



RMP
REGIONAL MONITORING
PROGRAM FOR WATER QUALITY
IN SAN FRANCISCO BAY

sfei.org/rmp

Calibration of the Multivariate AZTI Marine Biotic Index (M-AMBI) for Potential Inclusion into California Sediment Quality Objective Assessments in San Francisco Bay

Prepared by:

David J. Gillett, Ashley N. Parks, and Steven M. Bay
Southern California Coastal Water Research Project

Calibration of the Multivariate AZTI Marine Biotic Index (M-AMBI) for Potential Inclusion into California Sediment Quality Objective Assessments in San Francisco Bay

David J. Gillett, Ashley N. Parks, and Steven M. Bay
Southern California Coastal Water Research Project (SCCRWP)

SFEI Contribution #939
SCCWRP Technical Report #1070

Suggested Citation:

Gillet, D. J., A. N. Parks, and S. M. Bay. 2019. Calibration of the Multivariate AZTI Marine Biotic Index (M-AMBI) for Potential Inclusion into California Sediment Quality Objective Assessments in San Francisco Bay. SFEI Contribution #939. San Francisco Estuary Institute, Richmond, CA.

EXECUTIVE SUMMARY

Benthic infauna are the most common faunal assemblage used to assess habitat quality across the globe as they demonstrate relatively predictable community-scale changes in both structure and function when exposed to stress. Due to these characteristics, benthic infauna are used as one of three lines of evidence in California's Sediment Quality Objectives (CASQO) to assess the impact of toxic chemicals to biotic resources in the State's enclosed bays and estuaries. The changes in benthic community composition related to disturbance are quantified using benthic assessment indices.

However, benthic communities also change along natural environmental gradients like salinity and consequently many benthic assessment indices do not work across different habitats and salinity regimes. As a local example, the difficulty in disentangling community changes due to variation along environmental gradients from those due to anthropogenic disturbances contributes to the absence of robustly calibrated and validated benthic indices for the mesohaline, oligohaline, and tidal freshwater parts of San Francisco Bay.

Recently, Pelletier et al. (2018) have validated the performance of a newly re-calibrated benthic index that has been modified for application along the environmental gradients of estuaries in US coastal waters – including San Francisco Bay. This M-AMBI (Multivariate AZTI Marine Biotic Index) index could provide a solution to evaluate the benthic communities in the lower salinity habitats of San Francisco Bay that heretofore could not be assessed.

Our goals for this study were to: 1.) Test the applicability of the M-AMBI for analysis of benthic assemblages in the different habitats of San Francisco Bay; and 2.) Create condition thresholds for the M-AMBI that were compatible with the CASQO assessment framework. The applicability of the M-AMBI was evaluated with sample-specific usage criteria for the M-AMBI (% of abundance in a sample assigned a pollution tolerance value), as well as habitat-specific measures of the index's recognition of taxonomic identity (% of all taxa recorded in habitat) and infaunal abundance (% of all individuals enumerated in a habitat). Condition categories were created by comparing M-AMBI scores to those of the CASQO benthic indices within the Polyhaline Habitat of San Francisco Bay.

The M-AMBI was applicable for use on more than 96% of benthic samples collected across each of the habitats within San Francisco Bay. The index recognized between 82 and 97% of the benthic abundance in each habitat except the Oligohaline Habitat, where only 37% were recognized. The index recognized between 23 and 71% of the taxa observed in the different habitats, with the worst performance in the Tidal Freshwater Habitat.

Different iterations of the thresholds were examined to maximize the categorical agreement between the new M-AMBI categories and the existing CASQO index condition categories. The final M-AMBI thresholds were: Reference – ≥ 0.58 , Low Disturbance – < 0.58 and ≥ 0.48 ,

Moderate Disturbance – < 0.48 and ≥ 0.39 , and High Disturbance – < 0.39 . These thresholds yielded relatively good agreement between the M-MAMBI and CASQO benthic indices.

Taken together, the results of this study indicate that the modified M-AMBI can be used with confidence to evaluate the condition of the benthic infauna of the Polyhaline, Coarse Saline Sediment, Mesohaline, and Tidal Freshwater habitats of the San Francisco Bay in a monitoring and assessment context. The index can similarly be applied to assess the habitat condition of the Oligohaline Habitat of the estuary, though caution should be taken when interpreting the results, as some of the most abundant and dominant taxa there were not recognized by the index.

Within the Polyhaline Habitat of the San Francisco Bay Estuary, the new M-AMBI condition categories were concordant and correlated with the condition categories produced by the existing, fully vetted CASQO benthic indices. These patterns indicate that the two assessment approaches are similar and could potentially be interchangeable in the CASQO context, but only with additional validation studies.

Based upon its applicability and the success of the new thresholds mirroring existing CASQO benthic assessment results, the modified M-AMBI could potentially be used as the benthic index line of evidence for CASQO assessments of the samples collected in the Mesohaline, Oligohaline (with caution), and Tidal Freshwater habitats of San Francisco Bay. However, the incorporation of the M-AMBI would be contingent upon further validation of the index in these habitats and comparisons to the CASQO sediment chemistry and sediment toxicity lines of evidence.

INTRODUCTION

Benthic infauna are the most common faunal assemblage used to assess habitat quality across the globe due to their sessile lifestyle, taxonomic diversity, and that they integrate exposure to stress (e.g., Dauer 1993; Warwick 1988; Gray et al. 2002). These characteristics result in relatively predictable community-scale changes in both structure and function when benthic infauna are exposed to stressors (e.g., Rhoads et al. 1978; Pearson and Rosenberg 1978; Gillett 2010). Within the state of California, macrobenthic fauna are used in a variety of monitoring programs and have been incorporated into different regional and statewide environmental management policies as an indicator of disturbance. A prime example is California's Sediment Quality Objectives, which require macrobenthic community evaluation as part of the assessment framework for implementation, as described in the Inland Surface Waters and Enclosed Bays and Estuaries Water Quality Control Plan (SWRCB 2009).

California's Sediment Quality Objectives (CASQO) are used to assess the impacts of toxic, sediment-bound chemicals to the biological resources and beneficial uses of the State's enclosed bays and estuaries (SWRCB 2009). The CASQO framework uses a multiple lines of evidence approach, comprising measures of sediment chemistry, sediment toxicity, and macrobenthic community composition (Bay et al. 2014). Evaluating impacts on macrobenthic community structure is a complex task because perturbations are manifest as changes in abundance and, eventually, changes in the taxa that comprise the community. For ease of communication and repeatability, those community-scale changes have traditionally been summarized with a benthic assessment index. Assessment indices distill complex ecological information into a single value easily understood by environmental managers, while still grounded in sound ecological theory.

San Francisco Bay is the largest embayment in California and is an ecologically complex system that spans the full range of the estuarine salinity gradient from the tidal freshwater of the Sacramento-San Joaquin Delta to the euhaline waters in the Central Bay. These natural gradients make consistent evaluation of macrobenthic fauna potentially problematic because macrobenthic community composition naturally changes along the estuarine salinity gradient (e.g., Ranasinghe et al. 2012; Thompson et al. 2013). The difficulty in disentangling community changes due to variation along environmental gradients from those due to anthropogenic disturbances contributes to the lack of robustly calibrated and validated benthic indices for the mesohaline, oligohaline, and tidal freshwater parts of San Francisco Bay.

The problems in application and interpretation of benthic indices along an estuarine habitat gradient are not unique to San Francisco Bay. In Europe, the European Water Framework Directive (EU WFD) has adopted a benthic assessment approach using the Multivariate AZTI (Technical Institute of Fisheries and Food) Marine Biotic Index (M-AMBI) (Muxika et al. 2007) – a tool for consistent evaluation of benthic communities as they change across natural gradients.

