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 R. 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: Mon Jul 9 14:14:03 2018 -0700

# Abstract

Management goals for stream condition may be difficult or impossible to achieve where large-scale changes on the landscape (e.g., urbanization) impose constraints on the upper limit of biological integrity. We developed a statewide landscape model that provides context for a macroinvertebrate-based bioassessment index by predicting ranges of likely scores that are typical at a site for the observed level of landscape alteration. With this approach, a relative site score can be identified to evaluate if a site is scoring above, below, or within the range of scores that are expected given the level of landscape alteration. We applied the model at a local scale to classify sites and prioritize different management actions using guidance from a local stakeholder group from the San Gabriel River Regional Monitoring Program (SGRRMP, 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. We observed model predictions that were consistent with the clear land use gradient from the upper to lower watershed, 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)). A bioassessment index may be difficult to incorporate into management through regulatory statutes if thresholds for biological objectives are impractical within site-specific conditions. Managers require explicit information that 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 stream miles in the USA are estimated to be in poor biotic condition based on macroinvertebrate bioassesssment and the observed condition can be strongly associated with commonly observed in-stream stressors, such as excess phosphorus, nitrogen, or altered physical habitat (USEPA (US Environmental Protection Agency) [2016](#ref-USEPA16)). These proximate and immediate causes of poor biological condition are often linked to landscape-level factors that occur in the watershed. Consistent and empirical links between land use thresholds and poor biotic integrity have been identified in many cases (Allan, Erickson, and Fay [1997](#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))), but the precise link between land use and instream condition is not 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.

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)). 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 been incentivized or enforced through regulation. Although these approaches can lead to improvements in stream health, there is no universal remedy that applies to streams. 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)). A confounding factor for managing stream health in developed landscapes is the extensive modification to streams for flood control or water conveyance. Channel modification has been used as a basis for redefining water quality criteria that is site specific or for re-evaluating use attainability goals. For example, the Los Angeles River (California, USA) is heavily engineered as a concrete-lined channel and recreational uses that apply nationally are suspended under high flow conditions (California Regional Water Quality Control Board, Los Angeles Region [2014](#ref-CRWQCB14)). Other states have recommended a tiered aquatic life use or alternative use designations to account for baseline shifts in ecosystem health from environmental modification (Florida Department of Environmental Protection [2011](#ref-FLDEP11); US Environmental Protection Agency, Region 10 [2013](#ref-USEPA13); Midwest Biodiversity Insitute [2016](#ref-MBI16)).

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, or conversely, where landscape alteration might impose constraints on biological integrity that limits potential for management success. 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 a bioassessment context 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 as 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)). Applying these concepts in a predictive framework could facilitate an expectation of bioassessment and management potential relative to a site-specific context.

Previous modelling efforts for bioassessment 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 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. 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, specifically focusing on the potential of urban and agricultural land use to constrain biological condition. We include a case study that demonstrates how the statewide model can be used to classify and prioritize in a regional context using guidance from a local stakeholder group from a heavily urbanized watershed. An interactive software tool is also described that was developed to help choose management priorities from the landscape model.

# 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 with extreme variation in altitude, geology, 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. 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. Much of this development is recent such that 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) (McKay et al. [2012](#ref-McKay12)) 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 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. Many of the metrics in StreamCat were derived from the 2006 National Land Cover Database (Fry et al. [2011](#ref-Fry11)).

The California Stream Condition Index (CSCI) (Mazor et al. [2016](#ref-Mazor16)) was used as a measure of biological condition in California streams. 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 values 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 index score at a site can vary from 0 to ~ 1.4, with higher values indicating less deviation from reference state. Because the index was developed to minimize the influence of natural gradients, the index scores have consistent meaning across the state (Mazor et al. [2016](#ref-Mazor16)). Benthic macroinvertebrate data were used to calculate 6270 individual CSCI scores at nearly 3400 unique sites between 2000 and 2016. Field data were collected during base flow conditions typically between May and July following methods in Ode et al. ([2016](#ref-Ode16)).

## Building and validating the landscape model

A prediction model was developed to estimate likely ranges of CSCI scores associated with urban and agricultural land use gradients, collectively defined as developed landscapes. 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 S1). These variables were chosen specifically to model scores only in relation to constraints on biological condition that are typically beyond the scope of management intervention or where costs to mitigate are likely prohibitive. Potential effects on biological condition that may vary through time or from stressors not associated with urban or agricultural land use were not captured by the model. Similarly, potential differences in the magnitude of potential effects for the chosen variables were also not explicitly evaluated. Within these limits, deviation of observed scores from the model predictions were considered diagnostic of variation not related to landscape effects at a coarse level, in addition to potential 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 conversely 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 at five percent intervals (i.e., 5th, 10th, etc.) from the 5th to 95th percentile of expectations.

Calibration data for the landscape model were obtained from a random selection of 75% of reaches with observed CSCI scores and where sufficient data were available in StreamCat (n = 1965 reaches). The random selection was stratified by major regions in California (Figure 1), with each further divided into quartiles based on watershed imperviousness to adequately represent landscape gradients in each region. The remaining sites were used for model validation (n = 655). 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 by comparing differences between observed CSCI scores and median predictions at the same locations. Differences 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. Collectively, the performance metrics were chosen to evaluate both predictive ability of the landscape model and potential for bias which may vary depending on different land use gradients across the state.

## 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 1):

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

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 used for comparison with the modelled expectations (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 1).

CSCI scores from biomonitoring data were then compared to the stream reach classification to identify a “relative site score” (Figure 3d). This comparison is a fundamental product of the landscape model that can be used to evaluate if a site is scoring above, below, or within the range of scores that are expected given the level of landscape alteration. 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 considered “under-scoring”. If neither “over-scoring” nor “under-scoring”, the relative site score was considered 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 among 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. The possibly and likely constrained classes were evaluated together, as were the possibly and likely unconstrained classes. However, it was recognized that retaining the possibly or likely classes could be useful to refine the priorities described below.

## 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 low or high scores (Figure 3c). The 10th and 90th percentiles 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 changing this range (e.g., 25th to 75th percentile) was evaluated to determine the sensitivity of reach classifications on this decision point. Similary, the CSCI threshold was also changed to assess effects of relaxing or increasing flexibility in defining low or high scores. A default threshold of 0.79 as a measure of the 10th percentile of scores at reference (non-impacted) sites was used above, but this value can be increased to examine effects of a more conservative 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 extent in each classification.

## Defining management priorities in the San Gabriel River watershed

A framework for identifying SGR site priorities for management actions using results from the landscape model was developed through engagement with a local stakeholder group. The San Gabriel River Regional Monitoring Program (SGRRMP, Los Angeles County, California) includes stakeholders from water quality regulatory agencies, municipalities, and non-govermental organizations that cooperatively work to manage aquatic resources in the SGR watershed and improve coordination of compliance and ambient monitoring efforts. The workgroup met monthly over a six-month period to discuss model application and to refine the interpretation of results.

