltration rate 4-8 mm/h)
Riparian Scale Variables		
Percent urban land use	Rip_Urban	Percent buffer area in urban land use (NLCD 2000 categories 21, 22, 23, and 24)
Percent agricultural land use	Rip_Ag	Percent buffer area in agricultural land use (NLCD 2000 category 82)
Sum of percent Ag + Urban	Rip_Ag + Urb	Sum of percent buffer area in urban (NLCD 2000 categories 21, 22, 23, and 24) and agricultural (NLCD 82) land use
Percent forest	Rip_Forest	Percent buffer area in forest land use (NLCD 2000 categories 41, 42, 43)
Percent pasture	Rip_Pasture	Percent buffer area in pasture land use (NLCD 2000 category 81)
Percent shrub/scrub	Rip_Shrub	Percent buffer area in shrubland, shrub/scrub (NLCD 2000 category 52)
Road density	Rip_RdDens	Road density in buffer = Road length (km)/watershed area (km^2)
Mean population density	Rip_PopDens	Buffer area mean population density based on 2000 census (persons/km^2)
Mean slope percent	Rip_Slope	Mean percent buffer slope
Maximum elevation	Rip_Max-Elev	Maximum buffer elevation (m)
Response Variables: Invertebrate Metrics		
Observed/expected	O/E	Ratio of number of observed taxa at a site over the expected taxa based on modeled reference sites
Tolerant richness	RICHTOL	Average USEPA tolerance values for sample based on richness

2004), depending on which data source was closer to the macroinvertebrate sample date for that watershed. Watersheds and riparian zone buffers were used to define zones for analysis and calculate summary statistics. The 1992 and 2001 land cover datasets used slightly different classification schemes. Uniform codes based on the 2001 classification scheme were assigned to all land cover classes in the final summary statistics table (Fry *et al.*, 2009). We did not assess the distribution pattern of land use/land cover within the watershed though this can be important in some situations.

Description of Modeling Regions

The Coastal Southern California (SoCal; Southern and Central California Chaparral and Oak Woodlands Ecoregion) region has a Mediterranean climate of hot, dry summers and cool, moist winters (Ode *et al.*, 2005). Average precipitation at each site ranges from 25 to 50 cm/year. The geology of the ecoregion is dominated by recently uplifted and poorly consolidated marine sediments. Vegetative cover in this region consists mainly of chaparral and oak woodlands, though grasslands occur in some lower elevations and patches of pine are found at higher elevations (open low mountains or foothills). The landscape is currently dominated by urban development; the human population is approximately 19 million and is projected to exceed 28 million by 2025 (Ode *et al.*, 2005). Outside the urban centers, much of this region was historically grazed by domestic livestock or cultivated for fruits and vegetables, but most of this land has since been converted to urban uses.

The Blue Mountains (Blue_Mt) are the westernmost range of the Middle Rocky Mountains and, like the Cascade Range, are largely volcanic, with fertile plateaus and deeply fissured river valleys. Carved by two rivers (the John Day and Grande Ronde Rivers) the landscape has steep hillsides, bluffs and rimrock faces. Temperature and precipitation are highly correlated with elevation. Precipitation ranges from 22 to 45 cm/year along the river valleys and is >150 cm/year in the nearby mountains. This region is dominated by coniferous forests in mid to higher elevations and shrub and grassland in lower elevations, though much of the latter has been displaced by agriculture and grazing. The region has no large cities and urbanization is limited to scattered smaller cities and small towns.

The Willamette Valley (Will_V) ecoregion contains a mixture of rolling prairies, mixed forests, and extensive lowland valley wetlands. With temperate, dry summers and cool, wet winters, the Willamette River basin and surrounding area is characteristic of the Pacific Northwest climate. About 90% of the annual precipitation (100-130 cm/year) occurs during October through May (Uhrich and Wentz, 1999), falling as rain in the valley and snow in the mountains. The land use/land cover in the valley plains and foothills is primarily cultivated crops, pasture, and grasslands. Urbanization ranges from minimal to extensive (Waite *et al*., 2008). Centered on the confluence of the Columbia and Willamette Rivers, Portland is the most populous city in Oregon, with 539,000 people in city limits and nearly 3 million people in the Portland metropolitan area (U.S. Census Bureau, 2000). The population in the metropolitan area increased almost 30% from 1990 to 2000, with some suburban populations increasing more than 80% during the same period (U.S. Census Bureau, 2000). The drainage network in the Willamette Valley combines natural tributaries, complex networks of canals in agricultural areas, and stormwater canals and groundwater infiltration wells in cities.

The three geographic regions modeled in this study have differing natural settings and the extent and type of human disturbance in each respective region. SoCal has the driest climate, intermediate mean stream site elevation (Min-Elev) and percent agriculture, and the highest population density. Blue_Mt has the highest mean site elevation and mean watershed slope, intermediate mean precipitation, and the lowest population density, percent urban, and percent agriculture. Will_V has the greatest precipitation, lowest minimum site elevation, and the highest percent agriculture.