The M-AMBI is a weighted tolerance index that uses habitat-specific expectations of community metrics (e.g., diversity, richness), along with the relative abundance of pollution tolerant and sensitive taxa in a sample to evaluate the biological condition of the location from where it was collected. Because of the way it was constructed, the M-AMBI does not require a large calibration/validation data set and can therefore be applied in most estuaries with relative ease. With these attributes, the M-AMBI represents an attractive potential option for conducting benthic condition assessments along the whole estuarine salinity gradient of San Francisco Bay.

However, when the M-AMBI has been applied in systems outside the European continent where it was originally developed, its performance has produced mixed results at best (e.g., Borja et al. 2008; Borja and Tunberg 2011; Teixeira et al. 2012). To improve the performance of the index, both the AMBI (a univariate version of the M-AMBI) and the M-AMBI have been recalibrated for application in coastal and estuarine waters of the continental United States using data from the US Environmental Protection Agency (EPA) National Coastal Assessment (NCA) surveys (including data and taxonomic experts from the San Francisco Bay area) (Gillett et al. 2015; Pelletier et al. 2018). This new, revised version of the M-AMBI was designed to serve as a single benthic index with applicability to all coastal and estuarine waters of the United States. It was used in the benthic assessment component of the 2015 US EPA NCA survey and will most likely be used as the standard benthic index for NCA surveys in the future.

Given the aforementioned improvements, the revised US M-AMBI (Pelletier et al. 2018) may be a reasonable solution to our current inability to conduct benthic community-based assessments of many embayments in California, including the mesohaline and lower salinity portions of San Francisco Bay. Using the M-AMBI to assess the condition of benthic community samples collected by the San Francisco Bay Regional Monitoring Program for water quality (RMP) throughout the entire estuary would provide a number of potential benefits: 1.) It would be a single applicable to all parts of the estuary; 2.) It would enable direct, meaningful comparisons of ecosystem health in San Francisco Bay to other regions of the country; and 3.) It would allow for easier integration of local data into national assessment programs.

It should be noted though, that the Pelletier et al. (2018) calibration of the M-AMBI was conducted on a broad, regional scale with goals of its use as an assessment tool that could be applied across biogeographical boundaries or in individual locales without a pre-existing benthic index. If one was interested in using the M-AMBI (or any other index) in a local application, it is always prudent to evaluate its performance and adjust the index to best suit local needs (Muxika et al. 2007; Gillett et al. 2015). As an example, if the M-AMBI were to be used as a benthic index within the CASQO assessment framework (e.g., in those areas without a benthic index), the M-AMBI condition thresholds (five categories) would need to be re-calibrated to match those of the existing CASQO benthic indices mathematically (they use only four categories) and in interpretation of any potential disturbance (Ranasinghe et al. 2009).

For this study, the applicability and appropriateness of the M-AMBI for analysis of benthic assemblages in different habitats of San Francisco Bay were evaluated. We then used the index to assess the condition of macrobenthic samples in the polyhaline regions of San Francisco Bay (i.e., Habitat D in Ranasinghe et al. 2012) and recalibrated the thresholds of the index to more closely match the benthic line of evidence tools used in the CASQO (Ranasinghe et al. 2009; Bay et al. 2014). Lastly, we have provided an R-based function to apply the M-AMBI as a tool for assessment of benthic data collected in future surveys.

Methods

M-AMBI – The M-AMBI is a macrofauna-based condition index that assesses the health and biological integrity for a given location using a combination of pollution sensitive and pollution tolerant taxa, species richness, species diversity, and oligochaete abundance (Pelletier et al. 2018; Muxika et al. 2007). Over 6,200 estuarine and marine taxa from around the world have been assigned one of five pollution tolerance values, ranging from very sensitive to very tolerant, which are then used to calculate the relative abundance of sensitive and tolerant taxa (Gillett et al. 2015). Species richness, species diversity, oligochaete abundance, and relative abundance of tolerant/sensitive taxa values are integrated together via a factor analysis and compared to modelled expectations of reference and highly degraded measures based upon the location (U.S. West Coast vs. Gulf and East coasts) and bottom water salinity observed at the time of sample collection. M-AMBI scores can range from 0 (highly degraded) to 1 (not degraded) and are divided into five categories: Good Condition, High Condition, Moderate Condition, Poor Condition, and Bad Condition. Samples with extremely high diversity can sometimes exceed a score of 1 and are included in the Good Condition category.

CASQO Tool – The CASQO Benthic Line of Evidence (BLOE) tool is a macrofauna-based assessment tool that combines four separate benthic indices calibrated for use either in the saline embayments of Southern California or the Polyhaline Habitat of San Francisco Bay (Bay et al. 2014; Ranasinghe et al. 2009). The four benthic indices consist of the Benthic Response Index (BRI, a pollution tolerance index), the Relative Benthic Index (RBI, a multi-metric index), the Index of Biotic Integrity (IBI, a multimetric index), and the RIVPACS index (an observed to expected assemblage composition index). Each of these indices scores a sample independently between 1 (Reference) and 4 (Highly Disturbed). A final CASQO BLOE category is determined by calculating the median value of the scores from the four component indices and rounding up, where needed, to place a sample into one of four categories: Reference (1), Low Disturbance (2), Moderate Disturbance (3), and High Disturbance (4).

Data – The data for this San Francisco Bay-centric analysis was obtained from the CASQO development dataset (US EPA 2004; Hunt et al. 2001; Bay Area Dischargers Association 1994; Thompson et al. 1999) and San Francisco Bay RMP samples collected between 2008 and 2012 (Bay et al. 2013, Willis-Norton et al. 2013). In total there were 6,857 samples of macrobenthic community composition (taxa names and abundances) and station information (salinity, sediment

composition, depth, latitude, and longitude) from 486 stations across all habitats of the estuary. Samples were collected during all seasons. Taxonomy was standardized and updated where possible to reflect taxonomic standards of SCAMIT edition 5 (SCAMIT, 2008), which were used in the development the CASQO benthic indices (Ranasinghe et al. 2009).

Approach – The applicability of the M-AMBI was evaluated with three measures:

- 1.) M-AMBI usage criteria (following Borja and Muxika 2005) – if less than 50% of the total abundance in a sample has an M-AMBI tolerance value assigned to it, the index should not be used for that sample. If between 50 and 80% of the abundance in a sample has a tolerance value, caution should be used in interpreting results. If more than 80% of the abundance has a tolerance value, results can be used without reservation.
- 2.) Taxonomic coverage – a measurement of the relative amount of taxa observed in each of the estuarine habitats that are used by the M-AMBI. A good benthic index should use a large proportion of the taxa observed in a system to yield a holistic assessment of ecosystem condition.
- 3.) Abundance coverage – a measurement of how many organisms observed in each habitat are used by the M-AMBI. A good benthic index should be applicable to the dominant taxa in a system to yield a wholistic assessment of ecosystem condition.

To provide context for the evaluation of the taxonomic and abundance coverage of the M-AMBI, values were compared to the coverage of the CASQO Benthic Response Index (BRI [Ranasinghe et al. 2009]). The BRI is an abundance weighted tolerance index similar to the M-AMBI but calibrated and validated specifically for application in the Polyhaline Habitat of the San Francisco Bay Estuary.

Evaluation and adjustment of the condition thresholds for the M-AMBI (Pelletier et al. 2018) into four CASQO-like categories were conducted using a subset of the data from the Polyhaline Habitat of San Francisco Bay (Habitat D of Ranasinghe et al. 2012), the only habitat within the estuary with the four validated CASQO benthic indices (CASQO BLOE). There were 162 samples from the Polyhaline Habitat that had the required data characteristics for calculation of the indices (latitude, longitude, water depth, overlying salinity at time of sampling, and $\geq 50\%$ of their abundance with M-AMBI tolerance value). Eleven samples with anomalous CASQO benthic index scores and low salinity values (<18 PSU) for this habitat were removed as outliers, yielding a final polyhaline data set of 158 samples (noted in Appendix B). M-AMBI scores were calculated for each sample using the MAMBI.DJG function (Appendix A) in R, which was adapted from Pelletier et al. (2018). CASQO BLOE index scores were calculated following the guidelines of Bay et al. (2014) and calculators available at <http://www.sccwrp.org/about/research-areas/sediment-quality/sediment-quality-assessment-tools/>.

