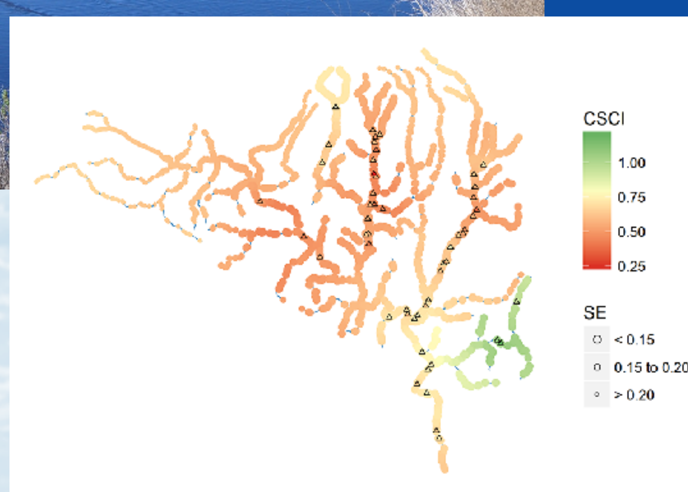
Spatial statistical network models to estimate the spatial representativeness of bioassessment samples

Raphael Mazor
Peter R. Ode
Andrew C. Rehn
Eric D. Stein



Southern California Coastal Water Research Project

SCCWRP Technical Report 979

Spatial statistical network models to estimate the spatial representativeness of bioassessment samples

Raphael Mazor[1], Peter R. Ode[2], Andrew C. Rehn[2], Eric D. Stein[1]

[1]Southern California Coastal Water Research Project, Costa Mesa, CA

[2]Aquatic Bioassessment Lab, Department of Fish and Wildlife, Rancho Cordova, CA

April 2017
SCCWRP Technical Report 979

ACKNOWLEDGEMENTS

This research was funded by the California Stormwater Quality Association. We thank Abel Santana, Liesel Tiefenthaler, and our technical panel: Ryan Hill (Environmental Protection Agency), Kevin Lunde (Regional Water Quality Control Board of San Francisco), Chris Sommers (EOA, Inc.), Clint Boschen (TetraTech, Inc.), and Michael Paul (TetraTech, Inc.).

EXECUTIVE SUMMARY

- Stream management and regulatory decisions are typically applied to large reaches or whole watersheds. However, assessment of stream condition relative to management or regulatory targets is based on discrete sampling reaches that are sparsely distributed across a watershed. Therefore, there is a need to extrapolate measured variables, such as bioassessment scores, to unmeasured reaches, and to understand uncertainty inherent to extrapolations.
- Spatial statistical network (SSNs) models allow estimation of bioassessment index scores at unsampled locations based on their proximity to sampled locations within a stream network. They facilitate the search for patterns at large spatial scales appropriate for biological indicators.
- We developed SSNs for 6 watersheds in northern and southern California to explore their utility in extrapolating scores for the California Stream Condition Index (CSCI).
- SSN models did not support a general distance limit to extrapolation that works in all settings (e.g., "a site score represents the condition of 800 m of a stream"), given the large variability observed among and within watersheds. On the contrary, the limits on extrapolation are more appropriately a site-specific determination.
- SSN models varied greatly among watersheds due to differences in watershed properties, patterns of degradation, and distribution of sampling locations. Models built for one watershed may not generalize to other watersheds, even if they share many environmental characteristics.
- SSN models offer a way to make site-specific and spatially explicit determinations by creating maps of extrapolated CSCI scores along a drainage network, that can support management decisions:
    - Maps may provide confidence in decisions about stream health, and whether they apply to upstream tributaries or downstream reaches.
    - Maps can identify regions where more sampling is required to improve confidence in estimates of condition.
    - Maps can be customized for diverse applications to reflect different levels of confidence they require.
    - Maps can be redrawn to incorporate new data as they become available.
- The drawbacks of a map-based approach are that the models will need to be generated for each watershed. However, the process could be streamlined to reduce the resources required for routine applications.
- Future efforts should explore the ability to create regional maps of similar watersheds to reduce effort and increase confidence in extrapolated scores. Models can and should be

developed for other management endpoints, such as algal indices of biotic integrity, hydromodification, and riparian wetland condition.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

INTRODUCTION

Biological indices such as the California Stream Condition Index (CSCI, Mazor et al. 2016) are increasingly being used in regulatory, management, and other monitoring programs focusing on stream health relative to beneficial use objectives. As the adoption of the CSCI in the regulatory and management environment expands in California, questions are increasingly arising over the best way to interpret scores. Among these questions is uncertainty about spatial representativeness. For example, how much distance along a reach does a score from a single site represent? Because bioassessment samples are typically collected from relatively short lengths of a stream (generally 150 m in California, Ode 2007) that are sparsely distributed throughout a watershed, there is often a need to extrapolate the results to adjacent reaches.

The field of spatial statistics offers a way to extrapolate scores from a sampled location to an unsampled location. By estimating the variance in chemical, physical, or biological measurements as a function of distance between sampling locations, spatial statistical models can estimate variables at intervening unsampled locations. Spatial statistical network models (SSNs), recently developed by Ver Hoef et al. (2016), are appropriate for fluvial systems because they incorporate up- and down-stream flow connectivity in characterizing distance between sampling locations.

Until recently, efforts to estimate biological condition at unsampled reaches made use of models based on land cover or other measured stressors (e.g., Carlisle et al. 2009, Falcone et al. 2010, May et al. 2015, Hill et al. in review). California has recently released statewide maps of estimated biological condition modeled from landuse and other stressors as part of its Healthy Watersheds Initiative (http://www.mywaterquality.ca.gov/monitoring_council/healthy_streams/index.html). Although these models are effective, they may be inappropriate for certain management applications, as they presuppose impacts from stressors without accounting for mitigating factors, such as habitat restoration or stormwater treatment. That is, these models use stressors as proxies for condition. In contrast, spatial statistical models make estimates from nearby sites with known conditions. Thus, they provide a different (and more direct) line of evidence about the condition of streams where bioassessment samples are lacking.

Spatial statistical models are not a new discipline, and are widely used in mining and other geological applications (Cressie 1993). They take advantage of the fact that samples close together tend to be similar (i.e., they are autocorrelated). Thus, the condition of an unsampled site can be estimated if sites nearby are sampled and scored (and estimates will be better if the sampled sites are closer). However, spatial models that incorporate the unique features of stream network topology have only recently become available (Ver Hoef and Peterson 2010, Ver Hoef et al. 2014, Ver Hoef et al. 2016). Spatial statistical network (SSN) models can assess autocorrelation in upstream and downstream directions independently, as well as "overland" autocorrelation among adjacent tributaries. SSNs have been used to model physico-chemical parameters in streams, such as temperature (e.g., Steel et al. 2016). SSNs have also been used to estimate biological parameters, such as fish population density (Isaak et al. 2017), as well as benthic macroinvertebrates bioassessment indices (Frieden et al. 2014).

We developed SSN models for 6 watersheds with at least 30 bioassessment sites in California to see if they could provide guidance on how to extrapolate CSCI scores to unsampled reaches. We developed models with both a “tail-up” (i.e., upstream) and “tail-down” (i.e., downstream) component, and evaluated the “ranges” (i.e., maximum distance of autocorrelation) for each. For one watershed (i.e., Alameda Creek), the influence of covariates was explored by adding developed land cover within a 500-m radius as a predictor in the model (this watershed was selected for exploring land cover covariates at random, prior to the development of models for any watershed). We subsequently applied the models to a dense array of prediction points in each watershed to generate maps of predicted condition at unsampled locations. We identified unsampled sites that could be designated with known confidence as either healthy or altered (i.e., sites whose prediction intervals were entirely above or below a threshold of biological condition). We then explored potential applications of SSNs and maps in different management scenarios.

METHODS

Sampling locations

Bioassessment data were collected at 3254 unique sites in California under a variety of programs (Figure 1). Benthic macroinvertebrates were sampled according to Ode 2007, and scored with the California Stream Condition Index (CSCI) following Mazor et al. (2016). The CSCI is a predictive index that compares observed taxa and metrics to values expected under reference conditions based on site-specific environmental variables, such as watershed area, geology, and climate (Ode et al. 2016). CSCI scores were classified as indicating “intact” or “altered” condition, using the normal approximation of the 10th percentile of CSCI reference calibration scores as a threshold (i.e., a score of 0.79).

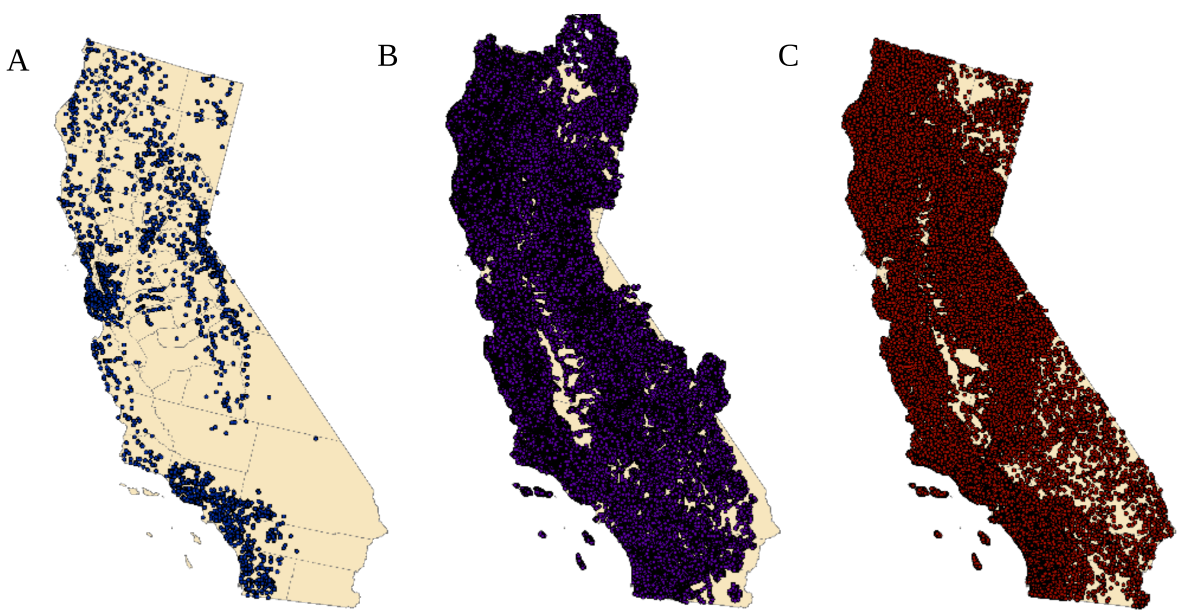


Figure 1. Sampling locations (A), prediction points (B) from the National Stream Internet (NSI), and evaluation points (C) from probabilistic surveys.

Stream network hydrography

The National Stream Internet Hydrography Network (NSI, Isaak et al. 2013) was used as a stream network for all analyses. The NSI is derived from the widely used National Hydrography Network Plus (NHD-Plus Version 2, medium resolution), but has been modified to facilitate spatial network analyses. For example, discontinuities are removed, and complex multi-thread channels are replaced with a single line representing the predominant direction of flow (Figure 2).

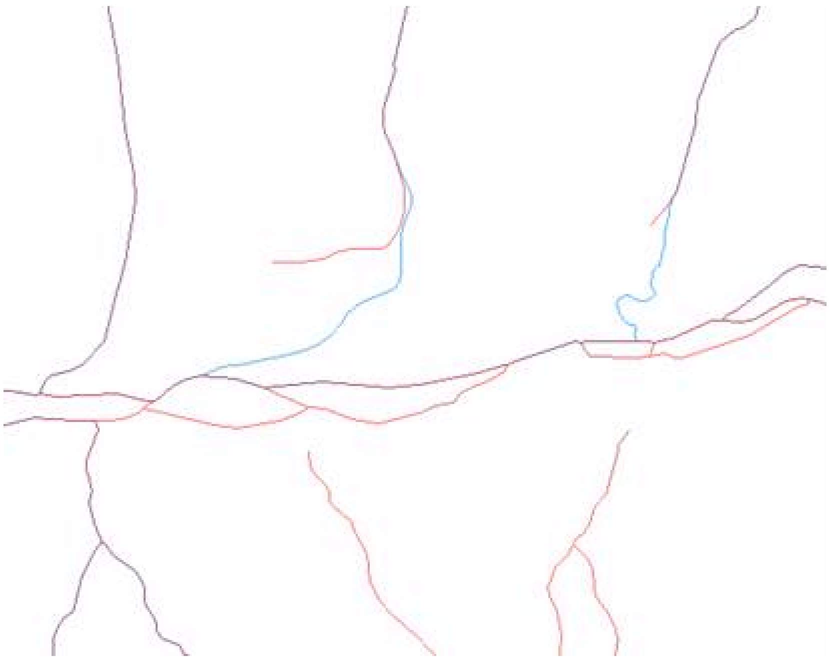


Figure 2: A portion of the mainstem of the Santa Clara River, as represented in the NHD-Plus (red) and the NSI (blue) hydrography networks. Although identical in many areas (purple), the NSI has fewer flow discontinuities, and represents multi-thread channels as a single line, which is more appropriate for spatial analyses.

Prediction points

Unsampled points for predicting bioassessment scores were derived from two sources: 1) a set of 134,696 points distributed as part of the NSI, specifically for use in spatial statistical network modeling; and 2) a set of 114,566 evaluation sites generated for probabilistic bioassessment surveys throughout the state, such as the statewide Perennial Streams Assessment, the Stormwater Monitoring Coalition in southern coastal California, and the Regional Monitoring Coalition in the Bay Area. The NSI prediction points are derived from the National Hydrography Dataset and have a density of approximately 1 site per stream-km in most regions of California. Because of the high number of surveys designed for southern California, the density of evaluation sites is substantially greater than the density of NSI prediction points in this region. Hereafter, both sets of points are collectively referred to as “prediction points”. (Figure 1)

Watersheds

A number of watersheds from throughout the state were selected for inclusion in the study (Figure 3, Table 1): Feather River, Alameda Creek, Coyote Creek, Malibu Creek, Los Angeles River, and Santa Margarita River. Watersheds were selected if they contained at least 30 bioassessment sites, and if a portion of the watershed was suspected to be affected by urban stormwater runoff. Notably, the watershed for Coyote Creek includes many hydrologically isolated creeks that drain directly to San Francisco Bay; other watersheds were largely comprised of single drainage networks with one outlet in each. In two watersheds (i.e., the Feather and Santa Margarita Rivers), the majority of sites had high CSCI scores (i.e., above the threshold of 0.79), but in the others, scores were typically lower.

