Modeling to Improve Precision, Accuracy, and Sensitivity of Biological Indices: Nevada as a case study

Charles Hawkins & Jake Vander Laan
chuck.hawkins@usu.edu
www.cnr.usu.edu/wmc

Western Center for Monitoring and Assessment
of Freshwater Ecosystems
Department of Watershed Sciences
Utah State University
Logan, Utah

California Aquatic Bioassessment Workgroup
18th Annual Meeting
9-10 November 2011

UtahState University
DEPARTMENT OF WATERSHED SCIENCES

The Western Center for Monitoring & Assessment of Freshwater Ecosystems

Outline

- Biological indices, sources of index variability, and the need to model.
- Index development for Nevada
  - O/E, BC, and MMI
    - Focus today on MMI development
    - Accounting for metric variation across natural environmental variation.
  - Evaluation and comparison of index performance
    - Precision
    - Bias
    - Responsiveness
    - Sensitivity

Predicting the reference condition

- Aim of bioassessment is to compare observed biological conditions with those expected under reference conditions.
- The main technical challenge in producing accurate bioassessments is estimating the appropriate biological reference condition.
  - Biota vary markedly among sites with natural differences in temperature, hydrology, water chemistry, geomorphology, and streambed substrates.
- Historical (pre-impact) data not available, so we must predict the reference condition.

How do we make predictions?

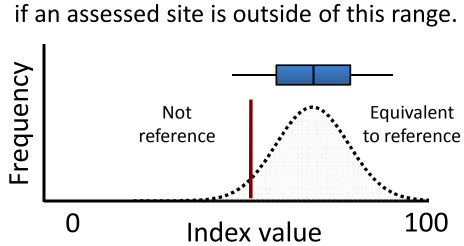
- Collect biological samples from reference sites.
- Match each assessed site with the most appropriate reference sites, which are used to derive the '*reference condition*' for each site.
- Two possible approaches:
  - Regionalizations or typologies: most IBIs and MMIs.
  - Continuous modeling: most O/E indices.

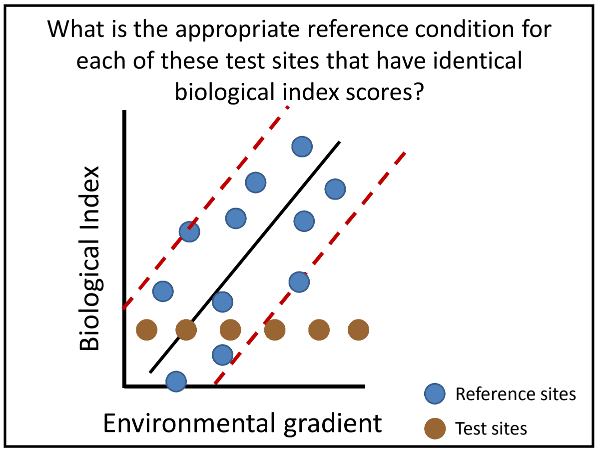
The case for modeling

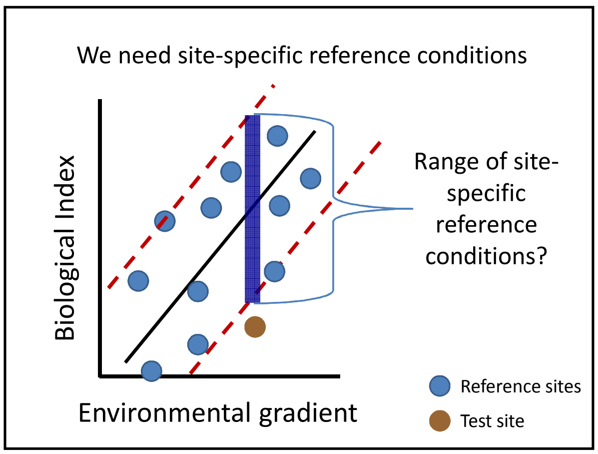
- Estimates of the biological reference condition contain multiple sources of variation.
  1. Sampling method – control by standardizing methods.
  2. Random sampling error (RSE).
  3. Systematic variation with season and year (temporal).
  4. Systematic variation with environmental differences within and among sites (spatial).
- All sources affect the risk of both Type I and Type II errors of inference, which we want to minimize.
- Minimize RSE by adjusting sampling protocols.
- Temporal variation at each site might be considered an estimate of the true range of natural variation, but may want to adjust assessments for, say, wet years versus dry years.
- Dealing with systematic variation across sites is probably the biggest problem.
  - Well known that regionalizations do not account for much of this variation, so need to model!

