Available online at www.sciencedirect.com
ScienceDirect
Marine Pollution Bulletin 54 (2007) 1598–1606
MARINE POLLUTION BULLETIN
www.elsevier.com/locate/marpolbul

Performance comparison of two biotic indices measuring the ecological status of water bodies in the Southern Baltic and Gulf of Lions

Dirk Fleischer [a,*], Antoine Grémare [b], Céline Labrune [b], Heye Rumohr [c], Edward Vanden Berghe [d], Michael L. Zettler [e]

[a] Institute for Polar Ecology of the University of Kiel, Wischhofstr. 1-3, 24148 Kiel, Germany
[b] Laboratoire d'Océanographie Biologique, Université Pierre et Marie Curie Paris 6, CNRS, UMR7621, F66651 Banyuls-sur-Mer cedex, France
[c] Leibniz Institute of Marine Sciences at the University of Kiel (IFM-GEOMAR), Düsternbrooker Weg 20, 24105 Kiel, Germany
[d] Flanders Marine Institute – VLIZ, Wandelaarkaai 7, 8400 Oostende, Belgium
[e] Baltic Sea Research Institute Warnemünde, Seestrasse 15, 18119 Rostock, Germany

Abstract

Two biotic indices, ATZI Marine Biotic Index (AMBI) and Benthic Quality Index (BQI) have been recently introduced within the EC Water Framework Directive to assess the quality of marine habitats: both are based on sensitivity/tolerance classification and quantitative information on the composition of soft-bottom macrofauna. Their performance, especially with regard to sampling effort was assessed based on two data sets collected in Southern Baltic and one from the Gulf of Lions Mediterranean. AMBI was not affected by sampling effort but BQI was. Two modifications were proposed for BQI (i.e., BQI) (1) the removal of the scaling term (i.e., BQI_W), and (2) the replacement of the scaling term by different scaling term (i.e., BQI_{ES}). Both modified BQIs were largely independent of sampling effort. Variability was slightly lower for BQI_W than for BQI_{ES}. BQI was highly correlated with BQI_W and with BQI_{ES} both in the Southern Baltic and in the Gulf of Lions. However, the proportions of stations, which were not attributed the same ecological quality status (EcoQ) when using BQI and its two modified forms were always high. Differences in ecological classification were mostly due to the scales used to infer EcoQ. Based on this study we recommend to use BQI_{ES} in future studies because it apparently constitutes the best compromise in (1) being independent of sampling effort, (2) limiting the variability in computation in relation with sampling effort, (3) being correlated with BQI and corresponding EcoQ.

Keywords: Biotic indices; AMBI; BQI; Sampling effort; Benthic macrofauna; Water Framework Directive; Southern Baltic; Gulf of Lions; Mediterranean

1. Introduction

The European Water Framework Directive (WFD; 2000/60/EC) establishes a basis for the protection of ground, continental, transitional and coastal waters. Its overall goal is to achieve an at least 'Good Ecological Status' for all water bodies defined within the WFD by 2015. The assessment of the status of each water will be based on a large variety of parameters including hydromorphological, physico-chemical and biological ones. Together with phytoplankton, macroalgae and fishes, benthic macrofauna is one of the biological compartments considered by the WFD. The WFD will first include the assessment of the currents EcoQ of each water body and then in the monitoring of these 'Ecological Quality status' (EcoQ). In order to unravel possible artefacts due to natural changes, the WFD recommends the definition of a reference per water

* Corresponding author. Tel.: +49 431 600 1263; fax: +49 431 600 1210.
E-mail address: dfleischer@ipoe.uni-kiel.de (D. Fleischer).

0025-326X/$ - see front matter
doi:10.1016/j.marpolbul.2007.06.011

body. This reference may either correspond to expert judgement, modeling, historical data or data collected at a reference site, which is known to be undisturbed. The term 'Ecological Quality Ratio' (EQR) defines the ratio of the values of the biological parameters at the monitored site by the values of the biological parameters at the reference site. EQR is supposed to vary between 0 and 1. It can be transformed in EcoQ using an appropriate scale (Borja et al., 2007). Temporal changes in EQR are supposed to reflect anthropogenic impacts on EcoQ of the water body irrespective of possible natural changes.

Benthic macrofauna has long been used as an index of habitat quality due to its rapid responses to natural and/or anthropogenic disturbances (Grall and Glemarec, 1997; Borja et al., 2000; Gesteira and Dauvin, 2000; Simboura and Zenetos, 2002; Rosenberg et al., 2004). The theoretical basis for this is the secondary succession theory, which describes spatio-temporal changes in the macrofauna composition related to a disturbance (Pearson and Rosenberg, 1978). The quantitative analysis of benthic macrofauna typically results in species/abundance tables, which can be analysed using a large variety of mathematical procedures including multivariate analyses (Field et al., 1982), ABC curves (Warwick and Clarke, 1994), and biotic indices (Borja et al., 2000; Rosenberg et al., 2004). Biotic indices clearly correspond to an extreme in term of data reduction since their computations involve the transformation of the whole data set in a single number. The use of biotic indices is however clearly favoured for the interpretation of benthic macrofauna data within the WFD (Borja et al., 2000; Rosenberg et al., 2004) because these indices are easier to translate in terms of EcoQ and EQR than the results of multivariate analyses. Biotic indices can be used on their own (Borja et al., 2003; Muxika et al., 2005) but also in conjunction with several other elements to assess the quality of marine habitats (Prior et al., 2004; Muxika et al., 2007).

Two biotic indices have been recently developed in view of the implementation of the WFD: (1) the AZTI Marine Biotic Index (AMBI) (Borja et al., 2000), and (2) the BQI (Rosenberg et al., 2004) (Eq. (1)). Both are largely based on the same paradigm: sensitive species tend to become dominant relative to tolerant species during the secondary succession process. Sensitive species are thus dominant in undisturbed environments, whereas tolerant species dominate in disturbed areas. The basis of calculation for these two indices however is completely different. In AMBI, the level of sensitivity/tolerance of a given species is based on the compilation of existing knowledge and its translation in a discrete value (i.e., Ecological group) between 1 and 5. Sensitive species are attributed a low value conversely to tolerant ones. This results in a single species list, which is available on the web (www.azti.es) and can thus be used for all data sets irrespective of their size.

