

M-AMBI revisited: looking inside a widely-used benthic index

Marco Sigovini · Erica Keppel ·
Davide Tagliapietra

Received: 8 February 2013 / Revised: 4 May 2013 / Accepted: 18 May 2013 / Published online: 8 June 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract M-AMBI is a multimetric index for assessing the ecological quality status of marine and transitional waters. It is based on benthic macroinvertebrates and integrates AMBI, a biotic index based on species sensitivity/tolerance, with diversity and richness, making it compliant with the European Water Framework Directive. The success of AMBI paved the way for the introduction of M-AMBI, which was subsequently incorporated into the regulations of several European countries. The M-AMBI algorithm integrates the metrics by means of factor analysis (FA). In this paper, we first reproduced the algorithm using the open source R software. This enabled us to point out that FA is not functional to M-AMBI, and its omission does not appreciably change the results. We then enhanced the applicability of the index, making it independent of the number of samples. In this way, M-AMBI is closely approximated by the simple mean of the normalised metrics with no need for multivariate techniques. Finally, we further simplified the

approach, presenting a bivariate version that is still highly correlated with M-AMBI, in which the constitutive metrics are reduced to a diversity measure and a species sensitivity index. The properties of this bivariate version include simplicity, transparency, robustness, and openness.

Keywords Biotic index · Diversity · Macrozoobenthos · Multimetric index · Open source · Water Framework Directive

Introduction

The European Water Framework Directive 2000/60/EC (hereafter WFD; European Community, 2000) aims to protect and improve the quality of European water bodies. Since its introduction, the WFD has encouraged the development of new tools to evaluate ecological quality status on the basis of given biological quality elements (Hatton-Ellis, 2008). Biological elements to be monitored in coastal and transitional waters include benthic macroinvertebrates in terms of their diversity, abundance and the presence of disturbance-sensitive taxa.

M-AMBI ('Multivariate AMBI', Bald et al., 2005; Muxika et al., 2007) is a multimetric index, based on the macrozoobenthic community, designed to be compliant with WFD requirements. It adopts a multivariate (trivariate) approach, integrating the response

Handling editor: Pierluigi Viaroli

Electronic supplementary material The online version of this article (doi:10.1007/s10750-013-1565-y) contains supplementary material, which is available to authorized users.

M. Sigovini (✉) · E. Keppel · D. Tagliapietra
CNR—National Research Council of Italy, ISMAR—
Marine Sciences Institute, Arsenalate-Tesa 104, Castello
2737/F, 30122 Venice, Italy
e-mail: marco.sigovini@ve.ismar.cnr.it

of three selected metrics, i.e. species richness, the Shannon diversity index (Shannon & Weaver, 1949) and the biotic index AMBI (Borja et al., 2000). AZTI Marine Biotic Index (AMBI) relies on the calculation of the biotic coefficient (BC), which is based in turn on the proportion of disturbance-sensitive taxa and is expressed on a continuous scale ranging from 0 (best status) to 6 (worst status). The AMBI approach follows a model (Glémarec & Hily, 1981; Grall & Glémarec, 1997) which categorises benthic invertebrates into five ecological groups (EGs), depending on their dominance along a gradient of organic enrichment and oxygen depletion. This conceptual model builds on the classical work of Pearson & Rosenberg (1978), which forms the basis of a number of other biotic indices (e.g. Simboura & Zenetos, 2002; Grall & Glémarec, 2003; Rosenberg et al., 2004; Dauvin & Ruellet, 2007). For an overview of the subject, see Tagliapietra et al. (2012).

The success of AMBI paved the way for the introduction of M-AMBI, which has been officially incorporated into the regulations of several European countries in the context of WFD implementation (Bulgaria, France, Germany, Italy, Romania, Slovenia and Spain; Carletti & Heiskanen, 2009). The BAT index ('Benthic Assessment Tool'; Teixeira et al., 2009), which is officially adopted in Portugal, relies on the same approach, with the Margalef index (Margalef, 1958) replacing richness among the constituent metrics. There have also been preliminary attempts to apply M-AMBI to the assessment of benthic community conditions as an indicator of sea-floor integrity in the context of the European Marine Strategy Framework Directive (2008/56/EC, European Community, 2008) (Borja et al., 2011). Recently, M-AMBI has also been applied further afield (Borja et al., 2008a; Bakalem et al., 2009; Costa-Dias et al., 2010; Borja & Tunberg, 2011).

The adoption of the index has also been favoured by a user-friendly software tool developed by AZTI-Tecnalia for calculating AMBI and M-AMBI. The calculation performed by the software is based on a list of macroinvertebrate taxa, mostly at the rank of species, which is revised from time to time and is now composed of about 6,500 entries (including synonyms; updated March 2012). The software is freely available from <http://ambi.azti.es/> (currently version 5.0; Borja et al., 2012b). However, the software code, based on a MATLAB routine, is not open source. We maintain that the user should be able to fully

understand and control the calculation procedures. Therefore, we reproduced and checked each step of the algorithm using the open source R software. In this way, we were able to deal with the method's assumptions and constraints and propose amendments.

