SCCWRP #1062

Accepted: 24 October 2017

DOI: 10.1111/maec.12471

ORIGINAL ARTICLE

WILEY marine ecology

Development of a biological condition assessment index for shallow, subtidal rocky reefs in Southern California, USA

Julia H. Coates[1,2] | **Kenneth Schiff**[2] | **Raphael D. Mazor**[2] | **Daniel J. Pondella II**[3] | **Rebecca Schaffner**[2] | **Elizabeth Whiteman**[1]

[1]California Ocean Science Trust, Oakland, CA, USA

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

[3]Vantuna Research Group, Occidental College, Los Angeles, CA, USA

Correspondence
Julia H. Coates, California Ocean Science California Department of Fish and Wildlife, Santa Barbara, CA, USA.
Email: julia.coates@wildlife.ca.gov

Abstract

Quantitative biological assessment indices overcome many of the challenges faced when trying to convey the status or trends of complex biological communities with large natural variability, particularly when attempting to evaluate the impacts from human influences. In this paper, we developed a biological condition index for shallow (<30 m) rocky reefs of the Southern California Bight, evaluated its ability to distinguish healthy from stressed sites, and then applied the index by examining relative correlations with fishing and water quality as ecosystem stressors. We utilized a multivariate, predictive index based on the ratio of observed-to-expected taxa (O/E). O/E indices are relatively common in freshwater environments, but rarely utilized in marine systems and never before applied to rocky reefs. Based on expectations drawn from region-wide reference reefs with the least fishing or water quality stress, the O/E index predicts expected taxa at a new site based on environmental factors such as sea surface temperature, reef area, and slope, among others. The observed taxa at that site are then compared to the predicted taxa to generate index scores; values near unity indicate intact, reference-like communities. Overall, the accuracy of the index was high, with minimal bias, and precision exceeded the performance of an index based on null models (i.e., indices that did not account for natural gradients). The mean index score was significantly higher among reference sites than stressed sites; however, sensitivity was low, as 84% of stressed sites had scores within the range of reference sites. Ultimately, fishing pressure was more correlated with changes in index scores from the non-reference data set than was water quality pressure. This study demonstrates that a multivariate predictive index is feasible in rocky reef assessment and illuminates additional investigative work to continue to advance index development.

KEYWORDS

bioassessment, ecosystem-based management, predictive index, rocky reef, water quality

1 | INTRODUCTION

The concept of measuring the health of whole ocean ecosystems is receiving increasing attention (Halpern et al., 2012, 2014), yet the best biological characteristics for defining healthy marine ecosystems remain uncertain. The uncertainty arises from the complexity of ecosystem responses to stressors, which are not always monotonic but instead often produce a complex mix of positive and negative effects on the number and abundance of species (Pearson & Rosenberg, 1978). For example, complex trophic linkages can produce unexpected gains within one fishing sector resulting from overexploitation of another fishing sector (Andersen & Pedersen, 2010; Steneck & Wahle, 2013). Additionally, increased runoff from agriculture can serve as a useful nutrient subsidy to some nearshore organisms (Gorman, Russell, & Connell, 2009; Page, Reed, Brzezinski, Melack, & Dugan, 2008; Russell, Thompson, Falkenberg, & Connell, 2009), but smother others through sedimentation (Airoldi, 2003; Balata, Piazzi, & Cinelli, 2007; Connell, 2005; Irving & Connell, 2002; Schiel, Wood, Dunmore, & Taylor, 2006; Walker, 2007). Determining ecosystem health in the context of these numerous and sometimes compound influences is not straightforward and may require a multivariate bioassessment approach that encompasses many ecosystem components.

In this paper, we developed a predictive multivariate index of ecosystem health for shallow rocky reefs of the Southern California Bight (SCB). This highly productive and diverse ecosystem supports intense consumptive and non-consumptive use. To address management concerns regarding consumptive use, the state promulgated the Marine Life Protection Act (CDFW 2014) in 1999, producing a spatially integrated network of marine protected areas (MPAs). Water quality stress is a primary concern among human impacts in Southern California because of its dense coastal human population centers. For example, trace metals measured in the water column off the coast of Los Angeles remain orders of magnitude greater than background crustal levels (Smail, Webb, Franks, Bruland, & Sañudo-Wilhelmy, 2012). The efficacy of the MPA network will need to be assessed in the context of variable natural gradients and stressor gradients. We aimed to perform a novel application of an index technique that has been well tested in freshwater systems, to this marine system. Our focus was not on MPAs specifically, but rather on testing an index technique that might allow for future assessment of ecosystem status across sites and time. Therefore, our study sites were not restricted to MPAs. However, a successful index might be applied to MPA management by providing a means to (i) track changes in ecosystem health following implementation and (ii) assess the relative importance of not only extraction of living resources but also other influences on ecosystem health, anthropogenic or not. Our specific objectives were to (i) develop a multivariate bioassessment index, (ii) assess its performance and ability to distinguish healthy from stressed sites, and (iii) apply the index in an effort to distinguish the relative effects of fishing and water quality as ecosystem stressors.

Historically, California has relied on changes in abundance and size structure of a series of target species (CalOST and CDFW 2013; CDFW 2008) to judge changes in ecosystem status over time following MPA implementation. This approach requires contentious judgments among professionals, who must value certain species over others for assessment purposes, and there is currently little consensus on which taxa are the most effective indicators of a healthy marine ecosystem. Moreover, protection resulting from MPAs is expected to differentially affect species occurring within their boundaries depending on their life-history characteristics and dispersal capabilities. Thus, single species assessments of MPA impacts may be variable and not provide a complete picture of ecosystem status. Finally, the strength of interactions among species may change under MPA protection and therefore unexpected ecosystem changes can occur that are not detected in single species assessments.

Many assessments of MPA effectiveness address ecosystem degradation due to consumptive use, but do not account for non-consumptive uses or stressors such as water quality. The potential for non-consumptive impacts to marine resources is unclear. Many studies have demonstrated the effect of sedimentation on settlement and survival of rocky reef organisms (Airoldi, 2003; Balata et al., 2007; Connell, 2005; Irving & Connell, 2002; Schiel et al., 2006; Walker, 2007). However, ecosystem impacts due to storm water runoff (Ghedini, Klein, & Coleman, 2011) and sewage discharge (O'Connor, 2013; Tegner et al., 1995) have not always been detected. Improved methods, including multivariate-based assessment tools for understanding impacts due to poor water quality may help managers determine whether substantive impacts that warrant management action are occurring.

Predictive bioassessment indices are increasingly used by the water quality management community as tools for assessing the ecosystem-level impacts of anthropogenic pollutants in primarily freshwater or freshwater-influenced ecosystems such as streams, wetlands and estuaries (Cao & Hawkins, 2011; Yoder & Rakin, 1998). These indices typically establish an expected biological community composition within reference sites defined by a relative absence of anthropogenic stress (Hawkins, 2006; Stoddard, Larsen, Hawkins, Johnson, & Norris, 2006). Predictive models describe variability in biological communities due solely to natural environmental characteristics, such as climate and abiotic habitat features, at these reference sites. Once natural environmental variability is taken into account with the predictive model, deviations in index scores among non-reference sites can then be associated with gradients in anthropogenic influences, thus quantifying the direction and magnitude of those influences on ecosystems (Vander Laan, Hawkins, Olson, & Hill, 2013). Similar indices have rarely been developed for marine ecosystems (Smith et al., 2001), perhaps because of a long tradition of single-species focus (Guerry, 2005; Mace, 2004), a lack of spatial management (Douvere, 2008), or disagreement on the most important ecosystem components (de Jonge, Pinto, & Turner, 2012; Ward, 2014).

2 | MATERIAL AND METHODS

We constructed an observed-to-expected (O/E) index to describe the condition of shallow-rock reef sites with the aim of using index scores to assess the relative impacts of fishing pressure and water quality on this ecosystem. The study habitat is defined as subtidal, generally kelp forested, areas of hard substratum ranging from appoximately 5–30 m depth. The geographic extent of our study covered the SCB, which ranges from Point Conception to the US/Mexico border including the Channel Islands. Approximately 49,000 hectares of shallow rocky reef area is distributed along the mainland and island coastlines in this region (Pondella et al., 2015).

2.1 | Aggregation of data

2.1.1 | Biological data

Rocky reef biological data was aggregated from three separate monitoring programs conducted in 4 years: (i) the California Department of Fish and Wildlife's Cooperative Research and Assessment of

Nearshore Ecosystems (CRANE) in 2004, (ii) the Southern California Coastal Water Research Project's Bight Regional Monitoring Program in 2008, and (iii) California Ocean Science Trust's South Coast Marine Protected Area Baseline Program in 2011 and 2012. These programs use nearly identical protocols to each other and to those used by the Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO; http://www.piscoweb.org/research/science-by-discipline/ecosystem-monitoring/kelp-forest-monitoring/subtidal-sampling-protocol) and described by Gillett et al. (2012). Briefly, three types of transect protocols were used to collect information on fish, invertebrates, benthic algae and physical habitat characteristics. Transects were placed within four reef sections (inner, middle, outer, deep). Fish, as well as mobile macroinvertebrates and algae (swath), were counted along replicate 30×4-m bottom transects within those sections. Additional fish transects were performed midwater and immediately below the kelp canopy. The uniform point contact (UPC) method assessed percent cover of invertebrates and algae that are colonial or blanketing and physical habitat characteristics. Most organisms were identified to species except approximately half of the UPC taxonomic units (e.g., red algal turf, colonial tunicate) and a few taxonomic units in the fish and swath data sets (e.g., juvenile rockfish were combined). Abundances on replicate transects for each survey method (fish, swath, UPC) were averaged within each site and sample year then converted to a presence/absence matrix. Therefore, a biological sample reflects a unique site–year combination. Monitoring was conducted at 140 unique reefs during between 1 and 4 sample years, resulting in 299 biological samples.

