Prioritizing management goals for stream biological integrity within the context of landscape constraints

Marcus W. Beck ([marcusb@sccwrp.org](mailto:marcusb@sccwrp.org)), Raphael D. Mazor ([raphaelm@sccwrp.org](mailto:raphaelm@sccwrp.org)), Scott Johnson ([scott@aquaticbioassay.com](mailto:scott@aquaticbioassay.com)), Karin Wisenbaker ([karin@aquaticbioassay.com](mailto:karin@aquaticbioassay.com)), Joshua Westfall ([jwestfall@lacsd.org](mailto:jwestfall@lacsd.org)), Peter D. Ode ([peter.ode@wildlife.ca.gov](mailto:peter.ode@wildlife.ca.gov)), Ryan Hill ([hill.ryan@epa.gov](mailto:hill.ryan@epa.gov)), Chad Loflen ([Chad.Loflen@waterboards.ca.gov](mailto:Chad.Loflen@waterboards.ca.gov)), Martha Sutula ([marthas@sccwrp.org](mailto:marthas@sccwrp.org)), Eric D. Stein ([erics@sccwrp.org](mailto:erics@sccwrp.org))

# Abstract

Many streams are failing to achieve desired biological condition and require management decisions to restore designated uses. Some management goals may be impractical with limited resources, particularly in streams where large-scale changes on the landscape (e.g., urbanization) impose constraints on the upper limit of biological integrity. A statewide landscape model was developed that sets reasonable expectations for observed conditions within landscape constraints to prioritize management actions. The model provides a context for what is likely to be achieved at a given site independent of an actual bioassessment score. With this approach, sites can be ranked as over- or under-scoring relative to an expectation that is typical for the observed level of landscape alteration. We developed a visualization tool to compare observed bioassessment scores with modelled expectations to rapidly identify reaches that were scoring better or worse than expected. Using this tool, a group of regulators, dischargers, stormwater agencies, and environmental advocates from the San Gabriel River watershed (Los Angeles County, California) identified regions in the watershed with consistent patterns in bioassessment scores relative to expectations. Based on these patterns, they prioritized different management actions for each region. Sites in both developed and undeveloped areas that scored below expectations were prioritized for restoration; in contrast, restoration was not a priority at developed sites where scores were low but within expected ranges. Sites scoring better than expected were prioritized for enhanced protection, as well as additional monitoring. Interactive tools that connect landscape models with observed data can help set management goals appropriate for stakeholder needs and likely constraints on biological integrity. These tools can easily be applied to other locations where biological data are used to assess environmental condition.

# Introduction

* Degraded biological condition in streams can occur from individual or multiple stressors acting at different scales (Novotny et al. [2005](#ref-Novotny05); Townsend, Uhlmann, and Matthaei [2008](#ref-Townsend08); Leps et al. [2015](#ref-Leps15)).
* In many urban and agricultural areas the majority of stream miles are not healthy and in need of some level of management (cite SWAMP, SMC, NRSA)
* Unfortunately, there are not sufficient resources to restore all streams to reference conditions, nor is it practical (e.g., varying costs and challenges of urban stream restoration (Kenney et al. [2012](#ref-Kenney12); Shoredits and Clayton [2013](#ref-Shoredits13)))
* Need a way to comprehensively evaluate streams across large spatial scales for “management potential”. Biological filters act at different scales (Poff [1997](#ref-Poff97)) and we can use this information to describe an expectation for prioritization that is scale-specific. This allows establishment of reasonable expectations and prioritization of limited resources most effectively. Landscape-level constraints are particularly relevant for macroinvertebrate communities in streams (Sponseller, Benfield, and Valett [2001](#ref-Sponseller01))
* Once these large spatial scales are understood, sites can be prioritized by local managers to ensure resources are wisely allocated.
* Goal: demonstrate application of a landscape model to classify and prioritize stream monitoring sites using estimated constraints on biological integrity.
  + Build on knowledge and relationships developed through existing monitoring programs and apply that in a predictive manner across entire landscapes to inform decisions
  + Statewide application of the model - 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 by watershed using guidance from a regional stakeholder group. Specific questions that were addressed through the case study. Active stakeholder involvement was critical in applying the landscape models to define a framework for decision-making because priorities varied with management objectives.