Conversely the BQI uses a variable concept to integrate the sensitivity/tolerance of a given species in a certain region together with the species richness. The species richness is incorporated directly (Eq. (1)). The estimated species number (ES_n) calculation is the expected number of species within an hypothetical sample of n individuals (e.g., 50 individuals is ES_{50}) based on the composition and the abundance distribution within the original sample (Sanders, 1968; Hurlbert, 1971). The ES_n concept allows to compare the species richness between samples of different sample size. BQI incorporates the sensitivity/tolerance of a species based on the analysis of the studied data set itself. The in BQI used $ES50_{0.05}$ value derives from the function of ES_{50} and the abundance of a single species. The lower 5% of the abundance distribution is defined by Rosenberg et al. (2004) as the sensitivity/tolerance measure and determines a specific ES_{50} value. This ES_{50} value is defined as the species specific sensitivity/tolerance measure $ES50_{0.05}$. Disturbed stations tend to show low ES_{50} values because only few species dominate the species composition with high abundances. $ES50_{0.05}$ indeed constitutes an index of species sensitivity/tolerance levels, with low values associated with tolerant species and high values with sensitive ones (Rosenberg et al., 2004).

$$\mathrm{BQI} = \left(\sum_{i=1}^{n}\left(\frac{A_i}{A_{\mathrm{tot}}} \times \mathrm{ES50}_{0.05i}\right)\right) \times \log(S+1) \qquad (1)$$

A_i Abundance of individuals of species i at the considered station;
A_{tot} Sum (at the considered station) of the abundances of individuals of all species for which it is possible to compute an $ES50_{0.05}$;
$ES50_{0.05i}$ $ES50_{0.05}$ of species i;
S Species richness at the considered station.

The computation of $ES50_{0.05}$ causes a severe limitation to the spread of the use of BQI, which is in practice restricted to large and often heterogeneous data sets characterized by a non uniform sampling effort (Rosenberg et al., 2004; Labrune et al., 2006). Heterogeneity in sampling effort may also be associated with the use of historical data as reference within the WFD. Another important difference between AMBI and BQI is that the later is taking into account species richness through a $\log(S+1)$ term, which is known to increase with sampling effort (Rumohr et al., 2001).

No specific study has been devoted to the effect of sampling effort on either AMBI or BQI. The aims of the present study were (1) to test the sensitivity of AMBI and BQI to sampling effort based on the very large number of replicated macrofaunal samples collected at the same station in the Kiel Bay by Rumohr et al. (2001), (2) hence to aim 1, to propose changes in the computation of these indices to make them independent of sampling effort, (3) to assess the relationships between original and modified indices based on two data sets collected in the Southern Baltic Sea (Zettler et al., 2007) and in the Gulf of Lions (Labrune et al., 2006), and (4) to assess the causes and consequences of differences in EcoQ derived from these original and modified indices.

2. Material and methods

2.1. Data sets

The first macrofauna data set was collected for an ICES/HELCOM intercalibration study (Rumohr et al., 2001). The sampled station 'Millionenviertel 14' is located in the northern part of the Kiel Bay (Fig. 1'A') at a depth of 23 m and has a sandy/mud sediment. Seventy replicated samples were collected using a 0.1 m^2 van Veen grab, sieved on board through a 1 mm mesh, preserved in 4% formaldehyde and the macrofauna was identified to species level if possible.

The second data set was from Zettler et al. (2007) from the Southern Baltic Sea (Fig. 1'B'). This data set consists of composition and abundance of macrofauna at 625 stations sampled by the Baltic Sea Research Institute during the past 10 years (see Zettler et al. (2007) for details). Most samples were taken with a 0.1 m^2 van Veen grab and were sieved with 1 mm meshsize. The sampling included two to three replicates at each station. The macrofauna was identified to species level if possible.

The third data set (Fig. 1'C') was slightly modified from the one collected in the Gulf of Lions by Labrune et al. (2006). The main difference is that the stations sampled by Guille (1971) were not included. Our third data set is based on quantitative data regarding the composition and the abundance of macrofauna collected at 195 sites during 10 surveys between 1994 and 2003. Sampling depths were between 3.8 and 70.0 m. Some sites were sampled several times leading to a total number of 215 stations. All samples were taken with a 0.1 m^2 van Veen grab. For calibration all taxa synonyms and names were checked following the European Register of Marine Species (Costello et al., 2004).

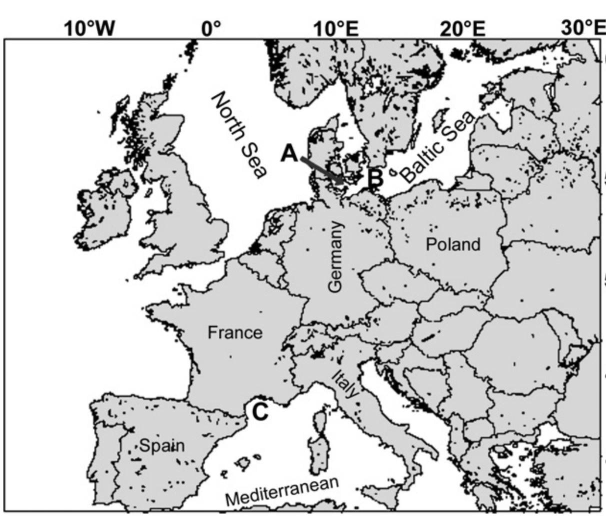

Fig. 1. Map showing the locations of the three dataset used during this study. 'A' indicates the first data set from 'Millionenviertel 14' (Rumohr et al., 2001), 'B' indicates the Baltic Sea data set from Zettler et al. (2007) and 'C' indicates the Gulf of Lions data set from Labrune et al. (2006).