A strong land-use gradient occurs in the SGR watershed that creates challenges for managing stream health (Figure 2). Headwaters begin in the San Gabriel mountains where the land is primarily undeveloped or protected for recreational 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. Spreading grounds are present in the middle of the watershed for groundwater recharge during high flow. As a result, the upper and lower watersheds are hydrologically disconnected when annual rainfall is normal. Nearly all of the stream reaches in the lower half of the watershed are channelized with concrete or other reinforcements. The majority of flow in the lower watershed is provided to San Jose Creek, Coyote Creek, and the mainstem of the San Gabriel River by wastewater treatment plants releasing tertiary treated effluent. Approximately half of the monitored sites in the watershed are in poor biological condition, all of which are in the lower watershed.

Pioritizing among the many sites that require some management intervention was a critical objective of the stakeholder group in applying the landscape model. The stakeholder group identified priorities in the watershed by first describing the types of management actions that were desired. Stakeholders identified their relevant priorities by evaluating the different site types that were possible from the landscape model relative to the stream class (under-scoring, expected, over-scoring for each of four reach classes). The priorities were also defined by considering if an observed CSCI score at a site was above or below a potential biological threshold (e.g., CSCI of 0.79), in addition to the scoring expectation for the stream class (Table 2). To facilitate the process, a template was used that showed the site scores relative to the reach classifications (Figure S3, left side). The priorities defined by the group were generalized into three categories:

* Investigate: Suggest additional monitoring or review of supplementary data (e.g., field visits, review aerial imagery) to characterize why a site is scoring above or below an expectation, or above or below a biological objective given the expectation;
* Protect: Require additional scrutiny of any proposed development and/or projects for sites that score above the expectation;
* Restore: Pursue targeted action for causal assessment and/or restoration activity for sites that score below the expectation.

One to many priorities were assigned as low, medium, or high priorities for the scoring possibilities that could occur at a site. The assignments were also made with the explicit recognition that any priority recommendations were in addition to baseline monitoring and maintenance that is currently provided by existing management programs (Surface Water Ambient Monitoring Program [2017](#ref-SWAMP17)).

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 S1, <http://shiny.sccwrp.org/scape/>, (**???**)). This application provided maps of the extent and type of classification for reaches in the watershed, deviation of observed CSCI scores from the expectation, and maps of recommended priority actions that were assigned to each of the scoring possibilities. The SCAPE tool also allowed the stakeholders to modify key decision points in the model (i.e., range of expectations that were used from the model, selected biological threshold) to evaluate how these changes propogated to changes in recommended priorities for each site.

# Results

## Model performance

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. Model performance statewide indicated generally good agreement between observed CSCI scores and the median prediction for the associated stream reach (Table S2). Agreement between observed and predicted values for the entire calibration dataset was r = 0.75 (Pearson) and RMSE = 0.17. The intercept and slope for a regression between observed and predicted values were 0.34 and 0.60, suggesting a slight negative bias of predictions at lower scores and slight positive bias at higher scores. The statewide validation data showed similar results, with slightly smaller correlation (r = 0.72) and larger RMSE (0.18) estimates.

The landscape model performed 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. Model performance was best in regions wih extensive urban development. Performance for the Chaparral 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 0.57 with observed values in the calibration dataset and 0.53 in the validation dataset. Model performance was weakest for the Sierra Nevada region, where timber harvesting, rather than urban or agricultural development, is the most widespread stressor. Overall, model performance was strongly associated with land use gradients in each region (Figure S2).

## Statewide patterns

Statewide patterns in stream constraints were apparent from the results of the landscape model that were consistent with land use (Figure 4). 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 3). Large rivers across the state were more commonly classified as possibly constrained (e.g., Klamath, Sacramento, Colorado rivers). By region, the most reaches classified as likely unconstrained reaches were in the Sierra Nevada (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. Overall, stream reaches were more often constrained for biotic integrity in regions with more 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 4). 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 4). 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.

## Associated drivers of biological constraints and sensitivity analysis

Importance measures from random forest models identified key variables that were associated with differences between constrained and unconstrained reaches between each region (Figure S4, see Figure S5 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 Chaparral region, second most important for the North Coast, and third most important for the Sierra Nevada region. 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 Sierra Nevada region. 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.

Sensitivity analyses underscored the potential impact of key decision points of the landscape model on estimates of the extent of streams in each class (Figure 7). 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.

## San Gabriel River Case study

Engagement of the SGRRMP stakeholders demonstrated how management actions can be prioritized through application of the landscape model. About 750 reaches in the SGR were identified and classified from NHD-plus, of which 10% were visited over a ten-year period 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 2a). 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 (Figure 5). Stream reaches in the upper watershed were a mix of likely and possibly unconstrained (40% and 28%), whereas stream reaches in the lower watershed were classified as likely and possibly constrained (25% and 7%). Several reaches in the lower watershed had ranges that were right-skewed toward very low CSCI scores consistent with extreme landscape pressures (bottom left, Figure 5b).

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 wide 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) (top, Figure 6). 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.

In general, the stakeholder group assigned high priority recommendations to over- and under-scoring sites in likely unconstrained reaches or those below the biological threshold with possibly unconstrained classification (Figure S3). Continuing current practices were generally recommended at constrained sites or restoration actions were recommended as a lower priority despite low CSCI scores. Recommended actions to investigate were more common for both over-scoring and under-scoring sites, protect was given a high priority exclusively 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 6). Sites in the lower watershed were lower priority if an action was recommended, whereas the five high priority sites were in the upper watershed (multiple recommendations were assigned to the sites). The distinction between lower and higher priorities between the lower and upper watershed was driven exclusively by the reach classifications, where constrained reaches were in the lower watershed and unconstrained reaches 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.

# Discussion

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 diagnosis of causes of impairment, nor does it provide an exemption 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.

Results from our analyses 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 exclude sites from management actions 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.

## 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. Instead, they can help identify patterns among monitoring sites where more intensive analyses may be appropriate. This application was effectively demonstrated through engagement of our local stakeholder group. Rather than identifying individual sites in need of specific management actions, the group used the landscape model to characterize patterns on the landscape that were consistent with the recommended managagment priorities. In doing so, the group was able to synthesize a large volume of bioassessment data to explore potential management actions relative to the landscape contexts of the watershed.

Our approach has the advantage of relating strictly to biological condition, as compared to physical or chemical endpoints to assess constraints. This approach 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 relatively low expected range of CSCI scores is an indication that stressors originating from the landscape may 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.

