Ecological Indicators 119 (2020) 106421

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Predictive biological indices for algae populations in diverse stream environments

Susanna Theroux[a,*], Raphael D. Mazor[a], Marcus W. Beck[a,1], Peter R. Ode[b], Eric D. Stein[a], Martha Sutula[a]

[a] Southern California Coastal Water Research Project, 3535 Harbor Boulevard, Suite 110, Costa Mesa, CA 92626, USA
[b] Aquatic Bioassessment Laboratory, California Department of Fish and Wildlife, 2005 Nimbus Road, Rancho Cordova, CA 95670, USA

ARTICLE INFO

Keywords:
Bioassessment
Algae
Multimetric index
California
Wadeable streams

ABSTRACT

Predictive biological indices have transformed the bioassessment landscape by allowing universal indices to be applicable across diverse environments. The successful development of a predictive benthic macroinvertebrate index for California wadeable streams helped to demonstrate the power of these tools in complex geographic settings. However, previous efforts to develop predictive algal indices for California were limited by poor performance and were ultimately unsuccessful. For this study, we leveraged a robust statewide dataset to develop two different types of predictive algal indices for California wadeable streams: an index of observed-to-expected taxa (O/E) to measure taxonomic completeness and a multimetric index (MMI) to evaluate ecological structure. We developed multiple versions of each index, including one for diatoms, one for soft-bodied algae, and a hybrid index using both assemblages. We evaluated index performance using a series of screening criteria for precision, accuracy, responsiveness, and regional bias. We found that final index performance varied among all assemblages: the best performing O/E index was a diatom-only index, whereas the predictive diatom and hybrid MMIs out-performed all other indices with excellent responsiveness and precision. We found that in comparison to benthic macroinvertebrates, algal communities were characterized by high beta diversity across reference sites and low average species richness per site, resulting in disparate algal populations that were challenging to model with predictive approaches, particularly for soft-bodied algae assemblages. While all O/E indices were considered to have weak performance, the predictive diatom and hybrid MMIs are accurate, responsive, and precise indices that will provide a powerful assessment of biological condition for statewide applications.

1. Introduction

The advent of predictive biological indices has allowed for the expansion of these powerful bioassessment tools to diverse, complex environments (Hawkins et al., 2010b; Mazor et al., 2016; Vander Laan and Hawkins, 2014). Predictive indices evaluate the quality of a waterbody as the degree of alteration of the biological community in comparison to site-specific, reference-based expectations, also known as the reference condition approach (RCA; Reynoldson et al., 1997). The use of predictive indices for evaluating benthic macroinvertebrate (BMI) communities has seen broadscale application, including in the development of the River Invertebrate Prediction and Classification System (RIVPACS) that measures the taxonomic completeness of a sample as the ratio of observed-to-expected (O/E) taxa (Wright, 2000, 1995). The extension of predictive modeling techniques to multimetric indices (MMIs) has allowed for the accounting of geographic variability in the ecological structure of biological communities. Like predictive O/E indices, successful predictive MMIs have been shown to improve the accuracy, precision, and sensitivity of MMIs in diverse environmental settings (Cao et al., 2007; Hawkins et al., 2010a; Mazor et al., 2016; Vander Laan and Hawkins, 2014).

While predictive benthic macroinvertebrate indices have become increasingly popular, efforts to develop predictive algal indices have had varying success. Mazor et al. (2006) successfully developed a RIVPACS-type index for periphyton communities in the Fraser River, although the performance of the predictive algal index was worse than the benthic macroinvertebrate index. Cao et al. (2007) successfully developed both an O/E and a predictive MMI index for diatoms of Idaho streams. Feio et al., (2009) saw comparable performance between a predictive diatom index and non-predictive multimetric indices, and

* Corresponding author.
E-mail address: susannat@sccwrp.org (S. Theroux).
[1] Current address: Tampa Bay Estuary Program, 263 13th Avenue South, Suite 350, St. Petersburg, FL 33701, USA.

https://doi.org/10.1016/j.ecolind.2020.106421
Received 5 January 2020; Received in revised form 7 April 2020; Accepted 15 April 2020
Available online 16 September 2020
1470-160X/

Feio et al., (2012) successfully developed a diatom and macrophyte predictive model (AQUAFLORA) for Portuguese streams. Recently, a predictive diatom index was developed for streams and rivers in Northern Spain (Pardo et al., 2018) that demonstrated responsiveness to anthropogenic stressors. Additional attempts at a predictive diatom index have been less successful. A predictive, diatom-only index was attempted for the Central Coast of California but suffered from low precision and accuracy (Ritz, 2010), and a predictive diatom genus-level index for streams in Australia had weaker performance than a macroinvertebrate index, possibly due to temporal variability in diatom communities and limitations in predictor variables (Chessman et al., 1999).

California is home to a number of diverse ecoregions (Ode et al., 2016; Omernik, 1987) and these distinct environmental settings necessitate the development of biological indices that can adequately account for regional impacts on biological community structure. Fortunately, California also benefits from extensive annual sampling campaigns in wadeable stream environments (Ode et al., 2011), efforts that have resulted in rich taxonomic and ecological datasets for both algae (Fetscher et al., 2013, 2014a) and benthic macroinvertebrates (Mazor et al., 2016) that can be used to develop robust biological indices. Ode et al. (2016b) developed a definition of reference condition for California wadeable streams that incorporates a strict set of screening criteria to measure anthropogenic disturbance at a variety of spatial scales. Mazor et al. (2016) applied this reference definition to develop the California Stream Condition Index (CSCI), a predictive index for benthic macroinvertebrates that incorporates both a measure of taxonomic completeness (O/E) and a measure of ecological structure (MMI) for assessing biological condition in wadeable streams. The CSCI has been increasingly incorporated into statewide and regional monitoring programs and is now the primary way California reports on the biological condition of its streams to Congress as required by the US Clean Water Act (State Water Resources Control Board, 2017).

The success of the CSCI served as the motivation to develop a complementary line of evidence based on algal communities for assessing biological condition in California's wadeable streams. To address this need, we followed the approach used to develop the CSCI and constructed both a predictive algal O/E index and a predictive MMI. We developed each index using diatoms alone, soft-bodied algae alone, and a hybrid approach using both assemblages. We evaluated the performance of each index across both environmental and disturbance gradients, with a concerted focus on selecting final indices with high precision and low regional bias, thereby addressing a primary concern for using these indices in statewide water quality and management decisions. This paper reviews the performance of each resulting algal index and discusses both the biological and technical factors that contribute to index performance in diverse environmental settings.

2. Materials and methods

2.1. Study region

California is a species-rich biogeographic region with a complex history of geological change, modern-day geographic settings, and climate fluctuations (Calsbeek et al., 2003; Sork et al., 2016). California's North Coast is characterized by temperate rainforests, eastern portions of the state are dominated by desert, and the coastal regions have chaparral, oak woodlands, and grasslands with a Mediterranean climate (Omernik, 1987). Anthropogenic development has transformed the landscape in the agriculturally dominated Central Valley as well as the densely urban South Coast and San Francisco Bay Area (Sleeter et al., 2011). The state can be divided into six ecoregions based on modified ecoregional (Omernik, 1987) and hydrological boundaries (Ode et al., 2016), hereafter referred to as Perennial Stream Assessment (PSA) ecoregions (Mazor et al., 2016): the North Coast (NC), Central Valley (CV), Chaparral (CH), Desert Modoc (DM), Sierra Nevada (SN), and South Coast (SC).

2.2. Data compilation

We compiled algae taxonomic data from multiple federal, state, and regional monitoring programs in California, resulting in a dataset of 1941 sampling sites and 2586 unique sampling events in wadeable streams. All sampling events followed a standardized periphyton sampling protocol (Fetscher et al., 2009; Ode et al., 2016a). Briefly, a reach of 150 m was subdivided into eleven transects. A "multihabitat method" was employed to objectively collect subsamples of algal specimens quantitatively from a known surface area over a representative sample of stream substrata (Fetscher et al., 2009). An additional qualitative sample was collected by sampling visible mats of macroalgae (Fetscher et al., 2009) to help aid in the taxonomic identification of soft-bodied algae in the quantitative fraction. Algae samples were composited and proportioned into diatom and soft-bodied algae aliquots for laboratory analysis. Microscopy-based analyses for taxonomic identification followed Stancheva et al. (2015), and relied on statewide, harmonized master taxa lists for both diatoms and soft-bodied algae (Surface Water Ambient Monitoring Program, n.d.). For sites with multiple years of sampling, we selected the most recent sample for index development and reserved previous sampling events for performance analyses (e.g. within-site variability). The complete development dataset is provided at http://www.github.com/stheroux/asci.

2.3. Data curation

Samples with < 200 diatom valves were excluded from diatom and hybrid analyses but retained for SBA analysis. Only ~ 60% of samples contained a qualitative soft-bodied algal fraction so due to the inconsistent inclusion of this fraction, only quantitative taxonomy data was included in subsequent analyses. All taxonomy results were converted from count (diatoms) or biovolume (soft-bodied algae) data to presence/absence data in anticipation of future assemblage data being derived from DNA-based methods. Provisional species names and morphospecies names were subjected to a name-harmonization with AlgaeBase (algaebase.org) and Biodata species names (http://aquatic.biodata.usgs.gov) to remove ambiguous identifiers. All indices were tested using both species-level and genus-level taxonomy. Harmonized name lists are provided at http://www.github.com/stheroux/asci.

2.4. Classifying reference and high-activity sites

We assessed the influence of anthropogenic activity using measures of surrounding land use as well as local habitat data after Ode et al. (2016b). We followed a "least-disturbed" reference concept (Stoddard et al., 2006) for the identification of "reference" sites and identified high-activity sites as those that were presumptively stressed with high levels of human activity in the watershed or riparian zone that could potentially degrade in-stream biological condition (Mazor et al., 2016; Table 1). We used reference sites to determine the biological composition of sites with minimal human disturbance (Hawkins et al., 2010a; Mazor et al., 2016) and therefore to calibrate our O/E and MMIs models. Reference site screening thresholds closely followed those used in the construction of the CSCI (Mazor et al., 2016; Ode et al., 2016), with the exception that conductivity was not used to eliminate a site from the reference pool (Table 1). High-activity sites were used in scoring MMIs as well as in evaluating performance of both O/E and MMIs. Any sites that did not pass the reference or high-activity screening thresholds were included in the "intermediate" site pool. We further divided each dataset into a calibration (80%) and validation (20%) subset and stratified assignment by ecoregion for equal representation of different environmental settings.

Table 1
Stressor and human-activity gradients used to identify reference sites and high-activity sites. Sites that did not exceed the listed reference thresholds (ref threshold) were used as reference sites. Sites that exceeded at least one high-activity threshold (str threshold) were used as high-activity sites. Sources: A = National Landcover Data Set (http://www.epa.gov/mrlc/nlcd −2006.html), B = custom roads layer, C = National Hydrography Dataset Plus (http://www.horizon-systems.com/nhdplus), D = National Inventory of Dams (http://geo.usace.army.mil), E = Mineral Resource Data System (http://tin.er.usgs.gov/mrds); F = Field measured variables. WS = watershed; 5 km = watershed clipped to a 5-km buffer of the sampling point; 1 km = watershed clipped to a 1-km buffer of the sampling point; % Code 21 = land-use category that corresponds to highly managed vegetation, such as roadsides, lawns, cemeteries, and golf courses. W1_HALL = proximity-weighted human activity index (Kaufmann et al., 1999).

Variable	Scale	Ref threshold	Str threshold	Unit	Source
% agricultural	1 km, 5 km, WS	< 3	> 50	%	A
% urban	1 km, 5 km, WS	< 3	> 50	%	A
% agricultural + % urban	1 km, 5 km, WS	< 5		%	A
% Code 21	1 km, 5 km	< 7	> 50	%	A
	WS	< 10	> 50	%	A
Road density	1 km, 5 km, WS	< 2	> 5	km/km^2	B
Road crossings	1 km	< 5		crossings	B, C
	5 km	< 10		crossings	B, C
	WS	< 50		crossings	B, C
Dam distance	WS	> 10		km	D
% canals and pipelines	WS	< 10		%	C
Instream gravel mines	5 km	< 0.1		mines/km	C, E
Producer mines	5 km	0		mines	E
W1_HALL	Reach	< 1.5	> 5	NA	F

2.5. Environmental variables

We assembled environmental data from multiple sources, including GIS-derived variables from the National Hydrography Dataset Plus (NHD, http://www.horizon-systems.com/nhdplus), the National Landcover Data Set (http://www.epa.gov/mrlc/nlcd −2006.html), the National Inventory of Dams (https://nid.sec.usace.army.mil/), Mineral Resource Data System (http://tin.er.usgs.gov/mrds), and predicted specific conductance (background levels predicted for unimpacted sites) from Olson and Hawkins (Olson and Hawkins, 2012). This environmental data included measures of climate, elevation, geology, land cover, land use, road density, hydrologic alteration, and mining activities. We derived a measure of xeric/montane setting based on PSA ecoregion wherein the Chaparral, Central Valley, Desert Modoc, and the South Coast ecoregion were considered xeric and the North Coast and Sierra Nevada ecoregions were considered montane. We used environmental variables that characterize immutable natural gradients as candidate predictors for the O/E and MMI models (Table 2, Supplemental Table 1), whereas environmental variables influenced by anthropogenic factors were used for screening reference sites and for assessing index performance along stressor gradients (Supplemental Table 3).

2.6. O/E index development

2.6.1. O/E index construction

We constructed our O/E indices using the following steps: 1) cluster reference calibration sites based on their taxonomic composition; 2) develop a random forest model to predict site membership within a taxonomic cluster using candidate predictor variables that are minimally affected by human perturbation; 3) use the random forest model to predict cluster membership of test sites based on their environmental setting; 4) calculate the probability of observing a taxon at a test site, or capture probability (*Pc*), as the cluster-membership probability-weighted frequencies of occurrence summed across all clusters (Moss et al., 1987; Wright, 2000); 5) sum the capture probabilities as the expected number of taxa (E) in a sample from a site; 6) calculate the observed (O) and expected (E) number of taxa for each site.

