Ecological Indicators 12 (2012) 1–7

Contents lists available at ScienceDirect

Ecological Indicators

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

The importance of setting targets and reference conditions in assessing marine ecosystem quality

Ángel Borja[a,*], Daniel M. Dauer[b], Antoine Grémare[c]

[a] AZTI – Tecnalia, Marine Research Division, Herrera Kaia, Portualdea s/n, 20110 Pasaia, Spain
[b] Department of Biological Sciences, Old Dominion University, Norfolk, VI 23529, USA
[c] Station Marine d'Arcachon, UMR 5805, Université Bordeaux 1 - CNRS, 2 rue du Professeur Jolyet, F33120 Arcachon, France

ARTICLE INFO

Article history:
Received 2 February 2011
Received in revised form 17 June 2011
Accepted 19 June 2011

Keywords:
Reference conditions
Targets
Benthic status
Ecological status
Indicators
M-AMBI

ABSTRACT

Assessing benthic quality status of marine and transitional water habitats requires to set up both: (i) tools (i.e. indices) to assess the relative quality of the considered habitat, and (ii) reference conditions for which such indices can be computed and used to infer the absolute ecological status (ES) of the considered habitat. The development of indices, their comparison and the assessment of the causes of their discrepancies have been largely discussed but less attention has been paid to the methods used for the setting of adequate reference conditions, although this step is clearly crucial for the sound assessment of ES. This contribution reviews the approaches available in setting both reference conditions (pristine areas, hindcasting, modelling and best professional judgment) and targets (baseline set in the past, current baseline and directional/trends). We scored the use of pristine or minimally impacted conditions as the best single method; however, the other methods were judged as adequate then combined with best professional judgment. The case of multivariate AMBI (AZTI's Marine Biotic Index) is used to highlight the importance of setting correct reference conditions. Hence, data from 29 references, including 14 countries from Europe and North America, and both coastal (15 cases) and transitional (17 cases) waters, have been used to study the response of multivariate AMBI to human pressures. Results show that the inability of this index to detect human pressure is in most cases linked with the use of inappropriate methods for setting reference conditions.

1. Introduction

Marine legislation worldwide (Clean Water Act (CWA) or Oceans Act in USA, Australia or Canada; Water Framework Directive (WFD) or Marine Strategy Framework Directive (MSFD) in Europe, National Water Act in South Africa, etc.) is addressing ecological quality or integrity of ecosystems, due to increasing human pressures, that can degrade these ecosystems if no mitigation occurs (Halpern et al., 2008; Stelzenmüller et al., 2010). These legislations seek to define quality in an integrative way, by using several biological elements, together with physico-chemical and pollution elements. Such an approach allows for the assessment of ecological status at the ecosystem level ('ecosystem approach' or 'holistic approach' methodologies), rather than at species or chemical levels alone (Borja et al., 2008a, 2010a).

The need for integrative tools to assess ecosystem quality is very important, both from a scientific and stakeholder point of view. Politicians and managers need information derived from simple and pragmatic, but scientifically sound methodologies, to show clearly to society the change in the ecological quality of a given geographical zone (estuary, coastal area, etc. reflective of human pressures or recovery processes) (Borja et al., 2008a). These methodologies include: (i) multidisciplinarity, the inclusion of knowledge and understanding of all disciplines of the natural and social sciences; (ii) integration of biotic and abiotic factors; (iii) determination of ecological integrity using accurate and validated methods; (iv) ability to establish the relationships between drivers, their pressures (stressors) and impacts (effects) providing managers with a scientific basis upon which to build a consensus on restoring degraded ecosystems; and (v) adequate indicators to follow changes of the monitored ecosystems (a measure of the state of the ecosystem) until total recovery or attainment of targeted goals. These approaches are inherent elements of the well known DPSIR paradigm (Drivers-Pressures, State-Impacts-Response), used extensively in marine waters (Elliott, 2002, 2011; Aubry and Elliott, 2006; Borja et al., 2006; USEPA, 2006; Atkins et al., 2011).

One major difficulty consists in transforming the values of computed indicators in ecological status (ES), i.e. in assessing the position of the considered habitat along a disturbance gradient. This

* Corresponding author. Tel.: +34 943 004 800.
E-mail address: aborja@azti.es (Á. Borja).

1470-160X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ecolind.2011.06.018

is usually achieved in three steps: (i) the definition of a reference condition, i.e. the conditions of the indicator in absence of human pressure or pristine areas; (ii) the computation of the considered indicator for this reference conditions and the computation of an Ecological Quality Ratio (EQR) (generally the ratio of the value of the indicator for the considered habitat and for the reference condition, *sensu* the European WFD); and (iii) the use of a conversion scale for transforming the EQR into an ES. The development and comparison of indicators has been largely enhanced by the development of national and international legislation (see for example Labrune et al., 2006; Grémare et al., 2009). Hence, in the above mentioned legislation specific restoration approaches and anticipated rates of ecosystem recovery are hampered by the lack of understanding and/or identification of controlling factors, exemplified by the paucity of literature on this topic (see Borja et al., 2010b).

However, other legislation, such as the MSFD, is seeking for environmental targets rather than for reference conditions, in order to guide progress toward achieving good environmental status. In this particular case, it is acknowledged that humans are part of the marine ecosystem (as users) and, then, the activities can produce a certain impact that makes unable finding pristine areas. Hence, in order to get sustainable activities, compatible with the conservation of marine ecosystems, some environmental targets for a good status must be defined.

Without long-term data from the analysis of biological and chemical indicators, and a full understanding of historical human activities, it is difficult to assess where an ecosystem is positioned along a trajectory toward recovery (Latimer et al., 2003). This clearly underlines the requirement of sound references and/or targets corresponding to good ecological status. In the contact of the DPSIR approach, an accurate assessment of the ecological quality status of an ecosystem (Gibson et al., 2000; Heiskanen et al., 2004) demands that clear targets and reference conditions to compare actual ecosystem state with the final objectives to restoration (habitat type-specific reference conditions).

