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))

Version Date: Tue Jun 12 14:53:13 2018 -0700

# Abstract

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 provides context for a macroinvertebrate-based bioasssessment index by predicting a range of expectations that are typical at a site for the observed level of landscape alteration. With this approach, sites can be described as over- or under-scoring relative to an expectatation that is determined by landscape constraints on biological condition. A regional application of the landscape model was used to classify sites and prioritize different management actions by a local stakeholder group of regulators, dischargers, stormwater agencies, and environmental advocates from the San Gabriel River watershed (Los Angeles County, California). Stakeholder decisions were facilitated with the Stream Classification and Priority Explorer (SCAPE) tool that compares observed bioassessment scores with expectations from the landscape model to rapidly identify reaches that are scoring better or worse than expected. Of the 71 sites in the watershed where bioassessment occurred, over half (58%) were found to be constrained by landscape, of which nine were assigned a medium priority for further investigation by the stakeholder group. Similarly, 30% of sites were unconstrained and assigned medium or high priorities for enhanced protection or restoration depending on whether a site was scoring better or worse than expected. A clear gradient from the upper to lower watershed was observed in land use that was consistent with the model predictions, where landscape constraints were more common in the heavily urbanized lower watershed. Interaction with the local stakeholder group was critical in connecting the landscape model with observed data to help set management goals appropriate for the region. The availability of geospatial and bioassessment data at the national level suggests that these tools can easily be applied to inform management decisions at other locations where biological indices are used to assess environmental condition.

# Introduction

The widespread use of bioassessment data to assess ecological condition of aquatic environments is a significant advance over chemical or physical methods of assessment, yet managers require contextual information for synthesizing and interpreting biological information. The reference condition concept that is built into many biological indices is an explicit means of evaluating observed condition relative to unaltered habitats for a particular region (Reynoldson et al. [1997](#ref-Reynoldson97); Stoddard et al. [2006](#ref-Stoddard06)). However, in many cases the reference benchmark may not completely describe actual limits on biotic condition at spatial scales that can be effectively managed for biological health (Chessman and Royal [2004](#ref-Chessman04); Chessman [2014](#ref-Chessman14)). Landscape constraints in developed environments, such as modified urban streams, have led to alternative considerations on use attainability where reference conditions may be difficult or impractical to achieve given local conditions. Potential upper limits on biological health imposed by land use also have relevance for the application of bioassessment indices for biological criteria. A bioassessment index may be difficult to incorporate into regulatory statutes if thresholds for biological objectives are impractical within site-specific conditions. Managers require explicit information that can not only synthesize site level bioassessment data at the watershed scale, but also provide an assessment context that is sufficiently interpretable for prioritization.

The application of bioassessment data to inform management requires understanding the effects of multiple stressors acting at local, catchment, or watershed scales (Novotny et al. [2005](#ref-Novotny05); Townsend, Uhlmann, and Matthaei [2008](#ref-Townsend08); Leps et al. [2015](#ref-Leps15)). Nearly half of all streams and rivers in the USA are estimated to be in poor condition as related to the most commonly observed in-stream stressors, such as excess phosphorus, nitrogen, or altered physical habitat (USEPA (US Environmental Protection Agency) [2016](#ref-USEPA16)). These proximal and immediate causes of poor biological condition are often linked to landscape-level factors that occur in the watershed. In many urban and agricultural areas the majority of stream miles are not healthy and in need of some level of management (USGS (US Geological Survey) [1999](#ref-USGS99); Finkenbine, Atwater, and Mavinic [2000](#ref-Finkenbine00); Morgan and Cushman [2005](#ref-Morgan05)). Consistent and empirical links between land use thresholds and poor biotic integrity have been identified in many cases (Allan, Erickson, and Fay [2007](#ref-Allan97); Wang et al. [1997](#ref-Wang97); Clapcott et al. [2011](#ref-Clapcott11)). Mechanistic linkages between land use and degraded biological condition have been described (e.g., Allan ([2004](#ref-Allan04)), Riseng et al. ([2011](#ref-Riseng11))), whereas the precise link between land use and instream condition may not be clear for other causal pathways (e.g., Cormier et al. ([2013](#ref-Cormier13))). Regardless, land use has long been used as a proxy for environmental condition and an associative link can be sufficient to predict condition as a function of watershed activities.

Developing an expectation of biological potential as a function of land use constraints could help prioritize where management actions are most likely to achieve intended outcomes. Conventional approaches to restore or protect aquatic resources have commonly focused on direct improvements at the site-level to mitigate instream stressors (Carline and Walsh [2007](#ref-Carline07); Lester and Boulton [2008](#ref-Lester08); Roni and Beechi [2012](#ref-Roni12)), whereas upstream preventative measures have be incentivized or enforced through regulation. Although these approaches can lead to improvements in stream health, there is no universal remedy that applies to all stream types and conditions. For example, the required resources to restore a highly degraded stream in an urban or agricultural setting can be costly and it may be unreasonable to set regional reference conditions as a restoration target (Kenney et al. [2012](#ref-Kenney12); Shoredits and Clayton [2013](#ref-Shoredits13)). Understanding limits to biological potential is a fundamental concept in bioassessment that has recieved some attention. Methods for factor-ceiling analysis have been explored in bioassessment to characterize environmental factors that limit assemblage composition (Chessman, Muschal, and Royal [2008](#ref-Chessman08); Chessman [2014](#ref-Chessman14)). This approach is based on the limiting factor theory that proposes the most limiting biotic or abiotic factor will be the primary regulator of species abundance and distribution. Similar concepts have been applied in a landscape context to understand both variation in bioassessment data at different spatial scales and limits of bioassessment tools with land use gradients (Waite [2013](#ref-Waite13); Waite et al. [2014](#ref-Waite14)).

An approach to comprehensively evaluate the potential for success of alternative management scenarios for streams under landscape constraints could be developed with predictive models that estimate limits on expectations. Previous efforts have used geospatial data to predict biological condition at regional or national scales (Vølstad et al. [2004](#ref-Volstad04); Carlisle, Falcone, and Meador [2009](#ref-Carlisle09); Brown et al. [2012](#ref-Brown12); Hill et al. [2017](#ref-Hill17)), with the general purpose of characterizing biological condition at unsampled reaches. Landscape-level constraints are particularly relevant for stream macroinvertebrate communities (Sponseller, Benfield, and Valett [2001](#ref-Sponseller01); Waite [2013](#ref-Waite13)) and associations of constraints with conditions could be used to predict a range of expectations for biotic integrity. This approach could build on previous applications of landscape models by predicting a lower and upper estimate of what bioassessment index scores are likely relative to the landscape, in addition to estimating biological condition at unsampled reaches. Once the predicted response of macroinvertebrate communities to landscape changes at large spatial scales are understood, expectations can be compared to field samples and sites can be prioritized by local managers based on deviation from the expectation. As such, the development of contextual tools for understanding biological condition across landscape gradients could provide a powerful approach to informing the use of limited resources to manage stream integrity.

The goal of this study is to present the development of a landscape model to classify and prioritize stream monitoring sites and demonstrate its application to estimate constraints on biological integrity in California streams. This work builds on the knowledge and relationships developed through existing monitoring programs and applies that in a predictive manner across entire landscapes to inform management decisions. The specific objectives were to 1) demonstrate development of a landscape model to predict expected ranges of biotic condition, 2) classify stream reaches using modelling expectations, 3) assess the extent of stream classes and explore the sensitivity of the classifications to decision points in the model output, and 4) prioritize potential management decisions by comparing predicted ranges to observed bioassessment scores. The model was developed and applied to all stream reaches in California. A case study demonstrated how the statewide model can be used to classify and prioritize in a regional context using guidance from a local stakeholder group. An interactive software tool that was used by the stakeholder group to develop priority recommendations from the landscape model is also described.

# Methods

## Study area and data sources

The landscape model was 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 (Figure 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 (Howard and Revenga [2000](#ref-Howard09); 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. 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 Plus (NHD-plus) (USGS (US Geological Survey) [2014](#ref-USGS14)) were used to identify reaches in California for modelling biological integrity. The NHD-plus 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-plus 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-plus were used as the discrete spatial unit for modelling biological integrity. 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 (i.e., nearby landscape flowing directly into a stream segment) and the entire upstream watershed for each reach. The StreamCat Dataset was developed specifically for the NHD-plus to leverage the topology of stream connections to estimate cumulative landscape metrics of all reaches.

The California Stream Condition Index (CSCI) (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 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 values near 1 indicating an observed community similar to 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 the landscape model

A prediction model of the CSCI was developed to estimate likely ranges of scores associated with land use gradients. Measures of land use development were quantified for riparian, catchment, and watershed areas of each stream reach in California using the StreamCat database (Hill et al. [2016](#ref-Hill16)). CSCI scores were modelled using estimates of canal/ditch density, imperviousness, road density/crossings, and urban and agricultural land use for each stream reach (Table 1). These variables were chosen specifically to model scores only in relation to long-term constraints on biological condition that are typically beyond the scope of management intervention or where costs to mitigate are likely to be prohibitive. The remainder of the variation in scores not related to to these 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, although deviations could also be attributed to model error. The predictive performance of the models is described below.

Models were developed using quantile regression forests to estimate ranges of likely CSCI scores in different landscapes (Meinshausen [2006](#ref-Meinshausen06), [2017](#ref-Meinshausen17)). Random forests are an ensemble learning approach to predictive modelling that aggregates information from a large number of regression trees and 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); Fox et al. [2017](#ref-Fox17)). Random forest models provide robust predictions by evaluating complex, non-linear relationships and interactions between variables relative to more commonly-used modelling approaches (Breiman [2001](#ref-Breiman01); Hastie, Tibshirani, and Friedman [2009](#ref-Hastie09)). Quantile models, such as quantile regression forests, evaluate the conditional response across the range of values that are expected, in contrast to conventional models that provide only an estimate of the mean response (Cade and Noon [2003](#ref-Cade03)). This modelling approach allows use of 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. Quantile regression forests were used to predict CSCI scores in each stream reach from the 5th to the 95th percentiles of expectations at five percent intervals (i.e., 5th, 10th, etc.) up to the 45th and 55th percentiles.

Calibration data for the landscape model was based on a random selection of 75% of monitoring stations with observed CSCI scores. The random selection was stratified by major regions in California (Figure 1), with each further divided into quartiles based on watershed imperviousness. The stratification method was chosen to ensure sufficient representation of landscape gradients in each ecoregion. The remaining sites were used for model validation. Where multiple samples were available at a single site, one was selected at random for both calibration and validation purposes. Model performance was assessed for the statewide dataset and within each major region. Differences between observed CSCI scores and median predictions were evaluated using Pearson correlations and root mean squared errors (RMSE). Regression analysis between predicted and observed scores was used to assess potential bias based on intercept and slope values differing from 0 and 1, respectively.

## Statewide application of the landscape model

We applied the landscape model statewide and in major regions of California to estimate the extent of streams in different constraint classes. Using the model results, each stream reach was assigned to one of four classes (Table 2):

* Likely constrained
* Possibly constrained
* Possibly unconstrained
* Likely unconstrained

Classifications were based solely on the intersection of the modelled CSCI expectations at a reach with a chosen CSCI threshold, where expectations could be below, above, or overlapping the threshold. Two decision points were considered for their importance in defining reach classification:

1. The selected range of CSCI expectations at a reach from the model(Figure 3a,b), chosen as the 10th to 90th percentile of model predictions.
2. The CSCI threshold to define nominally low or high scores (Figure 3), chosen as 0.79 following previous examples (Mazor et al. [2016](#ref-Mazor16)).

Stream reaches with the range of CSCI score expectations entirely below the threshold were considered likely constrained, whereas those with expectations entirely above were considered likely unconstrained. Reaches with score expectations that overlapped the CSCI threshold were considered possibly constrained or possibly unconstrained, where distinction between the two was based on location of the median expectation of a reach relative to the threshold (Table 2).

CSCI scores from biomonitoring data were then used to define relative sites scores at a sample site given the stream reach classification (Figure 3d). This provided a definition to understand the observed score relative to the biological context (i.e., expectation) of a reach. For each of the four reach classifications (likely constrained, possibly constrained, possibly unconstrained, and likely unconstrained), relative site scores were defined based on location to the range of expected CSCI scores. Sites with observed scores above the upper limit of the reach expectation (e.g., above the 90th percentile of expected scores) were considered “over scoring” and sites below the lower limit (e.g., 10th percentile) were “under scoring”. Sites with CSCI scores within the range of expectations were scoring as “expected” within the context of the landscape model.

