



State Water Resources Control Board

Policy for Toxicity Assessment and Control

Public Review Draft
June 2012

Preamble

Toxicity testing is an essential component of an integrated approach to water quality-based toxics control. Aquatic toxicity tests (toxicity tests) utilize aquatic organisms to examine the adverse chronic or acute effects of a given discharge. The results from these tests are used to detect aggregate toxic effects of known pollutants, and provide meaningful data when specific pollutants may not be known.

Previously, the Regional Water Quality Control Boards (Regional Water Boards) developed toxicity provisions in their respective Regional Water Quality Control Plans (Basin Plans), while section 4 of the Policy for Implementation of Toxics Standards for Inland Surface Waters, Enclosed Bays, and Estuaries of California (2005) established minimum testing requirements. As a result, toxicity requirements varied widely among the Basin Plans and permits. This Policy for Toxicity Assessment and Control (Policy) improves regulatory consistency through the adoption of statewide numeric objectives for chronic and acute toxicity. In addition, this Policy establishes a uniform approach to toxicity monitoring, analysis, and remediation measures that fulfill the requirements of State Water Resources Control Board (State Water Board) Resolution No. 2005-0019.

Applicability of Policy

Part I of this Policy provides definitions applicable to the Policy. Part II of this Policy establishes water quality objectives for toxicity that apply to all inland surface waters, enclosed bays, and estuaries of the state, including both waters of the United States and surface waters of the state. This Policy does not apply to ocean waters, including Monterey Bay and Santa Monica Bay. Part III of this Policy establishes toxicity test implementation procedures and assessment methodology for dischargers subject to this Policy. This Policy does not apply to sediment toxicity testing.

This Policy supersedes the toxicity control provisions in section 4 of the Policy for Implementation of Toxics Standards for Inland Surface Waters, Enclosed Bays, and Estuaries of California (2005), and all toxicity testing provisions established in the Basin Plans. This Policy does not supersede the narrative toxicity objectives established in the Basin Plans. This Policy establishes minimum requirements to protect aquatic life beneficial uses including, but not limited to, warm freshwater habitat (WARM), cold freshwater habitat (COLD), wildlife habitat (WILD), estuarine habitat (EST), commercial and sport fishing (COMM), marine habitat (MAR), inland saline water habitat (SAL), and wetland habitat (WET). This Policy shall be reevaluated by the State Water Board five years from its effective date.

Part I: Definitions

The following definitions apply to this Policy:

- A. Acute toxicity tests** measure the adverse effect (usually mortality) of a waste discharge or ambient water sample on a group of test organisms during a short-term exposure (e.g. 24, 48, or 96 hours).
- B. Applicable water board or water boards** means the State Water Resources Control Board (State Water Board) or Regional Water Quality Control Board (Regional Water Board) that issues a National Pollutant Discharge Elimination System (NPDES) permit, Waste Discharge Requirements (WDR), or conditional waivers of WDRs to a qualifying discharger.
- C. California Environmental Quality Act (CEQA)** means the statute set forth in Public Resources Code section 21000 et seq.
- D. Channelized dischargers regulated exclusively under the Porter-Cologne Water Quality Control Act (channelized dischargers)** means dischargers subject to conditional waivers, conditional prohibitions, or nonpoint source waste discharge requirements (WDR) that direct their discharge through a channel into surface waters not regulated under the NPDES permit program (e.g. dischargers subject to the Irrigated Lands Regulatory Program).
- E. Chronic toxicity tests** measure the sub-lethal effects of a discharge or ambient water sample (e.g. reduced growth or reproduction). Certain chronic toxicity tests include an additional measurement of lethality.
- F. Continuous dischargers** are NPDES wastewater dischargers and point source WDR dischargers that discharge without interruption throughout the operating hours of the facility, except for infrequent shutdowns for maintenance, process changes, or other similar activities.
- G. Individual industrial storm water dischargers** are industrial facilities that are issued an individual NPDES permit to discharge storm water and do not discharge to a municipal separate storm sewer system (MS4).
- H. Insignificant dischargers** are discharging entities that are deemed a very low threat to water quality by the applicable Water Board.
- I. Instream waste concentration (IWC)** is the concentration of a toxicant or effluent in the receiving water after mixing (the inverse of the dilution factor). A discharge of 100 percent effluent will be considered the IWC whenever mixing zones or dilution credits are not authorized by the applicable Water Board.
- J. Non-continuous dischargers** are NPDES wastewater and point source WDR dischargers that do not discharge on a continuous basis, and include facilities that discharge on an intermittent and seasonal basis.

- K. Maximum daily effluent limit (MDEL)** is, for the purposes of this Policy, an effluent limit based on the outcome of the Test of Significant Toxicity (TST) approach and the resulting percent effect at the IWC (see Part I (O)). The MDEL is exceeded when a toxicity test results in a “fail,” and the percent effect is greater than or equal to 0.50 for chronic toxicity tests or the percent effect is greater than or equal to 0.40 for acute toxicity tests.
- L. Median monthly effluent limit (MMEL)** is, for the purposes of this Policy, an effluent limit based on the median results of three independent toxicity tests, conducted within the same calendar month, and analyzed using the TST. The MMEL is exceeded when the median result (i.e. two out of three) is a “fail.”
- M. Municipal separate storm sewer system (MS4)** is any conveyance or system of conveyances used for collecting and conveying storm water runoff or a system of discharges from municipal separate storm sewers.
- N. NPDES wastewater dischargers** are dischargers subject to NPDES permitting requirements, but are not in the storm water program including, but not limited to, publicly owned treatment works (POTW).
- O. Percent effect** is the value that denotes the difference in response between the IWC and the control, divided by the mean response, and multiplied by 100 (see the equation in Step 6 of Appendix A).
- P. Point source WDR dischargers** means entities that discharge from industrial or wastewater treatment facilities into surface waters of the state, pursuant to WDRs that do not serve as NPDES permits (i.e. the receiving water is not a water of the United States, as that term is defined pursuant to Code of Federal Regulations, title 40, section 122.2).
- Q. Publicly owned treatment works (POTW)** are facilities owned by a state or municipality that store, treat, recycle, and reclaim municipal sewage or industrial wastes of a liquid nature.
- R. Reasonable potential** is a designation used for a waste discharge that is projected or calculated to cause or contribute to an excursion above a water quality standard. For the purposes of this Policy, reasonable potential is demonstrated when the IWC of a discharge produces a test result of “fail” or when the percent effect at the IWC is greater than 0.10.
- S. Regulatory management decision (RMD)** is the decision that represents the maximum allowable error rates and thresholds for chronic and acute toxicity (and non-toxicity) that would result in an acceptable risk to aquatic life.
- T. Replicates** are two or more independent organism exposures of the same treatment (i.e. effluent concentration) within a toxicity test. Replicates are typically conducted with separate test chambers and test organisms, each having the same effluent concentration.
- U. Response** is the measured biological endpoint(s) (e.g. survival, growth, and reproduction) used in a toxicity test method. The responses from the control and the IWC are quantified using statistical approaches to determine if toxicity is present.

- V. Small communities** are communities with populations of 20,000 or less, and a median household income below 80 percent of the statewide median household income. Communities with a population of 20,000 persons or less, and an MHI above 80 percent of the statewide median may also be considered disadvantaged on a case-by-case basis by the applicable Water Board, provided that a minimum of four percent of their MHI is paid toward wastewater infrastructure.
- W. Test of Significant Toxicity (TST)** is a statistical approach used to analyze toxicity test data. The TST incorporates a restated null hypothesis, Welch's t-test, and biological effect thresholds for chronic and acute toxicity.
- X. Toxicity** means the aggregate toxic effect of a waste discharge measured directly by a chronic or acute toxicity test. This aggregate effect is frequently referred to as "whole effluent toxicity."
- Y. Toxicity Reduction Evaluation (TRE)** is a site-specific study, conducted in a stepwise process that is designed to mitigate toxicity.

PART II: Toxicity Objectives

The following numeric toxicity objectives apply to all inland surface waters, enclosed bays, and estuaries of the state, including waters of the United States:

Chronic Toxicity

The chronic toxicity objective is expressed as a null hypothesis and a regulatory management decision (RMD) of 0.75 for chronic toxicity methods, where the following null hypothesis shall be used:

$$H_0: \text{Mean response (IWC)} \leq 0.75 \cdot \text{mean response (control)}$$

Attainment of the water quality objective is demonstrated by rejecting this null hypothesis in accordance with the statistical approach described in Appendix A.

Acute Toxicity

The acute toxicity objective is expressed as a null hypothesis and an RMD of 0.80 for acute toxicity methods, where the following null hypothesis shall be used:

$$H_0: \text{Mean response (IWC)} \leq 0.80 \cdot \text{mean response (control)}$$

Attainment of the water quality objective is demonstrated by rejecting this null hypothesis in accordance with the statistical approach described in Appendix A.

PART III: Implementation Procedures

Implementation procedures for National Pollutant Discharge Elimination System (NPDES) wastewater dischargers and point source Waste Discharge Requirements (WDR) dischargers are contained in Part III (A). Implementation procedures for storm water dischargers regulated pursuant to NPDES permits are contained in Part III (B). Implementation procedures for channelized dischargers regulated exclusively under the Porter-Cologne Water Quality Control Act (channelized dischargers) are contained in Part III (C).

A. NPDES Wastewater Dischargers and Point Source WDR Dischargers

1. Reasonable Potential Analyses and Species Sensitivity Screening

Except as otherwise provided in Part III (A)(9), prior to any permit issuance, reissuance, or reopening to address toxicity requirements that occurs after the effective date of this Policy, all NPDES wastewater dischargers and point source WDR dischargers shall conduct a reasonable potential analysis pursuant to the procedures established in this section to determine if their waste discharge has the reasonable potential to cause or contribute to an excursion above the chronic toxicity objective established in Part II. The applicable Water Board shall have the discretion to require reasonable potential analyses for acute toxicity. If a reasonable potential analysis for acute toxicity is required, the applicable Water Board will document the rationale in the NPDES fact sheet or WDR information sheet (or equivalent document).

POTWs that discharge one million gallons a day or more are classified as having reasonable potential to cause or contribute to an excursion above the chronic toxicity objective established in Part II. This is due to the steady and voluminous flow of influent these facilities receive from a variety of dischargers. Accordingly, the applicable Water Board shall ensure that publicly owned treatment works (POTW) that discharge one million gallons a day or more use the procedures of this section only to identify or confirm the most sensitive test species for routine monitoring use.

Test method selection is determined by salinity and tier classification (refer to Table 1 of Appendix A). Freshwater test methods shall be used for receiving waters with salinity less than 1,000 mg/L and marine test methods shall be used for receiving waters with salinity equal to or greater than 1,000 mg/L. However, NPDES wastewater and point source WDR dischargers that discharge freshwater effluent to marine waters may use freshwater test methods as determined by the applicable Water Board. Tier I test species are preferred for marine test methods, but the applicable Water Board may allow the use of Tier II test species if Tier I species are not available.

At a minimum, reasonable potential analyses and species sensitivity screenings for chronic toxicity shall include one vertebrate, one invertebrate and one aquatic plant. If the applicable Water Board requires a reasonable potential analysis/species sensitivity screening for acute toxicity, one vertebrate and one invertebrate shall be used. A minimum of four single-concentration toxicity tests utilizing the IWC and control shall be performed for each test species used. In addition, each set of tests conducted during the species sensitivity screening shall be conducted concurrently, or with discharge samples collected at the same time, or during overlapping times. The test methods

established in Code of Federal Regulations, title 40, section 136.3 and Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms, First Edition (EPA-600-R-95-136) shall be followed when conducting a reasonable potential analysis or species sensitivity screening. Test results shall be calculated using the Test of Significant Toxicity (TST), as described in Appendix A. Toxicity test data generated during a permit term active on or after the effective date of the Policy, or any valid data submitted for permit renewal may be used for reasonable potential analyses provided that the data meet all of the requirements established in Part III (A)(1). Reasonable potential is demonstrated if the effluent, at the IWC, produces a test result of “fail” as described in Appendix A. Toxicity test data that produce a test result of “pass” shall be further evaluated by the NPDES wastewater or point source WDR discharger to determine both reasonable potential and the most sensitive test species for use in routine monitoring. This evaluation shall be carried out by calculating the percent effect at the IWC, for each test result, using the following equation:

$$\% \text{ Effect at IWC} = \frac{\text{Mean Control Response} - \text{Mean IWC Response}}{\text{Mean Control Response}} \cdot 100$$

Based upon the foregoing, a waste discharge has reasonable potential to cause or contribute to an excursion above the toxicity objectives established in Part II if the effluent at the IWC produces a test result of “fail,” or if the percent effect at the IWC is greater than 0.10. A waste discharge does not have reasonable potential if the IWC passes each toxicity test and exhibits a percent effect at or below 0.10.

2. Numeric Effluent Limitations in Permits

If the applicable Water Board determines that reasonable potential exists for any NPDES wastewater discharger or point source WDR discharger, in accordance with Part III (A)(1), the applicable Water Board shall include numeric effluent limitations for chronic toxicity in any permit issued, reissued, or reopened to address toxicity requirements after the effective date of the Policy. The effluent limitations for chronic toxicity shall be expressed as a maximum daily effluent limitation (MDEL) and a median monthly effluent limitation (MMEL), in accordance with Code of Federal Regulations, title 40, section 122.45(d)(1)-(2).¹ The applicable Water Board has the discretion to include a numeric effluent limitation for acute toxicity. If acute toxicity limitations are included in the permit, the applicable Water Board shall document the need for acute limitations in the NPDES fact sheet or WDR information sheet (or equivalent document). If numeric effluent limitations for acute toxicity are imposed, they shall also be expressed as an MDEL and an MMEL.

The MDEL for chronic and acute toxicity shall be expressed as the outcome of the TST approach and the resulting percent effect. The MMEL for chronic and acute toxicity shall be expressed as the median result of three independent toxicity tests conducted within the same calendar month. Compliance with the MDEL and MMEL shall be determined using the TST approach, as described in Appendix A.

¹ Establishing an average monthly effluent limitation is impracticable using the TST statistical method, because the primary output of the approach is a non-numeric result of “pass” or “fail.”

Appropriate monitoring frequencies for chronic toxicity effluent limitations are established in Part III (A)(4)(a). Compliance with chronic and acute numeric effluent limitations shall be determined according to the TST approach, as described in Appendix A, and Part III (A)(6). Mixing zones and dilution credits, as established in an appropriate plan or policy, may be applied to these numeric effluent limitations. (Refer to Appendix B for an example of permit limitation language and Appendix C for examples of compliance determination.)

3. Test Methods

NPDES wastewater and point source WDR dischargers shall follow the methods for chronic toxicity tests as established in Code of Federal Regulations, title 40, section 136.3 using a single-concentration test design for routine monitoring, or a five-concentration test design for accelerated monitoring. The United States Environmental Protection Agency (U.S. EPA) method manuals referenced therein include Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, Fourth Edition (EPA-821-R-02-013), and Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms, Third Edition (EPA-821-R-02-014). Additional methods for chronic toxicity monitoring are outlined in Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms, First Edition (EPA-600-R-95-136). Dischargers required to monitor acute toxicity shall follow the toxicity test methods established in Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms, Fifth Edition (EPA-821-R-02-012).

4. Routine Monitoring

NPDES wastewater and point source WDR dischargers that demonstrate reasonable potential, as determined in Part III (A)(1), are required to conduct routine chronic toxicity monitoring at a frequency no less than that established in Part III (A)(4)(a) of this Policy. If the applicable Water Board determines that a discharger demonstrates reasonable potential to exceed the acute toxicity objective, the discharger shall conduct routine acute toxicity monitoring, in addition to chronic toxicity monitoring. The test species that exhibits the highest percent effect at the IWC during a reasonable potential analysis/species sensitivity screening (i.e. the most sensitive species) shall be utilized for routine monitoring during the permit cycle. Routine toxicity test design shall, at a minimum, include a single-concentration analysis of the IWC compared to a control. Results shall be analyzed using the TST method as described in Appendix A. In the absence of reasonable potential, the applicable Water Board has the discretion to require NPDES wastewater and point source WDR dischargers to conduct periodic monitoring for chronic or acute toxicity.

a. Monitoring Frequency

NPDES wastewater and point source WDR dischargers that are continuous dischargers that discharge at a rate equal to or greater than one million gallons per day, or POTWs with a dry weather design capacity of one million gallons per day, shall conduct one chronic toxicity test, every calendar month, for the duration of the permit. NPDES wastewater and point source WDR dischargers that are non-continuous dischargers that discharge at a rate equal to or greater than one million gallons per day shall conduct one chronic toxicity test every calendar month during which a discharge lasting more than two days occurs for the duration of the permit, but only during each period of discharge.