Macroinvertebrate Data

Macroinvertebrate data from 1994 to 2005 assembled for this study were considered to be comparable in terms of sampling protocols (sampled habitat, number of composite samples, and total sampled area) and laboratory procedures, including sorting, subsample count level, and taxonomic resolution (personal communication state agency personnel, 2005; Waite *et al.*, 2010). In general, all macroinvertebrate samples were collected in similar habitats using kick-net techniques from five to eight separate areas and combined for a composite sample (Moulton *et al.*, 2002; Peck *et al.*, 2006; Hubler, 2008). Extensive review of the data was completed to make sure aggregated data from disparate sources included the same taxonomic groups, followed the same nomenclature, and had appropriate taxonomic resolution before data analysis was attempted. The Invertebrate Data Analysis System software (Cuffney, 2003) was used to resolve by region all taxonomic issues (taxonomic identification level and nomenclature), to remove ambiguous taxa (Cuffney *et al.*, 2007), and to randomly subsample raw counts to an equal 300 (Will_V) or 500 specimen count (the highest possible based on the data in each region) across all study regions. In general, data for dominant aquatic insect orders were resolved at genus level. Less common orders were often aggregated to family level. Rare organisms or those with difficult taxonomy were sometimes aggregated to order or higher. The dipteran family Chironomidae is considered an important bioindicator group, yet historically a difficult group to identify to genus or species. As a result, data for this group were assigned to six taxa levels (five subfamilies plus Chironomidae) from the various family to genus level identifications within the original data. Tolerance and functional group metrics were calculated using values from Barbour *et al.* (1999), supplemented with values from Wisseman's tolerances for the Pacific Northwest (Wisseman, 1996, unpublished data). Macroinvertebrate O/E values were estimated using two existing regional models (East and West of the Cascade Mountains) that were developed by Oregon Department of Environmental Quality (Hubler, 2008). O/E models were not ready for the SoCal region at the time of analysis so we were not able to test O/E values for this area.

MODEL DEVELOPMENT

Details of MLR model development procedures are outlined in Waite *et al.* (2010). In brief, model performance was assessed using a variety of statistics, including adjusted mean sum of squares (R^2), root mean squared error, AIC, predicted sum of squares, and regression coefficients in Waite *et al.* (2010). We adopted a model fitting approach for each response variable. We used a step-wise selection based on AIC for all models ranging from 1 to 5 environmental variables, as appropriate by region. If necessary, variables were transformed to improve their distributions to better adhere to assumptions of linearity. Models were developed for each geographic region separately due to the large spatial separation between each region and as described above, because the climatic and disturbance regimes were distinct. Model residuals, potential outliers, and interaction terms were evaluated. A description of variables used in model development is provided in Table 1. A MLR model was developed for the response variable RICHTOL for all three regions; it included two predictor variables (population density and riparian road density) for SoCal, three predictor variables for Blue_Mt (percent shrubs, percent agriculture, and mean annual precipitation in the watershed) and three predictor variables for the Will_V region (percent agriculture plus urban land use in the watershed, mean annual precipitation, and percent agriculture plus urban land use in the riparian zone) (Waite *et al.*, 2010). As a comparison to the MLR models developed by Waite *et al.* (2010) for RICHTOL, new models were developed for O/E for the Blue_Mt and Will_V regions.

To gain additional insight into these data and as a comparison against the MLR models, single regression trees, RF, and BRT models were developed for each region individually. Regression trees are one type of technique within the commonly used CART or decision tree family, and their use and technical details have been described extensively in the literature (e.g., Breiman *et al.*, 1984; De'ath and Fabricius, 2000; Prasad *et al.*, 2006); therefore, we will only provide a brief overview. Trees attempt to explain variation in one categorical (classification) or continuous (regression) response variable by one or more explanatory variables, the resultant output being a dendogram, or tree, with varying numbers of branches or nodes. Trees are developed following a hierarchical binary splitting procedure that attempts to find the best single explanatory variable that minimizes the within group and maximizes the among group dissimilarity in the response variable at each split. It does this for each explanatory variable entered into model development and can thus provide a list of the explanatory or predictive power of the variables. We used R statistics scripts and software (R Development Core Team, 2007, version 2.10.0) following the procedures outlined by Therneau and Atkinson (1997) to determine the proper single regression tree and the appropriate pruning of branches (De'ath and Fabricius, 2000; Prasad *et al.*, 2006). Trees have a few properties that are highly desirable for ecological data analysis: (1) they can handle numeric and categorical variables (2) they are not affected by explanatory variables that follow non-normal distributions (i.e., skewed, Poisson, or bi-modal), and (3) they can model complex interactions simply (De'ath, 2007).

Random forests and BRT are among a family of techniques used to advance single classification or regression trees by averaging the results for each binary split from numerous trees or forests thus reducing the predictive error and improving overall performance (De'ath, 2007; Elith *et al.*, 2008). In BRT, after the initial tree has been generated, successive trees are grown on reweighted versions of the data giving more weight to those cases that are incorrectly classified than those that are correctly classified within each growth sequence. Thus, as more and more trees are grown in BRT, the large number of trees increases the chance that cases that are difficult to classify initially are correctly classified, thus representing an improvement to the basic averaging algorithm used in RF (De'ath, 2007). Boosted trees and RF models retain the positive aspects of single trees seen in CART models, yet have improved predictive performance, nonlinearities and interactions are catered to or easily assessed, and they can provide an ordered list of the importance of the explanatory variables (Cutler *et al.*, 2007; De'ath, 2007). Though RF and BRT offers improved modeling performance over CART, the simple single tree obtained from CART is lost, making it more difficult to visualize the results. Partial dependency plots (PDP) are a way to visualize the effect of a specific explanatory variable on the response variable after accounting for the average effects of all other explanatory variables (De'ath, 2007; Elith *et al.*, 2008); these are presented in this paper for select models as examples (e.g., Figures 2 and 3). Random forest models were developed using the rpart library in R following methods outlined in Cutler *et al*. (2007) and BRT models were run using the gbm library in R and specific code from Elith *et al.* (2008). We used R^2 values for assessing the amount of variation explained among the four modeling techniques since it is a common and well understood measure that allowed us to put each model on the same measurement currency; other model performance measures such as confidence intervals and *p*-values are not included for simplicity.

RESULTS

In general, the four modeling techniques selected the same primary explanatory variables within each

TABLE 2. Explanatory Variables in Order of Importance in the Models for Four Modeling Methods for Two Macroinvertebrate Metrics for Each of Three Study Regions (SoCal, Southern California; Will_V, Willamette Valley; Blue_Mt, Blue Mountains, Oregon).