Our approach to predict biological condition using landscape constraints also has advantages over other methods that define constraints based only on channel modification. Several states have recommended alternative use designations for applying bioassessment criteria in modified channels (Florida Department of Environmental Protection [2011](#ref-FLDEP11); US Environmental Protection Agency, Region 10 [2013](#ref-USEPA13); Midwest Biodiversity Insitute [2016](#ref-MBI16)). Our results generally support this approach, although defining constraints based only on channel modification may be insufficient. Physical habitat quality can be 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 in rural landscapes (e.g., the mainstem of the Klamath and Russian rivers in the North Coast region) were also identified by the model, as well as many streams in agricultural areas (e.g., Salinas River). 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. Modified channels may also be present in undeveloped landscapes and high bioassessment scores have been observed in armoured streams within national forest lands (Stein et al. [2013](#ref-Stein13)). A classification framework for biological constraints using only channel modification would provide incomplete information relative to an approach using landscape information. These results are well supported by other landscape studies, particularly for macroinvertebrates (May et al. [2015](#ref-May15)).

The utility of landscape models in supporting watershed management has applications outside of California. Our use of national geospatial datasets (i.e., NHDPlus, McKay et al. ([2012](#ref-McKay12)); StreamCat, Hill et al. ([2016](#ref-Hill16))) means that these methods could be applied elsewhere in diverse bioassessment contexts. The CSCI was developed for macroinvertebrate assessment in California, but this approach could be applied with other methods, such as a multi-metric index (the most common bioassessment approach within the US; Buss et al. ([2014](#ref-Buss14))), O/E assessments (Moss et al. [1987](#ref-Moss87)), biological condition gradients (Davies and Jackson [2006](#ref-Davies06)), or with other biological endpoints (e.g., fish or diatoms). In addition, extension of the landscape model could be explored to develop a national scale product of constraints on biological condition to complement recent work that predicted probable biological conditions with the National Rivers and Streams Assessment (Hill et al. [2017](#ref-Hill17)).

Extension of the landscape models beyond California should also consider landscape stressors that are predictive of biotic condition in other regions. For example, urban and agricultural gradients were sufficient to characterize constraints in many regions of California, whereas Hill et al. ([2017](#ref-Hill17)) found that the volume of water stored by dams was an important predictor of biological condition in the Northern Appalachian and Northern Plains regions of the US. In their paper, Hill et al. ([2017](#ref-Hill17)) provided an example of how predictive models could be used to identify potential sites for restoration or conservation, however, their illustration did not explicitly identify sites that were over- or under-scoring relative to a biological endpoint. Doing so in California provided stakeholders with important context that helped establish management priorities, demonstrating the potential utility of this approach in other states.

# Model assumptions and limitations

There are several characteristics of the landscape model that could affect its performance when applied outside of urban and agricultural settings. First, the model was developed with a focus on the needs of managers that apply bioassessment tools in developed landscapes. As such, landscape variables were chosen to capture the effects of development on CSCI scores in these areas (Table S1). This could lead to erroneous conclusions in regions where different stressors have strong impacts on stream health. For example, our results suggest that streams in the North Coast and Sierra Nevada regions are largely unconstrained, but model performance was poor in these areas. The dominant stressors likely to affect stream condition in these regions originate from sources that are less common in developed landscapes, e.g., silviculture, cannabis cultivation, water extraction and hydrologic alteration. The current landscape model does not adequately capture these impacts outside of urban and agricultural environments. Accurate data for quantifying these potential stressors are much less readily available, but this is an area where investments in improving spatial data could yield significant improvements.

An additional assumption is that the landscape model and the CSCI can adequately discriminate between intractable constraints on biology that are spatially and temporally pervasive relative to more tractable constraints. This assumption applies to any stressor gradient that could be used to develop the model. For example, our model adequately described urban constraints but there was no context for temporal or spatial scales that have management relevance. Pervasive and profound alteration to groundwater and hydrology is common in highly developed areas and stream communities may not ever be able to be restored to reference conditions even in the most extreme management interventions. Similar conditions likely exist for other land use gradients. For example, a landscape model developed to describe constraints from timber harvesting practices may not provide adequate information on the long-term impacts of siltation on stream integrity if the input data does not describe these impacts at spatial and temporal scales that are relevant for management. The potential legacy impacts of large-scale alterations of the natural environment are not well-captured by the current model, neither from a spatial nor temporal perspective. Our analysis of landscape factors associated with constraints using additional StreamCat variables provided a preliminary means of addressing this concern, although a more refined application of the landscape model would be necessary to evaluate different scales of impact. This could include developing separate models for each region, as well as more careful selection of model inputs to capture scales of interest for potential impacts on stream health.

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. A correspondence between habitat limitations and channel modification is likely in many cases but data are insufficient to evaluate biological effects statewide relative to land use constraints. Moreover, bioassessment scores can be similar in modified channels compared to natural streams independent of watershed land use, i.e., concordance between poor stream health and channel modification may not always be observed (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 using the landscape model 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 biological attributes of stream health (e.g., fish assemblages) 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. The focus of the landscape model on a specific taxa 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.

## Engagement of local stakeholders is critical for regional application

Application of the landscape model to define potential management actions was effectively tested with the SGRRMP stakeholder group. The final decision by the group to prioritize management actions for the different sites in broad categories of protect, restore, and investigate was based on an iterative process where ideas were discussed and shared freely among stakeholders. 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 SGRRMP. Application of the statewide model to other watersheds must include the engagement of local stakeholders in a similar process to develop recommendations that are adequate to meet regional needs.

The development of the SCAPE tool 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. The tool provided a means of demonstrating core concepts of the landscape model and allowed stakeholders to explore the key decision points that affect the model output, specifically related to changing certainties in the CSCI score predictions and the ability to explore alternative thresholds for biological objectives. This functionality allowed the stakeholders to develop recommendations that were completely independent of the model, i.e., decisions were not hard-wired into the model nor SCAPE. Because of this tool, this stakeholder group has a better understanding of the potential impacts of biointegrity policies currently under review in California. Additionally, the SCAPE tool provided assurance to the prioritization process by correctly identifying sites where dicrepancies between CSCI scores and other measures of stream condition had been observed. The SCAPE tool prioritized a site for restoration in the upper watershed that was unconstrained and under-performing. This confirmed a discrepancy identified by the stakeholders where good physical habitat conditions were observed from field visits, but the observed CSCI score was below the chosen threshold. As such, application of the landscape modelling approach to other regions will benefit from similar tools that actively engage managers with bioassessment data.