NPDES wastewater and point source WDR dischargers that are continuous dischargers discharging at a rate less than one million gallons per day shall conduct one chronic toxicity test each calendar quarter (e.g. January–March, April–June, etc.) for the duration of the permit. NPDES wastewater and point source WDR dischargers that are non-continuous dischargers discharging at a rate less than one million gallons per day shall conduct one chronic toxicity test each calendar quarter of the discharge period. A calendar quarter shall be counted whenever the discharge period lasts seven or more days during a calendar month. If required, acute toxicity monitoring shall be conducted at intervals determined by the applicable Water Board.

The rate of discharge shall be determined by the average of the daily discharge rates for a representative period of time prior to permit reissuance or reopening to address toxicity requirements. New permits issued to POTWs after the effective date of the Policy shall use the dry weather design capacity to determine flow rate, while all other NPDES wastewater and point source WDR dischargers shall use the highest expected rate of discharge. The rate of discharge for non-continuous dischargers shall be determined based only on the days a discharge occurs (i.e. no zeroes shall be used to calculate rate of discharge).

5. Statistical Method

Results obtained from single-concentration chronic and acute toxicity tests shall be analyzed using the TST approach. (Refer to Appendix A for step-by-step instructions for using the TST.)

6. Compliance Determination

NPDES wastewater and point source WDR dischargers shall report the results of reasonable potential analyses, species sensitivity screenings, and routine toxicity tests to the applicable Water Board as either a “pass” or a “fail” at the IWC, in accordance with the TST approach and provide the calculated percent effect at the IWC. (Refer to Appendix C for examples of compliance determination.)

a. Pass

A test result indicating a “pass” is interpreted as meeting the MDEL, MMEL and the objectives established in Part II. If a test results in a “pass,” dischargers shall continue routine monitoring in accordance with Part III (A)(4).

b. Fail

A chronic toxicity test result indicating a “fail” with a percent effect at or above 0.50 is an exceedance of the chronic MDEL. An acute toxicity test result indicating a fail with a percent effect at or above 0.40 is an exceedance of the acute MDEL. Dischargers will have the opportunity to implement corrective action, upon exceedance of the MDEL, if the source of toxicity is known (e.g. operational upset). Verification of the corrective action(s) will be determined with an additional toxicity test conducted within the same calendar month that the exceedance occurred. This verification test must result in a “pass.” If the verification test results in a “fail” at any percent effect, the discharger shall implement an accelerated monitoring schedule in accordance with Part III (A)(6)(c).

If an initial toxicity test (i.e. not a verification test) results in a “fail,” but the percent effect is below the MDEL, the discharger shall conduct two additional toxicity tests within the same calendar month in order to determine compliance with the MMEL. If either of these two additional tests results in a “fail,” the median monthly result is “fail” and the discharger will be in

exceedance of the MMEL. Exceeding an MDEL or MMEL will result in a violation and will require the implementation of an accelerated monitoring schedule in accordance with Part III (A)(6)(c).

c. Accelerated Monitoring

NPDES wastewater permits and point source WDRs issued, reissued, or reopened to address toxicity requirements after the effective date of this Policy shall include an accelerated monitoring schedule to be implemented following an exceedance of the MDEL or MMEL. At a minimum, an accelerated monitoring schedule shall consist of four, five-concentration chronic toxicity tests, conducted at approximately two-week intervals, over an eight-week period. All toxicity tests conducted during an accelerated monitoring schedule shall, at a minimum, include the IWC and four additional concentrations. The additional effluent concentrations are to provide useful information regarding the intensity of the toxic effect(s), should the discharger progress to a TRE. An IWC that results in a “fail” and exhibits a percent effect equal to, or greater than 0.25 for chronic toxicity tests or 0.20 for acute toxicity tests will require initiation of a TRE, as described in Part III (A)(6)(d)

d. Toxicity Reduction Evaluation

Any toxicity test conducted by an NPDES wastewater or point source WDR discharger, during accelerated monitoring, that results in a “fail” and exhibits a percent effect equal to, or greater than 0.25 for chronic toxicity tests or 0.20 for acute toxicity tests will be required to conduct a Toxicity Reduction Evaluation (TRE). A discharger shall conduct a TRE in accordance with a TRE Work Plan developed pursuant to the requirements of the applicable Water Board. When TREs are required of multiple facilities that discharge to the same water body, the facilities may coordinate the TREs with the approval of the applicable Water Board. Dischargers shall continue to conduct routine monitoring, in accordance with Part III (A)(4), for the duration of the TRE.

7. Violations

An exceedance of the MDEL or MMEL during routine monitoring is a violation. Any exceedances occurring during a required accelerated monitoring period and, if appropriate, a TRE period shall not constitute additional violations provided that: (1) the discharger proceeds with the accelerated monitoring and TRE (if required) in a timely manner; and (2) the accelerated monitoring and TRE are completed within six months of the initial exceedance. The applicable Water Board has the discretion to impose additional violations and initiate an enforcement action for toxicity tests that result in a “fail” after six months from the initial violation. Additionally, a discharger’s failure to initiate an accelerated monitoring schedule or conduct a TRE, as required by an NPDES wastewater permit or point source WDR, will result in all exceedances being considered violations of the MDEL or MMEL and may result in the initiation of an enforcement action.

8. Compliance Schedules

The applicable Water Board has the discretion to grant a compliance schedule to NPDES wastewater and point source WDR dischargers in order to achieve the objectives established in Part II. Compliance schedules shall be consistent with the State Water Board’s Policy for Compliance Schedules in National Pollutant Discharge Elimination System Permits, adopted on April 15, 2008 (Resolution No. 2008-0025), with the exception that the duration of the compliance schedule may not exceed two years from the date of permit issuance, reissuance,

or reopening to address toxicity requirements after the effective date of this Policy. The discretionary authority to grant compliance schedules, however, will expire ten years after the effective date of this Policy. In addition, dischargers with existing toxicity monitoring requirements are not eligible to receive a compliance schedule.

9. Exceptions

a. Small Communities

Small communities, as defined in Part I (Q), are exempt from the provisions of Part III unless the applicable Water Board finds them to have an impact on receiving water quality. Nothing in this section, however, precludes the applicable Water Board from requiring periodic toxicity testing for small communities.

b. Insignificant Dischargers

The Water Boards are authorized to exempt certain NPDES wastewater dischargers and point source WDR dischargers from the provisions of Part III (A) if the applicable Water Board finds that the discharge will have an insignificant impact on receiving water quality. Eligible dischargers must discharge less than one million gallons per day on a non-continuous basis.

c. Categorical Exceptions

The Water Boards may, after compliance with the California Environmental Quality Act (CEQA), allow short-term or seasonal exceptions from meeting the objectives established in Part II if determined to be necessary to implement control measures either:

1. For resources or pest management (e.g. vector or weed control, pest eradication, or fishery management) conducted by public entities or mutual water companies to fulfill statutory requirements including, but not limited to, those in the California Fish and Game, Food and Agriculture, Health and Safety, and Harbors and Navigation codes; or
2. Regarding drinking water conducted to fulfill statutory requirements under the federal Safe Drinking Water Act or the California Health and Safety Code. Such categorical exceptions may also be granted for draining water supply reservoirs, canals, and pipelines for maintenance; draining municipal storm water conveyances for cleaning or maintenance; or, draining water treatment facilities for cleaning or maintenance.

For each project, the discharger shall notify potentially affected public and governmental agencies. Also, the discharger shall submit to the Executive Officer of the applicable Water Board for approval:

- i. A detailed description of the proposed action, including the proposed method of completing the action;
- ii. A time schedule;
- iii. A discharge and receiving water quality monitoring plan (before project initiation, during the project, and after project completion, with the appropriate quality assurance and quality control procedures);
- iv. CEQA documentation;
- v. Contingency plans;
- vi. Identification of alternate water supply (if needed); and
- vii. Residual waste disposal plans.

Additionally, upon completion of the project, the discharger shall provide certification by a qualified biologist that the receiving water beneficial uses have been restored. To prevent unnecessary delays in taking emergency actions, or to expedite the approval process for expected or routine activities that fall under categorical exceptions, the discharger should file with the applicable Water Board, in advance of seeking approval from the applicable Water Board, the information required in Part III (A)(9)(c)(2)(i-vii), to the extent possible.

d. Case-by-Case Exceptions

Where site-specific conditions in individual water bodies or watersheds differ sufficiently from statewide conditions, and those differences cannot be addressed through other provisions of this Policy, the State Water Board may, in compliance with CEQA, subsequent to a public hearing, and with the concurrence of the U.S. EPA, grant an exception to meeting the objectives established in Part II or any other provision of this Policy where the State Water Board determines:

1. The exception will not compromise protection of enclosed bay, estuarine, and/or inland surface waters for beneficial uses; and
2. The public interest will be served.

B. Storm Water Dischargers Regulated Pursuant to NPDES Permits

This section applies to the monitoring of storm water discharges from municipal separate storm sewer systems (MS4), and individual industrial storm water discharge permits as defined in Part I (G). These provisions may be applied to the California Department of Transportation General Permit at the discretion of the State Water Board.

1. Application of TST Methodology to Existing Toxicity Monitoring Requirements

Within one year of the effective date of this Policy, the applicable Water Board shall issue Water Code section 13383 letters to MS4 and individual industrial storm water dischargers with existing toxicity monitoring requirements. These section 13383 letters shall require all toxicity data collected one year from the postmarked date of the section 13383 letter to be analyzed using the TST approach, as described in Appendix A.

If, after the effective date of the Policy, the applicable Water Board requires toxicity monitoring within new or reissued MS4 or individual industrial storm water permits, then the applicable Water Board shall also require the use of the TST approach for toxicity data analysis.

Results obtained from toxicity tests shall be reported to the applicable Water Board as either a “pass” or a “fail.” Dischargers that lack toxicity monitoring requirements shall be exempt from the provisions of Part III (B)(1).

2. Toxicity Monitoring Program Recommendations

It is recommended that all MS4 dischargers and individual industrial storm water dischargers implement a chronic toxicity monitoring program, if they are not currently required to do so. It is also recommended that these toxicity monitoring programs consist of four single-concentration toxicity tests, conducted each calendar quarter (e.g. January–March, April–June, etc.), during each year of the permit cycle. Dischargers are recommended to use samples from two storm events and two non-storm event flows, if non-storm event flows are present. Remediation measures are encouraged for MS4 and individual industrial storm water discharges that result in a “fail” at a percent effect equal to, or greater than 0.25 for chronic toxicity tests and 0.20 for acute toxicity tests. Additional recommendations are provided in Appendix E.

C. Channelized Dischargers Regulated Exclusively Under the Porter-Cologne Water Quality Control Act

This section applies to the monitoring of discharges from channelized dischargers regulated exclusively under the Porter-Cologne Water Quality Control Act (channelized dischargers) as defined in Part I (D).

1. Application of TST Methodology to Existing Toxicity Monitoring Requirements

Within one year of the effective date of this Policy, the applicable Water Board shall issue Water Code section 13267 letters to channelized dischargers required to monitor toxicity under existing requirements established in a conditional waiver, conditional prohibition, or nonpoint source WDR. These section 13267 letters shall require all toxicity data collected one year from the postmarked date of the section 13267 letter to be analyzed using the TST approach, as described in Appendix A.

If, after the effective date of the Policy, the applicable Water Board requires toxicity monitoring within a new or reissued conditional waiver, conditional prohibition, or nonpoint source WDR, then the applicable Water Board shall also require the use of the TST approach for toxicity data analysis.

Results obtained from toxicity tests shall be reported to the applicable Water Board as either a “pass” or a “fail” at a percent effect equal to, or greater than 0.25 for chronic toxicity tests and 0.20 for acute toxicity tests. Channelized dischargers that lack toxicity monitoring requirements in their conditional waiver or nonpoint source WDR shall be exempt from the provisions of Part III (C)(1).

2. Toxicity Monitoring Program Recommendations

It is recommended that all channelized dischargers implement a chronic toxicity monitoring program, if they are not currently required to do. It is also recommended that these toxicity monitoring programs consist of four single-concentration toxicity tests conducted each calendar quarter (e.g. January–March, April–June, etc.), during each year of the permit cycle. Remediation measures are encouraged for channelized discharges that result in a “fail.”

APPENDIX A

Test of Significant Toxicity

The TST is the recommended statistical approach for all toxicity monitoring programs. Results obtained from single-concentration chronic and acute toxicity tests shall be analyzed as follows:

For each test endpoint, follow Steps 1 through 5.

Step 1: Prior to analysis: if the measured response is reported as a percentage (e.g. percent survival, percent fertilization) it must be transformed using the arc sine square root transformation below. If the measured response is not reported as a percentage, skip Step 1 and proceed to Step 2.

Calculate the response proportion (RP) for each replicate:

$$RP = \frac{\text{Number of Surviving or Unaffected Organisms}}{\text{Number Exposed}}$$

Transform each RP to arc sine based on the following scenarios:

For $0 < RP < 1$

$$\text{Angle (in radians)} = \text{arc sine } \sqrt{(RP)}$$

For $RP = 0$

$$\text{Angle (in radians)} = \text{arc sine } \sqrt{1/4n}$$

Where n = number of organisms used for each replicate

For $RP = 1$

$$\text{Angle} = 1.5708 \text{ rad} - (\text{radians for } RP = 0)$$

Step 2: Conduct Welch's t-test using the following equation:

$$t = \frac{\bar{Y}_t - b \cdot \bar{Y}_c}{\sqrt{\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c}}}$$

where:

\bar{Y}_c	=	Mean response for the control
\bar{Y}_t	=	Mean response for the IWC
S_c^2	=	Estimate of the variance for the control
S_t^2	=	Estimate of the variance for the IWC
n_c	=	Number of replicates for the control
n_t	=	Number of replicates for the IWC
b	=	0.75 for chronic tests; 0.80 for acute tests

Step 3: Adjust the degrees of freedom, v , using the following equation:

$$v = \frac{\left(\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c} \right)^2}{\frac{\left(\frac{S_t^2}{n_t} \right)^2}{n_t - 1} + \frac{\left(\frac{b^2 S_c^2}{n_c} \right)^2}{n_c - 1}}$$

For tests using Welch's t-test, the degrees of freedom are obtained from v in the equation above. Since v is most likely a non-integer, round v to the next lowest integer.

Step 4: Using the calculated t-value from Step 2, compare the calculated t-value with the critical t-value in Table 2 using the test method-specific alpha values shown in Table 1. To obtain the critical t-value, look across the table for the α error value that corresponds to the toxicity test method and then look down the table for the appropriate degrees of freedom.

Step 5: If the calculated t-value is less than the critical t-value, the IWC is declared toxic and the test result is a "fail" at the IWC. If the calculated t-value is greater than the critical t-value, the IWC is not declared toxic and the test result is a "pass" at the IWC.

Step 6: Calculate the percent effect at the IWC using the following equation:

$$\% \text{ Effect at IWC} = \frac{\text{Mean Control Response} - \text{Mean IWC Response}}{\text{Mean Control Response}} \cdot 100$$

Refer to U.S. EPA's National Pollutant Discharge Elimination System Test of Significant Toxicity Implementation Document (EPA-833-R-10-003) for additional guidance.

Table 1. Summary of alpha (α) levels for toxicity test methods.

