Ecological Indicators 60 (2016) 420–427

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Application of signal detection theory approach for setting thresholds in benthic quality assessments

Romualda Chuševė[a,*], Henrik Nygård[b], Diana Vaičiūtė[a], Darius Daunys[d], Anastasija Zaiko[a,c]

[a] Marine Science and Technology Centre, Klaipėda University, H. Manto 84, LT 92294 Klaipėda, Lithuania
[b] Finnish Environment Institute SYKE, Marine Research Centre, Mechelininkatu 34a, P.O. Box 140, FI-00251 Helsinki, Finland
[c] Coastal and Freshwater Group, Cawthron Institute, 98 Halifax Street East, 7010 Nelson, New Zealand
[d] Department of Biology and Ecology, Faculty of Natural Sciences and Mathematics, Klaipėda University, H. Manto 84, LT 92294 Klaipėda, Lithuania

ARTICLE INFO

Article history:
Received 9 January 2015
Received in revised form 14 July 2015
Accepted 19 July 2015
Available online 4 August 2015

Keywords:
Macrozoobenthos
GES
BQI
Sensitivity
Specificity
Baltic Sea

ABSTRACT

The European Marine Strategy Framework Directive requires EU Member States to prepare national strategies and manage their seas to achieve good environmental status (GES) by 2020. There are many multimetric indices proposed as indicators of the ecological quality of the benthic environment. Their functionality and utility are extensively discussed in the literature. Different frameworks are suggested for comparative assessments of indicators with no agreement on a standardized way of selecting the most appropriate one. In the current study, we apply signal detection theory (SDT) to evaluate the specificity and sensitivity of the Benthic Quality Index (BQI), its response to the eutrophication pressure, and its performance under the effects of estuarine water outflow. The BQI showed acceptable response to total nitrogen, total phosphorus and chlorophyll-*a* concentrations in the study area. Based on the results, we suggest using SDT for setting GES thresholds in a standardized way. This aids a robust assessment of the environmental status and supports differentiation between the quality classes.

1. Introduction

The European Marine Strategy Framework Directive (MSFD) requires EU Member States to align national legislative policies and appropriately manage their seas in order to achieve good environmental status (GES) by 2020 (MSDF; European Commission 2008/56/EC). GES is defined as 'clean, healthy and productive seas within their intrinsic conditions, and the sustainable use of the marine environment'. The directive requires application of a set of indicators for environmental status assessment. When GES criteria are not met, the corresponding measures for achieving them must be specified and undertaken.

Obviously, an adequate and efficient management strategy for the improvement of environmental status implies a robust and reliable status assessment. The crucial step here is the selection of appropriate indicators, therefore many research projects specifically address this issue (Ferreira et al., 2011; Rice et al., 2012; ICES, 2013). A few selection criteria have been suggested, including (but not limited to) scientific basis, responsiveness, range of applicability, data availability, practicality, harmonization, accuracy and confidence (Rice and Rochet, 2005; Niemeijer and de Groot, 2008; Elliott, 2011). Several evaluation methods and conceptual frameworks have been discussed to facilitate decision-making (Borja and Dauer, 2008; Kershner et al., 2011; ICES, 2013). The responsiveness of an indicator is often distinguished among the selection criteria (Rombouts et al., 2013). Once an indicator has been developed, its performance in terms of sensitivity (response to an existing disturbance), specificity (resistance to the noise or non-targeted disturbances) and the accuracy in relation to the actual response can be evaluated (Murtaugh, 1996).

It is assumed that benthic species and communities reflect natural and anthropogenic changes in marine ecosystems as they are unable to avoid unfavourable conditions, have a long reproductive cycle, accumulate changes over time and occur at various depths (Zettler et al., 2007). A series of multimetric indices have been proposed to supply synoptic information about the state and ecological quality of the benthic environment, e.g. the Benthic Quality Index (BQI; Rosenberg et al., 2004; Leonardsson et al., 2009), the AZTI Marine Biotic Index (AMBI; Borja et al., 2000), the Biotic Index (BENTIX; Simboura and Zenetos, 2002), the Benthic Opportunistic Polychaeta Amphipoda Index (BOPA; Dauvin and Ruellet, 2007) and the Benthic Opportunistic Annelida Amphipods Index (BO2A;

* Corresponding author. Tel.: +370 67403484.
E-mail address: romualda.chuseve@jmtc.ku.lt (R. Chuševė).

http://dx.doi.org/10.1016/j.ecolind.2015.07.018
1470-160X/

Dauvin and Ruellet, 2009). Yet the performance of these indicators is unlikely to be consistent across habitats and ecosystems, since bottom-dwelling organisms are not equally sensitive to different types of anthropogenic and natural disturbances (Buhl-Mortensen et al., 2009), or environmental conditions (Tagliapietra et al., 2009). Many authors agree that eutrophication, chemical pollution and mechanical disturbance of the sea bottom are the major anthropogenic pressures determining changes in macrofauna abundance, distribution and species composition (McQuatters-Gollop et al., 2009; Van Hoey et al., 2010; Rice et al., 2012). Among those, eutrophication is often emphasized as a particularly large-scale driving force of ecosystem changes, having multiple indirect effects and therefore not being easily quantifiable by direct measurements (Van Hoey et al., 2010). Therefore, detection of eutrophication effects relies mostly on the sensitivity of selected indirect measurements and synoptic indicators (such as benthic indices).