	MLR	CART	RF	BRT
SoCal				
RICHTOL	PopDen	PopDen	PopDen	PopDen
	Rip_RdDens	MmStreams	Min-Elev	Rip_Slope
		Min-Elev	Rip_Slope	Min-Elev
Will_V				
RICHTOL	Ag + Urb	Ag + Urb	Ag + Urb	Ag + Urb
	MnAnnPrecip	MnAnnPrecip	MnAnnPrecip	MnAnnPrecip
	Rip_Ag + Urb	Rip_Forest	Rip_Forest	Rip_Max-Elev
			Rip_Max-Elev	Rip_Forest
O/E	Ag + Urb	Forest	Forest	Ag + Urb
	MnAnnPrecip	Rip_Max-Elev	Rip_Max-Elev	Rip_Max-Elev
	Rip_Ag + Urb		Soil_Mod-Infil	MnAnnPrecip
			Rip_Forest	
Blue_Mt				
RICHTOL	Shrub	Shrub	Shrub	Shrub
	Ag	Slope	Slope	MnAnnPrecip
	MnAnnPrecip	MnAnnPrecip	MnAnnPrecip	Slope
O/E	MnAnnPrecip	Slope	Shrub	Slope
	Shrub	MnAnnPrecip	Slope	MnAnnPrecip
	Slope		MnAnnPrecip	Shrub

Notes: MLR, multiple linear regression; CART, classification and regression trees; RF, random forest; BRT, boosted regression trees; RICHTOL, average tolerance value for sample based on richness at a site; O/E, ratio of observed/expected taxa.

region with minor variation among model types (Table 2): (1) SoCal: population density, minimum elevation, and riparian slope, (2) Blue_Mt: percent shrub, mean annual precipitation (MnAnnPrecip), and watershed slope, and (3) Will_V: percent agriculture plus urban, MnAnnPrecip, riparian maximum elevation, and percent riparian forest (see Table 1 for definitions). Generally, the RICHTOL R^2 values for MLR were slightly higher than those for the CART and RF models for all three regions (Table 3); however, this was not the case for the O/E models for Blue_Mt. Nevertheless, these differences are probably not meaningful because the R^2 values for CART and RF models are determined by a cross-validation method that ensures no over-fitting and thus usually gives a lower, more conservative value than the MLR values. Interaction affects were tested for and found to not be significant in the models developed. Conversely, the BRT models showed considerable improvement in the R^2 values over all the other models for both response variables (i.e., RICHTOL and O/E). For example, the SoCal RICHTOL R^2 values for the MLR compared to the BRT model increased from 0.67 to 0.79, Blue_Mt showed an increase from 0.44 to 0.59 for RICHTOL and from 0.08 to 0.28 for O/E, and the Will_V R^2 values increased from 0.74 to 0.83 for RICHTOL and from 0.64 to 0.75 for O/E (Table 3).

TABLE 3. Comparison of R^2 Values for Four Modeling Methods for Two Macroinvertebrate Metrics for Each of Three Study Regions (SoCal, Southern California; Will_V, Willamette Valley; Blue_Mt, Blue Mountains, Oregon).

	MLR	CART	RF	BRT
SoCal				
RICHTOL	0.67 (2)	0.64 (3)	0.65 (3)	**0.79** (3)
Will_V				
RICHTOL	0.74 (3)	0.68 (3)	0.73 (4)	**0.83** (4)
O/E	0.64 (3)	0.62 (2)	0.61 (4)	**0.75** (3)
Blue_Mt				
RICHTOL	0.44 (3)	0.34 (3)	0.41 (3)	**0.59** (3)
O/E	0.08 (3)	0.15 (2)	0.07 (3)	**0.28** (3)

Notes: Number of variables in model in parentheses. MLR, multiple linear regression; CART, classification and regression trees; RF, random forest; BRT, boosted regression trees; RICHTOL, average tolerance value for sample based on richness at a site; O/E, ratio of observed/expected taxa. Highest R^2 value across all models is shown in bold.

The O/E metric derived from RIVPACS type models specifically calibrated for Oregon consistently had lower R^2 values than RICHTOL for the two regions tested (Table 3). Modeled O/E R^2 values were between 0.06 and 0.10 lower than RICHTOL values for each of the four modeling methods applied in the Will_V region and were between 0.19 and 0.36 points lower for the Blue_Mt region.

As mentioned above, all modeling procedures (i.e., MLR, CART, RF, and BRT) generally retained the same subset of explanatory variables. These variables, with some minor exceptions in the Blue_Mt study region, generally accounted for approximately a similar proportion of the variance in the

RICHTOL and O/E response models. R^2 values, however, do not provide a complete picture of the model response pattern, and the overall influence of a specific explanatory variable on the environmental system or process being modeled is typically lost when the model is fit to a linear or nonlinear form. Partial dependency plots, which are provided as a diagnostic tool in the BRT and RF model output, provide a way to more fully examine the relative influence of individual explanatory variables on the response variable given the modeled structure. As explained in De'ath (2007) and Elith *et al.* (2008), PDP provide a way to visualize the effect of a specific explanatory variable on the response variable after accounting for the average effects of all other explanatory variables. For example, PDPs for the four variables retained in the BRT model for Will_V are shown in Figure 2. In general, the plots show a near linear increase in RICHTOL as the amount of agriculture plus urban land use in the watershed increases (Figure 2A) and a decrease in RICHTOL as riparian maximum elevation increases (Figure 2D). However, the response in RICHTOL values flattens out at approximately 60% agriculture plus urban land use, then again increases rapidly from approximately 90 to 100%. Likewise, the PDP graph shows that there is rapid change in RICHTOL values from near 0 to 200 m in riparian maximum elevation followed by no response beyond 400 m. The pattern shown for mean annual precipitation (Figure 2B) follows the opposite pattern of the amount of agriculture plus urban land use in the watershed, RICHTOL values decrease rapidly from the lowest precipitation values until approximately 80 cm/year beyond which values show no response. As the amount of riparian forest cover declines (Figure 2C), RICHTOL values increase little until riparian forest values drop to about 30%, where there is a step-wise increase until the point when there is only about 5% riparian forest remaining, whereupon there is a rapid increase in tolerance values. The PDPs for O/E in the Will_V show remarkable similarity to that described above for RICHTOL except that, as one would expect due to the differences in the invertebrate metrics, the curves respond in opposite directions (Figure 3). There is a general linear decrease in O/E values as agriculture plus urban land use increases (Figure 3A), a sharp increase in O/E values as riparian maximum elevation increases to 200 m (Figure 3B) or when mean annual precipitation increases to about 70 cm/year (Figure 3C). As seen for RICHTOL, O/E showed an abrupt threshold-type response at low levels of riparian forest (Figure 3D) followed by a step