2.2. Calculations

We started with the computation of the Bray–Curtis similarity matrix (Bray and Curtis, 1957) from square root transformed raw data of the first data set. Based on the Bray–Curtis similarities we excluded 7 replicates from the data set due to less then 60% similarity. We reduced the data set homogenization reasons and to eliminate sampling heterogeneity from the analytical data set. The relationships between sampling effort and both AMBI and BQI (Eq. (1)) were assessed using a bootstrap procedure (Efron, 1979) with the remaining 63 replicates. We generated random combinations of 1–63 replicates, we created 63 samples with 0.1 m^2 and 100 different samples each with 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5 and 6.0 m^2 out of many more possible permutations and 1 sample including all 63 replicates. In total we created 1964 different sample combinations from the 63 replicates. By doing so, we stepwise increased the sampled area from 0.1 to 6.3 m^2 and computed both AMBI and BQI for all generated combinations. AMBI was calculated following the recommendations by Borja and Muxika (2005) with the software available at http://www.azti.es in the year 2006. BQI was calculated using a programmed MS-Excel spreadsheet based on the $ES_{50_{0.05}}$ currently available for the Southern Baltic Sea (Zettler et al., 2007) and for the Gulf of Lions (Labrune et al., 2006). The same combinations of replicates were used for both indices. The mean and standard deviations of both indices were computed at each sampling effort. In order to account for the consequences of variability in term of EcoQ classification, we computed a misclassification index (MCI) using the following formula:

$$\mathrm{MCI} = \frac{\mathrm{STD}}{\mathrm{CI}} \times 100 \quad (2)$$

STD is the standard deviation of the estimate of the considered index
CI is the class interval used for the transformation of the same index to EcoQ

For AMBI, MCI was taken as the average class interval in the scale proposed by Borja et al. (2004a). For BQI, MCI was computed based on the data set collected in the Southern Baltic Sea by Zettler et al. (2007).

We explored two possibilities to make BQI (Eq. (1)) independent of sampling effort. The first one consisted in removing the $\log(S+1)$ term (BQI_W, Eq. (3)) and the second one in replacing $\log(S+1)$ by $\log(ES_{50}+1)$ (BQI_{ES}, Eq. (4)).

$$\mathrm{BQI_W} = \sum_{i=1}^{n}\left(\frac{A_i}{A_{\mathrm{tot}}} \times \mathrm{ES50}_{0.05i}\right) \quad (3)$$

$$\mathrm{BQI_{ES}} = \left(\sum_{i=1}^{n}\left(\frac{A_i}{A_{\mathrm{tot}}} \times \mathrm{ES50}_{0.05i}\right)\right) \times \log(\mathrm{ES}_{50} + 1) \quad (4)$$

The relationships between these two modified indices and sampling effort were assessed as described above for AMBI and BQI. The relationships between BQI, BQI_W and BQI_{ES} were assessed based on the second data set from Zettler et al. (2007) and the third data set collected by Labrune et al. (2006).

Simple linear regression models were used to assess the relationships between original and modified BQIs in both areas. The proportions of stations classified differently by BQI, BQI_W and BQI_{ES} were also computed to assess the consequences of modifying BQI in terms of EcoQ classification.

3. Results

3.1. Effects of sampling effort

The AMBI results for the 63 replicates of the 'Millionenviertel 14' station in Kiel Bay converged to 2.96. This would result in the WFD classification 'good' according to the conversion scale proposed by Borja et al. (2004a). While the mean AMBI values were not affected by sampling effort (Fig. 2a, Table 1), the variation coefficients decreased from 4.04% down to 0.11% with increasing sampling effort (1–6.0 m^2) (Fig. 3a). The corresponding MCI decreased accordingly from 9.98% down to 0.27% (Fig. 3b).

Table 1
'Millionenviertel 14' – Comparison of values and misclassification index (MCI) for the biotic indices computed during the present study at two contrasted levels of sampling effort

Index	Value 0.1 m^2	Value 6 m^2	MCI 0.1 m^2 (%)	MCI 6 m^2 (%)
AMBI	2.95	2.96	9.98	0.27
BQI	9.91	14.12	14.8	0.87
BQI_W	8.08	8.11	14.16	0.36
BQI_{ES}	8.65	9.02	20.82	1.85

The same calculations for the three BQI equations showed different results. The mean BQI increased with sampling effort from 9.91 to 14.12 for maximal sampling area (Fig. 2b, Table 1). According to the results of Zettler et al. (2007) this result would be a 'moderate' WFD classification. The mean BQI_W (Eq. (3)) was not affected by sampling effort (Fig. 2c, Table 1) and was between 8.11 (0.1 m^2) and 8.09 (6.0 m^2). This would lead to a 'good' classification according to the range from 1.38 to 12.33 for calculation of BQI_W. The mean BQI_{ES} (Eq. (4)) were almost not affected by increasing sampling effort (Fig. 2d). The 0.1 m^2 value was 8.65 versus 9.02 for the 6.0 m^2 one. Hence to the calculated range (0.5–15.85) is this a 'moderate' classification.

The MCI and the variation coefficient for both derived BQI equations decreased with sample effort. The BQI_W variation coefficient decreased from 4.31% (0.1 m^2) down to 0.11% (6.0 m^2) and the corresponding MCI decreased from 14.16% to 0.36% (Fig. 3a, Table 1). The BQI_{ES} variation coefficient decreased from 7.63% to 0.65% and the MCI decreased 20.82% to 1.85% (Fig. 3b, Table 1).