The authors of this contribution organised an International Council for the Exploration of the Sea (ICES) Theme Session entitled “Benthic indicators: responding to different human pressures and assessing integrative quality status”, within the Annual Science Conference (ASC), held in September 2010, in Nantes, France. From the 37 communications presented there (some of them published in this special issue) it appeared that many methods currently used in assessing ES do not include sound reference conditions and targets. Hence, the present contribution, as introduction to this special issue, discusses the current situation of setting reference and target conditions, when assessing the status of marine ecosystems, and underpins the importance of a sound setting of such conditions.

2. Methods for setting reference conditions and targets

Reference conditions are optimally defined/described from data (i) best acquired from multiple sites with similar physical characteristics, within an ecoregion and habitat type; (ii) that ideally represent minimally impaired or undisturbed conditions (i.e. absence or minimal human pressure); and (iii) that provide an estimate of the variability in biological communities and habitat quality due to natural physical and climatic factors (see http://water.epa.gov/scitech/swguidance/waterquality/standards/criteria/aqlife/biocriteria/ref2.cfm).

When none of the traditional reference condition approaches are applicable (using pristine areas, hindcasting, modelling or best professional judgment), alternative concepts are necessary. For example, the European MSFD stipulates the need to establish a comprehensive set of environmental targets, that implicitly indicate conditions where the indicator in question is not adversely affected or is only slightly affected, i.e. a *de facto* reference condition (see Borja et al., 2010a, 2011b). Hence, this approach still requires us to know whether activities are sustainable and, if not, can this target be achieved by the removal of pressures (Mee et al., 2008).

Different legislation (WFD, CWA) and regional seas conventions (Oslo-Paris Convention-OSPAR, Helsinki Convention-HELCOM) have proposed assessment systems based upon reference conditions setting (see Gibson et al., 2000; CIS, 2003a; Nielsen et al., 2003; Heiskanen et al., 2004; Topcu et al., 2009; Backer et al., 2010; HELCOM, 2010), as shown in Table 1. Hence, for setting reference conditions there are several alternative approaches (Fig. 1 and Table 1), which are described below.

Reference conditions using pristine areas or least disturbed areas: this is the preferred method recommended by the European WFD and also by the US Environmental Protection Agency (EPA, Gibson et al., 2000). However, it is generally recognized that marine and estuarine pristine habitats are rare. Indeed rigidly defining a pristine system as having no significant human impact does not seem possible in the context of human impacts on global climate change. Hence, resource managers must decide on an acceptable level of disturbance to represent an achievable or existing reference condition, e.g. set in relation to what is considered to be unimpacted or minimally impacted condition within sites or similar ecosystems where human pressures are negligible. Such reference conditions do not equate to totally undisturbed, pristine conditions, but include a level of human pressure where there are no or only very minor ecological effects (CIS, 2003a,b). It must be noted that when reference conditions are based upon minimally impacted systems the approach includes the use of some level of best professional judgment in assessing acceptable levels of human pressure. In general, this approach scored high in all items studied in Table 1, because it is considered the most transparent, applicable, practical and confident method.

Historical reference conditions (hindcasting): the conditions are set in relation to what is considered to be unimpacted condition based upon available historical information (CIS, 2003a). The biggest problem in hindcasting is determining when is the baseline date, e.g. any studied area can be compared against a time when there was no population there but this is not sensible or reasonable; hence, the use of dates with low pressure as reference. Some hindcasting tools to search for historical changes in coastal water quality have been developed (e.g. Kauppila et al., 2005; Clarke et al., 2006; Billen and Garnier, 2007). One possible difficulty with this method is the occurrence of temporal changes in the value of the considered indicators due for example to climatic oscillations (Kröncke et al., 1998; Grémare et al., 1998; Tunberg and Nelson, 1998; Hagberg and Tunberg, 2000; Dippner and Ikauniece, 2001; Labrune et al., 2007). In this case, the use of hindcasting can be highly misleading unless: (i) the causes of cyclical oscillations are well established; (ii) reference conditions are available for the whole oscillation cycle; and (iii) the positioning of both the reference condition and the considered year are strictly identical relative to the oscillation cycle. Moreover, global change in recent decades could also alter the former reference conditions, making impossible any comparison with dataset from 50, 100 or 200 years ago (see Kröncke and Reiss, 2010). However, there is no objective criterion for the selection of the baseline year to use and even if we wanted to return to some earlier status of few pressures caused by fewer people this is not an achievable goal in most regions (Hering et al., 2010). Hence, this method scored the lowest values in Table 1. In many cases the absence of historical records makes this approach inapplicable, and the influence of climate change makes this approach of only moderate utility.

Modelling of reference conditions: if no pristine sites exist, then reference conditions can be characterized based upon extrapola-

Table 1
Evaluation of target and reference conditions setting methods, regarding different issues. Note that scores are high: 3, moderate: 2, and low: 1, except in the case of data needs, which are opposite (the lowest data need the highest score). BPJ: best professional judgment (note that it can be used both in setting reference conditions and targets); WFD: Water Framework Directive; HD: Habitats Directive; CWA: Clean Water Act; OSPAR: Oslo-Paris Convention. Based partially upon the discussions within the OSPAR/Marine Strategy Framework Directive Workshop on approaches to determining Good Environmental Status for biodiversity, held in Utrecht (Netherlands), 23–24 November 2010.

