Available online at www.sciencedirect.com
SCIENCE DIRECT®
Marine Pollution Bulletin 49 (2004) 728–739
MARINE POLLUTION BULLETIN
www.elsevier.com/locate/marpolbul

Marine quality assessment by use of benthic species-abundance distributions: a proposed new protocol within the European Union Water Framework Directive

Rutger Rosenberg [a,*], Mats Blomqvist [b], Hans C Nilsson [c], Hans Cederwall [d], Anna Dimming [a]

[a] Department of Marine Ecology, Göteborg University, Kristineberg Marine Research Station, 450 34 Fiskebäckskil, Sweden
[b] Hafok AB, 179 61 Stenhamra, Sweden
[c] Marine Monitoring AB at Kristineberg, 450 34 Fiskebäckskil, Sweden
[d] Department of Systems Ecology, Stockholm University, 106 91 Stockholm, Sweden

Abstract

The aim of this study is to develop a new method for classification of marine benthic quality according to the European Union Water Framework Directive. Tolerance values to environmental disturbance were determined in an objective analysis for benthic species along the Swedish west coast by using 4676 samples from 257 stations. Based on a combination of the species tolerance values, abundance and diversity, a benthic quality index (BQI) was calculated for the assessment of environmental status at a particular station. The qualification of BQI was evaluated in relation to known spatial and temporal gradients of disturbance.

Keywords: Diversity index; Pollution; *Amphiura*; *Capitella*; Skagerrak; Kattegat

1. Introduction

Human impact on living resources has escalated over the last century and threatened the balance of many parts of the ecosystem. Awareness of the changed ecological conditions has fostered a need to assess the consequences and to suggest measures to reverse this trend. In the sea, coastal urban areas in particular are subject to negative ecological changes frequently associated with eutrophication, oxygen deficiency, contaminants and over-fishing.

Quality assessment of ecological changes in the sea can be provided most effectively by studying the sedimentary habitat and the benthic fauna, as most of the ecological impact and pollution load sooner or later will end up on the seabed. A marine benthic community in a fairly stable environment undergoes only minor qualitative and quantitative changes over time. Through evolution, benthic species have adjusted to cope with predicted environmental variations and interspecific competition. A significant disturbance will, however, introduce changes in the species' composition, their abundance and biomass. Such successional changes in benthic community structure are often predictable and, with increased perturbation the diversity, abundance and biomass will show a general decline (Pearson and Rosenberg, 1978).

1.1. Classification of various degrees of disturbance

Classification of aquatic systems into different degrees of pollution was first developed for freshwater in the saprobic system (Kolwitz and Marsson, 1909). The pioneer for classifying benthic marine systems was Reish (1955), who mainly used the distribution pattern of pollution-tolerant polychaetes, particularly *Capitella capitata*, to assess the spatial impact of pollution in California. Similar studies of pollution assessment were done in Finnish and Swedish waters by Leppäkoski (1975) and Rosenberg et al. (1975), and in the Mediterranean by Bellan (1985, and references therein). The

* Corresponding author. Tel.: +46-2531-8500; fax: +46-5231-8502.
E-mail address: r.rosenberg@kmf.gu.se (R. Rosenberg).

0025-326X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.marpolbul.2004.05.013

classification used by Leppäkoski (1975) was: (1) very polluted bottoms, (2) polluted bottoms, (3) semi-polluted bottoms, (4) semi-healthy bottoms, and (5) healthy bottoms. Based on these and other studies, relations between magnitude of disturbance and temporal or spatial changes in the benthic faunal composition were summarised and formulated in the Pearson and Rosenberg (1978) model. The benthic faunal successional changes in this model were depicted as a response to organic enrichment and oxygen deficiency, but was later also shown to apply to physical disturbance (Boesch and Rosenberg, 1981; Rhoads and Germano, 1986) and to organic enrichment in association with contaminants (Swartz et al., 1985). The model seems to have universal application for most disturbed, sub-littoral, soft-bottom habitats (e.g. Heip, 1995). Based on other studies and own published scientific results, Pearson and Rosenberg (1978) listed 31 indicator species or taxa diagnostic of particular sets of ecological conditions, notably the ubiquitous indicators of severe disturbance: *C. capitata* and *Scolelepis* (*Malacoceros*) *fuliginosa*.