## Summary

The landscape model can be used to characterize the extent of biologically constrained channels in urban and agricultural landscapes. Our application to the SGR watershed demonstrated how the results of the model can be used at a spatial scale where many management decisions are implemented through close interaction with a regional stakeholder group with direct interests in the local resources. Overall, the model provides a tool to determine how managers can best prioritize 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 information from multiple sources to develop a context for biological assessment that provides an expectation of what is likely to be achieved relative to urban and agricultural stressors. 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.

# Supplement

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

# Figures

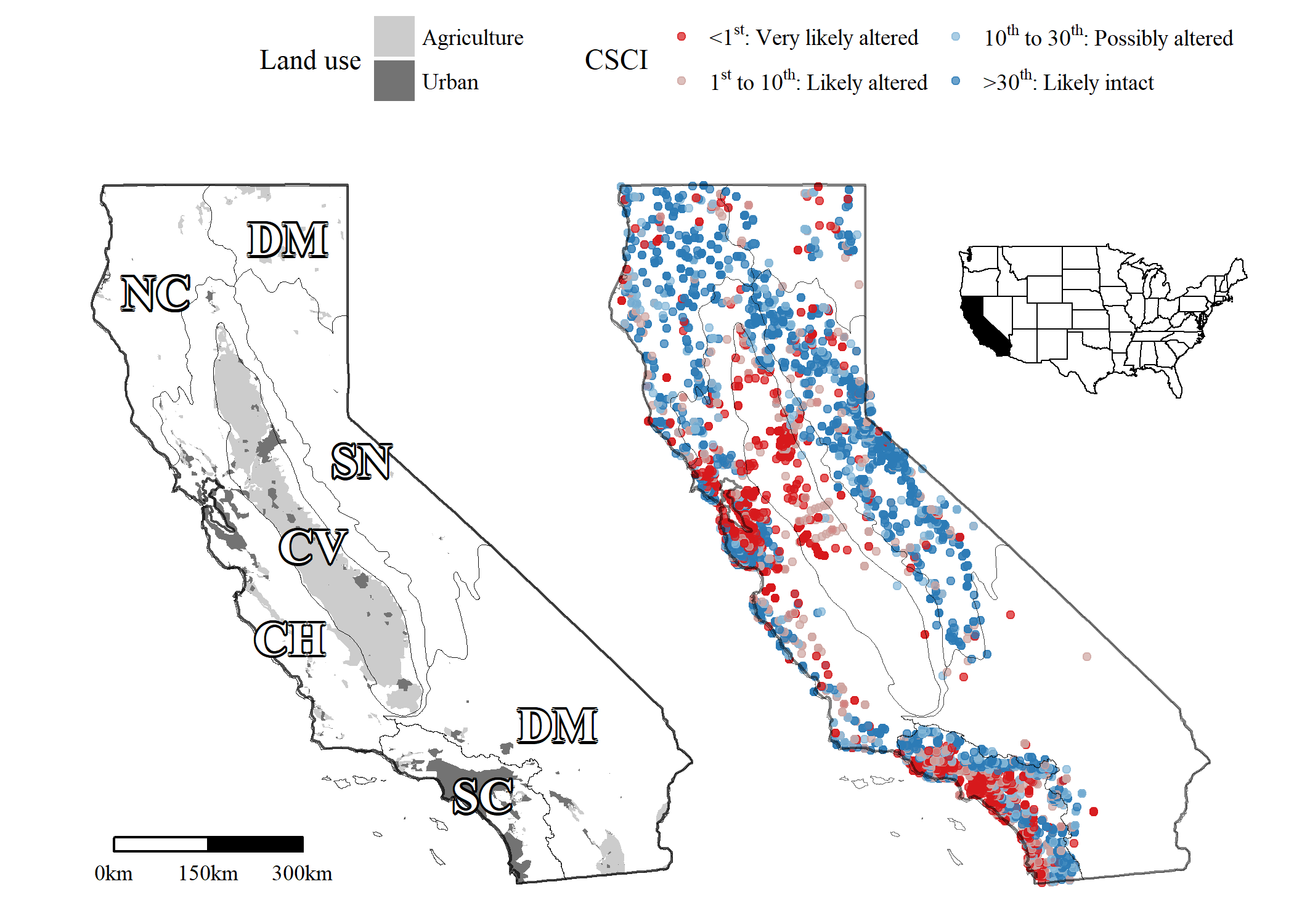


Figure 1 Urban and agricultural land use (left) and distribution of observed stream CSCI scores (right) in California. Cover of urban and agricultural land use in stream watersheds was used to develop a landscape model for stream reach expectations of bioassessment scores. Breakpoints for CSCI scores in the right plot are the 1st, 10th, and 30th percentile of scores at least-disturbed, reference sites throughout the state. Altered and intact refers to biological condition (Mazor et al. [2016](#ref-Mazor16)). Grey lines are major environmental regions in California defined by ecoregional and watershed boundaries, CV: Central Valley, CH: Chaparral, DM: Deserts and Modoc Plateau, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