## Landscape factors associated with constraints

Factors explaining variation between constrained and unconstrained stream reaches were evaluated for the major regions in California (Figure 1). Only a select subset of variables in StreamCat were used to develop the landscape model, with the purpose of describing long-term and broad scale constraints on biointegrity. Additional landscape measures in StreamCat were evaluated to provide additional insight into alternative factors within each region that were associated with constraints on stream integrity. Landscape and geological data in StreamCat at the riparian and watershed scale were used to model variation between reach classes using random forest models (Breiman [2001](#ref-Breiman01)). For each region, 1000 regression trees were created and the mean reduction in accuracy was estimated for the exclusion of each variable across all models. This created an estimate of importance of each variable for describing differences between constrained and unconstrained stream classes. Mean reduction in accuracy was estimated for all variables in each model to identify the top five important variables in each region. Reach classes as possibly or likely constrained (or unconstrained) were combined to evaluate the complete dataset.

## Sensitivity analysis of reach classifications

A sensitivity analysis was conducted to evaluate the influence of key decision points on the extent of reach classifications statewide and by major regions in California. Stream reach classifications depend on the chosen range of score expectations (or certainty) from the landscape model (Figure 3b) and the CSCI threshold for defining nominally low or high scores (Figure 3c). The 10th and 90th percentile of expected scores at a reach were used as a default range in which a high degree of certainty in the model output is assumed. The effect of reducing this range (e.g., 25th to 75th percentile) to assume less certainty in the model was evaluated. Similary, the CSCI threshold was also changed to assess effects of relaxing or increasing flexibility in defining low or high scores. A threshold of 0.79 was 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 (Mazor et al. [2016](#ref-Mazor16)). This value was increased to examine effects of a more conservative threshold and 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 extent in each classification.

## Unclassified reaches

Finally, some stream reaches were not classifed following application of the landscape model to the statewide hydrography dataset. Unclassified reaches occurred if StreamCat data were unavailable to estimate CSCI predictions or if a reach was excluded from the NHD-plus dataset (typically, small headwater streams). The former 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. An approach for assigning biological expectations to unclassified reaches is demonstrated for “typically” urban and agricultural reaches that relies on the range of expectations for reaches with similar land use by region. The approximate range of scores for reaches dominated by either urban, agricultural, and neither of the two (other) was identified by using kmeans clustering of percentage land use for each reach (MacQueen [1967](#ref-MacQueen67)). Typically urban and agricultural reaches were identified based on the largest centroid average of the clusters for each land use type, whereas the other category was based on the minimum sum of the centroid values for urban and agricultural land use. The expected ranges for each land use type in each region were based on averages of ranges across all reaches in the identified clusters.

## 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 model. 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.

A diverse group of stakeholders from local municipalities, water districts, water quality regulatory agencies, consulting groups, and non-government organizations have shared interests in managing aquatic resources in the watershed. 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 management. Individuals were selected for participation to include diverse management interests and based on willingness to adopt tools developed from the landscape model. The stakeholder workgroup met monthly over a six-month period to discuss model applications and to refine the interpretation of results.

## Prioritization

A framework for identifying site priorities for management actions was developed using….

Categories for relative sites scores were further split if they were above or below 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 for sites within each reach classification was added to define a site as above or below the threshold. For a likely unconstrained reach, under-scoring sites below the minimum expected score were additionally defined as being above or below the CSCI threshold. Similarly, over scoring 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 as expected were additionally defined as being below or above the CSCI threshold. In total, sixteen site types were defined for the reach classifications (Table 3).

Each site type was used to define a priority as a demonstration of how results from the landscape model can inform different stream management objectives. This final process relied exclusively on feedback from the stakeholder group. The interactive and online Stream Classification and Priority Explorer (SCAPE) tool 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 (e.g., monitoring, regulation, etc.). Stakeholders were tasked with identifying their relevant priorities by ranking each site type from high to low priority using a graphical template of Table 3 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., field visits, review aerial imagery);
* Protect: Additional scrutiny of any proposed development and/or projects;
* Restore: Targeted action for causal assessment and/or restoration activity.

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.

# Results

## State-wide patterns

The landscape model was used to predict an expected range of CSCI scores for 138716 stream reaches in California. The bioassessment dataset used to develop the model included 2620 unique field observations assigned to stream reaches in the NHD-plus dataset. By region, the most bioassessment samples were observed in the South Coast (n = 839), followed by the the Chapparal (n = 684) and Sierra Nevada regions (n = 548). Model performance statewide and by region indicated generally good agreement between observed CSCI scores and the median prediction for the associated stream reach (Table 4). Agreement between observed and predicted values for the entire calibration dataset was 0.84 and RMSE = 0.14. The intercept and slope for a regression between observed and predicted values were 0.24 and 0.72, suggesting a slight negative bias of predictions at lower scores and slight positive bias at higher scores. The statewide calibration data showed similar results, with slightly smaller correlation ( 0.72) and larger RMSE (0.18) estimates.

Model performance differed by region. Performance for the Chapparal and South Coast regions were comparable or slightly improved compared to the statewide dataset for both the calibration and validation datasets. Model predictions for the Central Valley, Desert Modoc, and North Coast regions had slightly lower performance compared to the statewide results, with correlations of approximately 0.75 with observed values in the calibration dataset and 0.55 in the validation dataset. Model performance was poor for the Sierra Nevada region where the lowest correlation between predicted and observed values was observed. Regression estimates for the validation data in this region showed an intercept and slope of approximately 1 and 0. Overall, model performance was strongly associated with land use gradients in each region (Figure 5). The landscape model peformed well in regions with a mix of urban, agricultural, and open land, such as the South Coast, where strong gradients occur in many watersheds. Conversely, the model did not perform well in regions where developed landscapes were less common, such as the Sierra Nevada region.

Statewide patterns in stream constraints were apparent from the results of the landscape model consistent with land use (Figure 6). Stream reaches were more often constrained for biotic integrity in regions with more watershed development, either as urban or agricultural land. For example, likely constrained reaches were apparent from the statewide map in coastal reaches of the South Coast where heavy urbanization occurs and in the Central Valley where agriculture is the dominant land use. Stream reaches were more likely to be unconstrained in regions with less development, with areas in the North Coast and the Sierra Nevada region visible on the map (right, Figure 6). A majority of reaches statewide were classified as possibly constrained (11% of all stream length) or possibly unconstrained (46%), whereas a minority were likely constrained (4%) or likely unconstrained (39%) (Table 6). By region, the most reaches classified as likely unconstrained reaches were in the Sierra Nevad (50%), North Coast (46%), and Desert Modoc (46%) regions, whereas the most reaches as likely constrained were observed in the Central Valley (22%) and South Coast (15%) regions. Relative CSCI scores compared to reach expectations were as expected for 80% of the sampled locations statewide, whereas a much smaller percentage of sites were equally under or over scoring (Table 6). Similar patterns were observed within regions, although a slightly larger percentage of sites in the Central Valley were under scoring compared to the other regions.

## Case study

Application of the landscape model to the San Gabriel River watershed and engagement with a stakeholder group demonstrated how the results can be used to locally prioritize actions for different stream reaches. About 750 reaches in the SGR were identified and classified from NHD-plus, of which 10% were visited for bioassessment sampling. CSCI scores ranged from 0.2 to 1.23 consistent with heavy urban development in the lower watershed and open land use at higher elevation in the upper watershed (Figure 7). Application of the landscape model results to the CSCI scores provided a context of expectations consistent with the strong land use gradient in the watershed. Stream reaches in the upper watershed were a mix of likely and possibly unconstrained, whereas all stream reaches in the lower watershed were classified as likely constrained. Several reaches in the lower watershed had ranges that were right-skewed toward very low CSCI scores consistent with extreme landscape pressures.

Most of the sampling stations in the SGR watershed were within the expected ranges of CSCI scores for the defined stream classes (top, Figure 8). Using a CSCI threshold based on the 10th percentile of reference calibration sites (i.e., 0.79, Mazor et al. ([2016](#ref-Mazor16))) and a relatively certain range of expected scores from the 10th to the 90th percentile of the model predictions, only six sites were under-scoring (two likely unconstrained and four likely constrained) and eight sites were over-scoring (five likely constrained, one possibly unconstrained, and two likely unconstrained). One of the under-scoring sites in the likely unconstrained class was below the hypothetical CSCI threshold. One site scoring as expected in the possibly unconstrained class was below the chosen CSCI threshold, whereas none of the constrained (possibly or likely) sites were above the threshold.

The three priority management actions identified by the stakeholder group (investigate, protect, restore) were ranked from low, medium, to high for each potential reach type (Figure 9). High priority recommendations were generally given to over and under scoring sites in likely unconstrained reaches or those below the hypothetical threshold with possibly unconstrained classification. Low priority actions were most often recommended for possibly and likely constrained sites, or no action was recommended where it was assumed baseline monitoring and maintenance was sufficient. Recommended actions to investigate were more common for both over scoring and under scoring sites, protect was more common at over scoring sites, and restore was more common at under scoring sites. A clear distinction between low and high priority actions was observed on the watershed map (bottom, Figure 8). Sites in the lower watershed were lower priority if an action was recommended, whereas the four high priority sites were in the upper watershed. Several sites that were scoring as expected for likely and possibly unconstrained reaches in the upper watershed were recommended as medium priority for protection.

## Drivers of biological constraints and sensitivity analysis

Importance measures from random forest models identified key variables that explained the differences between constrained and unconstrained reaches between each region (Figure 10, see Figure for importance measures of the selected measures in the statewide landscape model). Relative magnitudes of the importance measures between regions confirmed the estimates of model performance, such that regions where the model performed well (e.g., South Coast, Central Valley) had higher importance measures than those where the model did not perform well (e.g., North Coast, Sierra Nevada). The top five most important variables were similar between regions although some specific differences were observed. The amount of biological nitrogen fixation in watershed soils was ranked the most important variable for the Central Valley, Desert Modoc, and North Coast region and second most important for the Chapparal and Sierra Nevada regions. This variable was not in the top five for the South Coast region, which was exclusively described by imperviousness and urbanization. Soil erodibility was the most important variable in the Chaparral and Sierra Nevada regions. Other important variables that were shared between regions (excluding the South Coast) were fertilizer applications and the amount of crops and hay at the riparian and watershed scale.

The effects of changing certainty in the CSCI predictions and a hypothetical CSCI threshold were evaluated for twenty-seven scenarios to identify the change in stream reach classifications (Figure 12). Changes were evaluated as the percentage of stream length statewide and within major regions for each reach classification. Decreasing the certainty of predictions from the landscape model by choosing a narrower range of scores (5th/95th to 45th/55th at 5% intervals) increased the number of streams from the possible to likely category in both constrained and unconstrained reaches. Similarly, changing the CSCI threshold from relaxed to more conservative (0.63, 0.79, 0.92) increased the number of streams classified as possibly or likely constrained and decreased the number of streams as possibly or likely unconstrained. Changes by region with the different scenarios were also observed. For example, over 80% of reaches in the Central Valley were classified as likely constrained using a conservative CSCI threshold with low certainty of predictions, whereas less than 1% of reaches were in this category using a relaxed CSCI threshold with the highest level of certainty. Opposite trends were observed in regions with reduced land use pressures. For example, almost all stream reaches in the North Coast and Sierra Nevada regions were classified as likely unconstrained using a relaxed CSCI threshold and low certainty of predictions.

## Unclassified reaches

Ranges of expected CSCI scores for typical reaches in urban, agricultural, and neither of the two (other) are shown in table 7. These typical values are shown for more to less certainty in the range of predictions. Unclassified reaches can be defined by the dominant watershed land use as urban, agricultural, or other, and then matched to the appropriate values in the table. Between regions, the variation in expected scores also provides context for landscape pressures that differ by location. For example, the expected range of scores in regions with heavy urban development (e.g., South Coast) are much lower than streams that are neither urban nor agricultural. The North Coast region in contrast has an expected range of scores in urban streams that is similar to streams that are neither urban nor agricultural. Interestingly, the range of scores in urban and agricultural streams were similar in the Central Valley where agriculture is the dominant land use.

# Discussion

The landscape model was able to identify the location and extent of biologically constrained channels at the state level and major regions in California. Our application to the SGR watershed demonstrated how the results of the model can be used at the scale of an individual watershed through close interaction with a regional stakeholder group with direct interests in managing the local resources. Overall, the models provide tools that allow managers to determine how best to use limited resources for stream management by focusing on reaches where recommended actions are most likely to have the intended outcome of improving or protecting biological condition. The approach also leverages all available information to develop a context for biological assessment that provides an unambiguous expectation of what is likely to be achieved at any sampling location. This can facilitate more targeted management actions that vary depending on the identified context and can also inform decisions on extent and effort for future monitoring locations.

