Spring 2021

Development of Predictive Bioassessment Indices of Non-Perennial Streams and Rivers in the Arid Southwestern United States

Andrew M. Caudillo

Development of predictive bioassessment indices of non-perennial streams and rivers in the arid Southwestern United States

A Thesis

Presented to the

Faculty of the

Department of Applied Environmental Science

California State University Monterey Bay

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Environmental Science

By: Andrew M. Caudillo

Advisor: Dr. John R. Olson

Committee Members: Dr. Raphael D. Mazor and Dr. Jennifer M. Duggan

2021

The Undersigned Faculty Committee Approves the

Thesis of Andrew M. Caudillo:

Development of predictive bioassessment indices of non-perennial streams and rivers in the arid Southwestern United States

X____

John R. Olson, PhD

Environmental Science

X____

Jennifer M. Duggan, PhD

Environmental Science

X Raphael D. Mazor

Raphael D. Mazor, PhD

Southern California Coastal Watershed Research Project

X____________________________________ May 26, 2021

Daniel Shapiro, Dean

Dean of Graduate Students

Acknowledgements

We would like to thank the Environmental Protection Agency for funding for this project. We also want to thank Dr. Kai Chen for his insight and guidance.

Abstract

Freshwater systems are a limited resource and must be managed to maintain or restore their ecological health. Bioassessments, which use the biota at a site to draw conclusions on a system's ecological health, are commonly applied to freshwater systems. Freshwater bioassessment programs are typically only applicable in perennial systems (those which have surface water continually) and do not yet have an accepted role in assessing non-perennial streams in their dry phase. Although dry phase non-perennial streams that have a significant hydrological nexus with traditionally navigable waters navigable waters are protected under the state of California and national legislations, they cannot be assessed with bioassessment currently.

We sampled 106 dry streams in the arid southwestern United States and developed indices of taxonomic completeness (i.e., Observed to Expected or O/E indices) in dry streams to assess the effect of anthropogenic stress at these sites. We did this for three key assemblages in dry streams: channel-dwelling arthropods, riparian vegetation-dwelling arthropods, and bryophytes. We also explored different definitions ways of identifying reference sites and parameters related to index development to assess their effects on O/E indices' performance. O/E indices using channel arthropods were the most responsive and sensitive to human activity, regardless of which reference definition or probability of capture threshold was used. Channel arthropods were the most responsive likely because they respond in a predictable way to stress. Vegetation-dwelling arthropods and bryophytes were absent at up to one quarter of all sites, which negatively affected index performance. The absence of these assemblages at reference sites yielded less responsive and unsuccessful indices developed with these assemblages.

The ecological status of dry streams can be determined from indices of taxonomic completeness when using channel-dwelling arthropods. Other taxa may not occur consistently enough across sites to produce responsive indices. Vegetation-dwelling arthropods should be omitted from future O/E studies because of their lack of response to stressors in stream channels. Further research should be conducted on moss and their response to local stressors because of their response to stress is difficult to predict.

Table of Contents

Table of Figures

Table of Tables

Appendices

Introduction

Government agencies and land managers assess freshwater systems and draw conclusions on their ecological integrity to inform actions taken to protect freshwater resources. The inferences about ecological integrity are based on a comparison of an assessed system with a benchmark standard, referred to as the reference condition (Stoddard et al. 2006). A freshwater system is considered stressed when certain human activities affect an aspect of a system's ecological integrity, for example increased salts from agricultural runoff changing water chemistry from natural levels (Stoddard et al. 2006). Regulators and land managers need to assess the condition of a site to know appropriate management actions to take, depending on the intended uses of the waterbody.

Assessment of perennial systems is often performed via a bioassessment, which is the process of using biological data to estimate a system's ecological integrity (Barbour et al. 1999). Bioassessment informs decision-makers who enforce standards of the Clean Water Act, Porter-Cologne Act, and state Dredge-and-Fill policies (Bailey et al. 2004). A common class of bioassessment indices in the Western United States are predictive indices (Moss et al. 1987), which use models to estimate expected reference conditions at a site based on environmental characteristics. One predictive index, the predictive Multi-Metric Index (MMI), entails creating metrics and comparing their scores at reference and stressed sites to assess ecological structure (Karr 1981; Robinson et al. 2019). Another predictive index, the Observed to Expected Index (O/E), estimates taxonomic completeness, or the proportion of expected taxa at a site that are observed (Moss et al. 1987; Hawkins and Carlisle 2001). Bioassessments are very effective at estimating ecological integrity in perennial systems but are still being developed for dry non-perennial systems.

Non-perennial rivers and streams (NPRS) make up approximately 80% of global river networks (Larned et al. 2010) but a lack of accepted assessment methodologies for them has left NPRS largely understudied when dry (Larned et al. 2010) and are difficult to sample when flows stop. When flows of NPRS cease, these sites are often excluded from sampling efforts. As climates become increasingly arid, perennial systems will likely receive less surface water and shift to dry, non-perennial states ("dry streams"), which will be problematic for land managers when bioassessments are necessary (Steward at al. 2018; Datry et al. 2014). Although dry streams in California are protected under the Clean Water Act and Porter-Cologne Act, there is a lack of information on how dry streams are affected by stressors because they are ignored during bioassessments (Steward et al. 2018). The need to develop assessment methodologies for dry streams has been increasing (Mazor et al. 2014; Datry et al. 2016), especially in arid areas such as the North American Southwest.

Multi-Metric Indices have recently been developed for dry streams (Robinson et al. 2019), but MMIs have conceptual limitations**.** Although MMIs are effective bioassessment indices, they typically are expressed in terms that a broad audience cannot understand without background information on the relationship between index scores and ecosystem functions (Schoolmaster et al. 2012)**.** Observed to Expected scores are intuitive in their interpretation, given that they are simply ratios between the expected number of taxa and the taxa that are present, and can therefore be understood as measures of taxonomic completeness. Management agencies may be able to better understand, act on, and communicate O/E scores in reports rather than MMI scores because of their simplicity.

Observed to Expected indices provide an estimate of ecosystem integrity that complements MMI measurements of ecological integrity. Multi-Metric Indices estimate the effect of a stressor on the ecological structure of a site and can be more responsive to that stressor when the index is calibrated with sites that are affected by it. In contrast, O/E indices do not use disturbed sites for calibration but instead measure the impact of stress on the entire assemblage by directly measuring the deviation from reference condition, in terms of taxa lost. An O/E index is based on a predictive model that estimates the number of taxa expected at a site if that site was in reference condition. The number of taxa expected are then compared to the number actually observed to estimate taxonomic completeness at a site. The model in an O/E index uses natural environmental variation as predictors and grouped taxa based on taxonomic similarity as responses. The number of expected taxa that are observed at reference sites is then divided by the number of expected taxa to yield O/E index scores. The index is then applied to samples collected at sites to estimate their ecological condition where this is not known. O/E indices are effective when the expected number of taxa is high (e.g., 10), but when expected values are low (e.g., below 8), they typically do not perform well (Mazor et al. 2016).

The use of local-scale and landscape-level predictors, as opposed to only landscape level predictors in bioassessment indices has been considered in perennial bioassessments (Lunde et al. 2013). Although landscape-level predictors (e.g., watershed-level) are easily accessible for analyses, local-scale stressors may be a better choice for defining reference condition. When using landscape-level predictors, local stressors are ignored, but local predictors typically reveal if a site is truly within reference condition. The use of local-scale predictors may be more effective at defining the best minimally disturbed reference sites than landscape-level predictors (Lunde et al. 2013). For example, if a site's watershed integrity score were very high, but the concrete channel running through the site was ignored, this reference site would not truly be within reference condition. In dry streams, using local-scale predictors, along with landscape-level predictors, may be more suitable than using solely landscape-level predictors because landscape-level predictors that correlate to hydrology and flowing water will not always be pertinent to a dry stream.

In perennial systems, most O/E indices are developed with only a single assemblage, but multiple assemblages often can reflect ecological integrity (e.g., invertebrates and algae) because all assemblages respond to stress in some way (Theroux et al. 2020). Perennial predictive indices developed with multiple assemblages may be more responsive to stress than a single-assemblage index (Theroux et al. 2020), and this may also be true for dry stream O/E indices.

Using multiple bioassessment indices can lead to a more robust assessment than using just one index (Mazor et al. 2016). Multi-Metric Indices and O/E indices complement one another because they use biota to characterize ecological conditions differently, but the interpretation of site conditions based on their scores may not always agree. For example, an invasive species could alter taxonomic completeness of a site, but this species may not necessarily have a negative effect on the ecological structure there. Due to the ability of an O/E index to respond to deviations from reference conditions, along with an MMI's potential increase in sensitivity when calibrated with disturbed sites, both indices can be combined to develop a hybrid index (for example, the California Stream Condition Index [CSCI], Mazor et al. 2016). Using an approach similar to the CSCI (dry stream condition index, DSCI (Robinson et al. 2019)) in dry streams could better capture the ecological integrity than one MMI or O/E could capture because they measure ecological integrity in complementary ways.

Objective

The objective of this study was to develop and assess the performance of an O/E index for dry streams in the Southwestern United States. If a widely accepted assessment method for dry streams were to be developed, fewer streams would be excluded from assessments, resulting in more effective assessment by managers (Steward et al. 2018). To carry out this objective, we followed the methods established by Robinson et al. (2019) to sample macroinvertebrates and bryophytes in dry streams to provide the data needed to develop an O/E index. We aim to create a tool for management entities to use for assessing dry streams in the arid Southwest that could potentially be used in a hybrid index, like the CSCI, providing a more robust assessment than would be provided by using the MMI alone.

Methods

Study area

The study area was in the arid Southwestern United States ranging from Phoenix, Arizona to arid regions of California (e.g., southern and central coast, deserts, and central valley). (Figure 1). California is characterized as a combination of desert and Mediterranean climates depending on latitude, whereas Arizona is largely characterized by desert climates. The range of precipitation for sites from Arizona was 172mm to 583mm, and the range of precipitation for the sites from California was 85mm to 652mm. The ecoregions we sampled were a combination of deserts with an urban landscape surrounding them, arid regions with little disturbance. Both ecoregions were dominated by drought-resistant vegetation and woody shrubs and as coastal regions in southern California, all dominated by drought-resistant vegetation and woody shrubs. We chose 58 sites across the two states that were representative of the different habitats and topographies in the ecoregions.

Using previously collected data

In addition to the 58 sites we sampled, we also used the data from 60 sites collected in 2016 and 2017 from Robinson et al. (2019). The samples sites from Robinson et al. (2019) were in the coastal area of southern California, largely in the San Diego area.

Field data collection

We sampled during June to August of 2018. We followed the field collection protocol described in detail by Robinson et al. (2019). We combined the data collected for this study with Robinson et al. (2019).'s data into a single dataset that we used to develop an O/E index.

Following the methods of Robinson et al. (2019) we designated a representative 160-m reach at each site, which we separated into eight 20-m sections. In each section, we collected channel and vegetation-dwelling arthropods using ramped pitfall traps (Figure 2) and a canvas bag, respectively (Robinson et al. 2019). We used ramped pitfall traps because not only do they reduce disturbance to the habitat, they are also more suitable for sampling in stream beds because of the hard substrates typically found in stream channels (i.e., cobbles or concrete) that make digging pitfall traps impractical (Pearce et al. 2005; Patrick and Hanson 2013). The traps were filled with approximately 250 mL of propylene glycol (antifreeze) and left out for 24 hours to collect both diurnal and nocturnal macroinvertebrates. Samples were stored in jars, using the propylene glycol as a preservative.

Vegetation dwelling arthropods were collected on plants in or near the channel, following Robinson et al.'s methodology of choosing the plant that appeared to offer the best invertebrate habitat in each section. We wrapped a portion of the plant in a 1-m^2 canvas bag and hit it a total of 30 times (Robinson et al. 2019), using a plastic pipe to dislocate any vegetation-dwelling arthropods. The contents of the bag were placed in a jar and preserved with 70% ethanol.

Along with arthropods, we also collected bryophytes (moss) at each site, which were collected using a floristic approach (Newmaster et al. 2005; Robinson et al. 2019). We designated three mesohabitats at each site (i.e., right and left banks and the channel) and a variety of microhabitats within each mesohabitat (e.g., soil, rock, log, etc.) (Robinson et al. 2019). We followed Robinson et al. (2019)'s protocol of searching for and collecting moss with time-constrained searches (20 minute searches and 12 minute collection times in each meso-habitat). We collected up to a total of five samples of moss from each mesohabitat, collecting them by hand in a pattern from most diverse to least diverse patches in each microhabitat present (Robinson et al. 2019).