The correlation and concordance of the five M-AMBI condition categories and the four categories used with the CASQO BLOE were evaluated using a Pearson's Chi Squared test of independence ($\alpha = 0.1$). The relationship between M-AMBI scores and the CASQO condition categories was evaluated with an Analysis of Variance and subsequent post-hoc comparisons of mean M-AMBI scores within each category using a Tukey Honest Significant Difference test ($\alpha = 0.1$).

Population estimates (median, 25th, and 75th percentiles) of continuous M-AMBI scores of benthic samples within the defined CASQO BLOE condition categories were used as guidelines for creating a new four condition category-framework for M-AMBI. Using the percentiles as a starting point, categorical thresholds were adjusted to maximize categorical agreement between M-AMBI and the CASQO benthic indices. All calculations were conducted in R v3.5.1.

RESULTS

Based upon the usage criteria recommended for the AMBI and M-AMBI (Borja and Muxika 2005), the index was applicable for use on more than 96% of benthic samples collected from each of the habitats within the San Francisco Bay Estuary (Table 1). Following the usage criteria based upon percent of total abundance in a sample with a tolerance score assigned by the index, the M-AMBI was applicable to 74% of samples from the Polyhaline Habitat with no reservations and it could be applied to 22% of samples with some caution.

Table 1 *The percent of samples from each habitat within San Francisco Bay to which M-AMBI could be applied without reservation, cautiously, or not at all. Applicability was determined by the relative abundance (%) within a sample that could be assigned a tolerance value.*

Habitat Class	Is M-AMBI Applicable?		
	Yes	Cautiously	No
Polyhaline	74.0	22.3	3.7
Coarse Saline Sediments	70.4	25.9	3.7
Mesohaline	84.9	11.9	3.2
Tidal Freshwater	97.5	1.8	0.7
Oligohaline	69.8	27.0	3.1

The M-AMBI recognized between 23 and 71% of the taxa observed in the different habitats of the estuary (Table 2). Within the Polyhaline Habitat, 64% of the taxa in the dataset were

Table 2 *The percent of taxa from each habitat within San Francisco Bay that were recognized by the M-AMBI tolerance value taxa list. For context, the percent of taxa recognized by the BRI is also included for the Polyhaline Habitat.*

recognized by the M-AMBI and had an assigned tolerance value. In contrast, the BRI only recognized 16.1% of the taxa in the polyhaline portion of the dataset.

Habitat Class	Index	% of Taxa	
		Recognized	Not Recognized
Polyhaline	BRI	16.1	83.9
	M-AMBI	64.0	36.0
Coarse Saline Sediment	M-AMBI	71.4	28.6
Mesohaline	M-AMBI	57.1	42.9
Tidal Freshwater	M-AMBI	22.9	77.1
Oligohaline	M-AMBI	52.4	47.6

The M-AMBI recognized between 37 and 97% of the benthic abundance observed in the different habitats of the estuary (Table 3), with performance in all of the habitats except the oligohaline being greater than 82% recognition. Within the Polyhaline Habitat, 97% of the benthic abundance observed across all of the samples was recognized by the M-AMBI. Similarly, the BRI recognized 87% of the abundance.

Table 3 The percent of macrofaunal abundance from each habitat within San Francisco Bay that were recognized by the M-AMBI tolerance value taxa list. For context, the percent of abundance recognized by the BRI is also included for the Polyhaline Habitat.

Habitat Class	Index	% of Total Abundance	
		Recognized	Not Recognized
Polyhaline	BRI	86.6	13.4
	M-AMBI	96.8	3.2
Coarse Saline Sediment	M-AMBI	92.1	7.9
Mesohaline	M-AMBI	82.4	17.6
Tidal			
Freshwater	M-AMBI	93.7	6.3
Oligohaline	M-AMBI	37.1	62.9

There was a significant relationship ($p < 0.001$) among the condition categories assigned by the M-AMBI using the original five categories of Pelletier et al. (2018) and those assigned by the CASQO BLOE indices ($X = 108.89$, $d.f. = 12$). Both indices evaluated samples with a good degree of agreement with only four instances of highly discordant classification (i.e., M-AMBI = poor condition (category 4) and CASQO = Low Disturbance (category 2) or M-AMBI = Good

Table 4 Two-way contingency table comparing the condition category classification of benthic samples from the Polyhaline Habitat by the CASQO BLOE tool and the M-AMBI using condition thresholds of Pelletier et al. (2018).

Condition and CASQO= High Disturbance) (Table 4).

		CASQO Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance
Original M-AMBI Categories	High Condition	2	2	0	0
	Good Condition	25	37	6	1
	Moderate Condition	4	28	17	4
	Poor Condition	0	3	4	13

Bad Condition	0	0	0	8
------------------	---	---	---	---

Analysis of variance (ANOVA) was used to determine if samples grouped by their CASQO categories had different ($\alpha = 0.1$) M-AMBI scores. Overall, there were significant differences in M-AMBI scores with respect to the CASQO categorization of the samples ($p < 0.0001$, $F = 51.73$, d.f. = 3/150). Post-hoc comparison of the categories using Tukey Honest Significant Differences adjusted probabilities indicated there were significant differences between CASQO categories and M-AMBI scores for all categories (Figure 1).

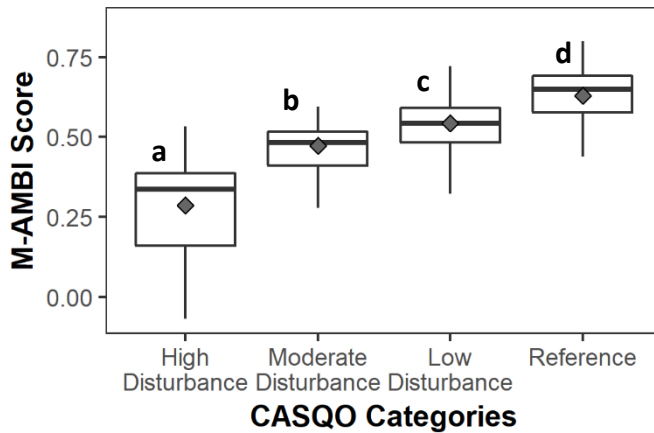


Figure 1 Schematic boxplot illustrating the distribution of M-AMBI scores within each of the four CASQO BLOE condition categories. The grey diamonds indicate the mean value. Different letters indicate statistically significant ($\alpha = 0.1$) differences in means between categories.

Given the significant differences in mean M-AMBI scores between CASQO categories noted above, initial thresholds to create new “CASQO compatible” thresholds for the M-AMBI were based upon 75th and 25th percentile M-AMBI scores within each CASQO category illustrated in Figure 1 (Table 5). Different iterations of the thresholds were examined to maximize the categorical agreement between the new M-AMBI categories and the existing CASQO BLOE condition categories. The final M-AMBI thresholds were: Reference – ≥ 0.58 , Low Disturbance – < 0.58 and ≥ 0.48 , Moderate Disturbance – < 0.48 and ≥ 0.39 , and High Disturbance – < 0.39 . These thresholds yielded relatively good agreement (Weighted Kappa = 0.4) between the M-AMBI and CASQO benthic indices (Figure 2, Table 6). The Reference and High Disturbance categories had the best agreement to the CASQO BLOE indices ($> 73\%$), while the Moderate Disturbance category was the weakest (30%).