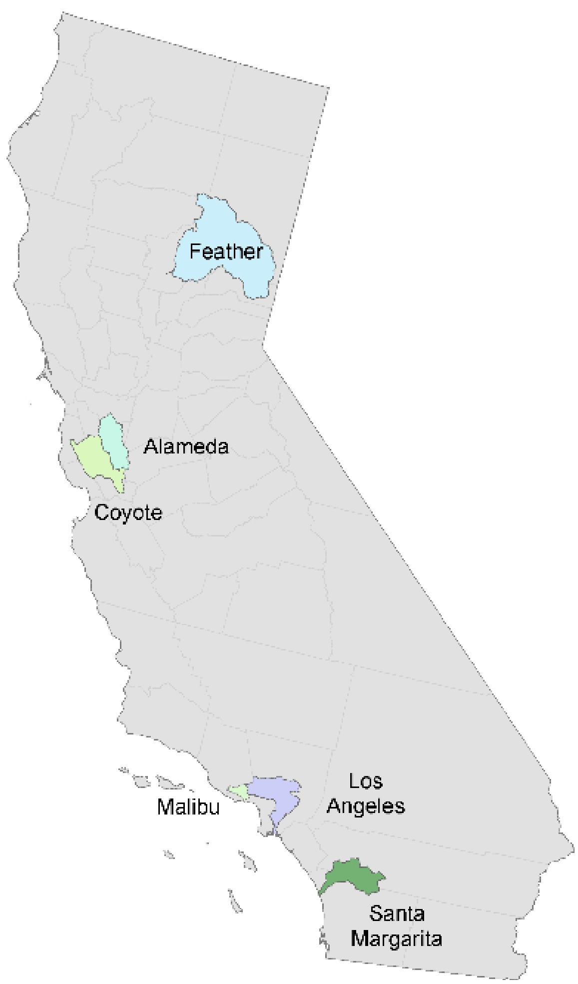


Figure 3. The six watersheds included in the analysis.

Table 1. Watershed summaries. NSI: National Stream Internet (Isaak et al. 2013). n: Number of sites sampled

	Watershed properties						Available data		
Watershed	Area (km^2)	Stream length in NSI (km)	% urban	% agricultural	% developed open space	Major cities	n	Mean CSCI	% CSCI ≥ 0.79
Feather River	9389	7377	0.3	1	1	Yuba City, Marysville	68	0.95	82
Alameda Creek	1638	1459	9	1	5	Pleasanton, Livermore	32	0.68	38
Coyote Creek	2174	1486	33	1	9	San Jose	101	0.70	40
Malibu Creek	285	260	12	2	13	Thousand Oaks, Agoura Hills	56	0.61	9
Los Angeles River	2160	1197	52	0	10	Los Angeles, Long Beach	91	0.63	31
Santa Margarita River	1924	1535	8	5	8	Temecula, Oceanside	32	0.87	69

Data processing to create spatial statistical networks

All methods follow Peterson (2015). The NSI layer was converted to a landscape spatial network (LSN) using the STARS package in ArcGIS (Peterson and Ver Hoef 2014). Network topology was checked for major errors (e.g., complex confluences or downstream divergences, as described in Peterson 2015), and corrected if any were found. Reach-contributing areas were created for each stream segment in the LSN, using NHD-Plus waterbodies to provide weights for additive function values (Ver Hoef and Peterson 2010). Sampling sites and prediction points were added to the LSN geodatabase, and all were exported as an R object for modeling. For one watershed (i.e., Alameda Creek), the area of developed land cover within a 500-m radius of the sampling or prediction point was also calculated for use as a covariate in models.

Modeling CSCI scores for spatial networks

The glmssn() function in the SSN package (Ver Hoef et al. 2014) was used to create generalized linear models that predict CSCI scores based on proximity to sampled locations. An SSN model can have up to three spatial autocorrelation components: a "tail-up" component (for autocorrelation in an upstream direction), a "tail-down" component (for autocorrelation in a downstream direction), and a Euclidean component (for autocorrelation over land, irrespective of the stream network). There are five types of functions that are appropriate for the tail-up and tail-down components (i.e., spherical, exponential, linear-sill, Mariah, and Epanechnikov), and four for the Euclidean components (i.e., spherical, exponential, Cauchy, and Gaussian). For details on these functions, refer to Ver Hoef and Peterson (2010). In addition to the three spatial components, a "nugget" component, representing the variability of multiple samples taken at a single location, is included in all models.

Because both up- and downstream autocorrelation may be important for extrapolating, we calibrated models for each watershed containing both a tail-up and a tail-down component (for all 25 possible combinations of correlation functions). We also calibrated three-component models containing Euclidean components (specifically, all-spherical and all-exponential) for comparison, but we did not consider a Euclidean component to be appropriate as it is more likely to reflect autocorrelation in land use rather than autocorrelation in biotic assemblages. Finally, we also calibrated a non-spatial model (also called "nugget-only" models), plus all one-component tail-up and tail-down models for comparison. These 3-component, 1-component, and nugget-only models provide benchmarks against which the selected 2-component models were compared. The InfoCritCompare() function in the SSN package was used to calculate a variety of measures of model performance, including the Aikake Information Criterion (AIC) value. The two-component model with the lowest AIC value was selected and refit using restricted maximum likelihood (restricted maximum likelihood is superior to maximum likelihood, but AIC comparisons are only valid when maximum likelihood is used when fixed effects change in different models, Verbeke and Molenberghs 2000). For Alameda Creek, the selected model was also fit with a covariate (i.e., development within a 500-m radius of the point). Bias was estimated using leave-one-out cross validation, as implemented by the InfoCritCompare() function.

Creation of maps

Selected models were used to predict CSCI scores at all prediction points in each watershed (up to 4277 sites in a single watershed). Because glmssn models also generate prediction standard errors (SE), prediction intervals could also be calculated. We calculated the percent of sites with small SE (i.e., SE $\leq$ 0.15) and medium SE (i.e., SE $\leq$ 0.2) for each watershed. The 75%, 90%, 95%, and 99% prediction intervals were calculated as the predicted score, plus or minus the z-score times the SE (e.g., 1.96 for the 95% prediction interval, 2.58 for the 99% prediction interval, etc.). Two types of maps were then created: One showing predicted CSCI score, and one highlighting sites where the prediction interval was above or below 0.79 (i.e., the threshold used to identify altered streams in recent condition assessments, e.g., Mazor 2015) for all 4 prediction intervals.

Results

Model calibration and performance

A total of 38 glmssn models were created for each watershed, in addition to two models that included land use covariates for Alameda Creek. One-component (i.e., tail-up or tail-down), two-component (i.e., tail-up and tail-down), and three-component (i.e., tail-up, tail-down, and Euclidean) models were all superior to nugget-only (non-spatial) models in terms of having lower root-mean squared prediction errors (RMSPE). In two watersheds (Coyote and Los Angeles), the three-component models had substantially lower errors than all two-component models, but in the other 4 watersheds, the addition of a Euclidean component did not lead to a major improvement (Figure 4). Overall performance varied widely among the watersheds. The lowest errors were for the Malibu Creek model (RMSPE: 0.11), followed by the Santa Margarita and Alameda Creek models (RMSPEs: 0.14); the Feather and Los Angeles Rivers models had larger errors (RMSPEs: 0.17), followed by Coyote Creek (RSMPE: 0.20).

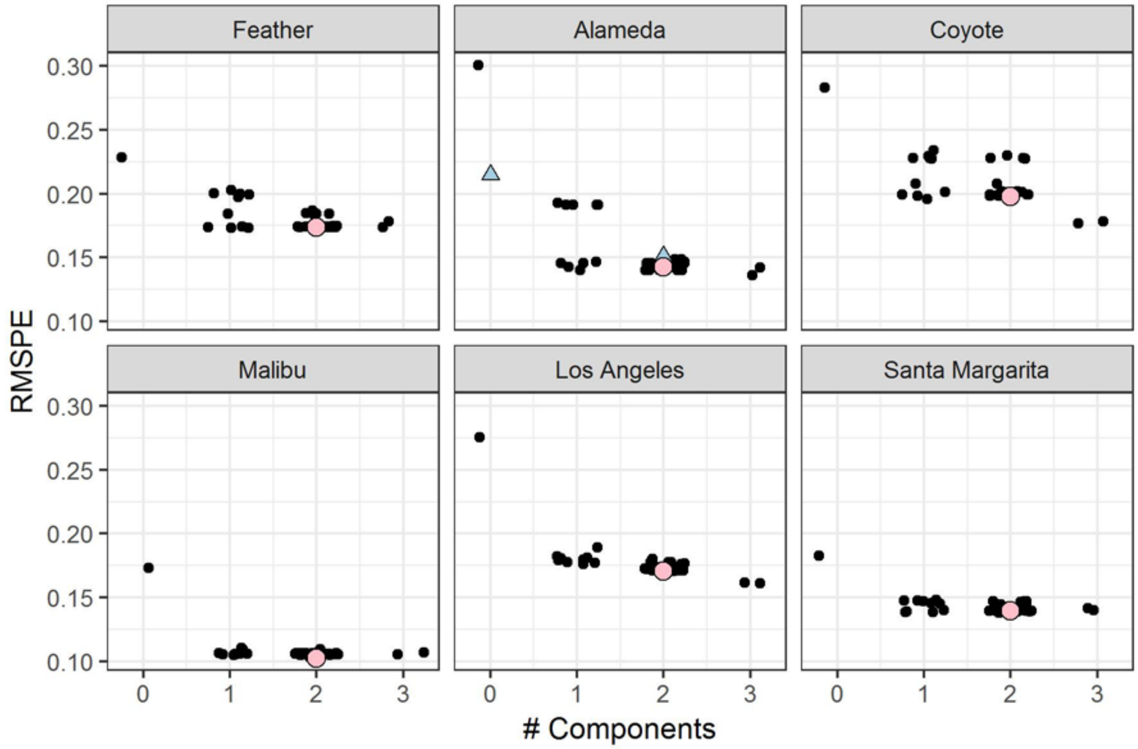


Figure 4. Root-mean squared prediction error versus the number of spatial components in glmssn models for each watershed. Each symbol represents a single model. Pink circles represent the selected two-component models with the lowest AIC. Blue triangles represent models with local landcover covariates (Alameda Creek alone). 0 component models: Nugget-only (nonspatial) models. 1-component models: Tail-up or tail-down models. 2-component models: Tail-up and tail-down models. 3-component models: Tail-up, tail-down, and Euclidean component models.

The selected models varied in terms of the importance of each spatial component, and the autocorrelation models they used. The taildown component was generally more important than the tailup component, although it was negligible (i.e., 0.01) for the Feather River, and less important than the tailup component

Table 2. Performance of selected models. n: Number of sites in model calibration. Pearson: Pearson correlation coefficient between observed values and values predicted by leave-one-out cross validation (LOOCV). GR2: Generalized R-squared of covariates. RMSPE: Root-mean squared prediction error. TU: Tailup. TD: Taildown. NU: Nugget. COV: Land-cover covariate. Sph: Spherical model. Lin: Linear-with-sill model. Epa: Epanechnikov model. Exp: Exponential model. --: Not applicable.

					Correlation function		Range (km)		Partial sill			Variance component			
Model	n	O vs. E r	GR2	RMSPE	TU	TD	TU	TD	NU	TU	TD	COV	NU	TU	TD
Two-component models															
Feather	68	0.64	--	0.17	Sph	Lin	1602	761	0.011	0.032	0	--	0.26	0.73	0.01
Alameda	32	0.87	--	0.14	Sph	Epa	2	36	0.009	0	0.054	--	0.15	0	0.85
Coyote	101	0.03	--	0.20	Exp	Epa	0	28	0	0.018	0.083	--	0	0.18	0.82
Malibu	56	0.78	--	0.10	Epa	Lin	30	15	0.007	0	0.039	--	0.16	0	0.84
Los Angeles	91	0.34	--	0.17	Epa	Epa	140	56	0.015	0.020	0.056	--	0.16	0.22	0.62
Santa Margarita	32	0.61	--	0.14	Epa	Epa	287	39	0.010	0.019	0.009	--	0.26	0.51	0.23
Alameda models with land use covariate															
Spatial	32	0.86	0.01	0.15	Sph	Epa	3	429	0.011	0	0.594	0.01	0.02	0	0.97
Nonspatial	32	0.68	0.52	0.21	--	--	--	--	0.044	--	--	0.52	0.48	--	--

The ranges of the models were often very large in at least one direction, frequently indicating that spatial autocorrelation could be detected throughout the entire drainage network. That is, the ranges were often close to or larger than the maximum distance between two flow-connected sites in the watershed (Table 2). For example, the tailup range for Malibu Creek was 30 km, while the maximum flow-connected distance was 37 km. However, despite the large ranges, the partial sills were small (i.e., <0.1), suggesting that spatial autocorrelation can be detected at great distances, although the effect of autocorrelation may be weak.

For most watersheds (with the exception of Coyote Creek, and to a lesser extent, the Los Angeles River), predicted values along the drainage network were close to observed values. For example, the Pearson's correlation coefficient between observed and cross-validated prediction values ranged from 0.34 (for the Los Angeles River) to 0.87 (for Alameda Creek). In contrast, Pearson's correlation coefficient was only 0.03 for Coyote Creek, meaning that the spatial models were effectively predicting the average CSCI score for all sites in the watershed (Figure 5, Table 2). Models were unbiased for four watersheds, but showed fairly strong bias for Coyote Creek and the Los Angeles River, where the model under-predicted high-scoring sites and over-predicted low-scoring sites (Figure 6).

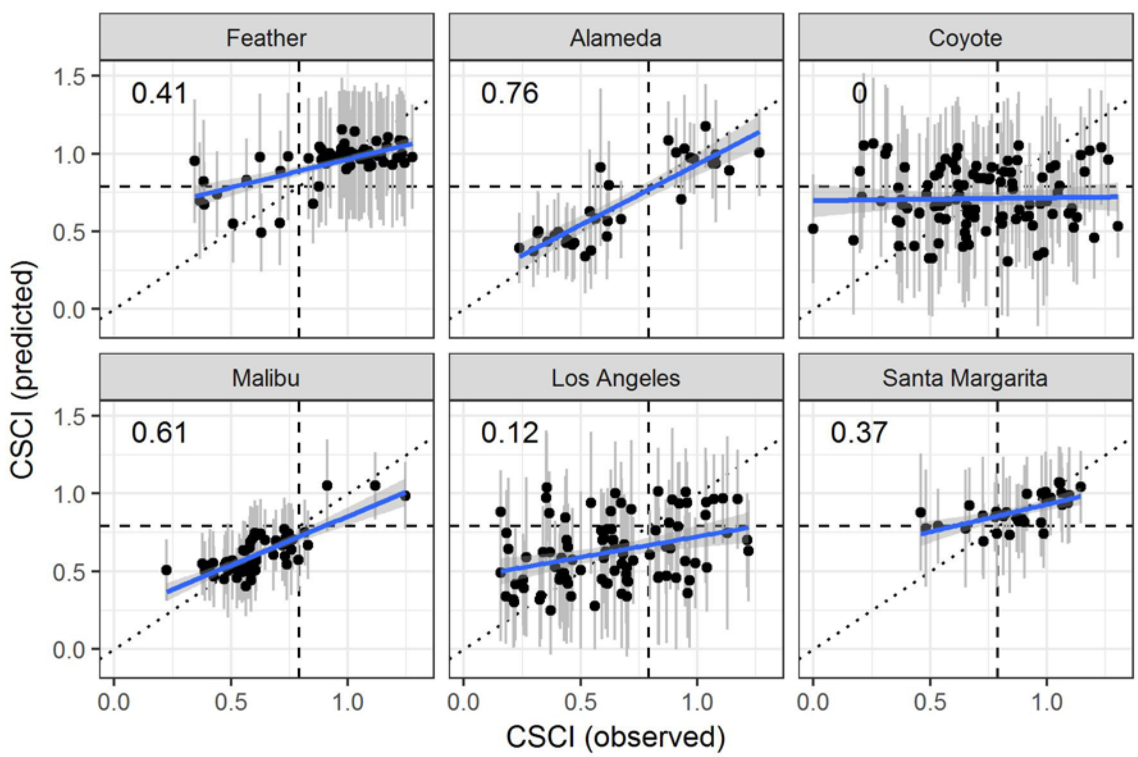


Figure 5. Predicted versus observed CSCI scores in each watershed. Numbers indicate Pearson's r-squared values. Black dots represent the prediction for each site, and the vertical gray lines represent the 95% prediction interval. Dashed lines represent the 0.79 threshold. The solid blue line represents a linear fit between observed and predicted values, and the gray band represents the 95% confidence interval of the fit. The dotted line represents perfect predictions.