We performed the clustering of reference site algae populations using a presence/absence transformed data matrix and excluded all taxa occurring in < 2.5% and > 95% of reference calibration sites (Hawkins et al., 2000; Van Sickle et al., 2007). The rare species excluded in the clustering step were retained for subsequent steps in the O/E development. For each of the O/E indices (diatom, soft-bodied algae, and hybrid), we determined an optimal number of clusters through a computational permutation with clusters ranging in size from 3 to 15 and a capture probability of either 0.4 or 0.5. We iterated the cluster numbers and capture probabilities with k-means clustering using the k-means function in the *stats* R package (R Core Team, 2013). We ensured that each cluster had at least five sites, as fewer sites per cluster has been shown to potentially omit taxa that are representative of reference conditions or under-represent stream types (Simpson and Norris, 2000; Wright et al., 1993). The cluster and capture probability combination of the best-performing O/E index (details below) were selected as the final parameters for that assemblage.

We used recursive feature elimination (RFE) as implemented in the *caret* package in R (Kuhn, 2008) to select environmental variables that were the best predictors of reference cluster group membership. In brief, we used RFE to identify the optimal number of environmental predictor variables whose model accuracy was within 1% of the best model. We then used the *randomForest* package (Liaw and Wiener, 2002) to construct a final 500-tree O/E model using the predictor variables identified by the RFE analysis. In addition to the statewide indices, we also generated an additional statewide O/E index for each algal assemblage that allowed for the inclusion of additional candidate predictor variables, including those influenced by human activity such as field-measured total nitrogen, total phosphorus, specific conductivity, and water temperature (Supplemental Table 1).

2.6.2. Evaluating O/E index performance

We evaluated O/E index performance by comparing the standard deviation (SD) of the predictive O/E index to the SD of the null O/E index, wherein all sites are in a single cluster and capture probabilities for each taxon are the same for all sites, to ensure that the predictive index had a lower SD than the null index. Additionally, we compared the predictive O/E index SD to the highest attainable precision possible based on estimates of the standard deviation among replicate samples (SDRS; (Van Sickle et al., 2005). We evaluated the model's ability to differentiate three condition classes (reference, intermediate, high-activity) using an ANOVA test as implemented in the R *stats* packages. We selected the best performing permutation of each algal O/E index by an averaged ranking of reference calibration site SD, predictive model performance improvement over the null model, and the ability of the model to differentiate reference from high-activity sites. All analyses were performed using modified scripts written by J. Van Sickle for evaluating O/E model performance (Van Sickle et al., 2005) in

Table 2
Selected candidate predictor variables for inclusion in both the O/E and MMI metric random forest models. When a predictor was selected, its importance is shown: Gini = mean decrease in Gini index, MSE = % increase in mean squared error for random forest models. Diatom metrics: m1 = cnt.spp.most.tol, m2 = EpiRho.richness, m3 = prop.spp.IndicatorClass_TN_low, m4 = prop.spp.Planktonic, m5 = prop.spp.Trophic.E, m6 = Salinity.BF.richness. Hybrid metrics: m1 = cnt.spp.IndicatorClass_TP_high, m2 = cnt.spp.most.tol, m3 = EpiRho.richness, m4 = OxyRed.DO_30.richness, m5 = prop.spp.Planktonic, m6 = prop.spp.Trophic.E, m7 = Salinity.BF.richness. Sources: A = PRISM climate database (http://www.prism.oregonstate.edu/); B = National Atmospheric Deposition Program National Trends Network (http://nadp.slh.wisc.edu/ntn/); C = Olson and Hawkins (Olson and Hawkins, 2012); D = National Elevation Dataset (http://ned.usgs.gov/). XerMtn: Xeric/Montane assignment derived from PSA ecoregion (see text for description).

		Diatom	SBA	Hybrid	Diatom MMI						Hybrid MMI							
Variable	Description	O/E	O/E	O/E	m1	m2	m3	m4	m5	m6	m1	m2	m3	m4	m5	m6	m7	Source
		Gini	*Gini*	*Gini*	*MSE*	*MSE*	*MSE*				*MSE*	*MSE*	*MSE*					
AREA_SQKM	Area	38.9	41.7	46.3	–	33.1	–	–	–	–	–	–	34.8	–	–	–	–	
AtmCa	Mean Ca + concentration	–	–	36.3	–	–	–	–	–	–	–	–	–	32.1	–	–	–	B
CondQR50	Median predicted conductivity	31.8	31.2	37.8	–	–	38.4	24.0	41.3	19.3	–	27.2	–	–	22.6	42.9	–	C
DayOfYear	Day of year	–	35.1	41.7	–	–	–	–	–	–	–	–	–	–	–	–	–	
KFCT_AVE	Catchment mean soil erodibility (K) factor	–	–	–	–	–	–	–	29.2	22.6	30.2	–	–	–	–	28.7	25.1	C
LPREM_mean	Catchment mean log geometric mean hydraulic conductivity	33.5	–	38.2	–	–	–	–	–	–	–	–	–	–	–	–	–	C
LST32AVE	Average of mean 1961–1990 first and last day of freeze	34.3	33.2	38.3	–	–	–	–	–	–	–	–	–	–	–	–	–	B
MAX_ELEV	Maximum elevation in catchment	39.9	33.3	37.8	–	–	30.7	–	–	–	–	–	–	–	–	–	–	D
Month	Month	–	12.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	
New_Lat	Latitude	32.6	35.7	36.1	–	–	–	–	–	–	–	–	–	–	–	–	–	
New_Long	Longitude	31.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	
PPT_00_09	10-y (2000–2009) average annual precipitation	–	–	–	25.4	–	–	–	–	–	45.6	–	–	26.0	–	–	–	A
SITE_ELEV	Site elevation	35.1	–	37.4	–	–	–	25.7	–	–	–	–	–	–	28.2	–	–	D
TEMP_00_09	10-y (2000–2009) average monthly temperature	29.9	33.3	33.8	–	–	–	–	–	–	–	–	–	–	–	–	–	A
TMAX_WS	Catchment mean of mean 1971–2000 maximum temperature	–	–	–	–	33.1	–	–	–	–	–	–	36.7	–	–	–	–	D
XerMtn	Xeric or Montane. PSA ecoregions CH, CV, DM, SC = = 0; NC, SN = = 1.	–	–	–	58.6	–	–	–	–	30.2	–	33.8	–	–	–	–	66.7	
XWD_WS	Catchment mean of mean 1961–1990 annual number of wet days	31.3	36.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	D

combination with custom scripts. All development scripts can be found at https://github.com/stheroux/asci.

As low numbers of expected taxa have been shown to impact O/E index performance (Hämäläinen et al., 2018; Mazor et al., 2016), we evaluated index sensitivity to low E values by calculating precision (standard deviation of reference calibration site scores for each index), sensitivity (proportion of high-activity sites not in reference condition), and accuracy (percent of reference calibration sites above 10th percentile of reference) of O/E indices for intervals ranging from 1 to 35 maximum E values. Additionally, we compared E values for diatom, soft-bodied algae, and hybrid O/E indices to predicted metric values for metrics that evaluate proportions of sensitive taxa (prop.spp.BCG12) and tolerant taxa (prop.spp.BCG4) at reference sites. This comparison allowed us to gauge how the expected proportion of sensitive/tolerant taxa varied along a range of E values and to test the hypothesis that greater proportions of sensitive and rare taxa would be predicted at sites with high E values. Additionally, we compared the E values for the hybrid O/E model with the combined sum of diatom and soft-algae E values for all sites. This comparison allowed us to evaluate if and when the hybrid O/E model resulted in higher E values than the combined E values for individual diatom and soft-algae O/E models.

2.7. Multi-metric index development

2.7.1. MMI construction

The MMIs were constructed in the following steps: 1) calculate observed (raw) metric values; 2) develop random forest models to predict metric values at reference sites and replace raw values with differences from predicted values, if appropriate; 3) score metrics; 4) select best-performing metrics and assemble proto-MMIs; 5) assemble final MMI with most frequent high-performing metrics.

2.7.2. Metric calculation

We assembled a custom dataset of species' trait attributes for calculating autoecological and phylogenic metrics. We obtained trait attributes from previously published algae attribute lists (Bahls, 1993; Porter et al., 2008; Potapova and Charles, 2007; Spaulding et al., 2010; van Dam et al., 1994), empirically-derived traits for southern California taxa (Fetscher et al., 2014b), recently developed traits related to species occurrence at reference vs. high-activity sites as derived by a panel of algal ecologists and taxonomists for the California Biological Condition Gradient (BCG; Paul et al., 2020). Additionally, we developed empirically-defined trait attributes by identifying "sensitive" and "tolerant" California taxa using an Indicator Species Analysis as implemented in the multipatt function in the *indicspecies* R package (Cáceres and Legendre, 2009) and classified sensitive taxa as enriched at reference sites and tolerant taxa enriched at high-activity sites. We generated metrics for species richness, Simpson and Shannon's diversity using the R package *vegan* (Oksanen et al., 2017).

We grouped metrics into thematic categories based on their autoecological, morphological, taxonomic, or species tolerance guilds (Table 3). As in the O/E model development, we performed all metric calculations on a presence/absence data matrix. For the hybrid MMI, we used a combined dataset of both diatom and soft-bodied algae taxa for calculating metric values. All metrics were calculated as both a proportion of total taxa and total count of taxa that met specific trait attributes and we assumed that changes in these metric values from reference expectations indicated degraded biological condition. We calculated all metrics using a combination of the R language vegan package (Oksanen et al., 2017) and custom R scripts (https://github.com/stheroux/asci).

Table 3
Metric classes for multimetric index. BCG taxa = taxa identified as indicative of a Biological Condition Gradient (BCG) Level indicators from the California BCG effort (Paul et al., 2020).

Category	Example metrics
Tolerance/sensitivity	Association with specific water-quality constituents (nitrogen, phosphorus, organic carbon, metals) Low/High oxygen preference Salinity tolerance
Autecological guild	Nitrogen-fixers Saprobic/heterotrophic taxa
Morphological guild	Sedimentation indicators
Relationship to reference	Taxa associated with reference vs. non-reference sites BCG taxa level
Taxonomic groups	Chlorophyta, Rhodophyta, Zygnemataceae, heterocystous cyanobacteria, Suriella, Cyclotella
Diversity	Richness

2.7.3. Prediction of metric values at reference sites

To account for environmental variability in reference site populations, we developed predictive MMIs wherein observed metric values are compared to expected values derived from models calibrated at reference sites. We developed random forest models to predict values for all metrics at reference calibration sites using the same candidate predictor variables that were used for O/E index development (Table 2). We again used an RFE approach to select the predictor variables for each metric and created a final 500-tree random forest model for each metric based on the predictors used in the model selected by RFE. We then used the final model to predict metric values for all sites. If the pseudo-R^2 of the model (calculated as 1 – mean squared error [MSE]/variance) was > 0.2 (following Mazor et al., 2016), we adjusted metric values by subtracting predicted values from observed values. For all other metrics, we used the observed metric values.

2.7.4. Metric scoring

Metric values were scored to account for differences in scale direction of response to stress among metrics (Blocksom, 2003). This scoring transformed metrics to a 0 to 1 scale, with lower scores indicative of more high-activity conditions and greater deviation from reference sites. We scored metrics after Cao et al. (2007). We scored metrics that decrease with human activity as:

$$(\text{Observed} - \text{Min})/(\text{Max} - \text{Min})$$

where Min is the 5th percentile of high-activity calibration sites and Max is the 95th percentile of reference calibration sites. We scored metrics that increase with human activity as:

$$(\text{Observed} - \text{Max})/(\text{Min} - \text{Max})$$

where Min is the 5th percentile of reference calibration sites, and Max is the 95th percentile of high-activity sites. We trimmed all scores outside the range of 0 to 1 by either rounding up to 0 or down to 1, respectively.

2.7.5. Metric selection

We selected metrics for possible inclusion in an MMI based on a series of performance screening criteria (Table 4). First, metrics were eliminated if they had inadequate range, which we defined as containing > 1/3 zero or one values (Stoddard et al., 2008) and median values at reference and high-activity sites > 0 (Stevenson et al., 2013). We evaluated the metric signal-to-noise (S:N) ratio as the ratio of the variance among all sites (signal) to the variance of repeated visits to the same site (noise) (Kaufmann et al., 1999). Metrics passed with a S:N ratio > 2 (Stoddard et al., 2008). Additionally, we eliminated metrics with a ratio of between-site to within-site variance < 3 (Mazor et al., 2016; Stoddard et al., 2008). We evaluated metric responsiveness by *t*-test of the metric scores in reference and high-activity sites. We assessed bias by determining whether metric values at reference sites varied among PSA ecoregions, using a threshold of an ANOVA F-statistic > 3.

2.7.6. Assembling the MMI

All metrics that met screening criteria for each algal assemblage were assembled into all possible combinations of proto-MMIs. We selected the top-performing proto-MMIs using screening criteria of regional bias (ANOVA F statistic < 3), precision (standard deviation of reference calibration scores < 0.2), and responsiveness in discriminating reference from high-activity sites (ANOVA F statistic > 100). We then calculated the most frequent metrics in these top performing proto-MMIs and grouped these metrics into thematic types, ensuring a distribution of final metrics across thematic categories (Table 3). The winning metrics from this step were again combined into all possible combinations, and the best performing final MMI was selected based on an averaged ranking of the regional bias, precision, and responsiveness screening criteria. We calculated scores for the final MMI by averaging metrics scores and rescaling by the mean of reference calibration site scores. This process resulted in an index that, like the O/E (null) index, has a mean of 1 at reference calibration sites.