We measured aspects of physical habitat such as channel bankfull depth, sediment size, and slope for each transect in the sample reach, following the Surface Water Ambient Monitoring Program (SWAMP) protocol (Robinson et al. 2019). We also recorded any stressors observed at each site. Using stressor categories such as fire breaks, walking paths, and other anthropogenic disturbances, we assigned a categorical value to each stressor based on how prevalent it is. These values ranged from “Not present” to or “over 75% of the reach”.

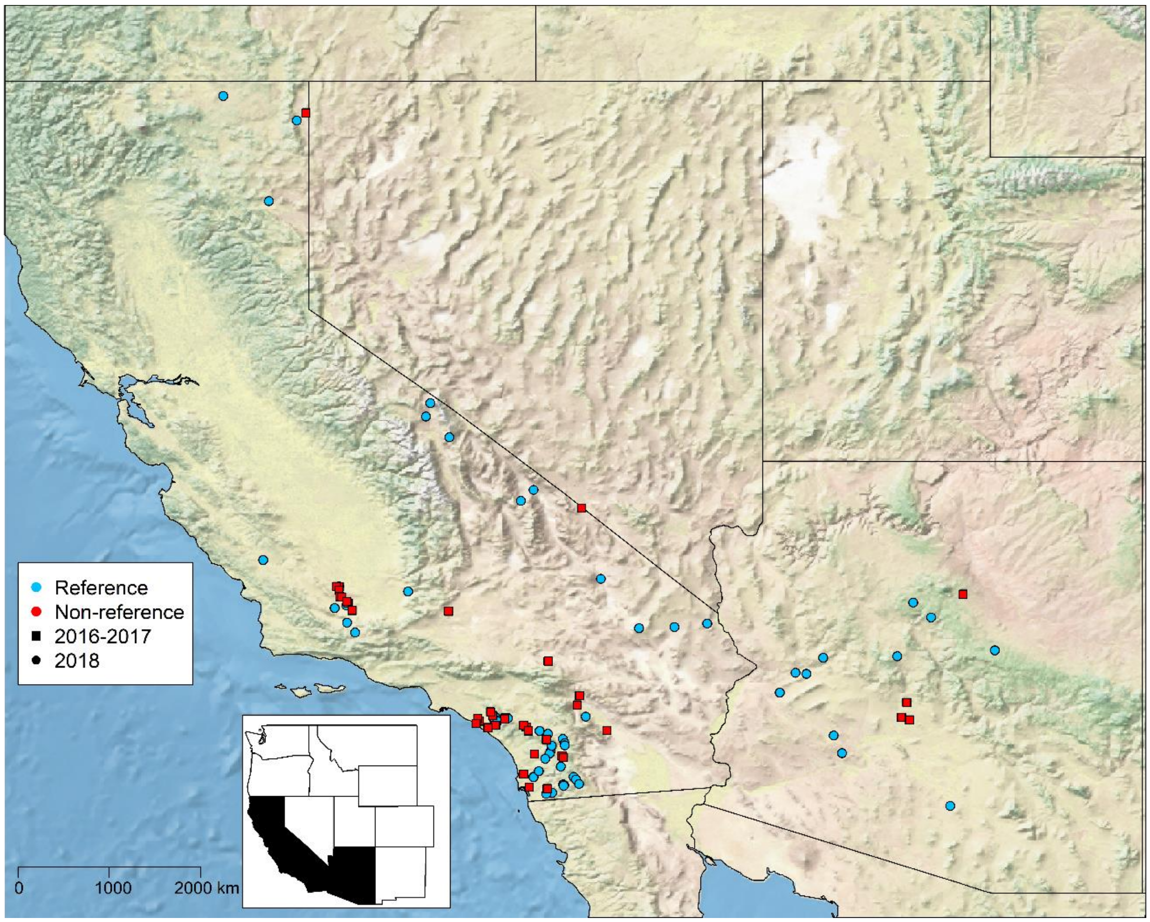


Figure 1. Map of study sites ranging from Modoc, California to Phoenix, Arizona.

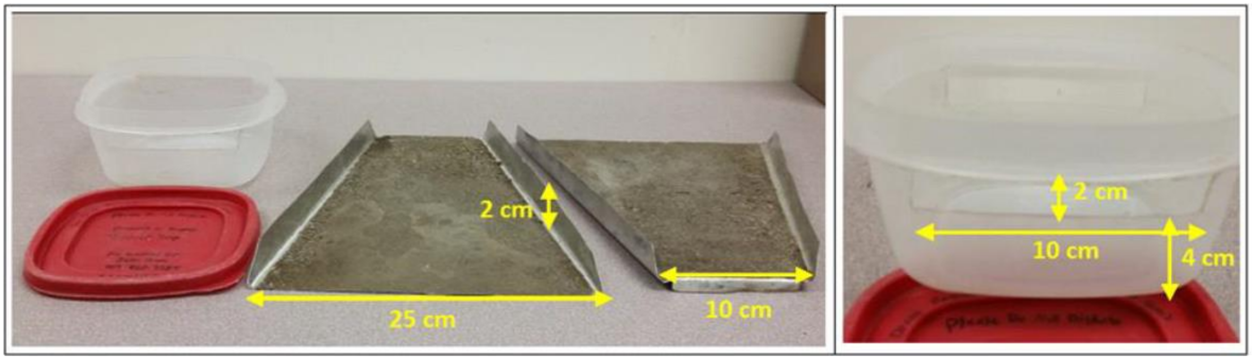


Figure 2. Schematics of pitfall traps

Laboratory methods

We identified the arthropods to taxonomic order or family and identified the moss to genus, which were all later characterized into morphospecies (Robinson et al. 2019). Morphospecies are characterized by physical differences and allow for greater taxonomic resolution than can be achieved with available keys. Morphospecies may be differences between genus or species and are commonly used in terrestrial studies (Oliver and Beattie 1996). For every distinct morphospecies, we took photographs and described the characteristics that made them unique for consistency and repeatability of identifications.

Reference sites and reference definitions

In predictive indices, reference sites are often chosen using watershed characteristics, but other spatial scales may be more influential on a site's ecological condition than the watershed scale in dry streams. For this study, we developed two forms of reference site definitions: 1) the catchment and watershed scales (referred to as watershed) and 2) the catchment, watershed, and local scales (referred to as local). We adapted the thresholds established by Mazor et al. (2019) for reference sites within each spatial scale we considered. For the local scale, we weighed local impacts based on the amount of harm we expected each type of impact to cause on dry stream biota and adapted the local thresholds of proximity and intensity established by Mazor et al. (2019). Local stressors were weighted by both distance from channel and a stressor category (e.g., walking path, armoring of channels, etc.) (Appendix A). Stressors that were innocuous, not very harmful, moderately impactful, or very impactful toward biota received a weight of 0, 1, 3, or 6, respectively (Appendix A). Stressors were weighted from 3 to 1 in increments of 0.5 based on the distance to the channel (in channel = 3, within 5-m of channel = 2, within 50-m of channel= 1.5, within 100-m of channel = 1, within 250-m of channel = 0.5) (Appendix B). Reference sites passed every threshold established by Mazor et al. (2019), regardless of the version of reference definition we used. Non-reference sites in both reference definitions did not pass at least one reference threshold. Degraded sites were classified as those having a local stress score more than 23 and an Index of Watershed Integrity score less than 0.6.

Index development

We standardized taxonomic information in the dataset by establishing operational taxonomic units (OTUs) for each assemblage (Hawkins et al. 2000). OTUs not only standardize taxonomic levels for each taxonomic group, but also are used to calculate the number of observed and expected taxa in an O/E index. The decision to lump or drop taxa depends on the relative amount of information at each taxonomic level. For example, if 100 individuals are identified to morphospecies and 10 are only identified to family, there is more taxonomic information in the morphospecies, thus the family level individuals should be dropped from analyses. Alternatively, if 100 individuals were identified to family and only 10 were identified to morphospecies, the 10 morphospecies could be lumped up to family. Ambiguous taxa (e.g., damaged or immature taxa) were left at taxonomic order or class.

Once we established OTUs, we performed a hierarchical cluster analysis on a Sørensen dissimilarity index of the reference site taxa. The type of cluster analysis we used was a flexible beta analysis, where beta was set at -0.5. The clusters represent groups of reference sites with similar taxonomic composition. We used these clusters to identify environmental patterns that likely cause taxa presence (or absence) in the predictive modeling steps.

The O/E index requires two probabilities to estimate the E value for each OTU (Moss et al. 1987; Hawkins and Carlisle 2001): the probability of membership of a site being in a cluster group (P_m) and the probability of an OTU occurring at a site within a cluster group (P_o). Once the P_m and P_o are determined, a third probability, the probability of capture (P_c) of that OTU at that site is calculated. To determine the P_m, we used a random forests model, where the clusters were the response variables and natural environment characteristics are the predictors. Every OTU's P_o for each cluster group is calculated as the number of sites within each cluster group that an OTU occurs in divided by the number of sites within that cluster group. The P_c for each OTU at a site is then calculated by summing the products of each site's P_m and P_o for all groups. Last, E is calculated as the sum of all P_cs per site that are greater than a predetermined threshold that removes rare and accidentally collected taxa. In many O/E indices, taxa that have a 50% chance or higher of being captured at a site are typically used (e.g., a threshold of 0.5) (e.g., Hawkins and Carlisle 2001; Ostermiller and Hawkins 2004; Chen et al. 2019), but there is no standard threshold. Additionally, thresholds of 0 are used to when developing O/E indices to determine how accounting for rare taxa effect the resulting index performance. We developed O/E indices using both thresholds to determine how threshold choice affected performance.

The predictors we used in the random forest models were a combination of environmental variables available from the Environmental Protection Agency's *StreamCat* dataset (Hill et al. 2016) and WorldClim's *BioClim* dataset (https://worldclim.org/data/v1.4/formats.html). The predictive models used in O/E indices were developed using the statistical program, R; specifically, the package "randomForests" (R Core Team 2020). The randomForests machine-learning algorithm is based on the generation of many decision trees (forest), each generated from different subsets of training data. A large number (specifically, 1500) of training subsets are randomly selected with replacement (bootstrapping), and an unpruned decision tree is developed for each subset. When used to predict categorical response data, every tree generates a vote for an outcome, and every row of data receives proportions of total votes for a response variable. The proportions of votes can be used in analyses like an O/E index as the probability of a site occurring in a cluster. Certain factors can be tuned within the randomForests function to create a more sensitive model, such as the number of decision trees to create and the number of variables to assess at each tree node, but except for the number of trees created we used the default values (Liaw and Weiner 2002).

Although single-assemblage models are traditionally used in O/E indices, they may not always capture all responses to stressors because different assemblages respond to stressors differently (Theroux et al. 2020). Multi-assemblage O/E indices are not commonly developed but may indicate ecological condition better than a single-assemblage index can (Theroux et al. 2020). Because different assemblages likely respond to different stressors, an O/E index using multiple assemblages may be more sensitive to a wider range of stressors than a single-assemblage index. We developed multi-assemblage O/E indices in two ways, by 1) summing all observed and expected numbers of taxa across each single-assemblage index and calculating O/E scores using their sums ("Combined index") and 2) applying the single-assemblage development method to OTUs from all assemblages composited together ("Multi-Assemblage index").

Evaluating model performance

When developing an O/E index, performance can be assessed via precision, accuracy, responsiveness, and sensitivity. Precision and accuracy determine how well an index is calibrated with reference sites, whereas responsiveness and sensitivity measure how well an index can respond to or detect deviations from reference condition. We developed a total of 40 O/E indices to explore the range of results as we varied each component of an O/E index (5 assemblages x 2 reference definitions x 2 P_c thresholds x (1 predicted model and 1 null model)) (Figure 3). To assess each index's precision, we 1) compared the standard deviation of scores at reference sites to both the standard deviation of scores from a null model (i.e., a model that does not account for natural environmental variation among sites) and the standard deviation of scores from the theoretical best possible model (Van Sickle et al. 2005) and 2) determined which indices' E values explained the largest amount of variation in O (i.e., which models have R^2 values closer to 1). The accuracy of the O/E indices were assessed by 1) comparing the mean of the index to 1 at reference sites because a score of very close to or at 1 indicates that a site is within reference condition (e.g., O and E are the same value) and 2) comparing the slope of a regression of O against E to 1, as well as comparing the y-intercept to 0. Another measure of accuracy was calculated via a random forest model where the reference site O/E scores were used as response variables with the same predictors during index development. An index was considered minimally biased if it had very low (5% or less) variation in scores at reference sites explained by the environmental predictors.