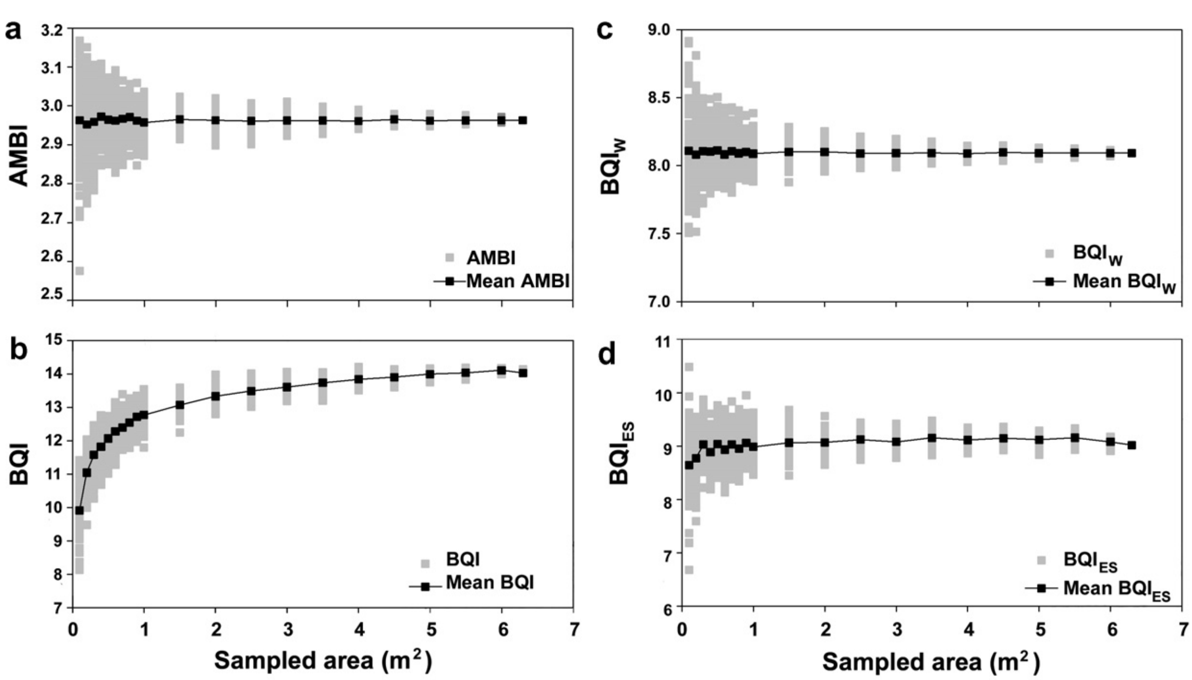

Fig. 2. The 63 independent replicates from the station 'Millionenviertel 14' in the western Baltic Sea (Fig. 1'A') were used for 1964 permutations to simulate increasing sampling effort. The sampling effort plotted against the four different quality indices AMBI (a), BQI (b), BQI_W (c) and BQI_{ES} (d).

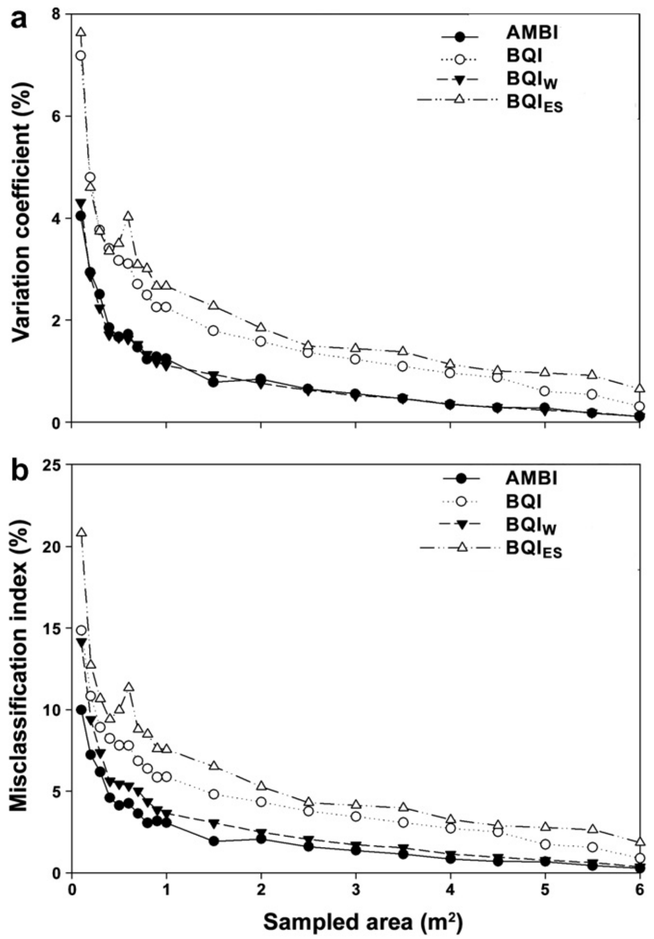

Fig. 3. Relationships between variation coefficients (a) and misclassification indices (b) of AMBI, BQI, BQI_W, and BQI_{ES} versus sampling effort at 'Millionenviertel 14' from the western Baltic Sea.

3.2. EcoQ classification

BQI, BQI_W and BQI_{ES} all correlated positively with each other in the Southern Baltic (Fig. 4a and Fig. 4b) and in the Gulf of Lions (Fig. 4c and Fig. 4d). In both cases, determination coefficients were slightly higher between BQI and BQI_{ES} than between BQI and BQI_W. Differences in the EcoQ obtained using these 3 indices are also highlighted in Fig. 4.

In the Southern Baltic, 58.7% of the stations were not classified in the same EcoQ class when using BQI and BQI_W. Corresponding EcoQs differed by more than one class in 0.3% of the stations. Only 13.1% of the stations were not classified in the same EcoQ class when using BQI and BQI_{ES} (Fig. 4a and Fig. 4c). The EcoQ of all these stations differed by one class only. Overall, BQI_W almost always resulted in a higher EcoQ than BQI (366 cases out of 367). This was also the case for BQI_{ES} (71 cases out of 82). 50.6% of the stations were not classified in the same EcoQ class when using BQI_W and BQI_{ES} (Fig. 5). Again, BQI_W resulted in most cases in a higher EcoQ than BQI_{ES} (311 cases out of 316).

In the Gulf of Lions the proportion of misclassification between BQI and BQI_W was much lower, only 24.7% of the stations were not classified in the same EcoQ class. Corresponding EcoQ differed by more than one class in 0.3% of the stations (Fig. 4a and Fig. 4c). 23.3% of the stations were not classified in the same EcoQ class when using BQI and BQI_{ES} (Fig. 4b and Fig. 4d). The EcoQs of all these stations differed by one class only. As for the Southern Baltic, BQI_W and BQI_{ES} almost always resulted in a higher EcoQ than BQI (51 of 53 cases and 47 cases out of 50, respectively). 14.8% of the stations were not classified in the same EcoQ class when using BQI_W and BQI_{ES} (Fig. 5). Here again, BQI_W almost always resulted in a higher EcoQ than BQI_{ES} (47 cases out of 50).