Table 5 Population estimates of M-AMBI scores within each CASQO BLOE category. Values correspond to Figure 1.

CASQO Category	M-AMBI Score Estimates			
	25 th Percentile	Median	75 th Percentile	Mean
Reference	0.578	0.649	0.692	0.630
Low Disturbance	0.482	0.542	0.590	0.544
Moderate Disturbance	0.411	0.483	0.517	0.472
High Disturbance	0.160	0.337	0.387	0.287

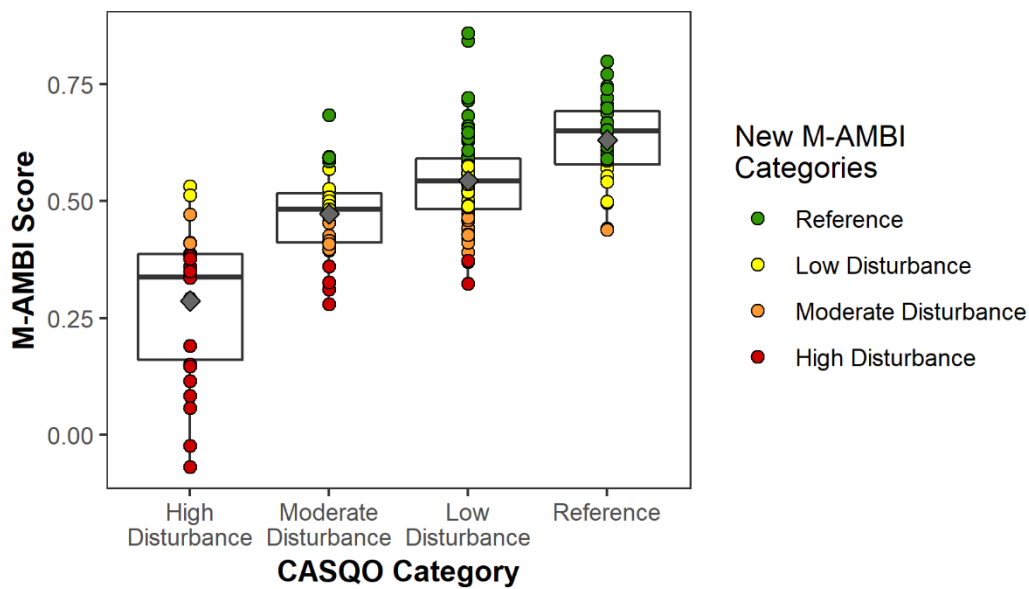


Figure 2 Schematic boxplot illustrating the distribution of M-AMBI scores within each of the four CASQO BLOE condition categories. The grey diamonds indicate the mean value. The circles represent M-AMBI scores of individual samples and the color indicates their condition using the new M-AMBI thresholds.

Table 6 Two-way contingency table comparing the condition category classification of benthic samples from the Polyhaline Habitat by the CASQO BLOE tool and the M-AMBI using the new condition thresholds.

		CASQO Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance
New M-AMBI Categories	Reference	23	21	5	0
	Low Disturbance	6	31	10	2
	Moderate Disturbance	2	15	8	5
	High Disturbance	0	3	4	19

As a point of comparison, M-AMBI scores and new condition categories were compared to those of the four individual indices that make up the CASQO BLOE: BRI, RBI, IBI, and RIVPACS. There was relatively good agreement of M-AMBI scores with the BRI and IBI scores, following a relatively linear pattern (Figure 3). Conversely, there was less agreement between M-AMBI scores and RBI or RIVPACS scores, following more of horizontal banding pattern. The

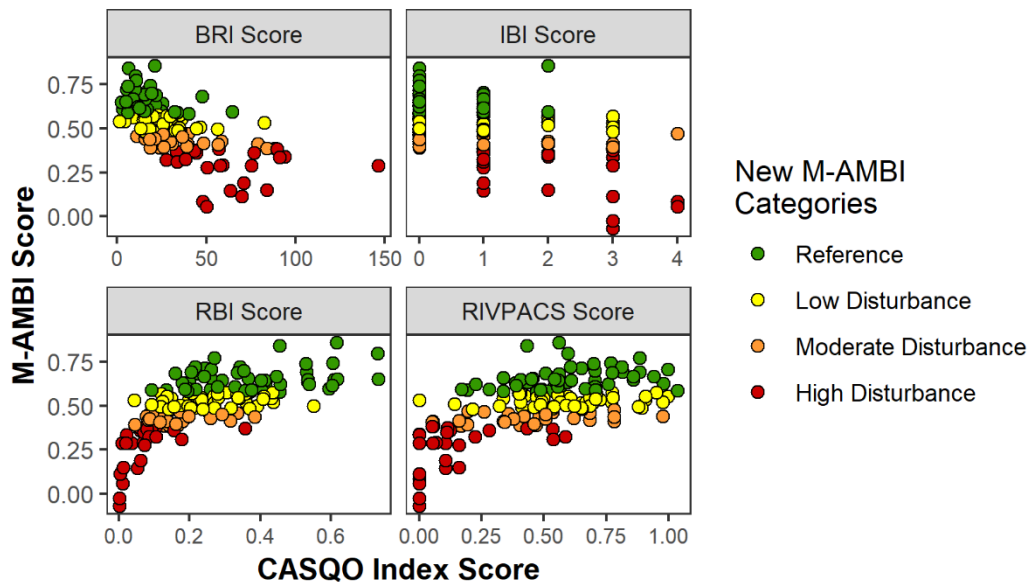


Figure 3 Scatter plots of M-AMBI scores with the scores of the four component benthic indices of the CASQO BLOE applied to samples from the Polyhaline habitat. Colors indicate the condition category based upon the new M-AMBI thresholds.

categorical agreement between the M-AMBI condition categories and those of the four component CASQO indices was weaker (Weighted Kappa between 0.1 and 0.3) than with the integrated CASQO BLOE (Table 7). Interestingly, where the M-MBI had the most disagreement, was with samples in the Moderate Disturbance category, which highlights a pattern common to biologically-based assessments, irrespective of the use of indices or local expert opinion. Most assessment frameworks are good at identifying clearly impacted or clearly unimpacted sites. Where they struggle, is classifying sites with intermediate levels of disturbance (Diaz et al. 2004; Pinto et al. 2008; Teixeira et al. 2010)

This pattern of reasonable agreement between the M-AMBI and the integrated CASQO BLOE, compared to the poor agreement with the component benthic indices, is consistent with the CASQO BLOE development process. As detailed in Ranasinghe et al. (2009), none of four benthic indices (BRI, RBI, IBI, or RIVPACS) were as good at correctly classifying validation sites of known condition as they were when combined together (i.e., the BLOE categorization). Each of the BLOE component indices have varying degrees of bias within them related to their conceptual approach. These mixed biases in turn contribute to the poor index-index agreement among the CASQO BLOE component pieces. As illustrated in Appendix D, the agreement of categorical classifications of the SQO indices with each other was often worse than with the M-AMBI.

Table 7 Two-way contingency tables comparing the condition category classification of benthic samples from the Polyhaline Habitat by the four component indices of the CASQO BLOE tool and the M-AMBI using the new condition thresholds.