2.8. Index performance evaluation

Index performance was evaluated using the following measures: a) accuracy, or the performance of the index as unbiased against environmental setting or time of sampling; b) precision, or the low variability of the index score within reference sites and among samples from repeated visits within sites; c) responsiveness, or the ability to show large responses to human activities and d) sensitivity, or the ability to score a non-reference site below the impairment threshold. We compared predictive index performance to its null counterpart and also compared performance scores for all indices with both calibration and validation datasets. Accuracy was assessed by comparing mean scores at reference sites. Additionally, we assessed regional bias by calculating the ANOVA *F* statistic for all reference site scores across ecoregions. To further assess regional bias and the influence of natural gradients, we created random forest models using all available environmental variables (Supplemental Table 1) to predict reference site scores, with a lower variance score indicative of a smaller influence of environmental variables on index scores at reference sites. For precision, we evaluated the standard deviation of index scores at reference sites and between sites with repeat sampling events. For responsiveness,

Table 4
Metric screening criteria for evaluating inclusion in the MMI. Description of criteria, statistical test, and threshold for passing.

Description	Test	Threshold	Reference
Regional bias	ANOVA of metric values at reference sites by ecoregion (PSA)	F statistic < 3	Mazor et al., 2016
Responsiveness	t-test comparing reference/high-activity site scores	t statistic > 10	Mazor et al., 2016
Frequency of Zero	Frequency of score = 0	< 33% of scores	Stoddard et al., 2008
Frequency of One	Frequency of score = 1	< 33% of scores	Stoddard et al., 2008
Range of Ref scores	Median score at reference sites	> 0	Stevenson et al., 2013
Range of Stress scores	Median score at high-activity sites	> 0	Stevenson et al., 2013
Signal to Noise	Variance across all sites / variance at repeat site visits	> 2	Stoddard et al., 2008
Repeat visit variation	ANOVA on repeat samplings of station codes	F statistic < 3	Mazor et al., 2016

we compared absolute *t*-statistic values for *t*-test analyses comparing reference and high-activity site scores. Additionally, we created random forest models using all available stressor gradient variables (Supplemental Table 3) to predict scores at all sites, with a higher variance score indicative of a larger response to stressor gradients. We calculated a series of Spearman's correlations between index scores and key natural and stressor gradients to assess sensitivity. Lastly, we reviewed all index performance in consultation with an expert science advisory panel.

3. Reference-based thresholds and application to a statewide assessment

We identified a series of reference-based thresholds to convert scores into condition classes following the approach used in Mazor et al. (2016). For each index, we calculated the 1st, 10th, and 30th percentile of reference calibration scores for all indices using the qnorm function as implemented in the *stats* package in R (R Core Team, 2013). We calculated the number of sites in each class for each PSA ecoregion: likely to be intact (≥30th percentile of reference sites), possibly altered (30th–10th percentile of reference sites), likely to be altered (1st–10th percentile of reference sites, 0.69), very likely to be altered (< 1st percentile of reference sites).

3.1. Comparison to benthic macroinvertebrate index performance

We compared the performance of the algal indices from this work to the benthic macroinvertebrate CSCI index (Mazor et al., 2016) by aligning samples from the same sampling station and sampling dates, resulting in a combined dataset of 1600 observations. The CSCI scores have a reference calibration mean of 1 as do the algal index scores and therefore could be compared directly. We used a reference condition threshold of the tenth percentile of reference site scores (0.79) for evaluating performance of the CSCI (Mazor et al., 2016). To compare CSCI and algal index responsiveness to stressor and natural gradients, we regressed the response of the indices using generalized additive models as implemented in the R package *mgcv* (Wood, 2017). To compare algal and benthic macroinvertebrate assemblage structure, we calculated Bray-Curtis distances among all reference calibration sites using "vegdist" as implemented in the R package *vegan* (Oksanen et al., 2017) on a presence/absence transformed species matrix for both the BMI and algae taxonomy datasets. We sourced benthic macroinvertebrate taxonomy data from the Stormwater Monitoring Coalition dataset (http://smc.sccwrp.org/) and calculated CSCI scores according to Mazor et al. (2016).

4. Results

4.1. Development dataset description

Of 1951 sampling locations, we classified 27% of sites as "Reference", 32% as "Intermediate", and 35% of sites as "High-activity" (Supplemental Fig. 1). The greatest number of Reference sites was found in the high elevation Sierra Nevada region (n = 199), and fewest in the Central Valley region (n = 2).

4.2. O/E index development

Genus-level taxonomy data resulted in stronger O/E indices than species-level data, and all final O/E indices were developed using genus-level taxonomy data. For diatoms, soft-bodied algae, and hybrid assemblages, the strongest performing O/E indices were comprised of biological clusters with 15, 4, and 15 clusters, respectively, and capture probabilities (P_c) of 0.5, 0.4, and 0.4, respectively. (Supplemental Fig. 2, Supplemental Table 2). Catchment area, predicted conductivity (i.e. background reference conductivity), average first/last day of freeze, maximum elevation in catchment, latitude, and average monthly temperature were selected as predictor variables for all three O/E indices (Table 2). The diatom, soft-bodied algae, and hybrid models explained 72, 54, and 65% of the variation in O, respectively, and had regression slopes (0.88, 0.80, 0.93) and y-intercepts (0.82, 0.08, −0.28) that were similar to those expected from unbiased predictions (slope = 1 and y-intercept = 0; Mazor et al., 2016).

We additionally generated statewide O/E indices with an expanded list of predictor variables, including those variables that can be influenced by human activity. With the expanded list of predictor variables, each assemblage saw modest improvements in index precision, with reference calibration SD values of 0.19 (SD null 0.21), 0.32 (SD null 0.36), 0.17 (SD null 0.18) for diatoms, soft-bodied algae, and hybrid, respectively. These O/E models selected predictor variables atmospheric SO_4, and field-measured water temperature and phosphorus concentrations (Supplemental Table 1).

4.3. MMI development

Species-level taxonomy data resulted in stronger metric performance than genus-level data, and all final indices were developed using species-level data. Of all metrics considered for inclusion in the MMIs, ~20% (29/142) of diatom metrics and ~18% (29/158) of hybrid metrics were selected as modeled metrics (random forest pseudo-R^2 value > 0.2). All remaining metrics, including all soft-bodied algae metrics, were therefore not adjusted to account for natural variability (Supplemental Table 4). Metrics were assessed for regional bias, precision, and accuracy, resulting in 17, 4, and 21 metrics considered for inclusion in the proto-MMIs for diatoms, soft-bodied algae, and hybrid MMIs, respectively. After all subsets of these proto-MMIs were compared, the final diatom, soft-bodied algae, and hybrid MMIs were comprised of 6, 4, and 8 metrics, respectively (Table 5), with the diatom and hybrid MMIs each containing 6 and 7 predictive metrics, respectively, and the soft-bodied algae MMI with no predictive metrics. Predicted conductivity (CondQR50) and soil erodibility (KFCT_AVE) were the most frequently selected predictor variables for all predictive metrics included in the final diatom and hybrid MMIs, followed by classification as xeric or montane (XerMtn) and mean annual precipitation (PPT_00_09) (Table 2).

The composition of each MMI varied with each assemblage (Table 5), although high performing metrics were shared across assemblages. The diatom MMI final metrics included two decreaser metrics for *Epithemia* and *Rhopalodia* species richness and low nitrogen indicator taxa, and four other increaser metrics for most tolerant indicator species, planktonic species, eutrophic species, and brackish/freshwater taxa. All diatom MMI metrics were predictive. The hybrid MMI consisted of two decreaser metrics as well, the same *Epithemia* and *Rhopalodia* metric as the diatom MMI in addition to a non-predictive metric for Zygnemataceae, heterocystous cyanobacteria and Rhodophyta (ZHR) taxa. The hybrid MMI also included metrics for high phosphorus indicators, low oxygen tolerant taxa, and four other metrics shared with the diatom MMI for tolerant taxa, planktonic species, eutrophic species, and brackish/freshwater taxa. Of these hybrid MMI metrics, two use both diatom and soft-bodied algae taxa (tolerant taxa and high phosphorous indicators), and one (ZHR) uses solely soft-bodied algae taxa. The soft-bodied algae MMI consisted of only non-predictive metrics, including three indicator taxa metrics for high dissolved organic carbon (DOC), high phosphorus, and non-reference condition, in addition to a metric for ZHR taxa (Table 5).

Multiple metrics across assemblages showed reduced regional bias when these metrics were modeled compared to unmodeled metrics. For example, the diatom metric for the proportion brackish-freshwater taxa (prop.spp.Salinity.BF) saw a large decrease in regional bias when corrected for natural gradients (ANOVA *F* statistic 41 to 0.39). This decrease in regional bias was at times accompanied by a decrease in responsiveness, i.e. the ability to discriminate reference versus high-

Table 5
Metrics selected in final diatom, soft-bodied algae, and hybrid MMIs. Type indicates predictive (P) or non-predictive (N). Response to stress indicated as increaser (inc) or decreaser (dec) metrics. Hybrid metric descriptions indicate whether they use diatom (d) or soft-bodied algae (s) taxonomy. ZHR = Zygnemataceae + heterocystous cyanobacteria + Rhodophyta. References: A = Fetscher et al., 2014; B = van Dam et al., 1994), C = Bahls, 1993), D = Stancheva and Sheath, 2016.

Metric name	Description	Diatom	SBA	Hybrid	Type	Stress	Ref.
cnt.spp.most.tol	Count of most tolerant indicator species	x			P	Inc	this work
EpiRho.richness	*Epithemia* and *Rhopalodia* richness	x			P	Dec	–
prop.spp.IndicatorClass_TN_low	Proportion of low total N indicator species	x			P	Dec	A
prop.spp.Planktonic	Proportion of planktonic species	x			P	Inc	C
prop.spp.Trophic.E	Proportion of eutrophic species	x			P	Inc	B
Salinity.BF.richness	Brackish/freshwater species richness (d)	x			P	Inc	B
cnt.spp.IndicatorClass_TP_high	Proportion of high total P indicator species (d,s)			x	P	Inc	A
cnt.spp.most.tol	Count of most tolerant indicator species (d,s)			x	P	Inc	this work
EpiRho.richness	*Epithemia* and *Rhopalodia* richness (d)			x	P	Dec	–
OxyRed.DO_30.richness	Richness of species requiring 30% oxygen (d)			x	P	Inc	B
prop.spp.Planktonic	Proportion of planktonic species (d)			x	P	Inc	C
prop.spp.Trophic.E	Proportion of eutrophic species (d)			x	P	Inc	B
prop.spp.ZHR	Proportion of species ZHR (s)			x	N	Dec	D
Salinity.BF.richness	Brackish/freshwater species richness (d)			x	P	Inc	B
prop.spp.IndicatorClass_DOC_high	Proportion of high DOC indicator species		x		N	Inc	A
prop.spp.IndicatorClass_NonRef	Proportion of non-reference indicator species		x		N	Inc	A
prop.spp.IndicatorClass_TP_high	Proportion of high total P indicator species		x		N	Inc	A
prop.spp.ZHR	Proportion of ZHR species		x		N	Dec	D

activity sites. For example, this same diatom metric (prop.spp.Salinity.BF) saw a decrease in responsiveness (*t*-statistic 29 to 25) with modeling. However, this was not true across all metrics as the hybrid MMI metric for oxygen preference (OxyRed.DO_30.richness) had both a decrease in regional bias (ANOVA *F* statistic 25 to 2.5) and an increase in sensitivity (*t*-statistic 21 to 26) with modeling (Supplemental Table 4). For the modeled metrics in the diatom and hybrid MMIs, the regional bias was decreased for all metrics with modeling, whereas sensitivity increased in four of the six diatom metrics and three of the six hybrid metrics with modeling (Supplemental Table 4).

4.4. Performance evaluation of the indices

We assessed all final indices for accuracy, precision, responsiveness, and sensitivity (Table 6). All final O/E and MMI indices were able to discriminate reference, intermediate, and high-activity sites (Fig. 1), although with varying degrees of accuracy and precision. All O/E and MMI indices had comparable regional bias performance, assessed as the distribution of reference site scores across PSA regions (Supplemental Fig. 3; Table 6). Similarly, all final predictive indices had low variance scores and substantial improvement over null model variance scores, indicating that predictive modeling had adequately accounted for the influence of natural gradients (Table 6; Fig. 2). MMI indices for each assemblage had better precision than their O/E counterparts, while all three O/E indices and the soft-bodied algae MMI had among-site precision scores above our target range (SD < 0.18), as well as higher within-site precision scores (Table 6). The three MMIs also had better responsiveness than the O/E indices, although the soft-bodied algae MMI had the weakest responsiveness of the three MMIs (Table 6; Fig. 3). The predictive diatom and hybrid MMIs had comparable among-site precision and responsiveness and were both able to accurately discriminate reference versus high-activity sites as well as variance in index scores that could be explained by human-activity gradients (Table 6).

All O/E and MMI indices classified between 89 and 95% of reference calibration sites correctly (> 10th percentile of reference). However, the predictive diatom and hybrid MMIs had much higher

Table 6
Performance measures to evaluate all final O/E and MMI indices at calibration (Cal) and validation (Val) sites. Type = null or predictive indices. Accuracy: mean score = mean score of reference sites (* indicates value is mathematically fixed at 1); F = F-statistic for differences in scores at reference calibration sites among five PSA regions (Central Valley excluded); Var = variance in index scores explained by natural gradients at reference sites. Precision: Among SD = standard deviation of scores at reference calibration and validation sites; Within SD = standard deviation of within-site scores for reference calibration and validation sites with multiple samples. Responsiveness: t = t-statistic for difference between mean scores at reference and high-activity (stressed) sites; Var = variance in index scores explained by human-activity gradients at all sites. Spearman's correlation Rho values for key stressor gradients total nitrogen (TN), total phosphorus (TP), specific conductivity (SC).