	Reference conditions			BPJ	Targets		
	Pristine areas	Historical data	Modelling		Baseline set in the past	Current baseline	Directional/trends target
Legislation using/proposing it	WFD, CWA	WFD, CWA, OSPAR	WFD, CWA	WFD, CWA	OSPAR	HD	OSPAR
Data needs	Moderate (2)	High (1)	High (1)	Low (3)	Moderate (2)	Moderate (2)	Moderate (2)
Scientific robustness	High (3)	Moderate (2)	Moderate/High (2.5)	High (3)	High (3)	High (3)	High (3)
Confidence of the method	High (3)	Moderate (2)	Moderate/high (2.5)	High (3)	Moderate (2)	High (3)	Moderate (2)
Applicability	High (3)	Low (1)	High (3)	High (3)	Moderate (2)	Moderate (2)	Moderate (2)
Practicality of the method within available time scales	High (3)	Moderate (2)	High (3)	High (3)	Moderate (2)	Moderate (2)	Moderate (2)
Transparency and comprehensibility	High (3)	High (3)	Low (1)	Low (1)	High (3)	High (3)	High (3)
Total scores	17	11	13	16	14	15	14

tion of the biological attributes representative of the aquatic biota expected to be found in the ecoregion or type, through quantitative models, i.e. using habitat suitability models, multiple regression models, etc. (CIS, 2003a). However, there are numerous issues in developing and applying such models: (i) selection of a conceptual ecological paradigm identifying the minimal number of dependent and independent variables of the model; (ii) identification and incorporation of linear and non-linear relationships; (iii) generalized models that are based primarily upon conceptual relationships may produce predictions of little applicability to different ecoregions or at smaller spatial scales necessary for restoration decisions; and (iv) models that have robust predictions based upon the data used for their creation may be inappropriate for predicting changes in Drivers-Pressures-States. In dynamic and highly variable marine environments our modelling capabilities are insufficient for defining reference conditions (Hering et al., 2010). This difficulty is overemphasized when natural oscillation occurs. Hence, this method scored the second value (Table 1). To achieve model predictions of high robustness, a large data set that is spatially and temporally intensive is required—in many eco-regions such data sets do not exist. The applicability of model predictions between eco-regions is of unknown validity. Most importantly, the complexities of this approach tend to make transparency and comprehensibility for stakeholders and policy makers low.

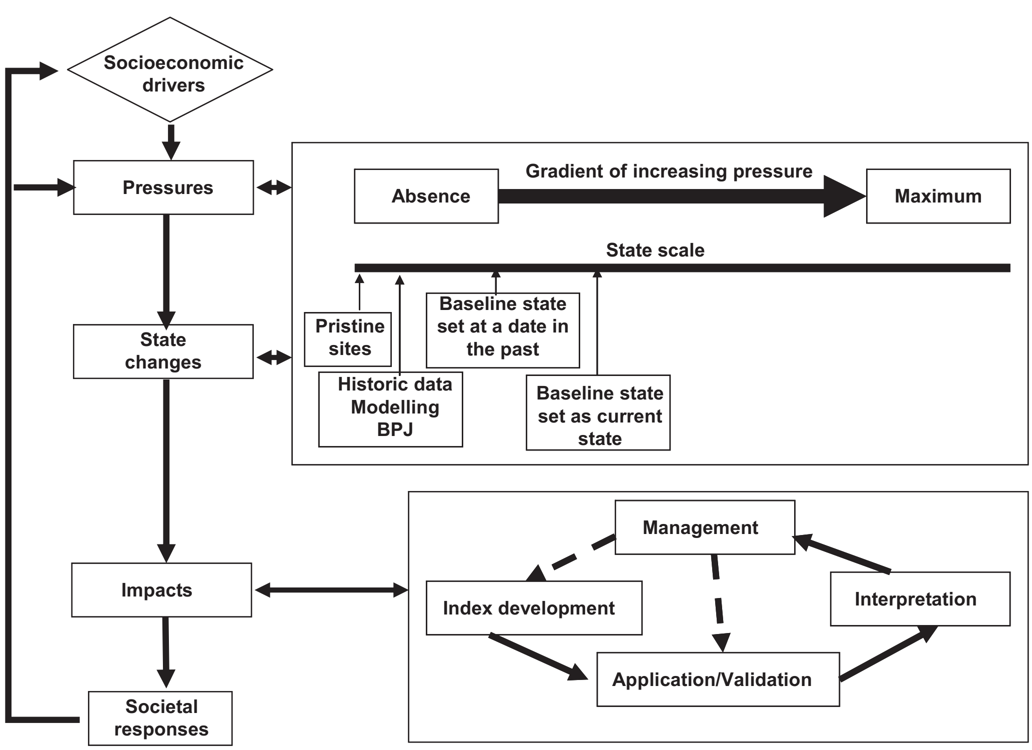

Fig. 1. The DPSIR approach showing relationships between Drivers-Pressure-State of change-Impact-Responses variables, in assessing environmental quality status in marine waters. Environmental status can be considered as a gradation from pristine conditions (high status in absence of human pressures) to an irrecoverable status (bad status, in a maximum human pressure). Assessment systems need to set reference conditions or baseline targets along the pressure (and subsequent state) gradient to assist in status assessment and for monitoring progress against time and actions. In this step the development and validation of impact assessment methods is needed.

Adapted from Borja and Dauer (2008) and Cochrane et al. (2010). BPJ—best professional judgment.

In addition, for setting targets there are also several alternative approaches (Fig. 1), which are described below.

Baseline set in the past: target set in relation to the first data point in a time series. However, this does not necessarily represent an unimpacted state, especially if the pressures in the past are not known. Another problem associated with this method is the absence of a target end-point, making difficult to know when the objectives are met in the process of restoration. Hence, this method scores in an intermediate position in Table 1, due to its moderate applicability, confidence or practicality.

Current baseline: target set in relation to the condition at the inception of a particular environmental policy—e.g. the European Habitats Directive approach where the state of the environment at 1994 was used as a baseline (i.e. the current state at the time of policy inception). We also scored this method in an intermediate position (Table 1) with moderate applicability, practicality, and reliability. The absence of an end-point target and sudden changes in the trend, without known pressures, is a problem in this method.

Directional/trends targets: target describes a desired trend in state in relation to the chosen baseline, i.e. a continuous improvement in state but where a final end point cannot be identified. We scored this method only slightly higher than the other target methods (Table 1). Even with moderate applicability or practicality, the method is robust and reliable, because: (i) it is easier to get good present data than past data, as in the previous method; and (ii) the method only requires relative assessments of ecological quality status, which makes it largely independent of the concept of reference conditions. Again, the absence of an end-point target is a problem in this method.