Materials and methods

M-AMBI integrates the biotic index AMBI with Shannon diversity and richness. The three metrics were originally selected by factor analysis (FA; see for example Thurstone, 1947; Mulaik, 1972; Harman, 1976; Hair et al., 1998). This multivariate statistical approach is mainly designed to explain the variance in the observed variables, and at the same time to reduce the dimensionality of the dataset using a smaller number of underlying variables ('latent factors'). However, after the selection of the metrics, this technique was retained within the M-AMBI algorithm (Muxika et al., 2007). The following steps were identified on the basis of published details (Muxika et al., 2007; Borja et al., 2012b) and reproduced with open source R software:

- (1) For each sample in the dataset, the species richness (number of species, S), diversity (Shannon index, H') and AMBI-BC are calculated. Diversity is calculated from numerical abundances, with logarithm set to base 2 (Shannon & Weaver, 1949). BC is calculated on the basis of the list of taxa with the assigned EGs supplied by AZTI-Tecnalia (<http://ambi.azti.es/>), with a null weighting given to the species that are not listed.
- (2) Minimum and maximum reference values for the three metrics are added to the dataset as fictitious samples. Typically, minimum reference values ('Bad reference conditions') are represented by the metrics' theoretical minima ($S = 0$, $H' = 0$, $BC = 6$) whereas maximum values ('High reference conditions') correspond to the highest values in the dataset (or the lowest in the case of BC). However, alternative reference values can be set (as occurs in the implementation of the WFD at the national level).
- (3) Each metric is standardised by subtracting its mean and dividing by its standard deviation.
- (4) Factor analysis is performed by means of an eigendecomposition (in practice, a PCA) of the

matrix of standardised metrics. In FA, the loadings are the eigenvectors rescaled by the square root of the eigenvalues (Revelle, 2012). All three factors are retained and a Varimax orthogonal rotation (Kaiser, 1958) is applied to the loadings. Borja et al. (2012b) indicate a generic regression-based method for estimating factor scores. However, we exactly reproduced the M-AMBI results obtained by the AZTI-Tecnalia software simply by multiplying the matrix of standardised metrics by the (rotated) loadings matrix (Grice, 2001).

- (5) The factor scores are orthogonally projected on to the line identified by the reference samples and normalised to the range 0–1 (by dividing by the square root of 2), with 0 corresponding to the fictitious ‘Bad’ sample and 1 corresponding to the ‘High’ one. The Euclidean distance between 0 and the score projection corresponds to the ‘ecological quality ratio’ (EQR) for that sample (European Community, 2000).

In addition to M-AMBI, we introduced two alternative algorithms. The first (hereafter identified as M-AMBI*) excludes point 4, i.e. the FA. The three standardised metrics are therefore projected directly on to the axis identified by the reference values. The second algorithm (M-AMBI*_(n)), in addition to dropping point 4, replaces metric standardisation (point 3) with minimum–maximum normalisation, i.e. the minimum is subtracted from each value and the result is divided by the total range, using ‘High’ and ‘Bad’ reference values as extrema. The latter approach was also applied to just two of the three metrics, combining the ‘sensitivity metric’ BC with either *S* or *H'* as a ‘diversity metric’.

M-AMBI and the alternative algorithms were calculated, and the Pearson correlation analysed, with reference to three datasets. The first is the famous Ekofisk dataset (Gray et al., 1990), which is available as an example in the PRIMER software for community analysis (Clarke & Gorley, 2006). It is based on a survey of the Ekofisk field (Norway) carried out in 1987. Each of 39 samples was sampled by a Day grab (0.1 m², three replicates) and sieved through a 1 mm mesh, yielding 13,883 individuals and 173 taxa (139 of which were identified to the rank of species). The second dataset (hereafter ‘Venice’, Tagliapietra et al., 1998, 2000) is representative of benthic assemblages

from a coastal transitional ecosystem. The samplings were performed in the Lagoon of Venice (Italy) in 1991 as part of a wider framework of studies under the patronage of UNESCO, known as ‘Sistema Lagunare Veneziano’ (Lasserre & Marzollo, 2000). 42 stations were sampled in Palude della Rosa, on the landward side of Venice Lagoon. Samples were collected by means of a Van Veen grab (0.1 m²) and sieved through a 1 mm mesh, yielding 34,732 individuals and 62 taxa (48 of which were identified to the rank of species).

A third dataset was artificially created, characterized by statistically independent metrics. Samples were generated independently of each other, with no gradient taken into account. The dataset comprises 50 samples, as recommended by Borja et al. (2008b). Species abundances follow a Poisson lognormal distribution (Grøtan & Engen, 2008), which occurs when sampling randomly from an assemblage with lognormal distribution (Bulmer, 1974), a widely recognised model in community ecology. The Poisson lognormal distribution is described by the parameters μ and σ^2 , which correspond to the mean and standard deviation (among species) of the log abundances. For each sample, μ and σ^2 were in turn randomly generated from a normal distribution (with parameters estimated on the basis of the two real datasets). The resulting matrix includes 120 ‘fictitious species’. Finally, the EG (1–5) for each of the simulated species was randomly generated to ensure the independence of AMBI.