2.1.2 | Stressor data

We characterized two stressor gradients: water quality and fishing pressure. Each was described by an index that aggregated multiple years of local-scale stressor data into an SCB regional-scale map. We used a risk-based approach to create the water quality index (WQI) detailed within Schaffner, Steinberg, and Schiff (2015). The risk was estimated as a function of pollutant load, magnitude, and duration of exposure. Pollutant loading was estimated for the SCB's two largest sources, publicly owned treatment works (POTWs) and rivers (Lyon & Stein, 2009). Spatial distribution of exposure duration was acquired from multiple sources that utilize advances in plume detection technology including high frequency radar and optical sensors for colored dissolved organic matter. Average annual loads of dissolved inorganic nitrogen (DIN), total suspended solids (TSS), and copper were calculated for 20 rivers based on 11 years of data (2000–2010; Howard et al., 2014) and four POTWs using 7 years of data (2003–2009; Lyon & Stein, 2009). Due to the absence of rivers and POTWs at the Channel Islands and because mainland plumes rarely if ever extend that far offshore, we assumed an absence of water quality pressure (WQI = 0) in that region of the SCB. Our WQI varied at spatial scales (250 m) finer than the biological sites (individual reefs from 6 to 5,000 hectares). The WQI incorporated data over a similar time frame to the biological data (2000–2010).

Data to describe fishing pressure were taken from a synthetic fishing pressure index derived for the SCB (Pondella et al., 2016). Development of the fishing index involved additively combining California Department of Fish and Wildlife commerial fisheries data with commercial passenger fishing vessel (CPFV) recreational data after converting both to pounds taken. Both fisheries report take with reference to a spatial grid or "fishery blocks" approximately 10 min square. The data were filtered to include only rocky reef associated species and trimmed to the time period 1980–2009. Some alterations were made for the purpose of the present study. These were weighting take within fishery blocks by the length of ocean-facing coastline within that block, removal of urchins from the aggregate index, and trimming the data to biolocial transect sample dates. Fishery take to be associated with a given biological transect sample reflected the most recent possible fishing data, not extending past the transect sample date (e.g., transect data collected in 2004 were associated with fishing data up to and not past 2004, transect data collected in 2012 were associated with fishing data up to 2009). The spatial scale of commercial fishing blocks is generally greater than that of individual reefs and therefore the index may not be able to distinguish fine variation in fishermen's behavior. While most of the MPAs in Southern California were established in 2012, after collection of the data used for this study, MPAs at the Northern Channel Islands were established in 2003. Reduced fishing pressure in those areas would be reflected in the commercial and CPFV harvest data by block relative to the amount of MPA overlap with the block grid.

2.2 | Designation of reference and non-reference sites

Reference sites were defined by the relative absence of stress, as measured by our two stressor indices (Figure 1a). Reference sites were defined as sites with low levels of stress from water quality and fishing pressure, consistent with the "best available" definition in Stoddard et al. (2006). This designation represents an objective reference definition. Best professional judgements of high-quality reef sites based on biological attributes like high species diversity were not considered, as they may perpetuate circular reasoning in developing a biological index (Ode et al., 2016; Yates & Bailey, 2010). The selected reference sites may be impacted by additional stressors not included in the two indices and therefore not truly represent high-quality sites. However, some of these unmeasured stressors may be correlated with those that were measured. For example, data on recreational fishing from private vessels are not included in the fishing pressure index because data are lacking but this type of fishing effort is likely to be correlated with CPFV effort.

Stress due to fishing was described in three ways: take within the 5 most recent years, within the 10 most recent years, and between 1980 and 2012. Stress due to water quality impacts was desribed using the synthetic WQI as well as each of its components (DIN, TSS, copper) (Schaffner et al., 2015). Criteria for identifying reference sites were set at fisheries take below the 30th percentile of take among all sites, and an absence of water quality stress (WQI = 0). Our goal was to have

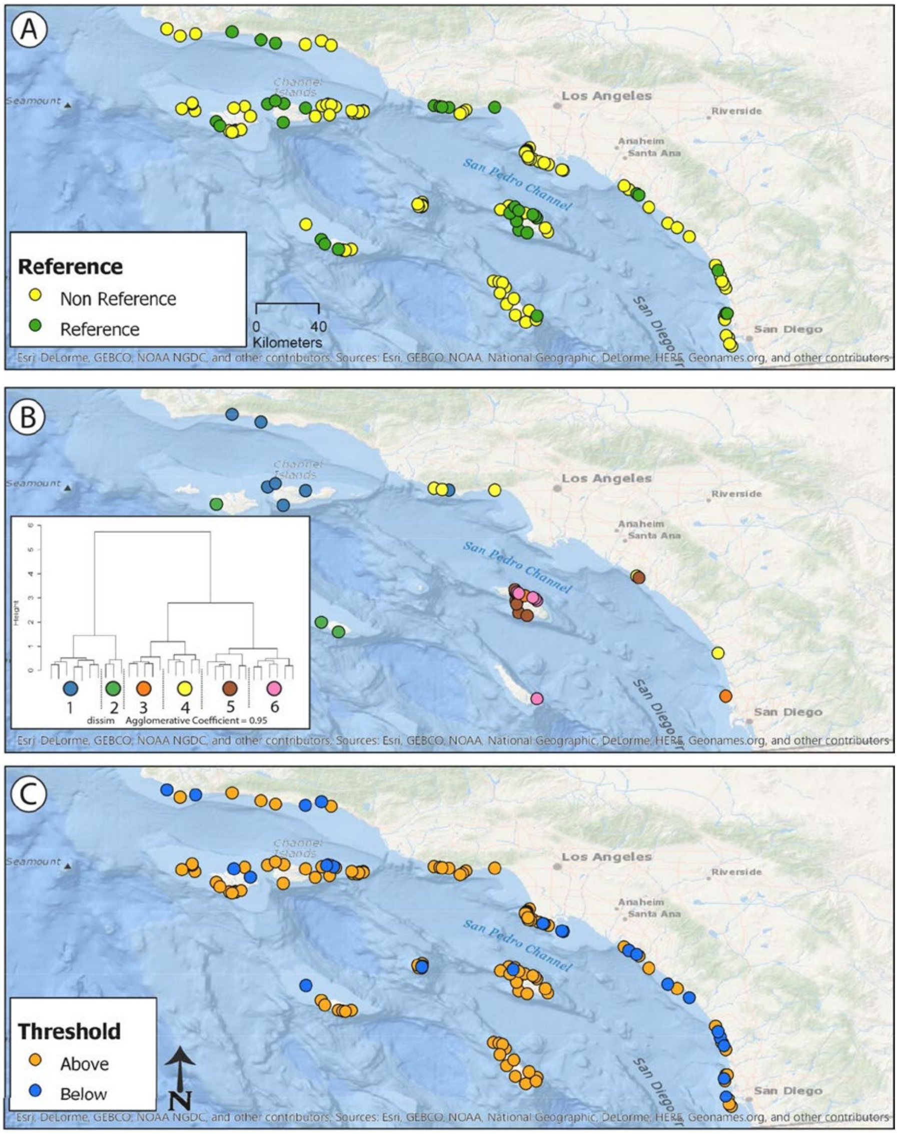

FIGURE 1 (a) Locations of reference sites experiencing little anthropogenic stress (green) and non-reference sites experiencing stress (yellow). (b) Reference sites falling within the six biotic clusters. Inset shows the dendrogram based on fish + swath community data among selected sample years at reference calibration sites. (c) Sites with index scores at or above those at reference sites (above threshold, orange) and below those at reference sites (below threshold, blue)

a sufficient number and geographic distribution of reference sites to reflect the full spectrum of natural habitat gradients, while minimizing excess in the system. To verify that reference sites could represent the environmetal gradients at stressed sites, we visually compared probability density plots of the environmental gradients represented within reference sites that met these thresholds and each of the three fisheries take time periods with environmental gradients within non-reference sites (Ode et al., 2016).

Reference sites were randomly divided into calibration (80%) and validation (20%) sets. Models were trained on a single sample year at each reference calibration site, chosen at random for sites with multiple years of samples (later referred to as selected samples). Index scores were generated for all sample years for reference calibration, reference validation, and non-reference sites. Some model evaluations were performed on selected sample years only.

2.2.1 | Habitat data

Each sample site was attributed with a suite of variables describing its habitat characteristics generally not associated with anthropogenic influence (Table 1). These were factors expected to influence the biological community on rocky reefs and inform the expected taxa of the O/E index. Some of these variables were expected to vary spatially, temporally or both. Many were measured during the transect surveys including the proportional area of bare rock, bare sand, bedrock, sand, sediment, boulders, and shell hash. Bedrock and sand substrate categories may or may not be covered by biota, while bare rock and bare sand were separately measured cover categories devoid of biota. Relief was also measured during transect surveys and categorized by the levels 0–0.1 m, 0.1–1 m, 1–2 m and >2 m. Exposure to sand transport or littoral drift was estimated from maps derived by Patsch and

TABLE 1 Habitat variables and temporal or spatial variability

	Variability		
Habitat gradient	**Spatial**	**Temporal**	**Citation**
Substrate (% cover)			
Bare rock	X	X	Pondella et al. (2011)
Bare sand	X	X	Pondella et al. (2011)
Bedrock	X	X	Pondella et al. (2011)
Boulder	X	X	Pondella et al. (2011)
Cobble	X	X	Pondella et al. (2011)
Sand	X	X	Pondella et al. (2011)
Sediment	X	X	Pondella et al. (2011)
Shell harsh	X	X	Pondella et al. (2011)
Relief (m)			
0–0.1	X	X	Pondella et al. (2011)
0.1–1	X	X	Pondella et al. (2011)
1–2	X	X	Pondella et al. (2011)
>2	X	X	Pondella et al. (2011)
Temperature (°C)			
Mean SST		X	
SST	X	X	
SST1	X	X	
SST2	X	X	
Geography & oceanography			
Latitude	X		
Longitude	X		
Island/mainland	X		
Deepest survey	X	X	
Site clustering	X		
Reef area	X		Pondella et al. (2011)
Slope	X		
ENSO		X	Wolter (2014); Wolter and Timlin (1993, 1998)
ENSO1		X	Wolter (2014); Wolter and Timlin (1993, 1998)
ENSO2		X	Wolter (2014); Wolter and Timlin (1993, 1998)
Littoral drift	X		Patsch and Griggs (2006)
Invertebrate settlement 1		X	Schroeter et al. (2012)
Invertebrate settlement 2		X	Schroeter et al. (2012)

ENSO, El Niño Southern Oscillation; SST, sea surface temperature.