Results from our analyses also have implications for managing biological impairments under state or federal water quality mandates. The landscape model could be used to evaluate the list of sites that are not meeting biological objectives by identifying locations where constraints may limit options for restoration. This can provide flexibility by focusing efforts in locations that can be most effectively targeted for actions such as TMDL limits or review of permitting. Further, the ability to evaluate the effects of changing proposed biological thresholds (e.g., tenth percentile of scores at reference sites) and certainty in the model conclusions (i.e., range of biological expectations) on the location and extent of constrained channels provides a means of choosing alternative scenarios for rule-making. A critical objective in allowing this flexibility is not to enable a discounts against sites that are less likely to achieve potential criteria, but rather to facilitate the decision-making process through a more transparent application of the model. Our results showing the change in stream length statewide and by region as related to potential thresholds and model certainty is a direct demonstration of this concept. This exercise could also be downscaled to an individual watershed to aid in rule-making.

What does the model mean vis a vis Pete’s comments - what variables did we include and how can we improve on this for more resolution depending on the region. Bias toward urban streams? What about time scale of impacts? What about possibility of model error?

## Landscape constraints can be described using biological endpoints

Our approach has several advantages compared to alternative means of identifying constraints on biological condition. The landscape model relies exclusively on publically available data using geospatial methods to identify potential constraints. An alternative is a field-based approach using site visits to characterize likely constraints at a location. While both approaches have strengths, the field-based approach is, by definition, labor-intensive and not appropriate as a screening tool to identify patterns on the landscape. More appropriately, the tools provided by the landscape approach could be used to inform decisions on where site visits may be needed to corraborate information from the model. The identification of constrained channels in urban areas that rely on field visits may also define the problem too narrowly. Physical habitat can be extremely limited in engineered channels and our models identified many of these locations in our case study. However, these channels were identified as constrained based on land use only. Constrained channels not in urban landscapes were also identified by the model, as for many streams in agricultural areas. The ability of the model to identify these locations was not accidental given the landscape variables that were used to develop the bioassessment predictions. In the context of the model, a constrained channel may or may not be engineered, but an engineered channel will typically be classified as constrained given the surrounding land use. These results are well supported by other landscape studies, particularly for macroinvertebrates (May et al. [2015](#ref-May15)).

Our approach also has the advantage of relating strictly to biological condition, as compared to physical or chemical endpoints to assess constraints. This has relevance from the perspective of ecological interpretation as well as potential implications for regulatory standards. The ability of the landscape model to predict the range of expected biological condition at a given site reflects an associative link between land use and stream biology. However, similar arguments that have been made for the use of biological indicators over chemical and physical indicators for assessment can be applied to identifying constraints with the landscape model. A limited range expectations at the lower end of the distibution of CSCI scores is an indication that stressors originating from the landscape have imposed habitat limits that constrain biology. A landscape model that is calibrated for physical or chemical endpoints may not sufficiently describe condition given that a constraint on either may not adequately characterize a constraint on biology. From a regulatory perspective, many states, including California, have explicit assessment requirements that relate to biology and the landscape modelling approach is well aligned with existing bioassessment tools. The use of biological endpoints in the landscape model will likely facilitate the development of biological standards as noted above. Landscape models could also be used to support conservation planning, particularly at the watershed scale. Ongoing work in California has focused on setting priorities for managing biodiversity that focus on watersheds within a conservation network (Howard et al. [2018](#ref-Howard18)). Results from the landscape model could be used to enhance this network by providing supporting information on constraints in an assessment framework.

## Engagement of local stakeholders is critical for regional application

Application of the landscape model to define potential management actions was effectively enabled through close interaction with our regional stakeholder group. The group was convened specifically to explore how the statewide model could be applied to inform decision-making at the watershed-level. The identified actions for monitoring sites in the San Gabriel watershed was chosen by discussion among the stakeholders that represented different interests (i.e., regulators, dischargers, consultants, and environmental advocates). The final decision by the group to prioritize management actions for the different sites in broad categories of protect, restore, and monitor was based on an iterative process where ideas were discussed and shared freely among stakeholders. Priorities were never assumed such that the recommended actions were defined through iterative discussion by the group. This approach ensured that stakeholders were generally in agreement with the final product and, therefore, more likely to adopt the recommendations provided by these tools in formal decision-making. The recommended actions have relevance only in the context of interests of the San Gabriel River Regional Monitoring Program. Application of the statewide model to other watersheds must engage local stakeholders in a similar process to develop recommendations that are adequate to meet regional needs.

The development of the SCAPE tool (see supplement) was also critical for engaging the stakeholder group. The tool was developed to achieve the dual purposes of demonstrating concepts applied by the model and allowing stakeholders to iteratively evaluate scenarios for defining stream classifications and priorities. Initial challanges in engaging stakeholders were encountered wherein a complete understanding of concepts, as well as limitations, of the model was insufficient for discussing regional application to the SGR. The tool provided a means of demonstrating these concepts that was not possible through more conventional methods, such as static presentations or simple text vignettes. SCAPE also allowed stakeholders to explore the different scenarios that influence the model output, specifically related to changing certainties in the CSCI score predictions and the ability explore alternative thresholds for biological objectives. This functionality is key in our approach to allow the stakeholders to develop recommendations that are completely independent of the model, i.e., decisions are not hard-wired into the model nor SCAPE. Application of the landscape modelling approach to other regions will benefit from similar tools that can easily be developed.

## The landscape model is a tool for exploring options

The primary objective of developing the landscape model was to provide a screening tool for exploring biological constraints to facilitate a discussion of management options relative to site contexts. These models are not intended for developing regulatory designations for individual sites, nor are they sufficient by themselves as a use attainability assessment at constrained sites. Specifically, the landscape model is associative by design and does not identify mechanistic links between biological constraints and proximal causes. Further, a distinction between constraints on biological condition and channel modification is implicit such that indication of the former by the model does not explicitly indicate presence of the latter. As noted above, our results consistently indicated that engineered channels are biologically constrained but the model is based on an a priori selection of land use variables to predict biotic integrity. We have explicitly not incorporated channel modification as a predictive mechanism by which biological constraints may occur, although habitat limitations imposed by modification may impose real limits on biology. Alternatively, bioasessment scores can be similar in modified channels compared to natural streams independent of watershed land use (Stein et al. [2013](#ref-Stein13)). The landscape model describes constraints at scales larger than instream characteristics as a necessary approach to accurately predict bioassessment scores. More comprehensive assessments at individual sites are needed to diagnose the immediate causes of degraded condition.

An additional consideration in applying the model to inform management decisions is the meaning of biologically constrained in the context of macroinvertebrate communities. Biologically constrained sites were considered those where landscapes were likely to limit the bioassessment index. The CSCI provides an indication of stream health by estimating the macroinvertebrate community that is likely to be present under factors that naturally influence biology for the region and comparing the observed community to the estimate. A constrained site as estimated by the landscape model is only constrained relative to the macroinvertebrate community. In many cases, poor biotic condition of the macroinvertebrate community translates to poor stream condition. However, a constrained macroinvertebrate community does not always mean other aspects diagnostic of stream health are also constrained. Many urban streams can support diverse algal assemblages such that algal-based measures of biotic condition may alternatively suggest good biotic condition relative to macroinvertebrate-based indices. This limitation of the landscape model is not unique to other bioassessment tools and application to other taxa as alternatives lines of evidence is needed for a more complete condition assessment.

## Additional applications of the landscape model

This work has immediate relevance for informing use of bioassessment indices in freshwater streams of California, both from the perspective of identifying management priorities and for developing an interpretation of broad scale factors that limit ecological integrity. Our use of geospatial datasets that are descriptive at the national scale suggests that these methods could be applied elsewhere and in unique bioassessment contexts. The NHD-plus and StreamCat datasets cover the entire hydrologic network of the conterminous United States and national-scale bioassessment tools are available through the National Rivers and Streams Assessment (USEPA (US Environmental Protection Agency) [2016](#ref-USEPA16); Hill et al. [2017](#ref-Hill17)). Statewide extension of the landscape model could be explored to develop a national scale product of constraints on biological condition. Although landscape stressors are ubiquituous at the contintental scale, extension of the landscape modelling approach beyond California should consider regional variation in landscape stressors that are predictive of biotic condition. Urban and agricultural gradients were sufficient to characterize constraints in many regions of California, whereas the model did not perform well in regions where thes gradients did not occur. A national level landscape model must consider alternative predictors provided by the StreamCat model that are sufficient to model biological condition.

## Summary

The prevalence of degraded stream sites in California requires the use of 1) assessment tools that can accurately evaluate condition, and 2) tools that can provide a context for evaluating assessment tools. The landscape model was developed to better inform application of the CSCI to inform decision-making in the context of landscape constraints on biological condition. Statewide development of the tool demonstrated where streams are likely constrained on a regional basis, whereas application to the SGR watershed demonstrated how the tool can be used by local stakeholders to prioritize management actions that are informed by landscape context. Most importantly, this tool does not provide a prescription for causes of impairment, nor does it discount sites from management intervention if constraints are high. The landscape model can inform the interpretation of biotic condition and is an exploratory tool that can help identify where management goals are most likely to be achieved.

# Supplement

The SCAPE model application website: <http://shiny.sccwrp.org/scape/>

# Figures

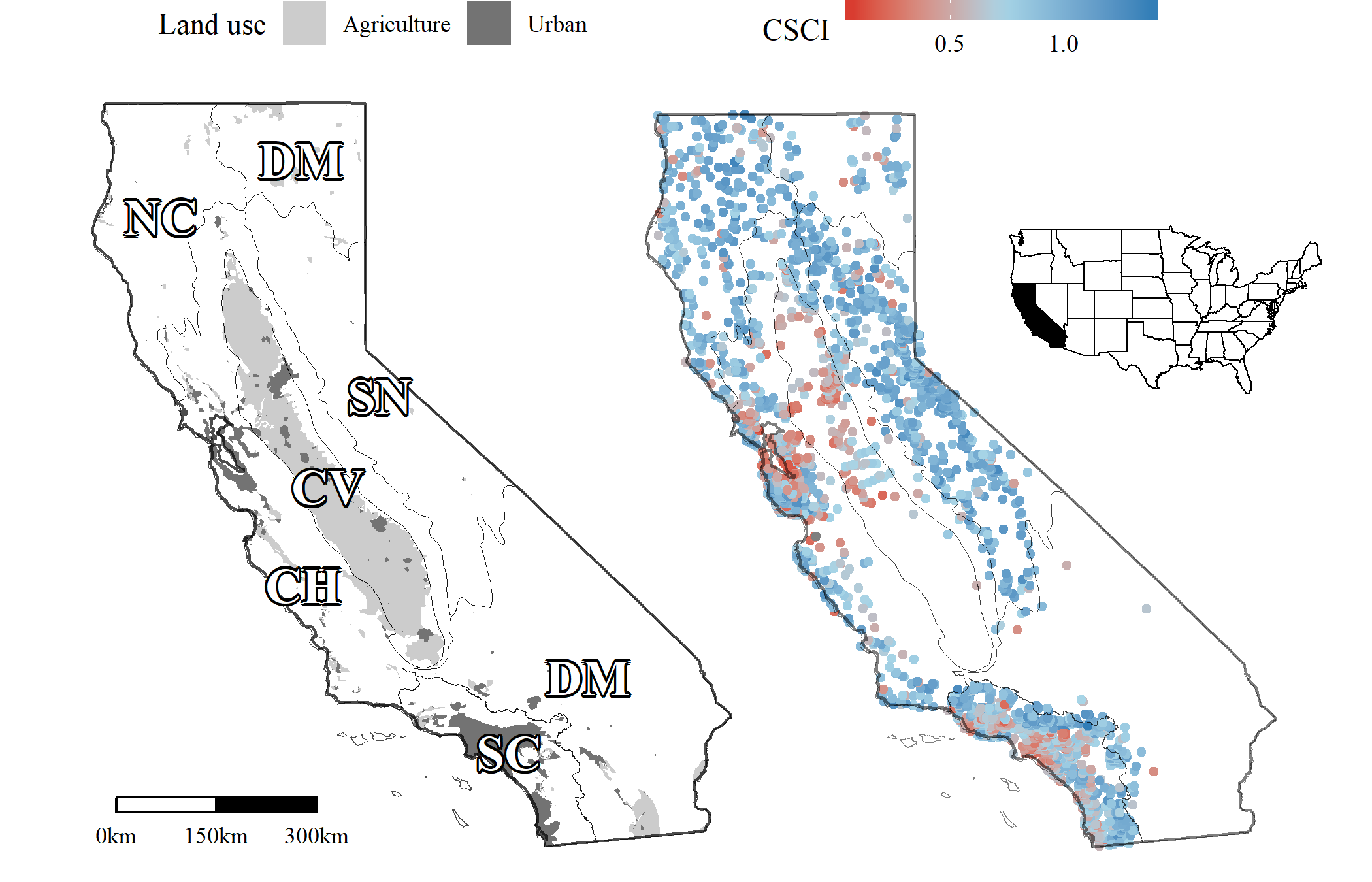


Figure 1 Urban and agricultural land use (left) and distribution of observed stream CSCI scores (right) in California. Cover of urban and agricultural land use in stream watersheds was used to develop a landscape model for stream reach expectations of bioassessment scores. Grey lines are ecoregions in California, CV: Central Valley, CH: Chaparral, DM: Deserts and Modoc Plateau, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