# Methods

## Study area and data sources

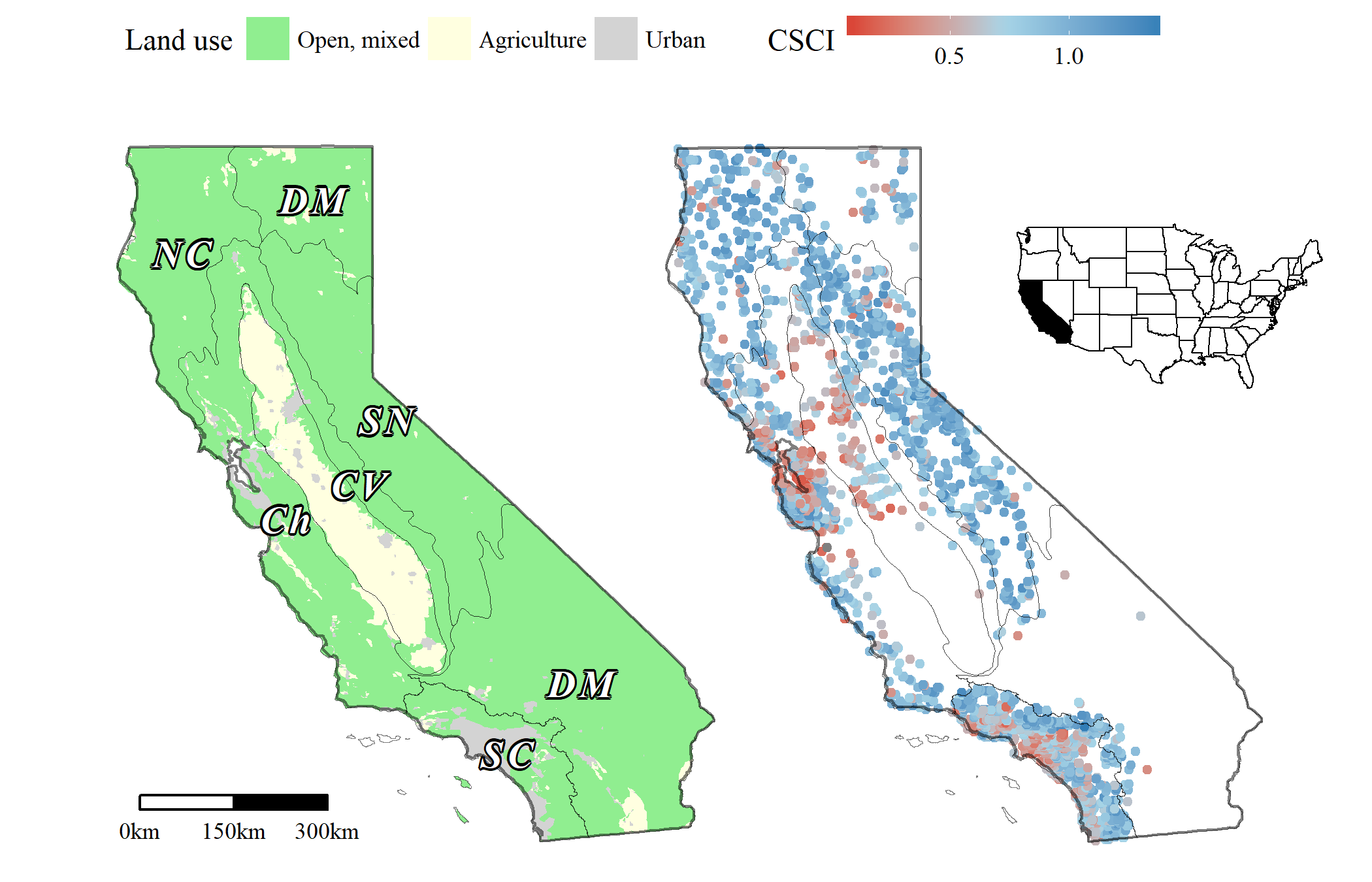


Figure 1 Land use (left) and distribution of stream CSCI scores (right) in California. Grey lines are ecoregions in California, CV: Central Valley, Ch: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

Landscape models were developed for California using land use data, stream hydrography, and biological assessments. California covers 424,000 km of land from latitudes 33 to 42N that includes extreme variation in altitude and climate (1). Temperate rainforests occur in the north, deserts in the northeast and southeast, and Mediterranean climates in coastal regions. California’s stream network is approximately 280,000 km in length and covers all of the major climate zones in the state. A high degree of endemism and biodiversity occurs in these streams including nearly 4000 species of vascular plants, macroinvertebrates, and vertebrates that depend on fresh water during their life history (J. Howard and Revenga [2000](#ref-Howard09); J. K. Howard et al. [2015](#ref-Howard15)). Approximately 30% of streams in California are perennial with the remaining as intermittent or ephemeral for portions of the year. Much of California is publicly owned and is used heavily for recreation. A large portion of the central region of the state is agricultural (i.e., Central Valley), whereas dense areas of urban development are in the southwest (i.e., Los Angeles and San Diego) and central (San Francisco Bay area) coast areas. Developed lands increased in California by 38% from 1973 to 2000 (Sleeter et al. [2011](#ref-Sleeter11)).

Stream data from the National Hydrography Dataset (NHD) (USGS (US Geological Survey) [2014](#ref-USGS14)) were used to identify reaches in California for modelling biological integrity. The NHD is a surface water framework that maps drainage networks and associated features (e.g., streams, lakes, canals, etc.) in the United States. Stream flow lines in the NHD are developed from flow accumulation models that estimate location of a stream given slope and elevation changes from existing elevation datasets. As such, flow lines in California represent both perennial, intermittent, and ephemeral streams that have wide variation in observed flow throughout the year. Stream reaches designated in the NHD were used as the discrete spatial unit for modelling biological integrity. A reach is defined as a continuous piece of surface water with similar hydrologic characteristics (USGS (US Geological Survey) [2014](#ref-USGS14)). Hydrography data were combined with landscape metrics available from the StreamCat Dataset (Hill et al. [2016](#ref-Hill16)) to estimate land use at the catchment (nearby landscape flowing directly into a stream) and the entire upstream watershed for each reach. The StreamCat Dataset was developed specifically for the NHD to leverage the topology of stream connections to estimate cumulative landscape metrics of all reaches.

The California Stream Condition Index (CSCI) (Ode et al. [2016](#ref-Ode16); Mazor et al. [2016](#ref-Mazor16)) was used as a measure of biological condition in California streams. Benthic macroinvertebrate data used to calculate CSCI scores were collected at nearly 3400 sites (6270 with repeat visits) between 2000 and 2016. Field data were collected during baseflow conditions typically between May and July following methods in Ode ([2007](#ref-Ode07)). The CSCI is a predictive index of stream health that compares the observed taxa and metrics at a site to those expected under reference conditions. Expected conditions at a site are based on models that estimate the likely macroinvertebrate community in relation to factors that naturally influence biology, e.g., watershed size, elevation, climate, etc. The CSCI score at a site is based on an observed-to-expected ratio of taxa and a predictive multimetric index composed of six individual metrics that describe the structure and function of the macroinvertebrate community. The index score at a site can vary from 0 to 1.4, with higher values indicating an observed community with less deviation from reference conditions. Because the index was developed to minimize the influence of natural gradients, the index scores have consistent meaning across the state (Reynoldson et al. [1997](#ref-Reynoldson97)). A threshold score based on a selected lower percentile of scores (e.g., 10%) at all reference sites is used to define nominally low and high scoring sites.