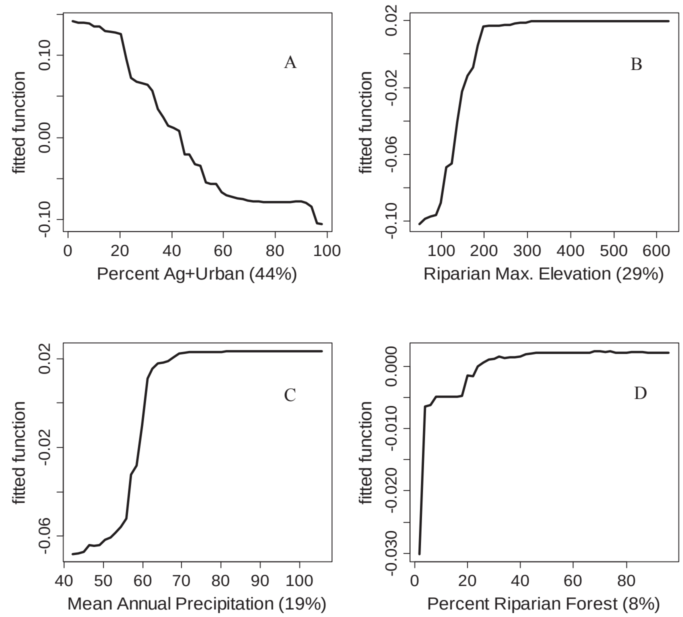


FIGURE 3. Partial Dependency Plots for Ag + Urb (A), Rip_Max-Elev (B), MnAnnPrecip (C), and Rip_Forest (D) in the Boosted Regression Model Developed for Observed/Expected (O/E) in Willamette Valley (Will_V). The *y*-axis represents the effect of the selected variable on the response variable O/E metric, the relative contribution of each explanatory variable is reported in parentheses. Refer to Table 1 for variable definitions.

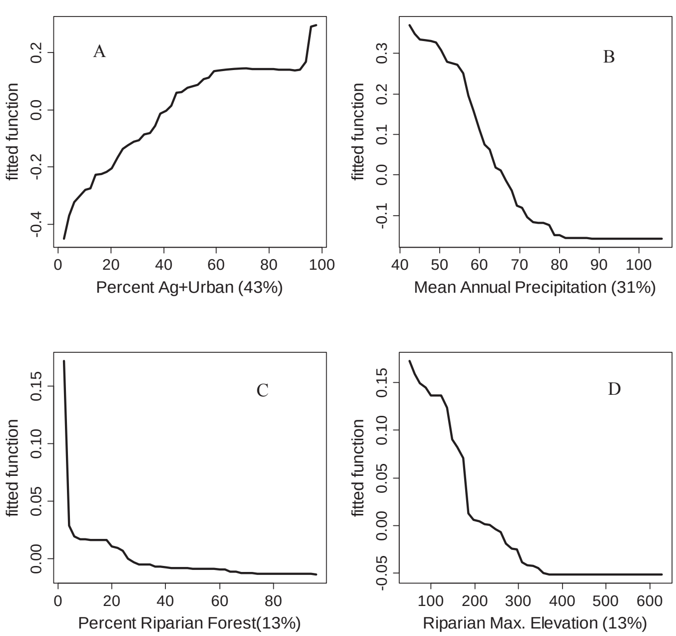


FIGURE 2. Partial Dependency Plots for Ag + Urb (A), MnAnnPrecip (B), Rip_Forest (C), and Rip_Max-Elev (D) in the Boosted Regression Model Developed for RICHTOL in Willamette Valley (Will_V). The *y*-axis fitted function represents the effect of the selected variable on the response variable RICHTOL; the relative contribution of each explanatory variable is reported in parentheses. Refer to Table 1 for variable definitions.

increase and a plateau above approximately 30% riparian forest cover.

DISCUSSION

It is encouraging that the MLR and the CART and RF (regression tree family) modeling techniques gave similar results selecting in general the same main explanatory variables (Table 2) and explaining similar amounts of variation (Table 3), which may indicate that the MLR methods used in this study are appropriate for these types of ecological data. The BRT models, however, did show notable improvement in model fit with increases in R^2 values ranging from 0.09 through 0.15 for RICHTOL to 0.11 through 0.20 for O/E compared to MLR models (Table 3). L. R. Brown, J. T. May, A. C. Rehn, P. R. Ode, I. R. Waite, and J. K. Kennen (personal communication), using a MMI for macroinvertebrates (i.e., BIBI) sampled across a strong urbanization gradient, also showed a notable improvement in model performance for BRT compared to MLR. De'ath and Fabricius (2000) suggest that for complex or messy data, even single regression trees will often outperform MLR and are preferred for determining variable selection and interaction effects due to the issue that MLR models with complex data are frequently difficult to interpret because they will often include too many variables with high order interactions. It was found that CART and RF models did not outperform the RICHTOL MLR models in this analysis which supports our overarching hypothesis that MLR will generally perform as well as many of the tree modeling techniques when data follows a general linear response or when, in the case of the three regions evaluated, there are few explanatory variables with no high order interactions. Maloney *et al.* (2009) found that CART models of land-use disturbance on macroinvertebrate IBI metrics provided results that were intuitive, but they did not classify sites any better than logistic regression models; however, unlike in this study, their RF models showed minor improvements in performance over CART and logistic regression models.