Griggs (2006). Sea surface temperatures were averages of measurements taken at 15-day intervals over multiple (Mean SST) or single year periods that were current with biological sampling or representative of 1 or 2 years prior to sampling (SST, SST1, SST2). Monthly El Niño Southern Oscillation (ENSO) index values were averaged within 1-year periods and relationships with mean ENSO values current with biological sampling as well as 1 and 2 years prior to sampling were examined (ENSO, ENSO1, ENSO2). The strength of settlement patterns was described using invertebrate settlement on artificial substrates at six Southern California sites collected by the Santa Barbara Coastal Long Term Ecological Research project with mean values for the sample year and 1 and 2 years prior examined (Settlement, Settlement1, Settlement2; Schroeter, Dixon, Ebert, & Richards, 2012). Values were averaged across all sites within 1-year periods and examined 1 and 2 years prior to biological sampling. The area of rocky reef at each sample site was taken from Pondella et al. (2015). Reef slope was calculated using a geographic information system (GIS) by calculating the slope of a line extending across the spatial extent of the reef and a bathymetric layer. GIS was also used to calculate a metric of reef clustering or proximity to other reefs. Depth was measured as the deepest transect survey depth. Finally, the year of biological sampling, latitude, and longitude were used as habitat variables.

2.3 | Development of the O/E

Degraded biological condition can be described as a loss of expected taxa, and quantified by the O/E ratio (Hawkins, 2006; Wright, 1995). Observed and expected values represent counts of taxa based on presence/absence data. We developed the reef O/E index following methods described by Moss, Furse, Wright, and Armitage (1987) and modified by Mazor et al. (2016). First, we used cluster analysis to identify biologically homogeneous groups of sites within the reference calibration set. Second, we developed a random forest model to relate group membership to the most important habitat variables. Third, this model was used to predict group membership of sites based on their habitat predictors, generating probabilities of taxa presence at those sites (capture probability). Fourth, the overall expected number of taxa (E) at a site was calculated as the sum of model-predicted probabilities of observing each taxon, including only taxa with capture probabilites >0.5. The observed number of taxa at a site (O) is the sum of all actually observed taxa within a potential set of taxa restricted to those with a probability >.5 of ocurring in that cluster group. Previous studies have shown that inclusion of rare species can decrease model performance (Hawkins, Norris, Hogue, & Feminella, 2000; Ostermiller & Hawkins, 2004; Van Sickle, Larsen, & Hawkins, 2007).

Not all habitat variables were retained in the final model. Habitat predictors of reference site group membership were evaluated based on mean decrease in accuracy. The final model was selected based on minimizing the number of predictors while also minimizing the *SD* of O/E scores for reference calibration samples. The final predictors were latitude, longitude, slope, depth, reef area, relief 1–2 m, Mean

SST, SST, SST1, SST2, Settlement1, bare rock, sand, and shell hash. While some variables were significantly correlated, particularly temperature variables, we retained them due to their potential to affect aspects of the ecosystem differently (e.g., temperature in the current year may cause some mobile species to migrate while temperatures in past years may have affected the settlement rates of other species). Recent research has shown that random forest models are robust to inclusion of large sets of variables and that in some cases models with reduced variable sets were less stable in their predictions (Fox et al., 2017).

In order to evaluate the relative improvement in index precision attributable to modeling reference site variability, an index is based on a null model. The null model is created by assigning all reference sites to the same group (i.e., cluster) and therefore the null model does not use habitat predictor variables to predict expected taxa. This allows an estimate of the lower limit of precision attainable by a model. A theoretical upper limit of precision was calculated as the *SD* of replicate samples (SDRS), as described by Van Sickle, Hawkins, Larsen, and Herlihy (2005). Thus, the *SD* of the index may be compared to the *SD* of the null index and the SDRS to evaluate the improvement caused by modeling natural gradients.

Five separate O/E indices were created using each of the three species assemblages (i.e., fish, swath and UPC), fish and swath combined, and all assemblages combined. When combining assemblages equal weight was given to each taxonomic unit regardless of assemblage. Different clusters of reference calibration sites were derived based on each species assemblage. Final random forest models with unique sets of habitat predictors were created for each species assemblage.

2.4 | Index performance evaluation

The performance of the O/E index was assessed with regard to accuracy, precision and responsiveness. Accuracy was evaluated as the mean score of reference sites. Mean scores close to 1 indiciate high accuracy. Accuracy was also evaluated as the number of validation reference sites with scores within 1 *SD* of the reference mean. We assessed the remaining influence of habitat variables (intended to already be accounted for by the modeling process) by regressing O/E scores with each of the habitat variables in the final random forest model at reference sites. Remaining influence of habitat on O/E scores suggests that the modeling process did not successfully normalize sites by their habitat characteristics and the score values cannot be considered pure reflections of stressors. Temporal bias with sample year was assessed using an analysis of variance of scores at reference sites by year. Finally, we used random forest models to examine the pseudo *R*-squared values associated with models relating O/E scores to habitat variables at reference sites; an unbiased index will have low variability associated with natural gradients at reference sites.

Precision was evaluated as the *SD* of scores at reference sites; low *SD* among reference sites indicates high precision (Ode, Hawkins, & Mazor, 2008). These were calculated for reference calibration and validation site sets. We also assessed precision by examining the within-site variability of index scores where sites were visited in multiple years. Responsiveness was evaluated using *t* tests comparing mean index scores at reference and non-reference sites. We assessed responsiveness by regressing index scores against continuously varying stressor data; negative slopes with high *R*-squared values indicate a responsive index.

Indices calibrated for different species assemblages (i.e., fish, swath, UPC, or combinations thereof) were first compared with regard to accuracy and precision as defined by the mean and *SD* of scores. Responsiveness was then evaluated. The best index based on these criteria was then further evaluated.

2.5 | Relative impacts of fishing pressure and water quality

Relationships between index scores and stressor levels were used to assess the relative impacts of stressors on site condition. A threshold to identify impaired biological communities was defined as the lowest reference O/E (fish + swath) score (0.847). Samples above this threshold were considered to be in reference condition and samples below below this threshold were considered degraded or in non-reference condition. We compared the means of stress and habitat variables for samples above and below the threshold using *t* tests. To examine the relative influence of stress and habitat variables on O/E index scores, we used the fishing pressure and water quality indices, along with habitat variables showing larger differences in samples above and below the threshold, as predictor varibles in a multiple regression with O/E (fish + swath) index scores as the response variable. A simplified model was produced using a step-wise procedure.

We examined the impact of stress on individual species in an effort to identify indicator species and their influence on the index. Those taxa with the potential to demonstrate the effects of stressors and that likely had an influence on O/E index scores were identified in two ways: (i) taxa with an average probability of capture among all samples that was greater than 50% and (ii) taxa that were observed at >25% of the survey sites. Random forest models were used to relate the abundance of those taxa to the fishing pressure index, water quality index, and their capture probability at each site. In this analysis, capture probability is a proxy for natural environmental factors influencing species occurrence.

3 | RESULTS

3.1 | Reference site set optimization

We identified a set of 41 reference sites with broad geographic distribution and good representation of the total range of habitat variability among all sites (Figure 1a). Probability density plots for 29 different habitat variables demonstrated agreement in the relative range of natural gradients among the reference and non-reference sample distributions (Figure 2). These 41 reference samples included reefs that scored zero in the water quality index and reefs that scored below the 30th percentile in the fishing pressure index dating back to 1980. To evaluate the potential effect of duration of fishing pressure on the

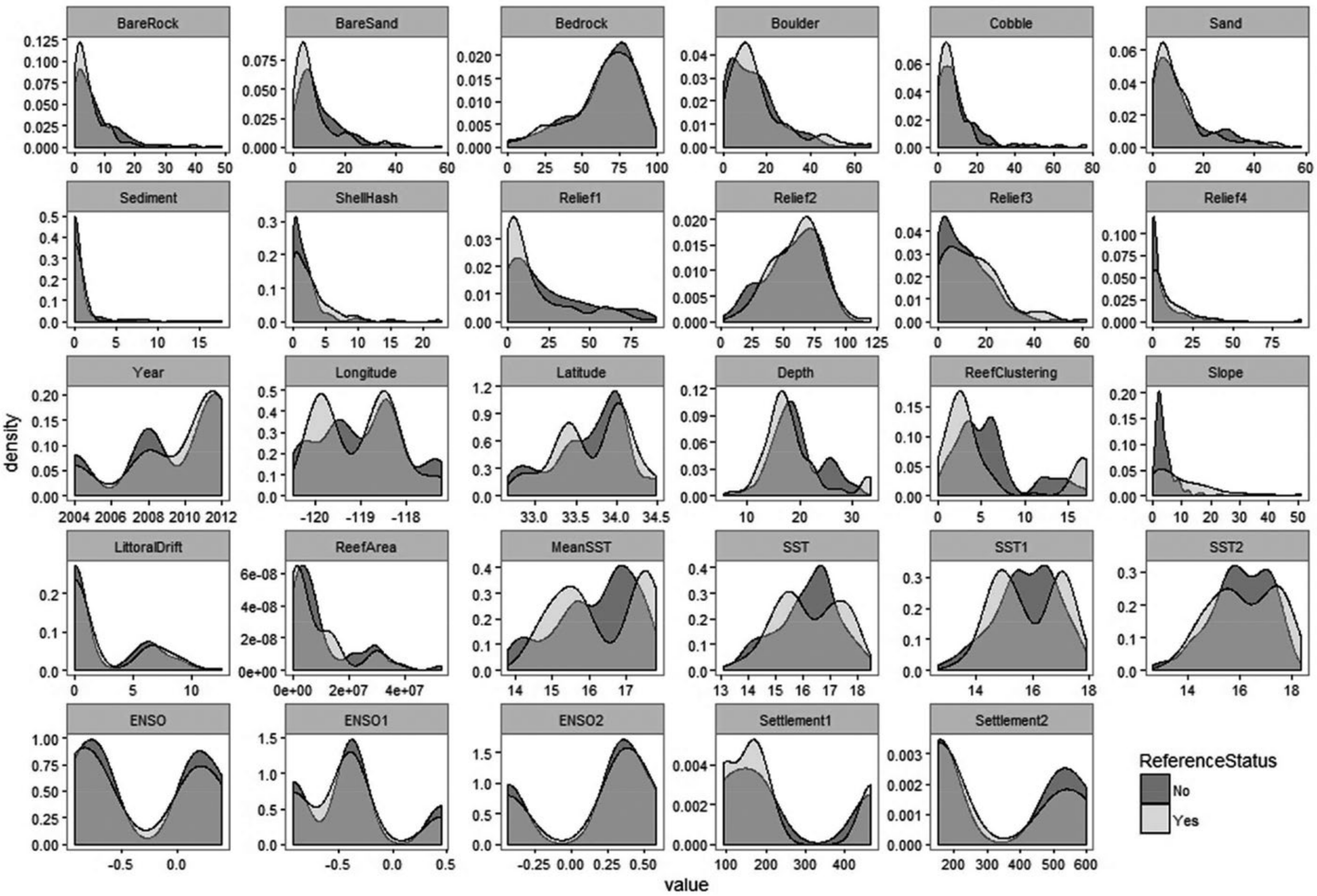

FIGURE 2 Probability density plots of habitat variable distributions among reference and non-reference samples. See Table 1 for descriptions of habitat variables

identities of reference samples, this analysis was re-run by defining fishery stress by the most recent 5 and 10 years of fishing pressure data for each sample (data not shown). Many, but not all, of the same sites were found. However, the reference sample population dating back to 1980 provided the greatest coverage of natural gradients.