When none of the previous information is available, best professional judgment (BPJ) can be used in setting both targets and/or reference conditions. It has been demonstrated as very useful in assessing the status of ample geographical areas within USA and Europe (Teixeira et al., 2010), with a common set of criteria among different experts. This method scored the second highest values in Table 1. This is because, as demonstrated by Teixeira et al. (2010), the method can be robust and reliable, applicable to any geographical area and very practical and pragmatic. The method, it is not very transparent and comprehensible for stakeholders. However, the abovementioned methods, used in combination with expert judgment, in many cases have been used for defining and determining reference conditions and targets (e.g. Carletti and Heiskanen, 2009). Of course, the BPJ may be affected also by global change as well, i.e. the naturalists of the early 1900s would certainly had a different view of pristine condition than today's ones.

3. The importance of setting correct reference conditions: the case of M-AMBI

Different benthic indices have been developed worldwide to assess marine quality status (see Díaz et al., 2004; Pinto et al., 2009). One of the currently most used indices is the AZTI's Marine Biotic Index (AMBI) (Borja et al., 2000), which, in addition to species richness and Shannon diversity, has produced a multivariate AMBI (M-AMBI; Borja et al., 2004; Muxika et al., 2007). M-AMBI has allowed for an application within the WFD in different countries (Borja et al., 2009a), and has been tested under different human pressures (Table 2). Although this is only one of the indices used in assessing benthos, the use of reference conditions in this index particularly gives lessons for the future for other indices. Hence, the use of this method requires the setting of reference conditions (Muxika et al., 2007), specific for each type or habitat, which can represent a limitation when the number of habitats is too high (de Paz et al., 2008; Teixeira et al., 2008). Table 2 includes the known publications for this method, together with the method used to set the reference conditions and, then, assess the status of different coastal and transitional locations, subjected to different human pressures. A total of 29 references were found, including 12 in coastal waters, 14 in transitional waters and 3 in both systems (hence, 15 coastal and 17 transitional, with a total of 32 cases).

From the 15 coastal cases, 11 presented significant response of M-AMBI to human pressures, 1 was unclear, 1 was not significant and 2 were focusing in spatial and temporal variability (showing low variability in both cases) (Table 2). In turn, from the 17 transitional cases, 10 presented a significant response, 4 were unclear, 2 presented not significant response and 1 was not specified by the authors (Table 2). Hence, from the 32 cases, 21 showed significant responses, 3 not significant, 5 had unclear responses, 1 was not specified and 2 focused on temporal and spatial variability.

However, in some of the cases, the authors do not clearly explain the methods they used in setting reference conditions for those areas, producing some problems when explaining results. For example, in 2 out of the 5 unclear cases, the setting method was that given by default within the AMBI software (http://ambi.azti.es), i.e. the maximum values of richness and diversity, and minimum value of AMBI, within the dataset used in the analysis. In another 2 out of the 5 unclear cases, the authors did not determine the reference conditions for their locations, but used those derived by other authors. In 2 out of the 3 cases which the response was not significant or not specified, the reference condition method was not specified. In turn, from the 21 significant responses, 11 used a combination of methods to set the reference conditions, including undisturbed areas, historical data, modelling, best professional judgment, and maximum – minimum values; 2 used only undisturbed conditions; 5 used minimum – maximum values (or a variation of the method); and 3 did not explain the method (Table 2).

Minimum – maximum values from AMBI software should not be used automatically. This is only a default resource given by the software to facilitate calculation. However, when using this software it is necessary to set adequate reference conditions for each biotope (Muxika et al., 2007). Setting adequate reference conditions is extremely important when assessing the status of areas with human pressures, because the whole monitored area can be affected and none of the stations (and the structural variables measured there and included in the M-AMBI analysis) can be used as reference. Reference conditions must be assessed independently for each habitat. This is true both for M-AMBI and for the analysis of reference stations. Doing otherwise is a major source of confusion in determining ES. One simple reason for that, is that ES within a given habitat becomes dependent of the sampling (or not) of other habitats. Hence, the use of unimpacted areas to set the reference conditions is a better way to determine the quality status, as derived from data in Table 2.

In the other hand, it seems that there are more significant results in coastal (73.3% of the cases) than in transitional (58.8%) waters. This could be related to the estuarine quality paradox (Elliott and Quintino, 2007; Dauvin, 2007), in which it is difficult to distinguish between natural variability associated to transitional waters and human pressures, although this index has been particularly advantageous in areas with high natural variability (see Borja and Tunberg, 2011). Some authors state that these differences are due the primary development of the index for coastal waters instead for transitional waters; but this is not true, because both AMBI (Borja et al., 2000) and M-AMBI (Borja et al., 2004; Muxika et al., 2007) were developed at the same time using coastal and transitional datasets. However, it is interesting to note that in 3 out of the 7 cases in which the results are not clear or not significant in transitional waters, many different estuaries and lagoons have been studied jointly. Moreover, in most of these 7 cases the same reference conditions were used for the whole dataset, without taking

Table 2
Use of M-AMBI in assessing benthic quality status, in different locations and water types, under different human pressures. The response of the method in assessing the status is included (significant: when the response to the pressure is as expected; not significant: when there is not response; unclear: sometimes works and in other cases not), together with the method used in setting reference conditions. NS: not specified; U: undisturbed; H: historical data; M: modelling; BPJ: best professional judgment; MM: minimum (AMBI) and maximum (richness, diversity) values provided by AMBI software (http://ambi.azti.es) within the dataset used in the analysis. SP: Spain, IT: Italy, FR: France, US: United States, PT: Portugal, D: Germany, AL: Albania, GR: Greece, CY: Cyprus, RO: Romania; BU: Bulgaria.