In general, regression trees allow the inclusion of more variables in the model building phase than MLR, allow for easier testing for interaction affects and produce a list of variables explaining the importance of variation in the response variable. In addition, the PDPs from BRT or RF can offer valuable insights into the pattern or form of the response variable based on select explanatory variables improving model interpretation. For example, the PDPs for Will_V (Figures 2 and 3) revealed that the response rate changed or flattened out and provided additional insight into potential thresholds along the range of the individual explanatory variables that are not easily depicted with MLR models.

The identification of thresholds (i.e., transition points in ecological condition) is of growing interest to the scientific and regulatory community, especially for forecasting the loss of biodiversity (Hilderbrand *et al.*, 2010) or for understanding system recovery (Clements *et al.*, 2010; Qian and Cuffney, 2012). More research is clearly needed to help better detect nonlinear and possible threshold responses (Dodds *et al.*, 2010) and new analytical tools are emerging (i.e., BRT results shown in this study) that can assist with identifying changes in taxa occurrence across an environmental gradient (Qian and Cuffney, 2012).

Even though we were able to successfully develop strong MLR models indicating that the primary responses were linear in nature (Waite *et al.*, 2010), the BRT PDPs reveal potential thresholds in the response variable in at least some of the regions (e.g., the Will_V PDPs shown for RICHTOL and O/E in Figures 2 and 3) that were not seen in the MLR models. It is possible that since MLR models assume linearity that they may sometimes miss nonlinear/thresholds in some explanatory variables. The response of RICHTOL and O/E for watershed agriculture plus urban (Ag + Urb) was primarily linear with a small step function at the end (Figures 2A and 3A). The two riparian variables, riparian maximum elevation (Rip_Max-Elev; Figures 2D and 3B) and riparian forest (Rip_Forest; Figures 2C and 3D) on the other hand showed potential thresholds. The response of the two invertebrate metrics to changes in Rip_Max-Elev showed no response from 600 to 400 m for RICHTOL and to 200 m for O/E, after which there was a steep increase or decrease to the lowest elevation (Figures 2D and 3B). It is likely that riparian elevation is acting as a surrogate for the natural climatic and geologic trend that occurs in the Willamette Valley, trending from the valley floor with low stream gradient and lower elevation and precipitation to higher values for these and other variables as one moves toward the foothills of the Coast or Cascade Ranges on either side of the valley. The response of RICHTOL and O/E to changes in Rip_Forest showed a slow but continuous linear increase or decrease as the amount of Rip_Forest decreased from 100% to approximately 5%, after which there appears to be a rapid change in either of the metric values, which may indicate a strong threshold at or near the 5% level. This suggests that as percent forest in the riparian zone along streams drops below approximately 5-10%

land cover, stream integrity degrades rapidly possible due to the reduction in natural buffering capacity seen in healthy riparian systems. L. R. Brown, J. T. May, A. C. Rehn, P. R. Ode, I. R. Waite, and J. K. Kennen (personal communication) found a similar response in the MMI they modeled (BIBI) against four explanatory variables across a strong urbanization gradient in some California streams. They showed that the amount of agriculture plus urban land use in the riparian zone and mean annual precipitation in the watershed showed approximate linear responses, though in opposite directions. They also found a threshold-type response in the BIBI to low values of population density (approximately 300 persons/km^2) in the watershed. Similar to the findings in this study, L. R. Brown, J. T. May, A. C. Rehn, P. R. Ode, I. R. Waite, and J. K. Kennen (personal communication) found that the BRT method appeared to be more sensitive for detecting nonlinear response patterns such as thresholds, for determining potential surrogate variables, and for model corroboration.

The overall poorer performance of the O/E metric compared to the single metric RICHTOL across all models was notable, yet the especially poor performance in the Blue_Mt region was particularly surprising (Table 3). When comparing the ability of O/E and a multimetric invertebrate IBI to differentiate between reference and degraded sites, Herbst and Silldorff (2006) found that the two methods were in close agreement for sites in eastern Sierra Nevada of California. Hawkins *et al.* (2010) compared the performance of a multimetric index and O/E for 225 sites from five ecoregions in the interior Columbia Basin, including many of the sites used in this study from the Blue_Mt ecoregions. They found that the O/E metric was better at distinguishing among the three disturbance classes, particularly between the intermediate and high disturbance classes than the multimetric index. The discrepancy between the poor performance of O/E in the Blue_Mt region in our study and the strong performance in their study may be due to a larger underlying disturbance gradient within their dataset, which resulted from the inclusion of data from multiple ecoregions. Models derived for the Will_V region, where there was a larger disturbance gradient than that found in the Blue_Mt region, showed relatively little difference in performance between the O/E and RICHTOL metrics. It is also possible that the lower R^2 for the O/E models may be because we are not able to model nor account for the error associated with estimation of the raw O/E metric values. Chessman *et al.* (2010) found that O/E values did not distinguish among site disturbance groups based on hydrologic alteration in Australia even though taxonomic richness and assemblage composition could. However, it is yet unclear why O/E performance would be inhibited in areas with a shorter disturbance gradient than that shown in Hawkins *et al.* (2010). One possibility is that because these O/E models are based on a subset of taxa that occur at 50% of the reference sites and therefore operate with a reduced taxa list, specifically with the relatively rare and arguably with the more sensitive portion of the taxa list removed, the resulting O/E values may be less able to distinguish the small more subtle differences among sites, such as that seen in the Blue_Mt study region. In contrast, the RICHTOL metric uses all the taxa that occur at a site and may be a more sensitive measure of changes in assemblage integrity in areas of low anthropogenic disturbance.