Flexible beta cluster analysis produced five dendrograms based on Bray–Curtis dissimilarity in species presence/absence data for each of the five species assemblages. We then visually chose five to six groups within those dendrograms. Sites within the same clusters grouped together geographically with some restricted to either island or mainland sites and others spanning both. For the fish + swath species assemblage, habitat features that differed most strongly among clusters included mean SST [$F(5, 32) = 34.73$, $p = 1.03 \times 10^{-10}$], reef area [$F(5, 32) = 9.931$, $p = 2.12 \times 10^{-5}$], and slope [$F(5, 32) = 4.206$, $p = .0062$], while differences in relief, littoral drift, and amount of bedrock were also observed. Samples grouped by colder temperatures were differentiated by reef area, and samples grouped by warmer temperatures were primarily differentiated by slope. Given that our reference validation sample set was small ($n = 8$), we considered discarding external validation and repeated the cluster analysis with the calibration and validation samples combined. Similar cluster groups were identified and similar habitat variables predicted group membership. We futher evaluated the average and *SD* of O/E scores for this larger calibration sample group with comparable results to the smaller calibration group and therefore retained external validation.

3.2 | Model performance

The predictive model produced similar accuracy and precision across the five different species assemblages (Table 2). The UPC index produced the highest accuracy as measured by reference site scores closest to 1. The indices using fish + swath or swath only were equally accurate; each type of species assemblage had only one reference validation sample falling outside 1 *SD* of the reference calibration mean. The indices using fish, UPC and all assemblages were less accurate, with eight, three and two validation samples outside 1 *SD* of the reference calibration mean, respectively. The index using all assemblages combined resulted in the highest precision as measured by the lowest *SD* of scores. At calibration sites, improvement in precision over the null model was greatest for fish, fish + swath, and swath assemblage index scores. Performance through predictive modeling was only impoved by 0.6% using the UPC assemblage in these samples. Reference site samples not used for model training showed the greatest index precision improvement when all

TABLE 2 Index score means and *SD* for each species assemblage and reference or non-reference category using predictive and null models. Reference calibration, reference validation, and non-reference data are shown for the randomly selected sample year at each site where multiple years were sampled. The "reference – all" category includes all sample years at both reference calibration and validation sites, but only those sample years not used for model calibration

	Fish + swath		Fish		Swath		UPC		All assemblages		
Data category	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	*n*
Predicted model											
Reference – calibration	1.02	0.13	1.01	0.13	1.01	0.18	1.03	0.18	1.03	0.13	32
Reference – validation	1.03	0.09	0.99	0.12	1.04	0.07	1.03	0.09	1.02	0.09	8
Reference – all	1.00	0.16	0.97	0.20	1.02	0.17	1.01	0.16	1.01	0.16	51
Non-reference	0.95	0.19	0.92	0.23	0.99	0.20	0.98	0.18	0.97	0.18	100
Null model											
Reference – calibration	1.00	0.16	1.00	0.16	1.00	0.21	1.00	0.18	1.00	0.13	32
Reference – validation	1.00	0.15	0.96	0.20	1.02	0.15	1.08	0.13	1.01	0.11	8
Reference – all	1.00	0.19	0.95	0.22	1.04	0.21	1.03	0.18	1.01	0.20	51
Non-reference	0.96	0.21	0.92	0.24	1.00	0.22	0.98	0.21	0.97	0.19	100

UPC, uniform point contact.

assemblages were used. Finally, the greatest improvement in precision at reference validation sites was achieved by the indices using fish and swath assemblages. The majority of reef sites were sampled in multiple years. The median of within-site *SD* of O/E (fish + swath) scores at reference and non-reference sites was 0.046 and 0.065, respectively.

Random forest models show that predictive modeling effectively reduced bias in selected calibration and validation samples by reducing the influence of habitat variables on fish + swath index scores (Table 3). Pseudo *R*-squared values for all reference samples and for validation and unselected calibration samples (not used for model calibration) were greatly reduced by predictive modeling but still remained at 21% (Table 3). Regressions of O/E index scores (fish + swath) at reference sites against the habitat variables remaining in the final random forest model showed minimal bias as shown by small slopes and low *R*-squared values. Figure 3 illustrates the four strongest relationships (i.e., worst examples of bias) between O/E scores at reference sites not used for model calibration and habitat variables. There was no significant bias by sample year among reference samples [*F*(3, 47) = 1.463, *p* = .237]. As another indicator of accuracy, index scores at reference validation sites were all greater than the 10th percentile of scores at reference calibration sites.

TABLE 3 Variance in index scores trained on the fish + swath species assemblage explained by habitat variables

	Pseudo *R*-squared	
	Predictive model	Null model
All reference samples	0.21	0.54
Reference (not for calibration)	0.21	0.44
Selected validation	−0.37	0.03
Selected calibration	−0.16	0.25

Index responsiveness to stress was greatest for the fish and fish + swath species assemblage, as measured by the largest difference between scores of reference (samples not used for model training) and non-reference samples (Table 4, Figure 4). Regressions of index scores, based on each assemblage, with stressor indices showed very small negative slopes and small *R*-squared values, indicating weak responsiveness of the index to stressors. Values were slightly higher for fishing pressure compared with the water quality variables. Figure 5 illustrates the two strongest relationships among stressors and indices based on different assemblages: fishing pressure and fish or fish + swath. While the indices using the fish and fish + swath assemblages showed comparable performance we chose to apply the fish + swath index to further investigations of the relative impacts of fishing and water quality because the larger suite of species might provide more insights into stressor effects.

3.3 | Relative impacts of fishing pressure and water quality

The threshold for the O/E (fish + swath) index was set at the value for the lowest performing reference sample (0.847). Non-reference samples scored as low as 0.173. Differences in means of habitat variables and stress indices in samples above and below the threshold showed that samples with low scores were associated with less bedrock, more cobble, low relief, shallower water, more sand transport, greater invertebrate settlement and higher fishing pressure (Table 5). Mean TSS and copper were higher while nitrate was lower at sites below the threshold, although differences were non-significant. From the multiple regression we identified both stress and habitat variability as important predictors of index scores. Stepwise variable reduction identified invertebrate settlement, sand transport, and fishing pressure as negatively associated with O/E (fish + swath) scores while

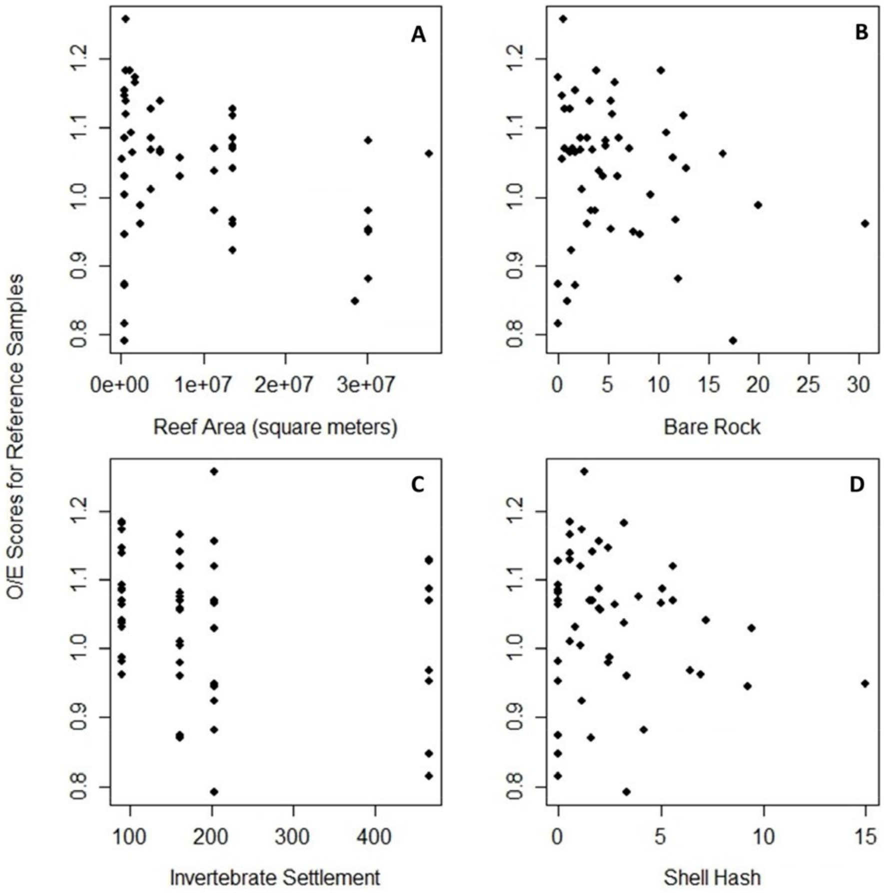

FIGURE 3 Scatter plots of observed to expected taxa (O/E) index scores at reference sites and four habitat variables ([a]: reef area, [b]: bare rock, [c]: invertebrate settlement, [d]: shell hash) with the strongest relationships to those scores

depth and the water quality index were positive influences, listed in order of predictive strength (Table 6).