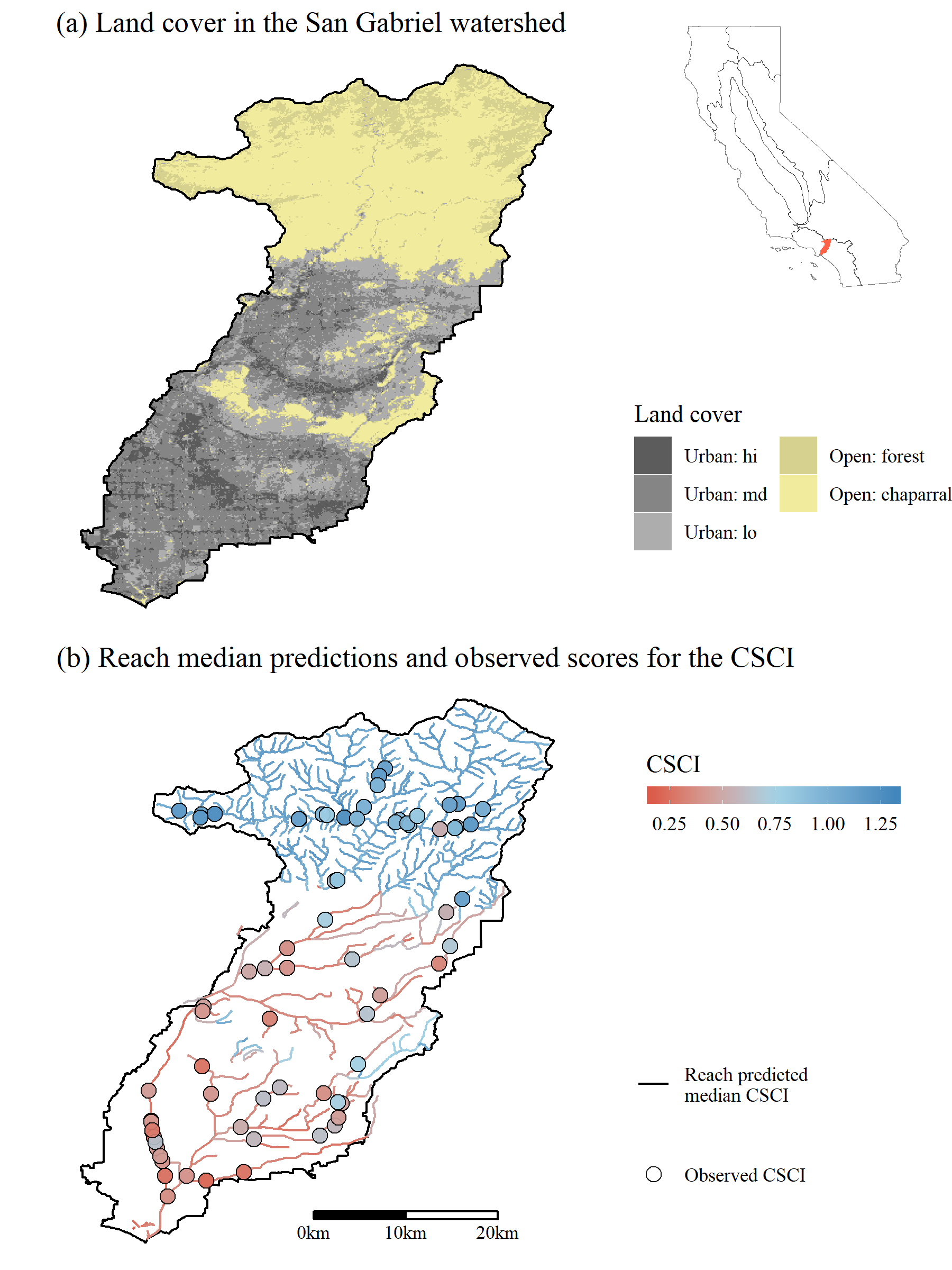


Figure 2 San Gabriel River watershed in southern California. Land cover is shown in plot (a) and the predicted median CSCI scores at each stream reach and observed CSCI scores are shown in (b). The watershed is largely undeveloped in the north and heavily urbanized in the south.

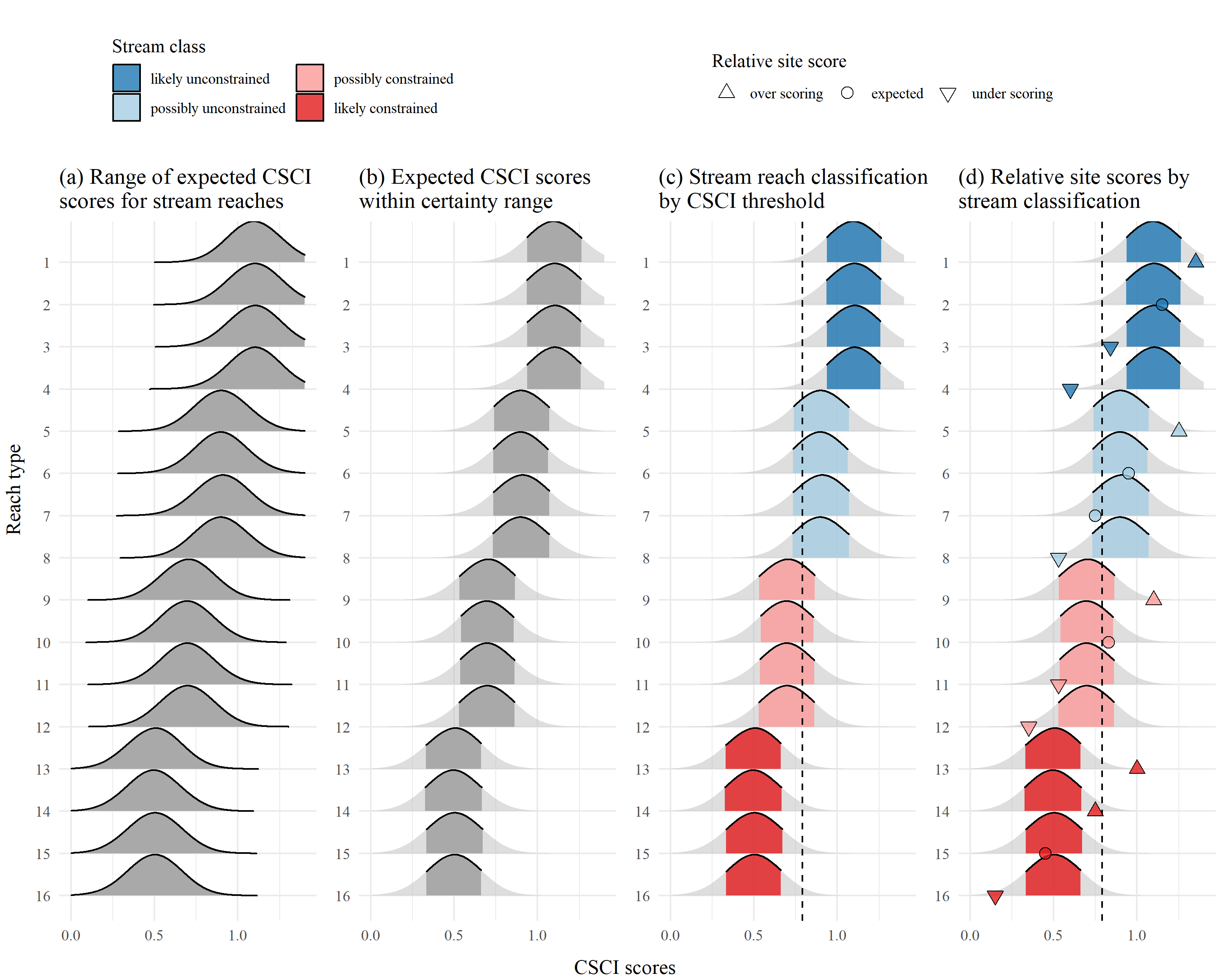


Figure 3 Application of the landscape model to identify site expectations and bioassessment performance for sixteen example stream reaches. A range of CSCI scores is predicted from the model (a) and the lower and upper limits of the expectations are cut to define a certainty range for the predictions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

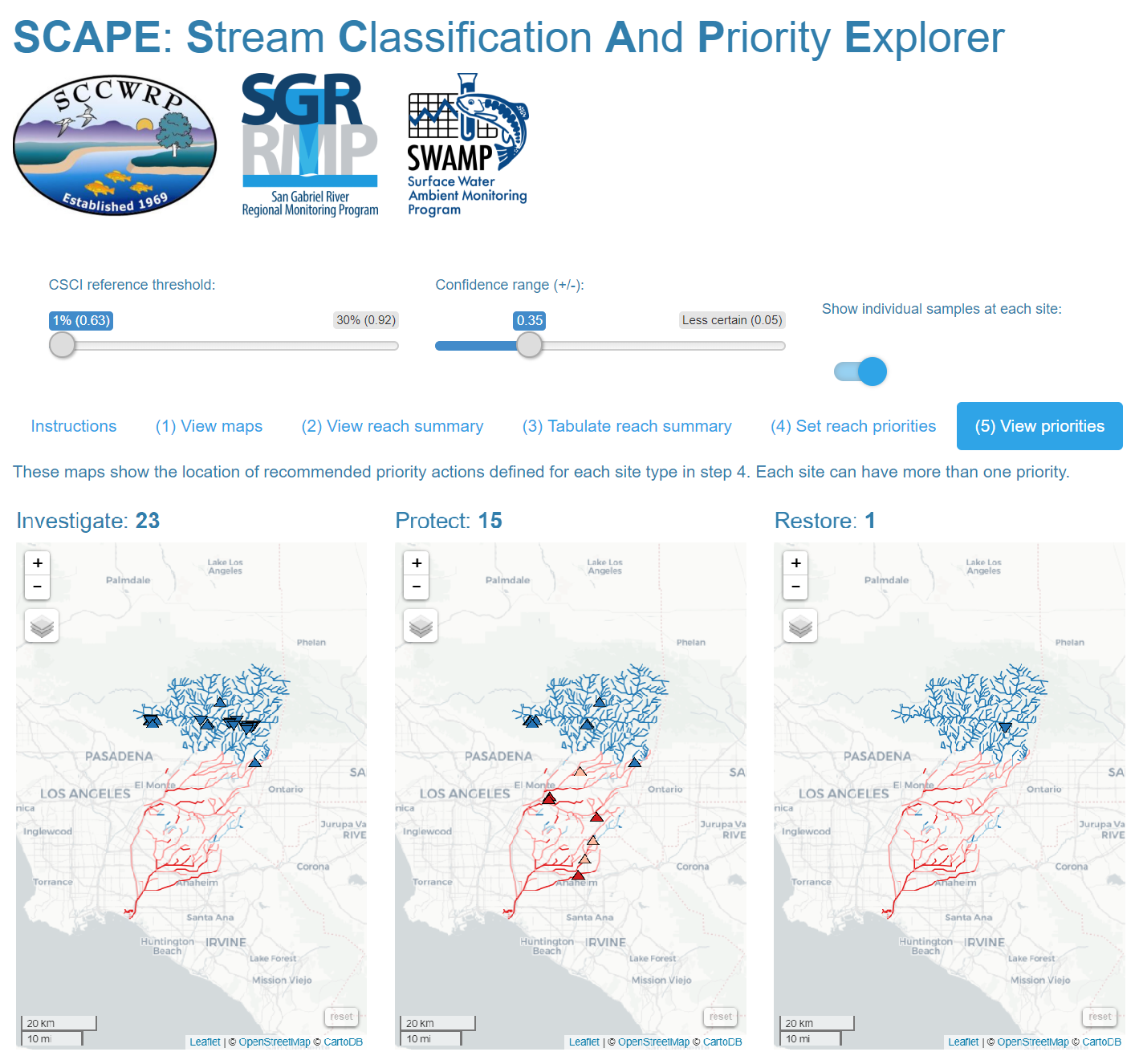


Figure 4 Screenshot of the Stream Classification and Priority Explorer (SCAPE) tool used by the stakeholder group to interact with and use results from the landscape model. The application allowed users to visualize results of reach classifications, relative site scores for the CSCI based on the expectation, and recommend management actions for each reach type. The app can be viewed in the supplementary material.