The analyses were performed using the R software environment for statistical computing, v. 2.15.0 (R Development Core Team, 2012), available as free software under the GNU General Public License.

Results

The mean values and total range of M-AMBI and the constituent metrics for the three datasets are shown in Table 1. The Ekofisk dataset is characterized by the lowest density (from 640 to 3,540 individuals m⁻²), the highest richness and diversity, the lowest BC (corresponding to the ‘best’ condition) and the highest M-AMBI. The dataset includes 14 taxa which are not assigned to any EG. In the Venice dataset, the density has a higher mean and a wider range (from 310 to 24,230 individuals m⁻²), whereas richness and *H'* have lower minima and maxima. The BC and

M-AMBI values are also shifted towards more stressed conditions (six taxa were not assigned to any EG). The simulated dataset is approximately characterized by intermediate values. Overall, M-AMBI ranges between 0.34 and 1.00.

The correlation among the metrics, M-AMBI and derived indices is shown in Table 2. In the Ekofisk dataset, S and H' are directly correlated ($r = 0.51$), BC and H' are inversely correlated ($r = -0.90$), and BC and S are uncorrelated ($r = -0.27$). In the Venice dataset too, S and H' , as well as BC and H' , are correlated ($r = 0.69$ and $r = -0.62$ respectively), with BC and S uncorrelated ($r = -0.30$). The simulated dataset shows no correlation among the three constituent metrics of M-AMBI. In the real datasets, M-AMBI generally shows good agreement with the three metrics, and particularly with H' , with values up to $r = 0.97$ in the Ekofisk dataset. For simulated data, the correlation between M-AMBI and its metrics is lower, with a minimum value recorded between M-AMBI and BC ($r = -0.45$).

The correlation coefficient between M-AMBI and both the simplified versions of the index (M-AMBI* and M-AMBI*_(n)) approximates to one for each of the three single datasets (Table 2), as well as for the whole set of samples. This means that the algorithms are basically equivalent. The mean absolute difference between the results of M-AMBI and M-AMBI* ranges from 0.006 to 0.008 depending on the dataset. The maximum absolute difference is about 0.02 in all datasets. The results of the M-AMBI*_(n) algorithm show still smaller deviation, with maximum absolute difference ranging between 0.011 and 0.002. A comparison of the results of the original M-AMBI algorithm and M-AMBI*_(n) is shown in Fig. 1a.

The correlation between M-AMBI and the index based on just S and BC ('S-AMBI') (Table 2) is highest for the Ekofisk and Venice datasets ($r = 0.99$ and $r = 0.97$, respectively), where the metrics are mutually correlated, but is still high for the simulated

data ($r = 0.85$). Taking into account the whole set of samples, the correlation is 0.95. An XY plot between the indices is presented in Fig. 1b. The correlation between M-AMBI and the index integrating H' and BC ('H'-AMBI') gives slightly lower results.

Discussion

The authors of M-AMBI explicitly state that the algorithm is based on FA (Muxika et al., 2007). The purpose of FA was originally to identify a latent structure within a set of observed variables and to reduce the dimensionality of the data. It was developed in the framework of social sciences and psychology studies (Spearman, 1904; Thurstone, 1931), but it subsequently found application in other contexts, such as quality assessment studies (e.g. Vega et al., 1998; Borja et al., 2004; Riba et al., 2004; Chainho et al., 2007). The FA has been the subject of extensive debate (see for example Armstrong, 1967; Stewart, 1981; Preacher & MacCallum, 2003; Henson & Roberts, 2006). Criticism has focused on the quality and meaningfulness of the results, as they are affected by the criteria used to choose the model, decide how many factors to retain, and select the rotation method, as well as by the tendency to not respect the assumption that measured variables are linearly related to latent variables (Preacher & MacCallum, 2003). FA was applied in the selection of the M-AMBI metrics by Muxika et al. (2007), but was also retained in the algorithm for the ordinary computation of the index. When M-AMBI is calculated, no factor is discarded after PCA is performed, and the Varimax rotation is applied to the original space. Since both of these transformations simply rotate the axes of the three-dimensional space, they have no effect at all on the cloud of points representing the samples, nor on the scalar product which orthogonally projects them on to the 'High'–'Bad' line. The actual role of FA in

Table 1 Mean values (and total range) of total abundance (N), richness (S), Shannon index (H'), AMBI benthic coefficient (BC) and M-AMBI for considered datasets

Dataset	N	S	H'	BC	M-AMBI
Ekofisk	356 (192 ÷ 1,062)	54 (36 ÷ 65)	4.62 (2.29 ÷ 5.23)	2.27 (1.81 ÷ 4.94)	0.86 (0.43 ÷ 0.99)
Venice	827 (31 ÷ 2,423)	15 (5 ÷ 38)	2.03 (0.80 ÷ 4.17)	3.46 (2.58 ÷ 4.29)	0.54 (0.34 ÷ 1.00)
Simulated	216 (15 ÷ 3,517)	32 (13 ÷ 55)	4.00 (1.38 ÷ 5.10)	2.99 (1.43 ÷ 4.34)	0.67 (0.46 ÷ 0.85)