Habitat was typically the strongest predictor of the abundance of common taxa when compared with the fishing pressure and water quality indices (Table 7). Of 29 common taxa, the abundance was most strongly related to habitat, fishing pressure and water quality in 19, eight and two cases, respectively. When comparing only the influence of stressors fishing pressure was a stronger predictor of abundance than water quality for all fish, most invertebrate, and most algal taxa. Those invertebrates most influenced by water quality were sessile. Of the four common algal taxa, only the important, habitat-forming giant kelp (*Macrocystis pyrifera*) was most strongly influenced by water quality.

4 | DISCUSSION

To our knowledge, this is the first O/E index constructed for a marine rocky reef ecosystem. While this pilot study employing a limited number of survey sites did not result in an assessment tool ready for use by scientists or managers, it illustrated initial concepts and demonstrated the potential utility of bioassessment index application to scientists and managers who invest in this habitat. The benefits of an O/E index overcome at least three challenges that have presented roadblocks to scientists and managers in the past. The first challenge the O/E overcomes is its ability to integrate across the entire biological community and not rely on single species (or groups of single species) assessments. Much research on marine community structure has focused on disentangling the effects of biotic or top-down forces (often impacted by fishing pressure) from abiotic or bottom-up forces (often impacted by water quality). Identification of the dominant force can be limited by ecological interactions and non-linear responses in single species assessments (Foster & Schiel, 2010; Heck & Valentine, 2007; Huntington, Karnauskas, Babcock, & Lirman, 2010; Sangil et al., 2013; Smith, Reed, Mohajerani, & Fong, 2004; Worm, Lotze, & Sommer, 2000). Multivariate approaches are common in assessments of rocky reef ecosystems (Hamilton, Caselle, Malone, & Carr, 2010;

TABLE 4 The difference in means of observed to expected taxa index scores for reference and non-reference samples (not used for model calibration) and 95% confidence intervals (CI) for each assemblage type

	Fish	Swath	UPC	Fish + swath	All assemblages
Difference in means	0.063	0.038	0.004	0.055	0.031
CI lower	0.017	−0.005	−0.035	0.019	−0.002
CI upper	0.109	0.081	0.042	0.091	0.065

UPC, uniform point contact.

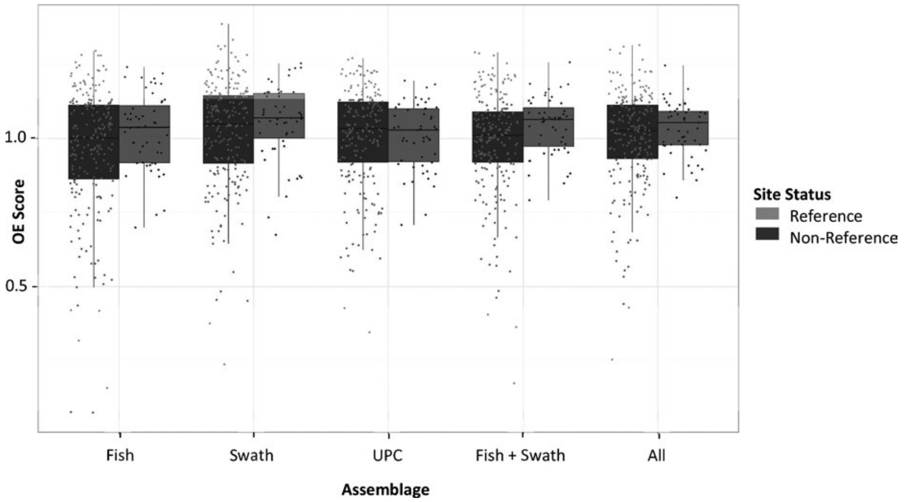

FIGURE 4 Boxplot of reference and non-reference sample index scores for each species assemblage. O/E index, observed to expected taxa index; UPC, uniform point contact

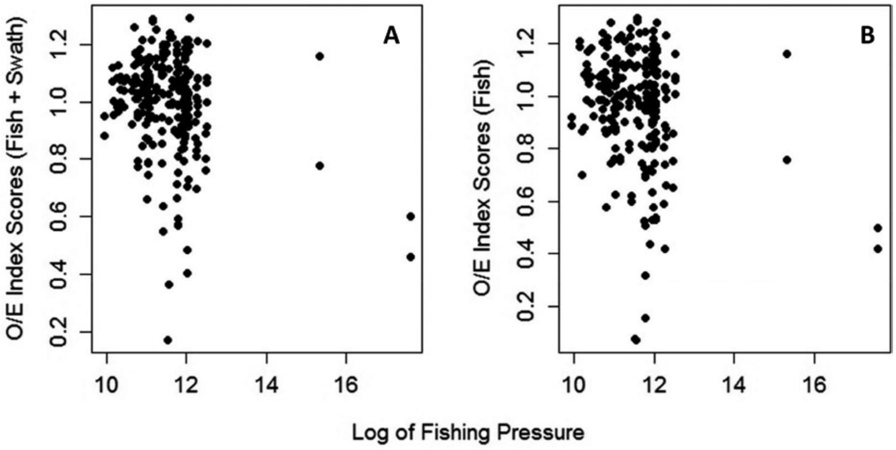

FIGURE 5 Scatter plots of observed to expected taxa (O/E) index scores based on (a): fish + swath and (b): fish assemblages with the log-transformed fishing pressure index

Horta e Costa et al., 2014; Sala et al., 2012; Sangil et al., 2013; Smith, Ajani, & Roberts, 1999; Smith & Simpson, 1993), but these studies are typically focused on single and spatially limited project issues. These single-focused multivariate approaches, unlike the O/E index, limit broader applicability such as when new sites or future time periods are sampled. In this paper, we trained five O/E models on five different species assemblages (including combinations of fish, invertebrates and algae), all of which performed similarly with regard to producing mean index scores close to 1 at reference sites (accuracy) and doing so consistently (precision). However, the O/E index achieved its greatest sensitivity to stress when fish and non-colonial/non-blanketing macroinvertebrates and algae were combined together (fish + swath). In freshwater habitats, the sensitivity of the O/E index to stress at the community level has been attributed to its ability to integrate differences in species tolerance and differences in life-history strategies, while accounting for natural species replacement (Hawkins, 2006). Our results show the mix of rocky reef assemblages in the selected O/E index has these important characteristics.

A second benefit an O/E index approach is that it mitigates the challenges of natural variability that confound many attempts to draw out changes due to human-induced stress. The fact that rocky reef communities respond to differences in temperature, reef size, substrate and other purely natural factors is not new (Dayton et al., 1984; Tegner & Dayton, 1991; Torres-Moye, Edwards, & Montaño-Moctezuma, 2013). However, the ecological modeling in the O/E index that predicts taxa occurrence is a novel approach in this habitat to 'normalize' these environmental gradients. The challenge of natural variability is perhaps maximized in California where the cold, arctic California Current meets the warm, sub-tropical Davidson Countercurrent, and rocky reef substrates range from low-lying cobbles with intermittent sand burial to tall pinnacles (Airame et al., 2003 and references therein). However, this challenge is no less than the dynamic freshwater habitat variability observed in California where O/E models have had to overcome natural gradients that range from the highest (Mt Whitney, 4,421 m) and lowest (Death Valley −86 m) elevations in the contiguous US (Mazor et al., 2016). The key to meeting the challenge of large natural variability is ensuring that our pool of reference sites captures the entire gradient of natural environmental conditions. In this study, we illustrated that our reference samples captured a similar range of environmental gradients as was observed in our non-reference samples.

The third benefit of O/E indices is their ability to identify stressed sites once natural variability is accounted for. In this study, 13% of the samples in Southern California not used for model calibration were

TABLE 5 Means and 95% confidence intervals (CIs) of habitat and stress variables for samples with observed to expected taxa index scores above and below the index threshold. A single randomly selected sample year for each site was used

Habitat or stress variable	CI lower	CI upper	Mean above	Mean below	*t*-value
Bare rock	−5.90	2.18	5.97	7.83	−1.0
Bare sand	−8.87	2.09	10.07	13.46	−1.3
Bedrock	4.32	27.85	64.39	48.31	2.8
Boulder	−11.72	3.86	14.79	18.72	−1.0
Cobble	−15.24	0.66	9.34	16.62	−1.9
Relief 0–0.1 m	−37.41	−9.59	23.43	46.93	−3.5
Relief 0.1–1 m	4.18	27.84	57.20	41.19	2.8
Relief 1–2 m	2.90	10.81	12.55	5.69	3.5
Relief > 2 m	−7.08	9.80	6.82	5.46	0.3
Sand	−12.21	2.49	11.49	16.35	−1.4
Sediment	−2.77	0.70	0.57	1.60	−1.2
Shell hash	−3.85	0.94	2.14	3.60	−1.3
Deepest survey	−0.13	4.80	19.15	16.81	1.9
Slope	−1.21	5.36	6.63	4.56	1.3
Reef area	−6,197,460	5,635,466	9,437,185	9,718,182	−0.1
Littoral drift	−4.36	−0.36	1.88	4.24	−2.4
Settlement 1	−225.76	−90.61	225.63	383.82	−4.8
Settlement 2	−200.35	−58.02	327.53	456.71	−3.7
Mean SST	−0.50	−0.55	16.39	16.36	0.1
SST	−0.40	0.78	16.37	16.18	0.7
SST1	−0.63	0.43	15.97	16.07	−0.4
SST2	−0.87	0.19	16.22	16.56	−1.3
ENSO	0.08	0.47	−0.27	−0.54	2.8
ENSO1	−0.25	0.06	−0.43	−0.33	−0.3
ENSO2	−0.28	−0.03	0.12	0.27	−2.4
WQ index	−0.13	0.28	0.55	0.48	−0.7
TSS	−3122.08	1872.81	1973.81	2598.45	−0.5
Copper	−1.11	0.88	0.77	0.88	−0.2
Nitrate	1.769	255.24	141.14	26.77	1.3
Log fishery	−1.06	0.14	11.43	11.89	−1.6

ENSO, El Niño Southern Oscillation; SST, sea surface temperature; TSS, total suspended solids; WQ, water quality.