The frequency distributions of the 5 EcoQ classes derived from AMBI, BQI, BQI_W, and BQI_{ES} are shown in Fig. 6 both for the Southern Baltic and the Gulf of Lions. In both cases, the use of AMBI resulted in the highest EcoQ. Most of the Southern Baltic Sea stations were classified as 'good' using AMBI versus 'poor' using BQI and BQI_{ES}, and 'moderate' using BQI_W. In the Gulf of Lions, all stations were classified as high and good using AMBI, whereas the distributions of EcoQ were much more similar and almost even using BQIs.

The relationships between AMBI ecological groups and the $ES_{50_{0.05}}$ values are provided in Fig. 7 both for the Southern Baltic and the Gulf of Lions. In the Gulf of Lions, there was no trend at all toward decreasing $ES_{50_{0.05}}$ with increasing ranks of AMBI ecological groups. In the Southern Baltic, this trend was not significant either (one-Way ANOVA, $p = 0.348$), although the maximal values of $ES_{50_{0.05}}$ for each ecological group seemed to decrease while the ranks of AMBI ecological group increased.

4. Discussion

The effects of sampling effort on BQI and AMBI are due to (1) changes in the mean values and the variability of these indices, and (2) the consequences on the assessment of EcoQ. Therefore, it is important to discuss the two proposed modifications of BQI to make it sampling effort independent while minimizing the alteration of the relationship between BQI and EcoQ.

4.1. Effect of sampling effort on average values of AMBI and BQI

Our results show that average values of AMBI is not affected by sampling effort, whereas BQI values are affected. As stated above, AMBI only reflects the average for the level of sensitivity/tolerance of the present species (Borja et al., 2000), whereas BQI is based on both sensitivity/tolerance and species richness (Rosenberg et al., 2004). Increase in sampling effort at a given station usually results in the continuous finding of new rare species (Rumohr et al., 2001). There is no *a priori* reason for why the levels of sensitivity/tolerance of those species would differ from

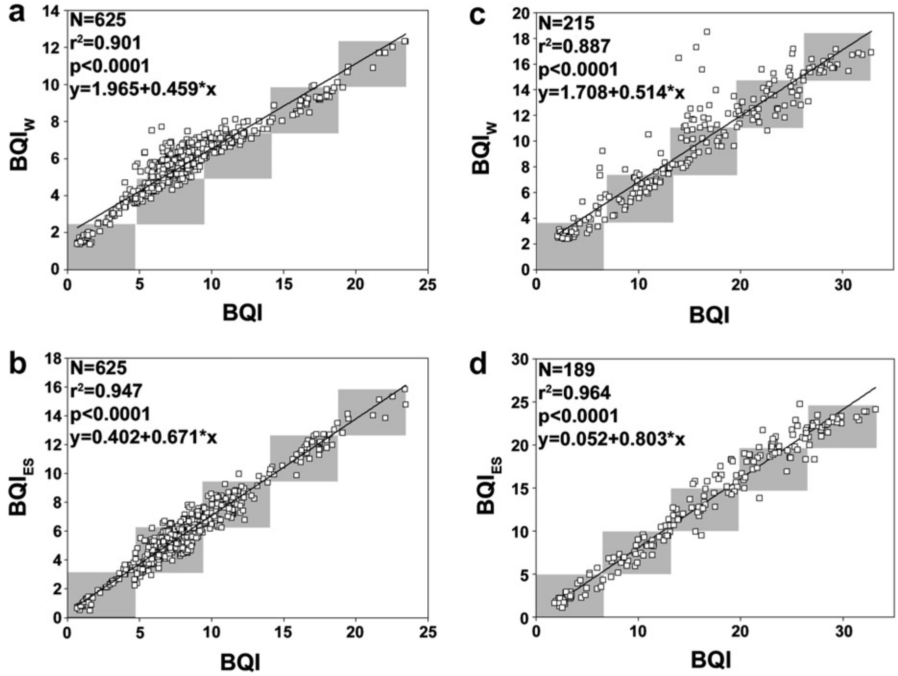

Fig. 4. Showing the large data sets from Southern Baltic (a and b) and Gulf of Lions (c and d). The EcoQ classification relationship between BQI and BQI_W (a and c), and BQI_{ES} (b and d). Shaded areas indicate that similar EcoQs are derived from the two considered indices.

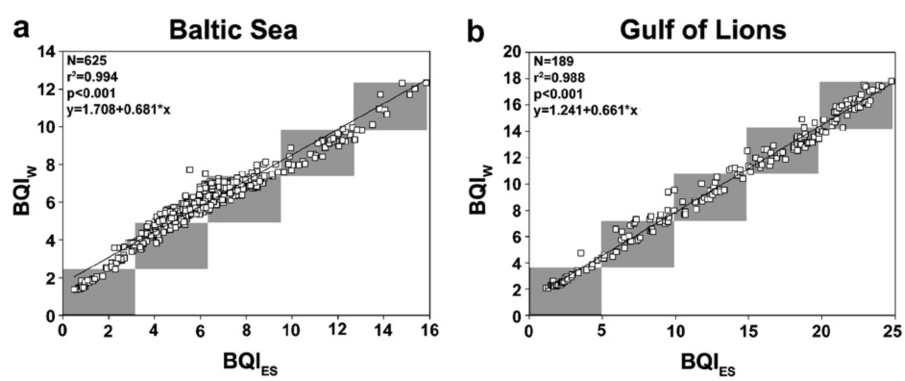

Fig. 5. Showing the large data sets from the Southern Baltic and Gulf of Lions. The EcoQ classification relationship between BQI_W and BQI_{ES}. Shaded areas indicate that similar EcoQs are derived from the two considered indices.

those of more common ones. Moreover, because AMBI is a weighed average of sensitivity–tolerance levels based on species abundance, rare species have a very low impact on the result. Therefore, it is not surprising that average values of AMBI were not affected by sampling effort. Conversely, species richness is well known to increase with sampling effort (Rumohr et al., 2001) and the species accumulation curve constitutes the basis for the computation of the total number of species present at a given site (Karakassis, 1995). The increase of species richness with sampling effort indeed accounted for the increase of average BQI values as indicated by the lack of relationship between sampling effort and BQI_W. The ES_n concept for species richness comparison between different sample sizes has the limitation that ES_n is tightly dependent on dominance pattern. Hurlbert (1971) thus recommended not to use single ES_n values but rather to compare whole species rarefaction curves as suggested by Sanders (1968). ES_{50} nevertheless constitutes the main basis of the computation of sensitivity/tolerance by BQI (Rosenberg et al., 2004). Since the WFD recommends the use of species richness a possible alternative to make BQI independent of sampling effort thus consists in removing $\log(S+1)$ (BQI_W) or replacing by $\log(ES_{50}+1)$ (BQI_{ES}). Our results show that the average values of these two modified indices were largely independent of sampling effort.