Table 2 Pearson correlation coefficients among richness (S), Shannon index (H'), AMBI benthic coefficient (BC), M-AMBI and derived indices for considered datasets

Dataset		S	H'	BC	M-AMBI	M-AMBI*	M-AMBI* _(n)	S-AMBI _(n)
Ekofisk	H'	0.512						
	BC	−0.265 ^a	−0.898					
	M-AMBI	0.660	0.968	−0.887				
	M-AMBI*	0.709	0.955	−0.856	0.998			
	M-AMBI* _(n)	0.627	0.972	−0.907	0.999	0.994		
	S-AMBI _(n)	0.679	0.928	−0.888	0.990	0.990	0.990	
	H-AMBI _(n)	0.385 ^b	0.968	−0.980	0.947	0.923	0.960	0.930
Venice	H'	0.693						
	BC	−0.303 ^a	−0.620					
	M-AMBI	0.811	0.938	−0.746				
	M-AMBI*	0.805	0.928	−0.763	0.999			
	M-AMBI* _(n)	0.818	0.935	−0.740	1.000	0.999		
	S-AMBI _(n)	0.842	0.816	−0.768	0.966	0.972	0.968	
	H-AMBI _(n)	0.578	0.924	−0.873	0.946	0.949	0.942	0.882
Simulated	H'	0.264 ^a						
	BC	−0.049 ^a	0.177 ^a					
	M-AMBI	0.763	0.596	−0.448 ^c				
	M-AMBI*	0.721	0.577	−0.510	0.997			
	M-AMBI* _(n)	0.765	0.600	−0.441 ^c	1.000	0.997		
	S-AMBI _(n)	0.778	0.086 ^a	−0.666	0.851	0.860	0.849	
	H-AMBI _(n)	0.248 ^a	0.663	−0.619	0.816	0.848	0.814	0.575

All coefficients have P value ≤ 0.001 , except when otherwise indicated

^a $P > 0.05$, ^b $0.01 < P \leq 0.05$, ^c $0.001 < P \leq 0.01$

M-AMBI can be assessed by comparing the results obtained by the original algorithm with the ones obtained by the M-AMBI* algorithm proposed in this study. The correlation approximates to one, showing no substantial difference when FA is excluded. The residual differences (with maximum values of about 2%) are not related to the PCA itself or to the Varimax rotation. Rather, they are due to the fact that, after PCA, the eigenvectors are rescaled by the square root of the eigenvalues, which is a typical step in FA (Revelle, 2012). We consider that this step does not justify the presence of FA in the index.

We suggest that FA in M-AMBI (as well as in the BAT index) is a vestigial component of the index development process and should be omitted. It is in fact a design feature that is no longer functional to the index calculation. According to the *lex parsimoniae*, better known as Ockham's Razor, the model with the least number of assumptions should be preferred, which will result in increased robustness.

One limitation of M-AMBI is that the results depend on the whole set of samples considered, and the addition of new data always leads to different results (Bald et al., 2005; Ruellet & Dauvin, 2008). However, the deviations are very small (Borja et al., 2008b). These differences have generally been interpreted as a drawback of FA; indeed, in the present case, in which all factors are retained and the dimensionality is thus not reduced, and assuming that the minimum and maximum reference values do not change, any differences should more precisely be ascribed to metric standardisation and eigenvector rescaling. To minimise this instability, Borja et al. (2008b) suggested using a sampling dataset of at least 50 samples, a condition which cannot always be fulfilled. When new samples are added to a dataset, discriminant analysis (DA) has been proposed for predicting which class they are most likely to fall into. This method avoids changing the classification of the original samples; however, it attributes the new cases

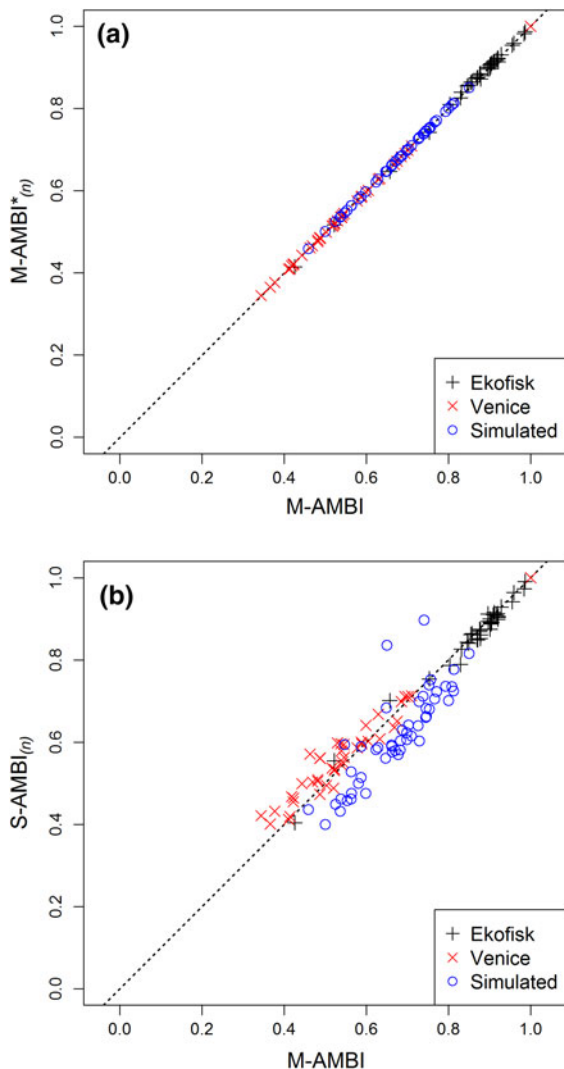


Fig. 1 Relationship between results of M-AMBI and simplified versions M-AMBI*_(n) (a) or S-AMBI_(n) (mean of normalised richness and AMBI-BC) (b) calculated for three datasets. Line with slope = 1 superimposed