## Building and validating landscape models

A prediction model of the CSCI was developed to estimate likely ranges of scores associated with land use gradients. Land use as urban and agricultural was quantified for the catchment of each stream reach in California using the StreamCat database (Hill et al. [2016](#ref-Hill16)). CSCI scores were modelled using only the estimates of urban and agricultural land use as the developed portion of the landscape within each stream reach. The model was incomplete by design to describe scores only in relation to large-scale constraints on biological condition that are not easily controlled by management actions or where costs to mitigate are likely to be excessive. The remainder of the variation in scores not related to landscape constraints could be attributed to additional, unmeasured environmental variables that influence stream biointegrity. Deviation of observed scores from the model predictions were considered diagnostic of variation not related to landscape effects.

Models were developed using quantile regression forests to estimate ranges of likely CSCI scores in different landscapes (N. Meinshausen [2006](#ref-Meinshausen06); Nicolai Meinshausen [2017](#ref-Meinshausen17)). Quantile models evaluate the conditional response across the range of values that are expected, such as the lower and upper percentiles of the distribution, as compared to only the mean response with conventional models (Cade and Noon [2003](#ref-Cade03)). This allows use of model predictions to describe where bioassessment targets are unlikely to be met or where streams are unlikely to be impacted by placing bounds on the range of expectations relative to landscape constraints. Random forest models also provide robust predictions by evaluating different subsets of observations from random splits of the predictor variables. The final predictions are the averaged response across several models. These models have been used extensively in bioassessment applications (Carlisle, Falcone, and Meador [2009](#ref-Carlisle09); Chen et al. [2014](#ref-Chen14); Mazor et al. [2016](#ref-Mazor16)) and can produce unbiased estimates that are relatively invariant to noisy relationships or non-normal distributions (Breiman [2001](#ref-Breiman01); Hastie, Tibshirani, and Friedman [2009](#ref-Hastie09)). Quantile regression forests were used to predict CSCI scores in each stream reach from the 5th to the 95th percentile of expectations at five percent intervals (i.e., 5th, 10th, etc.).

Landscape estimates for the catchments of all NHD stream reaches in California were separated into calibration and validation data.

## San Gabriel River watershed case study

Stream reach and bioassessment data from the San Gabriel River (SGR) watershed in southern California were used to develop reach classifications, site performance categories, and management priorities from the landscape models. A strong land use gradient occurs in the SGR watershed (Figure 2). Headwaters begin in the San Gabriel mountains where the land is primarily undeveloped or protected for reacreational use, whereas the lower watershed is in a heavily urbanized region of Los Angeles County. The San Gabriel river is dammed at four locations for flood control in the upper watershed and is hydrologically connected to the Los Angeles river to the west through the Whittier Reservoir in the lower watershed. Spreading grounds are present in the middle of the watershed for groundwater recharge during high flow. Nearly all of the stream reaches in the lower half of the watershed are channelized with concrete or other reinforcements.