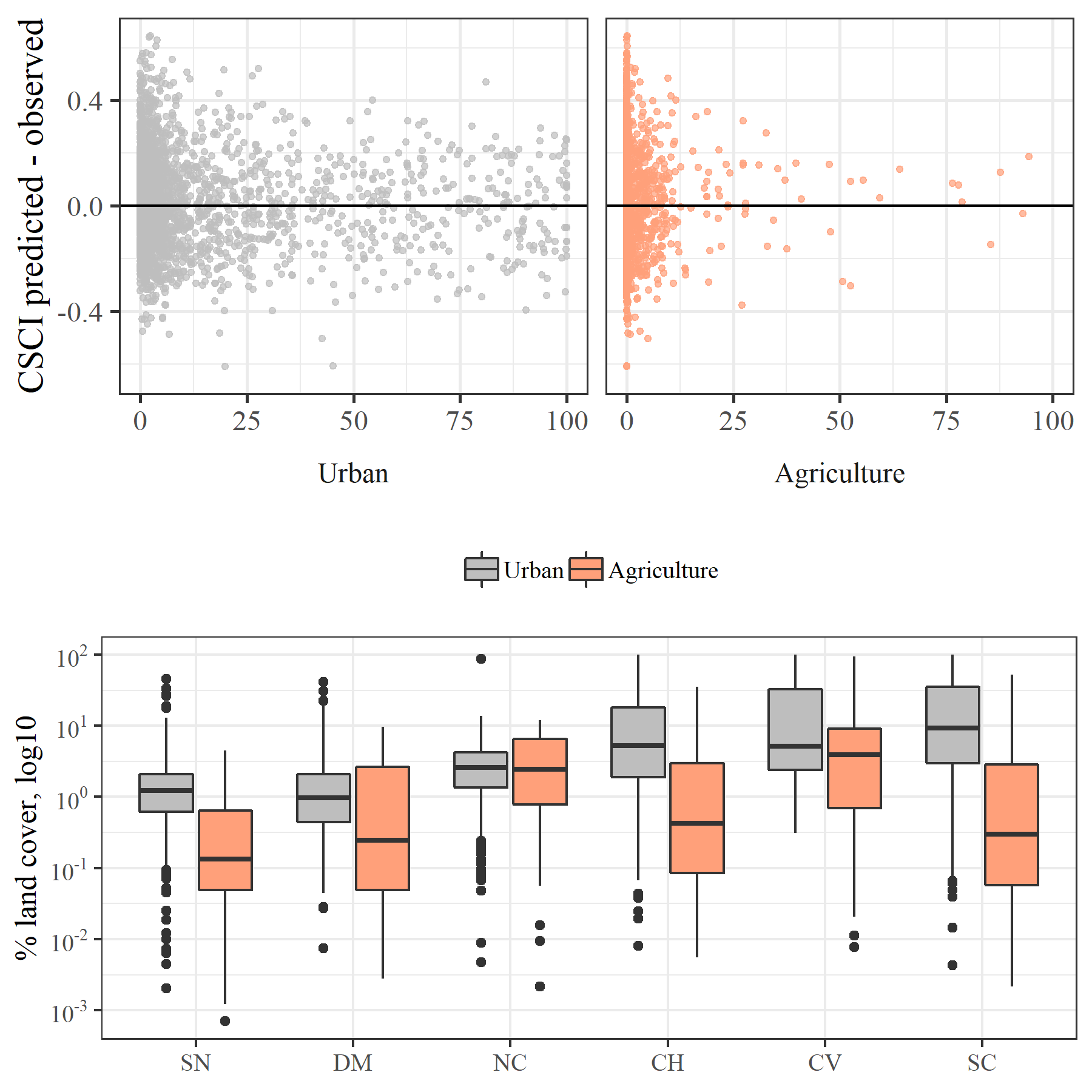


Figure 5 Model performance in relation to land cover and land cover by major regions in California. Model residuals (CSCI predicted - observed) were smaller in regions with more urban or agricultural land use (e.g., SC, CV) and larger in regions with less anthropogenic land use (e.g., SN, DM). CV: Central Valley, CH: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

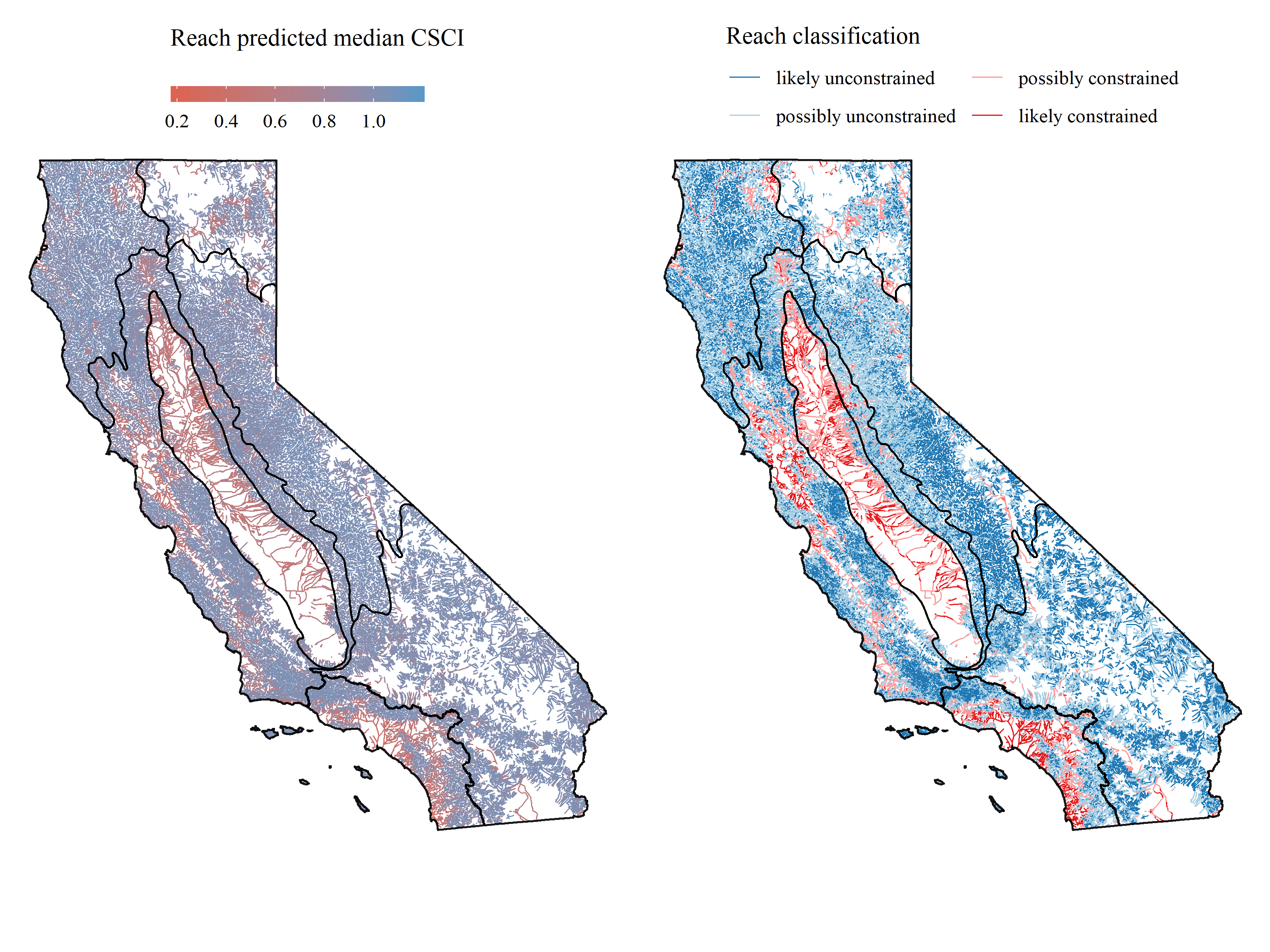


Figure 6 Statewide application of the landscape model showing the median predicted CSCI scores for each stream reach (left) and corresponding stream reach classifications (right). Major regional boundaries are also shown (see Figure 1).

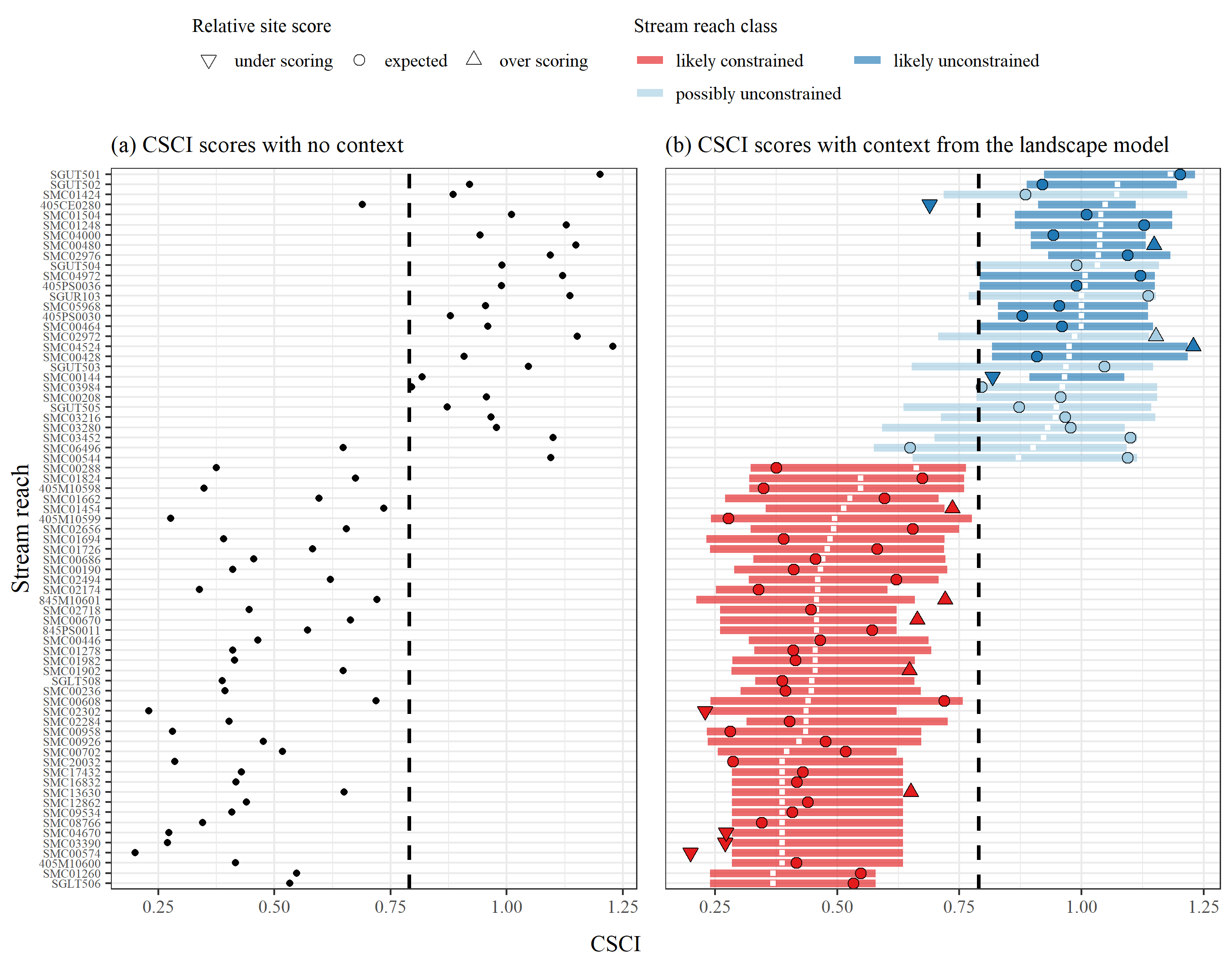


Figure 7 Application of the landscape model to stream reaches in the San Gabriel River watershed, Los Angeles, California. CSCI scores with no context from the model are on the left (a) and scores with context from the model are on the right (b). Relative site scores as under scoring, expected, or over scoring are based on observed scores given the reach class as likely constrained, possibly constrained, possibly unconstrained, and likely unconstrained. Reach classes are based on overlap of the expectations with a hypothetical biological threshold for the CSCI (dashed lined) and location of the median expectation (white ticks).