to discrete classes (Bald et al., 2005; Muxika et al., 2007). Excluding FA, as in M-AMBI*, also removes eigenvector rescaling. With M-AMBI*_(n), which involves substituting standardisation with minimum–maximum normalisation, the results become independent of other samples, so no minimum number of samples has to be enforced, and samples can freely be added to (or removed from) the dataset. Unlike BC, which ranges from 0 to 6, richness and diversity range from 0 to ∞ , and no theoretical maxima with which to normalise the two variables can be directly identified.

Therefore, for M-AMBI*_(n), we propose to normalise the metrics by means of reference values, such as the ‘High’ and ‘Bad’ values used in M-AMBI, which should not change when samples are added or removed.

Demonstrating that M-AMBI*_(n) is mathematically identical to the simple mean of the three normalised metrics, with equal weight given to each one, is fairly straightforward. This is the simplest way to integrate metrics into a multimetric index. M-AMBI closely approximates this value: $M-AMBI \approx M-AMBI^*_{(n)} = (S_{(n)} + H'_{(n)} + AMBI-BC_{(n)})/3$. Borja et al. (2008b), recalling the results of an ECOSTAT meeting for the implementation of the WFD, indicate a range of 0.05 EQR units as an acceptable deviation in the event of new samples being added to the dataset. In the analysed datasets, the deviations resulting from the proposed simplified algorithms M-AMBI* and M-AMBI*_(n) lie within this recommended range.

As well as formal considerations, there is another point that deserves attention. The software made available by AZTI-Tecnalia for the calculation of AMBI and M-AMBI (Borja et al., 2012b) is ‘freeware’, in the sense that it can be freely downloaded after registration, but is not ‘free and open source software’, since the source code is not available nor can it be unrestrictedly modified. In practice, the calculations performed by the software are hidden. Reproducibility is a main tenet of the scientific method (Stodden, 2011). The authors themselves advise the use of other statistical software for the calculation of the indices (Borja et al., 2008b). We performed all the analyses using the free software R (v. 2.15.0). The script we developed for the calculation of all the versions of M-AMBI is presented in the online resource 1. By reproducing the algorithm, we were able to evaluate the role of subjacent mathematical structures and propose a simplified version. In our opinion, open code and full transparency of the procedures would help to understand the methods and to interpret the results, allowing greater control, promoting methodological debate and, in the final analysis, contributing to the improvement of the methods. In this sense, it is desirable that the list of taxa with assigned EGs for the calculation of AMBI, which is now integrated into the AZTI-Tecnalia software (<http://ambi.azti.es/>), become fully accessible.

M-AMBI has been used to assess benthic quality status in different locations and habitats (a broad list of published works is reported in Borja et al., 2012a). The index inherits some of the limitations of its constituent

metrics. In coastal transitional ecosystems such as lagoons and estuaries, which are naturally characterized by high saprobity, low salinity and strong composite gradients (Tagliapietra et al., 2009, 2012), both AMBI and species richness and diversity can reflect natural stress (Muxika et al., 2005; Teixeira et al., 2008; Munari & Mistri, 2010). Specifically, as AMBI is derived from the Pearson & Rosenberg model (1978), there are some problems with detecting impacts that are not related to organic pollution such as physical impacts (Borja & Muxika, 2005) and with sites characterized by high hydrodynamics (Muxika et al., 2005). We consider that the behaviour of an index is more interpretable when the component metrics are mutually correlated as little as possible. In the real datasets, however, the three constituent metrics, S , H' and AMBI-BC, proved to be correlated. The mutual independence of the metrics in the simulated dataset served to remove the effects of collinearity. In this case too, the M-AMBI results were closely approximated by M-AMBI* and M-AMBI*_(n). On the other hand, this is not common in nature, where opportunistic species are usually characterized by high abundances and are associated with low evenness and diversity (Pearson & Rosenberg, 1978). Nevertheless, when these deviations from classical models occur, they deserve particular attention as they could indicate either inconsistent EG assignment or the presence of drivers other than organic pollution.