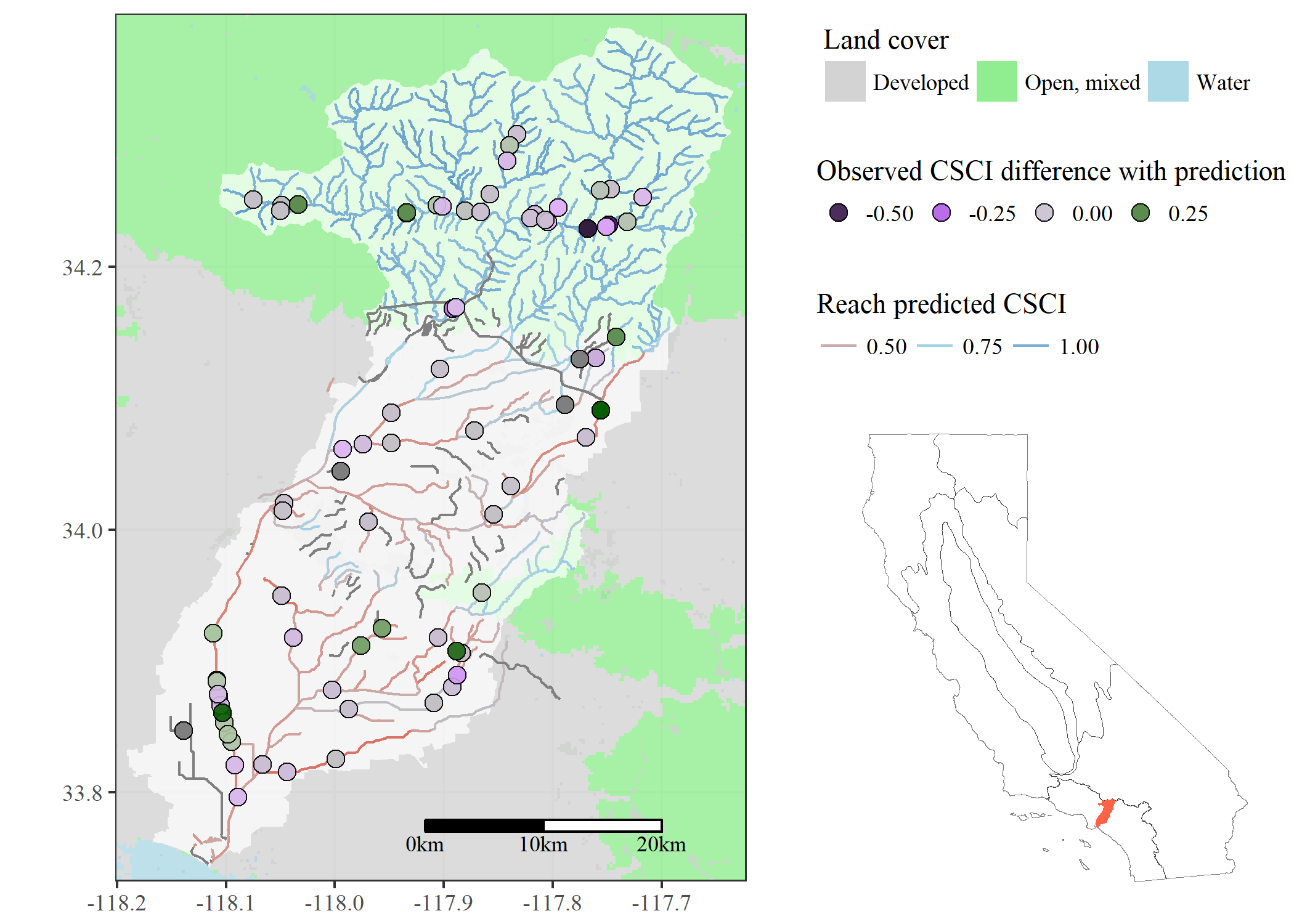


Figure 2 San Gabriel River watershed in southern California with predicted median CSCI scores at each stream reach and differences of observed CSCI scores with the predictions. The watershed is undeveloped in the north and heavily urbanized in the south. Gray stream reaches do not have model predictions because of insufficient data.

The SGR watershed contains a diverse group of stakeholders from local municipalities, water districts, water quality regulatory agencies, consulting groups, and non-government organizations. Collectively, the San Gabriel River Regional Monitoring Program (SGRRMP) includes stakeholders from these groups that cooperatively work to increase awareness of issues in the SGR watershed and work to improve coordination of compliance and ambient monitoring efforts. The stakeholder workgroup included individuals from the SGRRMP with interests in water supply, improvements to water quality, habitat protection or creation, and storm water permitting. Individuals were selected for partipation to include a variety of mangement interests and based on willingness to adopt tools developed from the landscape models. The stakeholder workgroup met monthly over a six-month period to discuss model applications and to refine the interpretation of results. Stakeholder involement was critical for developing an assessment framework that met the needs of all engaged parties and ensured that final products were more likely to be incorporated into formal processes of decision-making.

## Reach classification, site performance, and prioritization

A framework for identifying site priorities for management actions was developed using a three-step process. First, estimates of the range of expected CSCI scores at each stream reach in relation to land use were used to define reach classifications. Second, the relationship between observed CSCI scores and the reach classifications were then used to assign a relative performance value for each monitoring site. Third, site performance categories in relation to reach classification and bioassessment targets were used to define management priorities. This framework was developed through close interaction with the regional stakeholder group to demonstrate how the landscape model can be used as a management tool given that priorities will vary by interests and location. As such, the results are provided as a guide to facilitate decision-making rather than a prescription of targeted actions to manage stream health. The entire process is shown in Figure 3 and Table 1 and is described in detail below..

Identifying site priorities began with defining a classification framework for stream reaches to identify the possible or likely extent of biological constraints. Classifications were developed using the range of CSCI expectations at a reach (Figure 3a,b) relative to a chosen threshold for the CSCI to define nominally low or high scores (Figure 3c). The reach classification was based solely on the intersection of the CSCI expectations at a reach with chosen CSCI threshold, where expectations could be below, above, or overlapping the threshold. Stream reaches with a range of CSCI score expectations entirely below the thresholds were considered likely constrained, whereas those with expectations entirely above were considered likely unconstrained. Reaches with score expectations that included the CSCI thresholds were considered possibly constrained or possibly unconstrained, where the distinction was based on location of the median expectation of a reach relative to the threshold.