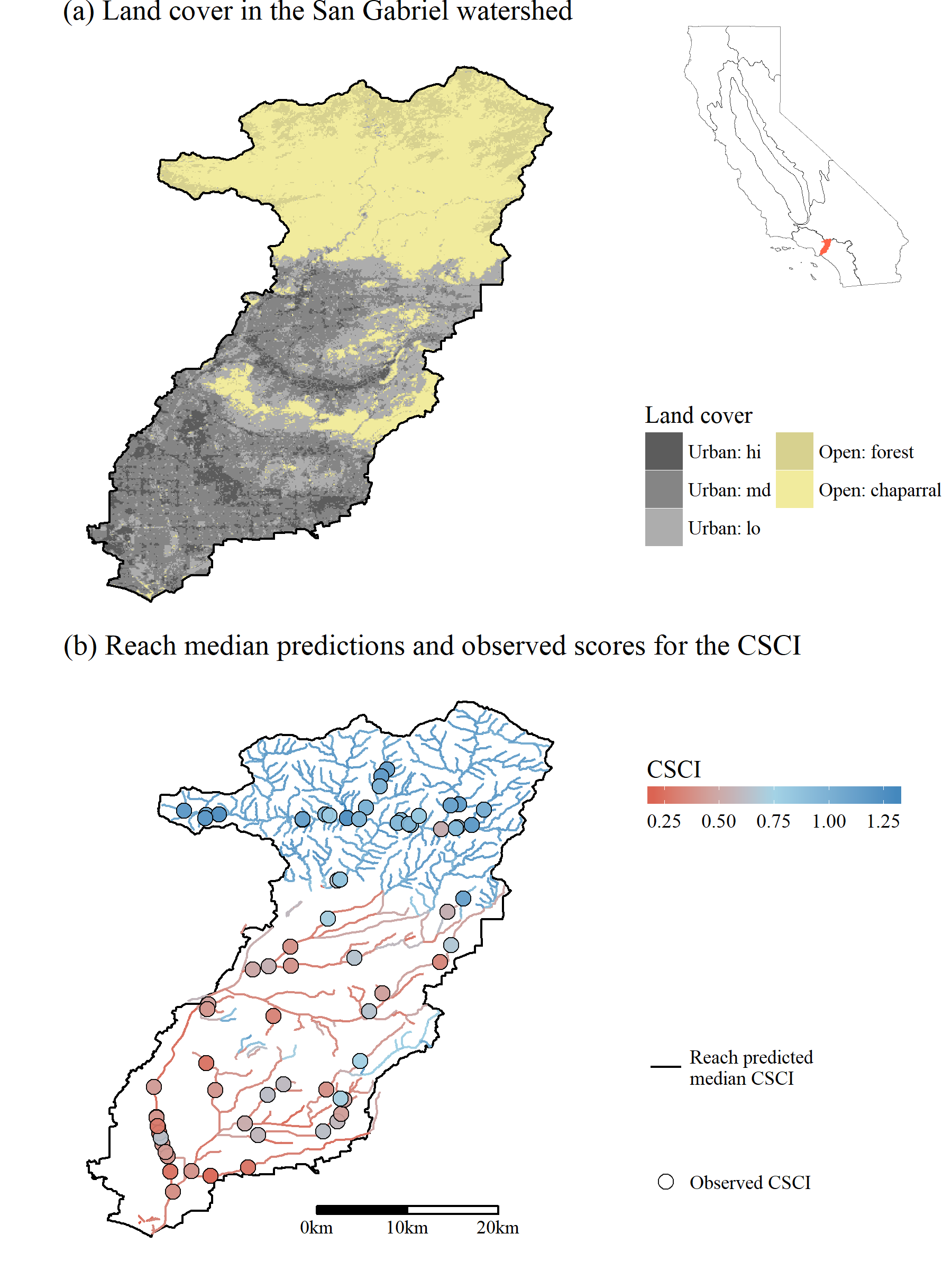


Figure 2 San Gabriel River watershed in southern California. Land cover is shown in plot (a) and the predicted median CSCI scores at each stream reach and observed CSCI scores are shown in (b). The upper watershed is largely undeveloped, whereas the lower watershed is heavily urbanized.

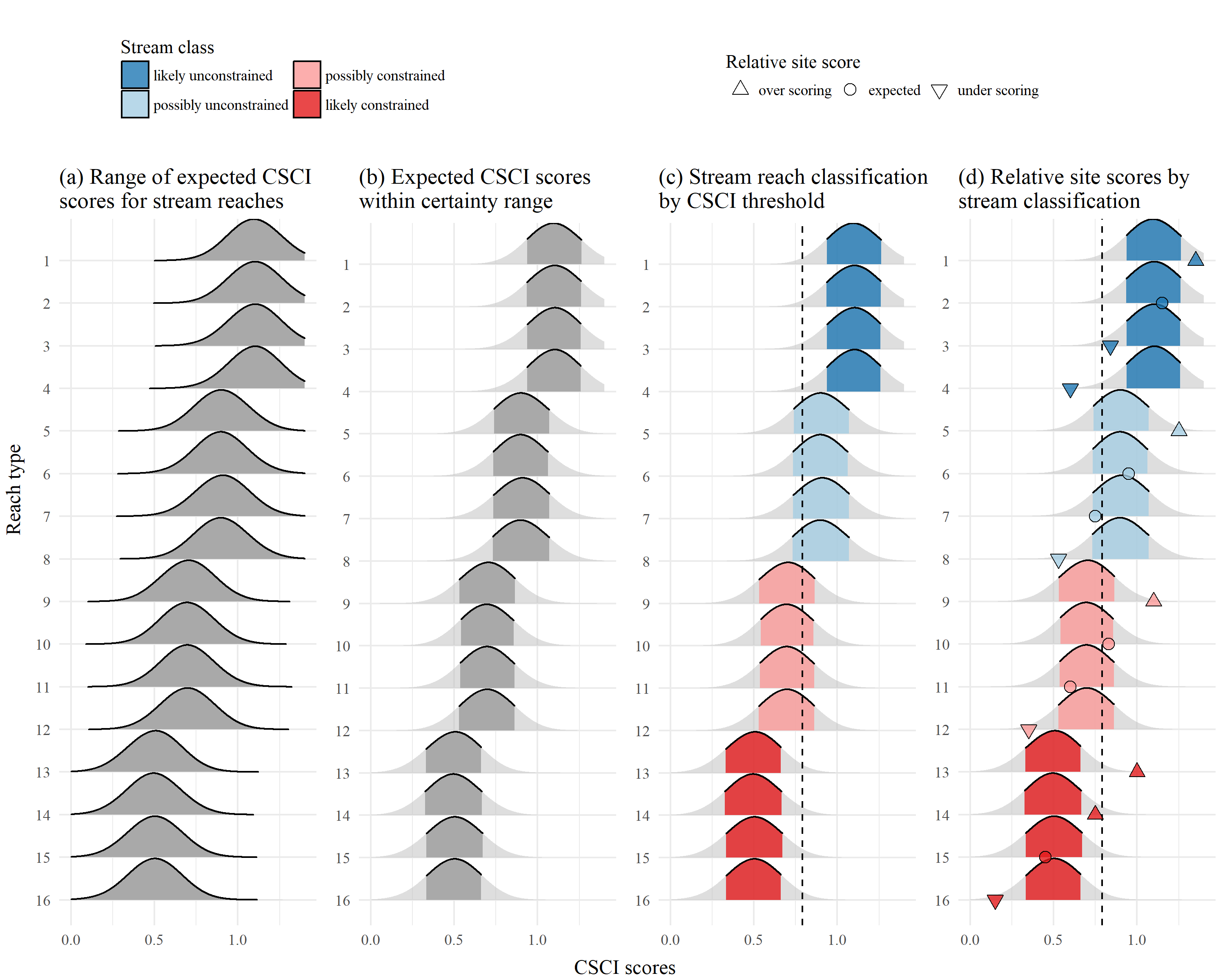


Figure 3 Application of the landscape model to identify site expectations and bioassessment performance for sixteen example stream reaches. A range of CSCI scores is predicted from the model (a) and the lower and upper limits of the expectations are cut to define a certainty range for the predictions (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).