Many studies have aimed to test and validate benthic indicators, applying different analytical frameworks and statistical approaches. For instance, the responsiveness of the BENTIX index (Simboura and Zenetos, 2002) to water quality parameters (dissolved oxygen, particulate and total organic carbon) was assessed using linear regression. Factorial analysis was used by Muxika et al. (2007) when validating benthic quality assessment performed with the AMBI Index (Borja et al., 2000). Diaz et al. (2004) assessed the functionality of 64 benthos-related indices applying qualitative comparison based on a comprehensive literature review.

Among different frameworks suggested for quality analysis of GES indicators, there is still little agreement on a uniform approach for a robust and standardized selection of appropriate metrics (Mazik et al., 2010; HELCOM, 2012). Here, we demonstrate the application of signal detection theory (SDT) to identify and quantify the indicator response to a particular anthropogenic pressure. This method has been extensively used in medical studies, but has also been considered for ecological application (Murtaugh, 1996; Hale and Heltshe, 2008). In the current study, we assess the specificity and sensitivity of the Benthic Quality Index (BQI), its response to the eutrophication pressure, and its performance in relation to the soft-bottom habitats affected by estuarine water outflow.

2. Materials and methods

2.1. Study area

The performance of the BQI was assessed in relation to the soft-bottom habitats in the Lithuanian coastal zone, south-eastern Baltic Sea (Fig. 1). Due to high wave exposure there is no oxygen deficiency in the near-bottom layer. Salinity in the study area varied from 6.3 to 7.4‰ outside the plume and decreased down to 3.3‰ in the areas exposed to a freshwater outflow from the Curonian Lagoon (the plume zone). Approximately 60 different benthic macrofauna species have been reported in this area (Olenin et al., 1996). Hard-bottom communities are dominated by the blue mussel *Mytilus edulis* and the barnacle *Amphibalanus improvisus*, whereas sandy bottoms are dominated by the spionid polychaetes *Pygospio elegans* and *Marenzelleria* sp. or the bivalve *Macoma baltica* (Bubinas and Vaitonis, 2003; Olenin and Daunys, 2004). Eutrophication is considered to be one of the main pressures affecting water quality in the study area (Olenin and Daunys, 2004).

2.2. Data collection

A long-term (May–September samplings between 1984 and 2012) benthic macrofauna data set covering six monitoring sites (Fig. 1) was used for assigning the species sensitivity values ($ES_{50\text{-}0.05}$), as described by Leonardsson et al. (2009). For the BQI calculation and responsiveness analysis, data (2005–2011) on macrofauna diversity and abundance (ind/m^2), and summer averages (June–August) of total phosphorus (TP mg/l), total nitrogen (TN mg/l) and chlorophyll-*a* (chl-*a* μg/l) concentrations were used. These parameters were chosen as "direct measures" of eutrophication, suggested among others within the MSFD (Ferreira et al., 2011).

Benthic samples were collected from the soft-bottom habitats at depths ranging from 13–20 m, sieved on-site through a 0.5 mm mesh and processed according to the standard HELCOM recommendations (COMBINE manual). Data on TP and TN were collected as part of the national monitoring programme (unpublished data, Environment Protection Agency), and chl-*a* data were retrieved from the MEdium Resolution Imaging Spectrometer (MERIS), the ENVISAT satellite of the European Space Agency.

The final data set used for the analysis consisted of 77 samples collected from six locations (Fig. 1) within the coastal zone.

2.3. BQI index calculations

When testing the responsiveness of the BQI to the eutrophication pressure (expressed by TP, TN and chl-*a* concentrations), a one-year lag was applied for the index values in respect of pelagic parameters. Instant effects (no lag) were less likely in our study due to the timing of pelagic and benthic samplings (June–August and May–September respectively). One-year lag was also supported by the best statistical response using multiple linear regression ($r = 0.30$, $p = 0.08$) of the BQI to environmental variables compared to no or two-year lag applications ($r = 0.06$, $p = 0.80$ and $r = 0.04$, $p = 0.86$ respectively).

Since the original version of the BQI (Rosenberg et al., 2004) is known to be sampling effort dependent (Fleischer et al., 2007), the adjusted calculation was applied (Fleischer and Zettler, 2009)

$$\mathrm{BQI} = \left(\sum_{i=1}^{n} \left(\frac{A_i}{A_{tot}} \times ES50_{0.05i} \right) \right) \times \log(ES_{50} + 1) \times \left(1 - \frac{5}{5 + A_{tot}} \right) \tag{1}$$

In the above equation, n denotes the observed species number. A_i stands for the abundance of the species i (ind m^{-2}) and A_{tot} is the sum of all individuals (ind m^{-2}). Finally, $ES50_{0.05i}$ is the sensitivity/tolerance value for the species i and ES_{50} denotes the estimated species number among 50 randomly picked individuals within a square metre (Hurlbert Index). The sensitivity value of a species was set to the 5th percentile of the ES_{50} ($ES50_{0.05i}$) in the samples where the species was present.

2.4. Signal detection theory

According to SDT, the sensitivity and specificity of an indicator can be calculated according to four possible outcomes – hits (correct interpretation of a true response – true positives), misses (inability to detect a true response – false negatives), false alarms (false detection of a response – false positives) and correct rejections (correctly interpreted missing response – true negatives) – given that the target condition ("gold standard") is known. Receiver operating characteristic (ROC) curves provide a visual tool for assessing the accuracy of an indicator, by plotting the probability of the true positives (sensitivity) against the probability of the true negatives (specificity). The area under the ROC curve (AUC) can be used as a measure of the indicator response. A perfect indicator should have an AUC of 1, whereas 0.5 is a measure of a non-informative indicator (Murtaugh, 1996). In ecological studies, AUC values ≥0.8 are considered to indicate an excellent and ≥0.7 an acceptable response (Hale and Heltshe, 2008).