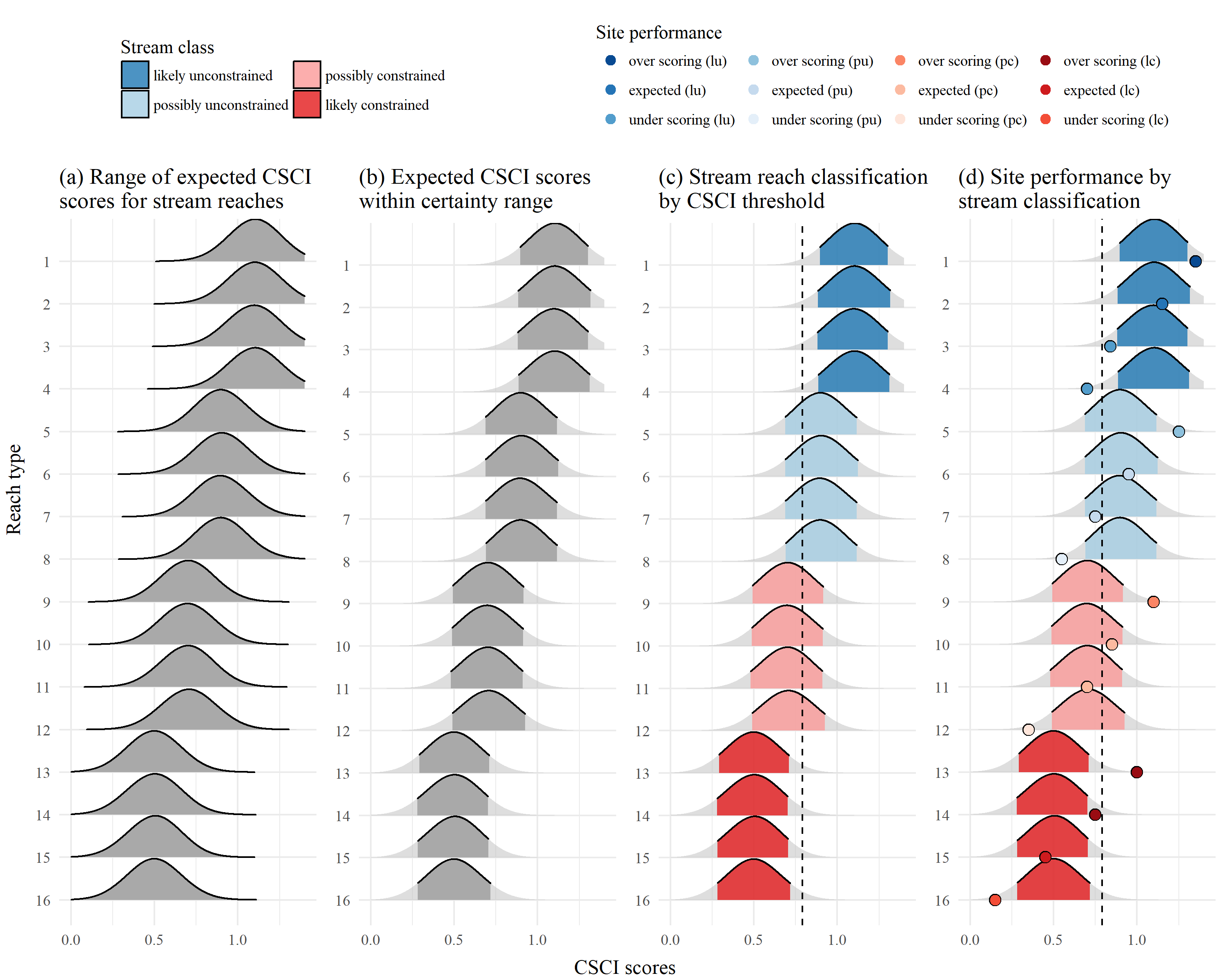


Figure 3 Application of landscape models to identify site expectations and bioassessment performance for sixteen example stream reaches. A range of CSCI scores is predicted from the model (a) and the lower and upper limits of the expectations are cut to define a certainty range for the predictions (b). Overlap of the certainty range at each reach with a chosen CSCI threshold (c) defines the stream reach classification as likely unconstrained (lu), possibly unconstrained (pu), possibly constrained (pc), and likely constrained (lc). The site performance from the observed bioassessment score is defined as over scoring (above the certainty threshold), expected (within), and under scoring (below) for each of four stream classes (d).

CSCI scores from biomonitoring data were used to define performance of a sample site relative to the stream reach classification (Figure 3d). For each of the four reach classifications (likely constrained, possibly constrained, possibly unconstrained, and likely unconstrained), the site performance was defined relative to the bounds of the expected CSCI scores. This provided a definition of site performance that can be used to understand the observed score relative to the biological context of a reach. Sites with observed scores above the upper limit of the reach expectation (e.g., above the 95th percentile of expected scores) were considered “over-performing” and sites below the lower limit were “under-performing”. Sites with CSCI scores within the range of expectations were as “expected”.

Site performance categories were further split relative to location to the selected CSCI threshold. This final split was created with the intent that description of site scores relative to a defined threshold (e.g., impairment threshold or restoration target) should also be considered. Specifically, a fourth category of site performance for each reach classifcation was added to define a site as above or below the threshold. For a likely unconstrained reach, underperforming sites below the minimum expected score were additionally defined as being above or below the CSCI threshold. Similarly, overperforming sites above the maximum expected score in a likely constrained reach were additionally defined as being below or above the CSCI threshold. For possibly constrained and possibly unconstrained reaches, sites that were performing as expected were addtionally defined as being below or above the CSCI threshold. In total, sixteen site types were defined for the three reach classification and three site performance classifications (Table 1).

Table 1 Possible site types based on stream reach classification, site performance, and observed CSCI score. The observed score column describes where a CSCI score is observed relative to the lower and upper percentiles (e.g., 5th and 95th) of expected scores for a reach and the chosen CSCI threshold (e.g., 10th percentile of scores at reference sites or 0.79) for nominally low or high values.

|  |  |  |  |
| --- | --- | --- | --- |
| Reach expectation | Site performance | Observed score | Type |
| **likely unconstrained** | over scoring | 95th | 1 |
|  | expected | 5th to 95th | 2 |
|  | under scoring | 0.79 to 5th | 3 |
|  | under scoring | < 0.79 | 4 |
| **possibly unconstrained** | over scoring | 95th | 5 |
|  | expected | 0.79 to 95th | 6 |
|  | expected | 5th to 0.79 | 7 |
|  | under scoring | < 5th | 8 |
| **possibly constrained** | over scoring | 95th | 9 |
|  | expected | 0.79 to 95th | 10 |
|  | expected | 5th to 0.79 | 11 |
|  | under scoring | < 5th | 12 |
| **likely constrained** | over scoring | 0.79 | 13 |
|  | over scoring | 95th to 0.79 | 14 |
|  | expected | 5th to 95th | 15 |
|  | under scoring | < 5th | 16 |

Each site type was used to define a priority as a demonstration of how results from the landscape model can help achieve different stream management objectives. This final process relied exclusively on feedback from the stakeholder group that represented interests in monitoring, regulation, restoration, and protection. An interactive online application was created for the stakeholder group to facilitate the recommendation of management actions for each site type (Figure 4). Priorities for each site type were defined accordingly with the expectation that site types will have different meanings for prioritization given the interest. Stakeholders from each sector were tasked with identifying their relevant priorities by ranking each site type from high to low priority using a blank template for reference. A brief description of the rationale for a site priority was also requested with the feedback. The final priorities were generalized into three categories to recommend actions in addition to baseline monitoring and maintenance. The final priorities also assumed that existing information available for each site was “true” following established practices to account for uncertainty or variation between assessments. A consensus was reached for the following definitions of each action:

* Investigate: Additional monitoring or review of supplementary data (e.g., aerial imagery);
* Protect: Additional scrutiny of proposed development and/or projects;
* Restore: Targeted action for causal assessment and/or restoration funds.

Each site type was ranked as high, medium, or low priority for each action. No priority assigned to an action for a stream type was indication that baseline monitoring and maintenance was sufficient for a site type.