Reference	Location	Water type	Pressure investigated	Setting method	Response
Muxika et al. (2007)	Basque Country (SP)	CW, TW	Multiple	H, M, BPJ	Significant
Forni and Occhipinti-Ambrogi (2007)	Adriatic Sea (IT)	CW	Eutrophication	MM	Significant
Pranovi et al. (2007)	Venice lagoon (IT)	TW	Eutrophication	NS	Not significant
Ruellet and Dauvin (2007)	Seine estuary (FR)	TW	Multiple	MM	Unclear
Borja et al. (2008b)	Chesapeake Bay (US)	TW	Multiple	15%>MM	Significant
Bigot et al. (2008)	Reunion Island (FR)	CW	Discharges	U, MM	Significant
Bouchet and Sauriau (2008)	Pertuis Charentais (FR)	CW	Oyster culture	U, MM	Significant
de Paz et al. (2008)	Eo estuary (SP)	TW	Oyster culture	Muxika et al. (2007)	Unclear
Ponti et al. (2008)	Adriatic & Black Sea (IT, AL, RO, BU)	TW	Several	NS	Significant
Puente and Diaz (2008)	Cantabria (SP)	TW	Multiple	Muxika et al. (2007)	Unclear
Simboura and Reizopoulou (2008)	Greece	CW, TW	Several	U, H, BPJ	Significant
Bakalem et al. (2009)	Algeria	CW	Discharges	MM	Significant
Borja et al. (2009b)	Basque Country (SP)	CW, TW	Multiple	H, M, BPJ	Significant
Callier et al. (2009)	Canada	TW	Mussel culture	U	Significant
Lavesque et al. (2009)	Arcachon Bay (FR)	TW	Several	U *Zostera*	Not significant
Prato et al. (2009)	Italy lagoons	TW	Discharges	H, MM	Significant
Simonini et al. (2009)	Adriatic Sea (IT)	CW	Eutrophication	H, BPJ	Significant
Tomassetti et al. (2009)	Tyrrhenian Sea (IT)	CW	Fish culture	U	Significant
Abbiati et al. (2010)	Adriatic lagoons (IT)	TW	Multiple	NS (MM?)	NS (significant?)
Costa-Dias et al. (2010)	Lima estuary (PT)	TW	Several	NS	Significant
Fitch and Crowe (2010)	Ireland	CW	Eutrophication	NS	Significant
Kröncke and Reiss (2010)	North Sea (D)	CW	Natural	Muxika et al. (2007)	Low variability
Munari et al. (2010)	Karavasta lagoon (AL)	TW	Several	MM	Significant
Munari and Mistri (2010)	Adriatic lagoons (IT)	TW	Several	MM	Unclear
Simboura and Argyrou (2010)	Eastern Mediterranean (GR, CY)	CW	Several	U, H, BPJ	Unclear
Tataranni and Lardicci (2010)	Tuscany (IT)	CW	Several	NS	Low variability
Borja and Tunberg (2011)	Florida estuary (US)	TW	Eutrophication	99th percentile	Significant
Forchino et al. (2011)	Sardinia (IT)	CW	Fish culture	U, MM	Significant
Sampaio et al. (2011)	Lisboa (PT)	CW	Discharges	MM	Not significant

into account differences in estuary and lagoon types or the salinity gradient, which produces different reference conditions, depending on the habitats present (Muxika et al., 2007; Borja et al., 2009b). Hence, adequate reference conditions for the natural gradient associated to transitional waters are needed for a correct ecosystem quality assessment, as stated by different authors (e.g. Muxika et al., 2007; Teixeira et al., 2008).

Some authors have claimed about the lack of resolution of these quality indices under study and their inability to show the human pressure, proposing alternative integrative approaches (Sampaio et al., 2011). However, some single and multimetric indices, used in benthic assessment under the WFD (including M-AMBI), have been tested successfully in different geographical areas and under different pressure gradients (Borja et al., 2011a).

The assessment of the importance of the reference conditions, carried out during the present study, is rather indirect because: (i) it is based on the assumption that M-AMBI is indeed able to detect always anthropogenic impacts; and (ii) not considering differences in the magnitude of such impact between studies. Further studies should focus on the direct assessment of the importance of reference conditions by extensively sampling an habitat for which historical data are available. Such studies could allow for the comparison of the magnitude of the effect of different indicators and different methods of setting reference conditions on ES assessment. It should not only consider reference conditions but also the different types of scales used to infer ES.

Finally, the absence of clear quantitative or semiquantitative human pressure data, when determining reference conditions or targets is also a major problem (Borja et al., 2011a). One possibility could be to assess those pressures spatially within a same community. The problem of taking into account all possible stressors is a major one. Additive effects of several stressors at low concentrations (i.e. 'cocktail effect') is also a major one not currently taken into account by multivariate approaches designed to relate environmental an biological parameters. Most areas have multiple pressures and so the change against reference conditions has to accommodate cumulative, synergistic and antagonistic effects (Crain et al., 2008).

4. Conclusions

Setting reference conditions for assessing benthic ecological status is a major issue in aquatic management. Sometimes this setting is not correctly addressed when using the method developed to assess the benthic health, producing wrong interpretation of the results. From the analysis undertaken it seems that using a combination of methods in setting reference conditions, is more adequate in obtaining final quality assessments related to the pressures than one method alone.

Acknowledgements

This paper is in part a result of the project WISER (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (contract No. 226273), www.wiser.eu. This is contribution number 542 from AZTI-Tecnalia (Marine Research Division).

References

Abbiati, M., Mistri, M., Bartoli, M., Ceccherelli, V.U., Colangelo, M.A., Giordani, G., Munari, C., Nizzoli, D., Ponti, M., Rossi, R., Ferrari, C.R., Viaroli, P., 2010. Trade-off between conservation and exploitation of the transitional water ecosystems of the northern Adriatic Sea. Chemistry and Ecology 26, 105–119.

Atkins, J.P, Burdon, D., Elliott, M., Gregory, A.J., 2011. Management of the marine environment: integrating ecosystem services and societal benefits with the DPSIR framework in a systems approach. Marine Pollution Bulletin 62, 215–226.

Aubry, A., Elliott, M., 2006. The use of environmental integrative indicators to assess seabed disturbance in estuaries and coasts: application to the Humber Estuary, UK. Marine Pollution Bulletin 53, 175–185.

Backer, H., Leppänen, J.-M., Brusendorff, A.C., Forsius, K., Stankiewicz, M., Mehtonen, J., Pyhälä, M., Laamanen, M., Paulomäki, H., Vlasov, N., Haaranen, T., 2010. HELCOM Baltic Sea Action Plan—a regional programme of measures for the marine environment based on the Ecosystem Approach. Marine Pollution Bulletin 60, 642–649.