CONCLUSIONS

Waite *et al.* (2010) were able to successfully develop MLR models for the three distinct and separate regional datasets presented in this study for individual macroinvertebrate metrics (e.g., RICHTOL, EPT). This study developed alternate models, CART, RF, BRT, for the same datasets and compared them to the MLR models previously developed. The O/E metric performed nearly as well as RICHTOL in the Will_V region where there was a strong disturbance gradient but performed poorly in Blue_Mt, a region with a relatively weak gradient. Though the data modeled in this study were not particularly noisy or complex, the BRT models, in all cases, outperformed the MLR methods and provided specific information on the form of the response function for each variable giving important insight into potential thresholds in the data. As a result of this ecological modeling comparison, BRT models may indeed represent a good alternative to MLR for modeling species distribution relative to environmental variables. Modeling results indicate that even when the response pattern is simple and strongly linear, BRT models not only markedly improve model fit, but can also help to corroborate results from other methods, provide additional information on potential interactions among variables, and support greater insight into understanding the response profile of a given metric, whether it be a linear, step, or a threshold function, across environmental gradients that may not be easily seen with MLR. Models like these can be used to better understand potential causal linkages between environmental drivers and stream biological attributes or condition and predict expected values of macroinvertebrate metrics at unsampled sites.

APPENDIX

TABLE A1. Sources of Geographical Information System (GIS) and Digital Data Used in Model Development.

Spatial Dataset	Data Source	Source Data Format	Processing Format	Resolution/ Scale	Reference
Hydrography	National Hydrography Dataset (NHD)	Vector	Vector	1:24,000	U.S. Geological Survey, National Hydrography Dataset, Digital data, *accessed* January 2007 at http://nhd.usgs.gov/data.html
Land Cover 1992	National Land Cover Dataset 1992 (NLCD)	Raster	Vector	30 m	U.S. Geological Survey, National Land Cover Dataset 1992, Digital data, *accessed* March 2003 at http://landcover.usgs.gov/natllandcover.php
Land Cover 2001	NLCD 2001	Raster	Vector	30 m	U.S. Geological Survey, National Land Cover Dataset 2001, Digital data, *accessed* January 2007 at http://www.mrlc.gov/
Elevation	National Elevation Dataset (NED)	Raster	Raster	10 m	U.S. Geological Survey, National Elevation Dataset, Digital data, *accessed* May 2007 at http://seamless.usgs.gov/
Slope	NED	Raster	Raster	10 m	U.S. Geological Survey, National Elevation Dataset, Digital data, *accessed* May 2007 at http://seamless.usgs.gov/
Road networks	U.S. Census Bureau Tiger	Vector	Vector	1:100,000	U.S. Census Bureau, TIGER line data, Digital data, *accessed* May 2007 at http://www.census.gov/geo/www/tiger/
	Ground Transportation Roads Publications Arc	Vector	Vector	1:24,000	Oregon BLM, Ground Transportation Roads Publication Arc, Digital data, *accessed* July 2007 at http://www.blm.gov/or/gis/
Soil infiltration capacity	USDA NRCS STATSGO	Vector	Vector	1:250,000	Natural Resource Conservation Service, STATSGO soils data, Digital data, *accessed* May 2007 at http://datagateway.nrcs.usda.gov/
Population density	U.S. Census Bureau Census 2000	Vector	Raster	30 m	U.S. Census Bureau, Census 2000, Digital data, *accessed* May 2007 at http://www.census.gov/main/www/cen2000.html
Precipitation	Oregon State University PRISM	Raster	Raster	30 arc-seconds	PRISM Group, Oregon State University, Precipitation data for the U.S., Digital data, *accessed* May 2007 at http://www.prismclimate.org
Dams	National Inventory of Dams	Vector	Vector	Various	U.S. Army Corps of Engineers, National Inventory of Dams, Digital data, Not publicly available

ACKNOWLEDGMENTS

We thank each of the individuals that provided the macroinvertebrate data and associated advice: Shannon Hubler, Oregon Department of Environmental Quality; Alan Herlihy, Oregon State University and Environmental Protection Agency; Pete Ode and Andy Rehn, California Department of Fish and Game. We thank the many dedicated USGS colleagues who acquired the biological samples used in this analysis. We also gratefully acknowledge Alan Herlihy, John Van Sickle, and Anthony Olsen (Environmental Protection Agency, Corvallis) for their insight and advice on issues related to modeling macroinvertebrates, and Dr. Yangdong Pan (Portland State University) for advice on R programming and statistics. Finally, we thank one USGS reviewer and one outside reviewer for their review of the manuscript prior to submission to the journal.

LITERATURE CITED

Aertsena, W., V. Kinta, J. Van Orshovena, K. Özkanb, and B. Muysa, 2010. Comparison and Ranking of Different Modelling Techniques for Prediction of Site Index in Mediterranean Mountain Forests. Ecological Modelling 221:1119-1130.

Allan, J.D., 2004. Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. Annual Review of Ecology, Evolution and Systematics 35:257-284.

Austin, M., 2007. Species Distribution Models and Ecological Theory: A Critical Assessment and Some Possible New Approaches. Ecological Modelling 200:1-19.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling, 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish (Second Edition). U.S. Environmental Protection Agency, Office of Water, EPA Report 841-B-99-002, Washington, D.C.

Booth, D.B., 2005. Challenges and Prospects for Restoring Urban Streams: A Perspective From the Pacific Northwest of North America. Journal of the North American Benthological Society 24:724-737.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, 1984. Classification and Regression Trees. Wadsworth and Brooks/Cole, Monterey, California.

Cabecinha, E., P. Silva-Santos, R. Cortes, and J. Cabral, 2007. Applying a Stochastic-Dynamic Methodology (StDM) to Facilitate Ecological Monitoring of Running Waters, Using Selected Trophic and Taxonomic Metrics as State Variables. Ecological Modelling 207:109-127.

Carlisle, D.M., J. Falcone, and M.R. Meador, 2008. Predicting the Biological Condition of Streams: Use of Geospatial Indicators of Natural and Anthropogenic Characteristics of Watersheds. Environmental Monitoring and Assessment 151:143-160.