4.2. Effects of sampling on the variability of AMBI and BQI

The use of biotic indices such as AMBI and BQI within the WFD is aiming at assessing the EcoQ of transitional and marine waters. Scales have therefore been established

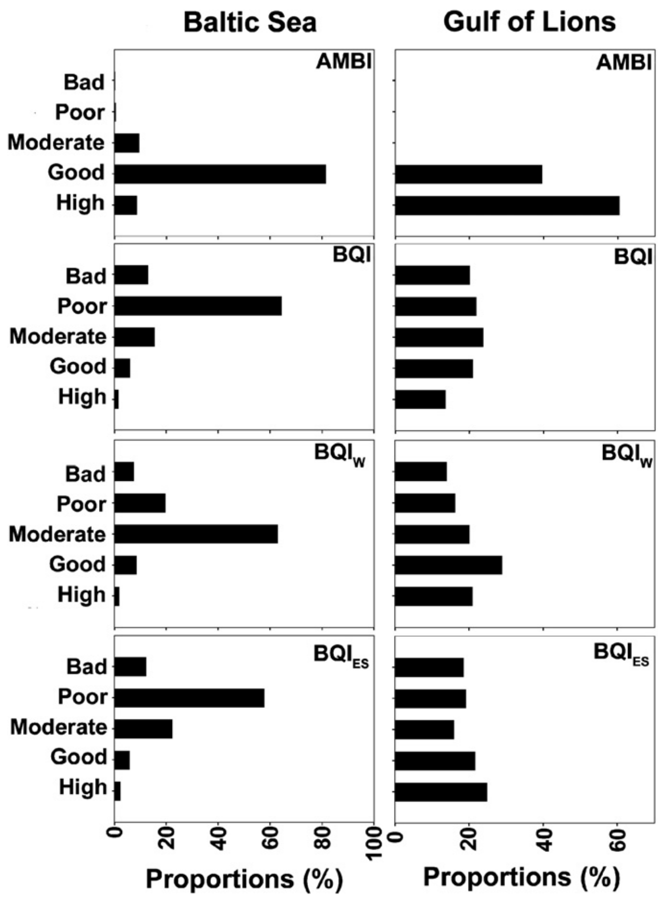

Fig. 6. Showing the large data sets from Southern Baltic Sea and Gulf of Lions and the frequency distribution of EcoQ based on AMBI, BQI, BQI_W, and BQI_{ES} in these two far apart areas.

to convert values of biotic indices in the 5 levels of EcoQ considered by the WFD. These scales largely differ between AMBI and BQI. AMBI is mostly using a single scale irrespective of the studied data set (Borja et al., 2004a), whereas BQI is using several scales (i.e., one per homogeneous habitat) within each studied data set (Rosenberg et al., 2004). The dynamic scales are responsible for intercalibration problems between BQI and other methodologies that maybe used in the WFD. It should however be underlined that it has also recently been proposed to use AMBI in conjunction with other biotic parameters to infer an operational scale (i.e., dependent on the reference conditions of each of the analysed typologies) based on the position of virtual 'good' and 'bad' reference sites along the main stretch within a factorial analysis (Borja et al., 2004b; Bald et al., 2005; Muxika et al., 2007). Sampling effort clearly affects the assessment of EcoQ through changes in the variability associated with the computation of each biotic index. Such differences in station classification are important to consider when evaluating the performance of biotic indices even though they probably preferentially affect stations at the margin between two consecutive EcoQ classes. Variation coefficients were minimal for AMBI and BQI_W and maximal for BQI and BQI_{ES}. Because of the use of different scales (with different class intervals) variation coefficients however did not fully account for the probabilities of changes in EcoQ due to the variability in the computation of each biotic index. These probabilities were assessed through the computation of a misclassification index (MCI). Our results showed that MCI was minimal for AMBI, maximal for BQI_{ES}, and intermediate for BQI and BQI_W. This reflects the fact that both AMBI and BQI_W are integrating variability associated with sensitivity/tolerance, whereas both BQI and BQI_{ES} integrate the variability associated both with sensitivity/tolerance and species richness.

4.3. Relationships between BQIs in the Southern Baltic and in the Gulf of Lions

We used two different data sets to assess the relationships between BQI, BQI_W, and BQI_{ES}. In all cases, there were significant correlations between these 3 indices. Both in the Southern Baltic and in the Gulf of Lions, correlations were slightly higher between BQI and BQI_{ES} than between BQI and BQI_W. The incorporation of $\log(ES_{50}+1)$ however resulted in an increase of only 4.6% and 7.7% of the variance

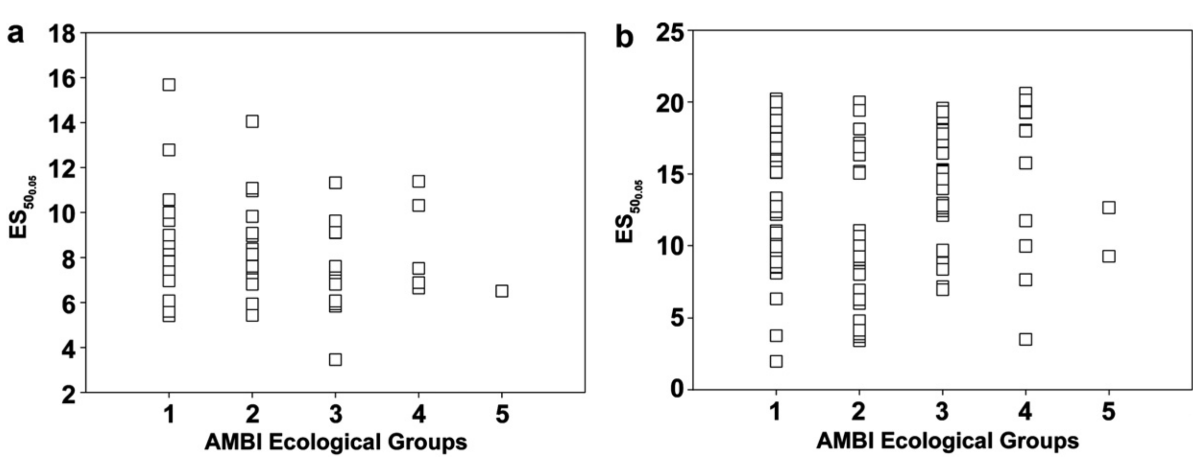

Fig. 7. Showing the relationship between the two different ways of species sensitifity/tolerance level calssification (AMBI ecological groups and $ES_{50_{0.05}}$ calculation) for the large data sets from Southern Baltic (a) and Gulf of Lions (b).