		BRI Categories				IBI Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance	Reference	Low Disturbance	Moderate Disturbance	High Disturbance
New M-AMBI Categories	Reference	38	6	5	0	44	4	1	0
	Low Disturbance	20	18	9	1	39	4	6	0
	Moderate Disturbance	7	12	10	2	23	2	4	1
	High Disturbance	0	2	17	7	10	4	10	2
		RBI Categories				RIVPACS Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance	Reference	Low Disturbance	Moderate Disturbance	High Disturbance
New M-AMBI Categories	Reference	12	17	7	13	18	27	4	0
	Low Disturbance	1	16	10	23	14	31	2	2
	Moderate Disturbance	0	2	5	23	4	18	5	3
	High Disturbance	0	1	0	25	0	4	4	18

CONCLUSIONS

- The US M-AMBI modified by Pelletier et al (2018) can be used with confidence to evaluate the condition of the benthic infauna of the Polyhaline, Coarse Saline Sediment, Mesohaline, and Tidal Freshwater habitats of the San Francisco Bay in a monitoring and assessment context. Our tests of index applicability illustrate that historical samples from those habitats meet the index's internal standards of applicability and that a large majority of the taxa that live in these habitats were taken into consideration by the index when evaluating habitat condition.
- The M-AMBI, as presently constructed, could tentatively be applied to assess the habitat condition of the Oligohaline habitat of the estuary in a monitoring and assessment context. A majority of those historical samples met the M-AMBI's acceptability criteria and the index is capable of evaluating most of the taxa found in this habitat. However, caution should be taken when interpreting the results, as some of the most abundant and dominant taxa of the Oligohaline Habitat were not assigned tolerance values by the M-AMBI.
- Within the Polyhaline Habitat of the San Francisco Bay Estuary, the M-AMBI categorical assessment of the condition of benthic samples was concordant and correlated with the condition categories produced by the CASQO Benthic Line of Evidence assessment tool. These patterns indicate that the two assessment approaches are similar and could potentially be interchangeable in the CASQO context, but only with additional validation studies.
- The new thresholds created for the M-AMBI to facilitate potential integration into the CASQO assessment framework were successful in attaining a good level of agreement with the patterns of the previously validated CASQO BLOE tool.
- Based upon its applicability and the success of the new thresholds mirroring existing CASQO BLOE results in the Polyhaline Habitat, the US M-AMBI could potentially be used as the benthic index/Benthic Line of Evidence for CASQO assessments of the samples collected in the Mesohaline, Oligohaline (with caution), and Tidal Freshwater habitats of San Francisco Bay. However, the incorporation of the M-AMBI would be contingent upon further validation of the index in these habitats and comparisons to the CASQO sediment chemistry and sediment toxicity lines of evidence.

RECOMMENDATIONS

- M-AMBI performance should be further validated in the Mesohaline, Oligohaline, and Tidal Freshwater habitats of the San Francisco Bay Estuary. Validation of index performance should be done in conjunction with known reference and degraded areas within each habitat.

- M-AMBI assessment results should be integrated with measures of sediment chemistry and sediment toxicity within the Mesohaline, Oligohaline, and Tidal Freshwater habitats to characterize patterns in overall CASQO assessments of these parts of the estuary.
- Additional macrobenthic taxa—especially the numerically dominant taxa—from the Oligohaline Habitat should be assigned M-AMBI tolerance values to improve the index’s applicability in that habitat.

AKNOWLEDGEMENTS

The authors thank Phil Trowbridge and the San Francisco Estuary Institute for guidance and funding of this work. The authors thank Melissa Foley, Jay Davis, Paul Work, and the San Francisco Public Utilities District for their comments on earlier versions of this report and their suggestions for improving the document. This report was funded by the Regional Monitoring Program under contract #1341 between SFEI and SCCWRP.

REFERENCES

- Alden, R.W., Dauer, D.M., Ranasinghe, J.A., Scott, L.C., Llansó, R.J., 2002. Statistical verification of the Chesapeake Bay benthic index of biotic integrity. *Environmetrics* 13, 473–498. doi:10.1002/env.548
- Bay, S.M., D.J. Greenstein, S.L. Moore, K.J. Ritter, Ranasinghe, J.A., 2013. Evaluation of sediment condition using California's sediment quality objectives assessment framework. Technical Report 764. Southern California Coastal Water Research Project. Costa Mesa, CA.
- Bay, S.M., Greenstein, D.J., Ranasinghe, J.A., Diehl, D.W., Fetscher, A.E., 2014. Sediment Quality Assessment Technical Support Manual. Southern California Coastal Water Research Project. Technical Report 777. Southern California Coastal Water Research Project. Costa Mesa, CA.
- Bay Area Dischargers Association, 1994. Local effects monitoring program, quality assurance project plan. Bay Area Clean Water Association, Oakland, CA.
- Borja, A. and Muxika, I., 2005. Guidelines for the use of AMBI (AZTI's Marine Biotic Index) in the assessment of the benthic ecological quality. *Mar. Pollut. Bull.* 50, 787-789.
- Borja, A., Dauer, D.M., Díaz, R., Llansó, R.J., Muxika, I., Rodríguez, J.G., Schaffner, L., 2008. Assessing estuarine benthic quality conditions in Chesapeake Bay: A comparison of three indices. *Ecol. Indic.* 8, 395–403. doi:10.1016/j.ecolind.2007.05.003
- Borja, A., Franco, J., Pérez, V., 2000. A marine Biotic Index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. *Mar. Pollut. Bull.* 40, 1100–1114. doi:10.1016/S0025-326X(00)00061-8
- Borja, A., Tunberg, B.G., 2011. Assessing benthic health in stressed subtropical estuaries, eastern Florida, USA using AMBI and M-AMBI. *Ecol. Indic.* 11, 295–303. doi:10.1016/j.ecolind.2010.05.007
- Diaz, R.J., Solan, M, and Valente, R.M., 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. *J Envi. Mang.* 73:165-181.
- Dauer, D.M., 1993. Biological criteria, environmental health, and estuarine macrobenthic community structure. *Mar. Pollut. Bull.*, 26:249-257.
- Gillett, D.J., 2010. Effects of Habitat Quality on Secondary Production in Shallow Estuarine Waters and the Consequences for the Benthic-Pelagic Food Web. Ph.D. Dissertation, The College of William and Mary.
- Gillett, D.J., Weisberg, S.B., Grayson, T., Hamilton, A., Hansen, V., Leppo, E.W., Pelletier, M.C., Borja, A., Cadien, D., Dauer, D., Diaz, R., Dutch, M., Hyland, J.L., Kellogg, M., Larsen, P.F., Levinton, J.S., Llansó, R., Lovell, L.L., Montagna, P.A., Pasko, D., Phillips,

- C.A., Rakocinski, C., Ranasinghe, J.A., Sanger, D.M., Teixeira, H., Van Dolah, R.F., Velarde, R.G., Welch, K.I., 2015. Effect of ecological group classification schemes on performance of the AMBI benthic index in US coastal waters. *Ecol. Indic.* 50, 99–107. doi:10.1016/j.ecolind.2014.11.005
- Gray, J.S., Wu, R.S., Or, Y.Y., 2002. Effects of hypoxia and organic enrichment on the coastal marine environment. *Mar. Ecol. Prog. Ser.*, 238:249-279.
- Hunt, J.W., Anderson, B.S., Phillips, B.M., Tjeerdema, R.S., Taberski, K.M., Wilson, C.J., Puckett, H.M., Stephenson, M., Fairey, R., Oakden, J.M., 2001. A large-scale categorization of sites in San Francisco Bay, USA, based on the sediment quality triad, toxicity identification evaluations, and gradient studies. *Environ. Tox. and Chem.* 20, 1252–1265.
- Muxika, I., Borja, Á., Bald, J., 2007. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. *Mar. Pollut. Bull.* 55, 16–29. doi:10.1016/j.marpolbul.2006.05.025
- Pelletier, M.C., Gillett, D.J., Hamilton, A., Grayson, T., Hansen, V., Leppo, E.W., Wesiberg, S.B., Borja, A., 2018. Adaptation and application of multivariate AMBI (M-AMBI) in US coastal waters. *Ecol. Indic.* 89, 818-827.
- Pearson, T.H. Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment and pollution of the environment. *Oceanogr. Mar. Biol.*, 16:229-311.
- Pinto, R., Patrício, J., Baeta, A., Fath, B.D., Neto, J.M., Marques, J.C., 2009. Review and evaluation of estuarine biotic indices to assess benthic condition. *Ecol. Indic.* 9, 1–25. doi:10.1016/j.ecolind.2008.01.005
- Ranasinghe, J.A., Weisberg, S.B., Smith, R.W., Montagne, D.E., Thompson, B., Oakden, J.M., Huff, D.D., Cadien, D.B., Velarde, R.G., Ritter, K.J., 2009. Calibration and evaluation of five indicators of benthic community condition in two California bay and estuary habitats. *Mar. Pollut. Bull.* 59, 5–13. doi:10.1016/j.marpolbul.2008.11.007
- Ranasinghe, J.A., Welch, K.I., Slattery, P.N., Montagne, D.E., Huff, D.D., Lee, H., Hyland, J.L., Thompson, B., Weisberg, S.B., Oakden, J.M., Cadien, D.B., Velarde, R.G., 2012. Habitat-related benthic macrofaunal assemblages of bays and estuaries of the western united states. *Integr. Environ. Assess. Manag.* 8, 638–648. doi:10.1002/ieam.62
- Rhoads, D.C., P.L. McCall, Yingst, J.Y., 1978. Disturbance and production on the estuarine seafloor. *Amer. Sci.*, 66:577-586.
- State Water Resources Control Board (SWRCB), 2009. Water Quality Control Plan for Enclosed Bays and Estuaries - Part 1: Sediment Quality. SWRCB. Sacramento, CA.