Two of the three metrics, species richness and shannon diversity, are closely related aspects of biological diversity, with the latter accounting for the number of species and the way individuals are distributed among species, i.e. species evenness. Accordingly, M-AMBI attributes more weight to the number of species than to the biotic index. It should also be noted that the number of species and related metrics, such as diversity indices, are dependent on sample size, in accordance with habitat specific species-area relationships. Therefore, comparison among datasets is meaningful only if they are characterized by the same sampling area. Reducing M-AMBI to a two-metric index highlights the contribution of the ‘sensitivity’ and ‘diversity’ components without substantially altering the results. Thus, we performed a bivariate M-AMBI*_(n) on AMBI and either S (‘S-AMBI’) or H' (‘H-AMBI’). In the Ekofisk and Venice datasets, both S-AMBI_(n) and H-AMBI_(n) are well correlated with M-AMBI, and there is still a good agreement in the simulated dataset, in which the

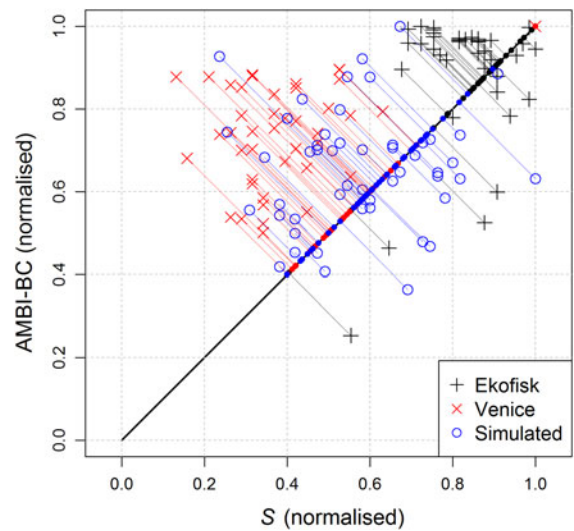


Fig. 2 Results of S-AMBI_(n) index based on richness (S) and AMBI-BC (both metrics are normalised). Value of index is represented by orthogonal projection of samples on line segment (identified by min and max reference samples and set to one)

constituent metrics are mutually uncorrelated. The M-AMBI and S-AMBI_(n) results are compared in Fig. 1b. To be compliant with the WFD, which is highly prescriptive in its terminology, diversity should probably be preferred to species richness, as the former is explicitly requested (European Community, 2000). When averaging two metrics, results and individual contributions can easily be plotted as in Fig. 2. The role of the two metrics in the overall score can be visually distinguished, enhancing the possibility of interpretation. Moreover, the rationale behind the procedure can be understood more easily. The R script in the online resource 1 allows for the calculation of bivariate ‘M-AMBI-like indices’ such as S-AMBI as well. As for BAT compared to M-AMBI (Teixeira et al., 2009), the same method of integration can be applied to different metrics.

Conclusion

M-AMBI is one of the most frequently applied benthic indices in Europe for coastal and transitional waters. A user-friendly free software is provided by the authors for direct calculation of the index. However, the exact program code is not accessible and the user is precluded from fully understanding and controlling the algorithm. A central role has customarily been attributed to the use of FA to integrate the three metrics. However, we argue that

FA in M-AMBI should be discarded, since the index does not benefit from it in any way. Moreover, by substituting standardisation of metrics with their min–max normalisation, the index is transformed into the simple mean of the three equally weighted normalised metrics, therefore, becoming independent of the number of samples. In the analysed datasets, the simplified versions of M-AMBI proposed in this paper produced results that closely approximated those of the original algorithm, with differences in EQR that fell within the accepted range of 0.05. However, we cannot exclude that larger differences would be obtained in other cases, particularly when dealing with a considerably higher number of samples or when the 'High' reference values are markedly lower than the maximum values of the metrics. Nevertheless, as the differences are not systematically negative or positive, it is very likely that the introduction of the proposed algorithms would have no meaningful effect on prior index calibrations or on comparisons with previous results. This approach would be a step towards simplification, stability, robustness, transparency, openness and falsifiability. Furthermore, we introduce a bivariate alternative to M-AMBI combining a diversity (or richness) index and a species sensitivity index, in this case AMBI. The reduction of redundancy is achieved with only a small deviation from the parent index. The latter solution is a distillation of the M-AMBI approach, further increasing generalisation, simplification and ecological interpretability, while at the same time remaining fully compliant with the WFD requirements.