of BQI in the Southern Baltic and in the Gulf of Lions, respectively. In this sense, our results suggest that the information relative to sensitivity/tolerance and species richness are largely redundant and thus support the elaboration of biotic indices only based on these information (Borja et al., 2000; Simboura and Zenetos, 2002).

In spite of such correlations, there were considerable differences in EcoQs computed based on AMBI and BQI as already reported by Reiss and Kröncke (2005) in the North Sea, Labrune et al. (2006) for the Gulf of Lions, Dauvin et al. (2007) for the Seine Estuary and Zettler et al. (2007) for the Southern Baltic. Such a discrepancy was not resulting from heterogeneity in sampling effort within data sets since EcoQs computed based on BQI_W and BQI_{ES} also largely differed from those computed based on AMBI. Differences in EcoQs thus resulted from differences between AMBI and BQI in (1) the assessment of sensitivity/tolerance, and/or (2) scales used to convert biotic indices in EcoQs.

Labrune et al. (2006) already reported an unexpected positive correlation between AMBI and BQI in the Gulf of Lions. This correlation mostly resulted from the strong dominance of a single species (i.e., the polychaete *Ditrupa arietina*), which was classified as a sensitive species by AMBI and as a tolerant one by BQI (Labrune et al., 2006). Our results confirmed the lack of correlation between AMBI ecological groups and $ES_{50_{0.05}}$, which probably did not result from the relatively low number of stations sampled in the Gulf of Lions, since there was also no correlation between these two variables in the Southern Baltic where the number of sampled stations was much higher. Reiss and Kröncke (2005) recently reported a similar result in the North Sea. AMBI and BQI are using very different approaches to assess sensitivity/tolerance (Fig. 7). AMBI is using a compilation of the information in the literature together with the knowledge of individual scientists. BQI is using the information within the studied data set through the $ES_{50_{0.05}}$ concept. Rosenberg et al. (2004) did not specify whether $ES_{50_{0.05}}$ should be computed separately for each habitat. However they stated that the degree of sensitivity/tolerance of a given species may differ between geographic areas–considering the location of the studied populations within the whole geographic range of the species. Along the same line, Zettler et al. (2007) have also underlined the interest of computing different $ES_{50_{0.05}}$ in areas with strong salinity gradient such as the Southern Baltic Sea. Therefore, one cannot exclude that part of the discrepancy between the two indices results from differences in the nature and the overall size of the compiled information. The question of the pertinence of establishing a single assessment of sensitivity/tolerance at the European level, as recommended by Borja and Muxika (2005) is still open to question and it seems now essential to compare the assessments of sensitivity/tolerance by AMBI and BQI based on a larger data set such as the soft-bottom macrofauna data base (MacroBen) recently built up by the European Network of Excellence MarBEF (Marine Biodiversity and Ecosystem Functioning).

The scale used to convert the values of biotic indices in EcoQ is clearly essential in deriving sound EcoQ. The approaches for this are very different between AMBI and BQI. AMBI is basically using a single scale (Borja et al., 2004a; Borja and Muxika, 2005; Muxika et al., 2007). Conversely, the scales used by BQI are tightly linked with the EQR concept and thus highly dependent on the studied data set. These scales are basically obtained by dividing the maximal value of BQI in a given habitat in 5 classes with equal ranges (Rosenberg et al., 2004), which is coherent with the EQR scale used by Denmark within the WFD (Borja et al., 2007), provided that appropriate reference stations are present in the studied data set (Labrune et al., 2006). Our results show that, in the Southern Baltic the frequency distributions of EcoQ derived from BQI and BQI_{ES} on one side, and BQI_W on the other side were different. More specifically EcoQ derived from BQI_W tended to be better than those derived from BQI and BQI_{ES}. This is simply a consequence of the approach used to convert BQI in EcoQ. For simplicity reasons consider that both scales used to derive EcoQ from BQI and BQI_W relies on a single reference site. The highest value of the considered indices is likely to be associated with the highest species richness. The term $\log(S+1)$ is maximal at this site, which results in the strongest possible diminution of each class interval associated with each EcoQ when switching from BQI to BQI_W. Conversely, at stations featuring a lower species richness, $\log(S+1)$ is lower so as the difference between BQI and BQI_W. Consequently, such stations have the tendency to be attributed a higher EcoQ when using BQI_W rather than BQI, which is exactly what was observed in the Southern Baltic. The removal of $\log(S+1)$ clearly modifies the signification of the linear conversion scale relatively to BQI. In order to preserve this relationship while making BQI sampling effort independent, it is therefore advisable to use BQI_{ES} rather than BQI_W. Another point is that BQI_W does not include species richness as specified by the WFD. Overall, there is no reason why the scale used to convert BQI in EcoQ should necessarily be linear and it would probably prove valuable to assess an empirical scale linking BQI and EcoQs as already achieved for AMBI (Borja et al., 2004a). Recently published adjustments to AMBI by developing m-AMBI (Muxika et al., 2007) fulfil the requirements of the WFD by including species richness and diversity. This m-AMBI now has to be evaluated on pan European scale like his forerunner in terms of EcoQ classification.