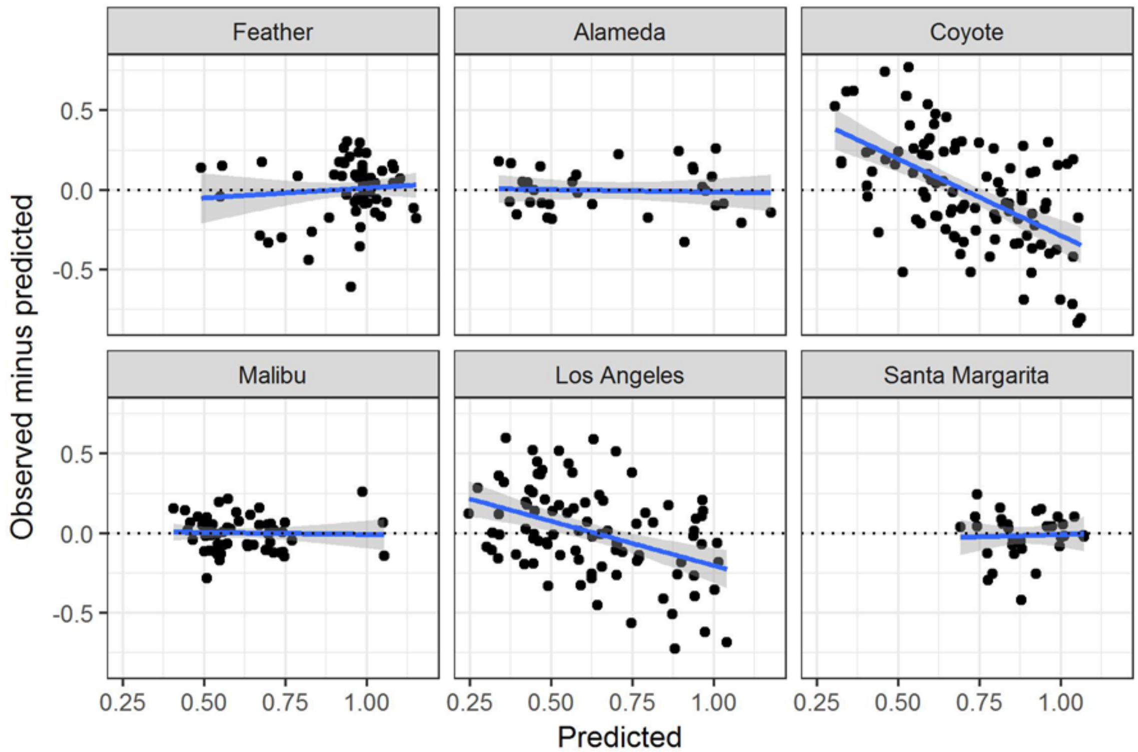


Figure 6. Differences between observed and predicted values versus predicted values. Black dots represent the prediction for each site. The solid blue line represents a linear fit between the difference and predicted values, and the gray band represents the 95% confidence interval of the fit. The dotted line represents perfect predictions.

Inclusion of a local landcover covariate had a negligible impact on the one spatial-only model where this effect was evaluated (i.e., the model for Alameda Creek). For example, the RMSPE was similar for both the spatial-only (0.14) and the spatial + covariate (0.15) models (Table 2). The generalized R-squared of the land use variable was only 0.01 in the spatial + covariate model, but 0.52 in the nonspatial model. Therefore, although local landcover could predict CSCI scores, it did not improve predictions based on spatial autocorrelation with nearby sampled sites. Notably, the Alameda Creek spatial-only model was the best of the 6 watersheds that were evaluated, likely limiting the potential for improvement.

Application to unsampled sites

The models were used to predict CSCI scores at all prediction points in each of the 6 watersheds (Figures 7A-7F). The majority of sites in three watersheds (Feather, Alameda, and Santa Margarita) were estimated to have CSCI scores above the 0.79 threshold, but less than half the sites in Coyote Creek and the Los Angeles River watersheds—and only 7% of sites in Malibu Creek—were estimated to have scores this high (Table 3).

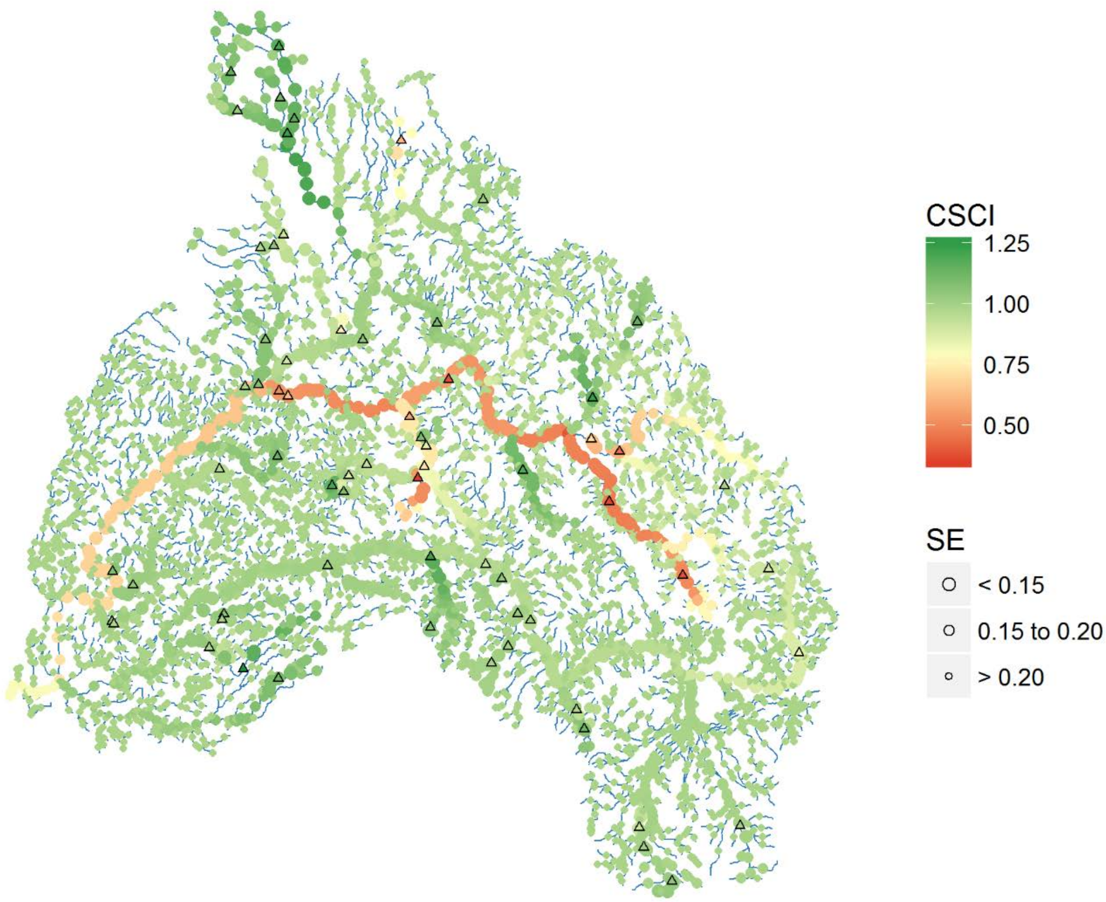


Figure 7A. Predicted CSCI scores for sites in the Feather River watershed. Large circles are more confidently estimated than small circles. Color indicates predicted score. Triangles indicate sampling locations, with color indicating observed score. SE: Prediction standard error.

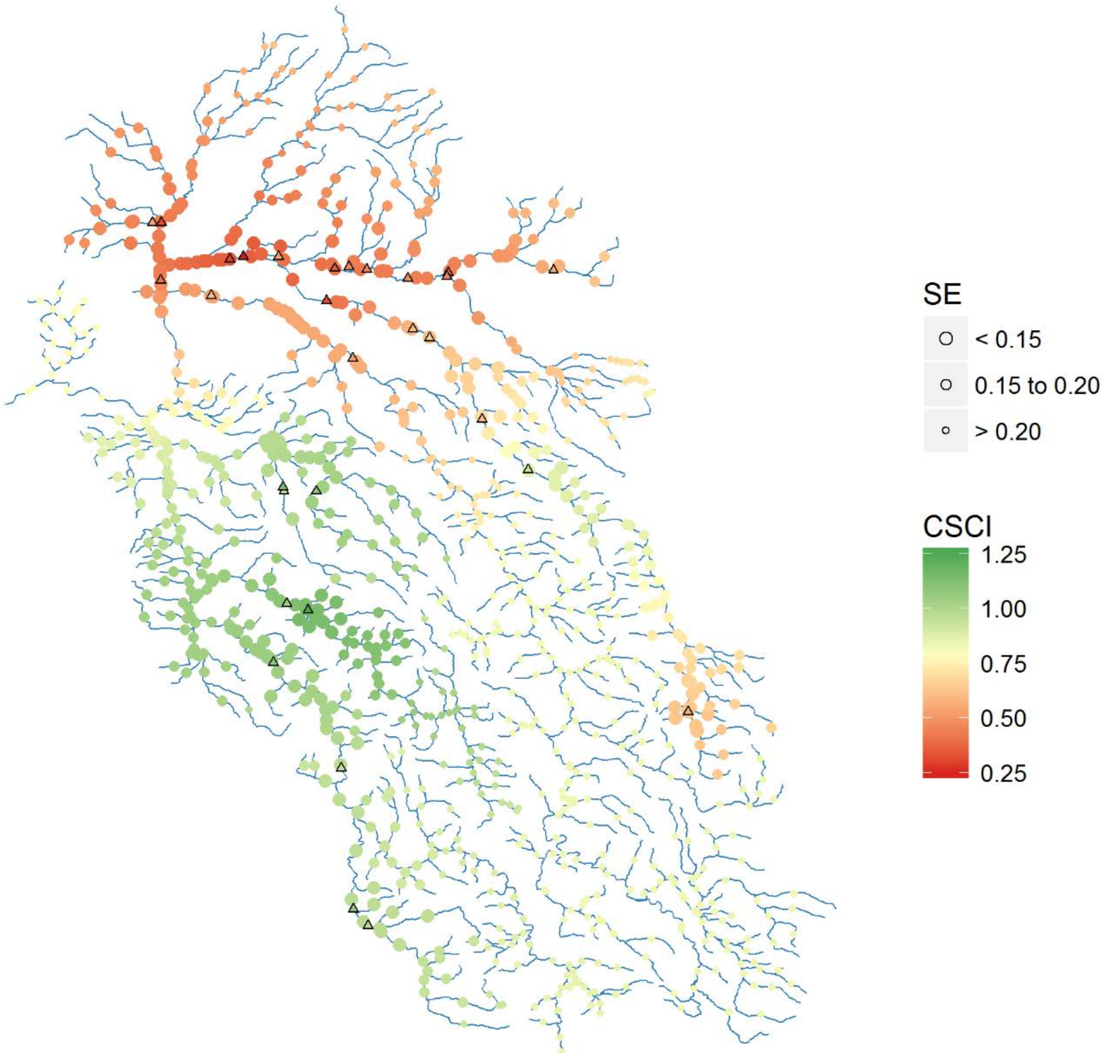


Figure 7B. Predicted CSCI scores for sites in the Alameda Creek watershed. Large circles are more confidently estimated than small circles. Color indicates predicted score. Triangles indicate sampling locations, with color indicating observed score. SE: Prediction standard error.

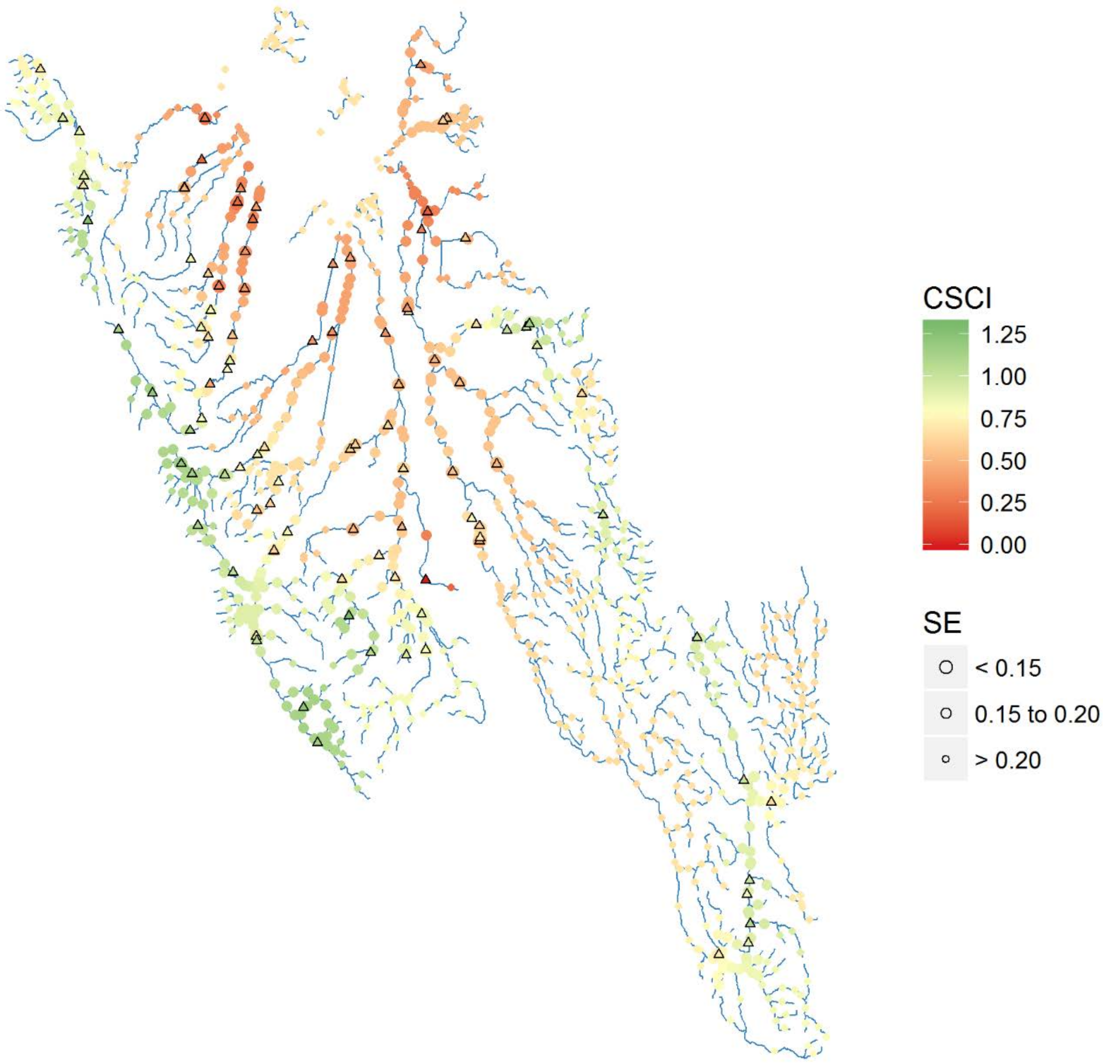


Figure 7C. Predicted CSCI scores for sites in the Coyote Creek watershed. Large circles are more confidently estimated than small circles. Color indicates predicted score. Triangles indicate sampling locations, with color indicating observed score. SE: Prediction standard error.