Chessman, B.C., H.A. Jones, N.K. Searle, I.O. Growns, and M.R. Pearson, 2010. Assessing Effects of Flow Alteration on Macroinvertebrate Assemblages in Australian Dryland Rivers. Freshwater Biology 55:1780-1800.

Clarke, R.T., 2000. Uncertainty in Estimates of River Quality Based on RIVPACS. *In*: Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques, J.F. Wright, D.W. Sutcliffe, and M.T. Furse (Editors). Freshwater Biological Association, Ambleside, pp. 39-54.

Clarke, R.T. and J.F. Murphy, 2006. Effects of Locally Rare Taxa on the Precision and Sensitivity of RIVPACS Bioassessment of Freshwaters. Freshwater Biology 51:1924-1940.

Clements, W.H., N.K.M. Vieira, and D.L. Sonderegger, 2010. Use of Ecological Thresholds to Assess Recovery in Lotic Ecosystems. Journal of the North American Benthological Society 29:1017-1023.

Cuffney, T.F., 2003. User's Manual for the National Water-Quality Assessment Program Invertebrate Data Analysis System (IDAS) Software: Version 3. U.S. Geological Survey Open-File Report 03-172, 103 pp.

Cuffney, T.F., M.D. Bilger, and A.M. Haigler, 2007. Ambiguous Taxa: Effects on the Characterization and Interpretation of Invertebrate Assemblages. Journal of the North American Benthological Society 26:286-307.

Cuffney, T.F., R.A. Brightbill, J.T. May, and I.R. Waite, 2010. Responses of Benthic Macroinvertebrates to Environmental Changes Associated With Urbanization in Nine Metropolitan Areas. Ecological Applications 20:1384-1401.

Cuffney, T.F., H. Zappia, E.M.P. Giddings, and J.F. Coles, 2005. Urbanization Effects on Benthic Macroinvertebrate Assemblages in Contrasting Environmental Settings: Boston, Birmingham, and Salt Lake City. *In*: Effects of Urbanization on Stream Ecosystems, L.R. Brown, R.H. Gray, R.M. Hughes, and M.R. Meador (Editors). American Fisheries Society Symposium 47: 361-408.

Cutler, D.R., T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler, 2007. Random Forests for Classification in Ecology. Ecology 88:2783-2792.

Davies, P.E., 2000. Development of a National River Bioassessment System (AUSRIVAS) in Australia. *In*: Assessing the Biological Quality of Freshwaters: RIVPACS and Other Techniques, J.F. Wright, D.W. Sutcliffe, and M.T. Furse (Editors). Freshwater Biological Association, Cumbria, United Kingdom, pp. 113-124.

Davis, W.S. and T.P. Simon, Editors. 1995. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, Florida.

De'ath, G., 2007. Boosted Trees for Ecological Modeling and Prediction. Ecology 88:243-251.

De'ath, G. and K.E. Fabricius, 2000. Classification and Regression Trees: A Powerful Yet Simple Technique for Ecological Data Analysis. Ecology 81:3178-3192.

Dodds, W.K., W.H. Clements, K. Gido, R.H. Hilderbrand, and R.S. King, 2010. Thresholds, Breakpoints, and Nonlinearity in Freshwaters as Related to Management. Journal of the North American Benthological Society 29:988-997.

Elith, J., J.R. Leathwick, and T. Hastie, 2008. A Working Guide to Boosted Regression Trees. Journal of Animal Ecology 77:802-813.

Fry, J.A., M.J. Coan, C.G. Homer, D.K. Meyer, and J.D. Wickham, 2009. Completion of the National Land Cover Database (NLCD) 1992-2001 Land Cover Change Retrofit Product. U.S. Geological Survey Open-File Report 2009-1379.

Grace, J.B., 2006. Structural Equation Modeling and Natural Systems. Cambridge University Press, New York, 365 pp.

Hawkins, C.P., 2006. Quantifying Biological Integrity by Taxonomic Completeness: Its Utility in Regional and Global Assessments. Ecological Applications 16:1277-1294.

Hawkins, C.P., Y. Cao, and B. Roper, 2010. Method of Predicting Reference Condition Biota Affects the Performance and Interpretation of Ecological Indices. Freshwater Biology 55:1066-1085.

Herbst, D.B. and E.L. Silldorff, 2006. Comparison of the Performance of Different Bioassessment Methods: Similar Evaluations of Biological Integrity From Separate Programs and Procedures. Journal of the North American Benthological Society 25:513-530.

Hilderbrand, R.H., R.M. Utz, S.A. Stranko, and R.L. Raesly, 2010. Applying Thresholds to Forecast Potential Biodiversity Loss From Human Development. Journal of the North American Benthological Society 29:1009-1016.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan, 2004. Development of a 2001 National Landcover Database for the United

States. Photogrammetric Engineering and Remote Sensing 70(7):829-840.

Hubler, S., 2008. Wadeable Stream Conditions in Oregon. Oregon Department of Environmental Quality Laboratory Division—Watershed Assessment Section, Portland, Oregon. DEQ07-LAB-0081-TR.

Kennen, J.G., M. Chang, and B.H. Tracy, 2005. Effects of Landscape Change on Fish Assemblage Structure in a Rapidly Growing Metropolitan Area in North Carolina, USA. *In*: Effects of Urbanization on Stream Ecosystems, L.R. Brown, R.H. Gray, R.M. Hughes, and M.R. Meador (Editors). American Fisheries Society Symposium 47: 39-52.

King, R.S. and M.E. Baker, 2010. Considerations for Analyzing Ecological Community Thresholds in Response to Anthropogenic Environmental Gradients. Journal of the North American Benthological Society 29:998-1008.

Leathwick, J.R., D. Rowe, J. Richardson, J. Elith, and T. Hastie, 2005. Using Multivariate Adaptive Regression Splines to Predict the Distributions of New Zealand's Freshwater Diadromous Fish. Freshwater Biology 50:2034-2052.