Bakalem, A., Ruellet, T., Dauvin, J.C., 2009. Benthic indices and ecological quality of shallow Algeria fine sand community. Ecological Indicators 9, 395–408.

Bigot, L., Grémare, A., Amouroux, J.-M., Frouin, P., Maire, O., Gaertner, J.C., 2008. Assessment of the ecological quality status of soft-bottoms in Reunion Island (tropical Southwest Indian Ocean) using AZTI marine biotic indices. Marine Pollution Bulletin 56, 704–722.

Billen, G., Garnier, J., 2007. River basin nutrient delivery to the coastal sea: assessing its potential to sustain new production of non-siliceous algae. Marine Chemistry 106, 148–160.

Borja, A., Dauer, D.M., 2008. Assessing the environmental quality status in estuarine and coastal systems: comparing methodologies and indices. Ecological Indicators 8, 331–337.

Borja, A., Tunberg, B.G., 2011. Assessing benthic health in stressed subtropical estuaries, eastern Florida, USA using AMBI and M-AMBI. Ecological Indicators 11, 295–303.

Borja, A., Franco, J., Pérez, V., 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Marine Pollution Bulletin 40, 1100–1114.

Borja, A., Franco, J., Valencia, V., Bald, J., Muxika, I., Belzunce, M.J., Solaun, O., 2004. Implementation of the European Water Framework Directive from the Basque Country (northern Spain): a methodological approach. Marine Pollution Bulletin 48 (3/4), 209–218.

Borja, A., Galparsoro, I., Solaun, O., Muxika, I., Tello, E.M., Uriarte, A., Valencia, V., 2006. The European Water Framework Directive and the DPSIR, a methodological approach to assess the risk of failing to achieve good ecological status. Estuarine, Coastal and Shelf Science 66, 84–96.

Borja, A., Bricker, S.B., Dauer, D.M., Demetriades, N.T., Ferreira, J.G., Forbes, A.T., Hutchings, P., Jia, X., Kenchington, R., Marques, J.C., Zhu, C., 2008a. Overview of integrative tools and methods in assessing ecological integrity in estuarine and coastal systems worldwide. Marine Pollution Bulletin 56, 1519–1537.

Borja, A., Dauer, D.M., Diaz, R., Llansó, R.J., Muxika, I., Rodriguez, J.G., Schaffner, L., 2008b. Assessing estuarine benthic quality conditions in Chesapeake Bay: a comparison of three indices. Ecological Indicators 8, 395–403.

Borja, A., Miles, A., Occhipinti-Ambrogi, A., Berg, T., 2009a. Current status of macroinvertebrate methods used for assessing the quality of European marine waters: implementing the Water Framework Directive. Hydrobiologia 633 (1), 181–196.

Borja, A., Muxika, I., Rodríguez, J.G., 2009b. Paradigmatic responses of marine benthic communities to different anthropogenic pressures, using M-AMBI, within the European Water Framework Directive. Marine Ecology 30, 214–227.

Borja, Á., Elliott, M., Carstensen, J., Heiskanen, A.-S., van de Bund, W., 2010a. Marine management—towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives. Marine Pollution Bulletin 60, 2175–2186.

Borja, Á., Dauer, D., Elliott, M., Simenstad, C., 2010b. Medium- and long-term recovery of estuarine and coastal ecosystems: patterns, rates and restoration effectiveness. Estuaries and Coasts 33, 1249–1260.

Borja, A., Barbone, E., Basset, A., Borgersen, G., Brkljacic, M., Elliott, M., Garmendia, J.M., Marques, J.C., Mazik, K., Muxika, I., Neto, J.M., Norling, K., Rodríguez, J.G., Rosati, I., Rygg, B., Teixeira, H., Trayanova, A., 2011a. Response of single benthic metrics and multi-metric methods to anthropogenic pressure gradients, in five distinct European coastal and transitional ecosystems. Marine Pollution Bulletin 62, 499–513.

Borja, Á., Galparsoro, I., Irigoien, X., Iriondo, A., Menchaca, I., Muxika, I., Pascual, M., Quincoces, I., Revilla, M., Germán Rodríguez, J., Santurtún, M., Solaun, O., Uriarte, A., Valencia, V., Zorita, I., 2011b. 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.

Bouchet, V.M.P., Sauriau, P.-G., 2008. Influence of oyster culture practices and environmental conditions on the ecological status of intertidal mudflats in the Pertuis Charentais (SW France): a multi-index approach. Marine Pollution Bulletin 56, 1898–1912.

Callier, M.D., Richard, M., McKindsey, C.W., Archambault, P., Desrosiers, G., 2009. Responses of benthic macrofauna and biogeochemical fluxes to various levels of mussel biodeposition: an in situ "benthocosm" experiment. Marine Pollution Bulletin 58, 1544–1553.

Carletti, A., Heiskanen, A.-S., 2009. Water Framework Directive Intercalibration Technical Report. Part 3: Coastal and Transitional Waters. JRC Scientific and Technical Reports, EUR 23838 EN/3, 240 pp. doi:10.2788/19561.

CIS, 2003a. Guidance Document No. 5 'Transitional and Coastal Waters—Typology, Reference Conditions, and Classification Systems'. Common Implementation Strategy of the Water Framework Directive, Available at: http://forum.europa.eu.int/Public/irc/env/wfd/library.

CIS, 2003b. Guidance Document No. 10. River and Lakes—Typology, Reference Conditions and Classification Systems. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). European Commission, Brussels.

Clarke, A.L., Weckström, K., Conley, D.J., Anderson, N.J., Adser, F., Andrén, E., De Jonge, V.N., Ellegaard, M., Juggins, S., Kauppila, P., Korhola, A., Reuss, N., Telford, R.J., Vaalgamaa, S., 2006. Long-term trends in eutrophication and nutrients in the coastal zone. Limnology and Oceanography 51, 385–397.