- Teixeira, H., Borja, A., Weisberg, S.B., Ranasinghe, J.A., Cadien, D.B., Dauer, D.M., Dauvin, J., Degraer, S., Diaz, R.J., Grémare, A., Karakassis, I., Llansó, R.J., Lovell, L.L., Marques, J.C., Montagne, D.E., Occhipinti-Ambrogi, A., Rosenberg, R., Sardá, Rafael, Schaffner, L.C., and Velarde, R.G., 2010. Assessing coastal benthic macrofauna community condition using best professional judgement – developing consensus across North America and Europe. *Mar. Poll. Bull.*, 60:589-600.
- Teixeira, H., Weisberg, S.B., Borja, A., Ranasinghe, J.A., Cadien, D.B., Velarde, R.G., Lovell, L.L., Pasko, D., Phillips, C.A., Montagne, D.E., Ritter, K.J., Salas, F., Marques, J.C., 2012. Calibration and validation of the AZTI's Marine Biotic Index (AMBI) for Southern California marine bays. *Ecol. Indic.* 12, 84–95. doi:10.1016/j.ecolind.2011.05.025
- Thompson, B., Anderson, B.S., Hunt, J.W., Taberski, K.M., Phillips, B.M., 1999. Relationships between sediment contamination and toxicity in San Francisco Bay. *Mar. Environ. Res.* 48, 285–395.
- U.S. Environmental Protection Agency, 2004. National Coastal Condition Report II. U.S. Environmental Protection Agency, Office of Research and Development, EPA-620/R-03/002, Washington, DC.
- Warwick, R.M., 1988. Effects on community structure of a pollutant gradient – summary. *Mar. Ecol. Prog. Ser.*, 46:207-211.
- Willis-Norton, E., Ranasinghe, J.A., Greenstein, D., Bay, S. 2013. Applying sediment quality objective assessments to San Francisco Bay samples from 2008-2012. Contribution No. 702. San Francisco Estuary Institute. Richmond, CA

APPENDIX A – M-AMBI R CODE

We have created a custom R function to calculate M-AMBI score in the MAMBI calculator-djg.R script. To run the M-AMBI R script, one will need to have base R – the analytical underpinnings – and R Studio – a good program to interface with R – installed on your machine. All of this work has been done on a PC running Windows.

R can be downloaded from: <https://cran.r-project.org/>

R Studio can be downloaded from: <https://www.rstudio.com/>

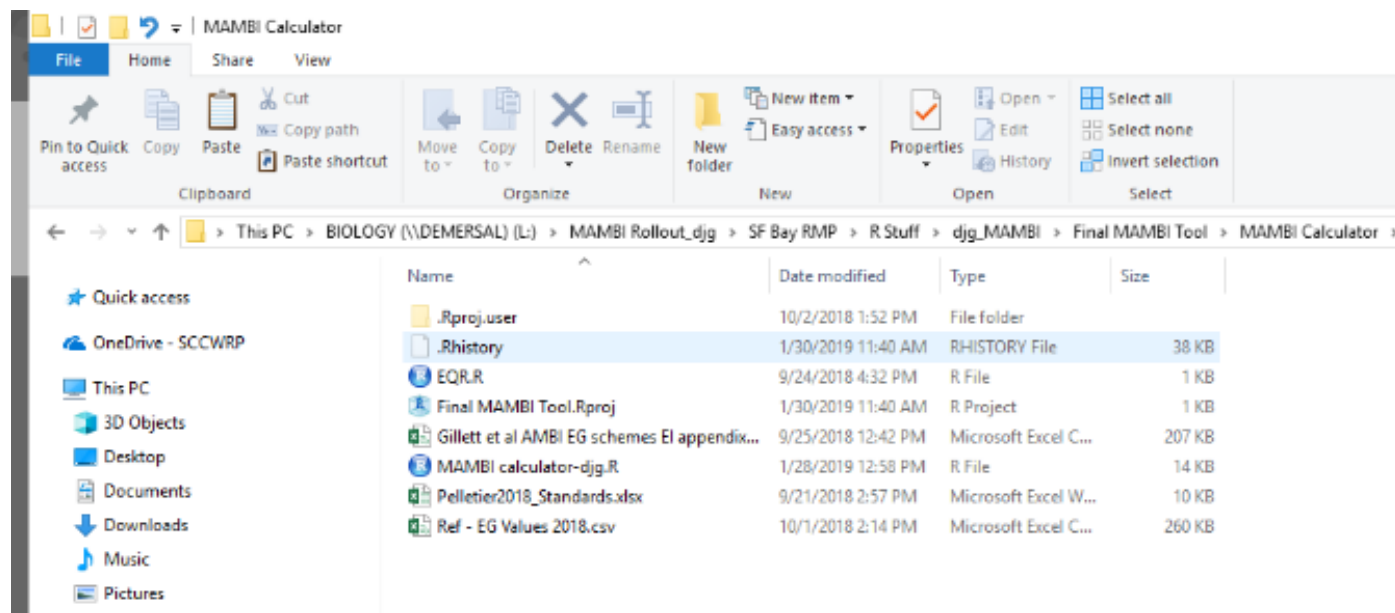
Once you have programs installed, you can start unpacking and working with the M-AMBI script. First unzip the file accompanying this report (or double click on the folder icon) onto your