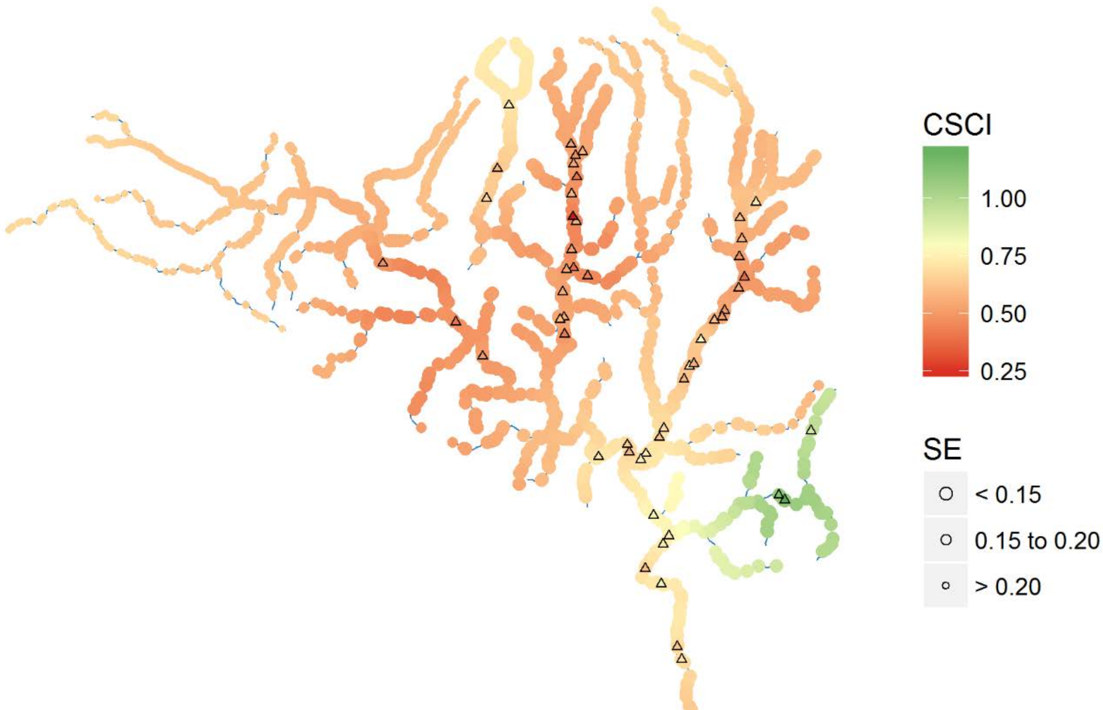


Figure 7D. Predicted CSCI scores for sites in the Malibu Creek watershed. Large circles are more confidently estimated than small circles. Color indicates predicted score. Triangles indicate sampling locations, with color indicating observed score. SE: Prediction standard error.

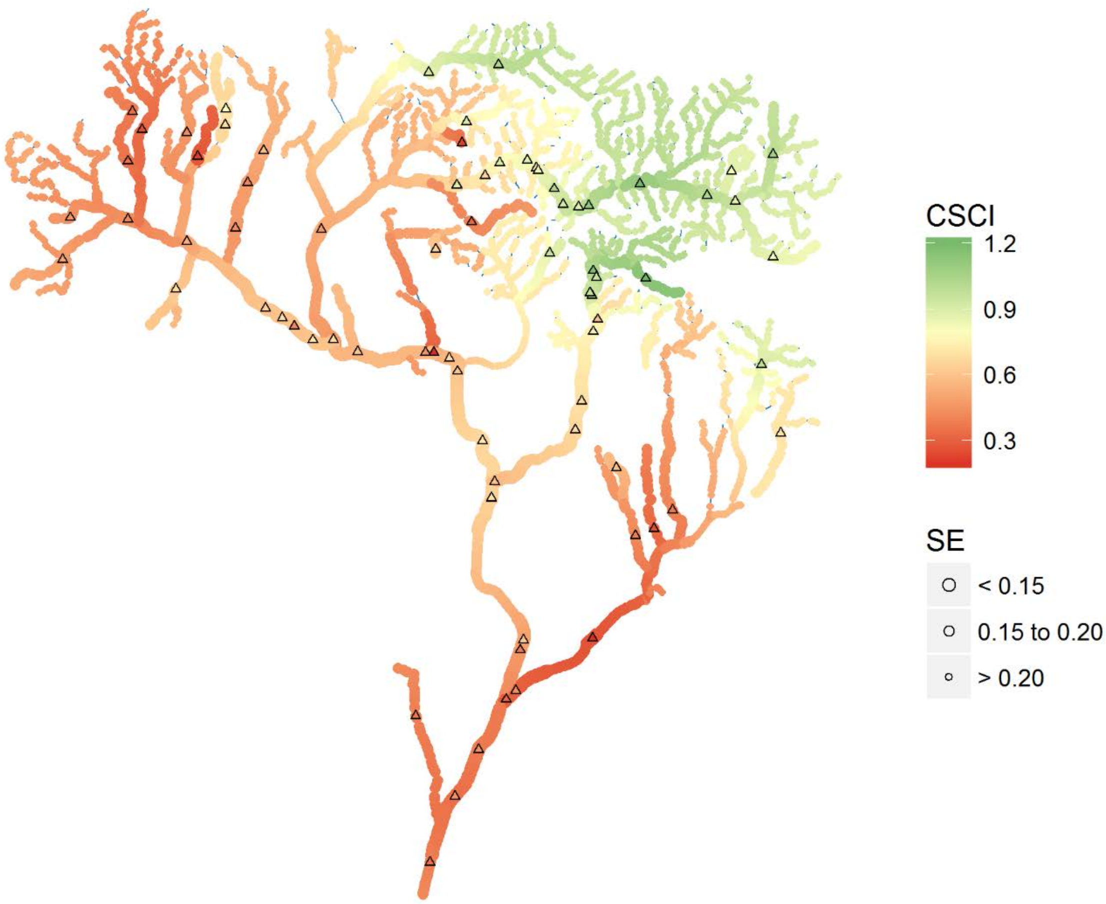


Figure 7E. Predicted CSCI scores for sites in the Los Angeles River watershed. Large circles are more confidently estimated than small circles. Color indicates predicted score. Triangles indicate sampling locations, with color indicating observed score. SE: Prediction standard error.

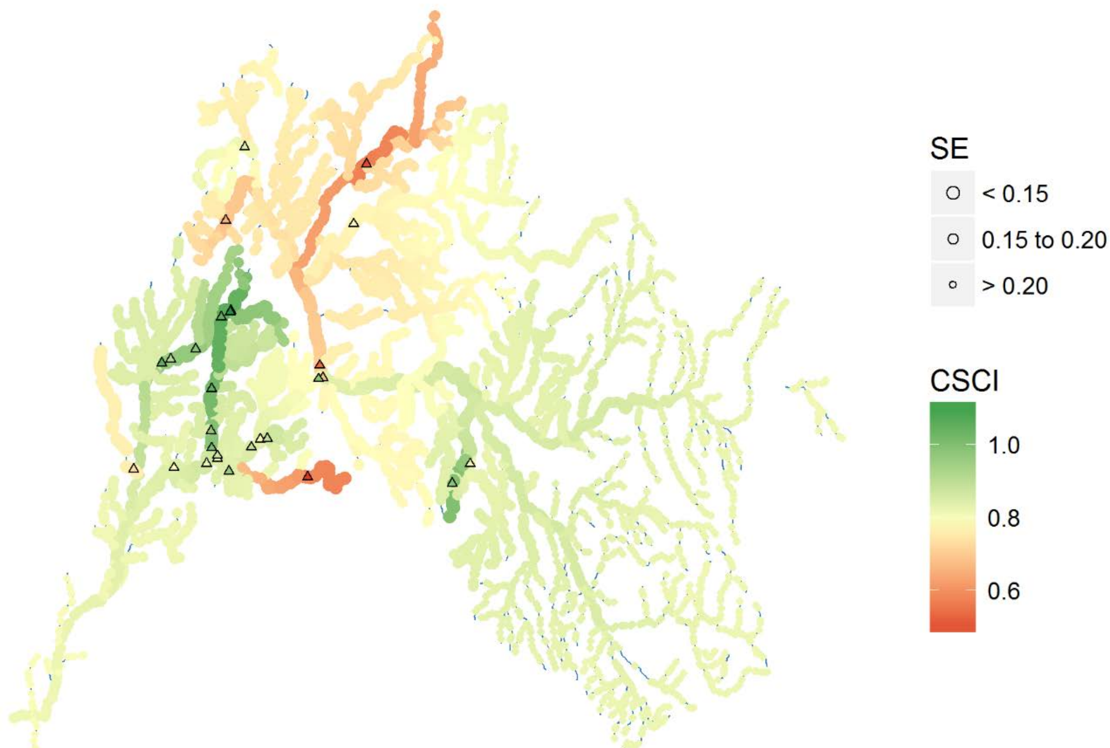


Figure 7F. Predicted CSCI scores for sites in the Santa Margarita River watershed. Large circles are more confidently estimated than small circles. Color indicates predicted score. Triangles indicate sampling locations, with color indicating observed score. SE: Prediction standard error.

Table 3. CSCI scores and prediction intervals (PI) at prediction points in each watershed. SD: standard deviation. SE: prediction standard error.

		CSCI			% sites with SE ≤		% Sites with PI ≥ 0.79				% Sites with PI < 0.79			
Watershed	# Prediction points	Mean	SD	Score ≥ 0.79	0.15	0.20	0.99	0.95	0.9	0.75	0.99	0.95	0.9	0.75
Feather	4277	0.97	0.09	4062	21%	54%	0%	2%	3%	13%	0%	1%	2%	2%
Alameda	817	0.82	0.17	558	4%	44%	1%	3%	8%	19%	4%	6%	9%	12%
Coyote	755	0.75	0.17	307	12%	26%	0%	1%	3%	9%	1%	3%	5%	11%
Malibu	142	0.62	0.12	10	17%	50%	1%	1%	4%	5%	10%	25%	39%	57%
Los Angeles	541	0.74	0.22	242	55%	78%	0%	1%	3%	7%	4%	8%	13%	24%
Santa Margarita	748	0.81	0.06	545	13%	72%	0%	1%	2%	3%	0%	0%	0%	1%

Precision of the predictions also varied among watersheds. For example, only 4% of prediction points in Alameda Creek had SE < 0.15, as opposed to 55% of sites in the Los Angeles River (Table 3). The best precision was for watersheds with a high density of samples per stream-km (e.g., Los Angeles River, Malibu Creek), or low variability in CSCI scores at sampled sites (e.g., Feather River, Santa Margarita). The number of sites whose prediction intervals were above or below the 0.79 threshold varied by watershed. For example, for the 75% prediction interval, 62% of the sites in Malibu Creek could be confidently designated as having scores above or (far more frequently) below the 0.79 threshold. In contrast, 96% of the Santa Margarita sites had a 75% prediction interval that straddled the threshold, meaning that only 4% could be confidently designated (Table 3, Figures 8A-F).

As the prediction intervals increased in size, the number of sites with confident designations decreased, as expected. At the highest level of confidence analyzed (i.e., 99% prediction interval), 1% or fewer sites were designated as healthy in any watershed, and only 10% were designated as altered in the Malibu watershed, where 90% of samples had CSCI scores below 0.79.

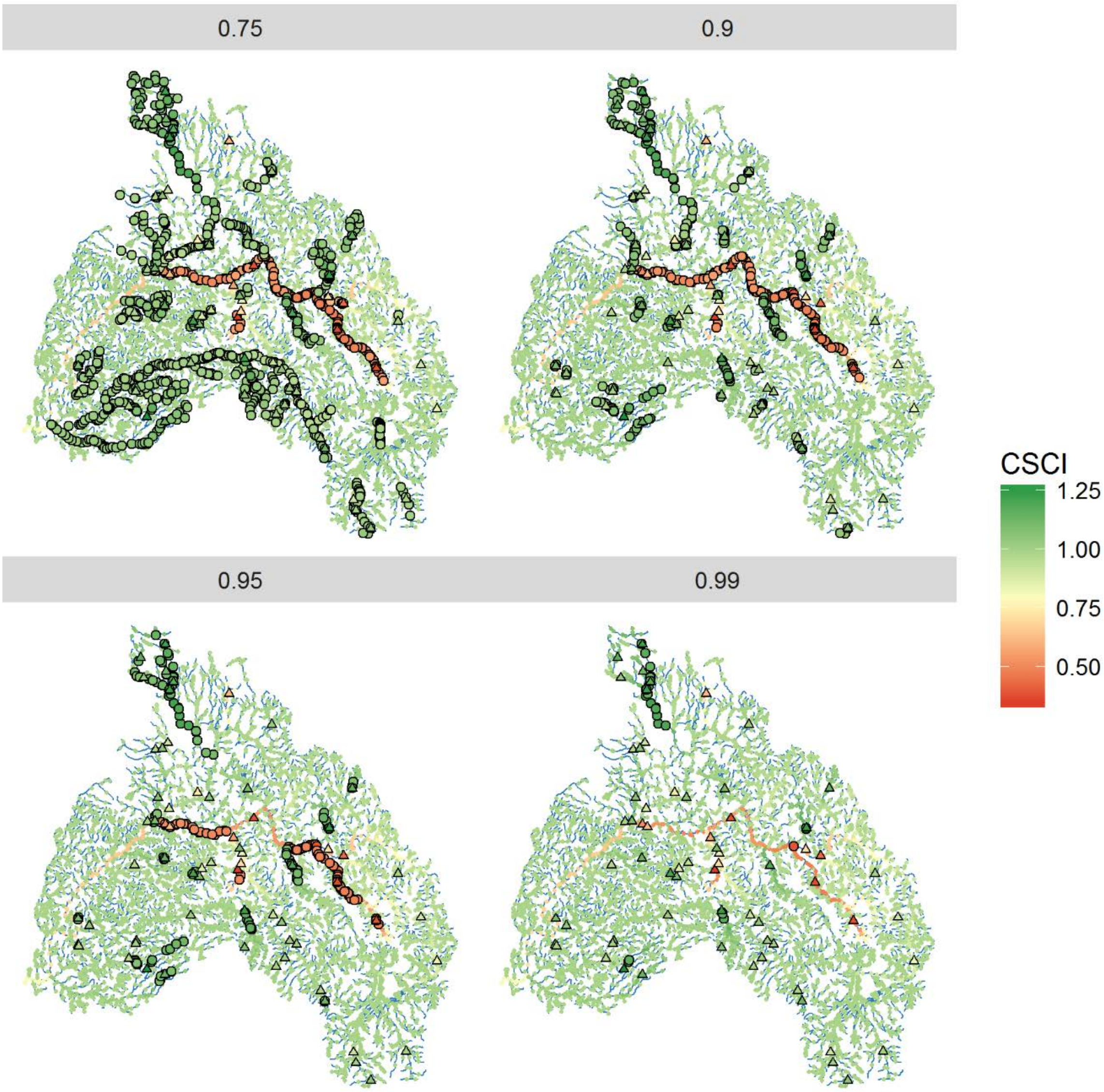


Figure 8A. Predicted CSCI scores in the Feather River watershed. Outlined circles represent sites whose prediction intervals are above or below the 0.79 threshold. The number above the panel indicates the prediction interval (i.e., 0.75, 0.90, 0.95, and 0.99). Triangles represent sampling locations.