1.2. Tolerant and sensitive species

Pearson and Rosenberg (1978) argued that unidirectional stress caused by a particular environmental disturbance will, as it increases in intensity, result first in adaptation by an individual within its abilities to respond, then it will be replaced by another better adapted individual able to respond to that particular stress. Beyond this level the species will be replaced by a group of species better adapted to the new conditions. Changes of faunal composition along a stress gradient may thus be viewed as a continuum interrupted by steps occurring at the points where the level of adaptability demanded exceeds the capacity of that level of organization. Based on these arguments it is suggested that tolerance should be analysed at the species level or highest possible taxonomic level as intraspecific adaptations cannot be accounted for in this kind of analysis. The use of lower organisation levels in taxonomy, as have been suggested for multivariate analysis of disturbance, e.g. Warwick (1988), is not recommended as different species within the same genera may show great discrepancies in tolerance.

Several scientists have tried different methods to classify the sensitivity or tolerance of benthic organisms to various degrees of disturbance. Gray and co-authors (Gray and Mirza, 1979; Gray, 1981; Pearson et al., 1983) compared the rank of abundance among species to obtain distinctions between patterns in polluted versus unpolluted conditions in north European waters. Warwick (1986) proposed that species ranked in order according to their abundance and biomass may be useful for determining if an area is unpolluted, moderately polluted or grossly polluted. Grall and Glémarec (1997) subjectively categorised benthic species into five groups of sensitivity for waters in French Brittany, which were used for benthic habitat quality assessment. Weisberg et al. (1997) used a multimetric approach testing 17 variables to objectively identify pollution-indicative and pollution-sensitive species in the Chesapeake Bay (see Table 1).

1.3. The water directive

The European Union Water Framework Directive (WFD) states, among other things, that the quality of all European coastal waters should be analysed regularly in the near future. One method to be used for this assessment is the species composition and abundance of benthic macrofauna. In addition, the concept of sensitive and tolerant species could complement such an evaluation. Based on the WFD, the European coastal habitats will be divided into different typologies based on, e.g. salinity, sediment characteristics, depth, and

Table 1
Summary of some published work dealing with sensitive and tolerant marine benthic species and benthic quality assessment based on benthic community data and sediment profile image (SPI)* analysis

Authors	Sensitive/tolerant species	Habitat assessment	Remarks
Reish (1955)	Own data	5 zones identified	Subjective
Leppäkoski (1975)	No	5 zones identified	Subjective
Peapson and Rosenberg (1978)	Literature data	4 successional stages	Subjective
Rhoads and Germano (1986)	No	4 successional stages	Subjective/objective
Gray and Pearson (1982)	Log-normal distributions	No	Objective
Grall and Glémarec (1997)	Literature, own experience	5 ecological groups	Subjective
Borja et al. (2003)	Literature, own experience	7 classes of Benthic Index	Subjective
Warwick (1986)	No	Based on abundance-biomass	Objective
Simboura and Zenetos (2002)	Literature, own experience	5 quality status groups	Subjective
Weisberg et al. (1997)	Probably literature data	17 variables used	Subjective/objective
Rygg (2002)	Diversity index	Indicator species index	Objective
This study	Diversity index	Faunal quality assessment	Objective
Rhoads and Germano (1986)*	No	Organism-sediment-index	Subjective/objective
Nilsson & Rosenberg (1997)*	No	Benthic habitat quality	Objective

water residence times. Recently, Borja et al. (2000), Simboura and Zenetos (2002) and Borja et al. (2003) listed >2000 indicator species in relation to their supposed tolerance to disturbance to be used within the WFD. The species lists are extensions of the work by Grall and Glémarec (1997) and based on literature data and own experiences; thus, the species are not assigned to different categories objectively. Based on these classifications of species, the authors calculated biotic indices that were used for environmental quality assessment (Table 1).

