Quantifying biological constraints on stream integrity for classification and priorization

Marcus W. Beck (marcusb@sccwpr.org), Raphael D. Mazor (raphaelm@sccwrp.org), Scott Johnson (scott@aquaticbioassay.com), Phil Markle (phil@lacsd.org), Peter D. Ode (peter.ode@wildlife.ca.gov), Ryan Hill (hill.ryan@epa.gov), Eric D. Stein (erics@sccwrp.org)

Introduction

- Degraded biological condition in aquatic environments can occur for several reasons. Some causes of poor condition can be effectively managed, others cannot. Fixing or restoring sites can have varying expenses and assurance of outcomes. This is why we prioritize.
- One approach to prioritize is to identify locations where efforts are likely to have desired outcomes. This requires identifying biological constraints or limits on the potential range of biological conditions. Identifying an appropriate context for observed conditions can be used to prioritize. Context can be defined by models, expert knowledge, and/or defined value sets.
- We don't have good constraint tools to develop a context of expectation of what's possible at a site. This can help prioritize locations where management efforts will or will not have the intended outcomes.
- The goal of this study is to demonstrate application of a landscape model to classify and prioritize stream monitoring sites using estimated constraints on biological integrity. The model provides an estimate of context for biological condition that provides an expectation of what is likely to be achieved at a given site relative to large-scale drivers of stream health. The model was developed and applied to all stream reaches in California. A case study is used to demonstrate how the model can be used to classify and prioritize using guidance from a stakeholder group.

Methods

Study area and data sources

- Brief description of CA, stream types and designated uses, PSA regions, management interests (e.g., southern vs. northern CA)
- Streamcat database used to quantify watershed land use at all sites (Hill et al. 2016)
- Streamcat data linked to National Hydrography Dataset Plus (NHD) (USGS (US Geological Survey) 2014), reach as individual unit for model output
- California Stream Condition Index (CSCI) as measure of stream integrity (Ode et al. 2016; Mazor et al. 2016), brief description of index

Building and validating landscape models

A prediction model of the CSCI was built to estimate likely ranges of scores associated with land use gradients. Land use parameters were urgan and agricultural land cover in the stream catchment (STREAMCAT). The model is incomplete by design such that CSCI scores were modelled only in relation to landscape-level variables that are not easily targeted by management. The model provided an explanation of variation in scores related to constraints on biology and unexplained variation was considered representative of additional, unmeasured environmental variables that influence stream biointegrity. Maybe describe modelling approach in (Mazor et al. 2016) - which variables were used to develop CSCI.

Models were developed using quantile random forests to estimate a range of likely CSCI scores in different landscapes (Liaw and Wiener 2002; D. R. Cutler et al. 2007). The model predictions were used to describe where bioassessment targets are unlikely to be met or where streams are unlikely to be impacted. Calibration and validation data were selected as xyz.

Classifying streams and prioritizing sites

- Description of SGRRMP and stakeholder group
- Methods for estimating stream class - possibly/likely constrained, possibly/likely unconstrained, certainty and CSCI threshold, some sites were unclassified
- Methods for estimating site performance - over, expected, underperforming, discussion of site types
- Sensitivity analysis - how do classes, performance categories change with thresholds and certainty
- Prioritization of types - stakeholder involvement

Results

State-wide patterns

Where does the model perform well, how does performance vary with validation and calibration datasets.

What is the consistency of patterns? For example, percent stream miles as xyz by PSA.

Figure Statewide map.

Case study

San Gabriel River Regional Monitoring Program

Extent, classification, prioritization - probabilistic assessment to make broader conclusions.

Relationships with environmental variables for constrained/unconstrained locations. Maybe apply to hardened/non-hardened reaches in constrained locations.

What to do with unclassified streams - typical urban, typical ag.

Tables Priority by type, by perspective

Discussion

- What do priorities really mean? Depends on your interests, needs, values, etc.
- Link with engineered channels study.

Supplement

Online application.

References

Cutler, D. R., T. C. Edwards Jr., K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. "Random Forests for Classification in Ecology." *Ecology* 88 (11): 2783–92. doi:10.1890/07-0539.1.

Hill, R. A., M. H. Weber, S. G. Leibowitz, A. R. Olsen, and D. J. Thornbrugh. 2016. "The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States." *Journal of the American Water Resources Assocation* 52: 120–28. doi:10.1111/1752-1688.12372.

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by randomForest." *R News* 2 (3): 18–22. http://CRAN.R-project.org/doc/Rnews/.

Mazor, R. D., A. C. Rehn, P. R. Ode, M. Engeln, K. C. Schiff, E. D. Stein, D. J. Gillett, D. B. Herbst, and C. P. Hawkins. 2016. "Bioassessment in Complex Environments: Designing an Index for Consistent Meaning in Different Settings." *Freshwater Science* 35 (1): 249–71.

Ode, P. R., A. C. Rehn, R. D. Mazor, K. C. Schiff, E. D. Stein, J. T. May, L. R. Brown, et al. 2016. "Evaluating the Adequacy of a Reference-Site Pool for Ecological Assessments in Environmentally Complex Regions." *Freshwater Science* 35 (1): 237–48.

USGS (US Geological Survey). 2014. "National Hydrography Dataset available on the World Wide Web."