Cochrane, S.K.J., Connor, D.W., Nilsson, P., Mitchell, I., Reker, J., Franco, J., Valavanis, V., Moncheva, S., Ekebom, J., Nygaard, K., Serrao Santos, R., Naberhaus, I., Packeiser, T., van de Bund, W., Cardoso, A.C., 2010. Marine Strategy Framework Directive—Task Group 1 Report Biological Diversity. EUR 24337 EN-Joint Research Centre. Office for Official Publications of the European Communities, Luxembourg, 110 pp.

Costa-Dias, S., Sousa, R., Antunes, C., 2010. Ecological quality assessment of the lower Lima Estuary. Marine Pollution Bulletin 61, 234–239.

Crain, C.M., Kroeker, K., Halpern, B.S., 2008. Interactive and cumulative effects of multiple human stressors in marine systems. Ecology Letters 11, 1304–1315.

Dauvin, J.C., 2007. Paradox of estuarine quality: benthic indicators and indices, consensus or debate for the future. Marine Pollution Bulletin 55, 271–281.

de Paz, L., Patrício, J., Marques, J.C., Borja, A., Laborda, A.J., 2008. Ecological status assessment in the lower Eo estuary (Spain). The challenge of habitat heterogeneity integration: a benthic perspective. Marine Pollution Bulletin 56, 1275–1283.

Díaz, R.J., Solan, M., Valente, R.M., 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. Journal of Environmental Management 73, 165–181.

Dippner, J.W., Ikauniece, A., 2001. Long-term zoobenthos variability in the Gulf of Riga in relation to climate variability. Journal of Marine Systems 30, 155–164.

Elliott, M., 2002. The role of the DPSIR approach and conceptual models in marine environmental management: an example for offshore wind power. Marine Pollution Bulletin 44, iii–iii10.

Elliott, M., 2011. Marine science and management means tackling exogenic unmanaged pressures and endogenic managed pressures—a numbered guide. Marine Pollution Bulletin 62, 651–655.

Elliott, M., Quintino, V., 2007. The Estuarine Quality Paradox. Environmental homeostasis and the difficulty of detecting anthropogenic stress in naturally stressed areas. Marine Pollution Bulletin 54, 640–645.

Fitch, J.E., Crowe, T.P., 2010. Effective methods for assessing ecological quality in intertidal soft-sediment habitats. Marine Pollution Bulletin 60, 1726–1733.

Forchino, A., Borja, A., Brambilla, F., Rodríguez, J.G., Muxika, I., Terova, G., Saroglia, M., 2011. Evaluation of the influence of off-shore cage aquaculture on the benthic ecosystem in the Alghero bay (Sardinia, Italy), using AMBI and M-AMBI. Ecological Indicators 11, 1112–1122.

Forni, G., Occhipinti-Ambrogi, A., 2007. Daphne: a new multimetric benthic index for the quality assessment of marine coastal environment in the Northern Adriatic Sea. Chemistry and Ecology 23, 427–442.

Gibson, G.R, Bowman, M.L., Gerritsen, J., Snyder, B.D., 2000. Estuarine and Coastal Marine Waters: Bioassessment and Biocriteria Technical Guidance. EPA 822-B-00-024. US Environmental Protection Agency, Office of Water, Washington, DC.

Grémare, A., Amouroux, J.M., Vétion, G., 1998. Long-term comparison of macrobenthos within the soft bottoms of the Bay of Banyuls-sur-mer (northwestern Mediterranean Sea). Journal of Sea Research 40, 281–302.

Grémare, A., Labrune, C., Vanden Berghe, E., Amouroux, J.M., Bachelet, G., Zettler, M.L., Vanaverbeke, J., Fleischer, D., Bigot, L., Maire, O., Deflandre, B., Craeymeersch, J., Degraer, S., Dounas, C., Duineveld, G., Heip, C., Herrmann, M., Hummel, H., Karakassis, I., Kedra, M., Kendall, M., Kingston, P., Laudien, J., Occhipinti-Ambrogi, A., Rachor, E., Sarda, R., Speybroeck, J., Van Hoey, G., Vincx, M., Whomersley, P., Willems, W., Wlodarska-Kowalczuk, M., Zenetos, A., 2009. Comparison of the performances of two biotic indices based on the MacroBen database. Marine Ecology Progress Series 382, 297–311.

Hagberg, J., Tunberg, B.G., 2000. Studies on the covariation between physical factors and the long-term variation of the marine soft bottom macrofauna in Western Sweden. Estuarine Coastal and Shelf Science 50, 373–385.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R., Watson, R., 2008. A global map of human impact on marine ecosystems. Science 319, 948–952.

Heiskanen, A.S., van de Bund, W., Cardoso, A.C., Nõges, P., 2004. Towards good ecological status of surface waters in Europe—interpretation and harmonisation of the concept. Water Science and Technology 49, 169–177.

HELCOM, 2010. Ecosystem health of the Baltic Sea 2003–2007: HELCOM initial holistic assessment. In: Baltic Sea Environmental Proceedings, vol. 122 , 68 pp.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C.K., Heiskanen, A.S., Johnson, R.K., Moe, J., Pont, D., Solheim, A.L., van de Bund, W., 2010. The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. Science of the Total Environment 408, 4007–4019.

Kauppila, P., Weckström, K., Vaalgamaa, S., Korhola, A., Pitkänen, H., Reuss, N., Drew, S., 2005. Tracing pollution and recovery using sediments in an urban estuary, northern Baltic Sea: are we far from ecological reference conditions? Marine Ecology Progress Series 290, 35–53.

Kröncke, I., Reiss, H., 2010. Influence of macrofauna long-term natural variability on benthic indices used in ecological quality assessment. Marine Pollution Bulletin 60, 58–68.

Kröncke, I., Dippner, J.W., Heyen, H., Zeiss, B., 1998. Long-term changes in macrofaunal communities of Nordeney (East Frisia, Germany) in relation to climate variability. Marine Ecology Progress Series 167, 25–36.

Labrune, C., Amouroux, J.M., Sardá, 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.