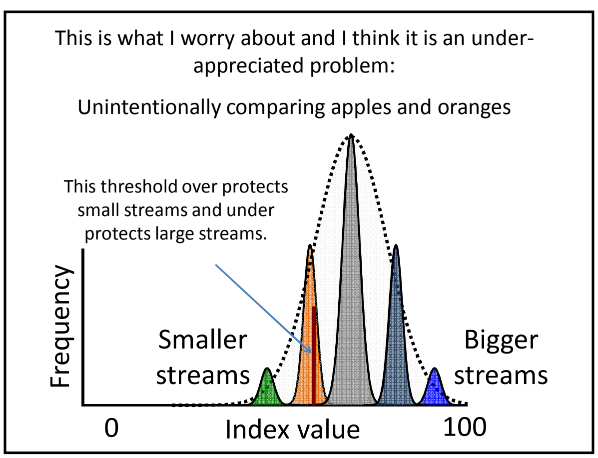
Bioassessment in Practice:
Observed variation among reference sites is often assumed to represent the range of natural variation.

We then select a threshold below which we infer if an assessed site is outside of this range.




What is the appropriate reference condition for each of these test sites that have identical biological index scores?
Biological Index
Environmental gradient
Reference sites
Test sites

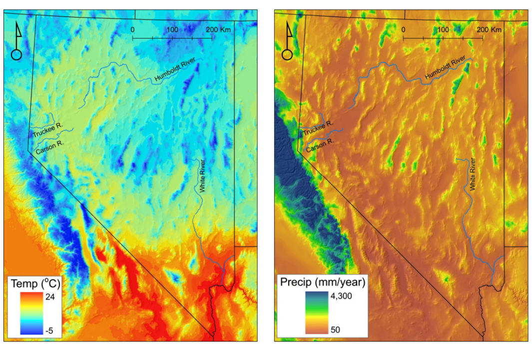

We need site-specific reference conditions
Range of site-specific reference conditions?
Biological Index
Environmental gradient
Reference sites
Test site


This is what I worry about and I think it is an under-appreciated problem:
Unintentionally comparing apples and oranges
This threshold over protects small streams and under protects large streams.
Frequency
Smaller streams
Bigger streams
0
Index value
100

The Nevada Indices: today's case study in maximizing index performance

- O/E: Index of taxonomic completeness.
  - Ratio based on the specific taxa predicted to occur at a site (E) that are observed (O).
  - Also calculated Van Sickle's BC (Bray-Curtis) index.
- MMI: Multimetric index (similar to an IBI)
  - Index based on the cumulative scores of several different types of individual metrics.

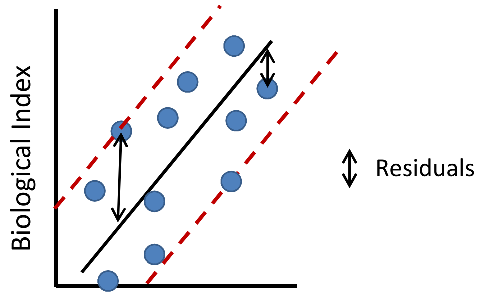
Spatial variation in Nevada climate



Developing the Indices: MMI

1. Identify reference sites.
2. Identify candidate metrics and calculate metric values from taxa counts.
3. Identify most degraded sites for index calibration.
4. Build Random Forest models for each metric to account for natural gradients and use residuals as metric values.
5. Conduct PCA on raw metrics or residuals to identify sets of statistically independent metrics.

Modeling out the effects of variation in natural environmental gradients



Developing the Indices: MMI

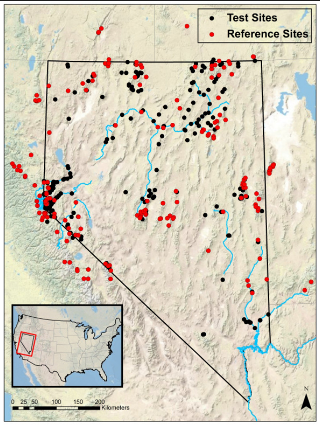
6. Determine discriminatory power for each metric (t-test based on mean ref and stressed site values).
7. Select the metric from each PC axis with greatest t-value.
8. Standardize metric values by scaling where the ceiling = 95th percentile value of reference sites and the floor = the 5th percentile of stressed site values.
9. Calculate the MMI value as Σ standardized metrics.
10. We also calculated the null MMI, i.e., metrics were not modeled.

Reference Site Selection

Variable	Small Watersheds	Medium Watersheds	Large Watersheds
Area (km^2)	≤ 164	164 - 3,380	> 3,380
% Agriculture in basin	0	1	≤ 2
% Urban in basin	0	≤ 1	≤ 2
% Agriculture within 3 km	0	≤ 3	≤ 5
% Urban within 3 km	0	≤ 3	≤ 5
NDPES discharges	0	NA	NA
Dams	0	0	NA
Mines	0*	0*	NA

The Nevada data

- Samples were collected by NDEP, EPA, and USU at ~500 sites.
- Selected EMAP comparable samples.
- 165 reference sites.
- 416 test (non-reference site) samples.



Selection of metrics

- Started with a list of > 100 metrics.
- We identified 30 candidate metrics that have been used in other western USA MMIs plus *Hydra* relative abundance (n = 31).
- We then selected those metrics that in either raw or modeled form were statistically independent and best discriminated between reference and stressed sites.
  - 12 final metrics across modeled and null MMI.