EPA Toxicity Test Method ¹	b Value	Tier	False Negative (α Error)
Chronic Freshwater Methods			
<i>Ceriodaphnia dubia</i> (water flea) reproduction ²	0.75	I	0.20
<i>Pimephales promelas</i> (fathead minnow) survival and growth	0.75	I	0.25
<i>Selenastrum capricornutum</i> (green alga) Growth	0.75	I	0.25
Chronic West Coast Marine Methods			
<i>Atherinops affinis</i> (topsmelt) survival and growth	0.75	I	0.25
<i>Dendraster excentricus</i> (sand dollar); <i>Strongylocentrotus purpuratus</i> (purple urchin) fertilization	0.75	I	0.05
<i>Dendraster excentricus</i> (sand dollar); <i>Strongylocentrotus purpuratus</i> (purple urchin) larval development	0.75	I	0.05
<i>Haliotis rufescens</i> (red abalone) larval development	0.75	I	0.05
<i>Mytilus sp.</i> (mussels); <i>Crassostrea gigas</i> (oyster) larval development methods	0.75	I	0.05
<i>Macrocystis pyrifera</i> (giant kelp) germination and germ-tube length	0.75	I	0.05
Chronic East Coast Marine Methods			
<i>Menidia beryllina</i> (inland silverside) larval survival and growth	0.75	II	0.25
<i>Americamysis bahia</i> (mysid) survival and growth	0.75	II	0.15
Acute Freshwater Methods			
<i>Ceriodaphnia dubia</i> ; <i>Daphnia magna</i> ; <i>Daphnia pulex</i> (water flea); <i>Hyalella azteca</i> (amphipod) Survival	0.80	I	0.10
<i>Pimephales promelas</i> (fathead minnow) <i>Oncorhynchus mykiss</i> (rainbow trout) <i>Salvelinus fontinalis</i> (brook trout) Survival	0.80	I	0.10
Acute Marine Methods			
<i>Atherinops affinis</i> (topsmelt); Survival	0.80	I	0.10
<i>Americamysis bahia</i> (mysid) Survival	0.80	II	0.10
<i>Menidia beryllina</i> (inland silverside) Survival	0.80	II	0.10

¹ The false positive rate (β error) is 0.05 for all test methods.

² The chronic *Ceriodaphnia dubia* test design for the survival endpoint is not amenable to a Welch's t-test, therefore the survival endpoint will be determined as a percent effect and a response equal or greater than 0.25 is considered toxic.

Table 2. Critical values of the t-distribution. One tail probability is assumed.

Degrees of Freedom (v)	α Error				
	0.25	0.20	0.15	0.10	0.05
1	1	1.3764	1.9626	3.0777	6.3138
2	0.8165	1.0607	1.3862	1.8856	2.92
3	0.7649	0.9785	1.2498	1.6377	2.3534
4	0.7407	0.941	1.1896	1.5332	2.1318
5	0.7267	0.9195	1.1558	1.4759	2.015
6	0.7176	0.9057	1.1342	1.4398	1.9432
7	0.7111	0.896	1.1192	1.4149	1.8946
8	0.7064	0.8889	1.1081	1.3968	1.8595
9	0.7027	0.8834	1.0997	1.383	1.8331
10	0.6998	0.8791	1.0931	1.3722	1.8125
11	0.6974	0.8755	1.0877	1.3634	1.7959
12	0.6955	0.8726	1.0832	1.3562	1.7823
13	0.6938	0.8702	1.0795	1.3502	1.7709
14	0.6924	0.8681	1.0763	1.345	1.7613
15	0.6912	0.8662	1.0735	1.3406	1.7531
16	0.6901	0.8647	1.0711	1.3368	1.7459
17	0.6892	0.8633	1.069	1.3334	1.7396
18	0.6884	0.862	1.0672	1.3304	1.7341
19	0.6876	0.861	1.0655	1.3277	1.7291
20	0.687	0.86	1.064	1.3253	1.7247
21	0.6864	0.8591	1.0627	1.3232	1.7207
22	0.6858	0.8583	1.0614	1.3212	1.7171
23	0.6853	0.8575	1.0603	1.3195	1.7139
24	0.6849	0.8569	1.0593	1.3178	1.7109
25	0.6844	0.8562	1.0584	1.3163	1.7081
26	0.684	0.8557	1.0575	1.315	1.7056
27	0.6837	0.8551	1.0567	1.3137	1.7033
28	0.6834	0.8546	1.056	1.3125	1.7011
29	0.683	0.8542	1.0553	1.3114	1.6991
30	0.6828	0.8538	1.0547	1.3104	1.6973
inf	0.6745	0.8416	1.0364	1.2816	1.6449

APPENDIX B

Example of Permit Effluent Limitation Language

The following is an example of a chronic effluent limitation that may be included in a permit pursuant to this Policy and is solely intended to serve as guidance for a permit writer:

The maximum daily effluent limitation (MDEL) for chronic toxicity is expressed as the outcome of the TST approach and the resulting percent effect at the instream waste concentration (IWC) for this discharge, and the monthly median effluent limit (MMEL) for chronic toxicity is expressed as the median result of three independent toxicity tests, conducted within the same calendar month. For this discharge, the IWC is [either 100 percent or an effluent concentration at the mixing zone to be determined at time of permit issuance or reissuance] percent effluent. To calculate either a “pass” or “fail” for a chronic toxicity test, follow the instructions in Appendix A of the State Policy for Toxicity Assessment and Control.

A “pass” result indicates no toxicity in the IWC and compliance with the MDEL, while a “fail” result with a percent effect at or above 0.50 indicates an exceedance of the MDEL and demonstrates toxicity at the IWC. If an exceedance of the MDEL occurs, the permittee will receive a violation, but will also have the opportunity to resolve the source of toxicity (if it is known) and return to routine monitoring the following calendar month. In order to verify the abatement of toxicity, a result of “pass” from one additional chronic toxicity test, conducted within the same calendar month that the exceedance occurred, is required. If the additional test results in a “fail,” the discharger will be required to implement an accelerated monitoring schedule [cite applicable monitoring section of the permit or cite Part III (A)(6)(c) of the State Policy for Toxicity Assessment and Control].

A “fail” result with a percent effect below 0.50 is not a violation of the MDEL, but will require the permittee to complete two additional chronic toxicity tests, within the same calendar month that the “fail” occurred, in order to determine compliance with the MMEL. Routine chronic toxicity monitoring will resume the following month if both of these additional tests result in a “pass.” Implementation of an accelerated monitoring schedule [cite applicable monitoring section of the permit or cite Part III (A)(6)(c) of the State Policy for Toxicity Assessment and Control] will be required if the permittee exceeds the MMEL. The permittee shall implement an approved Toxicity Reduction Evaluation (TRE) Work Plan if a chronic toxicity test results in a “fail” and exhibits a percent effect equal to or greater than 0.25 during an accelerated monitoring schedule.

APPENDIX C

Examples of Compliance Determination

a. Chronic *Ceriodaphnia dubia* reproduction test

Replicate/Statistic	Control	IWC
1	29	31
2	38	28
3	31	25
4	34	28
5	36	22
6	35	21
7	30	27
8	31	26
9	36	29
10	34	30
Mean	33.4	26.7
Standard Deviation	2.989	3.268
# of Replicates (n)	10	10

- 1) Transform data with arcsine square root transformation if applicable (not necessary for this type of data).
- 2) Conduct Welch's t-test.

$$t = \frac{\bar{Y}_t - b \times \bar{Y}_c}{\sqrt{\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c}}} = \frac{26.7 - (0.75 \times 33.4)}{\sqrt{\frac{10.68}{10} + \frac{(0.75)^2 (8.93)}{10}}} = 1.32$$

- 3) Adjust the degrees of freedom.

$$v = \frac{\left(\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c}\right)^2}{\frac{\left(\frac{S_t^2}{n_t}\right)^2}{n_t - 1} + \frac{\left(\frac{b^2 S_c^2}{n_c}\right)^2}{n_c - 1}} = \frac{\left(\frac{10.68}{10} + \frac{(0.75)^2 (8.93)}{10}\right)^2}{\frac{(10.68)^2}{10 - 1} + \frac{((0.75)^2 (8.93))^2}{10 - 1}} = 15$$

- 4) Compare the calculated t-value with the critical t-value:
Given 16 degrees of freedom and an alpha level set at 0.20, the critical t-value = 0.86 (obtained from Table 2 in this Policy).
- 5) $1.32 > 0.86 = \text{pass}$

6) Calculate the percent effect at the IWC

$$\% \text{ Effect at IWC} = \frac{33.4 - 26.7}{33.4} \cdot 100 = 20.1\%$$

This test result would be in compliance with the MDEL and the MMEL.

b. Acute fish survival test

Replicate/Statistic	Control	IWC
1	10	10
2	10	8
3	10	9
4	10	8
Mean	10	8.75
Standard Deviation	0.000	0.958
# of Replicates (n)	4	4

1) Transform data with arcsine square root transformation if applicable.

Replicate/Statistic	Control	Treatment
1	1.412	1.412
2	1.412	1.107
3	1.412	1.249
4	1.412	1.107
Mean	1.412	1.219
Standard Deviation	0.000	0.145
# of Replicates (n)	4	4

2) Conduct Welch's t-test.

$$t = \frac{\bar{Y}_t - b \times \bar{Y}_c}{\sqrt{\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c}}} = \frac{1.219 - (0.80 \times 1.412)}{\sqrt{\frac{0.021}{4} + \frac{(0.80)^2 (0.00)}{4}}} = 1.23$$

3) Adjust the degrees of freedom.

$$v = \frac{\left(\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c}\right)^2}{\frac{\left(\frac{S_t^2}{n_t}\right)^2}{n_t - 1} + \frac{\left(\frac{b^2 S_c^2}{n_c}\right)^2}{n_c - 1}} = \frac{\left(\frac{0.021}{4} + \frac{(0.80)^2 (0.00)}{4}\right)^2}{\frac{(0.021)^2}{4 - 1} + \frac{((0.80)^2 (0.00))^2}{4 - 1}} = 3$$

4) Compare the calculated t-value with the critical t-value:

Given 3 degrees of freedom and an alpha level set at 0.10, the critical t-value = 1.64 (obtained from Table 2 in this Policy).

5) $1.23 < 1.64 = \text{fail}$

6) Calculate the percent effect at the IWC.

$$\% \text{ Effect at IWC} = \frac{1.412 - 1.219}{1.412} \cdot 100 = 13.7\%$$

This test result did not exceed the MDEL, but the discharger would need to conduct two additional tests to determine compliance with the MMEL. Dischargers that do not conduct these two additional tests would be in violation of the MMEL.

APPENDIX D

Decision trees

Figure 1. Reasonable Potential Analysis for Wastewater Dischargers

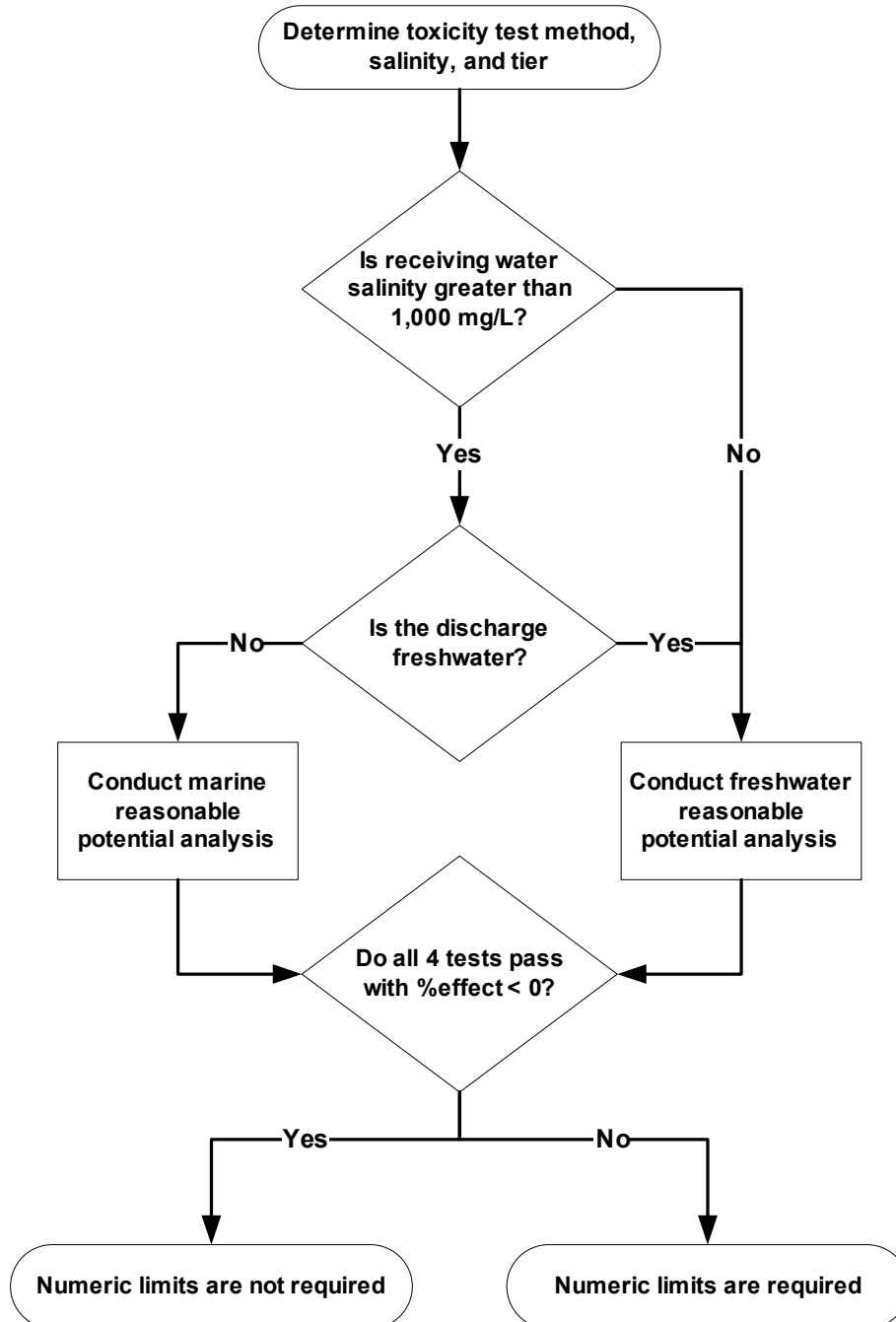
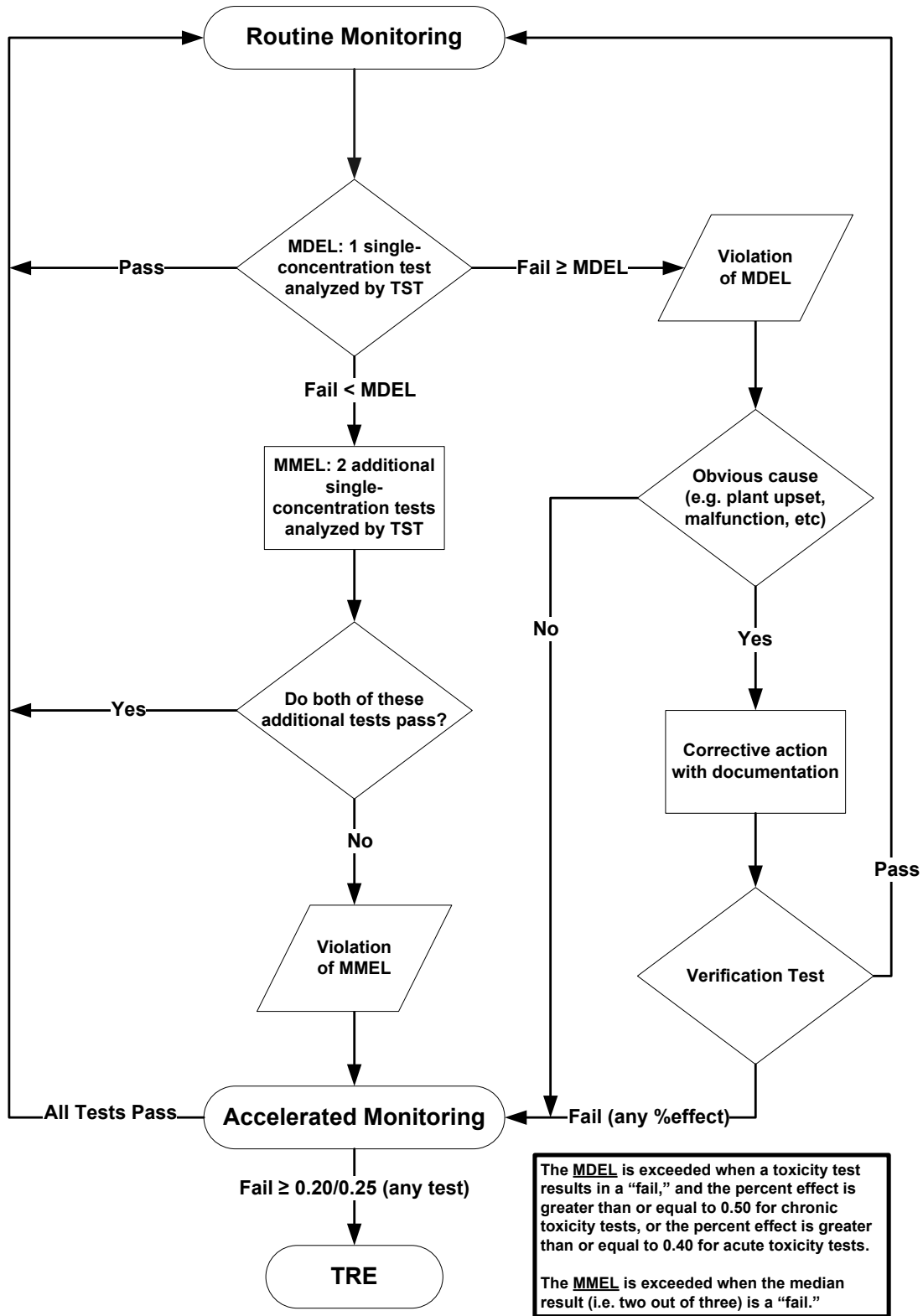


Figure 2. Compliance Determination for Wastewater Dischargers



APPENDIX E

Guidance for Storm Water Dischargers

The following is solely intended to serve as recommended guidance for Water Board staff and regulated stakeholders that implement toxicity testing in their MS4 NPDES permits. Minor edits have been made to the original draft of this guidance document in order to reflect revisions to the Policy for Toxicity Assessment and Control.