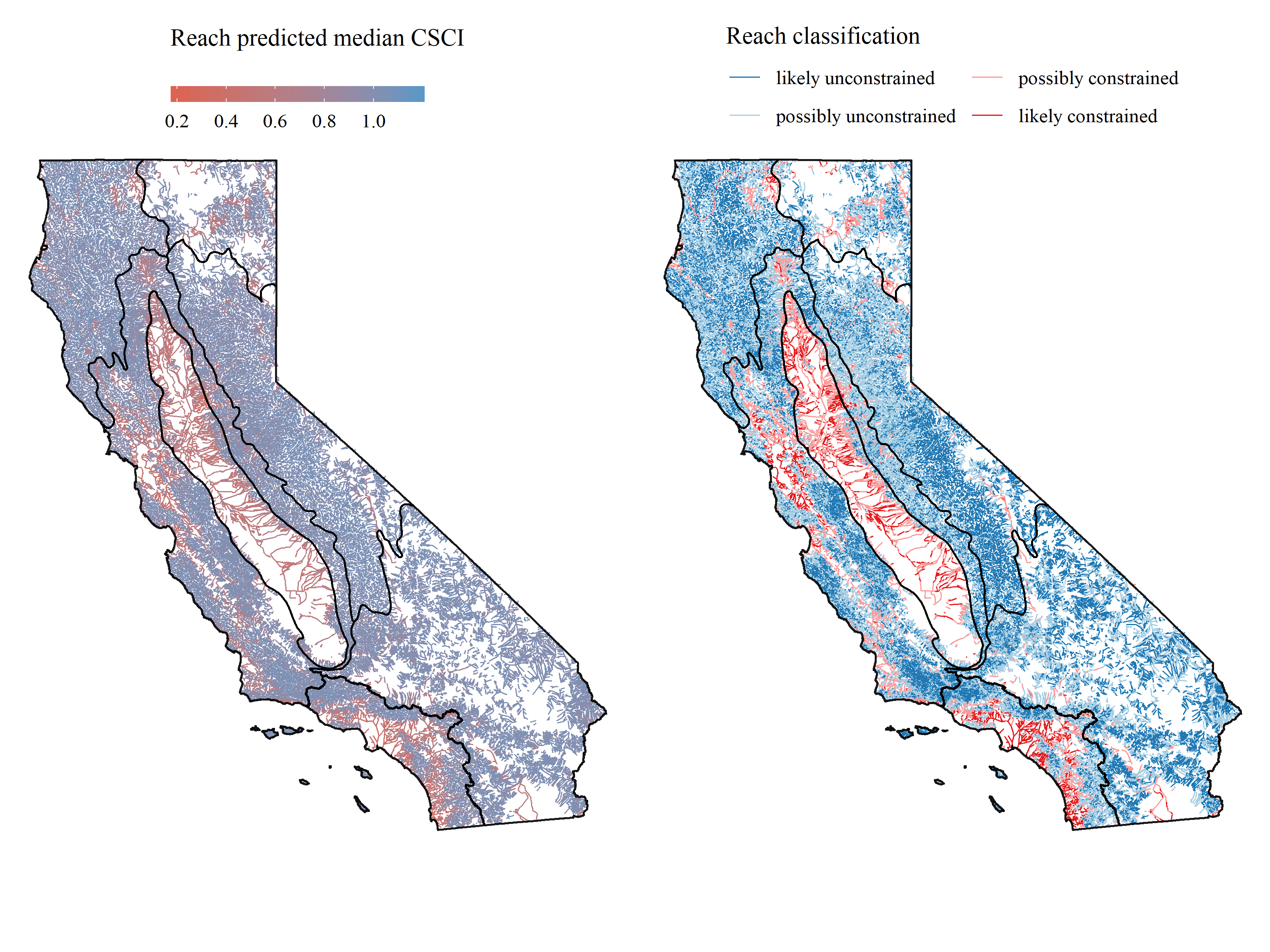


Figure 4 Statewide application of the landscape model showing the median predicted CSCI scores for each stream reach (left) and corresponding stream reach classifications (right). Major regional boundaries are also shown (see Figure 1).