To assess an index's response to deviations from reference condition, we calculated the model's sensitivity and responsiveness. The sensitivity of an O/E index was measured as the percent of test sites (those which are not reference condition, but not heavily degraded) classified as non-reference by the O/E index, where the threshold of reference was the 10^{th} percentile of reference site O/E scores. If a test site received an O/E score greater than the 10^{th} percentile of reference site scores, then the site was considered to be within reference condition. Responsiveness was assessed by applying the O/E index to degraded sites and evaluating the ability of the index to distinguish reference and degraded sites via comparing the means of the degraded and reference O/E scores via a Student's t-test. As another aspect of responsiveness, we used reference and degraded site O/E scores in a random forest model with anthropogenic predictors (via StreamCat, Hill et al. 2016) to determine the percent variation explained by the predictors.

Results

Ranges of expected values

Across all O/E indices, we saw expected values ranging from 0 to 40 (Figure 5). We only calculated expected values of 0 for the moss and vegetation-dwelling arthropod assemblages when using the P_c threshold of 0.5. When the P_c was very low for all OTUs at a site (e.g., less than 0.5), no expected taxa were predicted. The sites that received O/E scores of 0 were reference sites that passed all reference filters. When dividing an observed number of taxa by an expected value of 0, we overwrote the final O/E score as 0 because we never observed an O/E score of 0 when $E > 0$. We did not omit sites from our analyses that had O/E scores of 0 because we considered these sites to be within reference condition.

O/E indices can be developed for NPRS but not all are successful

Across all 40 indices, single-assemblage indices developed with channel arthropods were the best performing in responsiveness and sensitivity, regardless of the reference definition or P_c used (Table 1). Channel arthropod indices also performed well in terms of accuracy and precision. Indices for vegetation-dwelling arthropods and moss did not perform well with respect to responsiveness, sensitivity, or accuracy because the expected number of taxa was too low (Appendix I). Both multi-assemblage indices combatted the issue of low expected number of taxa and performed well in accuracy and precision. The multi-assemblage indices varied in performance when applied to non-reference sites. Indices developed with the local reference definition were slightly more responsive than indices developed with the watershed reference definition. When applied to degraded sites, approximately half of the indices were responsive and somewhat sensitive ($t > |2|$ and $x > 10\%$, respectively) (Table 1, Figure 5).

Vegetation-dwelling arthropod and moss indices were the only indices to be affected by the issue of low expected values. The Bray-Curtis dissimilarity of the moss and vegetation-dwelling arthropod assemblages was very high (e.g., 1), showing that many sites had no OTUs in common. Although the average species richness for moss and vegetation-dwelling arthropods was approximately 6 and 9, respectively (Appendix H), the expected values for these assemblages were less than 3 at 87% of reference sites used in indices developed with a $P_c \geq 0.5$. The low expected values were caused by the reference sites having P_ms between 30% and 49%, which resulted very few taxa being predicted as present.

Assemblage absence at sites affects index performance

The average richness of OTUs varied greatly across all sites (min = 2, max = 19, Appendix H). Moss and vegetation-dwelling arthropods were the only assemblages that were absent at some sites. Moss were absent from 24 of the 106 sites. Sites with no moss collected varied from very sandy substrate to very rocky sites with no soil in the channel. Vegetation-dwelling arthropods were absent from 5 of the 106 sites. Sites with no vegetation-dwelling arthropods were those where the healthiest vegetation was grass. Two sites were not sampled for vegetation-dwelling arthropods and three sites did not have vegetation-dwelling arthropods present. Unforeseen issues in the field, such as our equipment being stolen in Bakersfield, CA, also led to vegetation sampling being infeasible.

The absences of vegetation-arthropods and moss affected the performance of indices developed with these assemblages. Observed to Expected scores for vegetation-dwelling arthropod indices in degraded sites were largely greater than 1, which was because the expected values were always very low, usually half the number of observed taxa. Observed to expected scores for moss indices varied between the watershed and local reference definitions, but three of the four moss indices had 50% of their O/E scores between 0 and approximately 1.3, showing that degraded site expected values for moss varied.

Local reference definition yields better indices than watershed reference definition

Both reference definitions yielded approximately the same number of responsive indices (e.g., those with t- values > |2|) (Table 1). The responsiveness of indices developed with the local reference definition was generally higher than that of indices developed with the watershed reference definition. Indices developed with the local reference definition were also more sensitive than indices developed with the watershed reference definition. Although the level of sensitivity of indices varied between reference definitions, channel arthropod indices were always the most sensitive to deviations in reference condition.

Multi-assemblage indices and their performance

The multi-assemblage indices performed better than any single-assemblage index during calibration, but when we applied these indices to non-reference sites, they did not perform particularly well. The R^2 and slope of the multi-assemblage indices were generally higher than any single-assemblage index, showing that the expected values were closer to the observed values than the other indices (Figure 4). However, both multi-assemblage indices were largely unresponsive and insensitive when applied to non-reference sites (Table 1; Figure 5). The Combined Index was the only responsive multi-assemblage index and was responsive in both the watershed and the local reference definitions (Table 1).

Effect of probability of capture thresholds on index performance

Regardless of reference definition, indices developed with $P_c \geq 0.5$ performed better than $P_c \geq 0$ (Table 1). The precision of indices developed with a $P_c \geq 0$ were always higher than indices developed with a $P_c \geq 0.5$ (Table 1). The R^2 of indices developed with a $P_c \geq 0$ were also lower than those of indices developed with $P_c \geq 0.5$, except for channel arthropod indices (Table 1; Figure 4). Also, when applied to non-reference sites, indices developed with a $P_c \geq 0.5$ were generally more responsive than indices developed with a $P_c \geq 0$, regardless of reference definition (Table 1). However, indices developed with a $P_c \geq 0$ were generally more sensitive than those developed with a $P_c \geq 0.5$ (Table 1).

Null and predicted indices

Regardless of reference definition, all predicted indices performed better in the calibration stages than their null counterparts (Table 1). Although two of the null indices performed better when applied to non-reference sites than their predicted counterparts, these null indices were imprecise and less responsive than the predicted indices. Due to null indices being imprecise, we did not pursue them further.

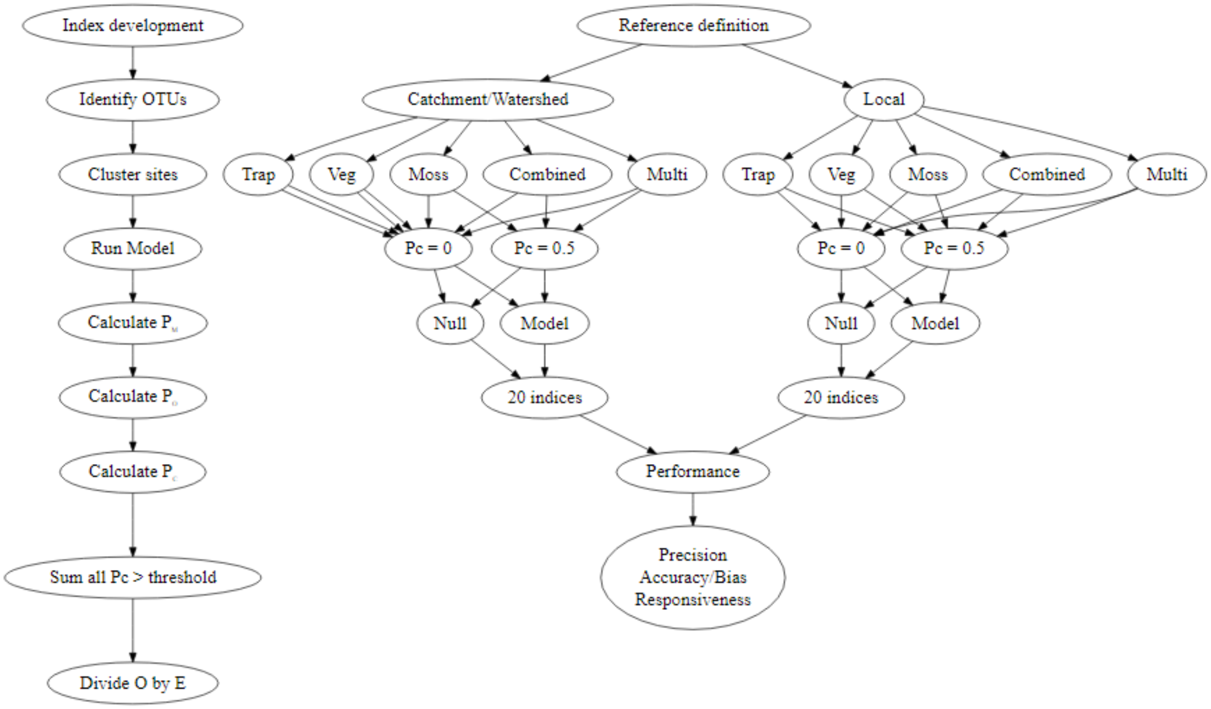


Figure 3. Flow chart describing the methods for developing an O/E index (left) and the methods we performed in this study (right).

Table 1. Calibration and application performance of each index. Bold values represent more precise and responsive indices for standard deviation and responsiveness, respectively. RSSDs with N/A are those assemblages that did not have enough expected taxa to generate an accurate value. Bold values are in the successful range of O/E indices

	Index	Mean	SD	RSSD	Responsiveness (p)	Responsiveness (R^2)	Sensitivity (%)	R^2	Slope
Catchment-Watershed	Channel Arthropod	1.01	**0.21**	**0.16**	**0.01**	**18.1**	**15**	**0.72**	**1**
	Vegetation-dwelling arthropod	0.92	0.53	N/A	**0.009**	**12.2**	0	0.61	**0.97**
	Moss	0.45	0.64	N/A	0.3	**25.7**	0	0.59	0.85
	Combined	1.01	**0.19**	N/A	0.97	0	8	**0.79**	**1**
	Multi-Assemblage	1.02	**0.19**	**0.14**	0.33	6.3	8	**0.78**	**1.1**
	Channel Arthropod Pc0	0.98	0.25	**0.16**	**0.004**	**18.1**	**35**	**0.70**	**1**
	Vegetation-dwelling arthropod Pc0	0.96	0.47	0.33	**0.01**	0	0	0.33	**1.3**
	Moss Pc0	1	0.85	0.27	0.79	0	0	0.24	**1**
	Combined Pc0	0.98	0.27	N/A	**0**	1.2	**12**	0.5	**1.2**
	Multi-AssemblagePc0	0.97	0.26	**0.13**	0.14	0.15	**15**	0.51	**1.2**
Catchment-Watershed-Local	Channel Arthropod	1.01	**0.2**	**0.16**	**0.01**	**15.7**	**22**	**0.70**	**1**
	Vegetation-dwelling arthropod	0.95	0.47	N/A	0.3	**20.4**	**30**	0.49	**0.98**
	Moss	0.7	0.7	N/A	0.88	**39.2**	0	0.53	0.92
	Combined	1.01	**0.21**	N/A	**0**	2.6	0	**0.75**	**1**
	Multi-Assemblage	1	**0.21**	**0.13**	0.43	0	8	**0.70**	**0.99**
	Channel Arthropod Pc0	0.98	0.27	**0.17**	**0.006**	**19.8**	**30**	**0.70**	**1**
	Vegetation-dwelling arthropod Pc0	0.97	0.45	0.30	**0.02**	0	8	0.23	**1.3**
	Moss Pc0	0.99	0.79	0.39	0.33	0	0	0.28	**1**
	Combined Pc0	0.98	0.27	N/A	**0**	3.8	**30**	0.47	**1.2**
	Multi-AssemblagePc0	0.85	**0.22**	**0.13**	0.16	1.3	**19**	0.22	**1.1**