Toxicity Testing Tool for Storm Water Dischargers

June 2012

ACKNOWLEDGEMENTS

The recommendations provided in this document were developed with a Committee that represents a diverse array of perspectives, needs, skills, and desires. The Committee includes a focus on the primary audience; municipal storm water program managers, storm water technical staff, Regional Water Quality Control Board staff, and independent toxicologists. Two members of each category were selected, one from northern and one from southern California, to ensure broad geographic representation within the state.

Committee Members

Representation	Northern California	Southern California
Storm water Manager Liaison	Jim Scanlin (Alameda Clean Water Program)	Arlene Chun (Riverside County Flood Control)
Storm water Technical Liaison	Kelly Moran (TDC Environmental)	Ted VonBitner (Orange County Public Works)
Regional Water Quality Control Board Liaison	Stephanie Fong (Regional Water Board 5)	Mike Lyons (Regional Water Board 4)
Independent Technical Resource	Bryn Phillips (University of California Davis, Dept. of Environmental Toxicology)	Steve Bay (Southern California Coastal Water Research Project)
State Water Resources Control Board		Paul Hann
Environmental Protection Agency		Debra Denton

LIST OF ABBREVIATIONS

AMC: Antecedent moisture condition
CEDEN: California Environmental Data Exchange Network
CEE: Carboxylesterase enzyme
DDT: Dichlorodiphenyl Trichlorethane
EC25: 25% Effects concentration
EC50: 50% Effects concentration
EDTA: Ethylene diaminetetraacetic acid
EPA: United States Environmental Protection Agency
LC₅₀: Median lethal concentration
MS4: Municipal separate storm sewer system
NOEC: No Observed Effect Concentration
NPDES: National Pollutant Discharge Elimination System
PCB: Polychlorinated biphenyl
PBO: piperonyl butoxide
POTW: Publicly owned treatment works
SPE: Solid phase extraction
TAC: Test acceptability criteria
TIE: Toxicity identification evaluation
TMDL: Total maximum daily load
TRE: Toxicity reduction evaluation
TST: Test of Significant Toxicity

EXECUTIVE SUMMARY

The objective of this document is to provide implementation guidance to the State Water Resources Control Board (State Water Board) on storm water applications to accompany their upcoming Policy for Toxicity Assessment and Control (Policy). The primary audience for this guidance is Regional Water Quality Control Board (Regional Water Board) staff that writes and interprets toxicity testing language in municipal separate storm sewer systems (MS4) National Pollutant Discharge Elimination (NPDES) permits, and regulated stakeholders that implement toxicity testing in their MS4 NPDES permits. Four major topics necessary for implementation of a toxicity monitoring program are addressed: 1) sampling; 2) testing; 3) data management; and 4) Toxicity Identification Evaluations (TIEs). The recommendations are not binding and not meant to be one size fits all. Instead, the recommendations are meant to serve as a starting point for developing a monitoring and reporting program. Default recommendations are provided (Table ES1), but alternative options are available for many decision points. Regulatory or regulated agencies may decide to do more (or less) monitoring depending on site-specific or agency-specific needs.

Table ES1. List of default recommendations for monitoring toxicity in storm water.

	Default Recommendation	Additional Options
Sampling		
Station location	Integrator site	Targeted site(s) for specific sources
Frequency	2 storms and 2 dry weather per year	Use power analysis to optimize trend detection
Storm trigger	Forecasted 0.25 inches with 50% probability of precipitation	Median storm event at nearest rain gauge; account for antecedent dry conditions
Collection method	Discrete sampling on the rising limb of the hydrograph	Composite sampling to reduce within storm variability
Sample containers	Amber glass with Teflon lined lid, kept in the dark	Alternate materials when compatible with the toxicant and its properties
Testing		
Species selection	Multi-species screening at new sites	Focused species at known sites or for known contaminants
Holding time	36 hr, in the dark at <6 C	Do not test samples >72 h
Renewals	≤48 hr using original storm sample	-
Test acceptability criteria	As prescribed in the EPA test methods	-
Data Management		
Minimum fields	Test summary information	California Data Exchange Network formats
Toxicity Identification Evaluation		
TIE Trigger	Create TIE work plan agreeable to both regulated and regulatory parties	≥50% effect on original sample; First storm of season
Treatments	Baseline, SPE, EDTA/cation exchange column, PBO	Other treatments and/or dilutions
Species selection	Same as original test species	-
Test methods	Follow EPA guidance	Modify volume or renewals when sample is limited
Quality assurance	Follow EPA guidance, increase blanks for TIE treatments	

TABLE OF CONTENTS

Acknowledgements.....27

List of Abbreviations.....28

Executive Summary29

Background.....32

 Objective of this Document32

Methods33

Monitoring Questions33

Sample Guidance34

 Sample Location.....34

 Frequency of Sampling35

 Storm Trigger35

 Storm Sampling.....37

 Container Type.....39

Methodological Guidance39

 Species Selection.....39

 Holding Time40

 Test Renewals.....41

 Test Acceptability Criteria.....41

 Review of Test Conditions.....41

 Review of Reference Toxicants.....42

Data and Information Guidance43

Toxicity Identification Guidance44

 TIE Triggers and Work Plan44

 Study Design45

 Test Species and Exposure Methods.....48

 Interpretation of Results49

 Identification and Confirmation of Toxicant50

Appendix 1 – Frequently Asked Questions.....57

Appendix 2 – Storm water Sampling and Testing Monitoring Inventory Questionnaire61

Appendix 3 – Flow Chart for Storm water Sampling Activities.....62

List of Tables

Table ES1. List of default recommendations for monitoring toxicity in storm water.....29

Table 1. Type of sampling location appropriate for each monitoring question.....35

Table 2. Considerations for grab versus composite sampling for storm events38

Table 3. TIE characterization treatments.47

Table 4. Expected response of selected TIE treatments to common storm water toxicants. ...48

Table 5. Example TIE results for dry weather flow sample tested using *Ceriodaphnia dubia*. ..48

Table 6. Effect level of typical storm water constituents for different test species.51

List of Figures

Figure 1. Long-term rainfall duration curves for three major metropolitan centers in California.36

Figure 2. Storm water toxicity of highway runoff to fathead minnows and *Ceriodaphnia dubia*.38

Figure 3. Storm water toxicity identification evaluation sequence.46

Figure 4. Example TIE results for storm water sample tested using the fathead minnow.50

BACKGROUND

Toxicity testing has been a part of National Pollutant Discharge Elimination System (NPDES) permit monitoring and reporting programs since the 1970's (Heber et al. 1996). Toxicity testing provides several advantages over chemical measurements alone. For example, toxicity tests can capture effects of unmeasured chemicals and variability in bioavailability. It is nearly impossible to measure all potential toxicants in a discharge, and to exacerbate the problem, hundreds of new chemicals are developed and potentially discharged into the environment each year (Muir and Howard 2006). Even if all the potential toxicants in a discharge could be chemically analyzed, such analysis would be expensive and would be inherently limited because they would have to be evaluated on a one-by-one basis. Yet, scientists know that many toxicants can interact to create synergistic or antagonistic effects on test organisms (Loureiro et al. 2010). Toxicity testing in NPDES monitoring and reporting programs has such value that there are now over a dozen different standardized toxicity test methods including freshwater, estuarine, and marine species measuring both lethal and sublethal endpoints.

Historically, point source facilities such as publicly owned treatment works (POTWs) and industrial facilities were the focus of toxicity tests for NPDES permit programs. Point source facilities are well-suited to toxicity testing because effluent flow and quality remains somewhat steady, changing little over time unless alterations in the treatment process occur (Lyon et al. 2005). These relatively static effluent conditions enable repeat testing for confirmation and follow up testing to identify and confirm the responsible toxicants. As such, standard permit language has evolved for point source facilities to include testing, toxicity identification evaluations (TIEs), and toxicity reduction evaluations (TREs) that include management actions responding to failed toxicity tests (see permit sample language in Denton et al. 2010).

Aquatic toxicity testing was first used in storm water NPDES permits in California during the 1990's (Katznelson and Mumley 1997; Skinner et al. 1998; Bailey et al. 2000; Fong et al. 2000; Larsen et al. 2000; SRWP 2000; Larsen and List 2002). The NPDES toxicity testing focused on waters receiving discharges from Phase I municipal separate storm sewer systems (MS4).

Unlike traditional point sources, toxicity tests are not applied, nor are the resulting information utilized, in a similar fashion among different MS4 permittees in California. Data interpretation can range from purely observational information to the primary data source used for placing a waterbody on the State's list of impaired waterbodies, which mandates regulatory oversight including a total maximum daily load (TMDL).

One reason that aquatic toxicity tests of urban runoff are applied less uniformly than point sources is due to runoff's unique challenges. Unpredictability in flow and water quality, particularly those associated with storms, makes sampling difficult. Runoff flows and contaminant concentrations can change orders of magnitude in less than an hour (Tiefenthaler et al. 2008). Moreover, the sources of toxicants in runoff are more diffuse than in point sources, making identifying and controlling toxicants more challenging.

Objective of this Document

The objective of this document is to provide implementation guidance to the State and Regional Water Boards on monitoring toxicity in storm water discharges. The primary audience for this implementation guidance is the Regional Water Board staff that must write and interpret toxicity

testing language in MS4 permits, and regulated stakeholders that must implement toxicity testing in their MS4 NPDES permits.

METHODS

Four major topics were addressed for implementing a storm water monitoring program: 1) sampling; 2) testing; 3) information management; and TIEs. There are several issues within each of these topic areas. Recommendations for these topics were developed in the context of the objectives of a monitoring program, expressed in terms of five key monitoring questions. The recommendations are not binding and not meant to be one size fits all. Instead, the recommendations are meant to serve as a starting point for developing a monitoring and reporting program. These recommendations are default options and either regulators or regulated may decide to do more (or less) depending on site-specific needs.

Three additional elements were provided to assist with implementation guidance. The first was a list of frequently asked questions (Appendix A), adapted from EPA guidance (Denton et al. 2010). The second was an unofficial survey of toxicity requirements conducted in March 2010 (Appendix B). A questionnaire was sent to 11 MS4 programs with 100 percent response. These programs represent a sampling of the majority of MS4 NPDES permits in California including the Phase I permits. The survey consisted of 28 questions covering each of the 4 topics. These results are referred to frequently throughout the document to give perspectives as to the level of effort, and level of comparability in effort, for existing monitoring and reporting programs in the state. The third was a flow chart of sampling activities (Appendix C), adapted from California Department of Transportation monitoring program (Caltrans 2009).

MONITORING QUESTIONS

Monitoring questions are a paramount element of a functional monitoring program. The questions drive all of the study design elements including what, where, when, and how sampling or laboratory measurements are to be made. Therefore, it is critical that any recommendations for monitoring be within the context of the question it is trying to answer. For this document, five questions appropriate to MS4 monitoring programs were addressed:

- 1) Do aquatic toxicity test organisms not respond to urban runoff?

This question is the most fundamental. Loosely translated, this question attempts to identify for regulators and regulated agencies if a potential for toxicity exists. If an aquatic toxicity test does not respond to urban runoff, then the intent of the toxicity Policy perseveres; there is no environmental problem. If, however, the toxicity test organism does respond to urban runoff exposure, then it is incumbent that environmental managers should take further steps to identify the scope of the potential problem. These next steps are expressed in questions two through five.

- 2) What is the temporal or spatial extent of toxic response by aquatic organisms?

Aquatic toxicity tests are not always conclusive. Particularly in situations such as wet weather urban runoff, where flows and concentrations fluctuate dramatically, one test

does not tell the whole story. When the first question does indicate an effect, managers should identify the scope of the toxicity. Before taking management actions, it will be important to ascertain if the toxicity was an isolated incident, or a consistent problem at this site. Similarly, assessing if toxicity occurs at many sites will help determine the level and direction of management actions necessary to resolve the toxicity.

3) What are the causes of aquatic toxicity in urban runoff?

When aquatic toxicity is pervasive, either over time at a site or across many sites, identifying the responsible toxicant(s) should become a priority. This detective work is typically conducted using a TIE. Successful TIEs include three phases: characterization, identification, and confirmation. It is the results of the TIEs that will most effectively focus managers for appropriate control actions.

4) What are the sources of aquatic toxicity in urban runoff?

Once the responsible toxicant(s) are identified, managers will need to identify where in the watershed the toxicant(s) originate. Sometimes the toxicant(s) arise from single locations while at other times toxicant(s) are diffuse and sources are difficult to pinpoint. The goal of this question is to determine the origins(s) of the toxicant so remedial action can occur.

5) Is the magnitude or extent of aquatic toxicity in urban runoff changing over time?

This question assesses trends. Typically, once a management action is undertaken, there is a desire to track progress to ensure the problem has been resolved. For both regulatory and regulated agencies, this question is critical to demonstrating successful compliance.

These questions are referred to frequently throughout this document. Technical decisions about recommended approaches are always developed in context of the question that is to be answered. These are not the only questions that could or should be asked. Based on experience, many Regional Water Boards ask additional questions that could lead to differing monitoring design recommendations. However, these questions, at a minimum, should be answered by an NPDES monitoring and reporting program.

SAMPLE GUIDANCE

There were five primary sampling issues that require guidance. These issues included station location, sampling frequency, storm trigger, collection method, and sample containers.

Sample Location

Two types of sampling locations exist; integrator sites and targeted sites. Integrator sites sample across an array of different sources and/or land uses. Integrator sites are often located at the bottom of a catchment or watershed, thus capturing many sources of pollutants found in urban runoff. Integrator sites are good for screening if a toxicity problem exists in a watershed or for determining if commingled toxicants from multiple sources result in toxic responses. A good example of integrator sites would be mass emission sites used for MS4 permit monitoring

programs. In contrast, targeted sites are often located immediately downstream of a discrete source to determine if discharges from that source are important contributors to toxicity. Targeted sites are valuable because they may capture only a single source for characterization and assessment. The types of sites used for construction or industrial storm water permit monitoring programs would be a good example of a targeted site. Both the integrator and targeted sites have their appropriate application depending on the question to be answered (Table 1). Integrator sites are the default recommendation for MS4 monitoring programs.

Table 1. Type of sampling location appropriate for each monitoring question.