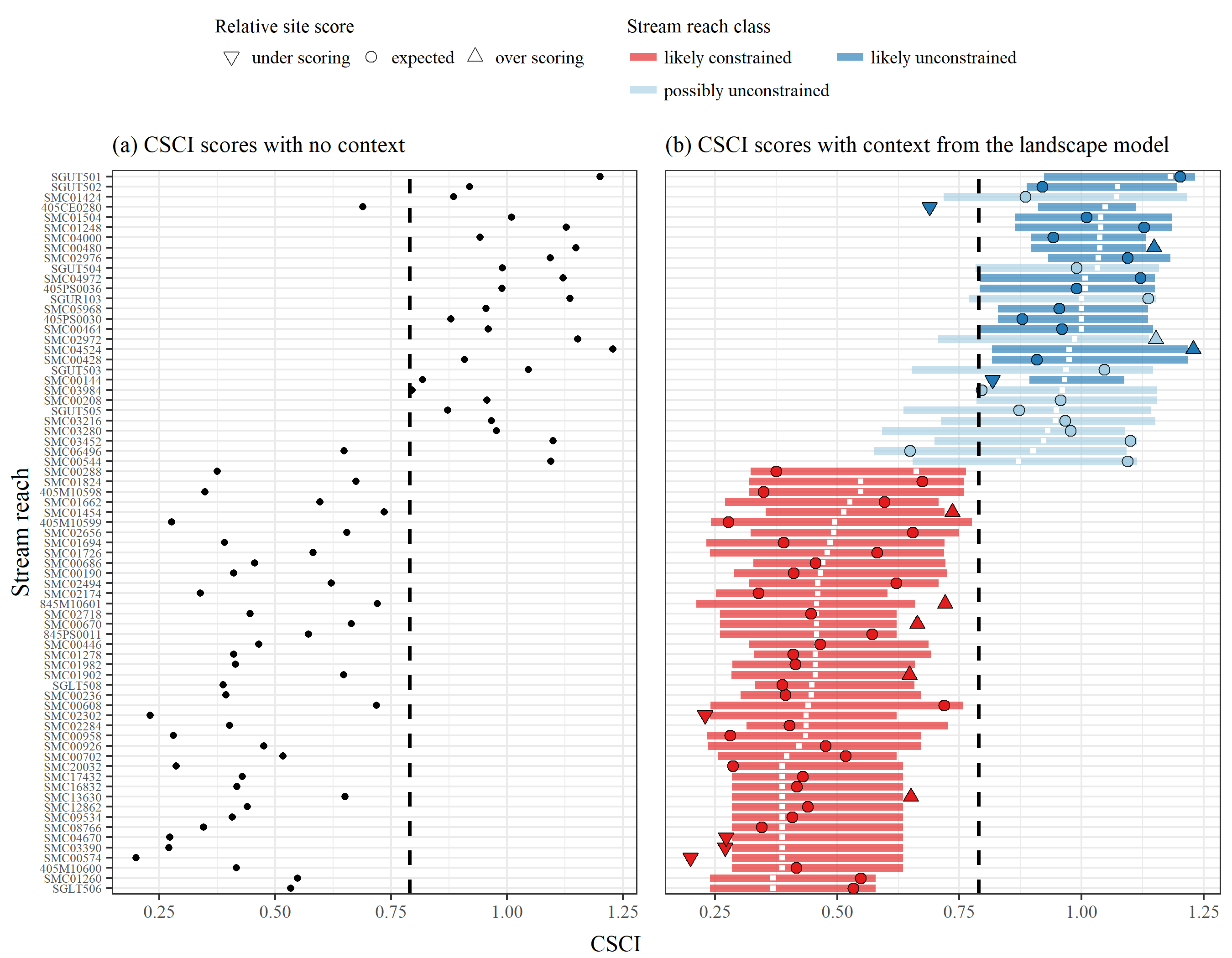


Figure 5 Application of the landscape model to stream reaches in the San Gabriel River watershed, Los Angeles County, California. CSCI scores with (a) no context from the model are on the left and (b) scores with context from the model are on the right. Relative site scores as under-scoring, expected, 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 biological threshold for the CSCI (0.79, dashed lined) and location of the median expectation (white ticks).

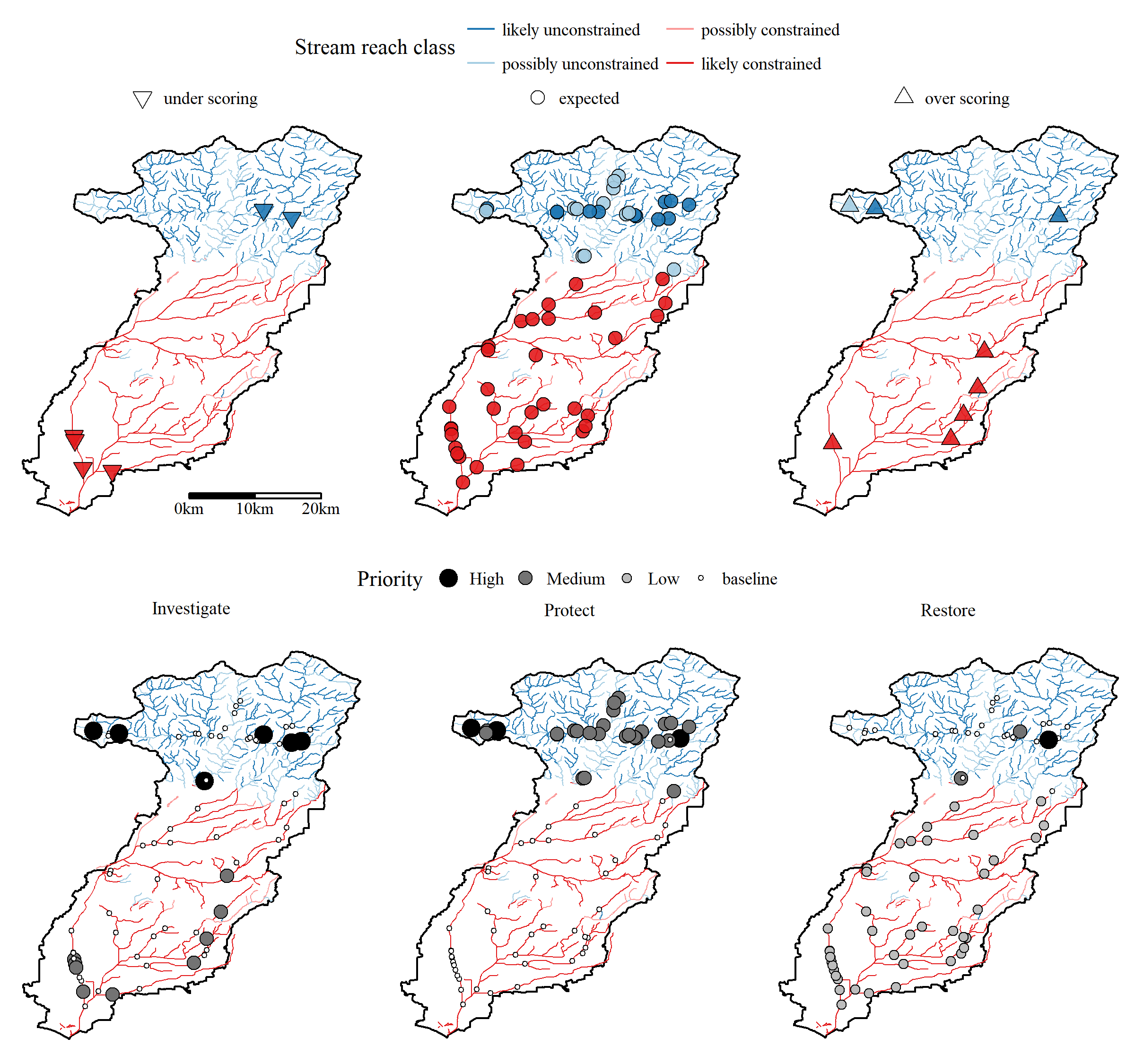


Figure 6 Relative site scores and recommended management actions for locations with CSCI scores in the San Gabriel River watershed. Relative site scores as under scoring, expected, or over scoring are based on observed scores given the reach class as likely constrained, possibly constrained, possibly unconstrained, and likely unconstrained. Recommended management actions are ranked by priority for actions to investigate, protect, and restore a site. No recommended actions assume baseline maintenance and monitoring is sufficient for a site. Recommended actions were defined by a local stakeholder group (see Figure S3).