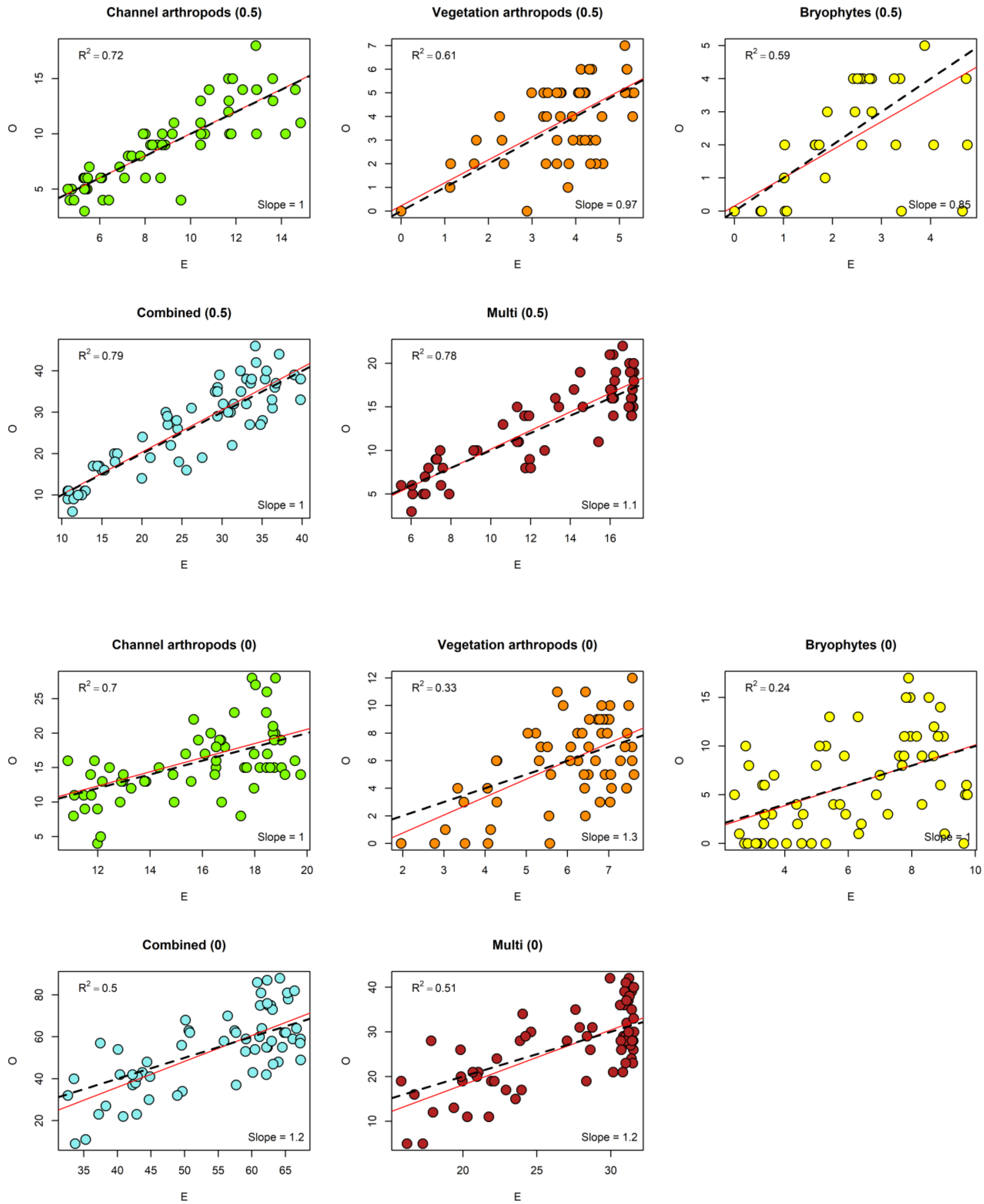


Figure 4. Calibration performance of all indices using the watershed reference definition. Assemblages used in each index are above each plot and the Pc threshold is shown in parentheses next to the index assemblage. Dotted lines represent a 1:1 line, while red lines represent slope.

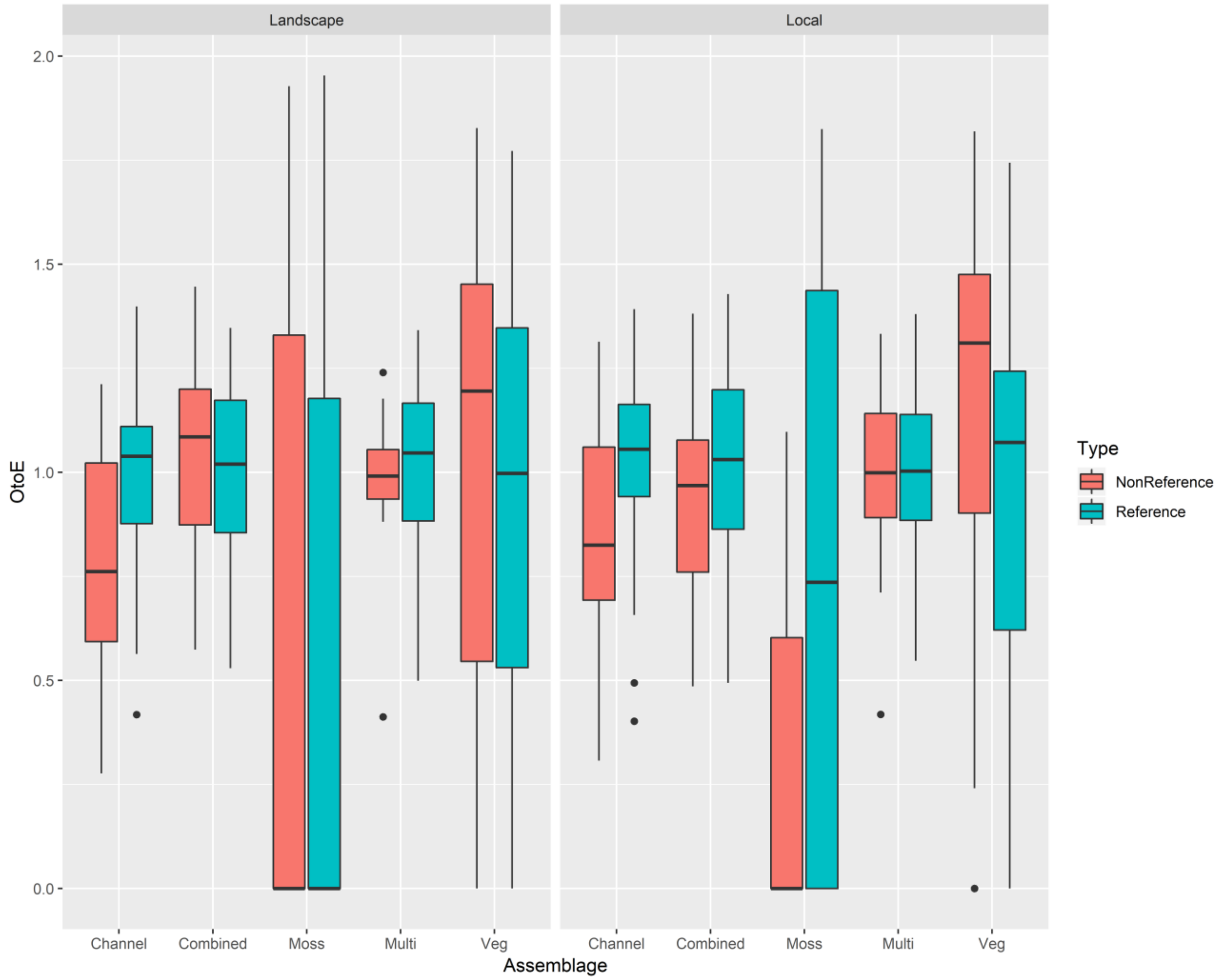


Figure 5. Distribution of reference and degraded O/E scores of indices developed with a $P_c \geq 0.5$.

Discussion

O/E indices developed with channel arthropods are effective for dry streams

We successfully developed and assessed O/E indices for dry streams and found that O/E indices for NPRS perform just as well as O/E indices for perennial streams. Multi-assemblage indices should be avoided because of the difficulty to predict multiple assemblages simultaneously, but certain single-assemblage indices should continue to be used. The channel arthropod indices were the most successful, regardless of reference definition, but other indices were also responsive. The standard deviations of the successful indices fell in the range of other responsive O/E indices used in perennial streams, showing that O/E indices can be developed for and applied to NPRS and are as successful as those in perennial streams. The standard deviations of the best channel arthropod indices were within the range of successful O/E indices for perennial streams (SD = 0.17 to 0.21, Chen et al. 2019).

The issues of low E and high O/E scores

The poor responsiveness of the vegetation-dwelling arthropod and moss indices was caused by imprecision in the two indices. The imprecision of the vegetation-dwelling arthropod and moss indices was likely caused by two factors 1) low expected values of taxa (between 0 and 6) and 2) some assemblages' lack of response to stress. Moss and vegetation-dwelling arthropods were absent at up to a quarter of reference sites, yielding low expected scores for these assemblages. Low expected scores in vegetation-dwelling arthropods and moss (e.g., between 0 and 0.55) resulted in high O/E scores (e.g., O=2, E=1.1). Reference sites with O/E scores higher than 1 can likely be dropped from future analyses because of low richness. The absence of moss may be caused by the loose sandy substrates and rocky, uneven substrates at some sites we sampled, which do not provide suitable substrate for moss. The fact that bryophytes were generally absent from these types of sites suggests that bryophytes are sensitive to local environmental conditions (e.g., median substrate size). Many of the sites we sampled were naturally sandy, suggesting that moss indices will be hard to develop because O/E indices developed with moss are only effective at sites where moss are expected. Observed to Expected indices developed for moss may be inaccurate in predicting taxonomic completeness because 1) observed moss richness was generally higher than the expected value and 2) moss may not be responding to stress in a consistent way (i.e., as taxonomic loss). Coysh et al. (2000) also found Observed to Expected scores greater than 1 at reference sites and suggested that more taxa are observed than expected because of a temporary increase in food or nutrient availability. For dry streams, increased nutrient availability may be degraded and reference sites, depending on the assemblage in question. While some of the disturbances we observed are likely to lead to loss of bryophyte taxa (e.g., small substrate size due to recent channel alteration), other disturbances appeared to increase richness by increasing the amount of suitable habitat (e.g., soil compaction and dried oil in the channel). Vegetation-dwelling arthropod indices were also imprecise, which may have been caused by variability in the quality of the vegetation present at site. For example, some plants, such as grasses, which may be naturally occurring at a site, do not provide a productive habitat for some arthropods. If a site has low-quality vegetation (e.g., grass or dead shrubs), the vegetation-dwelling arthropods will not be present there. When considering taxa responsiveness to stress, vegetation-dwelling arthropods may not be directly affected by the stressors present in a stream but may be influenced more by the quality of the vegetation they inhabit.

The low expected values in vegetation-dwelling and moss indices were caused by two potential issues. The first issue is that richness at reference sites for moss and vegetation-dwelling arthropods may be low naturally (Figure 6). For example, the richness at reference sites for moss and vegetation dwelling arthropods was less than 6 taxa. The second issue is that the cluster groups that we calculated may not have been very distinct (Appendices D and E), resulting in predictions of the random forest models to decreasing in precision. The expected value, which relies on P_m, was likely reduced because of poor predictions from the random forest models. Although random forests found important predictors for each assemblage, the important predictors may have been a proxy for predictors we did not measure (e.g., soil moisture).

The increased dissimilarity of moss and vegetation-dwelling arthropods between reference sites suggests little overlap of the composition of these assemblages at reference sites. Theroux et al. (2020) found that soft-bodied algae had very high Bray-Curtis dissimilarity values between reference sites, and that the resulting indices were less successful than others their study. Theroux et al. (2020) also suggests that high Bray-Curtis dissimilarity occurs in sites with high dispersal rates of individuals, which may be applicable for vegetation-dwelling arthropods, since many of their morphospecies we identified were Coleopterans and Hemipterans, which disperse by flight. The dissimilarity of moss between reference sites is likely caused by the habitats that moss prefer (e.g., rocky or compacted substrate) and the amount of various nutrients in soils at reference sites.

Multi-assemblage index benefits and drawbacks

The multi-assemblage indices were better calibrated than any single-assemblage index. The multi-assemblage indices are likely performing well for different reasons from one another. The first reason is the nature of the Combined index: which is a sum of all observed and expected scores across indices. Mazor et al. (2016) also saw that the combination of scores (e.g., O/E and MMI scores in the CSCI) accounts for individual variability across taxa to yield increased estimates of ecological integrity. The second reason is related to the Multi-assemblage index, which combats the issue of low number of expected taxa at a site. When all assemblages are combined at a site before being predicted, the number of expected taxa at a site would rarely be low because there are multiple assemblages at a site to contribute toward expected taxa. Theroux et al. (2020) also found that when combining assemblages, an index will be better calibrated than a single-assemblage index and have higher expected values than a single-assemblage index.

The multi-assemblage indices did not perform well when applied to non-reference sites for the same reason. When combining multiple assemblages into a single index, the responses of each assemblage to stress are treated the same, which is not necessarily true because bryophytes and arthropods respond to stress in different ways. Some disturbed habitats, like oil-fused sand, may be simultaneously suitable substrate for mosses, but harmful to arthropods.