<i>Question</i>	<i>Integrator Site</i>	<i>Targeted Site</i>
Is runoff toxic to aquatic life?	X	
What is the extent of aquatic toxicity?	X	
What is the cause of toxicity	X	X
What is the source of toxicity?		X
Is the aquatic toxicity in runoff changing?	X	X

Frequency of Sampling

Frequency of sampling is an important aspect of study design because the number of sampling events can be a large component of monitoring costs. The current draft toxicity policy recommends a default frequency of two wet weather events and two dry weather events. This frequency is similar to the frequency currently used for many MS4 programs in California. For example, 63 percent of the MS4 programs in our survey currently sample two storms per year. Specific recommendations for frequencies other than the default are a function of statistical power dependent on the question of interest. For example, the sample size (i.e. frequency) for the necessary statistical power to detect trends in Question 5 is a function of time (duration of trend), magnitude (amount of change), and variability (variance independent of trend such as storm to storm variability). If a frequency other than the default frequency is desired, then conducting power analysis is crucial to help guide optimal sample size.

Storm Trigger

Storm triggers are perhaps one of the most significant elements of storm water sampling. Weather forecasting is intrinsically challenging. Mobilizing for a storm that never materializes can be an expensive waste of effort. Don't mobilize for a storm that does materialize and you may have to wait another year to capture a similar event. Based on our survey of existing MS4 storm water monitoring programs, all of the respondents stated that storm triggers were based on quantity of predicted rainfall and that no program used a trigger greater than 0.25 inches. In addition, all of the monitoring agencies used a trigger of between three and seven antecedent dry days between storm events. Long-term rainfall duration curves from major metropolitan areas in California were compared to ascertain if variations in climate should factor into storm trigger recommendations (Figure 1). The median rainfall event in Los Angeles, San Francisco, and Sacramento were all approximately 0.25 inches.

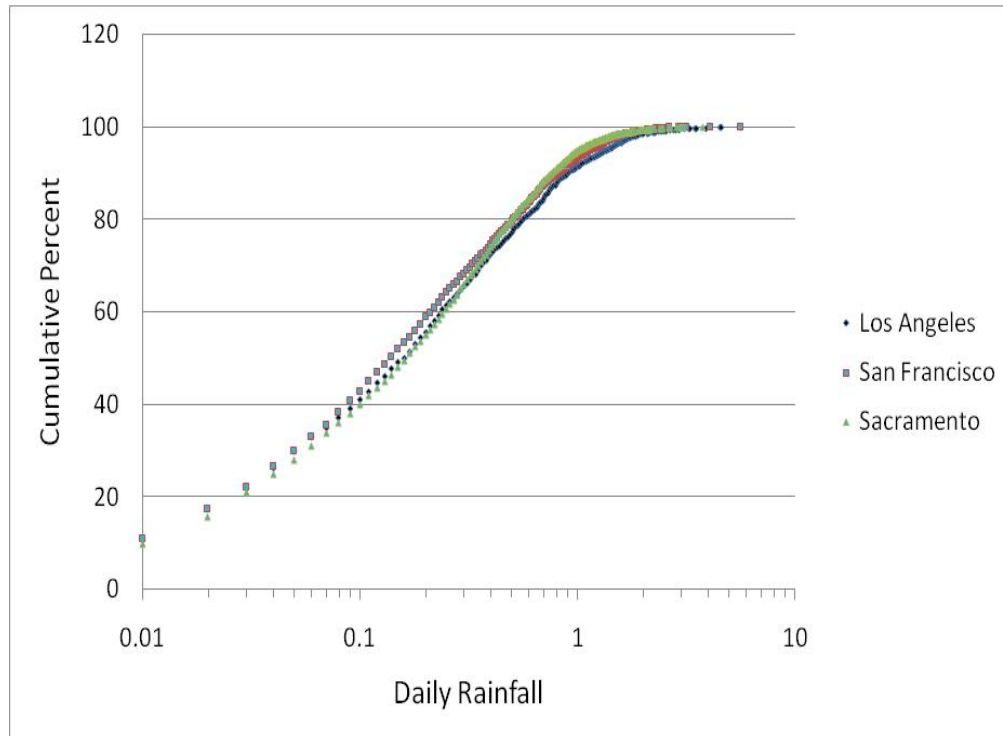


Figure 1. Long-term rainfall duration curves for three major metropolitan centers in California. Data covers a minimum of 65 years.

A rainfall forecast of ≥ 0.25 inches/24 hours with at least a 50 percent probability of precipitation by the National Weather Service is recommended as the default storm trigger. This trigger should be preceded by a minimum of 72 hours of no rain. This trigger is consistent with most monitoring programs and, as the median storm event across multiple major California municipalities, strikes a balance between storms that are not too small or too large. If a storm is collected, but the rainfall trigger is not surpassed, the sample(s) should be analyzed if ≥ 0.10 inch/24 hours of rainfall fell and there was an increase over baseline flow. It is preferable to resample storms < 0.1 inches/24 hours. Regulated and regulatory agencies should be given the option to change the rainfall trigger by selecting the median storm event for the long-term historical record at their locality. Another option would be the consideration of the antecedent moisture condition (AMC). The AMC, which is a function of soil infiltration capacity and time since previous rainfall, can be used to alter storm triggers (Heggen 2001; Sahu

Storm Water Sampling

Storm sampling is quite unlike sampling continuous discharges such as POTWs. Storms are unpredictable so sampling preparations require more attention and organization. Although only two storm samples are required each year, preparation and readiness may last six months. In addition, weather forecasts in two storms may be required in order to obtain an imperfect science and, because storm events may not materialize or equipment may not work properly, mobilization for more than sufficient samples for analysis. In order to assist storm water agencies and Regional Water Board's, an example flow chart has been provided in Appendix C to assist those unfamiliar with storm sampling.

et al. 2010).

Storm Sampling

There are two generic methods for sampling storm water; discrete sampling or composite sampling (Table 2). Discrete (sometimes called “grab”) sampling, usually collected by dipping a container in flowing water, is used to collect instantaneous, single samples of various volumes. Discrete sampling is also the most common of the sampling methods used in MS4 programs. Approximately two-thirds of regulated agencies employed discrete sampling based on our survey. However, because of the variability inherent in storm water concentrations, variation in toxicity data from sample-to-sample or storm-to-storm can be extreme (Figure 2). Composite samples, such as flow weighted composite samples, provide a much better estimate of a storm’s central tendency and, hence, less variation than discrete samples. However, composite sampling can require specialized equipment such as flow meters, peristaltic pumps, and electronic data storage devices. As a result, costs can increase significantly, especially if one or only a few storms are to be collected at that site (per sample costs decrease over time as the efficiencies of automated composite samplers are accrued across multiple events). Composite sampling may also have its logistical challenges if large volumes are required. The largest commercially available composite sample bottles are 20 L. While obtaining composite sample volumes greater than 20 L are possible, the challenges and costs increase at a greater rate than it would by discrete sampling alone. One important consideration is leveraging other monitoring data. In this instance, most programs collect composite samples for chemical analysis. If the intent is to compare the toxicity information to the chemistry information, then splitting the composite sample for chemistry and toxicity is the preferred alternative. Finally, toxicant loss in grab samples is presumed to be less than composite samples because of the instantaneous nature of discrete sampling. Direct filling of the sample bottle reduces the likelihood of contaminants volatilizing, sticking to sampling surfaces, or being cross-contaminated than a sample pumped through composite tubing, sometimes for long distances.

While discrete sampling and composite sampling are both acceptable, the recommended sampling method is a function of maximizing sampling criteria for the monitoring question you want to answer. If the objective is a simple screening to determine whether or not runoff is toxic to aquatic life, then discrete sampling is likely sufficient. Similarly, determining the causes and sources of toxicity can also be answered using discrete samples, especially if the most toxic times and/or locations are targeted. However, sampling to determine if toxicity is changing over time is best sampled using composites because variability is reduced and trends are more likely to be observed with fewer samples. Regardless of discrete or composite sampling, if toxicity and chemistry are to be compared, they should be subsampled from the same bottle.

Table 2. Considerations for grab versus composite sampling for storm events based on different decision sampling criteria.

Sampling Criteria	Discrete Sample	Composite Sample
Data Variability	Larger within and between storms	Smaller within and between storms
Ability to sample large volumes	Easier	Harder after 20 L
Cost	Lower cost	Higher cost
Ability to compare with chemistry	Split sample necessary	Split sample necessary
Sample Integrity	Better	Not bad

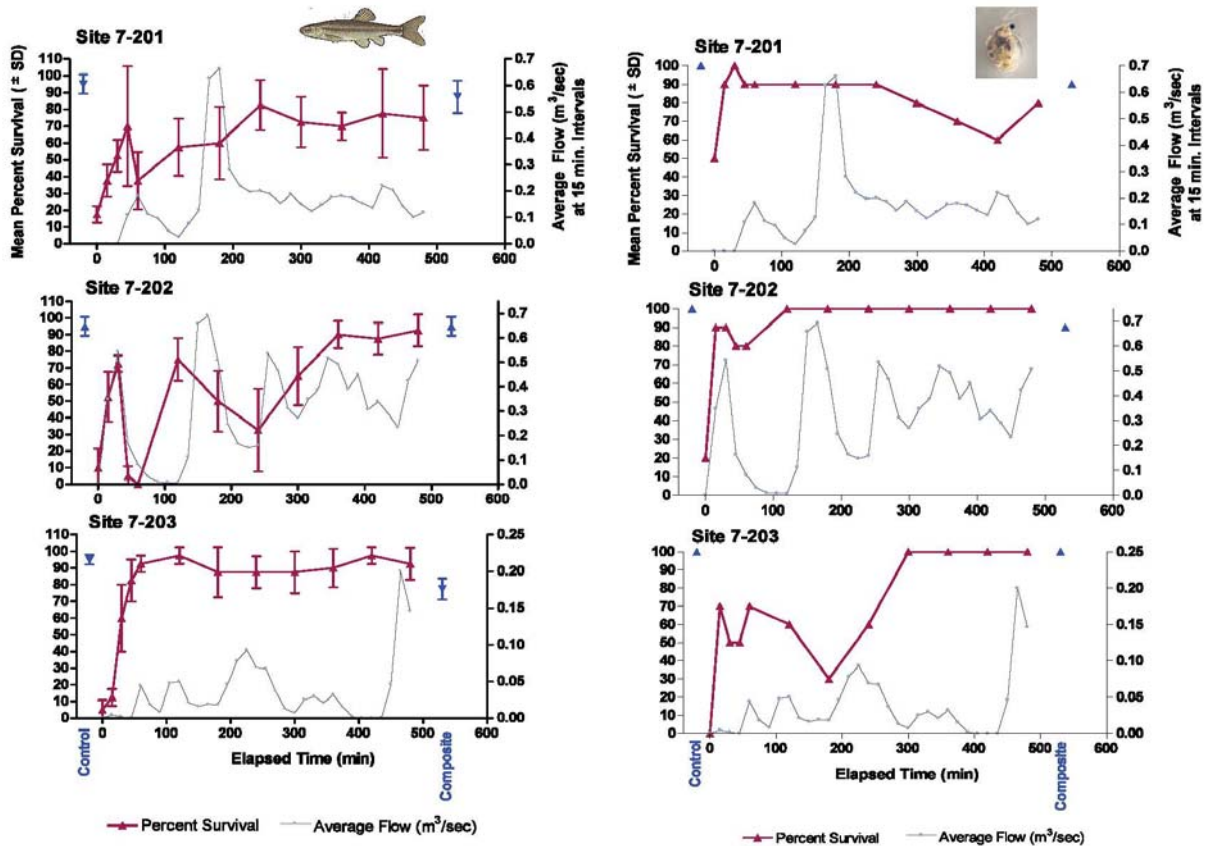


Figure 2. Storm water toxicity of highway runoff to fathead minnows and *Ceriodaphnia dubia* over time (Kayhanian et al. 2008). Eight discrete samples (in red) were collected and tested over the first eight hours of runoff. An equal weight composite was also prepared and tested for comparison (in blue following discrete sample results).

Container Type

The selection of container type is dependent upon the type of toxicant one is concerned about. The general rule is that glass containers should be used for hydrophobic contaminants, such as chlorinated organic compounds (i.e. Dichlorodiphenyl Trichlorethane (DDT), Polychlorinated Biphenyls (PCBs), insecticides (i.e. pyrethroids, organophosphorus), and herbicides (i.e. glyphosate). These compounds will sorb to sampling container walls made of organic materials such as plastic. In these cases, amber or dark colored glass is preferred to minimize photodegradation. The general rule also applies to hydrophilic contaminants, where plastic containers are recommended. In this case, trace metals (i.e. copper, zinc) or polar organic compounds (i.e. some surfactants) may stick to the surfaces of sample containers with charged surfaces. If the toxicant(s) is unknown, the default container type is glass. Containers should be filled to the neck to minimize headspace, which will reduce volatilization of aromatic components. All containers should have Teflon-lined lids and all containers should be pre-cleaned according to United States Environmental Protection Agency (EPA) guidance to avoid cross-contamination. Certified pre-cleaned glassware can be purchased from most suppliers. Additional techniques to help minimize loss of sample integrity especially when testing with hydrophobic compounds like pyrethroids (USGS 2009). These sampling and containers recommendations include: A) container composition affects the extent of aqueous pyrethroids loss: pyrethroids associate less to glass containers and plastic, and Telfon has the greatest pyrethroids loss caused by association to container surface, B) containers should be agitated vigorously for at least one minute before transfer to another sample container; C) maximize the volume to surface contact area ratio; and D) when using a filtration apparatus or autosampler, pump speeds should be at 500 ml/min. These are techniques to be implemented when handling compounds such as pyrethroids.

METHODOLOGICAL GUIDANCE

There are four primary methodological issues relevant to implementation. These issues include species selection, holding time, renewals, and test acceptability criteria (TAC). Most, but not all, of these issues are independent of the aforementioned monitoring questions.

Species Selection

Current strategies for selecting species are discussed in the toxicity policy. In brief, the requirement is to test multiple species from multiple phyla (i.e. one alga, one invertebrate, one vertebrate) to ascertain the most sensitive species. There are multiple factors to consider when applying this strategy to storm water testing including frequency, monitoring design and confounding factors.

Currently, POTWs conduct most sensitive species screenings approximately once per permit cycle. This frequency assumes that POTW effluent quality stays relatively homogeneous over time. In contrast, storm water discharges are far from homogeneous, with water quality potentially changing both within and between storm events. Therefore, options for selecting test species should be considered.

The recommendation for conducting most sensitive species screenings for storm water is dependent upon prior knowledge at a site. If there is no previous knowledge at a site, then the recommendation is to screen storm water samples with at least three species (a fish, an invertebrate, and a plant) for chronic testing and two species (a fish and an invertebrate) for acute testing (Denton et al. 2010). There are no acute test methods with plant species. This recommendation is based upon the fact that there are species sensitivity differences among different groups of organisms to different toxicants.

If many previous tests have been conducted at that site, then the recommendation is modified. If a sensitive species has already been determined, this information should be used to select test species. As a corollary to this recommendation, if there is prior knowledge of the potential toxicant(s) at a site, then selecting test species using known sensitivity to that toxicant(s) is recommended. For example, it is known that *Ceriodaphnia dubia* is a highly sensitive freshwater species to pesticides such as diazinon (see Table 5 for additional examples). If diazinon is the suspected contaminant of concern, then the monitoring program should select *C. dubia*.

Another consideration for species selection is monitoring design. For monitoring questions one through four that focus on determining whether or not toxicity exists and the toxicants of concern, varying test organism selection to target the most sensitive species is appropriate. However, if the monitoring question addresses trends (i.e. Question 5), then utilizing a common species over time is the more appropriate selection.

Storm water programs should pay particular attention to potential confounding factors when selecting test species. For example, high conductivity levels can occur naturally but will confound *C. dubia* tests. Conductivity will confound *C. dubia* tests because this species does not reproduce well, even in control exposures, at conductivity levels above 2,500 μmhos . Consistent with State recommendations (SWAMP 2008), *C. dubia* should be replaced with *Hyaella azteca* water phase tests above 2,500 μmhos conductivity.