TABLE 6 Results of multiple linear regression of stressors and habitat variables with observed to expected taxa (fish + swath) index scores

	Estimate	*t*
Intercept	1.40	7.8
Settlement 1	-3.55×10^{-4}	−4.1
Littoraldrift	-1.28×10^{-2}	−3.1
Logfishery	-4.33×10^{-2}	−2.7
Depth	6.19×10^{-3}	2.3
WQ index	4.41×10^{-2}	1.7
Bedrock	9.06×10^{-4}	1.5

WQ, water quality.

identified as non-reference or 'stressed'. O/E indices can provide an objective quantitative method for ranking and prioritizing future activities. This allows scientists to focus their efforts on the most stressed sites, identifying site-specific causative factors, and for managers to focus their efforts on restoration and/or mitigation. Moreover, once a management action is taken, an O/E index can provide a quantitative tool for measuring and communicating progress towards agreed upon endpoints of success.

Application of the O/E index in this study suggested that fishing was a more important stressor than water quality. In the step-wise regression of O/E scores versus stressor and habitat values, invertebrate settlement and sand transport were most correlated with decreasing O/E scores followed by fishing pressure (Table 6). The water quality index was positively related to O/E scores. The slopes and *R*-squared

TABLE 7 Increase in mean square error with removal of fishing pressure, water quality (WQ) index, or habitat from random forest models predicting abundance of individual taxa. The variable with the highest value (greatest influence on abundance) is shown in bold

Taxonomic group	Taxon	Common name	Fishing	WQ	Habitat
Fish	*Brachyistius frenatus*	Kelp perch	**0.00040**	0.00013	0.00039
	Chromis punctipinnis	Blacksmith	0.019	0.016	**0.049**
	Embiotoca jacksoni	Black surfperch	0.00005	0.00004	**0.00007**
	Girella nigricans	Opaleye	**0.00010**	−0.00006	−0.00023
	Halichoeres semicinctus	Rock wrasse	0.000015	0.000004	**0.000042**
	Hypsypops rubicundus	Garibaldi	0.00010	−0.00001	**0.00042**
	Medialuna californiensis	Halfmoon	0.000037	0.000005	**0.000123**
	Oxyjulis californica	Senorita	0.0012	−0.0022	**0.0031**
	Oxylebius pictus	Painted greenling	0.000007	0.000005	**0.000046**
	Paralabrax clathratus	Kelp bass	0.0005	0.0004	**0.0014**
	Rhacochilus vacca	Pile perch	**0.000023**	−0.000003	0.000020
	Sebastes atrovirens	Kelp rockfish	0.000014	0.000005	**0.000029**
	Semicossyphus pulcher	Sheephead	**0.000000**	−0.000004	0.000000
Invertebrate	*Anthopleura sola*	Sunburst anemone	−21.00	**51.18**	50.41
	Crassedoma gigantea	Rock scallop	0.488	0.912	**1.200**
	Kelletia kelletii	Kellet's whelk	**0.746**	0.084	−0.112
	Megastraea undosa	Wavy turban snail	**6.51**	2.00	2.53
	Megathura crenulata	Giant keyhole limpet	1.522	0.416	**1.614**
	Muricea californica	Golden gorgonian	42.26	80.20	**84.53**
	Parastichopus parvimensis	Warty sea cucumber	11.06	3.31	**30.42**
	Patiria miniata	Bat star	147.26	131.91	**644.13**
	Pisaster giganteus	Giant sea star	9.88	2.05	**16.01**
	Strongylocentrotus franciscanus	Purple urchin	**2398.60**	2029.04	1476.11
	Strongylocentrotus purpuratus	Red urchin	**9948.34**	4702.04	9264.58
	Tethya californiana	Orange puffball sponge	0.766	0.801	**9.683**
Algae	*Cystoseira osmundacea*	Cystoseira	206.04	18.38	**228.71**
	Eisenia arborea	Sea palm	49.62	28.69	**119.06**
	Laminaria farlowii	Laminaria	87.66	10.50	**118.87**
	Macrocystis pyrifera	Giant kelp	10.88	**15.23**	9.16

value of this regression indicated weak relationships with O/E scores for all predictors.

Species-specific responses to stressors also showed a greater influence from fishing than water quality pressure (Table 7). Disregarding the influence of habitat on taxa abundance, fishing was more strongly related to the abundance of all common fish species than water quality. Many of these taxa experience little to no fishing pressure. Abundance of these taxa might be related to changes in community interactions brought on by fishing pressure or there may be no causal link at all. This illustrates the potential for species interactions to complicate understanding of ecosystem response to stress and thus the need for an index. At a more local scale, Parnell, Dayton, Fisher, Loarie, and Darrow (2010) noted a relationship between areas of low kelp persistence and high fishing pressure within a single kelp forest, which may be mediated by trophic interactions between kelp, urchin grazers and harvested urchin predators. Our analyses showed a greater influence of water quality on giant kelp; however fishing was more strongly related to the abundance of all other common algae (Table 7). While sedimentation has been shown in many cases to be a driver of kelp decline (Foster & Schiel, 2010 and references therein; Spurkland & Iken, 2011), several other studies have noted little relationship between terrestrial inputs and kelp dynamics. These include a large sewage effluent spill event (Tegner et al., 1995) and relatively little uptake of terrrestrially sourced nutrients relative to upwelling (Foley & Koch, 2010).

While the O/E index approach has many advantages, certain limitations need to be addressed before it should be used. One limitation was the remaining influence of natural gradients, even after the ecological modeling was completed. While the relationship between O/E index scores at reference sites and habitat variables was weak (Figure 3), the step-wise regression analysis suggested that habitat features were more strongly related to index scores than stressors among reference and non-reference sites. While this may be a true reflection of weak ecosystem impacts from stressors, two outcomes could account for bias in the O/E model. The first is that the O/E model was not adequately trained, most likely due to insufficient reference sites with these natural characteristics. As previously illustrated, our reference pool covered all of the major natural gradients in Southern California rocky reefs found in our regionwide data set, and in the relative frequency of which they are found. However, the relative frequency of naturally occurring reefs with bare rock, bare sand or bare cobble reefs with high sand transport is low and, as a result, the number of reference sites with these natural characteristics was also low. The second outcome that could account for this bias is the co-variance of fishing pressure and reef characteristics. In essence, fishing pressure exerts a disproportionately large influence on reefs with these natural characteristics because resiliency is less. Archaeological evidence at the Channel Islands for geographic continuity in kelp forest-associated red abalone populations suggests resilience of kelp forests to dramatic ecological change over long time scales (Braje, Erlandson, Rick, Dayton, & Hatch, 2009). Southern California kelp forests are also thought to be particularly resilient due to resistance of urchin-mediated trophic cascades provided by a diverse suite of urchin predators (Steneck et al., 2002). Despite strong resilience at broad spatial and long temporal scales, local scale kelp forests and their associated organisms are more variable and subject to disturbance on unstable substrata (Dayton, 1985; Edwards, 2004).

A second limitation of the model is how reference sites were selected. We utilized a composite fishing pressure index and a composite water quality index as a screen for human influence. These were useful tools for us, as these were the two stressors we wished to evaluate, similar to how soft-bottom marine ecotoxicologists use sediment contamination as screens for reference sites for their quantitative condition assessment tool, the benthic response index (Smith et al., 2001). However, the fishing pressure and water quality composite indices each had their own assumptions that could miss potential impacts to reef communities. For example, the water quality index captured the largest pollution discharges and representative chemicals of concern in Southern California, but it did not capture every source and every chemical. Likewise, the fishing pressure index captured commercial fishing and CPFV recreational fishing, but it did not capture private vessel fishing effort (which remains unquantified in Southern California). Finally, both the fishing pressure and water quality indices were based on long-term averages of up to 10–30 years. However, the samples used for calibrating, validating and applying the O/E index are single point-in-time measurements. Clearly, further work in the development of screening tools should be evaluated for building future O/E indices because the inclusion of non-reference sites in the reference pool will decrease sensitivity of the model and potentially lead to false negative conclusions (or type II error of identifying a sample as reference when it really is not).

A third limitation of the O/E model is the basic construct of species presence or absence. Absence detections on surveys need not imply complete absence of a taxon from the area but rather reduced abundance or patchy distribution. Notwithstanding, it may be more difficult to detect absences in marine habitats. This is because relative to terrestrial species, marine species tend to be more abundant and have higher potential for subpopulation connectivity due to a pelagic larval phase. This may also be true relative to freshwater aquatic ecosystems. We would encourage future iterations of predicitive rocky reef ecological modeling to consider weighting factors for key biological community factors including abundance, biomass, taxa functional groups and/or diversity measures.

The O/E index we developed for shallow rocky reefs in Southern California showed a high degree of overlap in the range of scores at reference and non-reference sites (Figure 4). This indicates low sensitivity of the index to stressors and limits applicability for management in its current form. While this low sensitivty may accurately reflect an ecosystem minimally impacted by fishing and water quality stress, a combination of the limitations outlined above may be responsible for the high degree of overlap in the range of O/E index score values at reference and non-reference sites (Figure 4). We set the threshold for poor performing index scores at the lowest performing reference site (0.847) and this needs to be refined. Defining impact thresholds is partly technical, but also partly a policy task. The effects of the conservative index threshold we utilized can be seen in Figure 1a,c, which illustrates the relatively small proportion of non-reference sites that fell below the performance threshold. Other quantitative indices have used other population-based estimators, i.e., first percentile, fifth percentile (Mazor et al., 2016), 2 *SD* below the mean of the reference population (Ode, Rehn, & May, 2005), or community-based ecological cutoffs, i.e., loss of taxa groups, (Smith et al., 2001). The statistical approaches lend themselves to objective and repeatable thresholds, but the magnitude of these changes (and the management reactions to them) are ultimately a policy choice. Finally, further exploration of which taxonomic assemblages are included could lead to a better performing index as well as increased understanding of the relative effects of stressors on different ecosystem components.

ACKNOWLEDGEMENTS

We thank the Southern California Bight 2013 Regional Monitoring Program (Bight '13) Marine Protected Area Planning Committee for their guidance and review during this study. We also acknowledge the following organizations for the extensive field efforts and careful attention to consistent methodology that allowed for collation of the biological data used in this study: the Partnership for the Interdisciplinary Studies of Coastal Oceans, Vantuna Research Group, San Diego State University.