The unresponsiveness of most of the multi-assemblage indices is likely a function of how channel arthropods are responsive to stress, while moss and vegetation-dwelling arthropods are either less responsive to stress or not responsive to stress in a way we can predict, respectively. When combining a responsive index with two poor indices, the resulting multi-assemblage indices varied in responsiveness depending on the reference definition used. The varying responsiveness was cause by the different assemblages' response to stress because the three assemblages we used do not respond to stress in the same way. When combining multiple assemblages, the Multi-Assemblage O/E index cannot always distinguish reference and stressed sites, making it unreliable for estimating the condition of NPRS. Additionally, the vegetation-dwelling arthropods and moss assemblages had absences of all OTUs at approximately one quarter of all sites in this study, which negatively affected the Combined index's ability to distinguish reference from non-reference sites because its combined O/E scores were not always comprised of three, non-zero numbers for each site.

The effects reference definitions on index performance

Using the local reference threshold in addition to watershed reference threshold improved the performance of O/E indices. In both reference definitions, the channel arthropod indices were responsive to stress, but indices developed with the local reference definition were more sensitive and generally more responsive (Table 1), suggesting that local reference definitions, in addition to watershed definitions, should be used when choosing reference sites. The change in responsiveness between indices developed with the two different reference definitions is likely due to local stressors being a more accurate representation of site conditions, compared to conditions of watershed or catchment level stressors.

Assemblages in NPRS are likely largely affected by local stress because stress at a landscape-level could be less influential toward them than stressors present in the stream bed. Channel-dwelling arthropods are likely most responsive to local (and potentially catchment-level) stress because of how they enter a dry stream. Terrestrial arthropods colonize NPRS channels via their riparian habitats (Corti and Datry 2015) and are subject to the stressors present in these areas. The stress present in riparian areas could explain the loss of taxa at stressed sites because if terrestrial arthropods are not entering NPRS because of their absence at a stressed source habitat, their observed values in an NPRS channel would be lower than at a reference site NPRS channel.

Although the local reference definition produced more responsive indices, there were some costs to using this definition. Local stressors may not be known from a site unless the site is assessed in the field. If studies were to use landscape-level stressors without assessing a site, some sites that would be incorrectly classified as within reference condition because local stressors would be ignored. Local stressors removed over 15 sites from our watershed reference definition site pool. The loss of reference sites would be considered a disadvantage if the number of sites in a dataset are low, like in this study. If the number of reference sites in a study are low, then the local reference definition will remove too many sites, thus decreasing the representativeness of reference condition characteristics (Ode et al. 2016), making calibrating an O/E index difficult. To develop an index that is capable of accounting for natural environmental variables that influence assemblages, reference sites should be represented via a variety of geographic areas (i.e., the variety of sites in this study) and a variety of environmental conditions (i.e., the varying conditions at each site) (Ode et al. 2016; Chen et al. 2019)

Some local disturbances, like foot traffic, may impact a site less than others, such as All-Terrain Vehicle usage. To reflect these differences in impact on the channel arthropods, a weighting system should be used. When assigning weights to a stressor, overweighting a stressor should be avoided.

The effects of differing P_c thresholds on index performance

The O/E indices we developed cannot predict rare taxa well, but the omission of rare taxa lead to the indices being better at responding to disturbance. Indices developed with a $P_c \geq 0$ were more responsive than those developed with a $P_c \geq 0.5$ (as also found by Van Sickle et al. 2007 for perennial streams), except for the Combined indices developed with a $P_c \geq 0$ (Table 1). Vander Laan and Hawkins (2014) and Ostermiller ad Hawkins (2004) found the opposite of our results regarding P_c thresholds. Although indices developed with a $P_c \geq 0$ were more sensitive than those developed with a $P_c \geq 0.5$, they were not calibrated well. The increased sensitivity of indices developed with a $P_c \geq 0$ was due to rare taxa and common taxa being measured together at each site, resulting in more taxa responding to stress. Indices developed with a $P_c \geq 0$ lead to an imprecise representation of taxa at reference sites because the presence and absence of all taxa were being predicted.

The use of $P_c \geq 0.5$ is a common practice for perennial O/E indices, but for dry streams $P_c \geq 0$ is more effective for combatting the issue of low expected values than standard practice. When predicting assemblages where few OTUs were found at a site, like vegetation-dwelling arthropods and moss, many of the OTUs present had a P_c less than 0.5, yielding expected values of 0. If an assemblage does not have low expected values, like the channel-dwelling arthropods, then a $P_c \geq 0.5$ could continue being used, but for assemblages with low expected values, using 0.5 as a threshold for P_c should not be pursued.

Potential of the DSCI

The MMI scores and O/E scores for 98 sites largely agreed (Table 2). Following the methods of Mazor et al. (2016), a DSCI could be developed for NPRS and applied throughout the western United States. The benefits of using the DSCI would be effective for dry streams because an MMI would identify specific assemblages that respond to a stressor and O/E scores would estimate when a site is no longer taxonomically complete. The DSCI has the potential to be used by land managers when performing bioassessments of dry streams.

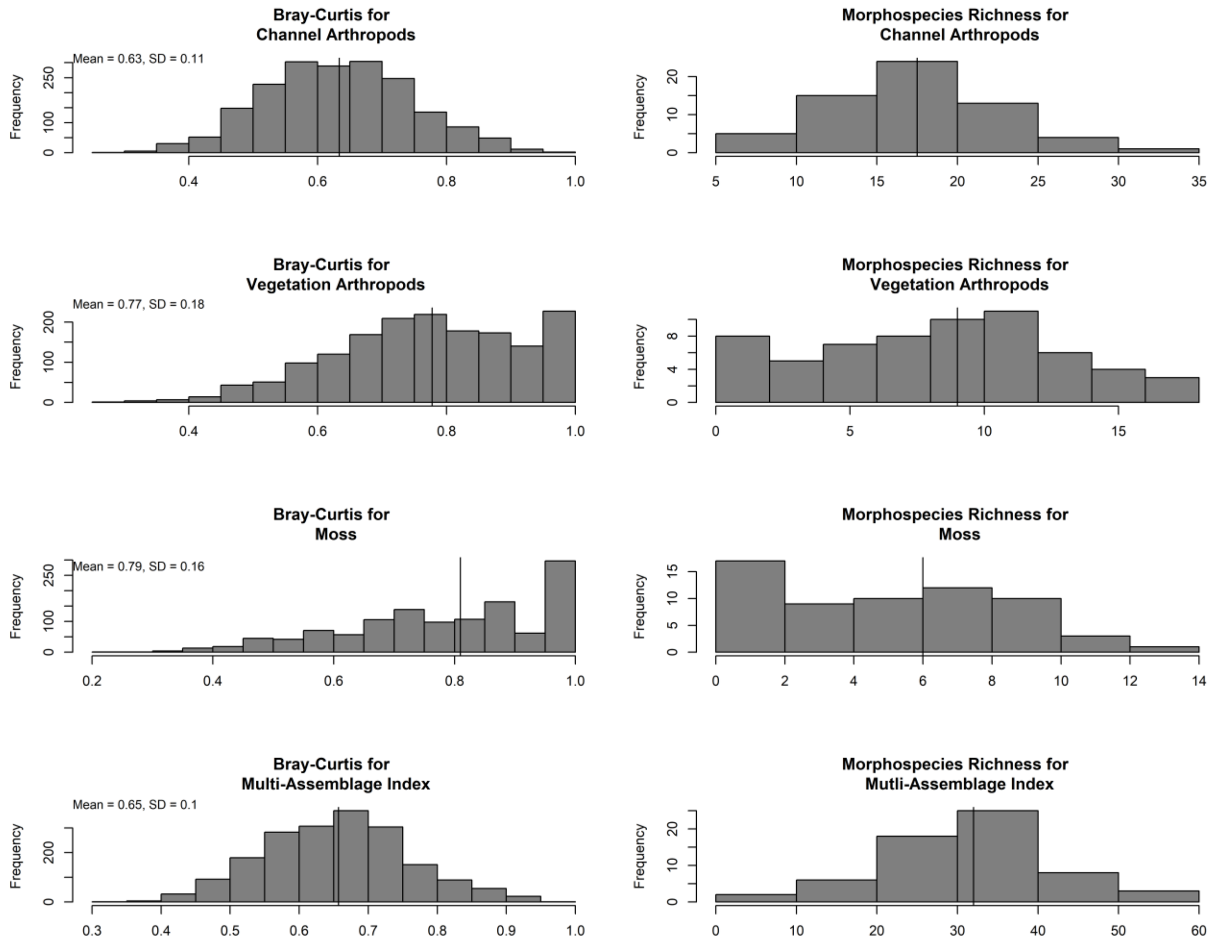


Figure 6. Bray-Curtis dissimilarity scores and morphospecies richness for channel arthropods, vegetation-dwelling arthropods, moss, and the Multi-Assemblage Index. Values shown for indices developed with a P_c > 0.5.

Conclusion

Our results indicate that indices of taxonomic completeness can be developed for NPRS, but not for every assemblage. Assemblages such as moss may need to either be omitted from future indices or they should be further studied to understand their response to stress. Future studies should omit vegetation-dwelling arthropods from indices of taxonomic completeness because of their inconsistent occurrence across sites or investigate sampling methods that will reduce the variability of this assemblage. Multi-assemblage indices should be avoided because of how different assemblages respond to stress. Single-assemblage indices should continue to be used in dry streams. Channel arthropod O/E indices are the best option to assess the ecological condition of dry streams because indices developed with them performed well in terms of precision, responsiveness, and sensitivity.

Literature Cited

Bailey RC, Norris RH, Reynoldson TB. Bioassessment of Freshwater Ecosystems. 2004:1–15. doi:10.1007/978-1-4419-8885-0_1

Barbour MT, Gerritsen J, Snyder, BD, Stribling, J.B. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99 002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Chen K, Olson JR, Vander Laan JJ, Hill RA, Wang B, Hawkins CP. Improving the performance of ecological indices by balancing reference site quality and representativeness. Hydrobiologia. 2019;837(1):177–194. doi:10.1007/s10750-019-3970-3

Corti R, Datry T. Terrestrial and aquatic invertebrates in the riverbed of an intermittent river: parallels and contrasts in community organisation. Freshwater Biology. 2015;61(8):1308–1320. doi:10.1111/fwb.12692

Coysh J, Nichols S, Ransom G, Simpson J, Norris R, Barmuta L, Chessman B. 2000. AUSRIVAS macroinvertebrate bioassessment: predictive modelling manual. CRC for Fresh-water Ecology, Canberra, Australia. (Available from: http://ausrivas.ewater.com.au/index.php/manuals-a-datasheets)

Datry T, Larned ST, Tockner K. Intermittent Rivers: A Challenge for Freshwater Ecology. BioScience. 2014;64(3):229–235. doi:10.1093/biosci/bit027

Datry T, Pella H, Leigh C, Bonada N, Hugueny B. A landscape approach to advance intermittent river ecology. Freshwater Biology. 2016;61(8):1200–1213. doi:10.1111/fwb.12645

Hawkins CP, Norris RH, Hogue JN, Feminella JW. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications. 2000;10(5):1456–1477.

Hawkins CP, Carlisle DM. Use of predictive models for assessing the biological integrity of wetlands and other aquatic habitats. Bioassessment and management of North American freshwater wetlands. 2001:59–83.

Hill RA, Weber MH, Leibowitz SG, Olsen AR, Thornbrugh DJ. The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. JAWRA Journal of the American Water Resources Association. 2016;52(1):120–128. doi:10.1111/1752-1688.12372

Karr JR. Assessment of Biotic Integrity Using Fish Communities. Fisheries Magazine. 1981;6(6):21–27.

Larned S, Datry T, Arscott DB, Tockner K. Emerging concepts in temporary-river ecology. Freshwater Biology. 2010;55(4):717–738. doi:10.1111/j.1365-2427.2009.02322.x

Liaw A, Weiner M. Classification and Regression by randomForest. R News. 2002;2(3):18–22.