Holding Time

Holding time is an important element of sample handling because contaminants can degrade prior to laboratory testing. The objective is to minimize this degradation to the maximum extent possible. This can be accomplished following these steps: 1) keep containers sealed to prevent volatilization and contamination; 2) keep containers cold ($<6^{\circ}\text{C}$) to prevent metabolic or thermal breakdown; and 3) keep samples in the dark to prevent photodegradation. Finally, samples should be tested within 36 hours of sampling. The 36-hour requirement is consistent with EPA test methods (USEPA 1995, 2002a, 2002b, 2002c) and is consistent across all of the MS4 programs in the state according to our survey. For composite samples, holding time begins after the last sample interval. The EPA has allowed exceptions to the 36-hour holding time, for example, when effluents are shipped overseas for testing (Denton et al. 2010). The primary reason for an extension of the holding time would be the consideration of the sampling and laboratory technicians safety (Burton and Pitt 2001; see page 255), and logistics of coordinating collection and transport of multiple samples within a short period. Since storm events are not pre-determined and typically are occurring rapidly throughout a watershed, many site samples must be coordinated with short notification. The 36-hour holding time for test initiation should

be targeted, but no more than 72 hours should elapse before initial use of a sample. This should all be followed up with personal communication to the appropriate regulatory agency for approval.

Test Renewals

Renewal of test chambers, where a portion of the exposure media is removed and replaced with fresh exposure media is a common laboratory practice documented in standard methods (USEPA 1995, 2002a, 2002b, 2002c). These renewals are necessary, especially in relatively long exposures (>4 days), to reduce the buildup of organism waste or uneaten food, and minimize associated confounding test factors.

Test sample renewals are a unique challenge for toxicity testing with storm water. Unlike POTW testing where renewal samples can be collected in the days following test initiation, most storms last less than a day and follow up samples for renewal would not be representative of the actual storm event. It is recommended that storm water monitoring programs conduct renewals, but with original storm samples. Renewals should occur no more infrequently than every 48 hours. This guidance is consistent with the existing monitoring programs based on our MS4 sampling and testing survey as well as standard practice for the State's Surface Water Ambient Monitoring Program (SWAMP 2008).

Test Acceptability Criteria

Test acceptability criteria set minimum requirements for performing toxicity tests. These minimum requirements are clearly identified in the EPA test method manuals. Both storm water and reference toxicant tests must meet these TAC. For example, the control for both the effluent test and the reference toxicant test must achieve 80 percent or greater survival and produce an average of 15 young per female for the chronic *C. dubia* survival and reproduction test method. These requirements are stated in the summary of test conditions and TAC table in each chapter for the test method manuals. Both the regulated and the regulatory authority should be familiar with these summary test conditions and TAC. Test data are reviewed to verify that TAC requirements for a valid test have been met. *Any test not meeting the minimum TAC is considered invalid.*

Review of Test Conditions

Test conditions should be reviewed and compared to the specifications listed in the summary of test condition tables provided for each method. Physical and chemical measurements taken during the test (e.g. temperature, pH, and dissolved oxygen) also are reviewed and compared to specified ranges. Any deviations from specifications should be documented and described in the data report.

The summary of test condition tables presented for each method identifies test conditions as required or recommended. For toxicity test data submitted under NPDES permits, all "required" test conditions must be met or the test is considered invalid and must be repeated with a newly collected sample. Deviations from "recommended" test conditions must be evaluated on a case-by-case basis to determine the validity of test results. Deviations from recommended test

conditions may or may not invalidate a test result depending on the degree of the departure and the objective of the test. The reviewer should consider the degree of the deviation and the potential or observed impact of the deviation on the test result before rejecting or accepting a test result as valid. For example, if dissolved oxygen is measured below 4.0 mg/L in one test chamber, the reviewer should consider whether any observed mortality in that test chamber corresponded with the drop in dissolved oxygen.

An individual test may be conditionally acceptable if temperature, dissolved oxygen and other specified conditions fall outside specifications, depending on the degree of the departure and the objectives of the tests (see test conditions and TAC specified for each test method). The acceptability of the test will depend on the experience and professional judgment of the laboratory investigator and the regulatory authority (see section on data evaluation in the test method manuals). Whereas slight deviations in test conditions may not invalidate an individual test result, test condition deviations that continue to occur frequently in a given laboratory may indicate the need for improved quality control in that laboratory.

Review of Reference Toxicants

The purpose of generating reference toxicant data is to: 1) assess the health and sensitivity of test organisms over time; and 2) document and demonstrate initially and ongoing acceptable laboratory performance. Satisfactory laboratory performance is demonstrated by performing at least one acceptable test per month with a reference toxicant for each toxicity test method conducted in the laboratory during a month. *For a given test method, successive tests must be performed with the same reference toxicant, at the same concentrations in the same dilution water, using the same data analysis methods.* Regardless of the source of test organisms (in-house cultures or purchased from external suppliers), the testing laboratory must perform at least one acceptable reference toxicant test per month for each type of toxicity test method conducted in that month (USEPA 2002a, 2002b). If a test method is conducted only monthly, or less frequently, a reference toxicant test must be performed concurrently with each effluent or storm water toxicity test. This requirement will document ongoing laboratory performance and assess organism sensitivity and consistency when organisms are cultured in-house. When organisms are obtained from external suppliers, concurrent reference toxicant test must be performed with each effluent sample, unless the test organism supplier provides control chart data from at least the last five months of reference toxicant testing. This requirement assesses organism sensitivity and health when organisms are obtained from external vendors.

The test review of a given effluent or receiving water should include review of the associated reference toxicant test and current control chart. The test reviewer should verify that a quality control reference toxicant test was conducted according to the specified frequency required by the regulatory authority or recommended by the method. The TAC, test conditions, concentration-response relationship, and test sensitivity of the reference toxicant tests are reviewed to verify that the reference toxicant tests conducted were valid. The results of the reference toxicant tests are then plotted on a control chart and compared to the current control chart limits. Reference toxicant tests that fall outside of the recommended control chart limits are evaluated to determine the validity of associated effluent and receiving water tests (see chapter on quality assurance of test method manuals). Reference toxicant tests should not be used as a *de facto* criterion for rejection of individual effluent or receiving water tests. An out of

control reference toxicant test does not necessarily invalidate the associated test results. The reviewer should consider the degree to which the reference toxicant test fell outside of the control chart limits, the width of the limits, the direction of the deviation (toward increasing test organism sensitivity or toward decreasing test organism sensitivity), the test conditions of both the effluent and the reference toxicant tests, and the objective of the test. More frequent and/or concurrent reference toxicant testing may be advantageous if recent problems (e.g., invalid tests, reference toxicant test results outside of control chart limits, reduced health of organism cultures, or increased within-test variability) have been identified in testing.

DATA AND INFORMATION GUIDANCE

Data and information management has two primary issues; required data for submittal and data analysis. At present, there is no standard for data submittals to the State Water Board. However, the consensus opinion of the Committee is that the default minimum data submittal should include:

- Unique Station Identifier
- Sample Date/Time
- Storm or Nonstorm Sample
- Discrete or Composite Sample
- Organism Name
- Test Duration
- Test Concentration
- Number of Replicates
- Mean Result Control
- Standard Deviation Control
- Mean Result Test Exposure
- Standard Deviation Test Exposure
- Units
- Test of Significant Toxicity (TST) (pass/fail)
- Any Quality Assurance Qualifiers
- TIE Follow-up (yes/no)
- Contact Information for Testing (including TIE follow-up)
- Comments

The State Water Board is currently developing a platform for electronic submittal of monitoring data. This platform, called the California Environmental Data Exchange Network (CEDEN), will include toxicity data (www.CEDEN.org). The purpose of the CEDEN network is to allow the exchange and integration of water and environmental data between groups and to make it accessible to the public. CEDEN is a system designed to facilitate integration and sharing of data collected by many different participants including state, regional, local, and private information. As such, CEDEN has specific templates for data entry, including toxicity information. The philosophy of CEDEN is to capture raw data and not calculated values. This means that toxicity data requirements will include results at the individual replicate level and be inclusive of all

quality assurance information (i.e. water quality results, batch process samples, etc.). While CEDEN is not a requirement for NPDES permittees at this point in time, it is assumed that this system, or something like it, will become a requirement in the future. While the level of data sharing is much greater, access and utility of the data will be vastly increased.

TOXICITY IDENTIFICATION GUIDANCE

Consistent with our third monitoring question to identify specific toxicants, a Toxicity Identification Evaluation (TIE) should be considered when toxicity testing demonstrates that the storm water or receiving water is toxic and exceeds triggers established by the regulatory agency. The pass/fail statistical endpoint derived from the TST (Denton et al. 2011) must be utilized to determine whether a sample is toxic. A TIE is an investigative process that uses laboratory modifications of test sample chemistry and resulting changes in toxicity to identify the likely causes of toxicity. This section describes the technical factors to be considered in conducting TIEs for storm water monitoring programs. Other aspects, including identifying and removing the sources of the toxicant(s), are important management actions that can follow from a TIE and should be developed by negotiation and consultation between stakeholder(s) and regulatory agencies.

The EPA and others have published extensive TIE technical guidance (see Study Design). In addition, numerous TIE research papers and case studies have been published which demonstrate the efficacy of the TIE process in identifying the cause(s) of toxicity. Most TIE method development and application has been applied to continuous point sources, such as POTWs and industrial discharges. Storm water differs fundamentally from such point sources in that its composition may be highly variable over time, obtaining sufficient sample volume is often difficult, and sample availability is limited by unpredictable weather events. The purpose of this section is to provide guidance on how to address such challenges, rather than a comprehensive description of existing TIE methods and literature.

TIE Triggers and Work Plan

The decision to conduct a TIE is based upon consideration of multiple factors such as the magnitude and persistence of toxicity. The magnitude of toxicity present in the storm water is an important consideration because a moderate to high level of toxicity typically yield more successful results. Usually, TIEs can be successfully conducted on samples producing at least a 50 percent effect (e.g., $\geq 50\%$ mortality or reduction in reproduction), and this value is recommended for general use in selecting samples for TIEs. However, effective TIEs can also be conducted with less toxic samples (e.g., $\geq 25\%$ effect), but there is a greater chance of the TIE being inconclusive due to changes in toxicity with storage or variability in response (Norberg-King et al. 2005).

Toxicity persistence refers to the detection of toxicity over multiple sampling events. In contrast to continuous point source discharges, where accelerated testing can be used to confirm the persistence of toxicity before initiating a toxicity reduction evaluation, unpredictable storm events and temporal variations in storm water quality provide few opportunities per year to investigate persistence and conduct TIEs. Once a decision has been made to conduct a TIE at a site, the monitoring agency should be prepared to conduct a TIE on the next available sample

that meets the selection criteria established in the TIE work plan. Storm water toxicity is often greatest in the first storms of the season, so early season storms should be targeted for TIEs as they will likely provide the best opportunities for a successful study. The default recommendation is that prior to starting a TIE study, the permittee should develop a TIE work plan in consultation with the regulatory authority. Establishment of such a plan will encourage the permittee and regulatory authority to address important design issues up front and thus will save time in initiating a study should the need arise to conduct a TIE. Decisions regarding whether and how to conduct a TIE must often be made on short notice in the midst of a monitoring event; a work plan that describes procedures and contingency plans is recommended to increase the potential for success. The TIE work plan should include the following elements:

- Criteria for initiating a TIE on a sample
- Roles and responsibilities of the TIE team
- Study design, sample treatments, and chemical analysis
- Data evaluation and communication
- Follow up actions

Study Design

The EPA has developed TIE procedures to determine the causes of acute and chronic toxicity to freshwater species (USEPA 1989a, 1989b, 1991, 1992, 1993a, 1993b, 2001) and to estuarine/marine organisms (USEPA 1996). A generic TIE consists of three phases: toxicity characterization (Phase I), toxicant identification (Phase II), and toxicant confirmation (Phase III). An overview of the TIE process is shown in Figure 3. These three phases are often conducted sequentially when investigating continuous point source discharges with stable effluent characteristics. The sporadic and variable nature of storm water may preclude a sequential approach, and it may be more effective for the storm water TIE design to incorporate greater flexibility and to combine elements of the characterization, identification, and confirmation phases to provide more information from each sample analyzed.

Many Phase I characterization treatments have been applied in TIEs (Table 3). These treatments are designed to classify the cause(s) of toxicity into several broad categories: metals, organics, ammonia, oxidants, particulates, and selected pesticides (organophosphates and pyrethroids). Many of these treatments are compatible with most commonly used toxicity test methods; however, pH manipulations are of limited utility in tests with marine species due to their limited tolerance of pH variation.

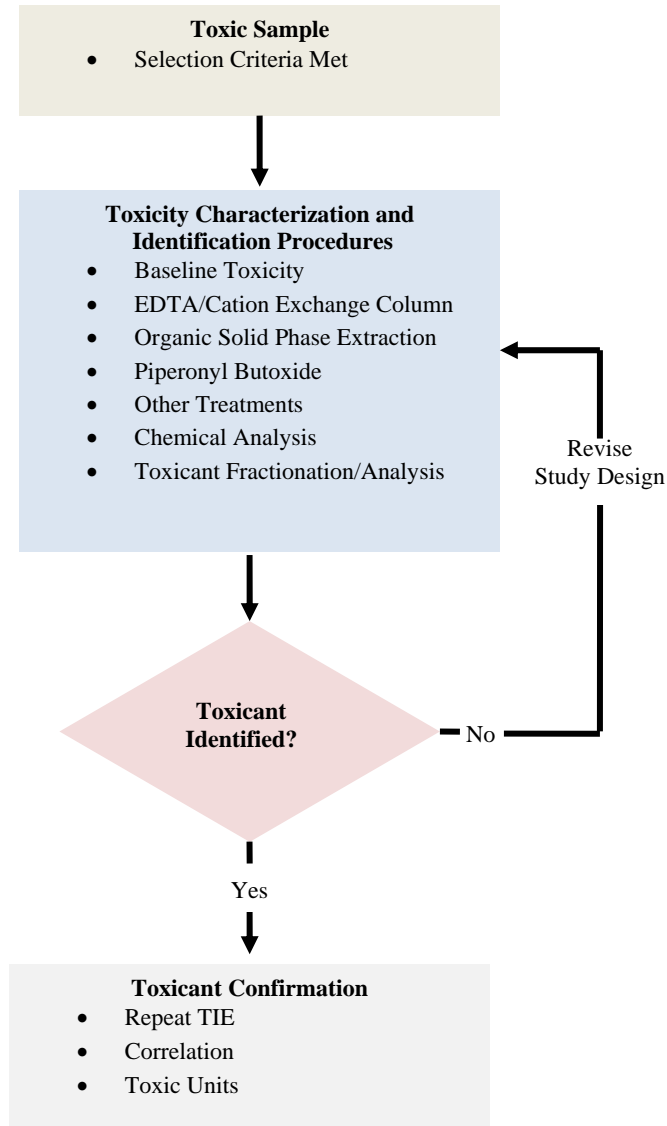


Figure 3. Storm water toxicity identification evaluation sequence. Suggested treatments and activities for each step are shown; actual treatments/activities should be determined based on program objectives and watershed characteristics.

Table 3. TIE characterization treatments. Highlighted treatments (shaded) are recommended for initial storm water testing. Additional treatments may be included depending on program objectives and previous knowledge.

<i>TIE Treatment</i>	<i>Treatment Identifies</i>
Initial toxicity (unaltered effluent)	Initial toxicity test demonstrating toxicity of sample
Baseline toxicity (unaltered effluent tested during TIE)	Results compared to TIE manipulations to assess effectiveness of TIE manipulations
Centrifugation/filtration	Particulate-bound toxicants
pH adjustment/filtration (pH 3 and pH 11) freshwater only	Particulate-bound toxicants
pH adjustment/aeration (pH 3 and pH 11) freshwater only	Ammonia and volatile, oxidizable toxicant
C-18 Solid-phase extraction (SPE) using columns	Polar and non-polar organic chemicals
pH adjustment/SPE (freshwater only)	Polar and non-polar organic chemicals
Sodium thiosulfate addition	Oxidants and some cationic metals
Ethylene diaminetetraacetic acid (EDTA) addition or Cation SPE exchange column	Cationic metals
Zeolite	Ammonia
Graduated pH adjustments (freshwater only)	Ammonia and pH-sensitive toxicants
Piperonyl butoxide (PBO)	Organophosphate and pyrethroid insecticides
Carboxylesterase enzyme (CEE)	Organophosphate and pyrethroid insecticides

Application of all possible characterization treatments is not recommended for storm water, as the sample volume requirements and cost could be excessive. Alternatively, it is recommended that a core suite of selected treatments be considered as the foundation for initial studies in a general storm water TIE. Most previous storm water TIEs have associated toxicity with relatively few types of toxicants: trace metals (copper and zinc), pesticides (organophosphates and pyrethroids), ammonia, and dissolved solids (i.e. water hardness). A reduced suite of treatments will be effective in determining whether these same constituents are the likely toxicants for the sample under investigation. Depending upon the outcome of the initial TIE tests, the treatments used in subsequent tests can be modified to either confirm or provide greater specificity to the results. The suggested initial TIE design is shown by highlighted rows in Table 3. In addition to these treatments, measurement of the sample for ammonia concentration and hardness will enable discrimination of the most likely contaminants, as shown in Table 4.