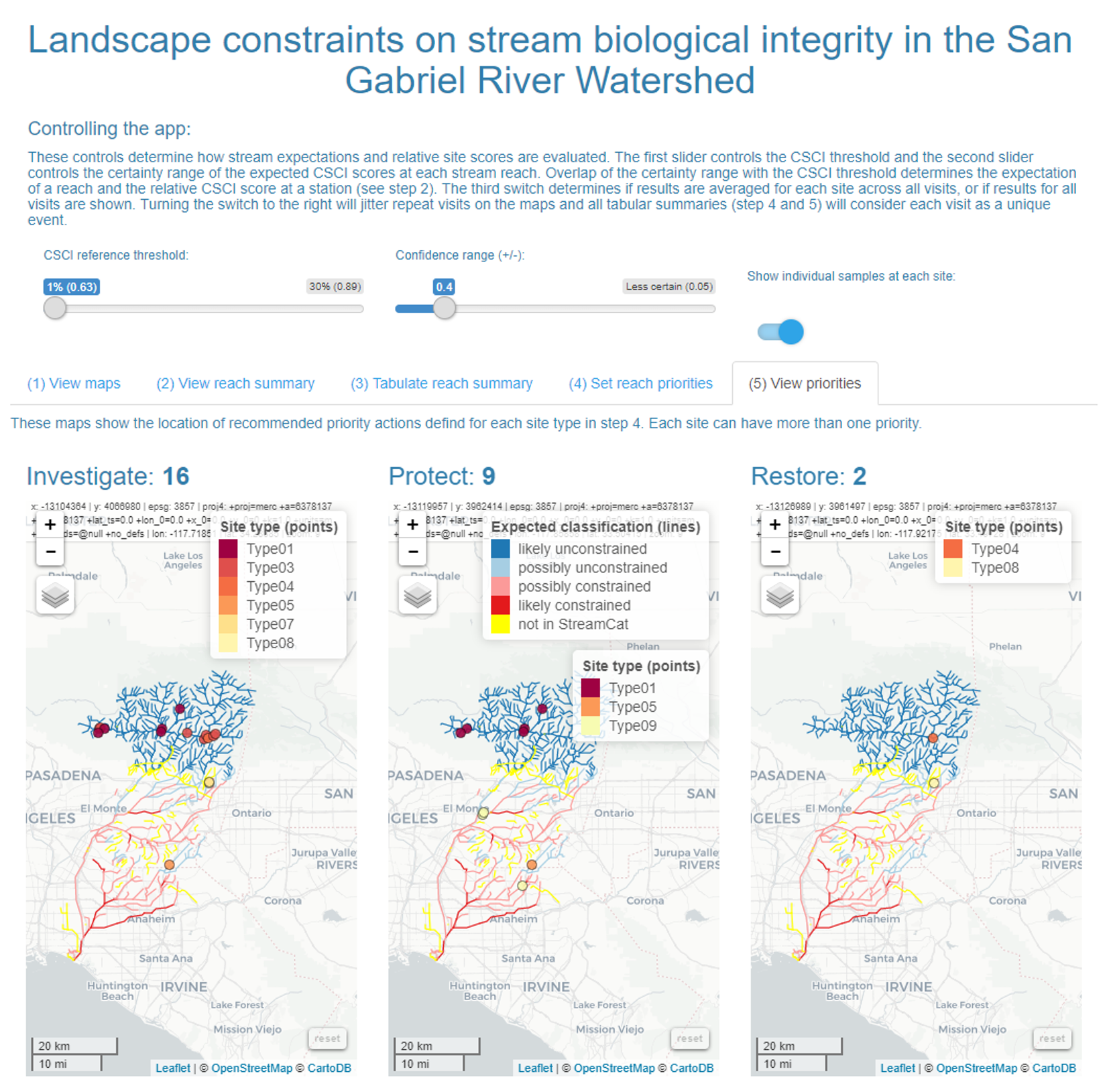


Figure 4 Screenshot of the online application used by the stakeholder group to interact with and use results from the landscape models. The application allowed users to visualize results of reach classications, relative site scores for the CSCI based on the expectation, and recommend management actions for each reach type. The app can be viewed at <https://beckmw.shinyapps.io/sgrrmp_classify/>

## Sensitivity analyses

Stream reach classifications and site performance categories depend on the range of score expectations from the landscape model (or certainty, Figure 3b) and the CSCI threshold for defining nominally low or high scores (Figure 3c). This framework for identifying priorities was developed to allow flexibility in how the model could be applied. First, the framework can accommodate degrees of certainty in the model by allowing variation in the range of scores that are used to define a stream reach classification. The 5th and 95th percentile of expected scores at a reach are used as a default range in which a high degree of certainty in the model output is assumed. The ability to reduce this range (e.g., 25th to 75th percentile) to assume less certainty in the model is provided. The CSCI threshold can also be changed to assess effects of relaxing or increasing flexibility in a potential definition of a regulatory standard. A threshold of 0.79 is used by default as a measure of the 10th percentile of scores at all reference (non-impacted) sites that were used to calibrate the CSCI index. This value can be increased to examine effects of a more stringent threshold or decreased for a more relaxed threshold. The combined effects of changing both the certainty in the model and the CSCI threshold were evaluated to estimate the changes in stream miles in each classification and the number of sites in each priority type.

## Unclassified reaches

Finally, some stream reaches were unclassifed following application of the landscape model to the statewide hydrography dataset. Unclassified reaches occurred when insufficient data in the StreamCat database were available to estimate CSCI predictions or if a stream catchment basin could not be defined for a particular reach. The latter was more common, particularly in developed areas where engineered channels or agricultural ditches were hydrologically removed from the natural stream network. Overall, unclassified reaches were not common in the statewide dataset but they may have regional importance depending on needs of local management groups. A preliminary approach for assigning biological expectations to unclassified reaches is demonstrated for ‘typically’ urban and agriculture reaches that relies on the range of expectations for reaches with similar land use by region.

# 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

* Extent, classification, prioritization
* Relationships with environmental variables for constrained/unconstrained locations. Maybe apply to hardened/non-hardened reaches in constrained locations.