			Accuracy						Precision				Responsiveness				Spearman's Rho		
			Mean score		*F* (PSA)		Var		Among SD		Within SD		*t*		Var		TN	TP	SC
	Index	Type	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val			
Diatoms	MMI	pred	1.00	1.00	0.42	0.73	−0.13	0.01	0.11	0.14	0.09	0.06	34.4	15.4	0.57	0.51	−0.56	−0.54	−0.51
Diatoms	MMI	null	1.00	1.01	13.48	4.88	0.15	0.17	0.19	0.16	0.09	0.07	33.8	18.2	0.61	0.59	−0.64	−0.60	−0.61
Hybrid	MMI	pred	1.00	1.00	0.07	0.87	−0.15	−0.01	0.11	0.13	0.09	0.06	34.9	16.7	0.54	0.48	−0.55	−0.53	−0.47
Hybrid	MMI	null	1.00	1.02	14.23	4.02	0.13	0.19	0.18	0.16	0.09	0.07	32.2	18.3	0.58	0.57	−0.63	−0.57	−0.57
SBA	MMI	null	1.00	0.97	1.00	0.98	0.02	0.00	0.27	0.29	0.10	0.13	16.2	6.0	0.28	0.25	−0.44	−0.42	−0.34
Diatoms	O/E	pred	1.01	0.99	0.50	0.74	−0.22	−0.08	0.18	0.19	0.10	0.07	9.1	3.8	0.20	0.22	−0.33	−0.19	−0.29
Diatoms	O/E	null	1.00	0.99	5.17	2.30	0.16	0.20	0.20	0.23	0.08	0.08	7.5	3.1	0.18	0.19	−0.28	−0.18	−0.25
Hybrid	O/E	pred	1.01	1.00	1.00	1.34	−0.10	0.00	0.20	0.21	0.14	0.13	16.4	6.8	0.26	0.25	−0.40	−0.33	−0.32
Hybrid	O/E	null	1.00	0.98	2.73	1.70	0.19	0.16	0.22	0.24	0.10	0.12	15.9	5.9	0.25	0.24	−0.39	−0.35	−0.31
SBA	O/E	pred	1.01	0.97	0.84	2.23	−0.12	0.05	0.37	0.42	0.27	0.15	13.5	4.8	0.19	0.17	−0.28	−0.33	−0.19
SBA	O/E	null	1.00	0.96	1.77	1.59	0.13	0.17	0.44	0.48	0.21	0.17	13.9	4.8	0.21	0.19	−0.28	−0.37	−0.20

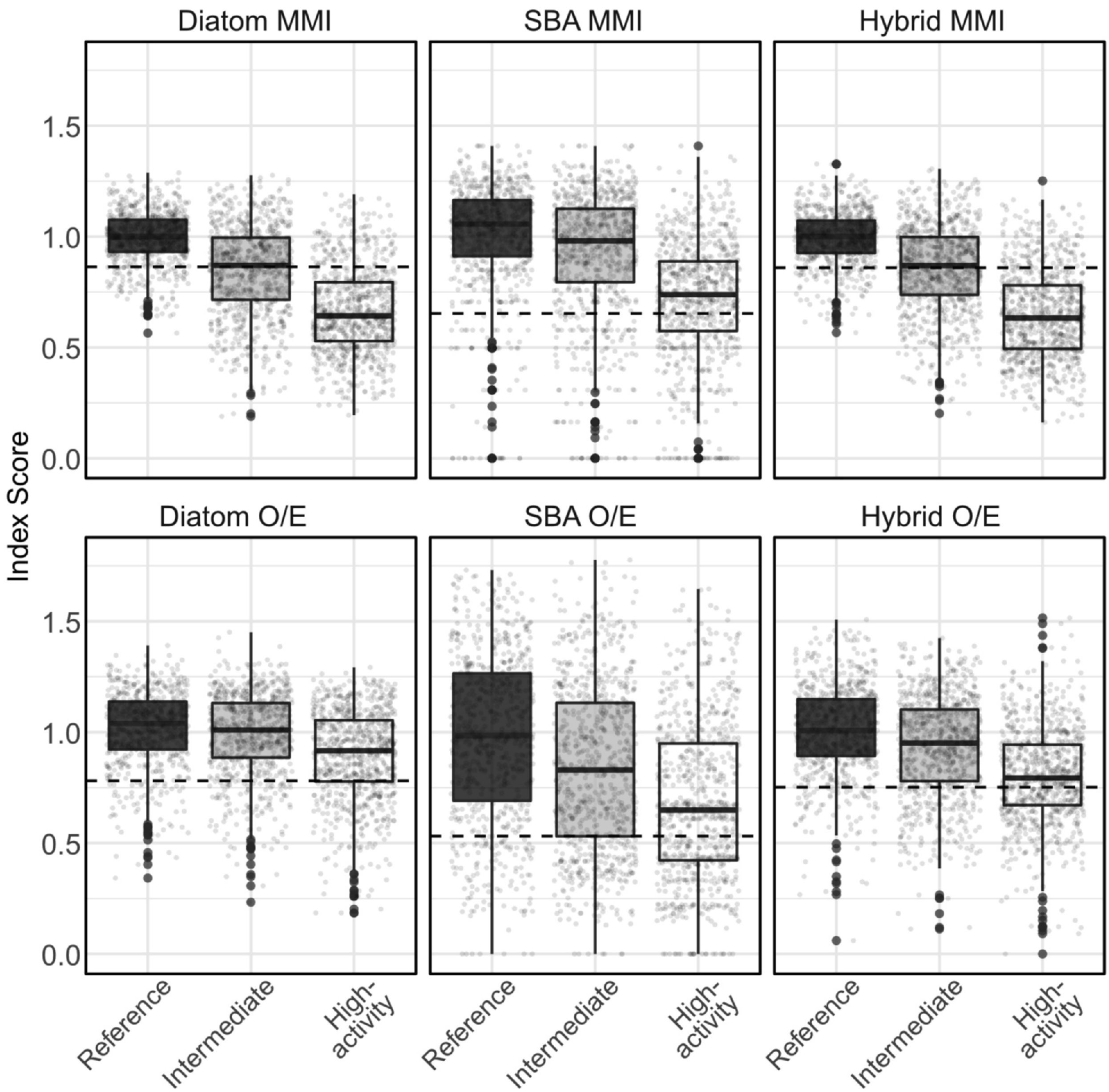

Fig. 1. Diatom, soft-bodied algae (SBA), and hybrid index scores by reference, intermediate, and high-activity sites for the O/E and MMI. Dashed line represents the 10th percentile of reference calibration site scores.

proportions (85% and 84%, respectively) of high-activity sites that were classified correctly (< 10th percentile of reference) in comparison to the O/E indices only classified between 26 and 41% of high-activity sites correctly (Table 7). Based on criteria developed with the expert science advisory panel, none of the three algal O/E indices are considered to have adequate performance, due to poor responsiveness, precision, and sensitivity, and are not recommended for future use. Likewise, the soft-bodied algae MMI is not recommended due to poor precision.

4.5. Effect of E on index performance

All three O/E indices saw improvements in performance with increasing E values. Accuracy and sensitivity of the O/E indices increased steadily with increasing E values (Fig. 4A), while precision was largely unaffected. Reference site expectations for proportions of sensitive (prop.spp.BCG12) and tolerant (prop.BCG4.taxa) taxa reflect a similar pattern, wherein at lower levels of E, a smaller proportion of sensitive taxa and a larger proportion of tolerant taxa are expected to occur (Fig. 4B). Additionally, we saw on average higher E values for hybrid O/E models than a combined sum of diatom and soft-algae E values for the same sites (Fig. 4C).

4.6. Comparison to benthic macroinvertebrate index and assemblage structure

The CSCI is now a widely accepted biological index for California wadeable streams, and thus we used CSCI performance as a benchmark against which to compare algal index performance criteria. Precision and accuracy of the diatom O/E was comparable to the benthic macroinvertebrate CSCI O/E index (Mazor et al., 2016), with standard deviation of reference site scores of 0.18 and 0.19, respectively, and lower within-site variability in the diatom O/E versus the CSCI O/E (0.10 and 0.16, respectively). In contrast, the diatom O/E from this study was less responsive in its ability to discriminate reference versus high-activity sites, although correlations to stressor gradients were comparable

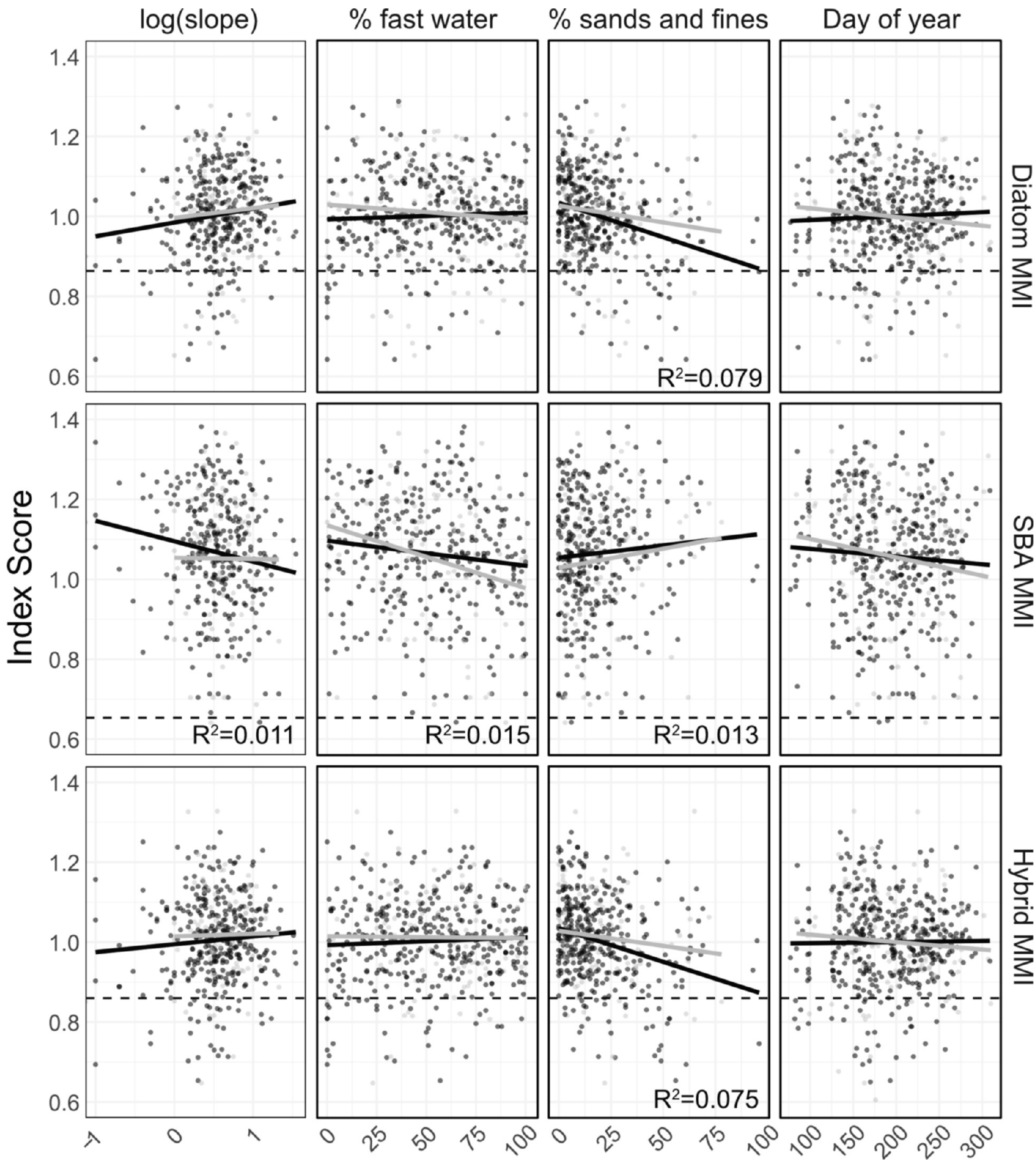

Fig. 2. Diatom, soft-bodied algae (SBA), and hybrid MMI scores across environmental gradients for reference sites. Linear regression lines included for calibration (black) and validation (grey) sites. R^2 values for linear regression of relationship with reference calibration scores as shown (for relationships $p < 0.05$). Dashed line indicates 10th percentile of reference calibration site scores for each index.

(Mazor et al., 2016). In contrast, the diatom and hybrid MMIs had comparable precision to the CSCI MMI and the combined CSCI (O/E + MMI), both for standard deviation of reference scores and within site variability. The diatom and hybrid MMIs also had equivalent accuracy and responsiveness to the CSCI MMI and combined CSCI (Mazor et al., 2016), whereas the soft-bodied algae MMI had worse performance than the CSCI for responsiveness.