ORCID

Julia H. Coates http://orcid.org/0000-0002-1847-8470

REFERENCES

Airame, S., Dugan, J. E., Airamé, S., Dugan, J. E., Lafferty, K. D., Leslie, H., … Warner, R. R. (2003). Applying ecological criteria to marine reserve design: A case study from the California Channel Islands. *Ecological Applications*, *13*(Supplement), 170–184. https://doi.org/10.1890/1051-0761(2003)013[0170:AECTMR]2.0.CO;2

Airoldi, L. (2003). The effects of sedimentation on rocky coast assemblages. *Oceanography and Marine Biology*, *41*(41), 161–236.

Andersen, K. H., & Pedersen, M. (2010). Damped trophic cascades driven by fishing in model marine ecosystems. *Proceedings of the Royal Society Biological Sciences*, *277*, 795–802. https://doi.org/10.1098/rspb.2009.1512

Balata, D., Piazzi, L., & Cinelli, F. (2007). Increase of sedimentation in a subtidal system: Effects on the structure and diversity of macroalgal assemblages. *Journal of Experimental Marine Biology and Ecology*, *351*(1–2), 73–82. https://doi.org/10.1016/j.jembe.2007.06.019

Braje, T. J., Erlandson, J. M., Rick, T. C., Dayton, P. K., & Hatch, M. B. (2009). Fishing from past to present: Continuity and relilience of red abalone fisheries on the Channel Islands, California. *Ecological Applications*, *19*(4), 906–919. https://doi.org/10.1890/08-0135.1

CalOST, CDFW (2013). *State of the California central coast: Results from baseline monitoring of marine protected areas 2007-2012*. California Ocean Science Trust and California Department of Fish and Wildlife, CA, USA. February 2013.

Cao, Y., & Hawkins, C. P. (2011). The comparability of bioassessments: A review of conceptual and methodological issues. *Journal of the North American Benthological Society*, *30*(3), 680–701. https://doi.org/10.1899/10-067.1

CDFW (2008). California Department of fish and game, partnership for interdisciplinary studies of coastal oceans, channel islands national marine sanctuary, and Channel Islands National Park. In S. Airame & J. Ugoretz (Eds.), *Channel islands marine protected areas: First 5 years of monitoring: 2003-2008*. 20 pp. www.dfg.ca.gov/marine.

CDFW (2014). 2850-2863 California Fish and Game Code (Marine Life Protection Act) Sections 2850-2863. https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=FGC&division=3.&title=&part=&chapter=10.5.&article=. (accessed on 16 January 2018).

Connell, S. D. (2005). Assembly and maintenance of subtidal habitat heterogeneity: Synergistic effects of light penetration and sedimentation. *Marine Ecology Progress Series*, *289*, 53–61. https://doi.org/10.3354/meps289053

Dayton, P. K. (1985). Ecology of kelp communities. *Annual Review of Ecology and Systematics*, *16*, 215–245. https://doi.org/10.1146/annurev.es.16.110185.001243

Dayton, P. K., Currie, V., Gerrodette, T., Keller, B. D., Rosenthal, R., & Tresca, D. V. (1984). Patch dynamics and stability of some California kelp communities. *Ecological Monographs*, *54*(3), 253–289. https://doi.org/10.2307/1942498

Douvere, F. (2008). The importance of marine spatial planning in advancing ecosystem-based sea use management. *Marine Policy*, *32*, 762–771. https://doi.org/10.1016/j.marpol.2008.03.021

Edwards, M. S. (2004). Estimating scale-dependency in disturbance impacts: El Ninos and giant kelp forests in the northeast Pacific. *Oecologia*, *138*, 436–447. https://doi.org/10.1007/s00442-003-1452-8

Foley, M. M., & Koch, P. L. (2010). Correlation between allochthonous subsidy input and isotopic variability in the giant kelp Macrocystis pyrifera in central California, USA. *Marine Ecology Progress Series*, *409*, 41–50. https://doi.org/10.3354/meps08600

Foster, M. S., & Schiel, D. R. (2010). Loss of predators and the collapse of southern California kelp forests (?): Alternatives, explanations and generalizations. *Journal of Experimental Marine Biology and Ecology*, *393*(1–2), 59–70. https://doi.org/10.1016/j.jembe.2010.07.002

Fox, W. E., Hill, R. A., Leibowitz, S. G., Olsen, A. R., Thornbrugh, D. J., & Weber, M. H. (2017). Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. *Environmental Monitoring and Assessment*, *189*, 316. https://doi.org/10.1007/s10661-017-6025-0

Ghedini, G., Klein, J. C., & Coleman, R. A. (2011). Potential effects of storm-water run-off on assemblages of mobile invertebrates. *Marine Ecology Progress Series*, *439*, 169–180. https://doi.org/10.3354/meps09307

Gillett, D. J., Pondella, D. J. II, Freiwald, J., Schiff, K. C., Caselle, J. E., Shuman, C., & Weisberg, S. B. (2012). Comparing volunteer and professionally collected monitoring data from the rocky subtidal reefs of Southern California, USA. *Environmental Monitoring and Assessment*, *184*(5), 3239–3257. https://doi.org/10.1007/s10661-011-2185-5

Gorman, D., Russell, B. D., & Connell, S. D. (2009). Land-to-sea connectivity: Linking human-derived terrestrial subsidies to subtidal habitat change on open rocky coasts. *Ecological Applications*, *19*(5), 1114–1126. https://doi.org/10.1890/08-0831.1

Guerry, A. D. (2005). Icarus and Daedalus: Conceptual and tactical lessons for marine ecosystem-based management. *Frontiers in Ecology and the Environment*, *3*(4), 202–211. https://doi.org/10.1890/1540-9295(2005)003[0202:IADCAT]2.0.CO;2

Halpern, B. S., Longo, C., Scarborough, C., Hardy, D., Best, B. D., Doney, S. C., … Samhouri, J. F. (2014). Assessing the health of the U.S. west coast with a regional-scale application of the ocean health index. *PLoS ONE*, *9*(6), e98995. https://doi.org/10.1371/journal.pone.0098995

Halpern, B. S., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J. F., Katona, S. K., … Zeller, D. (2012). An index to assess the health and benefits of the global ocean. *Nature*, *488*, 615–620. https://doi.org/10.1038/nature11397

Hamilton, S. L., Caselle, J. E., Malone, D. P., & Carr, M. H. (2010). Incorporating biogeography into evaluations of the Channel Islands marine reserve network. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18272–18277. https://doi.org/10.1073/pnas.0908091107

Hawkins, C. P. (2006). Quantifying biological integrity by taxonomic completeness: Its utility in regional and global assessments. *Ecological Applications*, *16*(4), 1277–1294. https://doi.org/10.1890/1051-0761(2006)016[1277:QBIBTC]2.0.CO;2

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. *Ecological Applications*, *10*, 1456–1477. https://doi.org/10.1890/1051-0761(2000)010[1456:DAEOPM]2.0.CO;2

Heck, K. L. Jr, & Valentine, J. F. (2007). The primacy of top-down effects in shallow benthic ecosystems. *Estuaries and Coasts*, *30*(3), 371–381. https://doi.org/10.1007/BF02819384

Horta e Costa, B., Assis, J., Franco, G., Erzini, K., Henriques, M., Gonçalves, E. J., & Caselle, J. E. (2014). Tropicalization of fish assemblages in temperate biogeographic transition zones. *Marine Ecology Progress Series*, *504*, 241–252. https://doi.org/10.3354/meps10749

Howard, M. D. A., Sutula, M., Caron, D. A., Chao, Y., Farrara, J. D., Frenzel, H., … Sengupta, A. (2014). Anthropogenic nutrrient sources rival natural sources on small scales in the coastal waters of the Southern California Bight. *Limnology and Oceanography*, *59*, 285–297.

Huntington, B. E., Karnauskas, M., Babcock, E. A., & Lirman, D. (2010). Untangling natural seascape variation from marine reserve effects using a landscape approach. *PLoS ONE*, *5*(8), e12327. https://doi.org/10.1371/journal.pone.0012327

Irving, A. D., & Connell, S. D. (2002). Sedimentation and light penetration interact to maintain heterogeneity of subtidal habitats: Algal versus invertebrate dominated assemblages. *Marine Ecology Progress Series*, *245*, 83–91. https://doi.org/10.3354/meps245083

de Jonge, V. N., Pinto, R., & Turner, R. K. (2012). Integrating ecological, economic and social aspects to generate useful management informatioin under the EU Directives 'ecosystem approach'. *Ocean & Coastal Management*, *68*, 169–188. https://doi.org/10.1016/j.ocecoaman.2012.05.017

Lyon, G. S., & Stein, E. D. (2009). How effective has the Clean Water Act been at reducing pollutant mass emissions to the Southern California Bight over the past 35 years? *Environmental Monitoring and Assessment*, *154*, 413–426. https://doi.org/10.1007/s10661-008-0408-1

Mace, P. M. (2004). In defence of fisheries scientists, single-species models and other scapegoats: Confronting the real problems. *Marine Ecology Progress Series*, *274*, 285–291.

Mazor, R. D., Rehn, A., Ode, P. R., Engeln, M., Schiff, K. C., Stein, E. D., … Hawkins, C. P. (2016). Bioassessment in complex environments: Designing an index for consistent meaning in different settings. *Freshwater Science*, *35*, 249–271. https://doi.org/10.1086/684130

Moss, D., Furse, T., Wright, J. F., & Armitage, P. D. (1987). The prediction of macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology*, *17*, 41–52. https://doi.org/10.1111/j.1365-2427.1987.tb01027.x

O'Connor, N. E. (2013). Impacts of sewage outfalls on rocky shores: Incorporating scale, biotic assemblage structure and variability into monitoring tools. *Ecological Indicators*, *29*, 501–509. https://doi.org/10.1016/j.ecolind.2013.01.020

Ode, P., Hawkins, C. P., & Mazor, R. D. (2008). Comparability of biological assessments derived from predictive models and multimetric indices of increasing geographic scope. *Journal of the North American Benthological Society*, *27*(4), 967–985. https://doi.org/10.1899/08-051.1

Ode, P., Rehn, A., & May, J. T. (2005). A quantitative tool for assessing the integrity of southern coastal California streams. *Environmental Management*, *35*, 493–504. https://doi.org/10.1007/s00267-004-0035-8

Ode, P., Rehn, A. C., Mazor, R. D., Schiff, K. C., Stein, E. D., May, J. T., … Hawkins, C. P. (2016). Evaluating the adequacy of a reference-site pool for ecological assessments in environmentally complex regions. *Freshwater Science*, *35*(1), 237–248. https://doi.org/10.1086/684003

Ostermiller, J. D., & Hawkins, C. P. (2004). Effects of sampling error on bioassessments of stream ecosystems: Application to RIVPACS-type models. *Journal of the North American Benthological Society*, *23*, 363–382. https://doi.org/10.1899/0887-3593(2004)023<0363:EOSEOB>2.0.CO;2

Page, H. M., Reed, D. C., Brzezinski, M. A., Melack, J. M., & Dugan, J. E. (2008). Assessing the importance of land and marine sources of organic matter to kelp forest food webs. *Marine Ecology Progress Series*, *360*, 47–62. https://doi.org/10.3354/meps07382

Parnell, P. E., Dayton, P. K., Fisher, R., Loarie, C., & Darrow, R. (2010). Spatial patterns of fishing effort off San Diego: Implications for zonal management and ecosystem function. *Ecological Applications*, *20*(8), 2203–2222. https://doi.org/10.1890/09-1543.1

Patsch, K., & Griggs, G. (2006) *Littoral cells, sand budgets, and beaches: Understanding California's shoreline*. In U.o.C. Institute of Marine Sciences, Santa Cruz, California Department of Boating and Waterways, California Coastal Sediment management WorkGroup (Ed.).