*Figure* Summary of extent of reach classification, site performance, selected examples

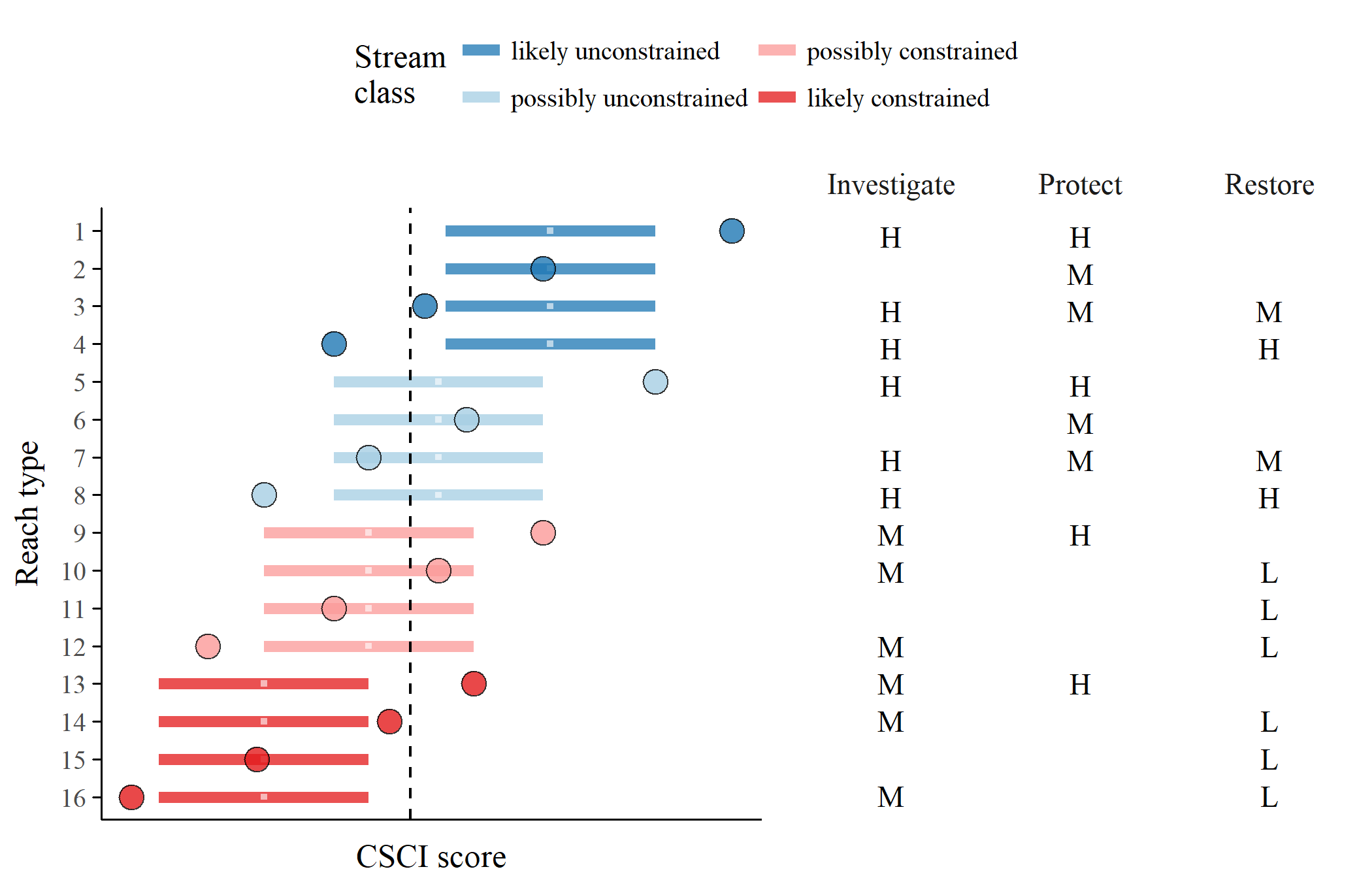


Figure 5 Template provided to stakeholders for priorization of recommended actions for each stream type. The stream types (Table 1) relate to the stream class for the biological expectation (lu: likely unconstrained, pu: possibly unconstrained, pc: possibly constrained, lc: likely constrained), site performance from observed CSCI score (os: over scoring, e: expected, us: under scoring), and location of the score relative to a hypothetical biological threshold (dashed line, above or below). Horizontal lines are the range of expected CSCI score for a site with tick marks for the median. Priority actions defined by stakeholders are shown on the right for each stream type. Actions are generalized as investigate, protect, or monitor as high (H), medium (M), or low (L) priority. Blank cells indicate that no additional measures are recommended beyond the baseline monitoring and maintenance practiced at all sites.

### Sensitivity analysis

* Statewide results - reach classification, site performance
* SGR application - where do priorities change related to which variables the model is most sensitive to? Do overall patterns remain? Maybe not do this.

*Table* Sensitivity results

### Unclassified reaches

* Extent of typical ag, typical urban statewide
* Framework for assigning unclassified reach to a class
* Statewide patterns, SGR patterns

*Table* Summary by location

# Discussion

* What is the value of identifying constrained channels?
  + Identification of constrained channels allows us to determine how best to spend limited resources and to focus on reaches where we have a decent chance of improving the biological condition.
  + Use of more data to develop context of assessment
  + Targeted management for desired outcomes
  + Informing decisions about future monitoring (i.e., prioritize future monitoring locations)
  + Listing implications, perhaps with a subheader, “Implications for determining biological impairments”. We should enlist Chad to help with this—after presenting him with LACSD’s proposal.
* What is useful about our approach compared to alternatives?
  + Field-based methods to identify constrained channels vs. landscape modelling
  + Related directly to biological condition and regulatory standards
  + Results are widely corraborated by other landscape studies - land use is big determinant of macroinvert assemblage
* What contributed to our success in defining priorities?
  + Stakeholder involvement guided process, contributed to achieving goals
  + An interactive/iterative approach was used - we provided tools to facilitate (web apps) and we did not assume priorities
* Caveats of our aproach
  + This analysis doesn’t truly tell us if a site can be fixed or if the conditions are truly constrained (key message, need to mention in intro)
  + What do priorities really mean? Depends on your interests, needs, values, etc.
  + Constrained may not always mean constrained - CSCI vs other biological indicators
  + Site-specific approaches are warranted in certain cases
  + Changing certainty or CSCI treshold - mechanistic effects and implications. Don’t cook the books.
* Future work
  + Ability to link with other assessment tools besides CSCI
  + Link with engineered channels study
  + Priorities statewide
  + Application to larger regions possible (national-scale), or how it can be applied in other areas

# Supplement

Online application.