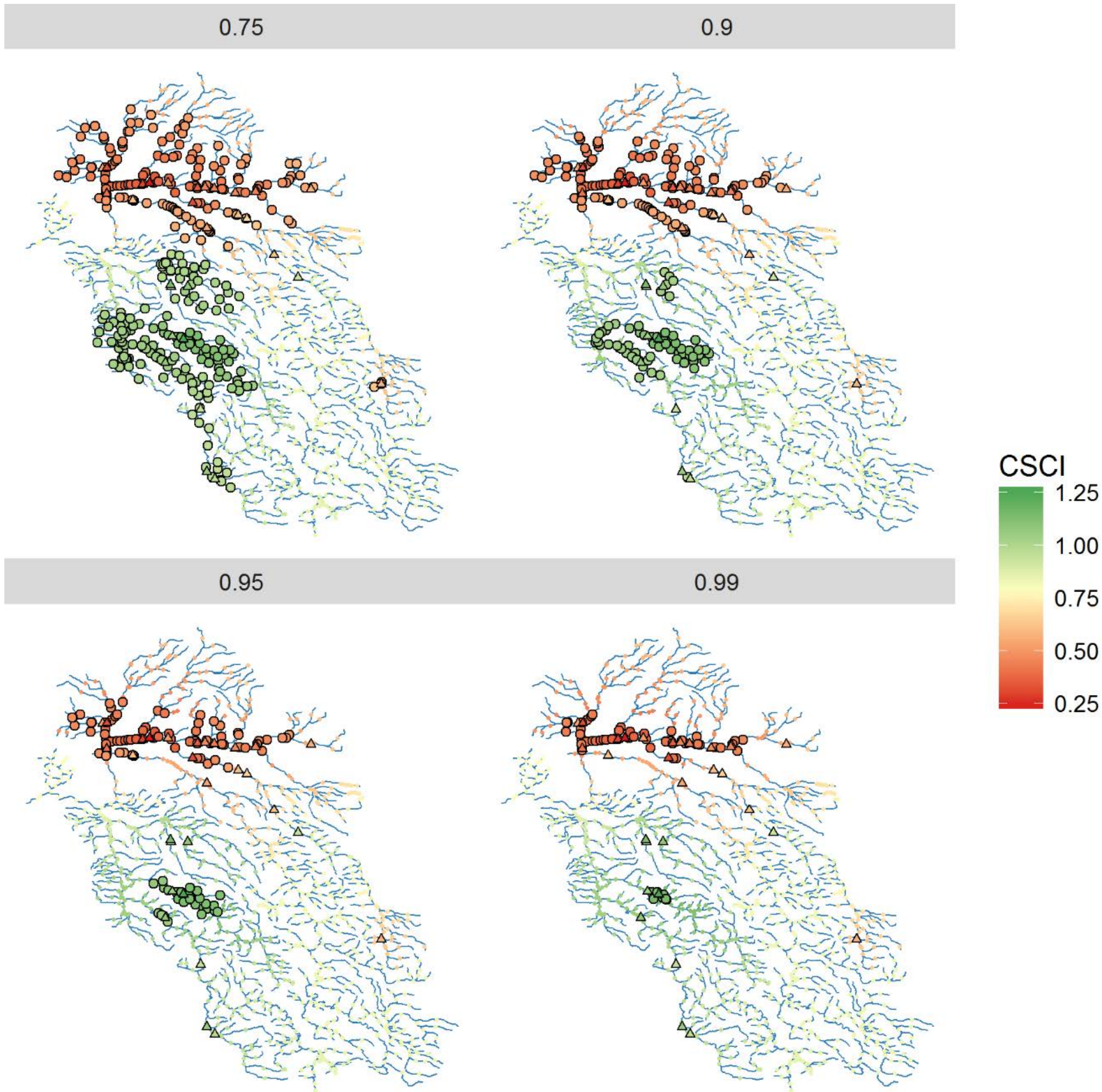


Figure 8B. Predicted CSCI scores in the Alameda Creek watershed. Outlined circles represent sites whose prediction intervals are above or below the 0.79 threshold. The number above the panel indicates the prediction interval (i.e., 0.75, 0.90, 0.95, and 0.99). Triangles represent sampling locations.

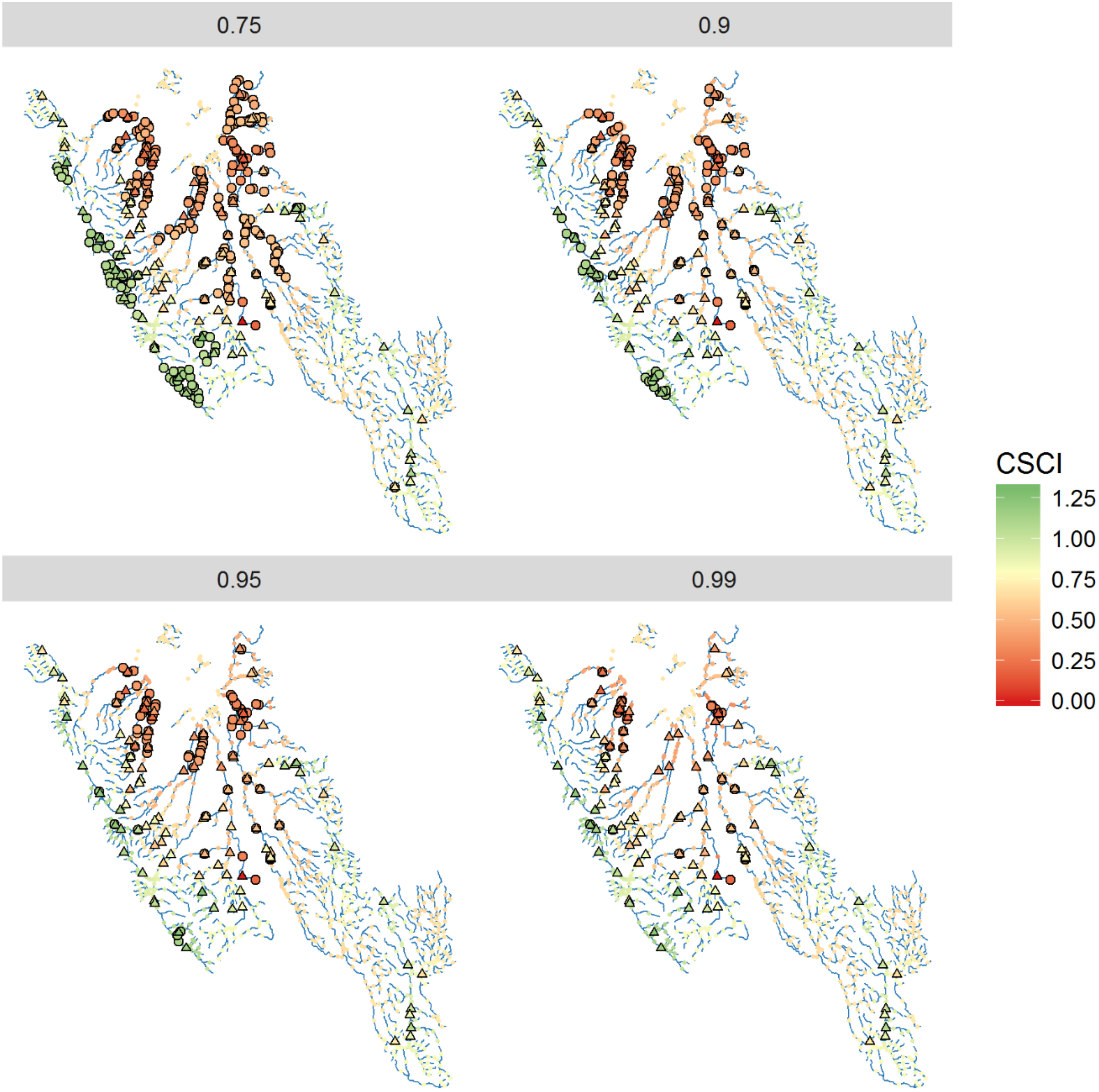


Figure 8C. Predicted CSCI scores in the Coyote Creek watershed. Outlined circles represent sites whose prediction intervals are above or below the 0.79 threshold. The number above the panel indicates the prediction interval (i.e., 0.75, 0.90, 0.95, and 0.99). Triangles represent sampling locations.

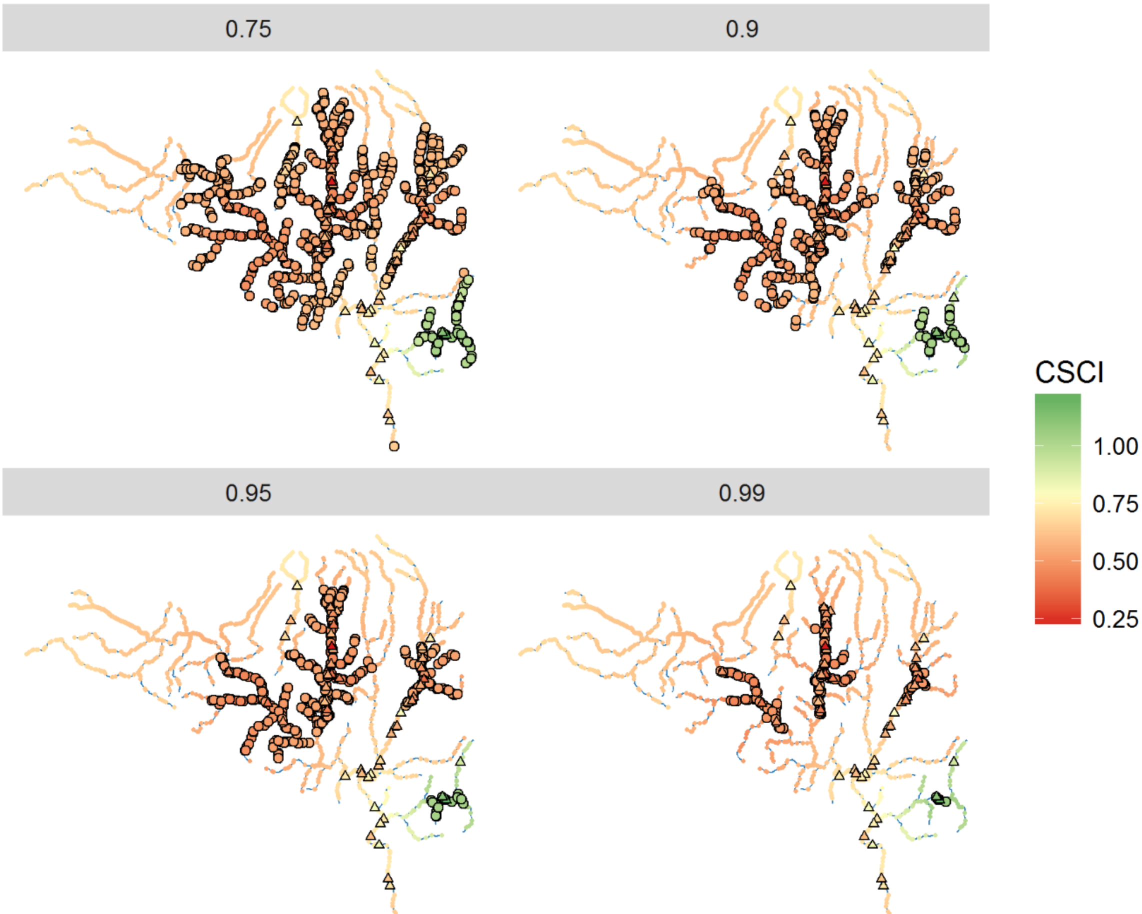


Figure 8D. Predicted CSCI scores in the Malibu Creek watershed. Outlined circles represent sites whose prediction intervals are above or below the 0.79 threshold. The number above the panel indicates the prediction interval (i.e., 0.75, 0.90, 0.95, and 0.99). Triangles represent sampling locations.