local machine. MAMBI Calculator 4-9-19.zip

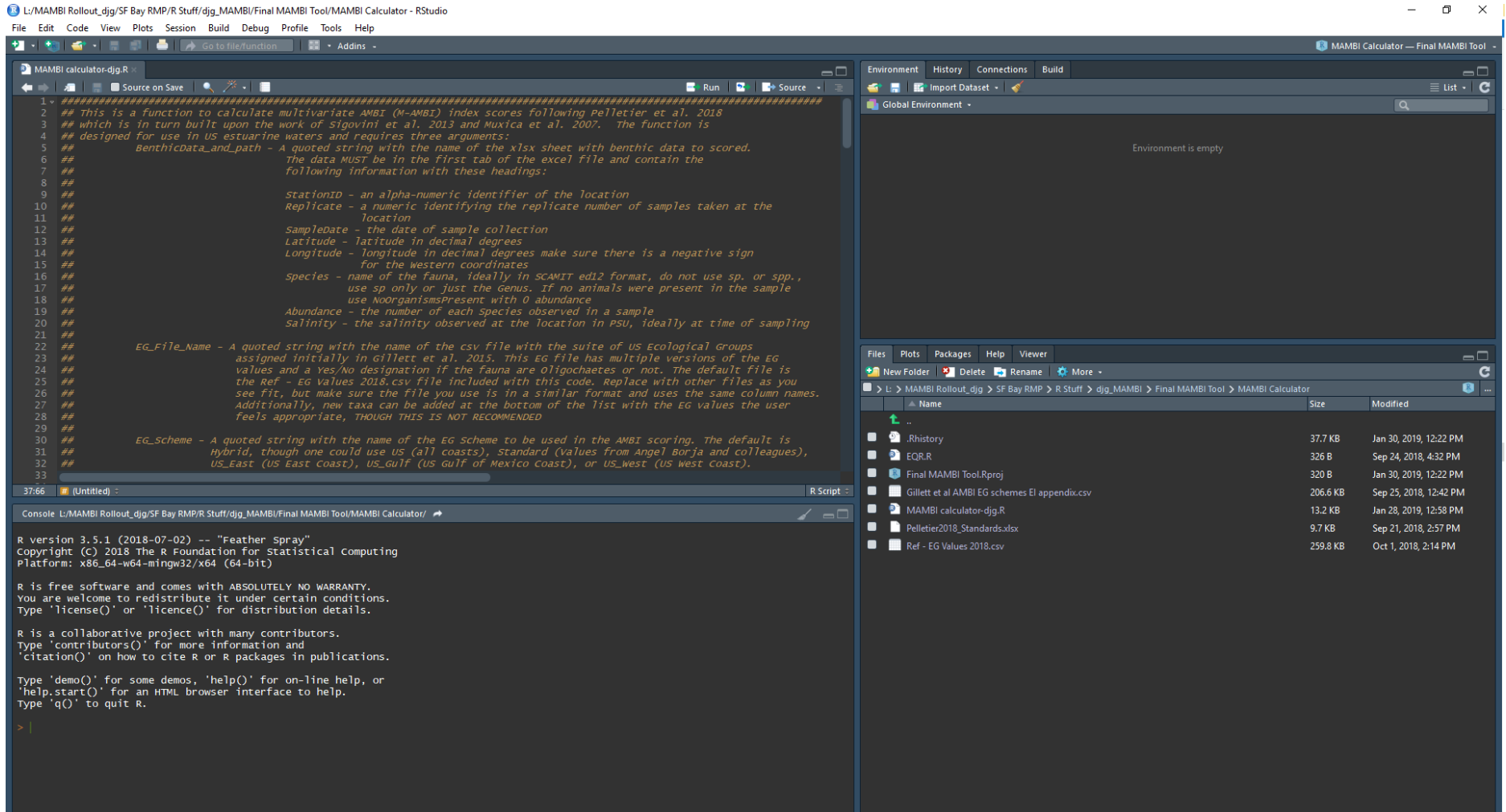
When unzipped, the contents should look like Figure A1.

Figure A1 Image of unzipped folder containing the R scripts and associated files



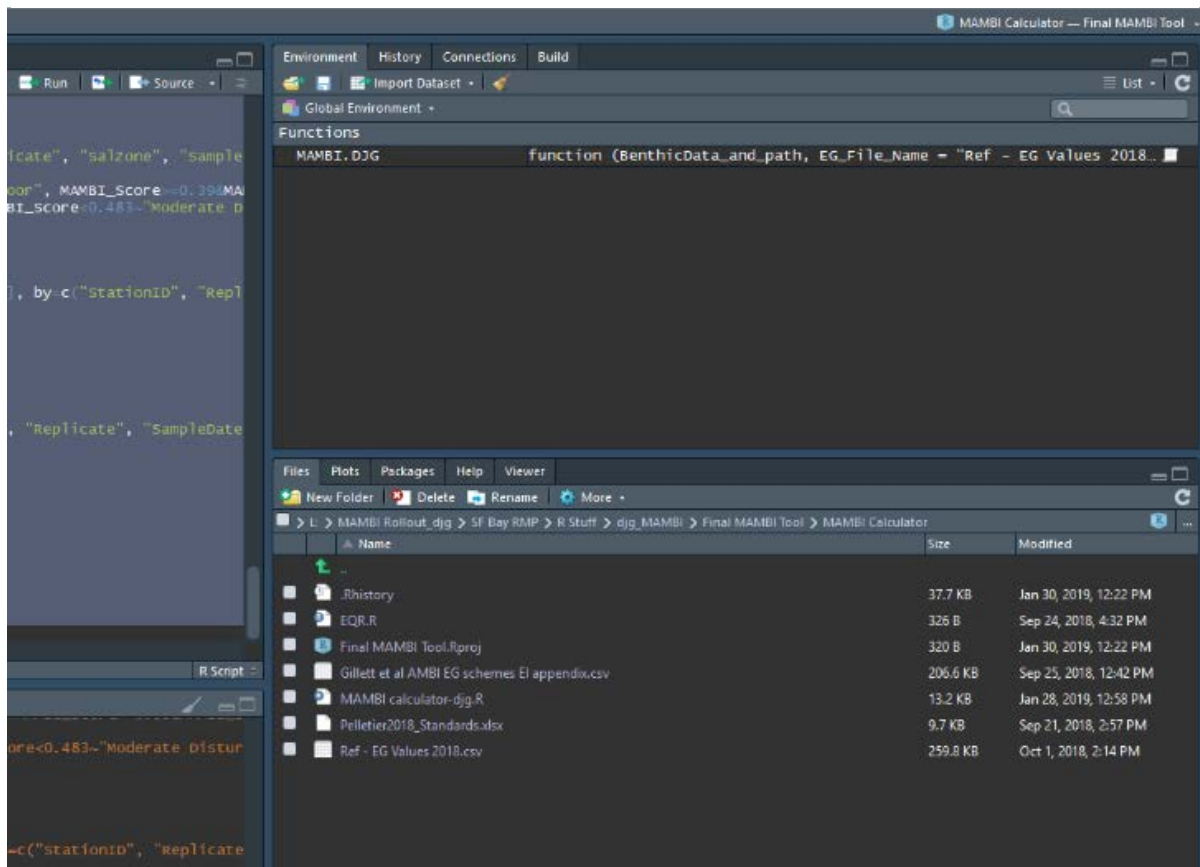
From this file folder, double-click the “Final MAMBI Tool.Rproj” file; this is an R Studio Project that will open R Studio and associate contents of the folder in the R Studio environment. In the lower right pane of R Studio (the file pane) double click on the “MAMBI calculator-djg.R” link; this will open the script into the Source pane (upper left), which should look like Figure A2. Instructions for application of the M-AMBI function, the input file format, and the data needed to calculate M-AMBI are at the top of the script (orange text in figure A2). The calculator will not work if the format is not followed exactly.

Figure A2 An image of R Studio with MAMBI Calculator loaded into Source pane (upper left)



As noted in the instructions, the following R packages must be installed on your machine for the MAMBI function to work: tidyverse, reshape2, vegan, and readxl. Packages can be installed via the tools tab at the top or directly using the function “install.packages()”. Once installed, highlight all of the text and run it. This will enable the MAMBI.DJG function into your R Studio session. Upon success, it should be listed in the Environment pane of R Studio (upper right)

Figure A3 An image of R Studio Environment pane once the MAMBI.DJG function has been enabled



(Figure A3).

Once the function has been enabled in your R Studio session, open a new script in R Studio. This is where you can start using the function to calculate M-AMBI. As noted in the instructions, the MAMBI.DJG function has three arguments. All three arguments must be present to calculate MAMBI scores.