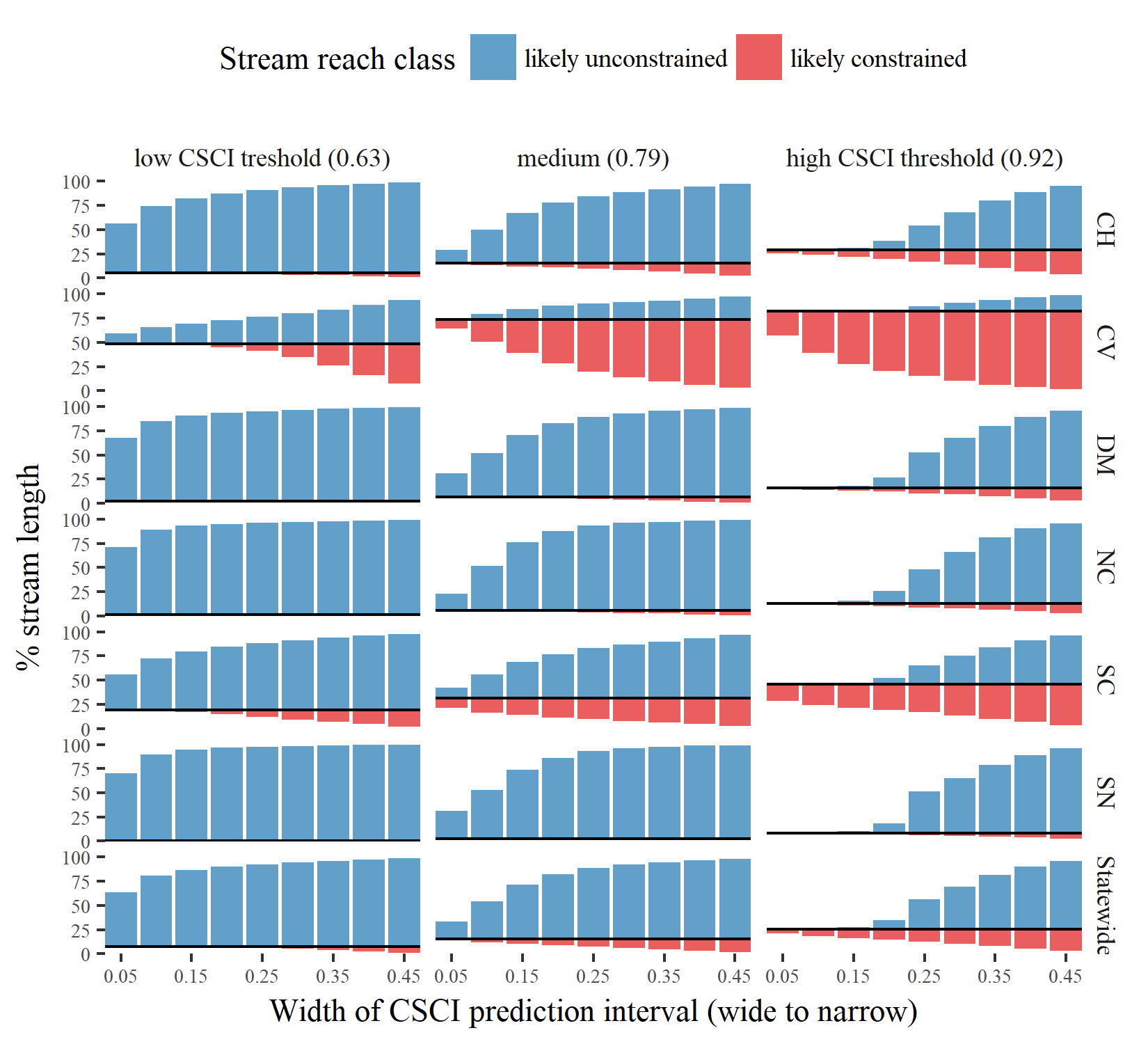


Figure 7 Changes in stream reach classes by region and statewide for different scenarios used to define biological constraints. The percentage of total stream length for likely unconstrained and likely constrained is shown for each scenario. Stream classifications as possibly unconstrained or possibly constrained are not shown but can be inferred form the area of white space above or below each bar. The solid black line indicates the percentage division between unconstrained and constrained classifications. Twenty-seven scenarios were tested that evaluated different combinations of certainty in the CSCI predictions (nine scenarios from wide to narrow prediction intervals as identified by the tail cutoff for the expected range) and potential CSCI threshold (three scenarios from low to high). CV: Central Valley, CH: Chaparral, DM: Deserts and Modoc Plateau, NC: North Coast, SN: Sierra Nevada, SC: South Coast.

# Tables

Table 1 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 2 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 defining 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 3: (#tab:clstot) Summary of stream length for each stream class statewide and major regions of California (Figures 1, 4). Lengths are in kilometers with the percentage of the total length in a region in parentheses. All lengths are based on a CSCI threshold of 0.79 and the 10th to 90th percentile of expected scores from the landscape model. CV: Central Valley, CH: Chaparral, DM: Deserts and Modoc Plateau, 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 4: (#tab:reltot) Summary of CSCI scores by relative expectations for each stream class statewide and in each major region of California (Figures 1, 4). Average (standard deviation) scores and counts (percent) 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 and Modoc Plateau, 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) |

# References

Allan, D., D. Erickson, and J. Fay. 1997. “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>.

Buss, D. F., D. M. Carlisle, T. -S. Chon, J. Culp, J. s. Harding, H. E. Keizer-Vlek, W. A. Robinson, S. Strachan, C. Thirion, and R. M. Hughes. 2014. “Stream Biomonitoring Using Macroinvertebrates Around the Globe: A Comparison of Large-Scale Programs.” *Environmental Monitoring and Assessment* 187:4132. <https://doi.org/10.1007/s10661-014-4132-8>.

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.

California Regional Water Quality Control Board, Los Angeles Region. 2014. “Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties.”

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>.

Davies, S. P., and S. K. Jackson. 2006. “The Biological Condition Gradient: A Descriptive Model for Interpreting Change in Aquatic Ecosystems.” *Ecological Applications* 16 (4):1251–66.

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>.

Florida Department of Environmental Protection. 2011. “Development of aquatic life use support attainment thresholds for Florida’s Stream Condition Index and Lake Vegetation Index.” DEP-SAS-003/11. Tallahassee, Florida: FDEP Standards; Assessment Section, Bureau of Assessment; Restoration Support.

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>.

Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham. 2011. “Completion of the 2006 National Land Cover Database for the Conterminous United States.” *Photogrammetric Engineering and Remote Sensing* 77 (9):858–64.

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>.

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.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Reah. 2012. “NHDPlus Version 2: User Guide.”

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>.

Midwest Biodiversity Insitute. 2016. “Identification of predictive habitat attributes for Minnesota streams to support tiered aquatic life uses.” MBI Technical Report MBI/OHPAN1518840. Columbus, Ohio: Midwest Biodiversity Institute, prepared on behalf of the Minnesota Pollution Control Agency.

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.

Moss, D., M. T. Furse, J. F. Wright, and P. D. Armitage. 1987. “The Prediction of the Macro-Invertebrate Fauna of Unpolluted Running-Water Sites in Great Britain Using Environmental Data.” *Freshwater Biology* 17 (1):41–52. <https://doi.org/10.1111/j.1365-2427.1987.tb01027.x>.

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., 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.

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.

Surface Water Ambient Monitoring Program. 2017. “SWAMP Quality Assurance Program Plan.” Final. Sacramento, California: Office of Information Management; Analysis, State Water Resources Control Board. <https://www.waterboards.ca.gov/water_issues/programs/swamp/qapp/swamp_QAPrP_2017_Final.pdf>.

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>.

US Environmental Protection Agency, Region 10. 2013. “Technical Support Docuemnt for EPA’s Action on the State of Oregon’s Revised Water Quality Standards for the West Division Main Canal.”

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.

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>.