Across all sites, the performance of the diatom and hybrid algal MMIs was correlated to the performance of the CSCI (Supplemental Fig. 4) with R^2 values of 0.26 and 0.30, respectively. The diatom and hybrid predictive MMIs, and the CSCI, all exhibited similar responsiveness to key stressor gradients, whereas the non-predictive soft-bodied algae MMI was not as responsive (Supplemental Fig. 5). In response to nutrient and physical habitat gradients, the algal MMIs and the CSCI had similar responses to nutrient (nitrogen and phosphorus) gradients, with the diatom and hybrid predictive MMI exhibiting slightly greater sensitivity to nitrogen and percent sands and fines gradients. In contrast, the CSCI exhibited a more sensitive response to physical habitat alteration as measured by California Rapid Assessment Method (CRAM) scores (Supplemental Fig. 5). In comparing CSCI and hybrid MMI scores, the two indices both ranked a site as “very likely intact” (> 10th percentile of reference calibration scores for both CSCI

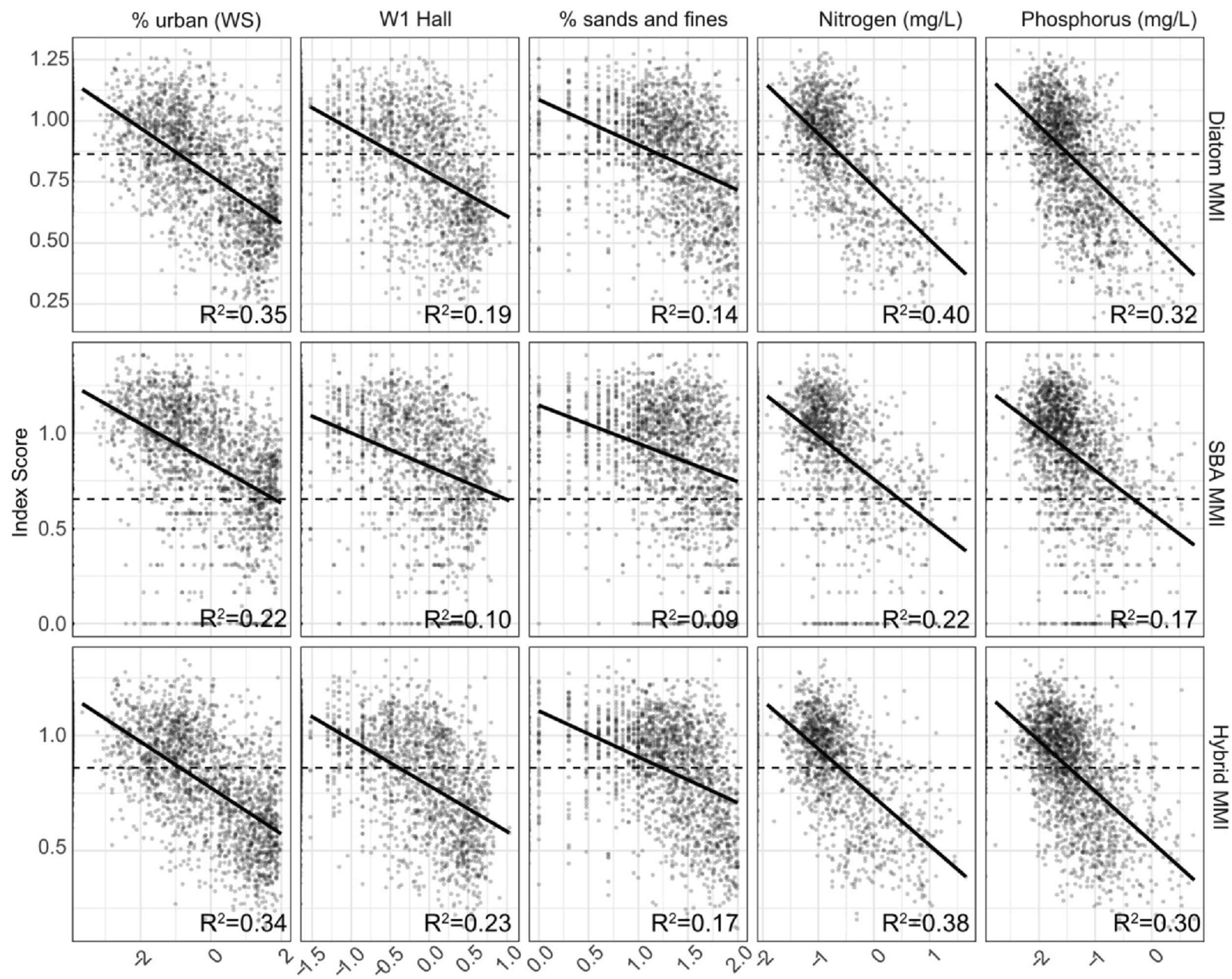

Fig. 3. Diatom, soft-bodied algae (SBA), and hybrid MMI scores across (log transformed) stressor gradients of percent urban development at the watershed scale (% urban WS), riparian activity (W1 Hall), percent sands and fines, total nitrogen, and total phosphorus. R^2 values for linear regression of relationship as shown ($p < 0.001$). Dashed line indicates 10th percentile of reference site scores for each index.

Table 7
O/E and MMI condition classes based on percentile of reference calibration site scores and percentage of sites correctly classified in corresponding condition classes.

	Diatom O/E	SBA O/E	Hybrid O/E	Diatom MMI	SBA MMI	Hybrid MMI
1st percentile of reference	0.60	0.14	0.54	0.75	0.37	0.75
10th percentile of reference	0.78	0.53	0.75	0.86	0.65	0.86
30th percentile of reference	0.92	0.81	0.91	0.94	0.86	0.94
% reference calibration above 10th percentile of reference	89	89	89	91	95	91
% reference validation above 10th percentile of reference	88	83	89	87	90	87
% high-activity below 10th percentile of reference	26	41	41	85	36	84

and hybrid MMI) 41% of the time. When the two indices did not agree, the CSCI ranked a site below the 10th percentile of reference 18% of the time, while the hybrid MMI did 9% of the time (Supplemental Fig. 6). The Sierra Nevada and North Coast regions had the greatest proportion of sites in agreement (> 80%), with both indices assessing a majority of sites as likely intact. The other four eco regions had roughly equivalent proportions of sites with CSCI and hybrid MMI scores in disagreement (~30%), with the hybrid MMI ranking a site lower more often in the highly urbanized South Coast region and the CSCI more likely to rank a site lower in the heavily agricultural Central Valley (Supplemental Fig. 6).

Across the same subset of reference sites, benthic macroinvertebrate and algal populations exhibited different alpha and beta diversity characteristics. We found that on average, pairwise Bray-Curtis dissimilarity between reference sites for algae assemblages averaged 0.75, whereas dissimilarities based on BMI assemblages averaged 0.70 (Fig. 5), indicating that BMI taxa had greater overlap in species assemblages within the reference pool than did algal populations. We identified similar numbers of species richness at reference sites, 691 for BMI taxa, 702 for diatoms, and 569 species for soft-bodied algae. However, we found that BMI taxa had an average per site species richness of 56 (SD 14), while diatoms and soft-bodied algae had average per site species richness values of 26 (SD 12) and 14 (SD 10), respectively. Even when compared with a combined diatom and soft-bodied algae assemblage, the average algal species richness per site was 41 (SD 18), still ~ 27% lower than the BMI species richness per site.

4.7. Distribution of biological condition classes across state

We established four biological condition classes based on the

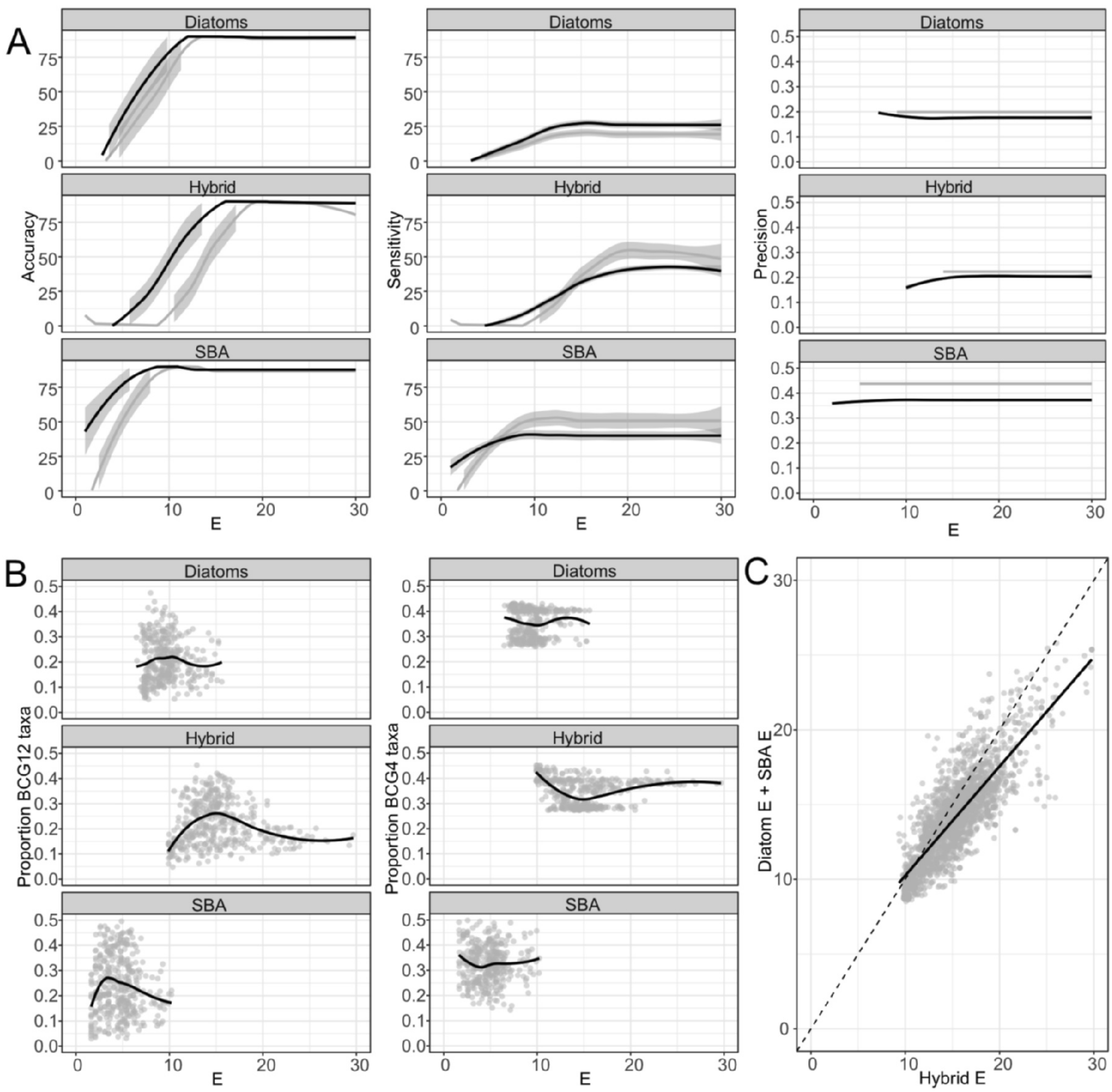

Fig. 4. Effects of expected taxa on index performance. A) Effect of the expected number of taxa (E) on the accuracy, sensitivity, and precision for null (grey) and predictive (black) O/E indices. Locally-weighted regression line and standard error as shown. Accuracy = proportion of reference calibration sites in reference condition (> 10th percentile of reference calibration scores) for each index. Sensitivity = proportion of high-activity sites not in reference condition (< 10th percentile of reference calibration scores). Precision = standard deviation of reference calibration sites for each index. B) Comparison between expected number of taxa (E) from O/E models versus expected proportions of sensitive (BCG12) or tolerant (BCG4) taxa from MMI development. Black line represents locally-weighted regression line. C) Comparison of hybrid O/E index expected taxa (E) versus the sum of diatom and soft-algae O/E index expected taxa (E). Black line represents best fit linear model, dashed line represents 1:1 relationship.

distribution of algal index scores at reference calibration sites using the hybrid MMI (Table 8). Statewide, 37% of stream sites were likely to be intact (hybrid MMI ≥ 0.94 [30th percentile of reference calibration sites]). Another 13% were possibly altered (hybrid MMI ≥ 0.86 [10th percentile]), 16% were likely to be altered (hybrid MMI ≥ 0.75 [1st percentile]), and 34% were very likely to be altered (< 1st percentile; Table 8). Although 69% of high-activity sites were very likely to be altered, this number varied considerably by region. The South Coast and the Chaparral had the highest percent of high-activity sites that were considered very likely to be altered (76 and 69%, respectively), while 26% of the Desert/Modoc and 8% of the Sierra Nevada high-activity sites were classified as likely altered. Statewide, about 71% of reference sites were classified as likely to be intact, with the highest percentages (74%) in the Sierra Nevada and North Coast regions, and lowest in the Central Valley where there were only two reference sites and they were considered likely or possibly altered (Fig. 6).

5. Discussion

Algal indicators provide a powerful line of evidence on biological condition. However, an algal index that is applicable across diverse natural landscapes, such as those found in California, relies on the ability to account for the influence of natural factors that vary among regions, such as climate. Predictive biological indices provide the necessary adjustments to account for environmental effects on biological assemblages. This study highlights the comparative strengths and weaknesses of two different types of predictive algal indices, an O/E and an MMI, and has implications for future development of both predictive and multi- assemblage algal indices in diverse landscapes.