1.4. Diversity indices

Different types of diversity indices have been widely used in ecology for the assessment of environmental quality; a high index value will indicate healthy conditions and a low index value bad conditions. One of the most widely used diversity indices is the Shannon–Wiener formula based on information theory (Pielou, 1969). However, the qualification of this index for environmental quality assessment has been repeatedly criticised (e.g. Gray, 1979). Another frequently used diversity index is Sanders' rarefaction technique (Sanders, 1968). In this index, the number of species is calculated in relation to a certain number of individuals, e.g. from a sample 100 individuals may be collected at random and among these individuals the number of species can be identified and counted. This will give the estimated number of species among 100 individuals (ES100). Sanders' method of calculation over-estimated the number of species and, therefore, to compensate for this, Hurlbert (1971) later modified the formula for this calculation. Rygg (2002) used this method to identify sensitive and tolerant benthic indicator species in Norwegian waters. The assumption was that sensitive species will only occur in samples with high diversity, and tolerant species will be found predominantly in samples with low diversity. Rygg calculated ES100 and selected the average of the five lowest values (ES100 min_5) at the studied stations to obtain a sensitivity value for each species.

1.5. Present study

In the present study, the rare faction technique according to Hurlbert's (1971) formula was used to categorise benthic species into different degrees of sensitivity to disturbance according to the WFD. Benthic fauna from predominantly silt–clay sediments in 4676 grab samples from the Swedish coasts of Skagerrak, Kattegat and Öresund, i.e. between the Norwegian border in the north and Copenhagen-Malmö in the south, were used for this categorisation (Fig. 1). The study focuses on the coastal areas out to one nautical mile outside the outer (westerly) islands as stated in the WFD, but some stations outside this border were also included to obtain a better coverage. Samples were obtained from areas with various degrees of disturbance, i.e. from severely stressed (organically enriched and oxygen stressed) areas to areas considered more or less undisturbed. Moreover, data were used both from benthic communities in decline because of increasing disturbance, and from communities at different successional phases of recovery.

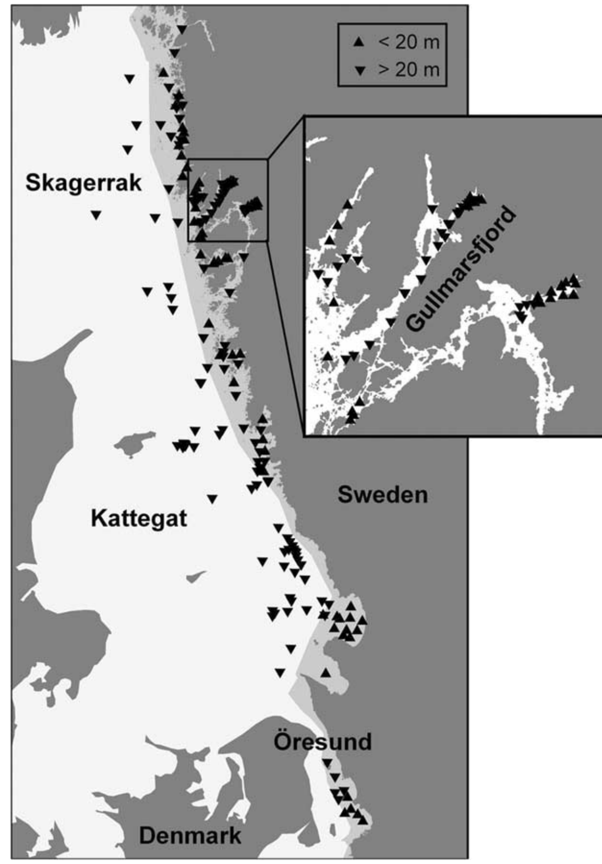

Fig. 1. Map of the benthic stations used in this analysis separated into depths >20 m and ⩽20 m. Information from one particular station could be used several times. The coastal area defined in the Water Framework Directive is shaded.

