Prioritizing Management Goals for Stream Biological Integrity Within the Developed Landscape Context.

This document describes a modeling approach to estimate the likely degree that 'large-scale, historical impacts from landscape alteration' constrains the condition of aquatic life at individual waterbodies defined here as reaches within the National Hydrography Dataset Plus. The goal of the project was to 'present the development and application of a landscape model to classify and prioritize stream monitoring sites based on probable ranges of bioassessment scores relative to landscape alteration'. Such a model could be of considerable value in that it could provide a rapid and inexpensive way of identifying streams that have been so greatly modified that they can no longer achieve biological integrity.

The charges to the science advisory panel were to:

- Comment on the adequacy of the data set and the analytical approaches to predict ranges of biointegrity scores associated with landscape development.
- Comment on the evaluation of performance and findings of the Channels in Developed Landscape Tool, including applicability of the Tool to the range of constructed or hydro-modified channels.
- Consider if there are technical ways to address stakeholder concerns?

Conceptual background

As acknowledged in the document, water resource managers need a way of estimating the degree to which individual waterbodies can support or achieve biological integrity goals. Specifically, we need objective ways of identifying waterbodies that are constrained in achieving biological integrity by historical land use and waterway alteration. This knowledge can then be used by policy makers to set practical goals for the protection or restoration of individual waterbodies and to prioritize management actions across waterbodies.

The conceptual framework underpinning the approach is not wholly new, and the document would be improved by at lease briefly pointing out that the approach is conceptually similar to ideas that have been discussed extensively elsewhere – e.g., the Driver-Pressure-State-Impact-Response (DPSIR) conceptual framework (e.g., Smeets and Wethering, 1999).

Smeets, E. and R. Weterings. 1999. Environmental indicators: typology and overview. European Environmental Agency Technical Report No. 25. 19 p

Clarity of the document

The document was generally well written and clear, but the clarity of some terms and sections could be improved. The need for improved clarity was noted by members of the science advisory panel, the external reviewers of the submitted manuscript, and the stakeholders. The authors define their use of the term 'constrained' to mean situations in which "reference conditions for the biological community may be difficult to achieve with limited resources because of large-scale, historical impacts from landscape alteration". In isolation, this definition is straightforward, but the term ‘constrained’ is used in the general stream ecology literature to mean that the location and morphology of the stream channel is physically constrained by natural landscape elements (e.g., canyon walls). The authors acknowledge this historical use of the term in the discussion, but it should be addressed early on. The authors should also strive to minimize the use of jargon and ensure that their use of terms is not ambiguous. For example, the authors use the term ‘context’ throughout the document, but its specific meaning is not always clear, and its meaning seems to vary.

The modeling approach

The objective of the modeling was to estimate the range of CSCI values likely to occur at individual waterbodies as a function of site-specific landscape alteration. The authors developed a statewide quantile random forest model to estimate upper and lower quantiles (and hence prediction intervals) that bracket the mean (or median) predicted values. The upper and lower prediction quantiles can be set by the user depending on how certain the user wants to be that the actual CSCI value falls within the prediction interval (e.g., 95%, 90%, 80%, etc.). The panel thought this was a technically sound approach to estimating site-specific prediction intervals for random forest models, which often describe complex, non-linear relationships. The data on which the model was calibrated and validated were robust (3,252 sites) stratified across 6 regions and 3 levels of watershed imperviousness. Analyses of model performance indicate that the model is reasonably precise (pseudo-r^2 = 0.62) and unbiased.

Although we understand why modeling prediction intervals is useful, it seems like this approach would complement a simpler approach of comparing observed CSCI values with the values predicted from a regular random forest regression based on these same predictors. The simple difference between observed and expected CSCI values represent the potential scope of improvement that could be achieved.

Miscellaneous questions in need of answers include:

1. Do the same set of predictor variables provide the same level of performance at all quantiles? If not, how would this information be used when interpreting the ‘contraints’ on upper and lower CSCI values.
2. Can you tease out what constraints can be mitigated and which ones cannot? This seems like a central issue vis-à-vis UAA or tiering decisions.

Comprehensiveness and limitations of the model

The current model uses 6 predictors: canal density, % impervious surface, % urbanized area, % area in agriculture, road density, and density of road stream crossings. The panel wondered if the model’s predictive accuracy and hence its utility may be compromised by environmental alterations that are not captured by STREAMCAT – e.g., hydrological modifications. The panel was somewhat confused regarding what breadth of land use factors (predictors) should ideally be included in the model. The degree of permanence of the land uses included in the model varies. To most accurately estimate the effects of truly permanently modified landscapes, the model should perhaps only include measures associated with urbanization and dams. But, if the intent is to better understand the relative effects of the suite of land use alterations that exist in California, it makes sense to include measures of agriculture, riparian condition, and forest management. As presented, the document seems to want the model to serve both functions, and the overall goal of its application and how it could be used for a variety of purposes needs more through development and explanation. The panel wondered how well the STREAMCAT coverage adequately represented alteration in riparian land cover, and whether it would be better measured by photo-imagery. We also thought STREAMCAT land use condition estimates should be validated with field (or Goggle Earth) observations to assess how much error in predictions could be associated with error in quantifying predictor values.

The panel could not tell how predictions, and hence interpretations and applications, might be affected by whether streams are perennial or nonperennial. This is an important issue because you say the model was developed for all streams and rivers in California, but differences in aquatic life associated with water permanence is not addressed in any formal way. This seems like it is a serious issue that needs resolving by better specifying the domain of streams the model applies to. If it truly applies to all streams and rivers, you need to specifically show that predictions and interpretations do not vary between perennial and nonperennial streams.

In general, we agree the model might be best used as a screening tool, but we still need to know how often it might mislead.

Miscellaneous questions related to model comprehensiveness and limitations:

1. What are the model limitations in both in terms of predictive performance and how it could inform policy in its current state of development?
2. Can the model be used in conjunction with the nutrient-stressor response models?

Value added and other uses of the tool

Uses for the on-line application beyond trying to predict biological condition are of interest. The tool might have utility in linking to flow-inundation models and identifying groundwater/surface water zones of gaining and losing stream reaches.

Can you add TN and TP to the suite of predictor variables in the model and examine partial dependence plots to help inform nutrient criteria? Perhaps classify stream reaches according to the level of constraint and fit TN/TP models within each class.

Stakeholder concerns

Most of the stakeholder concerns involved how the tool would be used within a policy context. Those questions are formally outside of the purview of the science advisory panel, but we note the following questions are important to address.

1. Where along the perennial-nonperennial continuum does the model not apply?
2. Will this tool ultimately be used for use attainability assessments? If so, how would it best be used for that purpose?

Miscellaneous

Some panel members thought the case study was interesting but may not be appropriate for journal article. However, the handling editor, while asking for significant revision, thought the combination of model development and case study was a strong aspect of the paper.