Pearson, T. H., & Rosenberg, R. (1978). Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology – An Annual Review*, *16*, 229–311.

Pondella, D. J., Claisse, J., Schaffner, R., Ritter, K., Schiff, K. (2011). Southern California Bight 2008 Regional Monitoring Program: Volume V. Rocky Reefs. Southern California Coastal Water Research Project, Costa Mesa, CA.

Pondella, D. J., Schiff, K., Schaffner, R., Zellmer, A., Coates, J. (2016). *Southern California bight 2013 regional monitoring program*: Vol. *II*. Rocky Reefs. Southern California Coastal Water Research Project.

Pondella, D. J. I., Williams, J. P., Claisse, J., Schaffner, B., Ritter, K., & Schiff, K. (2015). The physical characteristics of nearshore rocky reefs in the Southern California Bight. *Bulletin of the Southern California Academy of Sciences*, *114*(3), 105–122.

Russell, B. D., Thompson, J. A. I., Falkenberg, L. J., & Connell, S. D. (2009). Synergistic effects of climate change and local stressors: CO_2 and nutrient-driven change in subtidal rocky habitats. *Global Change Biology*, *15*(9), 2153–2162. https://doi.org/10.1111/j.1365-2486.2009.01886.x

Sala, E., Ballesteros, E., Dendrinos, P., Di Franco, A., Ferretti, F., Foley, D., … Zabala, M. (2012). The structure of mediterranean rocky reef ecosystems across environmental and human gradients, and conservation implications. *PLoS ONE*, *7*(2), 1–13.

Sangil, C., Martin-Garcia, L., Hernández, J. C., Concepción, L., Fernández, R., & Clemente, S. (2013). Impacts of fishing and environmental factors driving changes on littoral fish assemblages in a subtropical oceanic island. *Estuarine Coastal and Shelf Science*, *128*, 22–32. https://doi.org/10.1016/j.ecss.2013.04.023

Schaffner, B., Steinberg, S., & Schiff, K. (2015). *A GIS tool to compute a pollutant exposure index for the southern California bight*. Redlands, CA, USA: Esri Press. https://doi.org/10.17128/9781589483651

Schiel, D. R., Wood, S. A., Dunmore, R. A., & Taylor, D. I. (2006). Sediment on rocky intertidal reefs: Effects on early post-settlement stages of habitat-forming seaweeds. *Journal of Experimental Marine Biology and Ecology*, *331*(2), 158–172. https://doi.org/10.1016/j.jembe.2005.10.015

Schroeter, S. C., Dixon, J. D., Ebert, T., & Richards, J. (2012). *SBC LTER: Time series of settlement of urchins and other invertebrates*. Santa Barbara Coastal LTER.

Smail, E. A., Webb, E. A., Franks, R. P., Bruland, K. W., & Sañudo-Wilhelmy, S. A. (2012). Status of metal contamination in surface waters of the coastal ocea off Los Angeles, California since the Implementation of the Clean Water Act. *Environmental Science and Technology*, *46*(8), 4301–4311.

Smith, A. K., Ajani, P. A., & Roberts, D. E. (1999). Spatial and temporal variation in fish assemblages exposed to sewage and implications for management. *Marine Environmental Research*, *47*(3), 241–260. https://doi.org/10.1016/S0141-1136(98)00120-2

Smith, R., Bergen, M., Weisberg, S. B., Cadien, D., Dalkey, A., Montagne, D., … Velarde, R. G. (2001). Benthic response index for assessing infaunal communities on the Southern California mainland shelf. *Ecological Applications*, *11*, 1073–1087. https://doi.org/10.1890/1051-0761(2001)011[1073:BRIFAI]2.0.CO;2

Smith, J. R., Reed, B. J., Mohajerani, L., & Fong, P. (2004). Influence of abiotic factors on the persistence of kelp habitats along the north coast of Santa Monica Bay. *Bulletin Southern California Academy of Sciences*, *103*(2), 79–92.

Smith, S. D. A., & Simpson, R. D. (1993). Effects of pollution on holdfast macrofauna of the kelp *Ecklonia radiata*: Discrimination at different taxonomic levels. *Marine Ecology Progress Series*, *96*(2), 199–208. https://doi.org/10.3354/meps096199

Spurkland, T., & Iken, K. (2011). Kelp bed dynamics in estuarine environments in Subarctic Alaska. *Journal of Coastal Research*, *27*(6), 133–143. https://doi.org/10.2112/JCOASTRES-D-10-00194.1

Steneck, R. S., Graham, M. H., Bourque, B. J., Corbett, D., Erlandson, J. M., Estes, J. A., & Tegner, M. J. (2002). Kelp forest ecosystems: Biodiversity, stability, resilience and future. *Environmental Conservation*, *29*(4), 436–459.

Steneck, R. S., & Wahle, R. A. (2013). American lobster dynamics in a brave new ocean. *Canadian Journal of Fisheries & Aquatic Sciences*, *70*(11), 1612–1624. https://doi.org/10.1139/cjfas-2013-0094

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. *Ecological Applications*, *16*(4), 1267–1276. https://doi.org/10.1890/1051-0761(2006)016[1267:SEFTEC]2.0.CO;2

Tegner, M. J., & Dayton, P. K. (1991). Sea urchins, El Ninos, and the long-term stability of southern California kelp forest communities. *Marine Ecology Progress Series*, *77*(1), 49–63. https://doi.org/10.3354/meps077049

Tegner, M. J., Dayton, P. K., Riser, K. L., Chadwick, D. B., Dean, T. A., & Deysher, L. (1995). Effects of a large sewage spill on a kelp forest community: Catastrophe or disturbance? *Marine Environmental Research*, *40*(2), 181–224. https://doi.org/10.1016/0141-1136(94)00008-D

Torres-Moye, G., Edwards, M. S., & Montaño-Moctezuma, C. G. (2013). Benthic community structure in kelp forests from the Southern

California Bight. *Ciencias Marinas*, *39*(3), 239–252. https://doi.org/10.7773/cm

Van Sickle, J., Hawkins, C. P., Larsen, D. P., & Herlihy, A. T. (2005). A null model for the expected macroinvertebrate assemblage in streams. *Journal of the North American Benthological Society*, *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. *Journal of the North American Benthological Society*, *26*, 319–331. https://doi.org/10.1899/0887-3593(2007)26[319:EORTAP]2.0.CO;2

Vander Laan, J. J., Hawkins, C. P., Olson, J. R., & Hill, R. A. (2013). Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. *Freshwater Science*, *32*(3), 801–820. https://doi.org/10.1899/12-186.1

Walker, J. W. (2007). Effects of fine sediments on settlement and survival of the sea urchin *Evechinus chloroticus* in northeastern New Zealand. *Marine Ecology Progress Series*, *331*, 109–118. https://doi.org/10.3354/meps331109

Ward, T. J. (2014). The condition of Australia's marine environment is good but in decline – An integrated evidence-based national assessment by expert elicitation. *Ocean & Coastal Management*, *100*, 86–100. https://doi.org/10.1016/j.ocecoaman.2014.07.012

Wolter, K. (2014). *Multivariate ENSO Index (MEI)*. U.S. Department of Commerce, National Oceanic and Atomospheric Administration, Earth System Research Laboratory, Physical Sciences Division.

Wolter, K., & Timlin, M. S. (1993). *Monitoring ENSO in COADS with a seasonally adjusted principal component index*. In: N. NOAA/NMC/CAC, Oklahoma Climate Survey, CIMMS and the School of Meteorology, University of Oklahoma (Ed.), Proceedings of the 17th Climate Diagnostics Workshop: 52–57.

Wolter, K., & Timlin, M. S. (1998). Measuring the strength of ENSO events – how does 1997/98 rank? *Weather*, *53*, 314–324.

Worm, B., Lotze, H. K., & Sommer, U. (2000). Coastal food web structure, carbon storage, and nitrogen retention regulated by consumer pressure and nitrogen loading. *Limnology and Oceanography*, *45*(2), 339–349. https://doi.org/10.4319/lo.2000.45.2.0339

Wright, J. F. (1995). Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. *Australian Journal of Ecology*, *20*(1), 181–197. https://doi.org/10.1111/j.1442-9993.1995.tb00531.x

Yates, A. G., & Bailey, R. C. (2010). Selecting objectively defined reference sites for stream bioassessment programs. *Environmental Monitoring and Assessment*, *170*(1), 129–140. https://doi.org/10.1007/s10661-009-1221-1

Yoder, C. O., & Rakin, E. T. (1998). The role of biological indicators in a state water quality management process. *Environmental Monitoring and Assessment*, *51*, 61–88. https://doi.org/10.1023/A:1005937927108

How to cite this article: Coates JH, Schiff K, Mazor RD, Pondella DJ II, Schaffner B, Whiteman E. Development of a biological condition assessment index for shallow, subtidal rocky reefs in Southern California, USA. *Mar Ecol*. 2018;39:e12471. https://doi.org/10.1111/maec.12471