Lunde KB, Cover MR, Mazor RD, Sommers CA, Resh VH. Identifying Reference Conditions and Quantifying Biological Variability Within Benthic Macroinvertebrate Communities in Perennial and Non-perennial Northern California Streams. Environmental Management. 2013;51(6):1262–1273. doi:10.1007/s00267-013-0057-1

Mazor RD, Stein ED, Ode PR, Schiff K. Integrating intermittent streams into watershed assessments: applicability of an index of biotic integrity. Freshwater Science. 2014;33(2):459–474. doi:10.1086/675683

Mazor RD, Rehn AC, Ode PR, Engeln M, Schiff KC, Stein ED, Gillett DJ, Herbst DB, Hawkins CP. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshwater Science. 2016;35(1):249–271. doi:10.1086/684130

Mazor RD, Olson, JR, Robinson MD, Caudillo AM, Brown JS. 2019. Assessing the biological condition of dry ephemeral and intermittent streams: Technical Report 1089. Southern California Coastal Water Research Project. Costa Mesa, CA. Available from (https://ftp.sccwrp.org/pub/download/DOCUMENTS/TechnicalReports/1089_Ephemeral Streams.pdf)

Moss D, Furse MT, Wright JF, Armitage PD. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology. 1987;17(1):41–52. doi:10.1111/j.1365-2427.1987.tb01027.x

Newmaster SG, Belland RJ, Arsenault A, Vitt DH, Stephens TR. The ones we left behind: Comparing plot sampling and floristic habitat sampling for estimating bryophyte diversity. Diversity and Distributions. 2005;11(1):57–72. doi:10.1111/j.1366-9516.2005.00123.x

Ode PR, Rehn AC, Mazor RD, Schiff KC, Stein ED, May JT, Brown LR, Herbst DB, Gillett D, Lunde K, et al. Evaluating the adequacy of a reference-site pool for ecological assessments in environmentally complex regions. Freshwater Science. 2016;35(1):237–248. doi:10.1086/684003

Oliver I, Beattie AJ. Invertebrate Morphospecies as Surrogates for Species: A Case Study. Conservation Biology. 1996;10(1):99–109. doi:10.1046/j.1523-1739.1996.10010099.x

Ostermiller JD, Hawkins CP. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society. 2004;23(2):363–382. doi:10.1899/0887-3593(2004)023<0363:eoseob>2.0.co;2

Patrick LB, Hansen A. Comparing ramp and pitfall traps for capturing wandering spiders. Journal of Arachnology. 2013;41(3):404–406. doi:10.1636/hi12-52.1

Pearce JL, Schuurman D, Barber KN, Larrivée M, Venier LA, McKee J, McKenney D. Pitfall trap designs to maximize invertebrate captures and minimize captures of nontarget vertebrates. The Canadian Entomologist. 2005;137(2):233–250. doi:10.4039/n04-029

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson, MD. 2019. Development of biological indicators for the dry phase of non-perennial rivers and streams [*SNS Master's Theses]*. 41.

Schoolmaster DR, Grace JB, William Schweiger E. A general theory of multimetric indices and their properties. Methods in Ecology and Evolution. 2012;3(4):773–781. doi:10.1111/j.2041-210x.2012.00200.x

Steward AL, Negus P, Marshall JC, Clifford SE, Dent C. Assessing the ecological health of rivers when they are dry. Ecological Indicators. 2018;85:537–547. doi:10.1016/j.ecolind.2017.10.053

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH. SETTING EXPECTATIONS FOR THE ECOLOGICAL CONDITION OF STREAMS: THE CONCEPT OF REFERENCE CONDITION. Ecological Applications. 2006;16(4):1267–1276. doi:10.1890/1051-0761(2006)016[1267:seftec]2.0.co;2

Theroux S, Mazor RD, Beck MW, Ode PR, Stein ED, Sutula M. Predictive biological indices for algae populations in diverse stream environments. Ecological Indicators. 2020;119:106421. doi:10.1016/j.ecolind.2020.106421

Vander Laan JJ, Hawkins CP. Enhancing the performance and interpretation of freshwater biological indices: An application in arid zone streams. Ecological Indicators. 2014;36:470–482. doi:10.1016/j.ecolind.2013.09.006

Van Sickle J, Hawkins CP, Larsen DP, Herlihy AT. A null model for the expected macroinvertebrate assemblage in streams. Journal of the North American Benthological Society. 2005;24(1):178–191. doi:10.1899/0887-3593(2005)024<0178:anmfte>2.0.co;2

Van Sickle J, Larsen DP, Hawkins CP. Exclusion of rare taxa affects performance of the O/E index in bioassessments. Journal of the North American Benthological Society. 2007;26(2):319–331. doi:10.1899/0887-3593(2007)26[319:eortap]2.0.co;2

Appendices

Appendix A. Weights used to determine local stress scores.

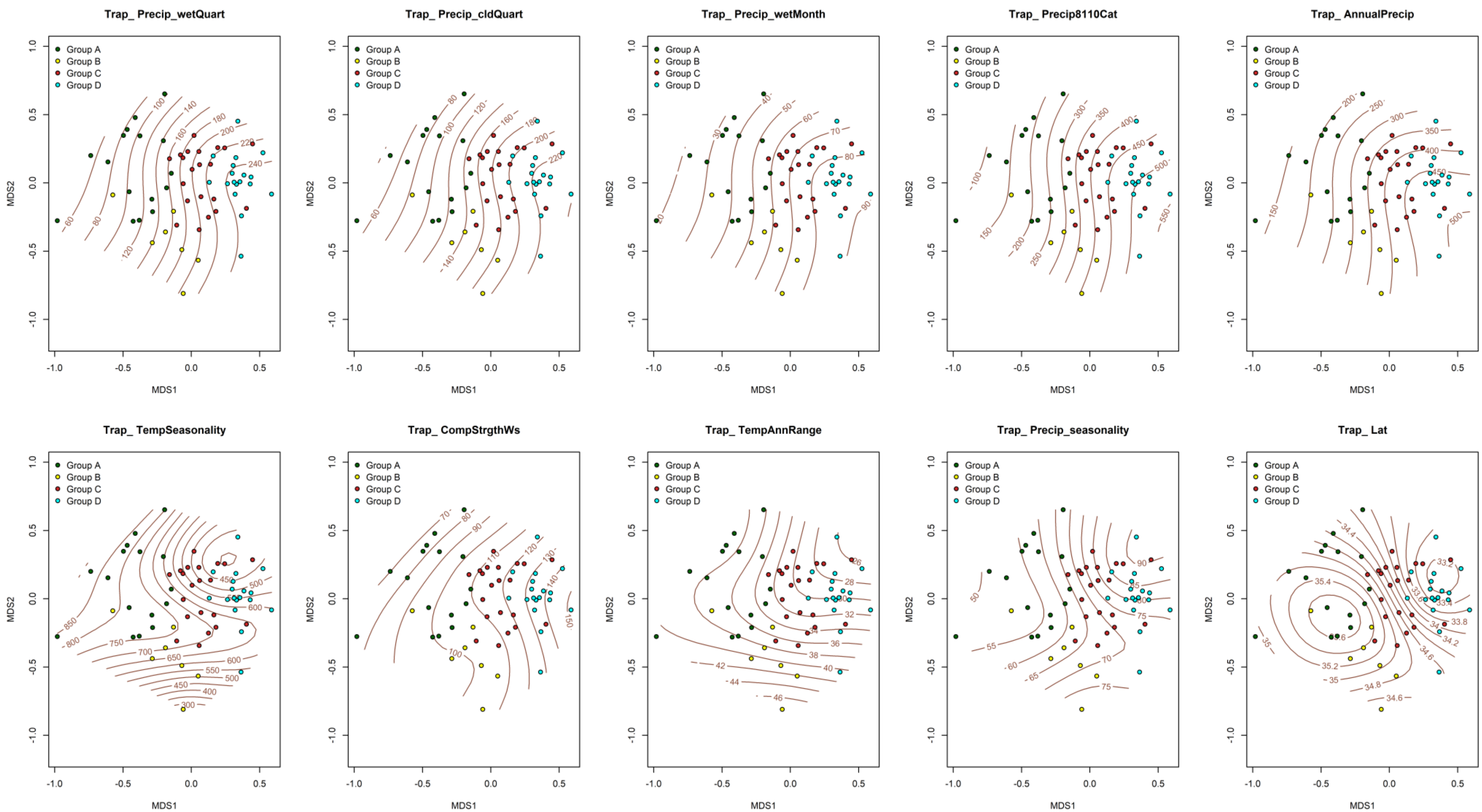
Analyte	Weight
Fire Breaks Proximity	1
Mowing/Cutting Proximity	1
Burns Proximity	0
Cattle Grazing Proximity	1
Invasive Plants Proximity	0
Animal Burrows Proximity	0
Industrial Proximity	6
Landfill Proximity	6
Mining Proximity	6
Military Land Proximity	6
Urban Commercial Proximity	6
Urban Residential Proximity	6
Heavy Urban Other Proximity	6
Suburban Residential Proximity	6
Rural Residential Proximity	1
Golf Course/Parks/Sports Fields Proximity	1
Excessive Human Visitation Proximity	1
Light Urban Other Proximity	6
Crops Irrigated Proximity	6
Crops Non-Irrigated Proximity	6
Vineyards Proximity	6
Timber Harvest Proximity	6
Orchards Proximity	6
Hay Proximity	1
Fallow Fields Proximity	1
Dairies Proximity	1
CAFOs Proximity	6
Pasture Proximity	1
Rangeland Proximity	1
Agricultural Other Proximity	6
Highway >2 lanes Proximity	6
Paved Roads Proximity	1
Unpaved Roads Proximity	1
Parking Lot/Pavement Proximity	6
Railroad Proximity	6
Air Traffic Proximity	0
Walking Path Proximity	0

Transportation Other Proximity	6
Point Source Discharges Proximity	6
Acid Mine Drainage Proximity	6
Noxious Chemical Odors Proximity	6
Industrial Water Quality Other Proximity	6
Non-Point Source Discharges Stormwater Proximity	6
Trash/Dumping Proximity	1
Vector Control Proximity	6
Urban Water Quality Other Proximity	6
Agricultural Runoff Proximity	6
Algal/Surface Mats/Benthic Algal Growth Proximity	0
Direct Septic/Sewage Discharge Proximity	6
Excess Animal Waste Proximity	6
Nutrient Related Water Other Proximity	6
High Concentration of Salts Proximity	6
Flow Diversions Proximity	1
Groundwater Extraction Proximity	1
Unnatural Inflows Proximity	1
Water Control Actions Other Proximity	1
Dike/Levee Proximity	1
Ditches/Canals Proximity	1
Dam Proximity	1
Weirs Proximity	1
Spring Boxes Proximity	1
Water Control Features Other Proximity	1
ATVs Proximity	1
Mountain Bikes Proximity	1
Horses Proximity	1
Excavation Proximity	1
Grading/Compaction Proximity	1
Feral Pig Disturbance Proximity	1
Sediment Disturbance Other Proximity	1
Passive Input (Construction/Erosion) Proximity	1
Debris Lines/Silt-Laden Vegetation Proximity	1
Excess Sediment Input Other Proximity	1
Rip-Rap/Armored Channel bed/bank Proximity	1
Obstructions (culverts, paved stream crossings) Proximity	1
Hardened Features Other Proximity	1

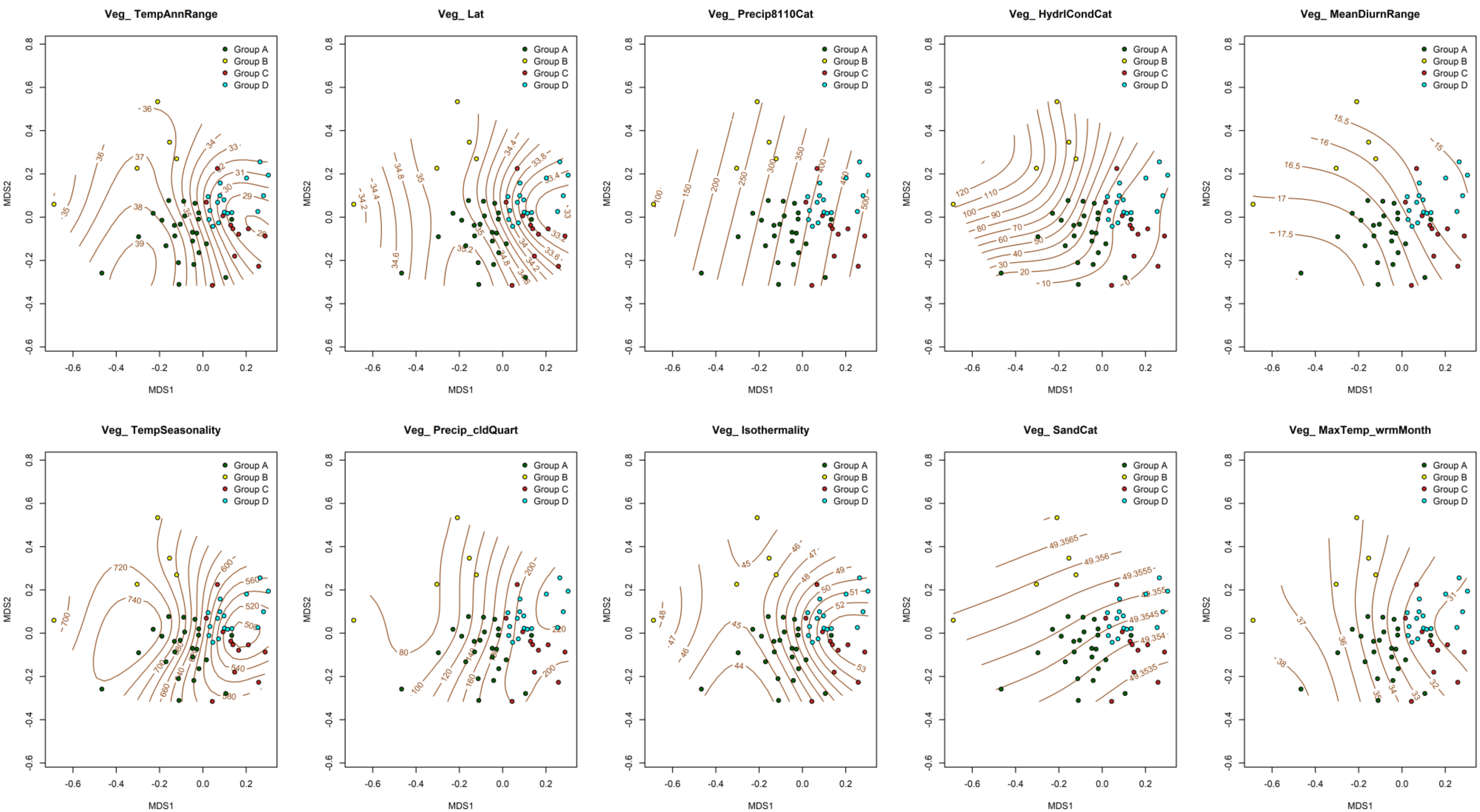
Appendix B. Proximity-based scores local stressors received.

