PEEIR

Pacific Estuarine Ecosystem Indicator Research Consortium

Limitations of Multi-Metric Indices, Including the Index of Biotic Integrity (IBI)

Issue

Biological integrity was a goal of the 1972 Clean Water Act, but was at first assumed to be best measured by physical and chemical variables. However, measuring these variables can be impractical and/or expensive and related biological effects can depend on their interactions, and on how they vary in time and space. Measurement of biological variables now plays a large role in environmental management. Nearly all entities with monitoring responsibilities (e.g., states, territories and tribes) use biological indicators. Comparing these indicators among sites, or to standards based on reference sites (i.e., relatively undisturbed), should help guide assessment and remediation efforts.

To facilitate such comparisons, many entities combine indicators into an overall multimetric index to summarize biological condition and distinguish degraded sites. At first, PEEIR considered developing such indices for West Coast wetlands. However, such standard indices require enough undisturbed sites in a region to estimate natural distributions of the indicators. West Coast wetlands are few in number, spread over many degrees of latitude, and are *all* disturbed. A new basis was needed. This is described in the Executive Summary, and involves multiple indices, each justified by its contribution toward an explicit goal.

This conclusion appears to have broad generality and is not particular to the wetlands studied by PEEIR. Appreciating this led us to re-evaluate the rationale behind the use of a single multimetric index to characterize biological integrity.

Approach and Rationale

We focused on the best-known standard index, the Index of Biotic Integrity (IBI). A large literature in journals, books and agency reports generally supports its use and applies it to one or more situations, usually freshwater stream sites, but also wetlands or other habitat types. From these articles we extracted a list of claims and supporting arguments made for IBI. We also made a list of criticisms from the few papers and related indices that criticize IBI, and also listed responses made to these critics by IBI proponents. We took care not to miss the virtues of IBI or of indices in general, or merely to repeat earlier criticisms. It is significant that the critiques, though seeming sensible and from respected authors, have had little effect on its use. Proponents might have invested too much in IBI to give it up, but a good critique must strive for a coherent argument against the strongest case for IBI.

IBI calculation involves choosing a set of biological indicator variables, such as numbers of species in various classes (like natives, insectivores, or "intolerant of degradation"). Their values are awarded numerical scores, based on comparison with the values at undisturbed sites. Metrics that approximate what biologists would expect at minimally disturbed sites are assigned a score of 5; those that deviate somewhat from such sites receive a score of 3; those that deviate strongly are scored 1. IBI is their sum.

Understanding the condition of a marsh ecosystem requires assessment of many aspects of this complex environment. A critical evaluation of how indicators are combined is essential. This is an aerial photo of Carpinteria Salt Marsh which, although relatively undisturbed, is subject to multiple stressors including habitat modification, excess nutrient inputs and urban runoff. PEEIR sampling stations are indicated by white dots.

A research partnership between **University of California, Davis, Bodega Marine Laboratory** and **University of California, Santa Barbara**
Funded by **U.S. EPA Science To Achieve Results (STAR) EAGLE Program** Grant No. R82867601

Findings and Impact

The value added by combining the indicator values into a single number is claimed to be:

1. more precise than single indicators because it combines them
2. quantitative
3. an effective way to communicate with non-specialists
4. analogous to other useful indices
5. able to summarize large datasets to reveal patterns in space and time
6. a synthesis of ecological theory and professional judgment
7. flexible
8. sensitive to many types of degradation
9. justified by its successful use in many studies.

Our main criticism is that combining indicators of different features or problems into a single number obscures information needed by managers, scientists and the public. Claim 1 is weak at best. Sites really needing assessment will have intermediate IBI scores that could result from thousands of possible combinations of indicator values. For example, in a study undertaken by PEEIR researchers that used 12 indicators, there were 8074 ways of obtaining a value of 48 for the IBI. These combinations carry different messages. Equal IBI's are not all equal for decision-making since unequal ones do not always imply that the site with the lower score is more degraded. Hence, claim 2 is false. Previous critiques of IBI usually list this criticism with several others, some of them arguable. Perhaps the potential significance of this primary criticism of IBI has been lost amidst a raft of other concerns.

We identified ways that an index combining diverse measurements can give useful information. It can be deduced or derived from mechanistic models or statistical methods like regression or factor analysis, or constructed intuitively from considerations of mechanisms, importance, etc. All are part of PEEIR's approach. The key is that it estimates or predicts an **independently defined** (possibly implicitly, as in factor analysis) target of interest.

IBI is not useless. It can identify extreme sites where almost all indicators are low, and since the indicator values still exist, calculating the IBI may seem harmless. However, we argue that it is likely to be actively harmful, especially in view of claim 3 above.

In summary, if all indicators indicate the same problem, the methods for combining them should have known properties and achieve clear objectives, preferably quantitative; if they indicate different problems, then combining them obscures their messages.

Applications

PEEIR advocates the development of indicators that are linked by key properties and clear objectives. We are working with several agencies to promote the use of Resident Species Portfolios based, in part, on statistical indicators for use in ecological risk assessment. In addition, we collaborated with The Bay Institute on publication of the San Francisco Bay Scorecard. The individual scores in the SF Bay Scorecard represent a situation where quantities to be aggregated relate to previously (independently) determined regulatory concerns – they would become meaningless if someone attempted to calculate a "GPA".

Publications

Please contact lead author for status of manuscripts in preparation.

Contact Names

University of California, Santa Barbara,
Allan Stewart-Oaten, stewart@lifesci.ucsb.edu
William W. Murdoch, murdoch@lifesci.ucsb.edu
Roger M. Nisbet, nisbet@lifesci.ucsb.edu
University of California, Santa Cruz,
Masami Fujiwara, fujiwara@ucsc.edu

A research partnership between **University of California, Davis, Bodega Marine Laboratory** and **University of California, Santa Barbara**
Funded by **U.S. EPA Science To Achieve Results (STAR) EAGLE Program** Grant No. R82867601

This research is funded by U.S. EPA - Science To Achieve Results (STAR) Program
Grant # R-82867601