# References

Breiman, L. 2001. “Random Forests.” *Machine Learning* 45: 5–32.

Cade, B. S., and B. R. Noon. 2003. “A Gentle Introduction to Quantile Regression for Ecologists.” *Frontiers in Ecology and the Environment* 1 (8): 412–20.

Carlisle, D. M., J. Falcone, and M. R. Meador. 2009. “Predicting the Biological Condition of Streams: Use of Geospatial Indicators of Natural and Anthropogenic Characteristics of Watersheds.” *Environmental Monitoring and Assessment* 151 (1-4): 143–60. doi:[10.1007/s10661-008-0256-z](https://doi.org/10.1007/s10661-008-0256-z).

Chen, K., R. M. Hughes, S. Xu, J. Zhang, D. Cai, and B. Wang. 2014. “Evaluating Performance of Macroinvertebrate-Based Adjusted and Unadjusted Multi-Metric Indices (MMI) Using Multi-Season and Multi-Year Samples.” *Ecological Indicators* 36: 142–51. doi:[10.1016/j.ecolind.2013.07.006](https://doi.org/10.1016/j.ecolind.2013.07.006).

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer.

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](https://doi.org/10.1111/1752-1688.12372).

Howard, J. K., K. R. Klausmeyer, K. A. Fesenmyer, J. Furnish, T. Gardali, T. Grantham, J. V. E. Katz, et al. 2015. “Patterns of Freshwater Species Richness, Endemism, and Vulnerability in California.” *PLOS ONE* 10 (7): e0130710. doi:[10.1371/journal.pone.0130710](https://doi.org/10.1371/journal.pone.0130710).

Howard, J., and C. Revenga. 2000. “California’s Freshwater Biodiveristy in a Continental Context. Science for Conservation Technical Brief Series.” San Francisco, CA: The Nature Conservancy of California.

Kenney, M. A., P. R. Wilcock, B. F. Hobbs, N. E. Flores, and D. C. Martínez. 2012. “Is Urban Stream Restoration Worth It?” *Journal of the American Water Resources Association* 48 (3): 603–15. doi:[10.1111/j.1752-1688.2011.00635.x](https://doi.org/10.1111/j.1752-1688.2011.00635.x).

Leps, M., J. D. Tonkin, V. Dahm, P. Haase, and A. Sundermann. 2015. “Disentangling Environmental Drivers of Benthic Invertebrate Assemblages: The Role of Spatial Scale and Riverscape Heterogeneity in a Multiple Stressor Environment.” *Science of the Total Environment* 536: 546–56. doi:[10.1016/j.scitotenv.2015.07.083](https://doi.org/10.1016/j.scitotenv.2015.07.083).

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.

Meinshausen, N. 2006. “Quantile Regression Forests.” *Journal of Machine Learning Research* 7: 983–99.

Meinshausen, Nicolai. 2017. *QuantregForest: Quantile Regression Forests*. <https://CRAN.R-project.org/package=quantregForest>.

Novotny, V., A. Bartosová, N. O’Reilly, and T. Ehlinger. 2005. “Unlocking the Relationship of Biotic Integrity of Impaired Waters to Anthropogenic Stresses.” *Water Research* 39 (1): 184–98. doi:[10.1016/j.watres.2004.09.002](https://doi.org/10.1016/j.watres.2004.09.002).

Ode, P. R. 2007. “Standard Operating Procedures for Collecting Benthic Macroinvertebrate Samples and Associated Physical and Chemical Data for Ambient Bioassessment in California.” Surface Water Ambient Monitoring Program. Sacramento, CA.

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.

Poff, N. L. 1997. “Landscape Filters and Species Traits: Towards Mechanistic Understanding and Prediction in Stream Ecology.” *Journal of the North American Benthological Society* 16 (2): 391–409.

Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. “The Reference Condition: A Comparison of Multimetric and Multivariate Approaches to Assess Water-Quality Impairment Using Benthic Macroinvertebrates.” *Journal of the North American Benthological Society* 16 (4): 833–52.

Shoredits, A. S., and J. A. Clayton. 2013. “Assessing the Practice and Challenges of Stream Restoration in Urbanized Environments of the USA.” *Geography Compass* 7 (5): 358–72. doi:[10.1111/gec3.12039](https://doi.org/10.1111/gec3.12039).

Sleeter, B. M., T. S. Wilson, C. E. Soulard, and J. Liu. 2011. “Estimation of the Late Twentieth Century Land-Cover Change in California.” *Environmental Monitoring and Assessment* 173 (1-4): 251–66. doi:[10.1007/s10661-010-1385-8](https://doi.org/10.1007/s10661-010-1385-8).

Sponseller, R. A., E. F. Benfield, and H. M. Valett. 2001. “Relationships Between Land Use, Spatial Scale and Stream Macroinvertebrate Communities.” *Freshwater Biology* 46 (10): 1409–24. doi:[10.1046/j.1365-2427.2001.00758.x](https://doi.org/10.1046/j.1365-2427.2001.00758.x).

Townsend, C. R., S. S. Uhlmann, and C. D. Matthaei. 2008. “Individual and Combined Responses of Stream Ecosystems to Multiple Stressors.” *Journal of Applied Ecology* 45 (6): 1810–9. doi:[10.1111/j.1365-2664.2008.01548.x](https://doi.org/10.1111/j.1365-2664.2008.01548.x).

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