Table 4. Expected response of selected TIE treatments to common storm water toxicants. Symbols indicate an increase (↑), decrease (↓), or no change (—) in toxicity relative to the baseline sample. Testing of multiple dilutions is recommended when feasible.

<i>Treatment</i>	<i>Metals</i>	<i>OP Pesticides</i>	<i>Pyrethroid Pesticides</i>	<i>Ammonia</i>	<i>Dissolved Solids</i>
EDTA/cation exchange column	↓	—	—	—	—
Organic SPE column	— or ↓	↓	↓	—	—
PBO	—	↓	↑	—	—
Water quality measurement				Above threshold ¹	Above threshold ¹

¹Comparison to species specific effect threshold (i.e. NOEC, EC25, or EC50)

Test Species and Exposure Methods

Toxicity Identification Evaluations differ fundamentally from compliance testing. Whereas compliance tests use standardized methods designed to provide reliable and comparable results, TIEs often use less standardized methods in order to investigate specific aspects of the test response. While the methods used in a TIE should be relevant for the situation (e.g. toxicity to a specific test organism) prompting the investigation, flexibility in the specifics of how the study is conducted is needed in order to increase the efficiency and chance for success of the TIE. Typically, TIEs should use the same species as those which prompted the investigation. The sensitivity to specific toxicants may vary markedly between species (Table 5) and the use of a different species in the TIE may result in misleading results. Use of an alternate species is acceptable in some cases, such as when the preferred species is unavailable or when a comparison of responses between species is part of the TIE design (e.g. confirmation studies).

When toxicity to more than one species has been detected in the monitoring program, the species showing the greatest relative response to the sample should be used for the TIE. The test method details should be comparable to those used in monitoring. However, modifications to factors such as replication, test volume, water changes, and sample storage time are acceptable and often necessary in order to apply the TIE work plan. The judgment of an experienced testing lab should be relied upon to determine which test method variations are suitable and data should be available to demonstrate that such changes do not substantially influence the effectiveness of the TIE.

Table 5. Example of TIE results for dry weather flow sample tested using *Ceriodaphnia dubia* at 0, 25, 50 and 100 percent concentrations. Toxic units were determined for each treatment using the calculated LC₅₀ from the dilution series. Organophosphate pesticides were concluded to be the most likely cause of toxicity (Phillips et al. 2010).

<i>Treatment</i>	<i>Mean Percent C. dubia Survival (SD)</i>				<i>Toxic Units</i>	<i>Chlorpyrifos (ng/L)</i>
	<i>Control</i>	<i>25%</i>	<i>50%</i>	<i>100%</i>		
Baseline	100 (0)	100 (0)	40 (20)	0 (0)	2.1	153
SPE Column	93 (12)	93 (12)	100 (0)	100 (0)	<1	ND
SPE Eluate	100 (0)	80 (0)	43 (6)	0 (0)	2.3	362
Carboxylesterase (CEE)	33 (58)*	93 (12)	93 (12)	0 (0)	1.4	137
Bovine Serum Albumin	100 (0)	100 (0)	0 (0)	0 (0)	2.8	158
Piperonyl Butoxide (PBO)	76 (8)	93 (12)	100 (0)	93 (12)	<1	147

Use of multiple test concentrations (e.g. 100%, 50%, 25%) is recommended, especially if toxicity in the original sample is high. Examining response to the TIE treatments over a concentration series enables a more confident interpretation of results for TIE analyses. Quality assurance/quality control procedures for TIEs also differ from those used in compliance testing. Reference toxicants and water quality limits (e.g. dissolved oxygen and pH thresholds) are not usually applied. However, greater use of controls and blanks is required in a TIE in order to detect unintended changes in toxicity due to sample handling or reagent toxicity. Usually, TIE treatment blanks consist of laboratory control water treated in the same manner as the test sample (e.g. addition of ethylene diaminetetraacetic acid (EDTA)).

Interpretation of Results

Several strategies are used in evaluation of TIE results, depending upon the types of treatments conducted. First, treatment effectiveness is determined by comparing the organism's relative response for each treatment compared to the untreated baseline sample. Several approaches have been used to conclude that a treatment has been effective at reducing toxicity including tests of significant difference between treatments, comparison of median lethal concentrations (LC₅₀) if dilutions are utilized, or a static rule of thumb such as a minimum difference of at least 15 percent between treatments and control. For example, the results shown in Figure 4 indicate that the EDTA, C18 solid phase extraction (SPE), aeration, zeolite, and pH increase treatments were effective in reducing toxicity.

Comparison of toxic units (calculated by dividing 100 by the LC₅₀) among the treatments and baseline sample provides a more quantitative measure of treatment effectiveness, but also requires dilution series for each treatment. The data shown in Table 5 indicated the greatest reductions in the sample's toxic units were produced by SPE and piperonyl butoxide (PBO), which is indicative of organophosphate pesticide toxicity. Table 5 also demonstrates the use of TIE treatment blanks (0% concentration). Blank toxicity was present in the carboxylesterase enzyme (CEE) treatment, indicating results for this treatment may be unreliable.

An inconclusive TIE is possible, such as when the toxicity is no longer present in the baseline sample or all TIE treatments are ineffective. Such occurrences should prompt a review and possible revision of the TIE study design before another TIE is conducted. Simple design modifications, such as reducing sample storage time or testing a different sample concentration

may result in a more effective TIE. Alternatively, new toxicants for which no treatment has been devised may be present in the sample.

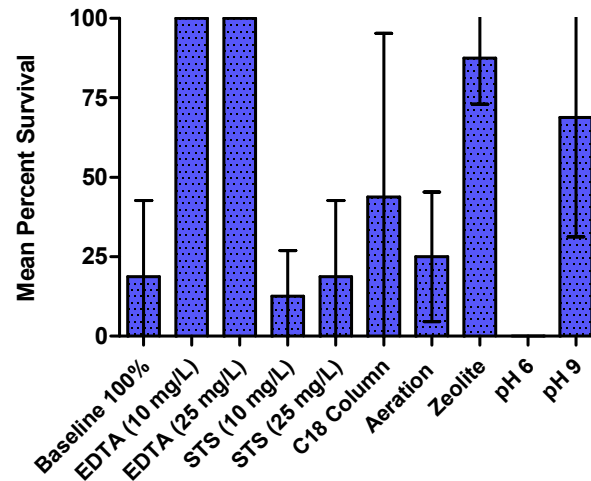


Figure 4. Example of TIE results for storm water sample tested using the fathead minnow. Metals were concluded to be the probable cause of toxicity (Kayhanian et al. 2008).

Identification and Confirmation of Toxicant

Toxicant identification is accomplished through additional sample treatment procedures such as elution/fractionation of constituents retained on SPE columns followed by toxicity testing and chemical analysis. These methods may require additional sample volume and time for testing. Because of the difficulty in obtaining additional sample from the study site that is representative of the initial sample, it is recommended that adequate volume be collected to support anticipated identification analyses at the time of initial sample collection. Toxicant identification methods are varied and often compound or sample specific. Examples of effective methods are found in EPA guidance documents (USEPA 1993a, 1999) and other publications (Norberg-King et al. 2005).

The confirmation phase of a TIE consists of using multiple lines of evidence to confirm the characterization and identification results. Several methods should be used in confirmation, including:

- Repeat characterization of subsequent samples;
- Comparing the mass balance of toxic units between chemical analysis and toxicity tests (requires dilution series);
- Reproducing toxicity by spiking suspected toxicants at similar concentrations; and
- Comparing sensitivity of different test species to that predicted based on threshold concentrations (e.g. Table 6).

Analysis of multiple test samples is needed to establish confidence in the TIE results. The type of toxicant may vary seasonally with land use practices or hydrological factors. At least three separate TIEs should be conducted before a conclusion is reached about cause of toxicity at a specific location. Where biological effects data are unavailable to assess organism sensitivity, such as when addressing constituents of emerging concern (e.g. some pesticides), collaborations among multiple agencies may be appropriate.

Table 6. Effect level (LC₅₀, except where noted) of typical storm water constituents for different test species.

<i>Species and Endpoint</i>	<i>Copper (µg/L)</i>	<i>Zinc (µg/L)</i>	<i>Unionized ammonia (mg/L)</i>	<i>Chlorpyrifos (µg/L)</i>	<i>Diazinon (µg/L)</i>	<i>Bifenthrin (ng/L)</i>	<i>Cyfluthrin (ng/L)</i>
<i>Ceriodaphnia dubia</i> (Survival)	28 (1)	360 (1)	1.49	53 (2)	320 (2)	142 (3)	344 (3)
<i>Pimephales promelas</i> Survival	503 (4)		0.61 (5)				
<i>Hyalella azteca</i> Survival	35 (6)	73 (6)	4.7 (7)	86(6)	6510 (8)	9.3 (9)	2.3 (10)
<i>Strongylocentrotus purpuratus</i> Development	15.3 (12)	96.9 (12)	0.07 (13)				
<i>Strongylocentrotus purpuratus</i> Fertilization	18 (13)	262 (14)	>1.4 (13)				
<i>Mytilus spp.</i> Development	7.8 (12)	178 (12)	0.12 (15)	4900 NOEC (16)			
<i>Haliotis rufescens</i> Development		64 (17)	0.082 (15)				
<i>Atherinops affinis</i> Survival	55.7 (18)		0.56 (15)				
<i>Americamysis bahia</i> Survival	181 (12)	499 (19)			4.2 (20)	3.97 (20)	6.37 (20)
<i>Holmesimysis costata</i> Survival		56 (21)	0.839 (15)				

1. Schubauer-Berigan et al., 1993
2. Bailey et al., 1997
3. Wheelock et al., 2004
4. Naddy et al., 2002
5. AMEC unpublished
6. Phipps et al., 1995
7. Ankley et al., 1995
8. Ankley and Collyard, 1995
9. Anderson et al., 2006
10. Weston and Jackson, 2009
11. Maund et al., 2002

12. Phillips et al., 2003
13. Bay et al., 1993
14. Dinnel et al., 1989
15. Phillips et al., 2005
16. Serrano et al., 1995
17. Hunt and Anderson, 1989
18. Anderson et al., 1994
19. Lussier et al., 1985
20. USEPA, 2000
21. Hunt et al., 1997

REFERENCES

- Anderson BS, Hunt JW, McNulty HR, Turpen, SLM. 1994. Off-season spawning and factors influencing toxicity test development with topsmelt *Atherinops affinis*. *Environ Toxicol Chem* 13:479-485.
- Anderson BS, Phillips BM, Hunt JW, Connor V, Richard N, Tjeerdema RS. 2006. Identifying primary stressors impacting macroinvertebrates in the Salinas River (California, USA): Relative effects of pesticides and suspended particles. *Environ Poll* 141:402-408.
- Ankley, G, Collyard, S. 1995. Influence of piperonyl butoxide on the toxicity of organophosphate insecticides to three species of freshwater benthic invertebrates. *Comp Biochem Physiol C* 110:149-155.
- Ankley GT, Schubauer-Berigan MK, Monson PD. 1995. Influence of pH and hardness on the toxicity of ammonia to the amphipod *Hyalella azteca*. *Can J Fish Aquat Sci* 52:2078-2083.
- Bailey HC, Miller JL, Miller MJ, Wiborg LC, Deanovic LA, Shed, T. 1997. Joint acute toxicity of diazinon and chlorpyrifos to *Ceriodaphnia dubia*. *Environ Toxicol Chem* 16:2304-2308.
- Bay SM, Burgess, RM, Nacci, D, 1993 Status and applications of Echinoid (Phylum Echinodermata) toxicity test methods. pp. 281-302 in: Landis WG, Hughes JS, Lewis MA (eds.), Environmental Toxicology and Risk Assessment. ASTM STP 1179. ASTM. Philadelphia, PA.
- Burton GA, Pitt RE 2001 Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers Lewis Publishers Washington, DC
- Caltrans. 2009. Caltrans Comprehensive Protocols Guidance Manual. CTSW-RT-03-1055142. California Department of Transportation, Sacramento, CA.
- Denton DL, Miller JM, Stuber RA. 2010. EPA Regions 8, 9, and 10 toxicity training tool, January 2010. United States Environmental Protection Agency. San Francisco, CA.
- Denton DL, Diamond J, Zheng L. 2011. Test of significant toxicity: A statistical application for assessing whether an effluent or site water is truly toxic. *Environ Toxicol Chem* doi:10.1002/etc.493.
- Dinnel PA, Link JM, Stober QJ, Letourneau MW, Roberts WE. 1989. Comparative sensitivity of sea urchin sperm bioassays to metals and pesticides. *Arch Environ Contam Toxicol* 18:748-755.
- Heber MA, Reed-Judkins DK, Davies TT. 1996. USEPA's whole effluent toxicity testing program: a national regulatory perspective. pp. 9-15 in: Grothe DR, Dickson KL, Reed-Judkins DK (eds.), Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts. SETAC Press. Pensacola, FL.

- Heggen, R. 2001. Normalized antecedent precipitation index. *J Hydrol Eng* 6:377-382.
- Hunt, JW, Anderson, BS. 1989. Sublethal effects of zinc and municipal effluents on larvae of the red abalone *Haliotis rufescens*. *Mar Biol* 101:545-552.
- Hunt JW, Anderson BS, Turpen SL, Englund MA, Piekarski W. 1997. Precision and sensitivity of a seven-day growth and survival toxicity test using the west coast marine mysid crustacean *Holmesimysis costata*. *Environ Toxicol Chem* 16:824-834.
- Kayhanian M, Stransky C, Bay SM, S Lau, Stenstrom MK. 2008. Toxicity of urban highway runoff with respect to storm duration. *Sci Total Environ* 389:386-406.
- Loureiro S, Svendsen C, Ferreira ALG, Pinheiro C, Ribeiro F, Soares AMVM. 2010. Toxicity of three binary mixtures to *Daphnia magna*: Comparing chemical modes of action and deviations from conceptual models. *Environ Toxicol Chem* 29:1716-1726.
- Lussier SM, Gentile JH, Walker J. 1985. Acute and chronic effects of heavy metals and cyanide on *Mysidopsis bahia* (Crustacea: Mysidacea). *Aquat Toxicol* 7:25-35
- Lyon G, Petschaur D, Stein, E. 2005. Effluent discharges to the Southern California Bight from large municipal wastewater treatment facilities in 2003 and 2004. Pp. 1-15 in: Weisberg S, Miller K (eds.), Southern California Coastal Water Research Project Biennial Report: 2003-2004. Southern California Coastal Water Research Project. Westminster, CA.
- Maund SJ, Hamer MJ, Lane MCG, Farrelly E, Rapley JH, Goggin UM, Gentle WE. 2002. Partitioning, bioavailability, and toxicity of the pyrethroid insecticide cypermethrin in sediments. *Environ Toxicol Chem* 21:9-15.
- Muir DCG, Howard PH. 2006. Are there other persistent organic pollutants? A challenge for environmental chemists. *Envir Sci Tech* 40:7157-7166.
- Naddy RB, Stubblefield WA, May JR, Tucker SA, Hockett JR. 2002. The effect of calcium and magnesium rations on the toxicity of copper to five aquatic species in freshwater. *Environ Toxicol Chem* 21:347-352.
- Norberg-King TJ, Ausley LW, Burton DT, Goodfellow WL, Miller JL, Waller WT. 2005. Toxicity Reduction and Toxicity Identification of Evaluations for Effluents, Ambient Waters, and Other Aqueous Media. SETAC. Pensacola, FL.
- Phillips BM, Anderson BS, Hunt JW, Siegler K, Voorhees JP. 2010. Santa Maria Watershed and Oso Flaco Creek Watershed TMDL Monitoring Study – Final Report. Central Coast Regional Water Quality Control Board. San Luis Obispo, CA.
- Phillips, BM, Nicely, PA, Hunt, JW, Anderson, BS, Tjeerdema, RS, Palmer, FH. 2005. Tolerance of five west coast marine toxicity test organisms to ammonia. *Bull Environ Contam Tox* 75:23-27.