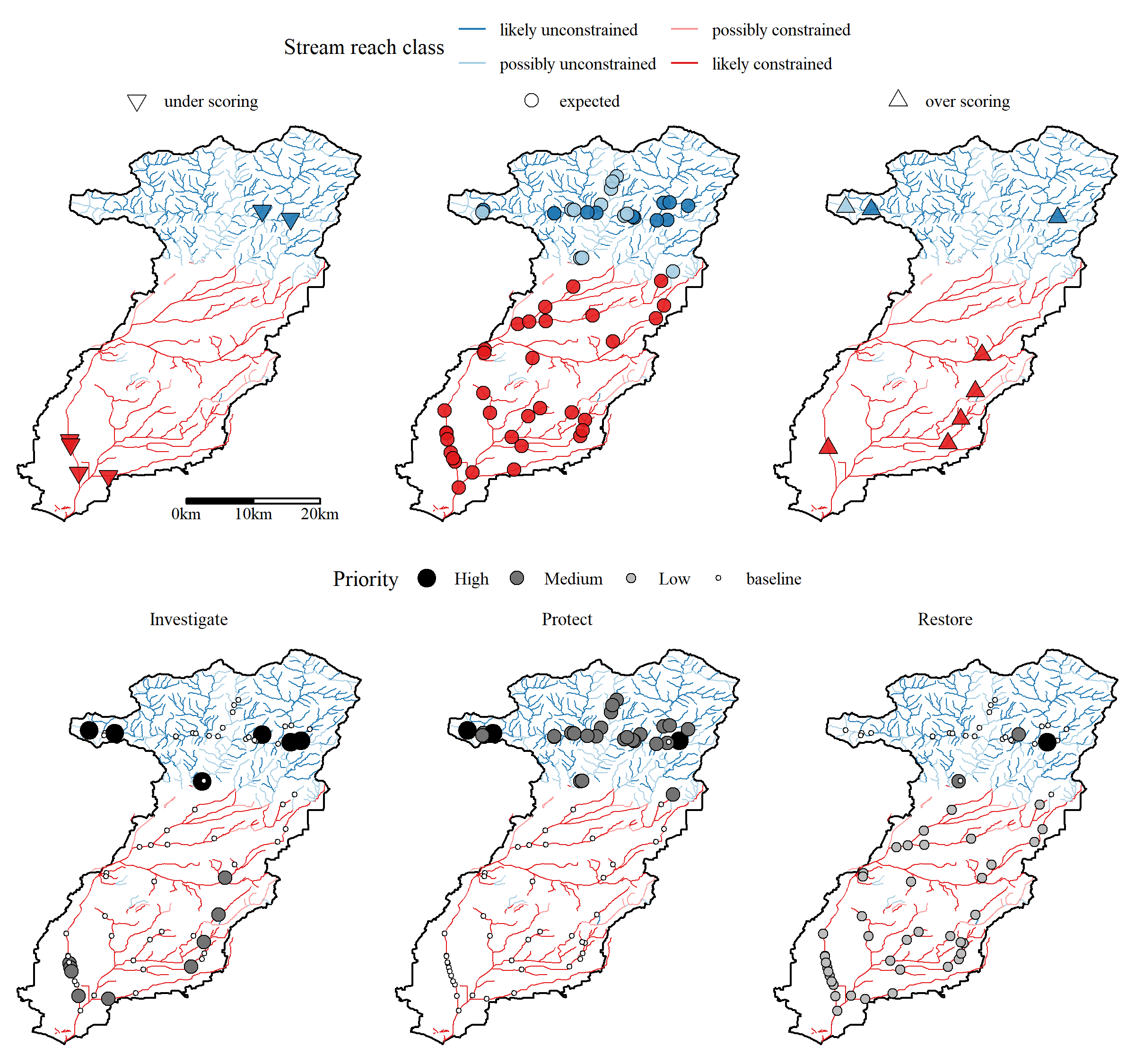


Figure 8 Relative site scores and recommended management actions for locations with CSCI scores in the San Gabriel River watershed. Relative site scores as under scoring, expected, or over scoring are based on observed scores given the reach class as likely constrained, possibly constrained, possibly unconstrained, and likely unconstrained. Recommended management actions are ranked by priority for actions to investigate, protect, and restore a site. No recommended actions assumes baseline maintainence and monitoring is sufficient for a site. Recommended actions were defined by a local stakeholder group (see Figure 9).

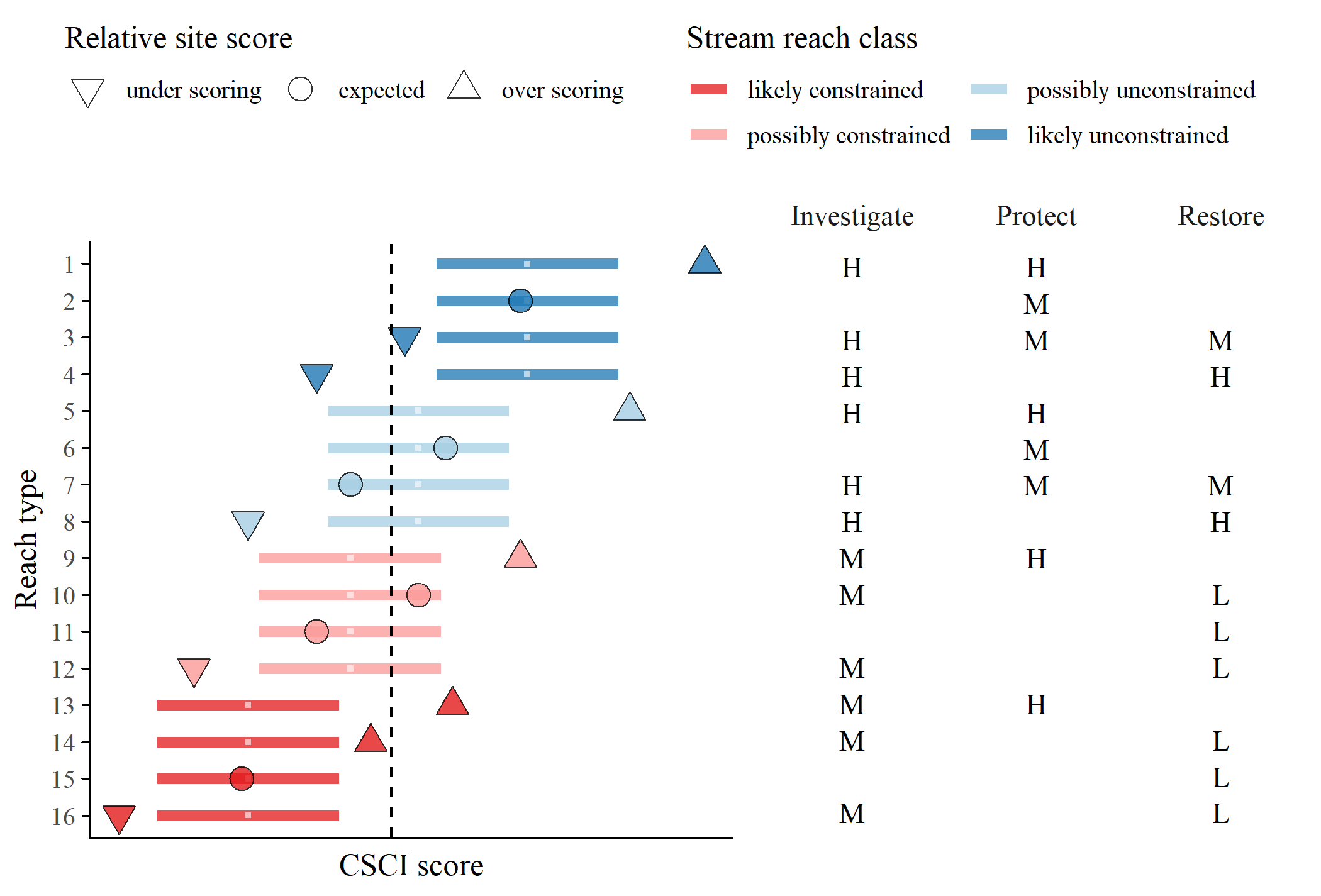


Figure 9 Template provided to stakeholders for priorization of recommended actions for each stream type. The reach types (Table 3) relate to the stream class for the biological expectation (likely unconstrained, possibly unconstrained, possibly constrained, likely constrained), relative site score for the observed CSCI (over scoring, expected, 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.

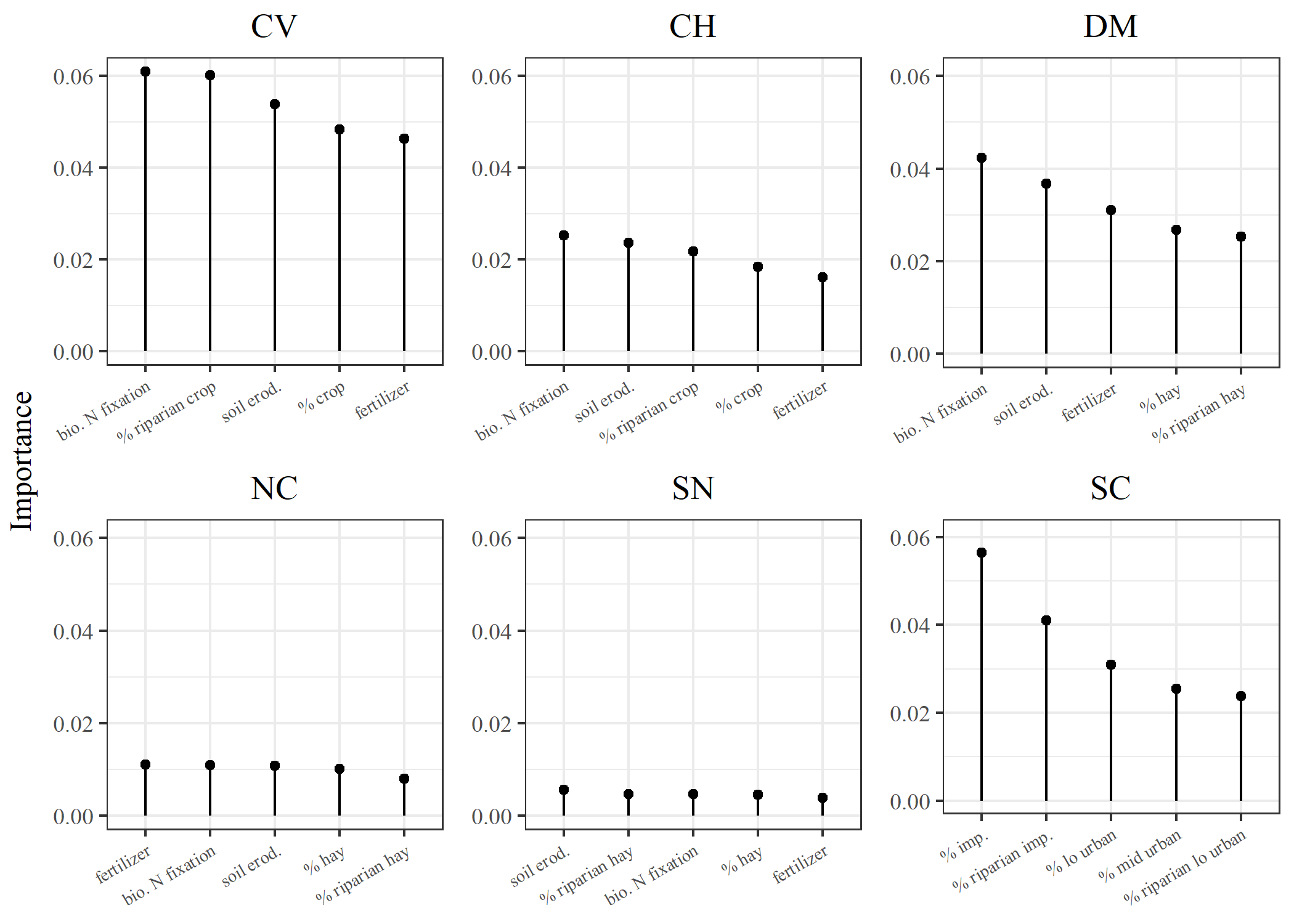


Figure 10 Factors associated with constrained and unconstrained stream reaches by major regions in California. Importance measures were obtained from random forest models of 130 watershed and riparian measures of landscape and geological characteristics from the StreamCat dataset (Hill et al. [2016](#ref-Hill16)). The top five variables for each region are shown. The importance measures describe the mean decrease in prediction accuracy with exclusion of a variable across 1000 random trees for each model. Stream reach classes as possibly or likely were combined for constrained and unconstrained to evaluate the complete dataset. CV: Central Valley, CH: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.