Distance	Scores
In Channel	3
Within 5m of channel	2
within 50m of channel	1.5
within 100m of channel	1
≥250m of channel	0

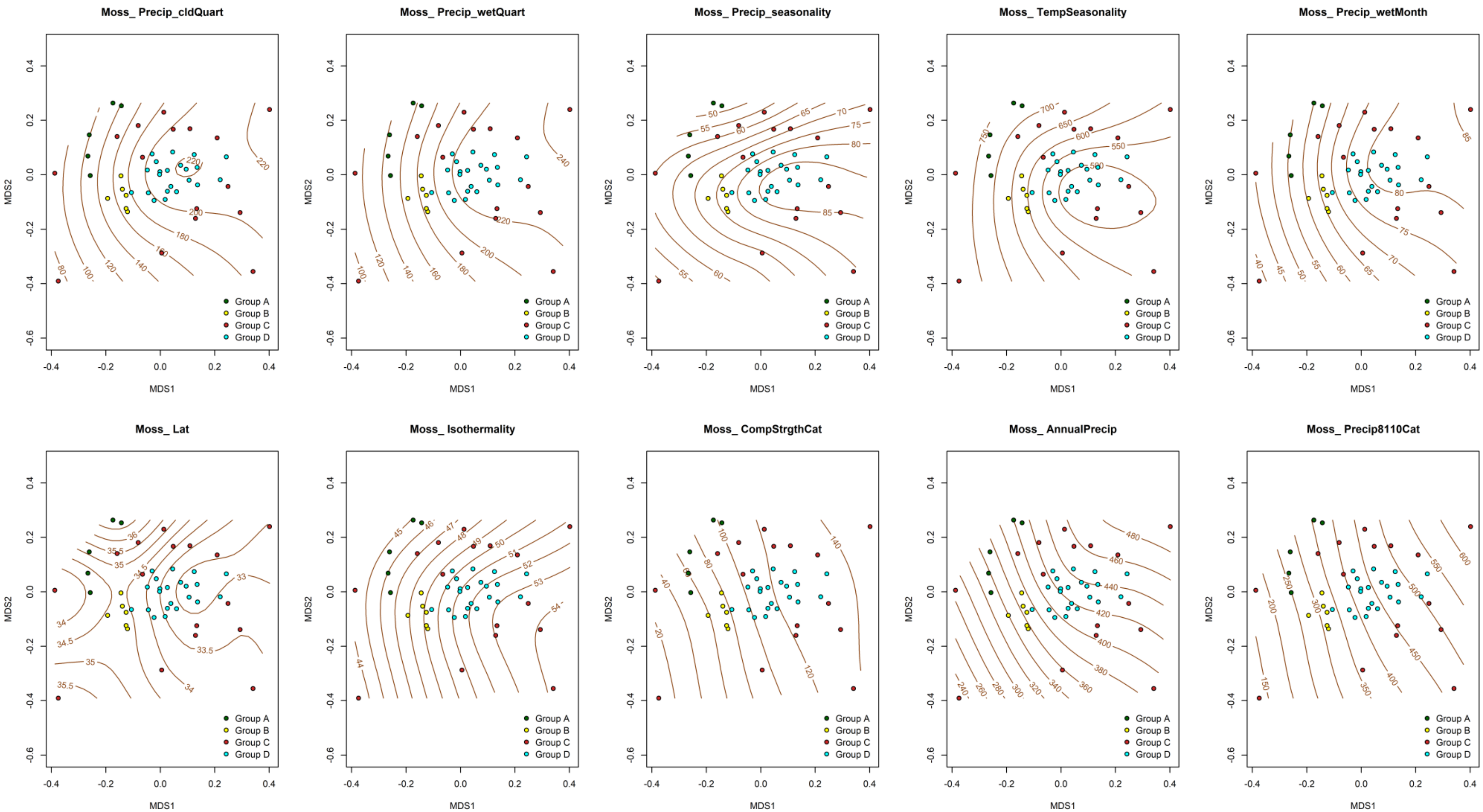
Appendix C. NMDS plots of channel arthropod reference site OTUs in the watershed reference definition. Fitted variables are the top 10 predictors via randomforest.



Appendix D. NMDS plots of vegetation-dwelling arthropods in the Watershed reference definition. Fitted variables are the top 10 predictors via randomforest.



Appendix E. NMDS plots of moss in the Watershed reference definition. Fitted variables are the top 10 predictors via randomforest.



Appendix F. Codes used in Appendix G.

StreamCat variable	**Code**
PctAg2006Slp20Cat	A
PctAg2006Slp10Cat	B
PctUrbLo2011Cat	C
PctUrbMd2011Cat	D
PctUrbHi2011Cat	E
Ag+Urban	F
RdDensCat	G
RdCrsCat	H
CanalDensCat	I
MineDensCat	J
PctAg2006Slp20Ws	K
PctAg2006Slp10Ws	L
PctUrbLo2011Ws	M
PctUrbMd2011Ws	N
PctUrbHi2011Ws	O
Ag+Urban	P
RdDensWs	Q
RdCrsWs	R
CanalDensWs	S
MineDensWs	T
Local Score	U
IWI	V

Appendix G. Catchment, watershed, and local thresholds used to determine reference sites. Red cells indicate values that made a site non-reference. See Appendix F for abbreviations. Asterisk denotes a site that we classified differently than its calculated status due to unmeasured stressors near the site.

Station Code	A	B	Total	C	D	E	Total	F	G	H	I	J	K	L	Total	M	N	O	Total	P	Q	R	S	T	Status (Watershed)	U	V	Status (Local)
105TCLL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.1	0.0	0.0	R	1.5	0.98	R
713CRST	0.0	0.0	0.0	6.8	0.2	0.0	7.0	7.0	2.0	0.4	0.0	0.0	0.0	0.0	0.0	1.4	0.1	0.0	1.5	1.5	0.6	0.1	0.0	0.0	R	6.5	0.96	D
558UnTrb09	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.2	0.1	1.3	1.3	2.3	1.6	0.0	0.0	T	4	0.64	T
558UNHY	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.0	0.0	0.0	R	0	0.55	D
558UBWH	0.0	0.0	0.0	2.4	0.1	0.0	2.6	2.6	2.2	2.0	0.0	0.0	0.0	0.0	0.0	3.6	0.3	0.0	3.9	3.9	2.3	1.6	0.0	0.0	T	10	0.76	T
558BPS01	0.0	0.0	0.0	2.5	1.0	0.4	4.0	4.0	3.4	0.4	0.0	0.0	0.0	0.0	0.0	2.5	1.0	0.4	4.0	4.0	3.4	0.4	0.0	0.0	D	31	0.55	D
557UBMR	0.0	0.0	0.0	2.9	0.1	0.2	3.2	3.2	4.6	2.5	0.0	0.0	0.0	0.0	0.0	1.4	0.1	0.1	1.5	1.5	3.1	1.1	0.0	0.0	D	56	0.76	D
557UCSR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	5.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	5.3	0.0	0.0	T	9	0.66	T
557SCAM	0.0	0.0	0.0	14.1	35.1	1.9	51.1	51.1	8.9	1.1	0.0	0.0	0.0	0.0	0.0	5.0	8.6	0.5	14.1	14.1	4.1	1.4	0.0	0.0	T	4	0.78	T
572UTSA	0.0	0.0	0.0	1.3	0.0	0.0	1.3	1.3	2.4	2.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	1.8	0.8	0.0	0.0	R	12	0.88	T
*557BITR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	3.2	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.2	1.1	0.5	0.0	0.0	D	30	0.78	D
557SVNC	0.0	0.0	0.0	1.9	0.8	0.0	2.7	2.7	1.4	1.8	0.0	0.0	0.0	0.0	0.0	5.4	0.3	0.0	5.6	5.6	2.0	1.8	0.0	0.0	T	22	0.79	D
557THPC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.3	0.0	0.0	R	18.5	0.47	D
312BLCW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.8	0.0	0.0	R	27	0.89	D
312APCY	0.2	0.8	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.0	0.0	R	4.5	0.94	R
317HOGC	0.0	2.6	2.6	0.0	0.0	0.0	0.0	2.6	2.8	0.6	0.0	0.0	0.0	0.2	0.2	0.0	0.0	0.0	0.1	0.2	1.2	0.5	0.0	0.0	R	39	0.64	D
609EMGW	0.0	0.0	0.0	0.0	6.8	0.0	6.8	6.8	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.4	0.4	0.4	0.2	0.0	0.0	R	0	0.97	R
609LSLT	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.2	0.2	0.5	0.1	0.0	0.0	R	4.5	0.98	R
609AMRG	0.0	0.0	0.0	2.3	0.2	0.0	2.6	2.6	1.7	0.4	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.3	0.3	0.7	0.4	0.0	0.0	T	0	0.96	T
605MLGC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.8	0.0	0.0	R	0.5	0.99	R
603QDCY	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	R	0	0.99	R
603CHCY	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.3	0.0	0.0	R	1	0.96	R

637RICE	0.0	0.0	0.0	0.2	0.2	0.0	0.4	0.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.6	0.0	0.0	0.0	R	12	0.97	T
641SAND	1.9	7.2	9.0	0.0	0.0	0.0	0.0	9.0	1.3	0.2	0.0	0.0	0.2	0.7	0.9	0.0	0.0	0.0	0.0	0.9	0.5	0.2	0.0	0.0	T	1.5	0.94	T
641GRNG	0.0	0.0	0.0	0.9	0.0	0.0	0.9	0.9	1.7	2.3	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	1.4	0.5	0.0	0.0	R	12	0.93	T
911TJPC2x	0.0	0.0	0.0	2.8	0.0	0.0	2.8	2.8	2.3	0.8	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.7	0.7	1.6	0.5	0.0	0.0	R	0	0.81	R
911NP9UCW	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.9	0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.9	0.2	0.0	0.0	R	0	0.89	R
911TJKC1x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.3	0.0	0.0	R	1.5	0.91	R
911NP9HTC	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.3	1.6	0.5	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.3	1.6	0.5	0.0	0.0	R	0	0.90	R
911NP9EPC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.5	0.5	0.2	0.0	0.0	R	0	0.91	R
911S01142	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.1	0.0	1.1	1.1	1.3	0.4	0.0	0.0	R	3	0.87	R
905DGCC2x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	R	0	0.89	R
905DGCC1x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.3	0.0	0.0	R	0	0.86	R
905DGSY1x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	1.4	0.5	0.0	0.0	R	9.5	0.84	T
907SRSD1x	0.0	0.0	0.0	0.2	0.3	0.0	0.5	0.5	1.9	0.6	0.0	0.0	0.0	0.0	0.0	0.2	0.3	0.0	0.5	0.5	1.9	0.6	0.0	0.0	T	5	0.79	T
905SDBDN9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.5	0.0	0.0	R	0.5	0.87	R
906SLCFGC	0.0	0.0	0.0	18.3	29.0	17.5	64.8	64.8	4.6	1.1	0.0	0.0	0.0	0.0	0.0	23.8	24.9	14.9	63.7	63.7	6.9	2.5	0.1	0.0	D	8	0.56	D
907NP9OSU	0.0	0.0	0.0	2.9	0.2	0.0	3.1	3.1	0.7	0.1	0.0	0.0	0.0	0.0	0.0	2.3	0.1	0.0	2.4	2.4	0.5	0.1	0.0	0.0	R	7.5	0.95	T
907NP9OSD	0.0	0.0	0.0	2.9	0.2	0.0	3.1	3.1	0.7	0.1	0.0	0.0	0.0	0.0	0.0	2.3	0.1	0.0	2.4	2.4	0.5	0.1	0.0	0.0	R	3.5	0.95	R
908CHI805	0.0	0.0	0.0	13.2	66.4	11.2	90.8	90.8	13.0	1.3	0.0	0.0	0.0	0.0	0.0	13.2	66.4	11.2	90.8	90.8	13.0	1.3	0.0	0.0	D	37.5	0.46	D
907NP9KLC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.1	0.0	0.0	R	6.5	0.81	T
910NP9RJT	0.0	0.0	0.0	5.3	0.7	0.0	6.0	6.0	2.4	3.5	0.0	0.0	0.0	0.0	0.0	5.3	0.7	0.0	6.0	6.0	2.4	3.5	0.0	0.0	T	0	0.72	T
910NP9CCN	0.0	0.0	0.0	0.5	0.2	0.0	0.8	0.8	0.8	0.6	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.2	0.2	0.5	0.2	0.0	0.0	R	0	0.83	R
910SYCAM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.5	0.0	0.0	R	0	0.84	R
910NP9ARP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	2.9	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.4	0.2	0.0	0.0	R	0	0.80	R
907SYCAM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.3	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.3	0.2	0.0	R	0.5	0.88	R
904ESCELN	3.6	5.6	9.2	26.8	40.0	2.4	69.2	78.4	11.2	1.9	0.2	0.0	1.8	2.7	4.5	10.8	15.4	0.9	27.1	31.6	6.1	1.2	0.2	0.0	D	50	0.68	D