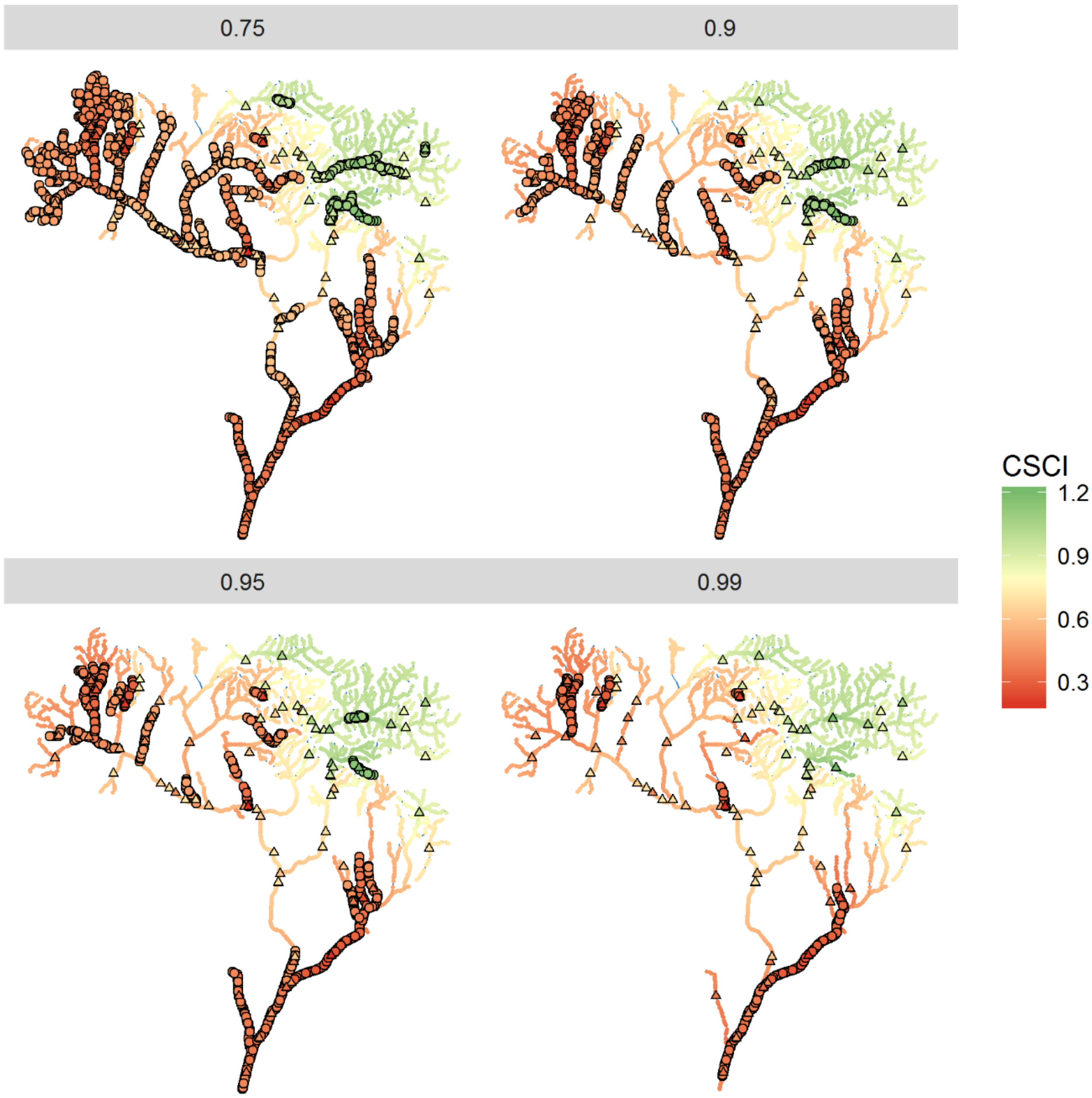


Figure 8E. Predicted CSCI scores in the Los Angeles River watershed. Outlined circles represent sites whose prediction intervals are above or below the 0.79 threshold. The number above the panel indicates the prediction interval (i.e., 0.75, 0.90, 0.95, and 0.99). Triangles represent sampling locations.

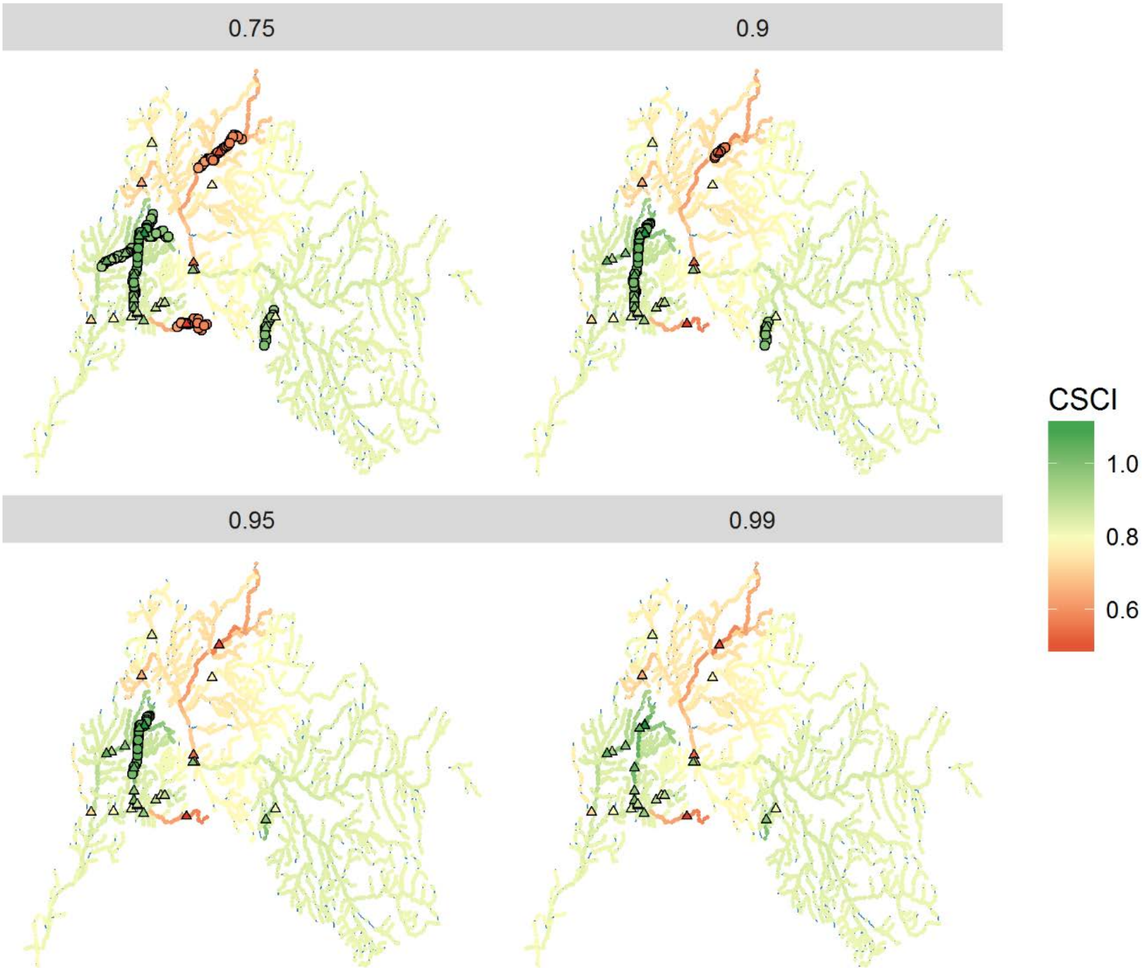


Figure 8F. Predicted CSCI scores in the Santa Margarita River watershed. Outlined circles represent sites whose prediction intervals are above or below the 0.79 threshold. The number above the panel indicates the prediction interval (i.e., 0.75, 0.90, 0.95, and 0.99). Triangles represent sampling locations.

Discussion

Factors affecting model performance

SSN models were effective at estimating CSCI scores at unsampled sites. Across the 6 watersheds, SSN models were able to estimate CSCI scores at 53% of prediction points with a SE of 0.2 or smaller. Furthermore, 11% of all evaluated prediction points across the 6 watersheds could be designated as healthy or altered relative to a threshold of 0.79, based on the 75% prediction interval. Some models, like Malibu creek and Alameda, were better than others, and the model for Coyote Creek was particularly ineffective. A high sampling density in a watershed with relatively homogenous CSCI scores is likely to contribute to successful predictions, as was the case in Malibu Creek. Although many sites were sampled throughout Coyote Creek, the spatial model was poor because most of the sites are not truly connected in a hydrologic sense. That is, most samples were on small isolated creeks draining directly to San Francisco Bay. In contrast, the other watersheds in the study were, with minor exceptions, true "watersheds" with many sites hydrologically connected to each other. In hydrologically connected watersheds, a purely spatial SSN approach works well to predict scores at unsampled sites (and better than a traditional non-spatial model based on surrounding land cover in the case of Alameda Creek, Table 2). In hydrologically disconnected locations (like Coyote Creek), predictions from purely spatial models that lack a Euclidean component are no better than a simple average watershed score.

Many of the models had a large tail-down component and a small or negligible tail-up component, although the reverse was true in the Feather River and the Santa Margarita (where, unlike in other watersheds, the majority of observed CSCI scores were in good condition; Table 1). This pattern is consistent with the way bioassessment scores are often informally interpreted to represent streams: high scores indicate that upstream conditions are also good, and poor scores indicate that downstream conditions are also poor. Although not evident in the data sets in the study, downstream recovery from impacts is possible (e.g., Rehn 2008). Therefore, the models may be reflecting the specific pattern of disturbance in a watershed and the representation of this pattern in sampled data, rather than the influence of intrinsic watershed properties on spatial autocorrelations of bioassessment scores, or properties of stream networks in general. The results of a model made for one watershed may not generalize to others, even if they have very similar catchment properties.

Applications to decision-making and monitoring design

Spatial models are an effective way to extrapolate scores to unsampled reaches. However, they do not directly provide general guidance on extrapolating scores that can be applied broadly. More specifically, the ranges of the tail-up or tail-down components should not be interpreted as limits on extrapolating scores. The SSN models in this study were strictly spatial, and (with the exception of Alameda Creek) did not account for covariates, such as land use, stream temperature, or perennial flow status. Including appropriate covariates in SSN models (such as temperature, habitat, or more meaningful measurements of contributing land cover) could decrease the amount of detectable spatial autocorrelation (D. Isaaks, personal communication). But even then, it may not be appropriate to interpret the ranges of these as limits for extrapolating bioassessment scores to unsampled reaches, because large ranges may not reflect the small (yet detectable) autocorrelation between sites separated by large distances. However, it may be possible to derive a general limit by estimating the maximum distance from a sampled location that a prediction point has an acceptably small SE.

Maps based on SSNs offer an alternative way for evaluating limits, and have the advantage of illustrating these limits in a site-specific and spatially explicit manner. Managers seeking to evaluate stream condition can identify with known confidence intervals where reaches are likely above or below thresholds of interest. For example, maps of the Malibu Creek watershed indicated that a large extent of this watershed is largely in poor condition, including areas far from sampled sites (e.g., Potrero Valley); however, confidence in the condition estimates for these remote tributaries is low. Moreover, certain portions of the watershed estimated with high confidence to be in good condition (e.g., Cold Creek). Maps like the ones shown in Figure 8D can support discussions among regulators and stakeholders when deciding how to designate an impaired waterbody, or identifying regions where confirmatory sampling is needed to support a designation. Additionally, they can easily be modified to reflect thresholds and prediction intervals (based on desired levels of confidence) that are appropriate to the questions at hand. This spatially explicit approach is an alternative to the traditional practice of watershed-wide classifications of good or bad condition.

A drawback of a map-based approach is that models require calibration for each relevant application, and they must be recalibrated to accommodate new data. This approach is considerably more complex than a simple limit that could be applied universally to all streams in California. But this complexity affords more flexibility, like the ability to incorporate new data, or to provide limits that are specific to individual reaches. The toolkits required to create the models are not simple (Peterson 2015, Ver Hoef et al. 2014), but routine application is feasible if the process is streamlined and automated, and training is provided to agency staff.

A watershed-by-watershed approach will work well in regions with ongoing bioassessment surveys (e.g., the Stormwater Monitoring Coalition in Southern California, or the Regional Monitoring Program in the Bay Area), and major watersheds in these regions are well suited for spatial modeling. Other parts of California may face greater challenges because few watersheds have the sampling density to make effective models (e.g., 30 sites per watershed in the present study). In these situations, it may be more efficient to develop models covering multiple watersheds in larger regions. Additionally, incorporating landcover (or similar) covariates into SSN models will likely yield more meaningful estimates of condition in watersheds with few to no existing samples.

Currently, watershed managers are pressed to extrapolate results from limited sampling locations to make decisions about larger reaches, or about entire catchments. Maps based on SSN models offer a way to support these inferences in a transparent and objective manner. Although maps and statistical models cannot substitute professional judgment or preempt local expertise, they provide a good foundation for stakeholders and regulators to evaluate available bioassessment data in a spatially explicit manner.

LITERATURE CITED

Carlisle, D., J. Falcone, and M. Meador. 2009. Predicting the biological condition of streams: use of geospatial indicators of natural and anthropogenic characteristics of watersheds. Environmental Monitoring and Assessment 151: 143-160.

Cressie, N. 1993. Statistics for spatial data. Revised Edition. New York. John Wiley and Sons.

Falcone, J.A., D.M. Carlisle, and L.C. Weber. 2010. Quantifying human disturbance in watersheds: Variable selection and performance of a GIS-based disturbance index for predicting the biological condition of perennial streams. Ecological Indicators 10: 264-273.

Frieden, J.C., E.E. Peterson, J.A. Webb, and P.M. Negus. 2014. Improving the predictive power of spatial statistical models of stream macroinvertebrates using weighted autocovariance functions. Environmental Modelling and Software 60: 320-330

Hill, R.A., E.W. Fox, S.G. Leibowitz, A.R. Olsen, D.J. Thornbrugh, and M.H. Weber. In review. Predictive mapping of the biotic condition of the conterminous-USA rivers and streams.

Isaak, D.J., E.E. Peterson, D.E. Nagel, J.M. Ver Hoef, and J. Kershner. 2013. A national stream internet to facilitate accurate, high-resolution status and trend assessments for water quality parameters and aquatic biotas. U.S. Fish and Wildlife Service, National Landscape Conservation Cooperative grant. Project website: www.fs.fed.us/rm/boise/AWAE/projects/NationalStreamInternet.html.

Isaak, D.J., J.M. Ver Hoef, E.E. Peterson, D.L. Horan, and D.E. Nagel. 2017. Scalable population estimates using spatial-stream-network (SSN) models, fish density surveys, and national geospatial database frameworks for streams. Canadian Journal of Fisheries and Aquatic Sciences 74: doi: 10.1139/cjfas-2016-0247.

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. Environmental Monitoring and Assessment 187: 4086. doi:10.1007/s10661-014-4086-x.

Mazor, R.D., 2015. Bioassessment of Perennial Streams in Southern California: A Report of the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey. Southern California Coastal Water Research Project Technical Report 844, May 2015. Available from http://ftp.sccwrp.org/pub/download/DOCUMENTS/TechnicalReports/844_SoCalStrmAssess.pdf

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): 249-271

Ode, P.R. 2007. Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Surface Water Ambient Monitoring Program. Sacramento, CA.

Ode, P.R., A.C. Rehn, R.D. Mazor, K.C. Schiff, E.D. Stein, J.T. May, L.R. Brown, D.B. Herbst, D. Gillett, K. Lunde, and C.P. Hawkins. 2016. Evaluating the adequacy of a reference-site pool for ecological assessment in environmentally complex regions. Freshwater Science 35: 237-248.

Peterson, E.E. and J.M. Ver Hoef. 2014. STARS: An ArcGIS toolset used to calculate the spatial information need to fit spatial statistical models to stream network data. Journal of Statistical Software 56(2).