In the present study, the Hurlbert (1971) diversity index was calculated to classify the benthic species according to tolerance and sensitivity of disturbance. The index values were then used in combination with the species abundance distribution pattern along a gradient of disturbance, and the total number of species at a particular station, to calculate a new benthic quality index (BQI) for that site (see Section 2). The BQI is used for the environmental assessment of an investigated area.

2. Material and methods

Information of the occurrence of species and their abundance from 257 stations (Fig. 1), sampled at 1114 occasions and encompassing 4676 grab samples has been used for analysis. Information from one particular station could be used repeatedly at different times of sampling. All samples were taken by a 0.1 m^2 a Smith–McIntyre grab, the material was sieved through 1 mm meshes and in most cases sorted at six times magnification. Replicate samples from one station or occasion were combined and averaged for abundance and species number. The samples originate from regional and national monitoring data as well as data from research projects from 6 to 300 m depth covering the period 1969–2002. The longitudinal distance over which samples were obtained was ~400 km. Trained experts made identification to species or higher taxa. Total number of individuals analysed was 1,549,479 and the number of taxa was 1234. Use of names of species and genera may change between taxonomists and over time. Synonyms have been checked and nomenclature used by ICES has been applied.

2.1. Diversity index

In the present study, calculation of the expected number of species (ES) was made among 50 individuals according to Hurlbert's (1971) formula, which is used in the computer software PRIMER (Clark and Warwick, 1994):

$$\mathrm{ES50} = \sum_{i=1}^{s} \frac{(N - N_i)!(N - 50)!}{(N - N_i - 50)!N!}$$

where N is the total number of individuals in a sample and "i" is the number of the "ith" species. The validation of the index is based on the individuals of each species being randomly distributed, which is not always the case. In order to exclude species occurring in a few samples only, the number of sample occasions where a species must be recorded was limited to $\geqslant 20$. We use ES50 instead of ES100 to include samples with abundances between 50 and 100 in the analysis, which could be useful in disturbed areas with abundances in this interval. A high correlation ($r^2 = 0.957$, $n = 382$) was found between ES50 and ES100. Thus, samples with <50 individuals were not included in the analysis.

2.2. Tolerant and sensitive species

Tolerant species are by definition predominantly found in disturbed environments. That means that they mainly occur at stations with low ES50. In contrast, sensitive species occur in areas with no or minor disturbance and would then be associated with high ES50. In an abundance frequency distribution of a particular species in relation to ES50 values at the stations where it has been recorded, as in Fig. 2, the most tolerant individuals of a species are likely to be associated with the lowest ES50 values. We selected that 5 % of the population will be associated to this category, and define this value as the species tolerance value: $\mathrm{ES50}_{0.05}$. The rest of the population may, for various reasons, have greater ES50 values and have been present in less disturbed environments. For clarity, examples of abundance distribution patterns in relation to ES50 values are shown for two species in Fig. 2. *Capitella capitata* is a rapid coloniser and was in some instances the only species present in some samples. This qualified for a low $\mathrm{ES50}_{0.05}$ of 1.5. The tolerance value for the ubiquitous brittle star *Amphiura filiformis* was calculated to 9.5.