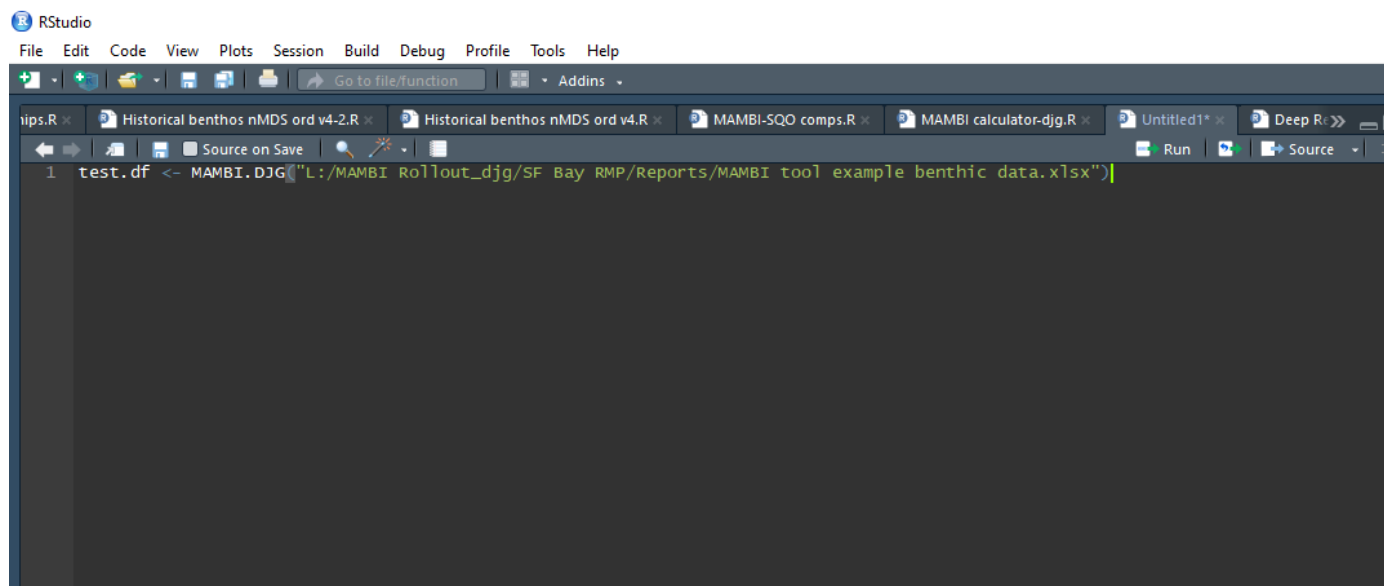
1. `BenthicData_and_path` – This is the path and `xlsx` file name of the benthic data you want to analyze contained in quotation marks. **This must be specified by the user.** Note that you have to change the slashes in a normal windows path from back slash (`\`) to forward

slashes (/) for R. Also note that if the data are not exactly in the format specified in the instructions, the function will not work.

2. EG_File_Name – this is the name of the csv file that contains the tolerance values the M-AMBI uses. **Default value is “Ref – EG Values 2018.csv” and is a file included in the zipped folder.** If the user wants to supply their own tolerance values, we would suggest editing this file or provide a path and file name for a new file of their own (not recommended for novice users).
3. EG_Scheme – this is the column name within the csv file containing the tolerance values. **The default value is “Hybrid”.** The hybrid scheme has the best set of tolerance values identified during the calibration and validation of this M-AMBI tool. If the user wants to use a different EG scheme, they can choose among the options in the csv file or supply their own (not recommended for novice users).

A small example benthic dataset from the San Francisco Bay Estuary has been included in appendix C. The following presents an example of how to calculate M-AMBI scores and condition categories using the MAMBI.DJG function. In short, the user has to name a dataframe into which the function will put its results. Begin by naming an object (test.df, in this illustration) and assigning the output of MAMBI.DJG function to it. In this example, we will choose to use the default settings and therefore only need to specify a quoted string with the name and associated path of the benthic data. The Source pane (upper left) should look like Figure A4,

Figure A4 An image of the R Studio Source pane with the code to calculate M-AMBI scores for the example data set



The image shows the RStudio interface with the Source pane open. The code in the Source pane is as follows:

```
1 test.df <- MAMBI.DJG("L:/MAMBI Rollout_djg/SF Bay RMP/Reports/MAMBI tool example benthic data.xlsx")
```

The Source pane also shows the file path: "L:/MAMBI Rollout_djg/SF Bay RMP/Reports/MAMBI tool example benthic data.xlsx".

albeit with your own pathway to the benthic data.

Highlight this line of code and run it. This will produce a data frame called test.df that contains the results of the M-AMBI scoring function. This test.df can be opened in R or exported for viewing in another program using a variety of different functions (e.g., write.table(), write.csv()). Exporting the data frame is the best way to save the results of the calculation. As noted in the MAMBI.DJG instructions, the output file will contain the following fields:

StationID, Replicate, and SampleDate – which combine to represent a unique sample

Latitude and Longitude – location of the sample

SalZone – the salinity zone the function assigned the sample to based upon its observed salinity and used to set index expectations of reference and highly disturbed (WPH= Western Polyhaline, MH = Mesohaline, OH = Oligohaline, TF = Tidal Freshwater)

AMBI_Score, S, and H – component metrics used to calculate M-AMBI

MAMBI_Score – the index score for that sample

Orig_MAMBI_Condition – condition category the M-AMBI score corresponds to using the scheme of Pelletier et al. (2018).

New_MAMBI_Condition – condition category the M-AMBI score corresponds to using the new scheme presented in this report

Use_MAMBI – Yes/No qualifier indicating if the M-AMBI was appropriate to apply to the sample

Use_AMBI – Yes/No/Cautiously qualifier indicating the confidence one should have in the M-AMBI score. It is based upon the % of the abundance in a sample that was assigned a tolerance value and following recommendations of Borja and Muxika (2005).

YesEG – the percent of the abundance in a sample that had a tolerance value assigned to it.

For inquiries about the performance of this script, please contact davidg@sccwrp.org

APPENDIX B

M-AMBI and CASQO BLOE scores, condition categories, and ancillary data produced from the dataset used in threshold creation within the Polyhaline Habitat of the San Francisco Bay Estuary. Note that the samples removed from the final analysis data set due to anomalous scores are marked with an X in the Data Flagged column.

Please see the attached “MAMBI-CASQO output appendix.xlsx” file.

APPENDIX C

Example input file of benthic data from San Francisco Bay that can be analyzed with the MAMBI.DJG R function.

Please see the attached “MAMBI tool example benthic data.xlsx” file.

APPENDIX D

Two-way contingency tables comparing the condition category classification of benthic samples from the Polyhaline Habitat by the four component indices of the CASQO BLOE tool to each other.

		BRI Categories				IBI Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance	Reference	Low Disturbance	Moderate Disturbance	High Disturbance
RBI Categories	Reference	12	1	0	0	13	0	0	0
	Low Disturbance	28	6	2	0	34	2	0	0
	Moderate Disturbance	11	8	3	0	16	4	2	0
	High Disturbance	14	23	36	10	53	8	19	3
		RIVPACS Category							
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance				
RBI Categories	Reference	4	8	1	0				
	Low Disturbance	14	21	1	0				
	Moderate Disturbance	9	11	2	0				
	High Disturbance	9	40	11	23				

e				
---	--	--	--	--

		RIVPACS Category				IBI Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance	Reference	Low Disturbance	Moderate Disturbance	High Disturbance
BRI Categories	Reference	23	38	4	0	61	3	1	0
	Low Disturbance	10	25	2	1	32	2	4	0
	Moderate Disturbance	3	17	7	14	22	7	9	3
	High Disturbance	0	0	2	8	1	2	7	0

		IBI Categories			
		Reference	Low Disturbance	Moderate Disturbance	High Disturbance
RIVPACS Category	Reference	34	2	0	0
	Low Disturbance	71	5	4	0
	Moderate Disturbance	8	3	3	1
	High Disturbance	3	4	14	2