Acknowledgments The authors wish to thank Dr. Stefano Guerzoni (CNR-ISMAR) for his kind support. The authors gratefully acknowledge the editor, P. Viaroli, and two anonymous referees, who helped us improve an earlier version of this manuscript. This work has been funded by the Flagship Project RITMARE—The Italian Research for the Sea—coordinated by the Italian National Research Council and funded by the Italian Ministry of Education, University and Research within the National Research Program 2011–2013.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Armstrong, J. S., 1967. Derivation of theory by means of factor analysis or Tom Swift and his magic factor analysis machine. *The American Statistician* 21: 17–21.
- Bakalem, A., T. Ruellet & J. C. Dauvin, 2009. Benthic indices and ecological quality of shallow Algeria fine sand community. *Ecological Indicators* 9: 395–408.
- Bald, J., Á. Borja, I. Muxika, J. Franco & V. Valencia, 2005. Assessing reference conditions and physico-chemical status according to the European Water Framework Directive: a case-study from the Basque Country (Northern Spain). *Marine Pollution Bulletin* 50: 1508–1522.
- Borja, Á. & I. Muxika, 2005. Guidelines for the use of AMBI (AZTI's Marine Biotic Index) in the assessment of the benthic ecological quality. *Marine Pollution Bulletin* 50: 787–789.
- Borja, Á. & B. G. Tunberg, 2011. Assessing benthic health in stressed subtropical estuaries, eastern Florida, USA using AMBI and M-AMBI. *Ecological Indicators* 11: 295–303.
- Borja, Á., J. Franco & V. Pérez, 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. *Marine Pollution Bulletin* 40: 1100–1114.
- Borja, Á., J. Franco, V. Valencia, J. Bald, I. Muxika, M. J. Belzunce & O. Solaun, 2004. Implementation of the European water framework directive from the Basque country (northern Spain): a methodological approach. *Marine Pollution Bulletin* 48: 209–218.
- Borja, Á., D. M. Dauer, R. Díaz, R. J. Llansó, I. Muxika, J. G. Rodríguez & L. Schaffner, 2008a. Assessing estuarine benthic quality conditions in Chesapeake Bay: a comparison of three indices. *Ecological Indicators* 8: 395–403.
- Borja, Á., J. Mader, I. Muxika, J. Germán Rodríguez & J. Bald, 2008b. Using M-AMBI in assessing benthic quality within the Water Framework Directive: some remarks and recommendations. *Marine Pollution Bulletin* 56: 1377–1379.
- Borja, Á., I. Galparsoro, X. Irigoien, A. Iriondo, I. Menchaca, I. Muxika, M. Pascual, I. Quincoces, M. Revilla, J. G. Rodríguez, M. Santurtún, O. Solaun, A. Uriarte, V. Valencia & I. Zorita, 2011. Implementation of the European Marine Strategy Framework Directive: a methodological approach for the assessment of environmental status, from the Basque Country (Bay of Biscay). *Marine Pollution Bulletin* 62: 889–904.
- Borja, Á., D. M. Dauer & A. Grémarec, 2012a. The importance of setting targets and reference conditions in assessing marine ecosystem quality. *Ecological Indicators* 12: 1–7.
- Borja, Á., J. Mader & I. Muxika, 2012b. Instructions for the use of the AMBI index software (version 5.0). *Revista de Investigación Marina, AZTI-Tecnalia* 19: 71–82.
- Bulmer, M. G., 1974. On fitting the Poisson lognormal distribution to species-abundance data. *Biometrics* 30: 101–110.
- Carletti, A. & A.-S. Heiskanen (eds), 2009. Water Framework Directive Intercalibration Technical Report. Part 3: Coastal and Transitional Waters. EUR 23838 EN/3.
- Chainho, P., J. L. Costa, M. L. Chaves, D. M. Dauer & M. J. Costa, 2007. Influence of seasonal variability in benthic invertebrate community structure on the use of biotic indices to assess the ecological status of a Portuguese estuary. *Marine Pollution Bulletin* 54: 1586–1597.
- Clarke, K. R. & R. N. Gorley, 2006. *PRIMER v6: User Manual/Tutorial*. PRIMER-E, Plymouth.
- Costa-Dias, S., R. Sousa & C. Antunes, 2010. Ecological quality assessment of the lower Lima Estuary. *Marine Pollution Bulletin* 61: 234–239.