Labrune, C., Grémare, A., Guizien, K., Amouroux, J.M., 2007. Long-term comparison of soft bottom macrobenthos in the Bay of Banyuls-sur-Mer (north-western Mediterranean Sea): a reappraisal. Journal of Sea Research 58, 125–143.
Latimer, J.S., Boothman, W.S., Pesch, C.E., Chmura, G.L., Pospelova, V., Jayaraman, S., 2003. Environmental stress and recovery: the geochemical record of human disturbance in New Bedford Harbor and Apponagansett Bay, Massachusetts (USA). The Science of the Total Environment 313, 153–176.
Lavesque, N., Blanchet, H., de Montaudouin, X., 2009. Development of a multimetric approach to assess perturbation of benthic macrofauna in *Zostera noltii* beds. Journal of Experimental Marine Biology and Ecology 368, 101–112.
Mee, L.D., Jefferson, R.L., Laffoley, D.d'A., Elliott, M., 2008. How good is good? Human values and Europe's proposed Marine Strategy Directive. Marine Pollution Bulletin 56, 187–204.
Munari, C., Mistri, M., 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.
Munari, C., Tessari, U., Rossi, R., Mistri, M., 2010. The ecological status of Karavasta Lagoon (Albania): closing the stable door before the horse has bolted? Marine Environmental Research 69, 10–17.
Muxika, I., Borja, Á., Bald, J., 2007. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. Marine Pollution Bulletin 55, 16–29.
Nielsen, K., Somod, B., Ellegaard, C., Krause-Jensen, D., 2003. Assessing reference conditions according to the European Water Framework Directive using modelling and analysis of historical data: an example from Randers Fjord, Denmark. Ambio 32, 287–294.
Pinto, R., Patricio, J., Baeta, A., Fath, B.D., Neto, J.M., Marques, J.C., 2009. Review and evaluation of estuarine biotic indices to assess benthic condition. Ecological Indicators 9, 1–25.
Ponti, M., Pinna, M., Basset, A., Moncheva, S., Trayanova, A., Georgescu, L.P., Beqiraj, S., Orfanidis, S., Abbiati, M., 2008. Quality assessment of Mediterranean and Black Sea transitional waters: comparing responses of benthic biotic indices. Aquatic Conservation: Marine and Freshwater Ecosystems 18, S62–S75.
Pranovi, F., Da Ponte, F., Torricelli, P., 2007. Application of biotic indices and relationship with structural and functional features of macrobenthic community in the lagoon of Venice: an example over a long time series of data. Marine Pollution Bulletin 54, 1607–1618.
Prato, S., Morgana, J.G., La Valle, P., Finoia, M.G., Lattanzi, L., Nicoletti, L., Ardizzone, G.D., Izzo, G., 2009. Application of biotic and taxonomic distinctness indices in assessing the Ecological Quality Status of two coastal lakes: Caprolace and Fogliano lakes (Central Italy). Ecological Indicators 9, 568–583.
Puente, A., Diaz, R.J., 2008. Is it possible to assess the ecological status of highly stressed natural estuarine environments using macroinvertebrates indices? Marine Pollution Bulletin 56, 1880–1889.
Ruellet, T., Dauvin, J.-C., 2007. Benthic indicators: analysis of the threshold values of ecological quality classifications for transitional waters. Marine Pollution Bulletin 54, 1707–1714.
Sampaio, L., Rodrigues, A.M., Quintino, V., 2011. Can biotic indices detect mild organic enrichment of the seafloor? Ecological Indicators 11, 1235–1244.
Simboura, N., Reizopoulou, S., 2008. An intercalibration of classification metrics of benthic macroinvertebrates in coastal and transitional ecosystems of the Eastern Mediterranean ecoregion (Greece). Marine Pollution Bulletin 56, 116–126.
Simboura, N., Argyrou, M., 2010. An insight into the performance of benthic classification indices tested in Eastern Mediterranean coastal waters. Marine Pollution Bulletin 60, 701–709.
Simonini, R., Grandi, V., Massamba-N'Siala, G., Lotti, M., Montanari, G., Prevedelli, D., 2009. Assessing the ecological status of the North-western Adriatic Sea within the European Water Framework Directive: a comparison of Bentix, AMBI and M-AMBI methods. Marine Ecology 30, 241–254.
Stelzenmüller, V., Lee, J., South, A., Rogers, S.I., 2010. Quantifying cumulative impacts of human pressures on the marine environment: a geospatial modelling framework. Marine Ecology Progress Series 398, 19–32.
Tataranni, M., Lardicci, C., 2010. Performance of some biotic indices in the real variable world: a case study at different spatial scales in North-Western Mediterranean Sea. Environmental Pollution 158, 26–34.
Teixeira, H., Salas, F., Neto, J.M., Patrício, J., Pinto, R., Veríssimo, H., García-Charton, J.A., Marcos, C., Pérez-Ruzafa, A., Marques, J.C., 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., Borja, Á., Weisberg, S.B., Ananda Ranasinghe, J., Cadien, D.B., Dauer, D.M., Dauvin, J.-C., Degraer, S., Diaz, R.J., Grémare, A., Karakassis, I., Llansó, R.J., Lovell, L.L., Marques, J.C., Montagne, D.E., Occhipinti-Ambrogi, A., Rosenberg, R., Sardá, R., Schaffner, L.C., Velarde, R.G., 2010. Assessing coastal benthic macrofauna community condition using best professional judgement—developing consensus across North America and Europe. Marine Pollution Bulletin 60, 589–600.
Tomassetti, P., Persia, E., Mercatali, I., Vani, D., Marussso, V., Porrello, S., 2009. Effects of mariculture on macrobenthic assemblages in a western Mediterranean site. Marine Pollution Bulletin 58, 533–541.
Topcu, D., Brockmann, U., Claussen, U., 2009. Relationship between eutrophication reference conditions and boundary settings considering OSPAR recommendations and the Water Framework Directive—examples from the German Bight. Hydrobiologia 629, 91–106.
Tunberg, B.G., Nelson, W.G., 1998. Do climatic oscillations influence cyclical patterns of soft bottom macrobenthic communities on the Swedish west coast? Marine Ecology Progress Series 170, 85–94.
USEPA, 2006. Indicator Development for Estuaries. United States Environmental Protection Agency, 138 pp.