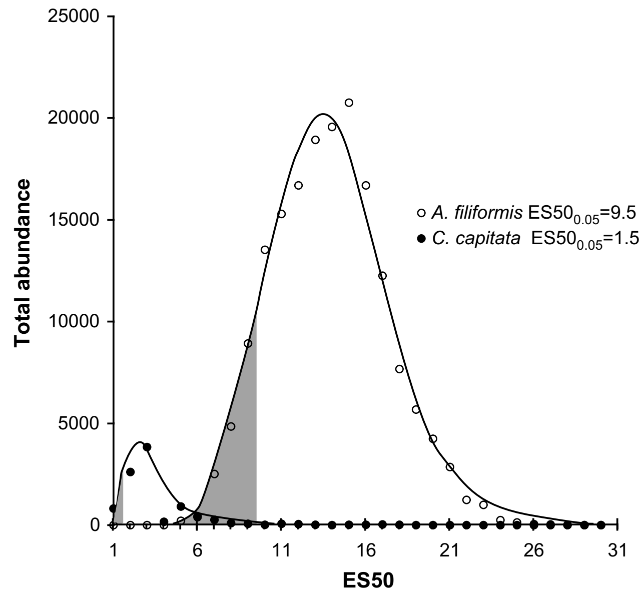

Fig. 2. Examples of total abundance frequency distributions of the pioneer coloniser *Capitella capitata* and the frequently occurring *Amphiura filiformis* in relation to their ES50 values. Shaded areas indicate the 5% abundance distribution in relation to the lowest ES50 values ($\mathrm{ES50}_{0.05}$); for *C. capitata* 1.5 and for *A. filiformis* 9.5.

2.3. Benthic quality assessment

For the assessment of the environmental quality at a particular station, a new benthic quality index (BQI) is proposed

$$\mathrm{BQI} = \left(\sum_{i=1}^{n} \left(\frac{A_i}{\mathrm{tot}A} \times \mathrm{ES50}_{0.05i} \right) \right) \times {}^{10}\log(S + 1)$$

The tolerance value ($\mathrm{ES50}_{0.05}$) of each species found at a station is multiplied with the mean relative abundance (A) of this species ("i") to put weight on common species in relation to rare species. Further, the sum is multiplied with ^{10}logarithm for the mean number of species (S) at the station, as high species diversity is related to high environmental quality. All information related to number of species and abundance at a station is used for this quality assessment. In this study, BQI varied between 0

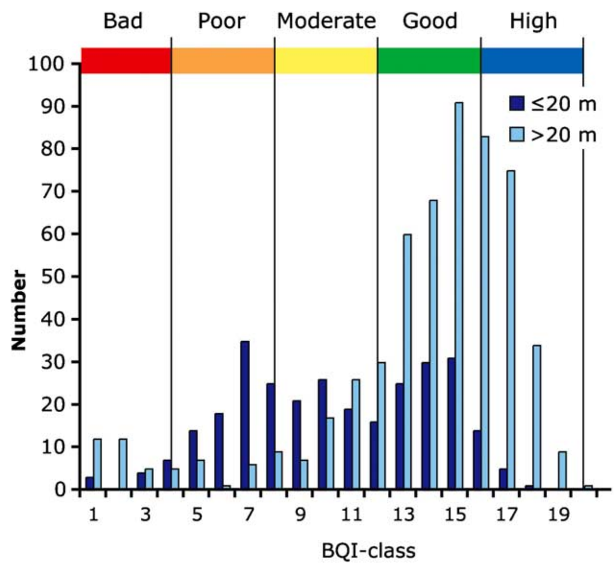

Fig. 3. Frequency distribution of the benthic quality indices (BQIs) for the coastal stations separated into depths >20 m and depths $\leqslant$20 m. The graph also shows the separation of BQI into five different classes of environmental status for depths >20 m according to the Water Framework Directive. The reference value is the greatest value, where BQI = 20 for depths >20 m and BQI = 18 for depths $\leqslant$20 m.

and 20 (reference value) for the coastal stations at depths >20 m, which are the endpoints between "bad" and "high" environmental status according to a total of five stages of classification within the WFD. The other official names for this classification are "poor", "moderate" and "good" (Fig. 3).

Similarity between benthic communities was calculated from $\sqrt{}$ transformed abundance data and presented as multidimensional scaling (MDS) based on Bray Curtis similarities according to Clark and Warwick (1994).