- Dauvin, J. C. & T. Ruellet, 2007. Polychaete/amphipod ratio revisited. *Marine Pollution Bulletin* 55: 215–224.
- European Community, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy. Official Journal of the European Communities L327, Bruxelles.
- European Community, 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 Establishing a Framework for Community Action in the Field of Marine Environmental Policy (Marine Strategy Framework Directive). Official Journal of the European Union L164, Bruxelles.
- Glémarec, M. & C. Hily, 1981. Perturbations apportées à la macrofaune benthique de la baie de Concarneau par les effluents urbains et portuaires. *Acta Oecologica, Oecologia Applicata* 2: 139–150.
- Grall, J. & M. Glémarec, 1997. Using biotic indices to estimate macrobenthic community perturbations in the Bay of Brest. *Estuarine, Coastal and Shelf Science* 44(Suppl. A): 43–53.
- Grall, J. & M. Glémarec, 2003. L'indice d'évaluation de l'endofaune côtière. In Alzieu, C. (ed.), *Bioévaluation de la qualité environnementale des sédiments portuaires et des zones d'immersion*. Edition IFREMER: 51–85.
- Gray, J. S., K. R. Clarke, R. M. Warwick & G. Hobbs, 1990. Detection of initial effects of pollution on marine benthos: an example from the Ekofisk and Eldfisk oilfields, North Sea. *Marine Ecology Progress Series* 66: 285–299.
- Grice, J. W., 2001. Computing and evaluating factor scores. *Psychological Methods* 6: 430–450.
- Grøtan, V. & S. Engen, 2008. *poilog: Poisson Lognormal and Bivariate Poisson Lognormal Distribution*. R package version 0.4.
- Hair, J. F. Jr., R. E. Anderson, R. L. Tatham & W. C. Black, 1998. *Multivariate Data Analysis*, 5th ed. Prentice Hall, Upper Saddle River, NJ.
- Harman, H. H., 1976. *Modern Factor Analysis*, 3rd ed. revised. University of Chicago Press, Chicago.
- Hatton-Ellis, T., 2008. The Hitchhiker's Guide to the Water Framework Directive. *Aquatic Conservation: Marine and Freshwater Ecosystems* 18: 111–116.
- Henson, R. K. & J. K. Roberts, 2006. Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educational and Psychological Measurement* 66: 393–416.
- Kaiser, H. F., 1958. The Varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23: 187–200.
- Lasserre, P. & A. Marzollo (eds), 2000. *The Venice Lagoon Ecosystem: Inputs and Interactions Between Land and Sea*. MAB series 25, UNESCO and The Parthenon Publishing Group, Paris, New York, Carnforth.
- Margalef, R., 1958. Information theory in ecology. *General Systems Bulletin* 3: 36–71.
- Mulaik, S. A., 1972. *The Foundations of Factor Analysis*. McGraw-Hill, New York.
- Munari, C. & M. Mistri, 2010. Towards the application of the Water Framework Directive in Italy: assessing the potential of benthic tools in Adriatic coastal transitional ecosystems. *Marine Pollution Bulletin* 60: 1040–1050.
- Muxika, I., Á. Borja & W. Bonne, 2005. The suitability of the marine biotic index (AMBI) to new impact sources along European coasts. *Ecological Indicators* 5: 19–31.
- Muxika, I., Á. Borja & J. Bald, 2007. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. *Marine Pollution Bulletin* 55: 16–29.
- Pearson, T. & R. Rosenberg, 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology Annual Review* 16: 229–311.
- Preacher, K. J. & R. C. MacCallum, 2003. Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics* 2: 13–43.
- R Development Core Team, 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org> (Accessed on 1/10/2012).
- Revelle, W., 2012. *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston. R package version 1.2.1.
- Riba, I., J. M. Forja, A. Gómez-Parra & T. Á. DelValls, 2004. Sediment quality in littoral regions of the Gulf of Cádiz: a triad approach to address the influence of mining activities. *Environmental Pollution* 132: 341–353.
- Rosenberg, R., M. Blomqvist, H. C. Nilsson & A. Dimming, 2004. Marine quality assessment by use of benthic species abundance distributions: a proposed new protocol within the European Union Water Framework Directive. *Marine Pollution Bulletin* 49: 728–739.
- Ruellet, T. & J. C. Dauvin, 2008. Comments on Muxika et al. "Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive" [*Marine Pollution Bulletin* 55 (2007): 16–29]. *Marine Pollution Bulletin* 56: 1234–1235.
- Shannon, C. E. & W. Weaver, 1949. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.
- Simboura, N. & A. Zenetos, 2002. Benthic indicators to use in ecological quality classification of Mediterranean soft bottoms marine ecosystems, including a new biotic index. *Mediterranean Marine Science* 3: 77–111.
- Spearman, C., 1904. General intelligence objectively determined and measured. *American Journal of Psychology* 15: 201–293.
- Stewart, D. W., 1981. The application and misapplication of factor analysis in marketing research. *Journal of Marketing Research* 18: 51–62.
- Stodden, V., 2011. Trust your science? Open your data and code. *AMSTAT News* 409: 21–22.
- Tagliapietra, D., M. Pavan & C. Wagner, 1998. Macrobenthic community changes related to eutrophication in Palude della Rosa (Venetian lagoon, Italy). *Estuarine Coast and Shelf Science* 47: 217–226.
- Tagliapietra, D., M. Pavan & C. Wagner, 2000. Benthic patterns in a salt marsh basin: a snapshot of Palude della Rosa (Venetian Lagoon, Italy). *Wetlands Ecology and Management* 8: 287–292.
- Tagliapietra, D., M. Sigovini & A. Volpi Ghirardini, 2009. A review of terms and definitions to categorise estuaries,

- lagoons and associated environments. *Marine and Freshwater Research* 60: 497–509.
- Tagliapietra, D., M. Sigovini & P. Magni, 2012. Saprobity: a unified view of benthic succession models for coastal lagoon. *Hydrobiologia* 686: 15–28.
- Teixeira, H., F. Salas, J. M. Neto, J. Patrício, R. Pinto, H. Veríssimo, J. A. García-Charton, C. Marcos, A. Pérez-Ruzafa & J. C. Marques, 2008. Ecological indices tracking distinct impacts along disturbance-recovery gradients in a temperate NE Atlantic Estuary – guidance on reference values. *Estuarine Coastal and Shelf Science* 80: 130–140.
- Teixeira, H., J. M. Neto, J. Patrício, H. Veríssimo, R. Pinto, F. Salas & J. C. Marques, 2009. Quality assessment of benthic macroinvertebrates under the scope of WFD using BAT, the Benthic Assessment Tool. *Marine Pollution Bulletin* 58: 1477–1486.
- Thurstone, L. L., 1931. Multiple factor analysis. *Psychological Review* 38: 406–427.
- Thurstone, L. L., 1947. *Multiple Factor Analysis*. University of Chicago Press, Chicago.
- Vega, M., R. Pardo, E. Barrado & L. Deban, 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research* 32: 3581–3592.