Peterson, E.E. 2015. STARS: Spatial analysis tool for the analysis of river systems version 2.0.4. A tutorial. Queensland University of Technology. Brisbane, QLD, Australia. Available from: http://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/downloads/STARS/STARS_tutorial_v204.pdf

Rehn, A.C. 2008. Benthic macroinvertebrates as indicators of biological condition below hydropower dams on west slope Sierra Nevada streams, California USA. River Research and Applications 25: 208-228.

Steel, E.A., C. Sowder, and E.E. Peterson. 2016. Spatial and temporal variation of water temperature regimes on the Snoqualmie River network. Journal of the Water Resources Association. 52: 769-787.

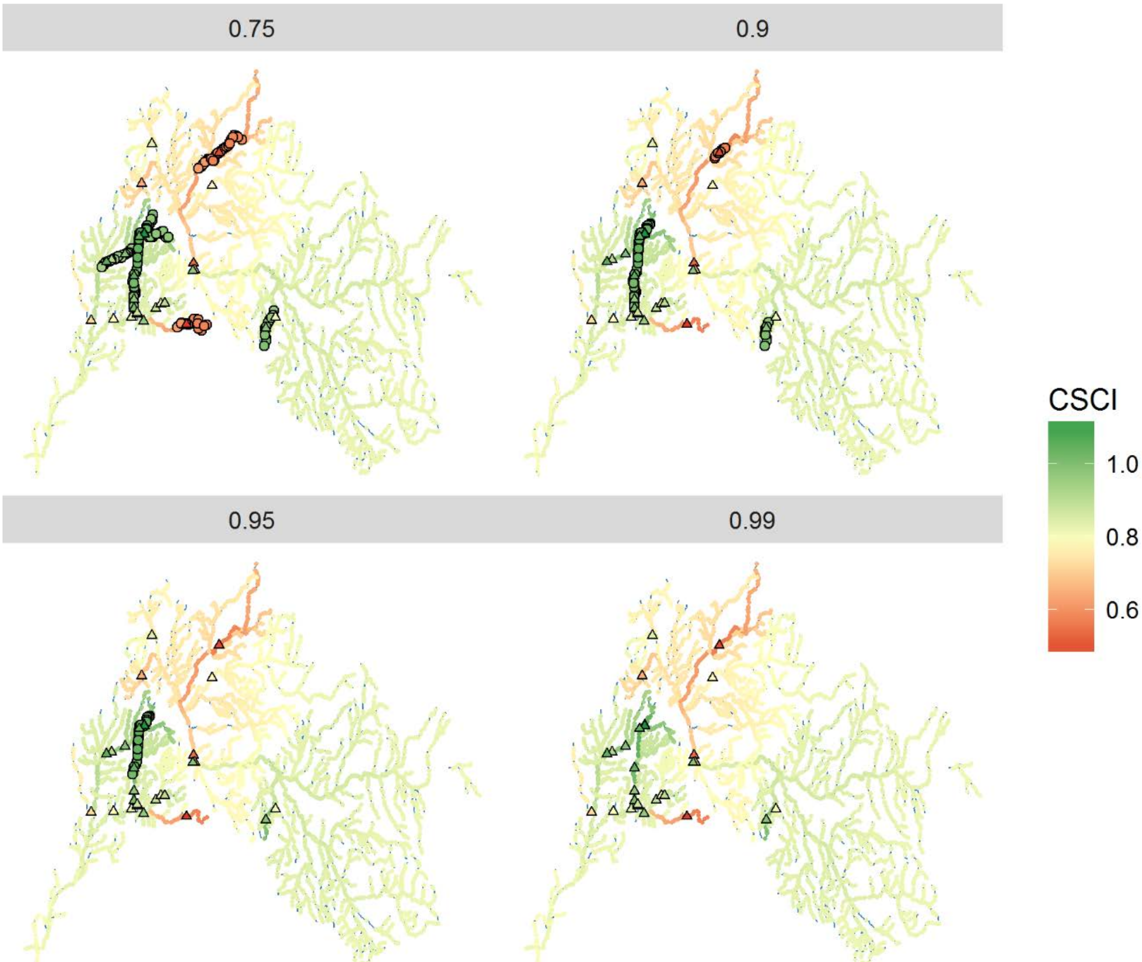
Ver Hoef, J.M., and E.E. Peterson. 2010. A moving average approach for spatial statistical models (with discussion). Journal of the American Statistical Association 105: 6-18.

Ver Hoef, J.M., E.E. Peterson, D. Clifford, and R. Shah. 2014. SSN: An R package for spatial statistical modeling on stream networks. Journal of Statistical Software 56: 1-43.

Ver Hoef, J.M., E.E. Peterson, and D. Theobald. 2016. Spatial statistical models that use flow and stream distance. Environmental and Ecological Statistics 13: 449-464.

Verbeke G., and G. Molenberghs. 2000. Linear Mixed Models for Longitudinal Data. Springer-Verlag Inc. ISBN 0-387-98222-1.

APPENDIX



Appendix Figure 1: Model performance and selection criteria for each watershed. TU: Tail-up. TD: Tail-down. EU: Euclidean. neg2LogL: -2 log-likelihood; low values are best. AIC: Aikake Information Criterion; low values are best. bias: mean of differences between observed and predicted values from leave-one-out-cross-validation (LOOCV); values close to zero are best. std.bias: standardized bias, calculated as the mean difference between observed and predicted values, divided by the prediction standard error from LOOCV; values close to zero are best. RMSPE: Root-mean-squared prediction error; low values are best. RAV: root-average variance, calculated as the square root of the mean of squared prediction errors; values close to RSMPE are best. std.MSPE: Standardized mean-squared prediction error; values close to 1 are best. cov.80: Number of times the predicted value was in the 80th percentile prediction interval; values close to 0.8 are best. cov.90: Number of times the predicted value was in the 90th percentile prediction interval; values close to 0.9 are best. cov.95: Number of times the predicted value was in the 95th percentile prediction interval; values close to 0.95 are best. Exp: Exponential autocorrelation model. Sph: Spherical autorelation model. Epa: Epanechnikov autocorrelation model. Lin: Linear-with-sill correlation model. Mar: Mariah correlation model. Asterisks (*) indicates selected models.

Appendix Table 1A. Model performance and selection criteria for the Feather River.

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
3	Exp	Exp	Exp	-33.6	-17.6	-0.0038	-0.0068	0.178	0.182	0.973	0.84	0.88	0.96
3	Sph	Sph	Sph	-34.0	-18.0	-0.0058	-0.0105	0.174	0.181	0.957	0.85	0.88	0.94
2	Epa	Epa	None	-34.0	-22.0	-0.0058	-0.0105	0.174	0.182	0.953	0.84	0.88	0.94
2	Epa	Exp	None	-34.0	-22.0	-0.0053	-0.0094	0.175	0.181	0.974	0.84	0.88	0.93
2	Epa	Lin	None	-34.1	-22.1	-0.0059	-0.0106	0.173	0.182	0.950	0.85	0.88	0.94
2	Epa	Mar	None	-34.1	-22.1	-0.0056	-0.0101	0.174	0.181	0.962	0.84	0.88	0.94
2	Epa	Sph	None	-34.1	-22.1	-0.0056	-0.0101	0.174	0.181	0.959	0.85	0.88	0.94
2	Exp	Epa	None	-33.8	-21.8	-0.0056	-0.0101	0.174	0.182	0.960	0.84	0.88	0.94
2	Exp	Exp	None	-34.1	-22.1	-0.0052	-0.0094	0.175	0.181	0.964	0.85	0.88	0.96
2	Exp	Lin	None	-33.8	-21.8	-0.0056	-0.0101	0.174	0.181	0.965	0.84	0.88	0.94
2	Exp	Mar	None	-33.7	-21.7	-0.0054	-0.0096	0.175	0.185	0.939	0.84	0.88	0.96
2	Exp	Sph	None	-33.8	-21.8	-0.0056	-0.0101	0.174	0.181	0.961	0.84	0.88	0.94
2	Lin	Epa	None	-34.0	-22.0	-0.0056	-0.0101	0.174	0.181	0.960	0.85	0.88	0.94
2	Lin	Exp	None	-34.1	-22.1	-0.0056	-0.0103	0.174	0.180	0.955	0.85	0.88	0.94
2	Lin	Lin	None	-34.1	-22.1	-0.0059	-0.0106	0.173	0.181	0.953	0.85	0.88	0.94
2	Lin	Mar	None	-34.1	-22.1	-0.0056	-0.0101	0.174	0.181	0.965	0.84	0.88	0.94
2	Lin	Sph	None	-34.0	-22.0	-0.0057	-0.0103	0.174	0.181	0.964	0.85	0.88	0.94
2	Mar	Epa	None	-30.3	-18.3	-0.0037	-0.0066	0.185	0.188	0.985	0.84	0.90	0.96
2	Mar	Exp	None	-29.7	-17.7	-0.0032	-0.0056	0.187	0.190	0.988	0.82	0.90	0.96
2	Mar	Lin	None	-30.3	-18.3	-0.0039	-0.0069	0.184	0.187	0.986	0.82	0.90	0.94
2	Mar	Mar	None	-30.4	-18.4	-0.0040	-0.0072	0.184	0.187	0.992	0.84	0.88	0.94
2	Mar	Sph	None	-30.4	-18.4	-0.0040	-0.0071	0.184	0.187	0.990	0.84	0.88	0.94
2	Sph	Epa	None	-34.0	-22.0	-0.0058	-0.0104	0.174	0.181	0.956	0.85	0.88	0.94
2	Sph	Exp	None	-34.0	-22.0	-0.0056	-0.0101	0.174	0.182	0.956	0.85	0.88	0.94

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
*2	Sph	Lin	None	-34.4	-22.4	-0.0053	-0.0094	0.174	0.187	0.910	0.85	0.91	0.96
2	Sph	Mar	None	-34.1	-22.1	-0.0059	-0.0106	0.174	0.181	0.962	0.85	0.88	0.94
2	Sph	Sph	None	-34.0	-22.0	-0.0058	-0.0104	0.174	0.181	0.955	0.85	0.88	0.94
0	None	None	None	-9.8	-5.8	0.0000	0.0000	0.228	0.227	1.015	0.85	0.90	0.93

Appendix Table 1B. Model performance and selection criteria for Alameda Creek.

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
3	Exp	Exp	Exp	-27.0	-11.0	0.0053	0.0088	0.142	0.139	1.017	0.75	0.94	1.00
3	Sph	Sph	Sph	-28.9	-12.9	0.0048	0.0081	0.136	0.137	1.001	0.72	0.94	1.00
2	Epa	Epa	None	-26.5	-14.5	0.0032	0.0043	0.143	0.142	1.016	0.78	0.94	1.00
2	Epa	Exp	None	-25.2	-13.2	-0.0003	-0.0022	0.146	0.144	1.024	0.78	0.91	1.00
2	Epa	Lin	None	-25.6	-13.6	0.0016	0.0017	0.140	0.144	0.992	0.78	0.94	1.00
2	Epa	Mar	None	-20.8	-8.8	-0.0020	-0.0038	0.149	0.156	0.953	0.81	0.94	0.97
2	Epa	Sph	None	-25.6	-13.6	0.0021	0.0024	0.145	0.145	1.019	0.81	0.94	1.00
2	Exp	Epa	None	-26.5	-14.5	0.0032	0.0043	0.143	0.142	1.016	0.78	0.94	1.00
2	Exp	Exp	None	-25.2	-13.2	-0.0003	-0.0022	0.146	0.144	1.024	0.78	0.91	1.00
2	Exp	Lin	None	-25.6	-13.6	0.0016	0.0017	0.140	0.144	0.992	0.78	0.94	1.00
2	Exp	Mar	None	-21.8	-9.8	-0.0033	-0.0066	0.147	0.154	0.965	0.81	0.94	0.97
2	Exp	Sph	None	-25.6	-13.6	0.0021	0.0024	0.145	0.145	1.019	0.81	0.94	1.00
2	Lin	Epa	None	-26.5	-14.5	0.0032	0.0043	0.143	0.142	1.016	0.78	0.94	1.00
2	Lin	Exp	None	-25.2	-13.2	-0.0003	-0.0022	0.146	0.144	1.024	0.78	0.91	1.00
2	Lin	Lin	None	-25.6	-13.6	0.0016	0.0017	0.140	0.144	0.992	0.78	0.94	1.00
2	Lin	Mar	None	-20.8	-8.8	-0.0020	-0.0038	0.149	0.156	0.953	0.81	0.94	0.97
2	Lin	Sph	None	-25.6	-13.6	0.0021	0.0024	0.145	0.145	1.019	0.81	0.94	1.00
2	Mar	Epa	None	-26.5	-14.5	0.0032	0.0043	0.143	0.142	1.016	0.78	0.94	1.00
2	Mar	Exp	None	-25.2	-13.2	-0.0003	-0.0022	0.146	0.144	1.025	0.78	0.91	1.00
2	Mar	Lin	None	-25.6	-13.6	0.0016	0.0017	0.140	0.144	0.992	0.78	0.94	1.00
2	Mar	Mar	None	-21.8	-9.8	-0.0033	-0.0066	0.147	0.154	0.966	0.81	0.94	0.97
2	Mar	Sph	None	-25.6	-13.6	0.0021	0.0024	0.145	0.145	1.019	0.81	0.94	1.00
2	Sph	Epa	None	-26.5	-14.5	0.0032	0.0043	0.143	0.142	1.016	0.78	0.94	1.00
*2	Sph	Exp	None	-25.2	-13.2	-0.0003	-0.0022	0.146	0.144	1.024	0.78	0.91	1.00
2	Sph	Lin	None	-25.6	-13.6	0.0016	0.0017	0.140	0.144	0.992	0.78	0.94	1.00

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
2	Sph	Mar	None	-21.8	-9.8	-0.0033	-0.0066	0.147	0.154	0.965	0.81	0.94	0.97
2	Sph	Sph	None	-25.6	-13.6	0.0021	0.0024	0.145	0.145	1.019	0.81	0.94	1.00
0	None	None	None	11.9	15.9	0.0000	0.0000	0.301	0.296	1.032	0.78	0.97	0.97

Appendix Table 1C. Model performance and selection criteria for Coyote Creek.