Figure 11 Importance measures for landscape variables used to develop the landscape model of expected stream bioassessment scores in California. Values were obtained from quantile regression models of twenty landscape measures shown in Table 1 obtained from the StreamCat dataset (Hill et al. [2016](#ref-Hill16)). The importance measures describe the percent increase in mean square error and the increase in node impurity with exclusion of a variable across all random trees for each model (Meinshausen [2017](#ref-Meinshausen17)).

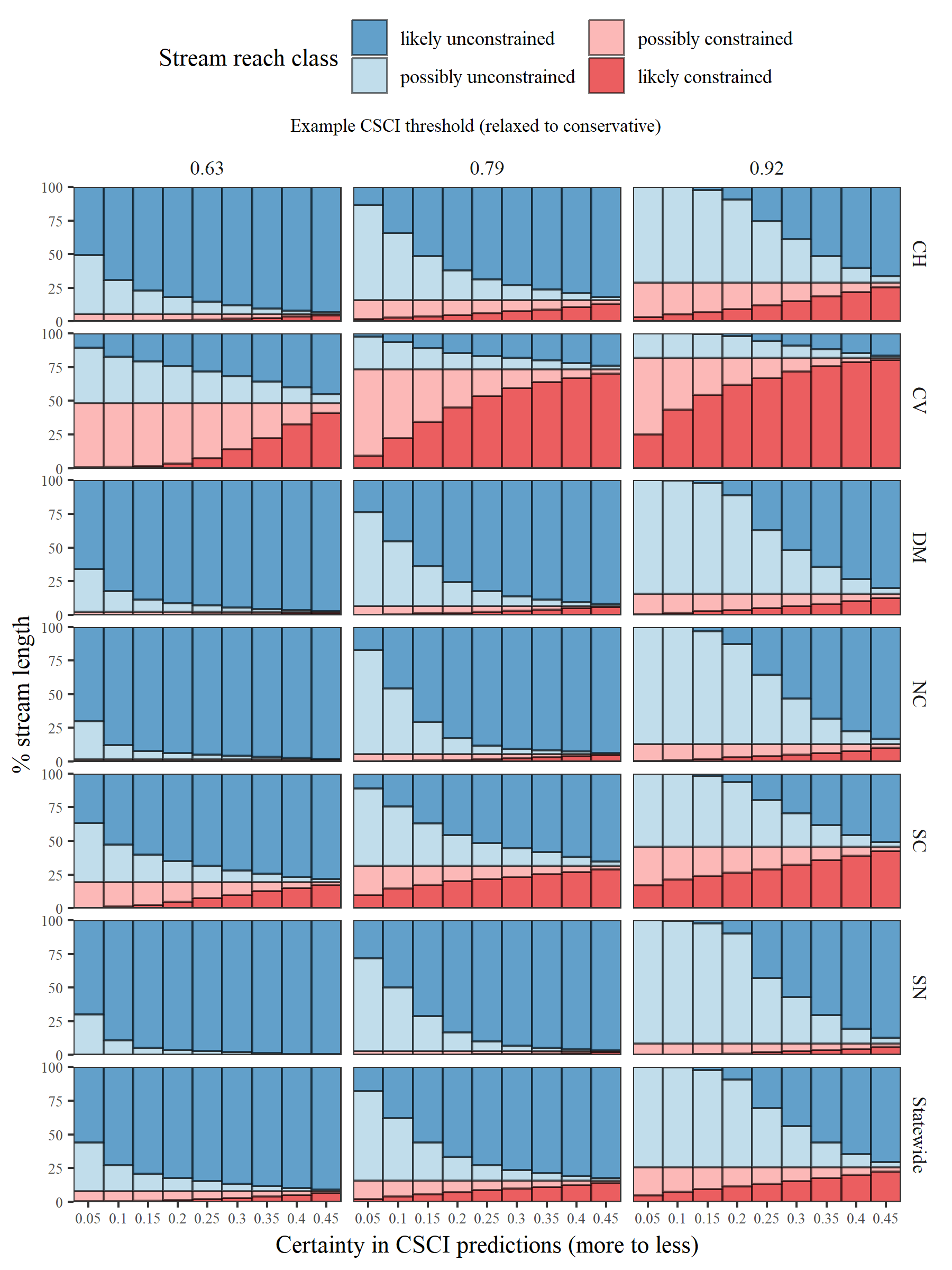


Figure 12 Changes in stream reach classes by region and statewide for different scenarios used to define biological constraints. Twenty-seven scenarios were tested that evaluated different combinations of certainty in the CSCI predictions (nine scenarios more certain to less certain as identified by the tail cutoff for the expected range) and potential CSCI threshold (three scenarios for relaxed to more conservative). The percentage of total stream length for each classification is shown for each scenario. CV: Central Valley, CH: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

# Tables

Table 1 Land use variables used to develop the landscape model of stream bioassessment scores. All variables were obtained from StreamCat (Hill et al. [2016](#ref-Hill16)). The measurement scale for each variable is at the catchment, watershed, and/or riparian scale (100 m buffer) relative to a stream reach. Total urban and agriculture land use variables were based on sums of indvidual variables in StreamCat as noted in the desciption.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Scale | Description | Unit |
| CanalDensCat | catchment | Density of NHDPlus line features classified as canal, ditch, or pipeline | km/sq km |
| CanalDensWs | watershed | Density of NHDPlus line features classified as canal, ditch, or pipeline | km/sq km |
| PctImp2006Cat | catchment | Mean imperviousness of anthropogenic surfaces (NLCD 2006) | % |
| PctImp2006Ws | watershed | Mean imperviousness of anthropogenic surfaces (NLCD 2006) | % |
| PctImp2006CatRp100 | catchment, riparian | Mean imperviousness of anthropogenic surfaces (NLCD 2006) | % |
| PctImp2006WsRp100 | watershed, riparian | Mean imperviousness of anthropogenic surfaces (NLCD 2006) | % |
| TotUrb2011Ws | watershed | Total urban land use as sum of developed open, low, medium, and high intensity (NLCD 2011) | % |
| TotUrb2011Cat | catchment | Total urban land use as sum of developed open, low, medium, and high intensity (NLCD 2011) | % |
| TotUrb2011WsRp100 | watershed, riparian | Total urban land use as sum of developed open, low, medium, and high intensity (NLCD 2011) | % |
| TotUrb2011CatRp100 | catchment, riparian | Total urban land use as sum of developed open, low, medium, and high intensity (NLCD 2011) | % |
| TotAg2011Ws | watershed | Total argricultural land use as sum of hay and crops (NLCD 2011) | % |
| TotAg2011Cat | catchment | Total argricultural land use as sum of hay and crops (NLCD 2011) | % |
| TotAg2011WsRp100 | watershed, riparian | Total argricultural land use as sum of hay and crops (NLCD 2011) | % |
| TotAg2011CatRp100 | catchment, riparian | Total argricultural land use as sum of hay and crops (NLCD 2011) | % |
| RdDensCat | catchment | Density of roads (2010 Census Tiger Lines) | km/sq km |
| RdDensWs | watershed | Density of roads (2010 Census Tiger Lines) | km/sq km |
| RdDensCatRp100 | catchment, riparian | Density of roads (2010 Census Tiger Lines) | km/sq km |
| RdDensWsRp100 | watershed, riparian | Density of roads (2010 Census Tiger Lines) | km/sq km |
| RdCrsCat | catchment | Density of roads-stream intersections (2010 Census Tiger Lines-NHD stream lines) | crossings/sq km |
| RdCrsWs | watershed | Density of roads-stream intersections (2010 Census Tiger Lines-NHD stream lines) | crosssings/sq km |

Table 2 Stream class definitions describing potential biological constraints. Classes are based on the overlap of the range of likely bioassessment scores with a potential threshold for a biological objective. Identifying stream classes requires selecting the cutoff range of likely scores from the landscape model and a chosen threshold for the objective.

|  |  |  |
| --- | --- | --- |
| Class | Definition | Example |
| Likely unconstrained | Lower bound of prediction interval is above threshold | 10th percentile > 0.79 |
| Possibly unconstrained | Median prediction is above threshold | 50th percentile > 0.79 |
| Possibly constrained | Median prediction is below threshold | 50th percentile < 0.79 |
| Likely constrained | Upper bound of prediction interval is below threshold | 90th percentile < 0.79 |

Table 3 Possible site types based on stream reach classification, relative site score, and observed CSCI score. The observed score column describes where a CSCI score is observed relative to the lower and upper percentiles (e.g., 10th and 90th) 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 | Relative site score | Observed score | Type |
| **likely unconstrained** | over scoring | 90th | 1 |
|  | expected | 10th to 90th | 2 |
|  | under scoring | 0.79 to 10th | 3 |
|  | under scoring | < 0.79 | 4 |
| **possibly unconstrained** | over scoring | 90th | 5 |
|  | expected | 0.79 to 90th | 6 |
|  | expected | 10th to 0.79 | 7 |
|  | under scoring | < 10th | 8 |
| **possibly constrained** | over scoring | 90th | 9 |
|  | expected | 0.79 to 90th | 10 |
|  | expected | 10th to 0.79 | 11 |
|  | under scoring | < 10th | 12 |
| **likely constrained** | over scoring | 0.79 | 13 |
|  | over scoring | 90th to 0.79 | 14 |
|  | expected | 10th to 90th | 15 |
|  | under scoring | < 10th | 16 |

Table 4 Performance of the landscape model by calibration and validation datasets in predicting CSCI scores. The statewide dataset (Figure 6) and individual regions of California (Figure 1 are evaluated. Averages and standard deviations (in parentheses) for observed and predicted values of each dataset are shown. Correlations, root mean squared errors (RMSE), intercept, and slopes are for comparisons of predicted and observed values to evaluate model performance.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dataset | Location | n | Observed | Predicted | Correlation | RMSE | Intercept | Slope |
| Cal | Statewide | 1965 | 0.82 (0.26) | 0.83 (0.2) | 0.75 | 0.17 | 0.34 | 0.60 |
|  | CH | 512 | 0.76 (0.27) | 0.79 (0.21) | 0.71 | 0.19 | 0.38 | 0.54 |
|  | CV | 116 | 0.51 (0.18) | 0.57 (0.15) | 0.66 | 0.15 | 0.29 | 0.54 |
|  | DM | 86 | 0.87 (0.22) | 0.91 (0.14) | 0.50 | 0.20 | 0.63 | 0.31 |
|  | NC | 208 | 0.92 (0.2) | 0.94 (0.13) | 0.55 | 0.17 | 0.61 | 0.36 |
|  | SC | 631 | 0.79 (0.24) | 0.78 (0.21) | 0.75 | 0.16 | 0.27 | 0.65 |
|  | SN | 412 | 0.98 (0.18) | 0.98 (0.09) | 0.45 | 0.16 | 0.75 | 0.23 |
| Val | Statewide | 655 | 0.82 (0.25) | 0.84 (0.2) | 0.72 | 0.18 | 0.36 | 0.58 |
|  | CH | 172 | 0.76 (0.27) | 0.81 (0.21) | 0.74 | 0.19 | 0.39 | 0.56 |
|  | CV | 40 | 0.52 (0.19) | 0.59 (0.16) | 0.49 | 0.19 | 0.38 | 0.40 |
|  | DM | 28 | 0.84 (0.17) | 0.93 (0.11) | 0.55 | 0.17 | 0.63 | 0.36 |
|  | NC | 71 | 0.94 (0.19) | 0.96 (0.11) | 0.55 | 0.16 | 0.67 | 0.31 |
|  | SC | 208 | 0.8 (0.24) | 0.78 (0.21) | 0.72 | 0.17 | 0.27 | 0.63 |
|  | SN | 136 | 0.97 (0.17) | 0.98 (0.09) | 0.21 | 0.17 | 0.88 | 0.11 |

*Table 5: (#tab:clstot) Summary of stream length for each stream class statewide and in major regions of California (Figures 1, 6). Lengths are in kilometers with the percentage of the total length in a region in parentheses. CV: Central Valley, CH: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.*

|  | constrained | | unconstrained | |
| --- | --- | --- | --- | --- |
| Region | likely | possibly | possibly | likely |
| Statewide | 8150 (4) | 24735 (11) | 101591 (46) | 85317 (39) |
| CV | 3356 (22) | 8010 (52) | 3202 (21) | 951 (6) |
| CH | 1642 (3) | 7840 (13) | 30693 (50) | 21206 (35) |
| DM | 255 (0) | 3395 (6) | 27194 (47) | 26479 (46) |
| NC | 108 (0) | 1442 (5) | 14152 (49) | 13286 (46) |
| SN | 20 (0) | 1067 (3) | 18228 (48) | 19032 (50) |
| SC | 2770 (15) | 2981 (16) | 8122 (45) | 4363 (24) |