Phillips BM, Nicely PA, Hunt JW, Anderson BS, Tjeerdema RS, Palmer SE, Palmer FH, Puckett HM. 2003. Toxicity of cadmium-copper-nickel-zinc mixtures to larval purple sea urchins (*Strongylocentrotus purpuratus*). *Bull Environ Contam Tox* 70:592-599.

Phipps GL, Mattson VR, Ankley GT. 1995. The relative sensitivity of three benthic test species to ten chemicals. *Arch Environ Toxicol Chem* 28:281-286.

Sacramento River Watershed Program (SRWP). 2000. Sacramento River Watershed Program toxicity testing data results summary: 1998-99. Staff report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.

Sahu RK, Mishra SK, Eldho TT. 2010. An improved AMC-coupled runoff curve number model. *Hydrologic Process* 24:2834-2839.

Schubauer-Berigan MK, Dierkes JR, Monson PD, Ankley GT. 1993. pH-dependent toxicity of Cd, Cu, Ni, Pb, and Zn to *Ceriodaphnia dubia*, *Pimephales promelas*, *Hyaella azteca*, and *Lumbriculus variegatus*. *Environ Toxicol Chem* 12:1261-1266.

Serrano R, Hernandez F, Pena JB, Dosda V, Canales J. 1995. Toxicity of bioconcentration of selected organophosphorus pesticides in *Mytilus galloprovincialis* and *Venus gallina*. *Arch Environ Contam Toxicol* 29:287-290.

Skinner L, dePeyster A, Schiff K. 1998. Developmental effects of urban stormwater in San Diego County, California in medaka (*Oryzias latipes*) and inland silverside (*Menidia beryllina*). *Arch Environ Contam Toxicol* 37:227-235.

Surface Water Ambient Monitoring Program (SWAMP). 2008. The Surface Water Ambient Monitoring Program Quality Assurance Program Plan. State Water Resources Control Board. Sacramento, CA. http://www.waterboards.ca.gov/water_issues/programs/swamp/docs/qapp

Tiefenthaler LL, Stein ED, Schiff KC. 2008. Watershed and land use-based sources of trace metals in urban stormwater. *Environ Toxicol Chem* 27:277-287.

United States Environmental Protection Agency (USEPA). 1989a. Methods for aquatic toxicity identification evaluations: Phase II toxicity identification procedures. EPA/600/3-88/035. USEPA Office of Research and Development. Washington, DC.

USEPA. 1989b. Methods for aquatic toxicity identification evaluations: Phase III toxicity confirmation procedures. EPA/600/3-88/036. USEPA Office of Research and Development. Washington, DC.

USEPA. 1991. Methods for aquatic toxicity identification evaluations: Phase I toxicity characterization procedures. Second edition. Norberg-King TJ, Mount DI, Durhan EJ, Ankley GT, Burkhard LP, Amato JR, Lukasewycz MT, Schubauer-Berigan MK, Anderson-Carnahan L (eds.). EPA/600/6-91/003. USEPA Office of Research and Development. Washington, DC.

USEPA. 1992. Toxicity identification evaluation: Characterization of chronically toxic effluents, Phase I. Norberg-King TJ, Mount DI, Amato JR, Jensen DA, Thompson JA (eds.). EPA/600/6-91/005F. USEPA Office of Research and Development. Washington, DC.

USEPA. 1993a. Methods for aquatic toxicity identification evaluations: Phase II toxicity identification procedures for samples exhibiting acute and chronic toxicity. EPA/600/R-92/080. USEPA Office of Research and Development. Washington, DC.

USEPA. 1993b. Methods for aquatic toxicity identification evaluations: Phase III toxicity identification procedures for acutely and chronically toxic samples. EPA/600/R-92/081. USEPA Office of Research and Development. Duluth, MN.

USEPA. 1995. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to West Coast marine and estuarine organisms. Chapman GA, Denton DL, Lazorchak JM (eds.). EPA/600/R-95/136. USEPA Office of Research and Development. Cincinnati, OH.

USEPA. 1996. Marine toxicity identification evaluation (TIE): Phase I guidance document Burgess R, Ho K, Morrison G, Chapman GA, Denton DL (eds.). EPA/600/R-95/054. USEPA Environmental Effects Research Laboratory. Narragansett, RI.

USEPA. 1999. Toxicity reduction evaluation guidance for municipal wastewater treatment plants. Second edition. EPA/833/B-99/002. USEPA Office of Water. Washington, DC.

USEPA. 2000. Pesticide Ecotoxicity Database (Formerly: Environmental Effects Database (EEDB)). USEPA Office of Pesticide Programs Environmental Fate and Effects Division. Washington, DC.

USEPA. 2001. Clarifications regarding toxicity reduction and identification evaluations in the national pollutant discharge elimination system program. USEPA Office of Wastewater Management and Office of Regulatory Enforcement. March 27, 2001 Washington, DC.

USEPA. 2002a. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms. Fifth edition. EPA/821/R-02/012. USEPA Office of Water. Washington, DC.

USEPA. 2002b. 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.

USEPA. 2002c. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms. Third edition. EPA/821/R-02/014. USEPA Office of Water. Washington, DC.

Weston D, Jackson C. 2009. Use of engineered enzymes to identify organophosphate and pyrethroid-related toxicity in toxicity identification evaluations. *Environ Sci Tech* 43:5514-5520.

Wheelock CE, Miller JL, Miller,MJ, Gee SJ, Shan G, Hammock BD. 2004. Development of toxicity identification evaluation procedure for pyrethroid detection using esterase activity. *Environ Toxicol Chem* 23:2699-2708.

APPENDIX 1 – FREQUENTLY ASKED QUESTIONS

(adapted from Denton et al. 2010)

Q: Have the toxicity test methods been used to assess agricultural, urban, and industrial storm water runoff toxicity? If so, what toxicant(s) have been identified?

A: Toxicity testing of storm water has been used as a monitoring tool for urban and agricultural storm water assessments in California. For example, researchers have identified the pesticides diazinon and chlorpyrifos in urban storm waters (Katznelson and Mumley 1997; Bailey et al. 2000; Fong et al. 2000; Larsen et al. 2000; SRWP 2000; Larsen and List 2002). Toxicity testing of storm water from agricultural settings has identified rice pesticides, diazinon, chlorpyrifos, carbofuran, and carbaryl as toxicants (SRWP 1998; Foe et al. 1998; Reyes et al. 2000; Werner et al. 2000).

Q: Are acute and/or chronic test method(s) used to assess storm and ambient waters?

A: Acute and short-term chronic tests are primarily being used to initially assess the toxicity of storm and ambient waters.

Q: What testing factors may need to be considered differently for storm water testing compared to testing effluent from a continuous discharge?

A: The main factors include: 1) sample collection and sample initiation holding time; 2) sample renewals; and 3) immediate responses to observed toxicity such as initiating toxicity identification evaluations.

Q: Can an exception to the 36-hour holding time for initiation of the test be allowed for storm and ambient water testing?

A: All tests should be conducted as soon as possible following sample collection. EPA has allowed exceptions to the 36-hour holding time, for example, when effluents are shipped overseas for testing (Denton and Narvaez 1996). The primary reason for an extension of the holding time would be the consideration of the sampling, laboratory technician safety (Burton and Pitt 2001; see page 255), and logistics of coordinating collection and transport of multiple storm water samples within a short period of time. Storm events are not pre-determined events and typically occur rapidly throughout a watershed; therefore, many site samples must be coordinated and processed with short notification to the toxicity testing laboratories. It is encouraged that the 36-hour holding time for test initiation be targeted; however, the Permitting Authorities may allow an exception beyond the 36-hours. However, no more than 72 hours should elapse before initial use of a sample. Any sample tested between 36 and 72 hours should be flagged with a data qualifier.

Q: How is the standard test renewal practices specified in the test method manuals followed, given that storm events may be of short duration?

A: EPA 5th edition acute test methods specify that test solutions be renewed after 48 hours for a 96-hour test. However, for storm events in short duration, this is not always feasible. A more realistic option, in cases when a second storm water sample may not be available, would be to collect sufficient volume during the storm event to use for the start of the test and at the 48-hour renewal.

Q: When would a multiple dilution test be performed if a single concentration test is initially conducted?

A: A single concentration is typically compared to a control to determine the effect in 100 percent storm water and ambient water exposures. A multiple concentration test could be considered for the next sampling event if toxicity is of significant magnitude in the 100 percent storm water (e.g. 100% mortality within 24 to 48 hours). The testing facility may consider testing the original sample (assuming sufficient volume collected) with a dilution series to more fully characterize the sample, for those samples which demonstrate high mortality within a short timeframe.

Q: What is meant by the term “first flush” when referring to collection of storm water samples?

A: “First flush” refers to the first waters released from a discharge point as a result of a storm event or runoff associated with ice and snow melt. Typically, constituent concentrations are highest in this “first flush” sample. “First flush” is operationally defined by a time-period in some states (e.g. waters discharged within the first 15 or first 30 minutes of a discharge event). However, the “first flush” may not always contain the highest concentrations of pollutants as this depends on the rain intensity, type of pollutant, and size of the watershed. The first flush phenomenon is more prevalent for rains with relatively constant intensities and small watershed size (Burton and Pitt 2001). Therefore, it is important to understand the watershed in order to determine if sampling of first flush in a storm event is critical. Another consideration is to capture the first seasonal flush (e.g. after an extended dry period) in arid areas.

Q: Is capturing the first flush important?

A: The precedent has been established for chemical-specific storm water sampling to sample first-flush discharges. Existing data suggests the potential for higher chemical-specific toxicity in first-flush samples. This “first flush” effect depends on the nature and form of the pollutant (Ward and Elliot 1995). The chemograph peak often precedes that of the hydrograph for sediments or sediment-bound pollutants (e.g. chlorpyrifos, phosphorus) entrained in the water column. However, for dissolved pollutants like diazinon, the chemograph peak sometimes follows that of the hydrograph.

Q: Is timing of sample collection to a flow measurement important?

A: Two types of samples are appropriate for collecting storm water to test toxicity; discrete or composite samples. Composite samples are collected throughout an entire storm event typically as a function of flow (i.e. flow-weighted compositing). Discrete sampling is instantaneous; hence, timing of discrete sampling is very important. Discrete samples should be collected on the rising limb of the hydrograph. In nearly all cases, flow data should be collected at the time of sampling.

REFERENCES

- Bailey HC, Miller JL, Miller MJ, Weber LC, Deanovic LA, Shed T. 2000. Diazinon and chlorpyrifos in urban waterways in Northern California. *Environ Toxicol Chem* 19:82-87.
- Burton GA, Pitt RE. 2001. *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers*. Lewis Publishers. Washington, DC.
- Denton DL, Narvaez M. 1996. EPA Regions 9 and 10 guidance for implementing whole effluent toxicity testing programs. USEPA Regions 9 and 10. Sacramento, CA.
- Denton DL, Diamond J, Zheng L. 2011. Test of Significant Toxicity: A statistical application for assessing whether an effluent or site water is truly toxic. *Environ Toxicol Chem* doi:10.1002/etc. 493.
- Foe CG, Deanovic LA, Hinton DE. 1998. Toxicity identification evaluations of orchard dormant spray storm runoff. Staff report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.
- Fong S, Deanovic L, Reyes E, Larsen K, Louie S, Hinton D. 2000. Sacramento-San Joaquin Delta toxicity monitoring project: 1998-1999 Final Report. Staff Report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.
- Katznelson R, Mumley TE. 1997. Diazinon in surface waters in the San Francisco Bay area: Occurrence and potential impact. State Water Resources Control Board Report. Sacramento, CA.
- Larsen KL, List K. 2002. Sacramento River Watershed Program Toxicity Monitoring Data Summary: 1999-2000. Staff report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.
- Larsen KL, McGraw MR, Connor VM, Deanovic LA, Kimball TS, Hinton DE. 2000. Cache Creek and Putah Creek Watersheds toxicity monitoring results: 1998-1999 Final Report. Staff report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.
- Reyes E, Foe C, Deanovic L, Hinton DE. 2000. Orchard in-season spray toxicity monitoring results: 1997. Staff report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.
- Sacramento River Watershed Program (SRWP). 1998. Sacramento river watershed project toxicity monitoring results: 1996-1997. Final Report. Larsen K, Connor V, Deanovic LA, Hinton DE, editors. Sacramento, CA.
- SRWP. 2000. Sacramento River Watershed Program toxicity testing data results summary: 1998-99. Staff report to the Central Valley Regional Water Quality Control Board. Sacramento, CA.

Ward AD, Elliot WJ. 1995. Environmental Hydrology. CRC Lewis Publishers. Washington, DC.

Werner I, Deanovic LA, Connor V, de Vlaming V, Bailey HC, Hinton DE. 2000. Insecticide-caused toxicity to *Ceriodaphnia dubia* (cladocera) in the Sacramento-San Joaquin River Delta, California, USA. *Environ Toxicol Chem* 19:215-227.

APPENDIX 2 – STORM WATER SAMPLING AND TESTING MONITORING INVENTORY QUESTIONNAIRE

MS4 Sampling and Testing Questionnaire

IDENTIFICATION INFORMATION

Agency name
RWQCB jurisdiction
Permit number
Year of last permit renewal

TOXICITY TESTING

Species
Endpoints
Number of replicates
Number of dilutions
Renewals (Yes/No)
Frequency of renewals
Reporting units
QA Criteria

SAMPLING METHODS

Total number of sites
Number sites/watershed (average)
Number wet samples per year
Number dry samples per year
Sample volume for toxicity testing (liters)
New sample for each renewal (yes/no)
New sample for TIE (Yes/No)
Storm trigger (rainfall quantity)
Storm trigger (antecedent dry)
Storm sample type (Grab, time weighted, flow weighted)
Storm sample collection method (surface, single point, depth integrated)
Holding time (hrs)
Container type and storage conditions

TIE TRIGGER

Number TIEs per year
TIE trigger
Initial TIE treatments routinely used
Additional treatments available/have been used
New sample/Extra volume from Initial sample
Dilutions tested
QA criteria and blanks

APPENDIX 3 – FLOW CHART FOR STORM WATER SAMPLING ACTIVITIES

Storm sampling is quite unlike sampling more steady discharges such as POTWs. Storms are unpredictable so sampling preparations require more attention and organization. Although only two storm samples are recommended each year, preparation and readiness may last six months. In addition, weather forecasting is an imperfect science and, because storm events may not materialize or equipment may not work properly, mobilization for more than two storms may be required in order to obtain sufficient samples for analysis. In order to assist storm water agencies and Regional Water Board's, an example flow chart modeled after the Caltrans (2009) monitoring program has been prepared to assist those unfamiliar with storm sampling. Budgeting, site selection, and sampling equipment preparation must occur well before the rainy season. Several criteria should be considered for selecting a site including safety of field crews, representativeness, capability to rate flow, flood hazards, and power/telecommunication for sampling equipment. Site set up has its own unique challenges including obtaining encroachment permits, selecting and purchasing appropriate equipment, and equipment security. Once installed and the wet season approaches, weather forecasts should be monitored daily. As storms approach, increase the frequency of weather forecasts and notify personnel and support laboratories within a minimum of three days. This notification is crucial so that toxicity laboratories have sufficient time to schedule staff, prepare brood stock, ready equipment, and clean glassware. As the storm approaches, additional communication with the toxicity laboratory is required to ensure that personnel are available to deliver/receive samples because samples may arrive at night, on weekends, or holidays. While field personnel are crucial for sampling success, their job is typically completed within 24 hours of storm end. Toxicity testing personnel, however, are needed for up to 10 days post-storm. Finally, toxicity testing laboratories need to communicate with storm water agency personnel at test completion to inform them of a successful test. If toxicity testing is unsuccessful, storm water agencies may wish to cancel chemistry analysis and prepare for sampling additional storms.