5.1. Limitations in O/E index performance

All three algal assemblages yielded O/E indices with moderate to weak performance. Two key factors may contribute to the difficulty of

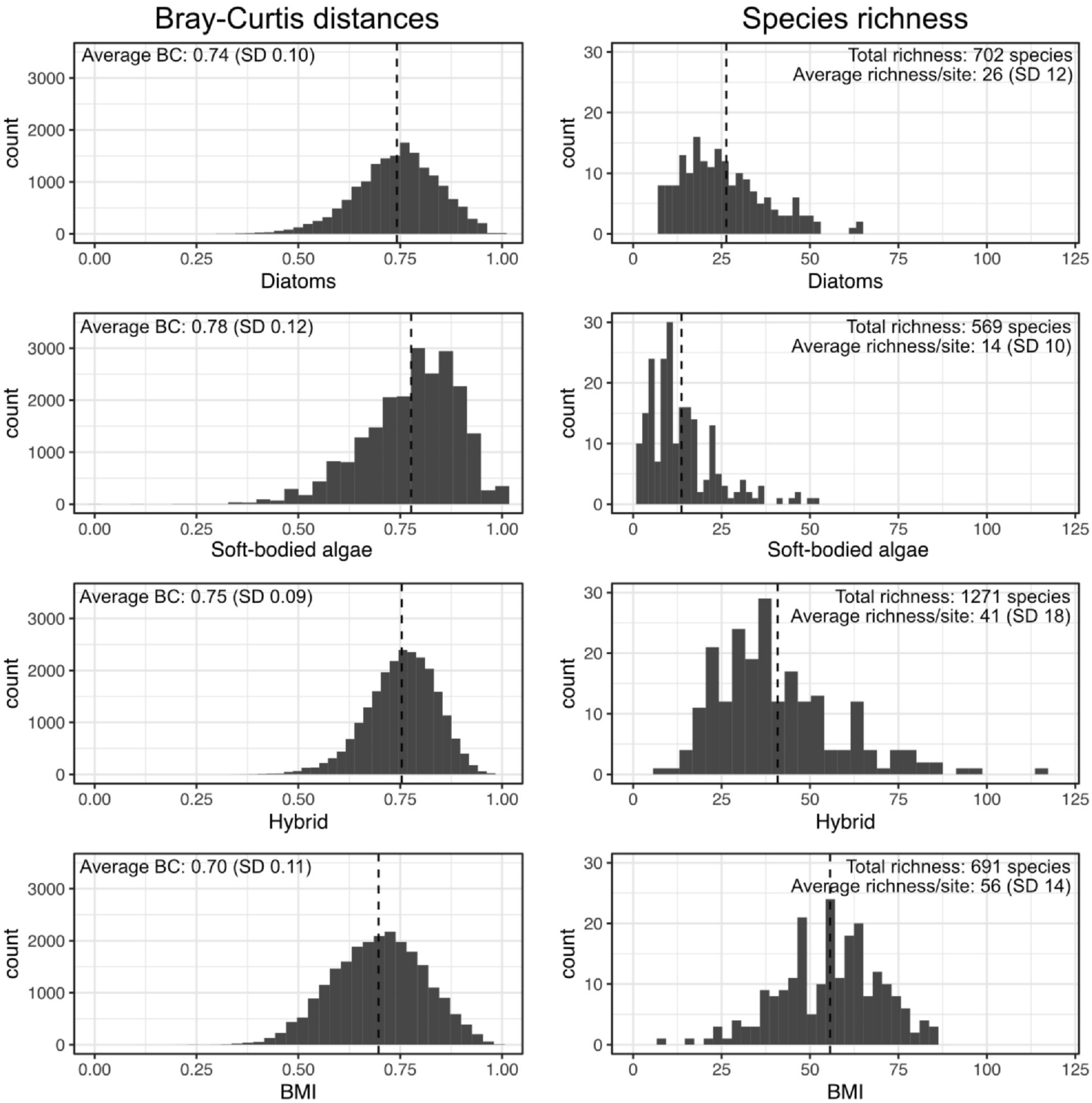

Fig. 5. Comparison of algae and benthic macroinvertebrate (BMI) Bray-Curtis distances and species richness at reference sites. Graphs show distribution of Bray-Curtis distances for comparison of all site × site comparisons (left panels) as well as distribution of species richness (right panels) across 216 reference sites with both algae and BMI data. Vertical dashed line represents the mean.

predicting algal species distributions, as is necessary in the development of O/E models: a) geographic parameters used as candidate predictor variables are largely not responsible for shaping algal species distribution patterns; b) high beta diversity between reference sites pose a challenge to modeling efforts. These factors are discussed below.

The inclusion of additional candidate predictor variables, including those influenced by human activity such as atmospheric SO_4, total phosphorus, and in-field temperature, resulted in modest improvements in O/E index precision for all three assemblages. While these improvements in performance were small, this finding demonstrates that additional predictor variables were able to help explain residual species distributions effects that were not accounted for with our initial candidate predictor list. Previous studies have likewise found that locally measured environmental variables, such as light availability and velocity, may improve species distribution model performance (Aguiar et al., 2011; Feio et al., 2012; Lamb and Lowe, 1987; Sabater, 2006; Soininen, 2005; Veraart et al., 2008). However, such models are less useful for bioassessment index development as they may be unable to separate the influence of natural variation in these factors from anthropogenic disturbance (Hawkins et al., 2010a; Mendes et al., 2014; Reynoldson et al., 1997), a key requirement of successful predictive indices for bioassessment applications. The expansive suite of candidate predictor variables used in this study included modeled background levels of conductivity (CondQR50, Table 2) a variable that allows for the accounting of a variable (conductivity) that can otherwise be subject to human influence. The continued development of candidate predictor variables that account for natural background levels, such as water temperature or phosphorus concentrations, will likewise benefit future predictive models.

High beta diversity of algae population at reference sites may be responsible for the poor performance of the predictive O/E indices. Our results found that in comparison to BMI populations, diatoms and soft-

Table 8
Percentage of sites in different condition classes by region and site status based on hybrid MMI scores. Percentiles refer to the distribution of scores at reference calibration sites.

		Total sites	Likely to be intact ≥ 30th percentile (MMI ≥ 0.94)	Possibly altered 30th– 10th percentile (MMI ≥ 0.86)	Likely to be altered 1st–10th percentile (MMI ≥ 0.75)	Very likely to be altered < 1st percentile (MMI < 0.75)
North Coast	Reference	100	74	16	7	3
	Intermediate	105	44	20	25	11
	High-activity	46	24	20	17	39
	Total	251	52	18	16	13
Desert/Modoc	Reference	49	63	20	16	0
	Intermediate	58	40	21	31	9
	High-activity	8	25	0	50	25
	Total	115	49	19	26	6
Chaparral	Reference	169	70	17	10	3
	Intermediate	192	38	17	24	21
	High-activity	214	7	6	18	69
	Total	575	36	13	18	34
Central Valley	Reference	2	0	50	50	0
	Intermediate	14	43	7	7	43
	High-activity	80	11	16	14	59
	Total	96	16	16	14	55
South Coast	Reference	216	68	14	11	7
	Intermediate	366	27	12	19	42
	High-activity	481	6	6	12	76
	Total	1063	26	10	14	50
Sierra Nevada	Reference	199	74	20	5	1
	Intermediate	115	50	19	23	8
	High-activity	12	58	8	25	8
	Total	326	65	19	12	4
Statewide	Reference	737	71	17	9	3
	Intermediate	850	36	16	22	26
	High-activity	841	9	7	15	69
	Total	2428	37	13	16	34

bodied algae have comparable species richness across all reference sites, but < 50% of average species richness at individual reference sites (Fig. 5). This results in higher average algal Bray-Curtis (beta diversity) distances among all reference sites, and therefore greater average dissimilarity between two reference sites. Taken together, these results suggest that algal populations at two sites of comparable ecological status and similar geographic setting may have only minimal overlap in species composition. Increased beta diversity has been shown to occur with decreased connectivity and dispersal between habitats (Forbes and Chase, 2002; Matthiessen et al., 2010; Vander Laan and Hawkins, 2014), and may be explained by the presence of intermittent streams subject to frequent drying events (Vander Laan and Hawkins, 2014) and the absence of overland dispersal seen in invertebrate communities (Cauvy-Fraunié et al., 2015). However, it is less clear why soft-bodied algae and diatoms would experience different responses to habitat fragmentation, but may be the result of differences in species resiliency and turnover rates, with previous studies finding that diatoms may have greater dispersal rates than soft-bodied algae (Schneider et al., 2012) and would therefore be expected to have greater similarity across sites than soft-bodied algae populations, as we observed.

Additionally, these dissimilar algal assemblages at reference sites may help explain the low E values that were common across the algal O/E models: with a greater number of taxa spread across reference sites, and few taxa shared among sites, the O/E models were only able to confidently predict a handful of taxa at each site. Likewise, this helps explain why genus-level taxonomy data resulted in higher-performing O/E indices, as aggregating to a higher taxonomic level helped to increase the common taxa shared between sites. For algal populations, E values averaged 10 for diatoms and 5 for soft-bodied algae, whereas for BMI populations, E values as low as 5 were rarely observed (Mazor et al., 2016). These low E values for algae taxa resulted in poor accuracy and sensitivity in each of the three algal O/E indices (Fig. 4A), consistent with previous observations that low E values limited benthic macroinvertebrate index performance (Mazor et al., 2016). Low E sites are those with fewer common taxa, a reflection of either a more disconnected reference site pool (Vander Laan and Hawkins, 2014) and/or lower species diversity at individual sites within the reference site pool. Additionally, low E sites have been found to be associated with more stressful environments, including those subject to drying events (Mazor et al., 2016; Vander Laan and Hawkins, 2014), and therefore may be dominated by more tolerant, resilient taxa, as suggested by the greater proportion of tolerant algae taxa that were predicted to occur at low E sites (Fig. 4C).

5.2. Performance of MMIs varied across algal assemblages

All three assemblages yielded MMIs that were responsive to stressor gradients, although the soft-bodied algae MMI was far less precise than either the diatom or hybrid MMIs. For diatom and hybrid assemblages, predictive modeling improved MMI index performance and resulted in indices with greater responsiveness and less regional bias, consistent with previous studies that have seen improvements in index performance with modeling for natural variability (Cao et al., 2007; Hawkins et al., 2010a; Mazor et al., 2016; Vander Laan and Hawkins, 2014).

The absence of predictive metrics in the soft-bodied algae MMI highlights the difficulties in building predictive models for soft-bodied algal assemblages. As highlighted above, the soft-bodied algae communities exhibited the lowest species richness and the greatest beta diversity across reference sites, therefore posing inherent challenges for identifying geographic parameters that were responsible for trait distributions. The poor precision of the soft-bodied algae MMI could be in part be attributed to the fact that 11% of the reference site pool had species richness values of ≤ 3 (Fig. 5), well below the average of 14 species for soft-bodied algae at reference sites, and a third of these low-richness reference sites earned a soft-bodied algae MMI score of zero (Fig. 1). Additionally, the qualitative soft-bodied algae fraction was omitted in the creation of the MMIs due to its absence in > 40% of all samples; however, previous index development in Southern California has shown that the qualitative fraction helps improve the diagnostic signal of soft-bodied algae assemblages by increasing the observed

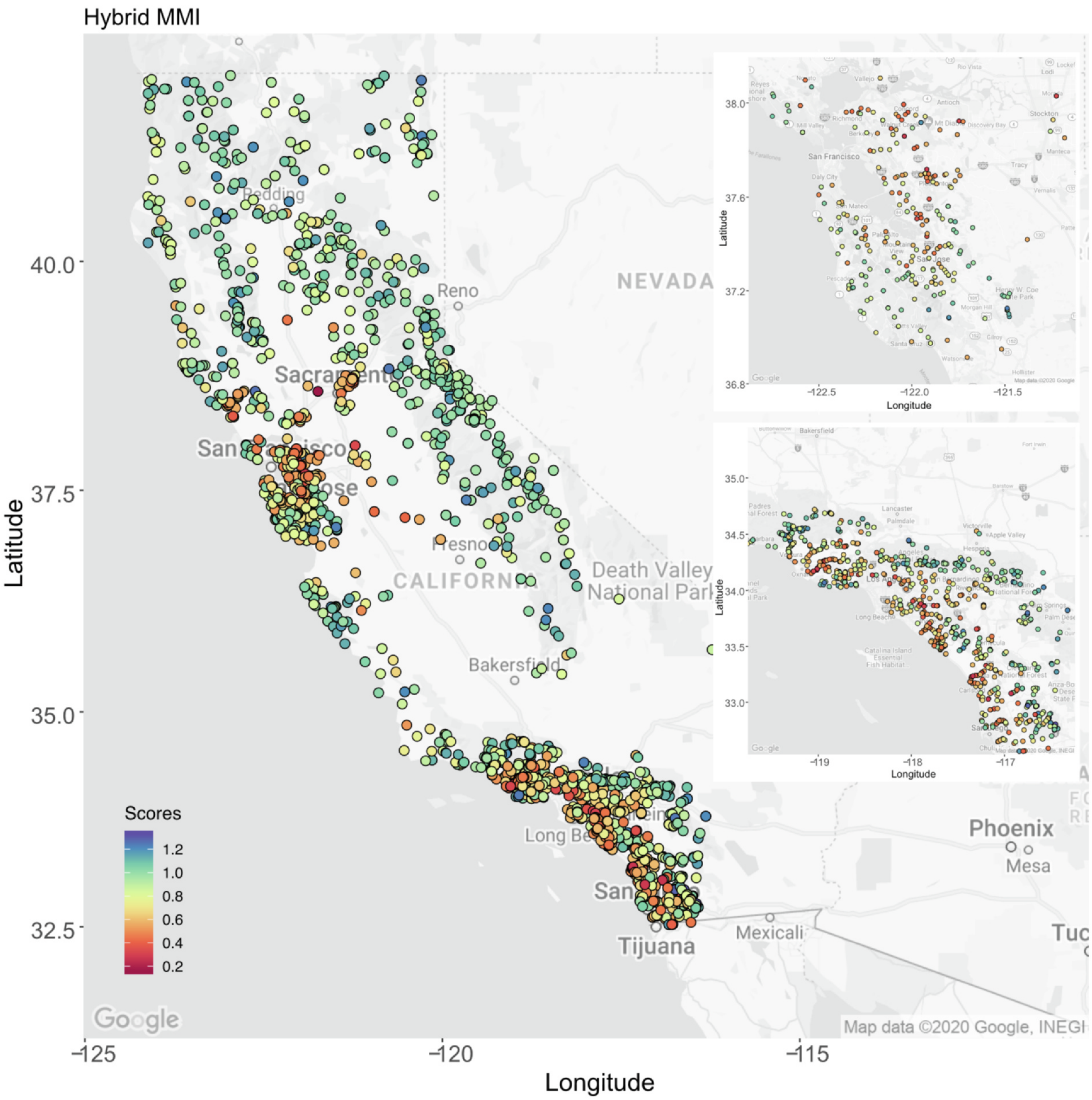

Fig. 6. Hybrid MMI results across California. Inset: detail view of hybrid MMI scores for the Bay Area (top) and Los Angeles (bottom).

species richness across sites (Fetscher et al., 2014a). Eliminating the qualitative sample from our analyses may have exacerbated the low species richness of the soft-bodied algae populations and further depressed the biological signal available for model development. Lastly, the paucity of available trait attribute information for soft-bodied algae (Fetscher et al., 2014a) effectively reduced the number of metrics for proto-MMI testing. Future efforts to expand the coverage of auto-ecological traits for California taxa are warranted, including the further development of trait attributes through empirical testing, targeted mesocosm studies, nonparametric techniques such as Threshold Indicator Taxa Analyses (TITAN, Baker and King, 2010) and machine-learning approaches (Cordier et al., 2018; Feio et al., 2020).