*Table 6: (#tab:reltot) Summary of CSCI scores by relative expectations for each stream class statewide and in each major region of California (Figures 1, 6). Average (standard deviation) scores and counts (percents) of the number of monitoring stations in each relative expectation and region are shown. Sites are over scoring if the observed scores are above the range of expectations at a reach, expected if within the range, or under scoring if below the range. CV: Central Valley, CH: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.*

|  | under scoring | | expected | | over scoring | |
| --- | --- | --- | --- | --- | --- | --- |
| Region | CSCI | n (%) | CSCI | n (%) | CSCI | n (%) |
| Statewide | 0.54 (0.21) | 267 (10) | 0.83 (0.23) | 2041 (80) | 1.08 (0.17) | 242 (9) |
| CH | 0.47 (0.18) | 89 (13) | 0.79 (0.24) | 535 (80) | 1.08 (0.17) | 45 (7) |
| CV | 0.34 (0.12) | 25 (17) | 0.54 (0.17) | 118 (81) | 0.63 (0.25) | 2 (1) |
| DM | 0.6 (0.17) | 15 (14) | 0.9 (0.17) | 89 (80) | 1.15 (0.08) | 7 (6) |
| NC | 0.66 (0.17) | 28 (10) | 0.93 (0.16) | 228 (82) | 1.15 (0.08) | 22 (8) |
| SC | 0.54 (0.22) | 56 (7) | 0.78 (0.22) | 656 (81) | 1.02 (0.2) | 97 (12) |
| SN | 0.67 (0.16) | 54 (10) | 0.99 (0.11) | 415 (77) | 1.16 (0.06) | 69 (13) |

Table 7 Ranges of expected CSCI scores for sites that are typically urban, agricultural, or other land uses by major regions in California and statewide. Ranges can be used to identify approximate expectations for stream reaches with insufficient data for application of the landscape model. CV: Central Valley, CH: Chaparral, DM: Deserts Modoc, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Region | grps | 5th - 95th (More certain) | 25th - 75th | 45th - 55th (Less certain) |
| Statewide | Urban | 0.3 - 0.89 | 0.45 - 0.72 | 0.56 - 0.61 |
|  | Ag | 0.33 - 1 | 0.47 - 0.77 | 0.57 - 0.63 |
|  | Other | 0.73 - 1.19 | 0.91 - 1.08 | 0.99 - 1.02 |
| DM | Urban | 0.43 - 1.11 | 0.65 - 0.96 | 0.79 - 0.84 |
|  | Ag | 0.34 - 1.03 | 0.48 - 0.79 | 0.59 - 0.65 |
|  | Other | 0.71 - 1.19 | 0.9 - 1.08 | 0.98 - 1.02 |
| SN | Urban | 0.43 - 1.12 | 0.63 - 0.96 | 0.78 - 0.84 |
|  | Ag | 0.35 - 1.08 | 0.53 - 0.87 | 0.65 - 0.72 |
|  | Other | 0.73 - 1.2 | 0.92 - 1.09 | 0.99 - 1.02 |
| NC | Urban | 0.62 - 1.19 | 0.86 - 1.09 | 0.95 - 1 |
|  | Ag | 0.35 - 1.08 | 0.51 - 0.86 | 0.63 - 0.69 |
|  | Other | 0.76 - 1.18 | 0.92 - 1.07 | 0.98 - 1.01 |
| CH | Urban | 0.3 - 0.92 | 0.47 - 0.74 | 0.58 - 0.63 |
|  | Ag | 0.37 - 1.07 | 0.53 - 0.86 | 0.64 - 0.7 |
|  | Other | 0.73 - 1.18 | 0.91 - 1.08 | 0.98 - 1.01 |
| CV | Urban | 0.33 - 0.96 | 0.5 - 0.78 | 0.62 - 0.68 |
|  | Ag | 0.32 - 0.97 | 0.45 - 0.73 | 0.54 - 0.6 |
|  | Other | 0.6 - 1.16 | 0.8 - 1.02 | 0.89 - 0.93 |
| SC | Urban | 0.27 - 0.82 | 0.41 - 0.66 | 0.51 - 0.56 |
|  | Ag | 0.35 - 1.06 | 0.54 - 0.9 | 0.67 - 0.74 |
|  | Other | 0.77 - 1.19 | 0.93 - 1.08 | 0.99 - 1.02 |

# References

Allan, D., D. Erickson, and J. Fay. 2007. “The Influence of Catchment Land Use on Stream Integrity Across Multiple Spatial Scales.” *Freshwater Biology* 37 (1):149–61. <https://doi.org/10.1046/j.1365-2427.1997.d01-546.x>.

Allan, J. D. 2004. “Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems.” *Annual Review of Ecology, Evolution, and Systematics* 35:257–64. <https://doi.org/10.1146/annurev.ecolsys.35.120202.110122>.

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

Brown, L. R., J. T. May, A. C. Rehn, P. R. Ode, I. R. Waite, and J. G. Kennen. 2012. “Predicting Biological Condition in Southern California Streams.” *Landscape and Urban Planning* 108 (1):17–27. <https://doi.org/10.1016/j.landurbplan.2012.07.009>.

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.

Carline, R. F., and M. C. Walsh. 2007. “Responses to Riparian Restoration in the Spring Creek Watershed, Central Pennsylvania.” *Restoration Ecology* 15 (4):731–42. <https://doi.org/10.1111/j.1526-100X.2007.00285.x>.

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. <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. <https://doi.org/10.1016/j.ecolind.2013.07.006>.

Chessman, B. C. 2014. “Predicting Reference Assemblages for Freshwater Bioassessment with Limiting Environmental Difference Analysis.” *Freshwater Science* 33 (4):1261–71. <https://doi.org/10.1086/678701>.

Chessman, B. C., M. Muschal, and M. J. Royal. 2008. “Comparing Apples with Apples: Use of Limiting Environmental Differences to Match Reference and Stressor-Exposure Sites for Bioassessment of Streams.” *River Research and Applications* 24 (1):103–17. <https://doi.org/10.1002/rra.1053>.

Chessman, B. C., and M. J. Royal. 2004. “Bioassessment Without Reference Sites: Use of Environmental Filters to Predict Natural Assemblages of River Macroinvertebrates.” *Journal of the North American Benthological Society* 23 (3):599–615. <https://doi.org/10.1899/0887-3593(2004)023%3C0599:BWRSUO%3E2.0.CO;2>.

Clapcott, J. E., K. J. Collier, R. G. Death, E. O. Goodwin, J. S. Harding, D. Kelly, J. R. Leathwick, and R. G. Young. 2011. “Quantifying Relationships Between Land-Use Gradients and Structural and Functional Indicators of Stream Ecological Integrity.” *Freshwater Biology* 57 (1):74–90. <https://doi.org/10.1111/j.1365-2427.2011.02696.x>.

Cormier, S. M., G. W. Suter II, L. Zhang, and G. J. Pond. 2013. “Assessing Causation of the Extirpation of Stream Macroinvertebrates by a Mixture of Ions.” *Environmental Toxicology and Chemistry* 32 (2):277–87. <https://doi.org/10.1002/etc.2059>.

Finkenbine, J. K., J. W. Atwater, and D. S. Mavinic. 2000. “Stream Health After Urbanization.” *Journal of the American Water Resources Association* 36 (5):1149–60. <https://doi.org/10.1111/j.1752-1688.2000.tb05717.x>.

Fox, E. W., R. A. Hill, S. G. Leibowitz, A. R. Olsen, D. J. Thornbrugh, and M. H. Weber. 2017. “Assessing the Accuracy and Stability of Variable Selection Methods for Random Forest Modeling in Ecology.” *Environmental Monitoring and Assessment* 189:316. <https://doi.org/10.1007/s10661-017-6025-0>.

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., E. W. Fox, S. G. Leibowitz, A. R. Olsen, D. J. Thornbrugh, and M. H. Weber. 2017. “Predictive Mapping of the Biotic Condition of Conterminous U.S. Rivers and Streams.” *Ecological Applications* 27 (8):2397–2415. <https://doi.org/10.1002/eap.1617>.

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. <https://doi.org/10.1111/1752-1688.12372>.

Howard, J. K., K. A. Fesenmyer, T. E Grantham, J. H. Viers, P. R. Ode, P. B. Moyle, S. J. Kupferburg, et al. 2018. “A Freshwater Conservation Blueprint for California: Prioritizing Watersheds for Freshwater Biodiversity.” *Freshwater Science* 37 (2):417–31. <https://doi.org/10.1086/697996>.

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. <https://doi.org/10.1371/journal.pone.0130710>.

Howard, J., and C. Revenga. 2000. “California’s Freshwater Biodiversity 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. <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. <https://doi.org/10.1016/j.scitotenv.2015.07.083>.

Lester, R. E., and A. J. Boulton. 2008. “Rehabilitating Agricultural Streams in Australia with Wood: A Review.” *Environmental Management* 42 (2):310–26. <https://doi.org/10.1007%2Fs00267-008-9151-1>.

MacQueen, J. 1967. “Some Methods for Classification and Analysis of Multivariate Observations.” In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, edited by L. M. Le Cam and J. Neyman, 281–97. Berkeley, California, USA: University of California Press.

May, J. T., L. R. Brown, A. C. Rehn, I. R. Waite, P. R. Ode, R. D. Mazor, and K. C. Schiff. 2015. “Correspondence of Biological Condition Models of California Streams at Statewide and Regional Scales.” *Envrionmental Monitoring and Assessment* 187:4086. <https://doi.org/10.1007/s10661-014-4086-x>.

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>.

Morgan, R. P., and S. E. Cushman. 2005. “Urbanization Effects on Stream Fish Assemblages in Maryland, USA.” *Journal of the North American Benthological Society* 24 (3):643–55.

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. <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.

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.

Riseng, C. M., M. J. Wiley, R. W. Black, and M. D. Munn. 2011. “Impacts of Agricultural Land Use on Biological Integirty: A Causal Analysis.” *Ecological Applications* 21 (8):3128–46. <https://doi.org/10.1890/11-0077.1>.

Roni, P., and T. Beechi. 2012. *Stream and Watershed Restoration Guide: A Guide to Restoring Riverine Processes and Habitats*. First. Hoboken, New Jersey: John Wiley & Sons.

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. <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. <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. <https://doi.org/10.1046/j.1365-2427.2001.00758.x>.

Stein, E. D., M. R. Cover, A. E. Fetscher, C. O’Reilly, R. Guardado, and C. W. Solek. 2013. “Reach-Scale Geomorphic and Biological Effects of Localized Streambank Armoring.” *Journal of the American Water Resources Association* 49 (4):780–92. <https://doi.org/10.1111/jawr.12035>.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. “Setting Expectations for the Ecological Condition of Streams: The Concept of Reference Condition.” *Ecological Applications* 16 (4):1267–76.

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. <https://doi.org/10.1111/j.1365-2664.2008.01548.x>.

USEPA (US Environmental Protection Agency). 2016. “National Rivers and Streams Assessment 2008-2009: A Collaborative Survey.” EPA-841-R-16-007. Washington, DC.

USGS (US Geological Survey). 1999. “The quality of our nation’s waters: nutrients and pesticides.” Reston, Virginia.

———. 2014. “National Hydrography Dataset available on the World Wide Web.”

Vølstad, J. H., N. E. Roth, G. Mercurio, M. T. Southerland, and D. E. Strebel. 2004. “Using Environmental Stressor Information to Predict the Ecological Status of Maryland Non-Tidal Streams as Measured by Biological Indicators.” *Environmental Monitoring and Assessment* 84 (3):219–42. <https://doi.org/10.1023/A:1023374524254>.

Waite, I. R. 2013. “Development and Application of an Agricultural Intensity Index to Intevertebreate and Algal Metrics from Streams at Two Scales.” *Journal of the American Water Resources Assocation* 49 (2):431–48. <https://doi.org/10.1111/jawr.12032>.

Waite, I. R., J. G. Kennen, J. T. May, L. R. Brown, T. F. Cuffney, K. A. Jones, and J. L. Orlando. 2014. “Stream Macroinvertebrate Response Models for Bioassessment Metrics: Addressing the Issue of Spatial Scale.” *PLOS ONE* 9 (3):e90944. <https://doi.org/10.1371/journal.pone.0090944>.

Wang, L., J. Lyons, P. Kanehl, and R. Gatti. 1997. “Influences of Watershed Land Use on Habitat Quality and Biotic Integrity in Wisconsin Streams.” *Fisheries* 22 (6):6–12. <https://doi.org/10.1577/1548-8446(1997)022%3C0006:IOWLUO%3E2.0.CO;2>.