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
3	Exp	Exp	Exp	-36.9	-20.9	0.0005	0.0007	0.178	0.180	0.999	0.82	0.93	0.96
3	Sph	Sph	Sph	-39.9	-23.9	0.0003	0.0004	0.177	0.178	1.002	0.81	0.93	0.96
2	Epa	Epa	None	-18.4	-6.4	0.0057	0.0093	0.198	0.194	1.016	0.80	0.91	0.96
2	Epa	Exp	None	-14.6	-2.6	0.0052	0.0083	0.201	0.200	1.003	0.82	0.90	0.97
2	Epa	Lin	None	-18.2	-6.2	0.0063	0.0106	0.199	0.194	1.019	0.80	0.92	0.96
2	Epa	Mar	None	-5.6	6.4	0.0015	0.0020	0.228	0.218	1.038	0.77	0.91	0.96
2	Epa	Sph	None	-17.7	-5.7	0.0057	0.0092	0.199	0.196	1.013	0.82	0.91	0.95
*2	Exp	Epa	None	-19.6	-7.6	0.0059	0.0097	0.198	0.193	1.017	0.81	0.91	0.96
2	Exp	Exp	None	-14.6	-2.6	0.0052	0.0083	0.201	0.200	1.003	0.82	0.90	0.97
2	Exp	Lin	None	-18.1	-6.1	0.0065	0.0109	0.199	0.194	1.022	0.81	0.92	0.95
2	Exp	Mar	None	-3.2	8.8	0.0015	0.0020	0.230	0.221	1.031	0.78	0.89	0.97
2	Exp	Sph	None	-18.7	-6.7	0.0058	0.0094	0.199	0.195	1.011	0.81	0.92	0.96
2	Lin	Epa	None	-18.5	-6.5	0.0057	0.0093	0.198	0.194	1.016	0.79	0.91	0.96
2	Lin	Exp	None	-14.6	-2.6	0.0052	0.0083	0.201	0.200	1.003	0.82	0.90	0.97
2	Lin	Lin	None	-18.2	-6.2	0.0063	0.0105	0.199	0.194	1.018	0.80	0.92	0.96
2	Lin	Mar	None	-5.8	6.2	0.0014	0.0019	0.227	0.217	1.037	0.77	0.90	0.96
2	Lin	Sph	None	-17.7	-5.7	0.0057	0.0092	0.199	0.196	1.013	0.82	0.91	0.95
2	Mar	Epa	None	-19.0	-7.0	0.0059	0.0096	0.198	0.194	1.018	0.81	0.92	0.96
2	Mar	Exp	None	-14.8	-2.8	0.0052	0.0083	0.202	0.200	1.003	0.81	0.90	0.97
2	Mar	Lin	None	-18.7	-6.7	0.0066	0.0111	0.199	0.194	1.026	0.81	0.92	0.95
2	Mar	Mar	None	-8.2	3.8	0.0037	0.0059	0.208	0.206	0.999	0.80	0.90	0.97
2	Mar	Sph	None	-18.1	-6.1	0.0057	0.0093	0.199	0.196	1.014	0.79	0.91	0.96
2	Sph	Epa	None	-18.5	-6.5	0.0058	0.0094	0.198	0.194	1.016	0.82	0.92	0.96
2	Sph	Exp	None	-14.7	-2.7	0.0050	0.0080	0.202	0.202	0.979	0.82	0.91	0.97
2	Sph	Lin	None	-18.2	-6.2	0.0064	0.0106	0.199	0.194	1.020	0.80	0.92	0.96

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
2	Sph	Mar	None	-5.1	6.9	0.0015	0.0020	0.228	0.219	1.034	0.77	0.90	0.97
2	Sph	Sph	None	-17.7	-5.7	0.0057	0.0092	0.199	0.196	1.013	0.82	0.91	0.95
0	None	None	None	29.8	33.8	0.0000	0.0000	0.283	0.282	1.010	0.79	0.90	0.97

Appendix Table 1D. Model performance and selection criteria for Malibu Creek.

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
3	Exp	Exp	Exp	-76.7	-60.7	-0.0020	-0.0046	0.107	0.108	0.995	0.80	0.95	0.95
3	Sph	Sph	Sph	-77.8	-61.8	-0.0021	-0.0050	0.105	0.109	0.974	0.80	0.95	0.95
2	Epa	Epa	None	-78.4	-66.4	-0.0024	-0.0054	0.105	0.107	0.991	0.86	0.93	0.95
2	Epa	Exp	None	-76.8	-64.8	-0.0023	-0.0052	0.106	0.108	0.988	0.80	0.95	0.95
2	Epa	Lin	None	-79.2	-67.2	-0.0017	-0.0039	0.103	0.106	0.979	0.82	0.95	0.95
2	Epa	Mar	None	-74.1	-62.1	-0.0041	-0.0092	0.106	0.110	0.983	0.84	0.95	0.95
2	Epa	Sph	None	-77.7	-65.7	-0.0021	-0.0049	0.106	0.108	0.985	0.80	0.93	0.95
2	Exp	Epa	None	-78.2	-66.2	-0.0024	-0.0054	0.105	0.106	1.013	0.84	0.91	0.95
2	Exp	Exp	None	-76.7	-64.7	-0.0022	-0.0051	0.106	0.109	0.985	0.80	0.95	0.95
*2	Exp	Lin	None	-79.1	-67.1	-0.0016	-0.0038	0.103	0.106	0.983	0.82	0.95	0.95
2	Exp	Mar	None	-74.0	-62.0	-0.0038	-0.0084	0.106	0.110	0.985	0.84	0.95	0.95
2	Exp	Sph	None	-77.7	-65.7	-0.0023	-0.0052	0.105	0.108	0.984	0.80	0.95	0.95
2	Lin	Epa	None	-78.1	-66.1	-0.0017	-0.0041	0.106	0.107	0.990	0.82	0.93	0.95
2	Lin	Exp	None	-76.8	-64.8	-0.0023	-0.0054	0.106	0.108	0.984	0.82	0.95	0.95
2	Lin	Lin	None	-78.5	-66.5	-0.0018	-0.0042	0.105	0.107	0.989	0.82	0.93	0.95
2	Lin	Mar	None	-73.9	-61.9	-0.0038	-0.0084	0.106	0.110	0.986	0.84	0.95	0.95
2	Lin	Sph	None	-77.7	-65.7	-0.0022	-0.0051	0.105	0.108	0.989	0.80	0.93	0.95
2	Mar	Epa	None	-78.7	-66.7	-0.0022	-0.0053	0.106	0.107	1.010	0.84	0.93	0.95
2	Mar	Exp	None	-77.3	-65.3	-0.0028	-0.0067	0.107	0.108	1.004	0.79	0.95	0.96
2	Mar	Lin	None	-79.1	-67.1	-0.0016	-0.0040	0.105	0.105	1.019	0.77	0.95	0.96
2	Mar	Mar	None	-73.2	-61.2	-0.0039	-0.0090	0.109	0.111	1.010	0.82	0.93	0.96
2	Mar	Sph	None	-78.2	-66.2	-0.0027	-0.0064	0.106	0.107	1.005	0.79	0.95	0.96
2	Sph	Epa	None	-78.2	-66.2	-0.0021	-0.0050	0.105	0.105	1.040	0.84	0.91	0.95
2	Sph	Exp	None	-76.7	-64.7	-0.0023	-0.0053	0.106	0.108	0.987	0.80	0.95	0.95
2	Sph	Lin	None	-78.6	-66.6	-0.0017	-0.0040	0.105	0.106	0.999	0.82	0.93	0.95

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
2	Sph	Mar	None	-74.0	-62.0	-0.0038	-0.0085	0.106	0.110	0.986	0.84	0.95	0.95
2	Sph	Sph	None	-77.6	-65.6	-0.0020	-0.0047	0.106	0.109	0.966	0.80	0.95	0.95
0	None	None	None	-39.6	-35.6	0.0000	0.0000	0.173	0.171	1.018	0.89	0.93	0.95

Appendix Table 1E. Model performance and selection criteria for the Los Angeles River.

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
3	Exp	Exp	Exp	-51.2	-35.2	-0.0028	-0.0051	0.161	0.166	0.972	0.84	0.91	0.97
3	Sph	Sph	Sph	-52.7	-36.7	-0.0022	-0.0041	0.161	0.165	0.978	0.84	0.91	0.97
*2	Epa	Epa	None	-48.0	-36.0	-0.0060	-0.0100	0.171	0.173	0.983	0.80	0.91	0.97
2	Epa	Exp	None	-44.8	-32.8	-0.0069	-0.0118	0.174	0.176	0.979	0.82	0.90	0.97
2	Epa	Lin	None	-47.8	-35.8	-0.0054	-0.0091	0.172	0.172	0.991	0.82	0.90	0.96
2	Epa	Mar	None	-39.7	-27.7	-0.0070	-0.0121	0.177	0.186	0.936	0.81	0.91	0.98
2	Epa	Sph	None	-47.0	-35.0	-0.0065	-0.0109	0.171	0.174	0.978	0.81	0.91	0.97
2	Exp	Epa	None	-47.9	-35.9	-0.0060	-0.0100	0.171	0.172	0.986	0.80	0.91	0.97
2	Exp	Exp	None	-44.7	-32.7	-0.0070	-0.0119	0.174	0.177	0.975	0.82	0.90	0.97
2	Exp	Lin	None	-47.7	-35.7	-0.0054	-0.0091	0.172	0.172	0.995	0.81	0.90	0.96
2	Exp	Mar	None	-39.1	-27.1	-0.0080	-0.0140	0.178	0.184	0.960	0.82	0.91	0.98
2	Exp	Sph	None	-46.8	-34.8	-0.0065	-0.0110	0.171	0.174	0.980	0.81	0.91	0.97
2	Lin	Epa	None	-47.9	-35.9	-0.0060	-0.0100	0.171	0.173	0.983	0.81	0.91	0.97
2	Lin	Exp	None	-44.8	-32.8	-0.0070	-0.0119	0.174	0.176	0.980	0.82	0.90	0.97
2	Lin	Lin	None	-47.8	-35.8	-0.0054	-0.0091	0.172	0.172	0.991	0.81	0.90	0.96
2	Lin	Mar	None	-39.1	-27.1	-0.0065	-0.0114	0.178	0.180	0.984	0.80	0.90	0.98
2	Lin	Sph	None	-46.8	-34.8	-0.0064	-0.0109	0.171	0.174	0.979	0.84	0.91	0.97
2	Mar	Epa	None	-46.0	-34.0	-0.0069	-0.0117	0.172	0.173	0.995	0.85	0.92	0.96
2	Mar	Exp	None	-42.7	-30.7	-0.0081	-0.0139	0.176	0.177	0.988	0.84	0.91	0.97
2	Mar	Lin	None	-46.1	-34.1	-0.0058	-0.0101	0.173	0.169	1.025	0.82	0.90	0.95
2	Mar	Mar	None	-36.8	-24.8	-0.0092	-0.0160	0.180	0.186	0.934	0.82	0.92	0.97
2	Mar	Sph	None	-44.6	-32.6	-0.0073	-0.0124	0.173	0.176	0.965	0.85	0.92	0.96
2	Sph	Epa	None	-47.9	-35.9	-0.0060	-0.0100	0.171	0.173	0.983	0.81	0.91	0.97
2	Sph	Exp	None	-44.8	-32.8	-0.0070	-0.0119	0.174	0.176	0.980	0.82	0.90	0.97
2	Sph	Lin	None	-47.7	-35.7	-0.0054	-0.0091	0.172	0.172	0.990	0.81	0.90	0.96

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
2	Sph	Mar	None	-38.6	-26.6	-0.0064	-0.0109	0.178	0.189	0.873	0.84	0.95	0.98
2	Sph	Sph	None	-46.9	-34.9	-0.0065	-0.0110	0.171	0.174	0.978	0.81	0.91	0.97
0	None	None	None	21.6	25.6	0.0000	0.0000	0.275	0.274	1.011	0.77	0.91	0.97

Appendix Table 1F. Model performance and selection criteria for the Santa Margarita River.

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
3	Exp	Exp	Exp	-32.1	-16.1	0.0141	0.0271	0.140	0.140	1.005	0.81	0.84	0.91
3	Sph	Sph	Sph	-32.3	-16.3	0.0155	0.0296	0.142	0.141	0.993	0.81	0.88	0.88
*2	Epa	Epa	None	-32.7	-20.7	0.0155	0.0300	0.139	0.138	1.008	0.81	0.84	0.88
2	Epa	Exp	None	-32.5	-20.5	0.0149	0.0287	0.139	0.140	0.979	0.81	0.84	0.91
2	Epa	Lin	None	-32.7	-20.7	0.0146	0.0281	0.140	0.140	0.987	0.81	0.88	0.88
2	Epa	Mar	None	-32.5	-20.5	0.0138	0.0264	0.138	0.139	0.990	0.81	0.84	0.94
2	Epa	Sph	None	-32.7	-20.7	0.0150	0.0290	0.140	0.138	1.008	0.81	0.88	0.88
2	Exp	Epa	None	-31.9	-19.9	0.0134	0.0261	0.143	0.137	1.013	0.81	0.84	0.91
2	Exp	Exp	None	-31.9	-19.9	0.0153	0.0293	0.141	0.141	1.000	0.81	0.88	0.88
2	Exp	Lin	None	-31.9	-19.9	0.0153	0.0291	0.143	0.143	0.960	0.84	0.84	0.94
2	Exp	Mar	None	-32.2	-20.2	0.0129	0.0244	0.139	0.139	0.997	0.81	0.84	0.91
2	Exp	Sph	None	-32.2	-20.2	0.0147	0.0283	0.141	0.140	0.997	0.81	0.88	0.88
2	Lin	Epa	None	-32.7	-20.7	0.0156	0.0303	0.140	0.137	1.024	0.81	0.84	0.88
2	Lin	Exp	None	-32.6	-20.6	0.0150	0.0291	0.139	0.139	1.004	0.81	0.84	0.88
2	Lin	Lin	None	-32.7	-20.7	0.0151	0.0294	0.139	0.136	1.039	0.81	0.84	0.88
2	Lin	Mar	None	-32.5	-20.5	0.0138	0.0264	0.138	0.139	0.992	0.81	0.84	0.94
2	Lin	Sph	None	-32.5	-20.5	0.0145	0.0282	0.139	0.139	1.006	0.81	0.84	0.94
2	Mar	Epa	None	-30.5	-18.5	0.0150	0.0285	0.146	0.140	1.029	0.84	0.88	0.91
2	Mar	Exp	None	-30.0	-18.0	0.0146	0.0276	0.147	0.142	1.022	0.84	0.84	0.91
2	Mar	Lin	None	-30.5	-18.5	0.0138	0.0261	0.145	0.141	1.004	0.84	0.84	0.91
2	Mar	Mar	None	-29.1	-17.1	0.0137	0.0259	0.147	0.140	1.064	0.84	0.84	0.91
2	Mar	Sph	None	-30.3	-18.3	0.0152	0.0287	0.147	0.140	1.037	0.84	0.88	0.91
2	Sph	Epa	None	-32.6	-20.6	0.0148	0.0288	0.140	0.138	1.013	0.81	0.88	0.88
2	Sph	Exp	None	-32.5	-20.5	0.0149	0.0288	0.140	0.139	1.006	0.81	0.84	0.88
2	Sph	Lin	None	-32.6	-20.6	0.0143	0.0278	0.140	0.138	1.011	0.81	0.88	0.88

Components	TU	TD	EU	neg2LogL	AIC	bias	std.bias	RMSPE	RAV	std.MSPE	cov.80	cov.90	cov.95
2	Sph	Mar	None	-32.4	-20.4	0.0142	0.0273	0.139	0.140	1.001	0.81	0.84	0.94
2	Sph	Sph	None	-32.4	-20.4	0.0142	0.0275	0.139	0.140	1.003	0.81	0.88	0.94
0	None	None	None	-20.0	-16.0	0.0000	0.0000	0.183	0.180	1.032	0.88	0.91	0.94