The diatom and hybrid MMIs had very comparable performance, although using a hybrid index that incorporates two distinct assemblages presents certain advantages and disadvantages. Previous studies have found that diatom and soft-bodied algae communities respond differently to nutrient conditions, with diatom richness generally increasing with nutrient availability in contrast to decreasing soft-bodied algae richness (Schneider et al., 2012; Stancheva and Sheath, 2016). Diatoms are credited with faster response times (Lavoie et al., 2008) and soft-bodied algae provide a more integrative response over longer time scales (Whitton, 2012). Although only three of the seven hybrid metrics include soft-bodied algae taxonomy (Table 5), by incorporating both assemblages the hybrid index may provide a more integrative assessment of biological condition than a single assemblage index. However, environmental managers must also consider the increased financial and technical burden of analyzing two algal assemblages. Soft-bodied algae can present unique challenges for identification with light microscopy (Stancheva and Sheath, 2016), and bioassessment programs such as the National Rivers and Streams Assessment (U.S. Environmental Protection Agency, 2016) and many countries within the EU Water Framework Directive (Almeida and Feio, 2012) have elected to focus solely on diatom communities. Likewise, California bioassessment program managers will have to determine if cost or technical limitations may necessitate the use of the diatom-only index for certain applications. A transition to a molecular, or DNA-based, approach may help alleviate both the financial and technical challenges as DNA-based analyses have the potential to offer considerable cost savings and may help circumvent technical barriers to implementation (Darling and Mahon, 2011; Valentini et al., 2016).

5.3. Implications for bioassessment applications

A primary motivation in developing algal indices for California was to create a complementary line of evidence of biological condition for streams to use alongside the benthic macroinvertebrate CSCI (Almeida and Feio, 2012). BMI and algae have been shown to have different sensitivities to environmental stressors (Carlisle et al., 2008; Johnson and Hering, 2009; Sonneman et al., 2001), with BMI communities in general responding more sensitively to physical habitat structure (Voss et al., 2012) and hydromorphological pressures (Pardo et al., 2014) and algal communities to nutrients and water chemistry (Hering et al., 2006; Johnson and Hering, 2009; Pardo et al., 2018; Rehn, 2016; Sonneman et al., 2001). In agreement with these previous studies, we see a stronger response to physical habitat alterations by the CSCI, while the algal indices had a stronger response to nutrient gradients (Supplemental Fig. 5). Across the state, the majority of CSCI- and the hybrid MMI-derived assessment endpoints were in agreement (Supplemental Fig. 6), but the instances when these two indices disagreed helps to highlight the situations wherein the BMI and algal indices are responding differently to environmental conditions. Given that both the CSCI and the predictive diatom and hybrid MMI indices presented here demonstrated consistent performance in assessing reference conditions across diverse ecoregions, both are appropriate for statewide application. Combining multiple lines of evidence from both algal and benthic macroinvertebrate indices will provide a more integrated perspective on biological condition across a broader spectrum of environmental stressors.

This study developed both diatom and a hybrid predictive multimetric indices whose performance achieved high precision, accuracy, responsiveness, and low regional bias, therefore qualifying them as suitable tools for statewide management applications. The development and application of multimetric indices for evaluating biological condition has grown increasingly popular in the past four decades (Ruaro et al., 2020), including in the creation of the Indice de Polluosensibilité Spécifique (IPS, Cemagref, 1982) and the Diatom Biological Index (IBD, Coste et al., 2009). MMIs incorporate multifaceted biological attributes (Chen et al., 2019; Stoddard et al., 2008), provide integrative assessments of biological assemblages that are understandable for a broad audience (Karr and Chu, 2000), and can help identify causal stressors on aquatic ecosystems (Hering et al., 2006; Lunde and Resh, 2012; Martins et al., 2020). Accurate, sensitive, and precise predictive multimetric indices, like those presented here, are calibrated on a site-specific, reference-based expectations. This approach is able to provide a robust assessment of deviation from reference condition as a means to evaluate impacts to biological communities and serve as a powerful indicator of ecological condition.

6. Conclusions

The 1972 Clean Water Act explicitly mentions the use of biological integrity as an ecological endpoint, and with this study, we endeavored to create a powerful bioassessment tool that could be used for assessing biological condition throughout California's complex ecoregions. In this effort, we found that the inclusion of predictive modeling greatly improved algal index performance by reducing regional bias and increasing index precision and sensitivity, yielding both a diatom and a hybrid predictive multimetric index that are suitable for statewide application. The limitations of predictive modeling of taxonomic completeness for the O/E indices could be credited to the disparate and diverse algal communities that severely limited numbers of expected taxa and therefore resulted in poor precision and accuracy. In contrast, by modeling ecological structure, we were able to capitalize on aggregated trait attribute information and therefore generated stronger predictive models and better-performing indices. This study adds to the growing body of evidence that predictive modeling improves index performance in complex environments and also helps demonstrate the biological and technical challenges that remain for predicting diverse species assemblages in disconnected environments. Future studies will benefit from a continued investment in the generation of both robust taxonomic datasets and environmental predictor portfolios that will help ensure the success of algal biological indices.

Funding

This work was supported by the California State Water Resources Control Board under agreements No. 17-070-270 and D15-15002.

CRediT authorship contribution statement

Susanna Theroux: Conceptualization, Formal analysis, Methodology, Software, Writing - original draft. **Raphael D. Mazor:** Conceptualization, Methodology, Software, Resources, Writing - original draft. **Marcus W. Beck:** Software, Visualization. **Peter R. Ode:** Conceptualization, Methodology. **Eric D. Stein:** Writing - review & editing, Project administration, Supervision. **Martha Sutula:** Writing - review & editing, Project administration, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the California Surface Water Ambient Monitoring Program (SWAMP), the Southern California Stormwater Monitoring Coalition, the Bay Area's Regional Monitoring Program, and other data providers. We thank SWAMP's bioassessment workgroup, especially A. Elizabeth Fetscher, Chad Loflen, Kevin Lunde, Andrew Rehn, and Rosalina Stancheva. Thanks to Aaron Jones and Kenny McCune for technical assistance. Special thanks to the Biostimulatory and Biointegrity Science Advisory Panel for their expert guidance: Chuck Hawkins, Jan Stevenson, Ken Reckhow, Lester Yuan, Paul Stacey, and Clifford Dahm. We dedicate this manuscript to Robert Sheath.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2020.106421.

References

Aguiar, F.C., Feio, M.J., Ferreira, M.T., 2011. Choosing the best method for stream bioassessment using macrophyte communities: Indices and predictive models. Ecol. Ind. 11, 379–388. https://doi.org/10.1016/j.ecolind.2010.06.006.
Almeida, S.F.P., Feio, M.J., 2012. DIATMOD: diatom predictive model for quality assessment of Portuguese running waters. Hydrobiologia 695, 185–197. https://doi.org/10.1007/s10750-012-1110-4.
Bahls, L., 1993. Periphyton bioassessment methods for Montana streams. Montana Water Quality Bureau, Department of Health and Environmental Science, Helena, USA.
Baker, M.E., King, R.S., 2010. A new method for detecting and interpreting biodiversity and ecological community thresholds. Methods Ecol. Evol. 1, 25–37. https://doi.org/10.1111/j.2041-210X.2009.00007.x.
Blocksom, K.A., 2003. A Performance Comparison of Metric Scoring Methods for a Multimetric Index for Mid-Atlantic Highlands Streams. Environ. Manage. 31, 0670–0682. https://doi.org/10.1007/s00267-002-2949-3.
Cáceres, M.D., Legendre, P., 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90 (12), 3566–3574. https://doi.org/10.1890/08-1823.1.
Calsbeek, R., Thompson, J.N., Richardson, J.E., 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Mol. Ecol. 12, 1021–1029.