Candidate Predictors

WS area	WS max wet days
WS min elev	Site max precipitation
WS max elev	WS max precipitation
WS mean elev	WS min precipitation
WS slope	WS hydro-stability
Elev CV at sample site	WS base-flow Index (BFI)
Predicted conductivity	WS max air temperature

Response of 12 metrics used in the MMIs to natural environmental variation

Metric	R^2	Type of Predictor in Rank Order of Importance						
Simpson diversity	0							
Shannon diversity	<10							
# Insect taxa	20	ELV	T	WSA	P	P		
% mayfly abundance	21	P	HYDR	ELV	T	P	BFI	
% *Hydra* abundance	23	T	P	WSA	ELV			
% CF abundance	24	P	ELV	P	Slope	T	WSA	HYDR
# Long-lived taxa	31	Slope	ELV	T	T	P	WSA	HYDR
# Mayfly taxa	33	P	WSA	P	P			
# Non-insect taxa	34	P	HYDR	P	T	P	ELV	
% Stonefly abundance	37	P	P	BFI	T	WSA		
# Clinger taxa	40	P	WSA	BFI	P			
# Intolerant taxa	40	P	T	Slope	WSA	ELV		

BFI	Base-flow index	P	Precipitation
ELV	Elevation	T	Temperature
HYDR	Hydrostability	WSA	Watershed area

Final metrics for the two MMIs

Modeled (7 metrics)	Null Model (6 metrics)
# Insect taxa	# Insect taxa
Mayfly relative abundance	# Intolerant taxa
Shannon diversity	Simpson diversity
Collector-filterer relative abundance	*Hydra* relative abundance
Stonefly relative abundance	# Mayfly taxa
# Non-insect taxa	# Long-lived taxa
# Clinger taxa	

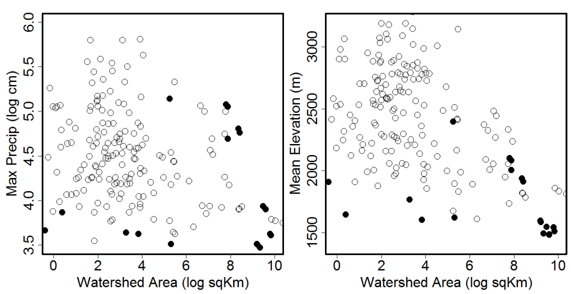
Putting it all together

- Guarding against inappropriate assessments.
- Applying appropriate statistical tests to determine impairment:
  - the combined use of interval and equivalency tests.
- Index performance.
- Index response to known stressors.

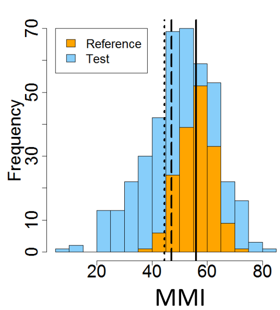
A neglected issue in MMI world: flagging problematic site matching

- Should not assess sites that are naturally different from the network of reference sites.
- Used a nearest-neighbor test to determine if a test site is different from its closest reference-sites in terms of 3 major environmental factors:
  - Watershed area
  - Mean watershed elevation
  - Long-term mean maximum watershed precipitation
- Sites outside of the 90th percentile ellipse calculated from the 10 nearest reference sites in 3-factor space are flagged.

The nearest-neighbor approach



Assessing Impairment: Interval and Equivalence Tests (threshold = 5th %tile of reference values)



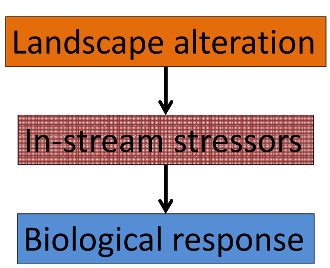
Evaluating Index Performance

- Precision: CV of reference site index values.
- Bias: % variation of reference site index values associated with natural gradients (RF model).
- Responsiveness: mean difference between reference and heavily-stressed sites.
- Sensitivity: % of non-reference site samples inferred as being in non-reference condition.

Index Performance

	Bias	Precision	Responsiveness		Sensitivity
	Post-model R^2	CV	All Test	Most Stressed	% Non-Ref
O/E	0.6	0.21	0.89	**0.57**	23.3
Null O/E	**11.7**	0.23	0.90	**0.57**	22.6
MMI	0	**0.11**	0.90	0.64	**34.9**
Null MMI	**21.7**	0.22	0.88	0.66	20.4

More on index response to stress (Jake Vander Laan's presentation)



Take Home Messages

- Effects of natural gradients on biology are pervasive – expect to see them (always).
- Landscape alteration is often confounded with natural gradients, which can confound assessments.
- To produce the most accurate and fair inference of biological condition, assessments should be based on site-specific biological potentials.
- Modeling is not voo-doo and can produce site-specific expectations for any biological index.

Contacts

- Charles Hawkins - chuck.hawkins@usu.edu
- Jake Vander Laan - jacob.vl@aggiemail.usu.edu