Maloney, K.O., D.E. Weller, M.J. Russell, and T. Hothorn, 2009. Classifying the Biological Condition of Small Streams: An Example Using Benthic Macroinvertebrates. Journal of the North American Benthological Society 28:869-884.

Morgan, R.P. and S.F. Cushman, 2005. Urbanization Effects on Stream Fish Assemblages in Maryland, USA. Journal of the North American Benthological Society 24:643-655.

Moss, D., 2000. Evolution of Statistical Methods in RIVPACS. *In*: Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques, J.F. Wright, D.W. Sutcliffe, and M.T. Furse (Editors). Freshwater Biological Association, Ambleside, pp. 25-37.

Moulton, II, S.R., J.G. Kennen, R.M. Goldstein, and J.A. Hambrook, 2002. Revised Protocols for Sampling Algal, Invertebrate and Fish Communities as Part of the National Water-Quality Assessment Program. U.S. Geological Survey Open-File Report 02-150, 75 pp.

Oberdorff, T., D. Pont, B. Hugueny, and D. Chessel, 2001. A Probabilistic Model Characterizing Fish Assemblages of French Rivers: A Framework for Environmental Assessment. Freshwater Biology 46:399-415.

Ode, P.R., C.P. Hawkins, and R.D. Mazor, 2008. Comparability of Biological Assessments Derived From Predictive Models and Multimetric Indices of Increasing Geographic Scope. Journal of the North American Benthological Society 27:967-985.

Ode, P.R., A.C. Rehn, and J.T. May, 2005. A Quantitative Tool for Assessing the Integrity of Southern Coastal California Streams. Environmental Management 35:493-504.

Paul, M.J. and J.L. Meyer, 2001. Streams in the Urban Landscape. Annual Review of Ecology and Systematics 32:333-365.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchak, F.H. McCormick, S.A. Peterson, S.A. Ringold, T. Magee, and M. Cappaert, 2006. Environmental Monitoring and Assessment Program – Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA/620/R-06/003. Office of Research and Development, US EPA, Corvallis, Oregon.

Prasad, A.M., L.R. Iverson, and A. Liaw, 2006. Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. Ecosystems 9:181-199.

Qian, S.S. and T.F. Cuffney, 2012. To Threshold or Not to Threshold? That's the Question. Ecological Indicators 15:1-9.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Roy, A.H., M.C. Freeman, B.J. Freeman, S.J. Wenger, W.E. Ensign, and J.L. Meyer, 2005. Investigating Hydrologic Alteration as a Mechanism of Fish Assemblage Shifts in Urbanizing Streams. Journal of the North American Benthological Society 24:656-678.

Stanford, J.A., M.S. Lorang, and F.R. Hauer, 2005. The Shifting Habitat Mosaic of River Ecosystems. Verhandlungen Internationaler Vereinigung fur Limnologie 29:123-136.

Therneau, T.M. and E.J. Atkinson, 1997. An Introduction to Recursive Partitioning Using the RPART Routines. Technical report no. 61. Mayo Clinic, Rochester, MN, 52 pp.

Turak, E., S. Ferrier, T. Barrett, E. Mesley, M. Drielsma, G. Manion, G. Doyle, J. Stein, and G. Gordon, 2011. Planning for the Persistence of River Biodiversity: Exploring Alternative Futures Using Process-Based Models. Freshwater Biology 56:39-56.

Uhrich, M.A. and D.A. Wentz, 1999. Environmental Setting of the Willamette Basin, Oregon. U.S. Geological Survey Water-Resources Investigations Report 97-4082-A, 20 pp.

U.S. Census Bureau, 2000. Census 2000 Redistricting Data Summary File: U.S. Census Bureau Technical Documentation Public Law, 94-171, 223 pp.

U.S. Geological Survey, 2007, National Hydrography Dataset (High Resolution), Digital Data, http://nhd.usgs.gov/data.html, *accessed* November 2007.

Van Sickle, J., J. Baker, A. Herlihy, P. Bayley, S. Gregory, and P. Haggerty, 2004. Projecting the Biological Condition of Streams Under Alternative Scenarios of Human Land Use. Ecological Applications 14:368-380.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel, 2001. Completion of the 1990s National Land Cover Data Set for the Conterminous United States From Landsat Thematic Mapper Data and Ancillary Data Sources. Photogrammetric Engineering and Remote Sensing 67:650-652.

Waite, I.R., L.R. Brown, J.G. Kennen, J.T. May, T.F. Cuffney, J.L. Orlando, and K.A. Jones, 2010. Comparison of Watershed Disturbance Predictive Models for Stream Benthic Macroinvertebrates for Three Distinct Ecoregions in Western US. Ecological Indicators 10:1125-1136.

Waite, I.R., S. Sobieszczyk, K.D. Carpenter, A.J. Arnsberg, H.M. Johnson, C.A. Hughes, M.J. Sarantou, and F.A. Rinella, 2008. Effects of Urbanization on Stream Ecosystems in the Willamette River Basin and Surrounding Area, Oregon and Washington. Chapter D of "Effects of Urbanization on Stream Ecosystems in six Metropolitan Areas of the United States". Scientific Investigations Report 2006-5101-D, 62 pp.

Walsh, C.J., T.D. Fletcher, and A.R. Ladson, 2005. Stream Restoration in Urban Catchments Through Redesignin Stormwater Systems: Looking to the Catchment to Save the Stream. Journal of the North American Benthological Society 24:690-705.

Wang, L.J., J. Lyons, P. Kanehl, and R. Bannerman, 2001. Impacts of Urbanization on Stream Habitat and Fish Across Multiple Scales. Environmental Management 28:255-266.

Wisseman, B., 1996. Version 1.0. Unpublished Ecological Coding Attributes for Freshwater Invertebrate Taxa in Western North America. Available from Aquatic Biology Associates, Inc., Corvallis, Oregon.