Cao, Y., Hawkins, C.P., Olson, J., Kosterman, M.A., 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. J. North Am. Benthol. Soc. 26, 566–585. https://doi.org/10.1899/06-078.1.
Carlisle, D.M., Hawkins, C.P., Meador, M.R., Potapova, M., Falcone, J., 2008. Biological assessments of Appalachian streams based on predictive models for fish, macroinvertebrate, and diatom assemblages. J. North Am. Benthol. Soc. 27, 22. https://doi.org/10.1899/06-081.1.
Cauvy-Fraunié, S., Espinosa, R., Andino, P., Jacobsen, D., Dangles, O., 2015. Invertebrate Metacommunity Structure and Dynamics in an Andean Glacial Stream Network Facing Climate Change. PLoS One 10. https://doi.org/10.1371/journal.pone.0136793.
Cemagref, 1982. Étude des méthodes biologiques quantitative d'appréciation de la qualité des eaux. Bassin Rhône-Méditerranée-Corse. Centre National du Machinisme Agricole, du Génie rural, des Eaux et des Forêts, Lyon, France.
Chen, K., Sun, D., Rajper, A.R., Mulatibieke, M., Hughes, R.M., Pan, Y., Tayibazaer, A., Chen, Q., Wang, B., 2019. Concordance in biological condition and biodiversity between diatom and macroinvertebrate assemblages in Chinese arid-zone streams. Hydrobiologia 829, 245–263. https://doi.org/10.1007/s10750-018-3836-0.
Chessman, B., Growns, I., Currey, J., Plunkett-Cole, N., 1999. Predicting diatom communities at the genus level for the rapid biological assessment of rivers. Freshw. Biol. 41, 317–331. https://doi.org/10.1046/j.1365-2427.1999.00433.x.
Cordier, T., Forster, D., Dufresne, Y., Martins, C.I.M., Stoeck, T., Pawlowski, J., 2018. Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. Mol. Ecol. Resour. 18, 1381–1391. https://doi.org/10.1111/1755-0998.12926.
Coste, M., Boutry, S., Tison-Rosebery, J., Delmas, F., 2009. Improvements of the Biological Diatom Index (BDI): Description and efficiency of the new version (BDI-2006). Ecol. Indic. 9, 621–650. https://doi.org/10.1016/j.ecolind.2008.06.003.
Darling, J.A., Mahon, A.R., 2011. From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environ. Res. Invasive Species 111, 978–988. https://doi.org/10.1016/j.envres.2011.02.001.
Feio, M.J., Aguiar, F.C., Almeida, S.F.P., Ferreira, M.T., 2012. AQUAFLORA: A predictive model based on diatoms and macrophytes for streams water quality assessment. Ecol. Ind. 18, 586–598. https://doi.org/10.1016/j.ecolind.2012.01.016.
Feio, M.J., Almeida, S.F.P., Craveiro, S.C., Calado, A.J., 2009. A comparison between biotic indices and predictive models in stream water quality assessment based on benthic diatom communities. Ecol. Ind. 9, 497–507. https://doi.org/10.1016/j.ecolind.2008.07.001.
Feio, M.J., Serra, S.R.Q., Mortágua, A., Bouchez, A., Rimet, F., Vasselon, V., Almeida, S.F.P., 2020. A taxonomy-free approach based on machine learning to assess the quality of rivers with diatoms. Sci. Total Environ. 722, 137900. https://doi.org/10.1016/j.scitotenv.2020.137900.
Fetscher, A.E., Busse, L., Ode, P.R., 2009. Standard operating procedures for collecting stream algae samples and associated physical habitat and chemical data for ambient bioassessments in California. Calif. State Water Resour. Control Board Surf. Water Ambient Monit, Program SWAMP Bioassessment SOP.
Fetscher, A.E., Stancheva, R., Kociolek, J.P., Sheath, R.G., Stein, E.D., Mazor, R.D., Ode, P.R., Busse, L.B., 2014. Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination. J. Appl. Phycol. 26, 433–450. https://doi.org/10.1007/s10811-013-0088-2.
Fetscher, A.E., Sutula, M.A., Busse, L.B., Stein, E.D., 2013. Condition of California perennial wadeable streams based on algal indicators. Surf. Water Ambient Monit. Program Tech. Rep.
Forbes, A.E., Chase, J.M., 2002. The role of habitat connectivity and landscape geometry in experimental zooplankton metacommunities. Oikos 96, 433–440. https://doi.org/10.1034/j.1600-0706.2002.960305.x.
Hämäläinen, H., Aroviita, J., Jyväsjärvi, J., Kärkkäinen, S., 2018. Dangerous relationships: biases in freshwater bioassessment based on observed to expected ratios. Ecol. Appl. 28, 1260–1272. https://doi.org/10.1002/eap.1725.
Hawkins, C.P., Cao, Y., Roper, B., 2010a. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshw. Biol. 55, 1066–1085. https://doi.org/10.1111/j.1365-2427.2009.02357.x.
Hawkins, C.P., Norris, R.H., Hogue, J.N., Feminella, J.W., 2000. Development and Evaluation of Predictive Models for Measuring the Biological Integrity of Streams. Ecol. Appl. 10, 1456–1477. https://doi.org/10.1890/1051-0761(2000)010[1456:DAEOPM]2.0.CO;2.
Hawkins, C.P., Olson, J.R., Hill, R.A., 2010b. The reference condition: predicting benchmarks for ecological and water-quality assessments. J. North Am. Benthol. Soc. 29, 312–343. https://doi.org/10.1899/09-092.1.
Hering, D., Johnson, R.K., Kramm, S., Schmutz, S., Szoszkiewicz, K., Verdonschot, P.F.M., 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshw. Biol. 51, 1757–1785. https://doi.org/10.1111/j.1365-2427.2006.01610.x.
Johnson, R.K., Hering, D., 2009. Response of taxonomic groups in streams to gradients in resource and habitat characteristics. J. Appl. Ecol. 46, 175–186. https://doi.org/10.1111/j.1365-2664.2008.01583.x.
Karr, J.R., Chu, E.W., 2000. Sustaining living rivers. Hydrobiologia 422, 1–14. https://doi.org/10.1023/A:1017097611303.
Kaufmann, P.R., Levine, P., Peck, D.V., Robison, E.G., Seeliger, C., 1999. Quantifying physical habitat in wadeable streams. USEPA [National Health and Environmental Effects Research Laboratory, Western Ecology Division].
Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. J. Stat. Softw. 28, 1–26. https://doi.org/10.18637/jss.v028.i05.
Lamb, M.A., Lowe, R.L., 1987. Effects of Current Velocity on the Physical Structuring of. Diatom (Bacillariophyceae) Communities.
Lavoie, I., Campeau, S., Darchambeau, F., Cabana, G., Dillon, P.J., 2008. Are diatoms good integrators of temporal variability in stream water quality? Freshw. Biol. 53, 827–841. https://doi.org/10.1111/j.1365-2427.2007.01935.x.
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.
Lunde, K.B., Resh, V.H., 2012. Development and validation of a macroinvertebrate index of biotic integrity (IBI) for assessing urban impacts to Northern California freshwater wetlands. Environ. Monit. Assess. 184, 3653–3674. https://doi.org/10.1007/s10661-011-2214-4.
Martins, I., Macedo, D.R., Hughes, R.M., Callisto, M., 2020. Are multiple multimetric indices effective for assessing ecological condition in tropical basins? Ecol. Indic. 110, 105953. https://doi.org/10.1016/j.ecolind.2019.105953.
Matthiessen, B., Mielke, E., Sommer, U., 2010. Dispersal decreases diversity in heterogeneous metacommunities by enhancing regional competition. Ecology 91, 2022–2033. https://doi.org/10.1890/09-1395.1.
Mazor, R.D., Rehn, A.C., Ode, P.R., Engeln, M., Schiff, K.C., Stein, E.D., Gillett, D.J., Herbst, D.B., Hawkins, C.P., 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshw. Sci. 35, 249–271. https://doi.org/10.1086/684130.
Mazor, R.D., Reynoldson, T.B., Rosenberg, D.M., Resh, V.H., 2006. Effects of biotic assemblage, classification, and assessment method on bioassessment performance. Can. J. Fish. Aquat. Sci. 63, 394–411. https://doi.org/10.1139/f05-222.
Mendes, T., Calapez, A.R., Elias, C.L., Almeida, S.F.P., Feio, M.J., 2014. Comparing alternatives for combining invertebrate and diatom assessment in stream quality classification. Mar. Freshw. Res.
Moss, D., Furse, M.T., Wright, J.F., Armitage, P.D., 1987. The prediction of the macroinvertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshw. Biol. 17, 41–52. https://doi.org/10.1111/j.1365-2427.1987.tb01027.x.
Ode, P.R., Fetscher, A.E., Busse, L.B., 2016a. Standard Operating Procedures (SOP) for the Collection of Field Data for Bioassessments of California Wadeable Streams: Benthic Macroinvertebrates, Algae, and Physical Habitat (No. Technical Report 835). State Water Resources Control Board Surface Water Ambient Monitoring Program. Sacramento, CA.
Ode, P.R., Kincaid, T.M., Fleming, T., Rehn, A.C., 2011. Ecological condition assessments of California's perennial wadeable streams: highlights from the Surface Water Ambient Monitoring Program's Perennial Streams Assessment (PSA) (2000-2007). A collaboration between the State Water Resources Control Board's Non-Point Source Pollution Control Program (NPS Program), Surface Water Ambient Monitoring Program (SWAMP), California Department of Fish and Game Aquatic Bioassessment Laboratory, and the U.S. Environmental Protection Agency.
Ode, P.R., Rehn, A.C., Mazor, R.D., Schiff, K.C., Stein, E.D., May, J.T., Brown, L.R., Herbst, D.B., Gillett, D., Lunde, K., Hawkins, C.P., 2016. Evaluating the adequacy of a reference-site pool for ecological assessments in environmentally complex regions. Freshw. Sci. 35, 237–248. https://doi.org/10.1086/684003.
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2017. vegan. Community Ecology Package.
Olson, J.R., Hawkins, C.P., 2012. Predicting natural base-flow stream water chemistry in the western United States. Water Resour. Res. 48. https://doi.org/10.1029/2011WR011088.
Omernik, J.M., 1987. Ecoregions of the Conterminous United States. Ann. Assoc. Am. Geogr. 77, 118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x.
Pardo, I., Delgado, C., Abraín, R., Gómez-Rodríguez, C., García-Roselló, E., García, L., Reynoldson, T.B., 2018. A predictive diatom-based model to assess the ecological status of streams and rivers of Northern Spain. Ecol. Ind. 90, 519–528. https://doi.org/10.1016/j.ecolind.2018.03.042.
Pardo, I., Gómez-Rodríguez, C., Abraín, R., García-Roselló, E., Reynoldson, T.B., 2014. An invertebrate predictive model (NORTI) for streams and rivers: Sensitivity of the model in detecting stress gradients. Ecol. Ind. 45, 51–62. https://doi.org/10.1016/j.ecolind.2014.03.019.
Paul, M., Jessup, B., Brown, L., Carter, J., Cantonati, M., Charles, D.F., Gerritsen, J., Herbst, D.B., Stancheva, R., Howard, J., Isham, B., Lowe, Rex L., Mazor, R., Mendez, P., Ode, P.R., O'Dowd, A., Olson, J.R., Pan, Y., Rhen, A., Spaulding, S., Sutula, M., Theroux, S., 2020. Characterizing benthic macroinvertebrate and algal biological condition gradient models for California wadeable Streams, USA. Ecol. Indic. 117. https://doi.org/10.1016/j.ecolind.2020.106618.
Porter, S.D., Mueller, D.K., Spahr, N.E., Munn, M.D., Dubrovsky, N.M., 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshw. Biol. 53, 1036–1054. https://doi.org/10.1111/j.1365-2427.2007.01951.x.
Potapova, M., Charles, D.F., 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. Ecol. Ind. 7, 48–70. https://doi.org/10.1016/j.ecolind.2005.10.001.
R Core Team, 2013. R: A Language and Environment for Statistical Computing [WWW Document]. URL http://www.R-project.org/. (Accessed 3.3.19).
Rehn, A.C., 2016. Using Multiple Biological and Habitat Condition Indices for Bioassessment of California Streams (Technical memo No. SWAMP-TM-SB-2016-0003). State Water Resources Control Board Surface Water Ambient Monitoring Program, Sacramento, CA.
Reynoldson, T.B., Norris, R.H., Resh, V.H., Day, K.E., Rosenberg, D.M., 1997. The Reference Condition: A Comparison of Multimetric and Multivariate Approaches to Assess Water-Quality Impairment Using Benthic Macroinvertebrates. J. North Am. Benthol. Soc. 16, 833–852. https://doi.org/10.2307/1468175.
Ritz, C., 2010. Effectiveness of RIVPACS predictive models to evaluate diatom response to nutrient stress in coastal California streams (Master of Science). California State University, Monterey Bay.

Ruaro, R., Gubiani, É.A., Hughes, R.M., Mormul, R.P., 2020. Global trends and challenges in multimetric indices of biological condition. Ecol. Ind. 110, 105862. https://doi.org/10.1016/j.ecolind.2019.105862.
Sabater, S., 2006. Hydrology, light and the use of organic and inorganic materials as structuring factors of biological communities in Mediterranean streams. Limnetica 25, 335–348. https://doi.org/10.23818/limn.25.23.
Schneider, S.C., Lawniczak, A.E., Picińska-Faltynowicz, J., Szoszkiewicz, K., 2012. Do macrophytes, diatoms and non-diatom benthic algae give redundant information? Results from a case study in Poland. Limnologica 42, 204–211. https://doi.org/10.1016/j.limno.2011.12.001.
Simpson, J.C., Norris, R.H., 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs. Assess. Biol. Qual. Fresh Waters RIVPACS Tech. Proc. Int. Workshop Held Oxf. UK 16-18 Sept. 1997 125–142.
Sleeter, B.M., Wilson, T.S., Soulard, C.E., Liu, J., 2011. Estimation of late twentieth century land-cover change in California. Environ. Monit. Assess. 173, 251–266. https://doi.org/10.1007/s10661-010-1385-8.
Soininen, J., 2005. Assessing the current related heterogeneity and diversity patterns of benthic diatom communities in a turbid and a clear water river. Aquat. Ecol. 38, 495–501. https://doi.org/10.1007/s10452-005-4089-3.
Sonneman, J.A., Walsh, C.J., Breen, P.F., Sharpe, A.K., 2001. Effects of urbanization on streams of the Melbourne region, Victoria, Australia. II. Benthic diatom communities. Freshw. Biol. 46, 553–565. https://doi.org/10.1046/j.1365-2427.2001.00689.x.
Sork, V.L., Gugger, P.F., Chen, J.-M., Werth, S., 2016. Evolutionary lessons from California plant phylogeography. Proc. Natl. Acad. Sci. 113, 8064–8071. https://doi.org/10.1073/pnas.1602675113.
Spaulding, S., Lubinski, D., Potapova, M., 2010. Diatoms of the United States [WWW Document]. Diatoms U. S. URL http://westerndiatoms.colorado.edu. (Accessed 10.10.17).
Stancheva, R., Busse, L., Kociolek, J.P., Sheath, R., 2015. Standard Operating Procedures for Laboratory Processing and Identification of Stream Algae in California (No. SWAMP-SOP-2015-0003). California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 0003.
Stancheva, R., Sheath, R.G., 2016. Benthic soft-bodied algae as bioindicators of stream water quality. Knowl. Manage. Aquat. Ecosyst. 15. https://doi.org/10.1051/kmae/2016002.
State Water Resources Control Board, 2017. 2014 and 2016 California Integrated Report, Clean Water Act Sections 303(d) and 305(b).
Stevenson, R.J., Zalack, J.T., Wolin, J., 2013. A multimetric index of lake diatom condition based on surface-sediment assemblages. Freshw. Sci. 32, 1005–1025. https://doi.org/10.1899/12-183.1.
Stoddard, J.L., Herlihy, A.T., Peck, D.V., Hughes, R.M., Whittier, T.R., Tarquinio, E., 2008. A process for creating multimetric indices for large-scale aquatic surveys. J. North Am. Benthol. Soc. 27, 878–891. https://doi.org/10.1899/08-053.1.
Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., Norris, R.H., 2006. Setting Expectations for the Ecological Condition of Streams: The Concept of Reference Condition. Ecol. Appl. 16, 1267–1276. https://doi.org/10.1890/1051-0761(2006)016[1267:SEFTEC]2.0.CO;2.
Surface Water Ambient Monitoring Program, n.d. SWAMP Data Checker Look Up Lists [WWW Document]. SWAMP Data Checker Look Lists. URL https://swamp.waterboards.ca.gov/swamp_checker/LookUpLists.aspx. (Accessed 4.1.20).
U.S. Environmental Protection Agency, 2016. National Aquatic Resource Surveys. National Rivers and Streams Assessment 2008-2009. [WWW Document]. URL https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys. (Accessed 4.1.20).
Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G.H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., Peroux, T., Crivelli, A.J., Olivier, A., Acqueberge, M., Brun, M.L., Møller, P.R., Willerslev, E., Dejean, T., 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Mol. Ecol. 25, 929–942. https://doi.org/10.1111/mec.13428.
van Dam, H., Mertens, A., Sinkeldam, J., 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. Netherland J. Aquat. Ecol. 28, 117–133. https://doi.org/10.1007/BF02334251.
Van Sickle, J., Hawkins, C.P., Larsen, D.P., Herlihy, A.T., 2005. A null model for the expected macroinvertebrate assemblage in streams. J. North Am. Benthol. Soc. 24, 178–191. https://doi.org/10.1899/0887-3593(2005)024<0178:ANMFTE>2.0.CO;2.
Van Sickle, J., Larsen, D.P., Hawkins, C.P., 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. J. North Am. Benthol. Soc. 26, 319–331.
Vander Laan, J.J., Hawkins, C.P., 2014. Enhancing the performance and interpretation of freshwater biological indices: An application in arid zone streams. Ecol. Ind. 36, 470–482. https://doi.org/10.1016/j.ecolind.2013.09.006.
Veraart, A.J., Romaní, A.M., Tornés, E., Sabater, S., 2008. Algal Response to Nutrient Enrichment in Forested Oligotrophic Stream1. J. Phycol. 44, 564–572. https://doi.org/10.1111/j.1529-8817.2008.00503.x.
Voss, K.A., Pohlman, A., Viswanathan, S., Gibson, D., Purohit, J., 2012. A study of the effect of physical and chemical stressors on biological integrity within the San Diego hydrologic region. Environ. Monit. Assess. 184, 1603–1616. https://doi.org/10.1007/s10661-011-2064-0.
Whitton, B.A., 2012. Changing approaches to monitoring during the period of the 'Use of Algae for Monitoring Rivers' symposia. Hydrobiologia 695, 7–16. https://doi.org/10.1007/s10750-012-1121-1.
Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, (2nd edition). Chapman and Hall/CRC.
Wright, J.F., 2000. Assessing the biological quality of fresh waters: RIVPACS and other techniques ; invited contributions from an international workshop held in Oxford, UK on 16-18 September 1997 ... FBA, Ambleside.
Wright, J.F., 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Aust. J. Ecol. 20, 181–197. https://doi.org/10.1111/j.1442-9993.1995.tb00531.x.
Wright, J.F., Furse, M.T., Armitage, P.D., 1993. RIVPACS – a technique for evaluating the biological quality of rivers in the U.K. Eur. Water Pollut. Control 3, 15–25.