903NP9SLR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.6	0.0	0.0	R	0	0.92	R
*903SLFRCx	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.4	0.0	0.0	D	12	0.80	D
903ACPCT1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.3	0.0	0.0	R	0	0.93	R
903CVPCT	0.0	0.0	0.0	0.4	0.1	0.0	0.5	0.5	1.5	0.1	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.2	1.2	0.4	0.0	0.0	R	0	0.85	R
903NP9PRC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	R	0	0.79	R
901TCTCR	0.0	0.0	0.0	24.9	12.0	0.0	36.8	36.8	7.2	3.2	0.0	0.0	0.0	0.0	0.0	1.1	0.4	0.0	1.5	1.5	1.3	0.7	0.0	0.0	D	34.5	0.94	D
901BELOLV	0.0	0.0	0.0	0.9	0.5	0.0	1.5	1.5	1.8	1.2	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.2	0.2	0.3	0.2	0.0	0.0	R	0	0.98	R
901NP9LCC	0.0	0.0	0.0	0.9	0.0	0.0	0.9	0.9	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	2.2	1.5	0.0	0.0	T	0	0.91	T
901SJOF1x	0.0	0.0	0.0	4.1	0.0	0.0	4.1	4.1	3.6	9.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	1.4	0.8	0.0	0.0	T	0.5	0.93	T
901NP9MRC	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	1.3	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.7	0.0	0.0	R	0	0.94	R
901AUDFOX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0.99	R
901NP9CSC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	R	1.5	0.97	R
901LAUREL	0.0	0.0	0.0	2.8	0.3	0.0	3.1	3.1	3.2	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.3	0.0	3.1	3.1	3.2	0.0	0.0	0.0	T	0	0.91	T
901AUDCRW	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	R	0	0.99	R
901SJMS1x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.5	0.0	0.0	R	0	0.95	R
901SJVERD	0.0	0.0	0.0	13.8	4.2	0.0	17.9	17.9	3.1	10.7	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.4	0.4	1.0	1.0	0.0	0.0	T	0	0.93	T
901SJLANV	0.0	0.0	0.0	31.0	26.0	0.3	57.2	57.2	8.1	0.2	0.0	0.0	0.0	0.1	0.1	3.6	2.7	0.1	6.4	6.5	1.7	0.8	0.0	0.0	T	8	0.85	T
801SHDCYN	0.0	0.0	0.0	2.0	1.3	0.0	3.3	3.3	1.4	2.2	0.0	0.0	0.0	0.0	0.0	2.0	1.3	0.0	3.3	3.3	1.4	2.2	0.0	0.0	T	0	0.78	T
902LNGCYN	0.0	0.0	0.0	37.4	41.3	2.6	81.4	81.4	8.7	2.3	0.0	0.0	0.0	0.1	0.1	27.8	28.9	1.0	57.7	57.8	7.8	3.3	0.1	0.0	D	30.5	0.53	D
902WRMSPC	0.0	0.2	0.2	7.9	26.7	8.6	43.2	43.4	7.2	1.1	0.0	0.0	0.0	0.0	0.0	4.5	9.2	0.8	14.4	14.5	3.5	1.2	0.2	0.0	T	17	0.70	T
902PECHNG	0.0	0.1	0.1	5.5	4.3	0.7	10.5	10.6	2.8	1.4	0.6	0.0	0.0	0.0	0.0	1.2	1.0	0.2	2.4	2.4	1.4	0.7	0.1	0.0	T	13.5	0.86	T
902SMAS2x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	R	0	0.98	R
902SMAS1x	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.2	0.0	0.0	R	3	0.94	R
902NP9CWC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	R	0	0.97	R
*801SANT1x	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	2.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.4	0.0	0.0	D	3.5	0.91	D

719DRMT	0.0	0.0	0.0	4.6	7.7	1.3	13.6	13.6	3.9	1.6	0.2	0.0	0.0	0.0	0.0	3.5	5.2	0.7	9.4	9.4	3.3	1.8	0.1	0.0	T	9.5	0.87	T
719COYC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0.99	R
719LOWD	0.0	0.0	0.0	12.5	1.4	0.0	13.8	13.8	4.8	4.4	0.4	0.0	0.0	0.0	0.0	1.9	0.2	0.0	2.0	2.0	2.1	2.2	0.1	0.0	D	7.5	0.49	D
710WAWT	0.0	0.0	0.0	0.9	0.0	0.0	0.9	0.9	1.4	0.9	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.9	0.2	0.0	0.0	R	1	0.95	R
701CUSH	0.0	0.0	0.0	8.0	1.7	0.0	9.7	9.7	3.5	1.6	0.0	0.0	0.0	0.0	0.0	1.4	1.1	0.1	2.6	2.6	1.6	0.7	0.0	0.1	T	4	0.92	T
628COHI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.3	0.0	0.0	R	3	0.99	R
626CAML	0.0	0.0	0.0	0.4	0.0	0.0	0.5	0.5	3.2	1.5	0.1	0.0	0.0	0.0	0.0	0.6	0.1	0.0	0.7	0.7	3.2	2.1	0.2	0.0	T	20	0.92	T
609SCDV	0.0	0.0	0.0	0.8	0.4	0.1	1.3	1.3	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.2	0.2	0.2	0.1	0.0	0.0	R	0.5	0.99	R
719WRDS	0.0	0.0	0.0	22.9	40.4	6.5	69.8	69.8	7.9	0.1	0.0	0.0	0.0	0.0	0.0	3.5	2.5	0.3	6.4	6.4	1.8	0.6	0.0	0.0	T	3.5	0.85	T
311UASL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.3	0.0	0.0	R	1.5	0.64	R
901EMRCYN	0.0	0.0	0.0	12.3	9.2	0.8	22.3	22.3	4.6	0.7	0.0	0.0	0.0	0.0	0.0	6.9	5.2	0.5	12.6	12.6	3.2	0.5	0.0	0.0	T	0	0.85	T
SCSBC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.7	0.4	0.0	0.0	R	3.5	0.97	R
LCWLW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.0	0.0	R	11	0.99	T
MGUCW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.1	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.5	1.7	0.4	0.2	0.0	R	22	0.85	D
BGDMW	0.0	0.0	0.0	5.5	4.1	0.7	10.2	10.2	4.4	0.6	0.0	0.0	0.0	0.0	0.0	14.3	15.4	1.8	31.4	31.4	5.8	0.3	0.0	0.0	T	6.5	0.77	T
MGSKB	0.0	0.0	0.0	40.2	42.5	9.3	91.9	91.9	9.4	1.1	0.0	0.0	0.0	0.0	0.0	8.7	7.3	1.3	17.3	17.3	3.9	0.7	0.0	0.0	D	24	0.84	D
MGUCV	0.0	0.0	0.0	19.7	39.8	2.8	62.3	62.3	8.3	1.1	0.0	0.0	0.0	0.0	0.0	27.7	32.0	1.4	61.1	61.1	7.7	1.4	0.0	0.0	T	3	0.60	T
BWIVW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	R	3	0.99	R
BWBUW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.2	0.0	0.0	R	3.5	0.99	R
BMPLA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.3	0.0	0.0	R	3	0.98	R
CLUCW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.3	0.0	0.0	R	21.5	0.84	T
MGBGBx	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.5	2.0	0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	1.4	0.2	0.0	0.0	R	8	0.96	T
VRUBS001.35	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.2	0.0	0.0	R	0.5	0.96	R
VRSEC002.23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.2	0.0	0.0	R	0	0.97	R

LCSFW	0.0	0.0	0.0	0.0	0.6	0.0	0.6	0.6	1.3	0.0	0.0	0.0	0.0	0.0	0.0	2.5	2.6	0.7	5.8	5.8	2.4	0.4	0.0	0.0	D	42.5	0.91	D
LCWIL018.74	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.1	0.0	0.0	R	15.5	0.94	T
911NP9ATC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	R	1.5	0.97	R

Appendix H. Richness across all sites

Site	Channel	Vegetation	Moss	Total	Average
BMPLA	16	9	0	25	8
BWIVW	13	17	7	37	12
BWBUW	14	10	7	31	10
SCSBC	17	11	2	30	10
LCWLW	10	4	0	14	5
MGUCW	13	0	2	15	5
CLUCW	11	4	6	21	7
719COYC	16	3	5	24	8
312APCY	17	9	6	32	11
311UASL	17	7	8	32	11
312BLCW	18	13	7	38	13
603QDCY	15	5	6	26	9
603CHCY	13	12	0	25	8
641GRNG	15	0	0	15	5
105TCLL	15	5	3	23	8
637RICE	10	11	0	21	7
609LSLT	7	2	0	9	3
713CRST	13	0	0	13	4
710WAWT	17	0	0	17	6
628COHI	17	9	4	30	10
609SCDV	6	1	0	7	2
609EMGW	13	5	3	21	7
605MLGC	15	16	0	31	10
317HOGC	17	9	9	35	12
VRSEC002.23	22	17	7	46	15
MGBGBx	26	11	2	39	13
LCWIL018.74	20	10	4	34	11
VRUBS001.35	23	13	9	45	15
558UNHY	13	1	4	18	6
557THPC	15	8	0	23	8
572UTSA	15	14	5	34	11
901AUDFOX	20	2	9	31	10
901AUDCRW	20	12	13	45	15
901BELOLV	23	8	7	38	13
901NP9MRC	34	11	12	57	19
902SMAS2x	21	5	11	37	12
902SMAS1x	18	7	9	34	11
902NP9CWC	22	9	11	42	14
903ACPCT1	30	15	9	54	18

903CVPCT	29	17	6	52	17
903NP9SLR	29	12	8	49	16
907SYCAM	24	7	8	39	13
907NP9OSD	24	16	10	50	17
907NP9OSU	21	11	6	38	13
910NP9CCN	17	16	3	36	12
910NP9ARP	14	12	3	29	10
910SYCAM	23	13	7	43	14
911NP9HTC	24	9	10	43	14
911NP9EPC	20	7	10	37	12
911S01142	18	8	0	26	9
911TJKC1x	18	8	4	30	10
911NP9UCW	20	5	2	27	9
911TJPC2x	22	11	5	38	13
907NP9KLC	17	12	4	33	11
905DGSY1x	22	6	8	36	12
903NP9PRC	16	14	7	37	12
905SDBDN9	22	13	1	36	12
901NP9CSC	18	9	9	36	12
905DGCC1x	18	4	9	31	10
905DGCC2x	20	4	6	30	10
901SJMS1x	18	5	6	29	10
911NP9ATC	8	10	8	26	9

Appendix I. Expected taxa at reference sites across assemblages for O/E indices with a Pc >0.5