5. Conclusion

We demonstrated that BQI is sampling effort dependant while AMBI is not. Additionally we propose modifications like BQI_W (Eq. (3)) and BQI_{ES} (Eq. (4)) to achieve sampling effort independent BQI varieties and tested these. The variability of BQI_W is lower than BQI_{ES} and both variations of BQI correlat with the original BQI. Still the EcoQ classification of stations changed, because the linear classification scale for BQI changed to a non linear scale when using BQI_W. For this reason we recommend to use BQI_{ES} in future studies instead of BQI because it apparently constitutes the best compromise in (1) being independent of sampling effort, (2) limiting the variability in computation in relation with sampling effort, (3) being correlated with original BQI values and corresponding EcoQ. The two approaches used by AMBI and BQI to assess sensitivity/tolerance nevertheless led to rather different and largely inconsistent results both in the Southern Baltic and in the Gulf of Lions. This underlines the necessity of comparing both AMBI ecological groups and $ES_{50_{0.05}}$ values on a much larger pan European data set. Especially since AMBI and BQI have to deal in this context with large ranges of natural environmental factors. Definitions of salinity ranges and/or subarea species lists probably have to be considered and tested on the pan European scale.

Acknowledgements

This work was carried out within the MarBEF network of excellence and was initiated during a workshop held in Crete during 2005. We would like to deeply thank Christos Arvanitidis (HCMR) for organizing this workshop.

References

Borja, A., Muxika, H., 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, A., Franco, J., Perez, V., 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, A., Muxika, I., Franco, J., 2003. The application of a marine biotic index to different impact sources affecting soft-bottom benthic communities along European coasts. Marine Pollution Bulletin 46, 835–845.

Borja, A., Franco, J., Muxika, I., 2004a. The biotic indices and the Water Framework Directive: the required consensus in the new benthic monitoring tools. Marine Pollution Bulletin 48, 405–408.

Borja, A., Franco, J., Valencia, V., Bald, J., Muxika, I., Belzunce, M., Solaun, O., 2004b. Implementation of the european water framework directive from the basque country (northern spain): a methodological approach. Marine Pollution Bulletin 48, 209–218.

Bald, J., Borja, A., Muxika, I., Franco, J., Valencia, V., 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, A., Josefson, A., Miles, A., Muxika, I., Olsgard, F., Phillips, G., Rodriguez, J., Rygg, B., 2007. An approach to the intercalibration of benthic ecological status assessment in the North Atlantic ecoregion, according to the European Water Framework Directive. Marine Pollution Bulletin 55, 42–52.

Bray, J., Curtis, J., 1957. An ordination of the upland forest of Southern Wisconsin. Ecological Monographs 27, 225–349.

Costello, M., Bouchet, P., Boxhall, G., Emblow, C., Vanden Berghe, E., 2004. European register of marine species. <http://www.marbef.org/data/erms.php>.

Dauvin, J.-C., Ruellet, T., Desroy, N., Janson, A.-L., 2007. The ecological status of the Bay of Seine and the Seine estuary: use of biotic indices. Marine Pollution Bulletin 55, 241–257.

Efron, B., 1979. Bootstrap methods: another look at the jacknife. Annals of Statistics 7, 1–26.

Field, J.G., Clarke, K.R., Warwick, R.M., 1982. A practical strategy for analysing multispecies distribution patterns. Marine Ecology Progress Series 8 (1), 37–52.

Gesteira, J., Dauvin, J., 2000. Amphipods are good bioindicators of the impact of oil spills on soft-bottom macrobenthic communities. Marine Pollution Bulletin 40, 1017–1027.

Grall, J., Glemarec, M., 1997. Using biotic indices to estimate macrobenthic community perturbations in the Bay of Brest. Estuarine Coastal And Shelf Science 44, 43–53.

Guille, A., 1971. Bionomie benthique du plateau continental de la côte Catalane Française. II – Les communautés de la macrofaune. Vie et Milieu 21, 149–280.

Hurlbert, S., 1971. The non-concept of species diversity: a critique and alternative parameters. Ecology 52, 577–586.

Karakassis, I., 1995. S_{∞}: a new method for calculating macrobenthic species richness. Marine Ecology Progress Series 120, 299–303.

Labrune, C., Amouroux, J., Sarda, R., Dutrieux, E., Thorin, S., Rosenberg, R., Grémare, A., 2006. Characterization of the ecological quality of the coastal Gulf of Lions (NW Mediterranean). a comparative approach based on three biotic indices. Marine Pollution Bulletin 52, 34–47.

Muxika, I., Borja, A., Bonne, W., 2005. The suitability of the marine biotic index (AMBI) to new impact sources along European coasts. Ecological Indicators 5, 19–31.

Muxika, I., Borja, A., 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. Marine Pollution Bulletin 55, 16–29.

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

Prior, A., Miles, A., Sparrow, A., Price, N., 2004. Development of a classification scheme for the marine benthic invertebrate component, water framework directive. Phase I & II- transitional and costal waters. Tech. rep., Environmental Agency, National Marine Service, pp. 1–85.

Reiss, H., Kröncke, I., 2005. Seasonal variability of benthic indices: an approach to test the applicability of different indices for ecosystem quality assessment. Marine Pollution Bulletin 50, 1490–1499.

Rosenberg, R., Blomqvist, M., Nilsson, H., Cederwall, H., Dimming, A., 2004. Marine quality assessment by use of benthic species-abundance distribution: a proposed new protocol within the European Union Water Framework Directive. Marine Pollution Bulletin 49, 728–739.

Rumohr, H., Karakassis, I., Jensen, J., 2001. Estimating species richness, abundance and diversity with 70 macrobenthic replicates in the Western Baltic Sea. Marine Ecology Progress Series 214, 103–110.

Sanders, H., 1968. Marine benthic diversity: a comparative study. American Naturalist 102, 243–280.

Simboura, N., Zenetos, A., 2002. Benthic indicators to use in ecological quality classification of Mediterranean soft bottom marine ecosystems, including a new biotic index. Mediterranean Marine Science 3, 77–111.

Warwick, R., Clarke, K., 1994. Relearning the ABC – taxonomic changes and abundance biomass relationships in disturbed benthic communities. Marine Biology 118, 739–744.

Zettler, M., Schiedek, D., Bobertz, B., 2007. Benthic biodiversity indices versus salinity gradient in the southern Baltic Sea. Marine Pollution Bulletin 55, 258–270.