3. Results and discussion

3.1. Species tolerance value

$ES50_{0.05}$ values for 308 species or taxa are available at: www.marine-monitoring.se. The 20 common species including highest and lowest $ES50_{0.05}$, respectively, are listed in Table 2. The span between the lowest ranked species *Capitella capitata* and the highest ranked *Ophelina cylindricaudata* was from 1.5 to 16.0. This means that 5% of these two species numerical distributions, closest to a presumed gradient of disturbance (Fig. 2), were found at stations where on average 1.5 and 16.0 species were found among 50 randomly selected individuals. Examples of hypothetical $ES50_{0.05}$ values of some common species in the area are: *Amphiura filiformis* 9.5, *A. chiajei* 10.6, *Abra nitida* 9.4, *Maldane sarsi* 9.4, *Melinna cristata* 10.4 and *Brissopsis lyrifera* 9.4. Thus, for these species the values are rather similar and slightly above the median value. Lowest values were calculated for *Capitella capitata* 1.5, *Polydora ciliata* 3.4, *Abra alba* 4.0 and *Corbula gibba* 4.7. Species given the highest values are not known to be dominants in any particular area or to have a wide distribution.

3.2. Benthic quality index (BQI)

Mean BQIs have been calculated for 1114 sample occasions along the Swedish west coast. The analysis has

Table 2
Tolerance value of some common species $ES50_{0.05}$

Taxa	$ES50_{0.05}$	Taxa	$ES50_{0.05}$
Capitella capitata	1.5	*Echinocardium cordatum*	9.3
Polydora ciliata	3.4	*Abra nitida*	9.4
Abra alba	4.0	*Maldane sarsi*	9.4
Corbula gibba	4.7	*Amphiura filiformis*	9.5
Macoma balthica	5.8	*Heteromastus filiformis*	9.7
Nephtys hombergii	6.9	*Lumbrineris fragilis*	9.8
Pectinaria koreni	7.0	*Chaetoderma nitidulum*	10.3
Scoloplos armiger	7.1	*Melinna cristata*	10.4
Owenia fusiformis	7.4	*Pectinaria auricoma*	10.6
Scalibregma inflatum	7.5	*Amphiura chiajei*	10.6
Thyasira sarsii	7.5	*Thyasira equalis*	11.0
Arctica islandica	7.5	*Labidoplax buski*	11.4
Ophiura ophiura	7.8	*Ophiura affinis*	11.7
Glycera alba	7.9	*Thracia convexa*	12.3
Ophiodromus flexuosus	8.0	*Ampelisca tenuicornis*	13.0
Pholoe baltica	8.2	*Callianassa tyrrhena*	14.2
Terebellides stroemi	8.3	*Lumbrineris gracilis*	14.7
Priapulus caudatus	8.7	*Ophelina norvegica*	15.0
Nephtys incisa	9.0	*Nephrops norvegicus*	15.6
Thyasira flexuosa	9.1	*Ophelina cylindricaudata*	16.0

Low values indicate tolerant species and high values indicate sensitive species.

been separated for stations >20 m and stations ⩽ 20 m depth as the conditions above and below the halocline has significant impact on the benthic fauna composition (see below). Frequency distributions of BQIs are shown for stations within the coastal zone according to the WFD in Fig. 3 with peaks of the highest BQIs at 15 for stations >20 m and at 14 for those ⩽ 20 m. Classification of the BQI classes in relation the WFD is discussed below.

In Fig. 4, a temporal change of BQI is shown at four stations sampled within the Swedish National Monitoring Programme where some stations have been sampled from 1974. The BQI were similar with minor temporal changes at three of the stations. The fourth station, Alsbäck at 118 m in the Gullmarsfjord, showed aperiodic strong declines related to low oxygen concentrations in the bottom water followed by rapid improvement. The fauna was eliminated in 1979 (Josefson and Widbom, 1988) and 1997 (Nilsson and Rosenberg, 2000).

The second example is from the Gullmarsfjord (Fig. 1) and shows BQI (Fig. 5A), number of species (Fig. 5B), and abundance (Fig. 5C) from five stations showing different successional changes to: (1) decreasing oxygen concentrations in the bottom water (June 1997–February 1998) (Nilsson and Rosenberg, 2000), and (2) re-oxygenated conditions of the bottom water (Rosenberg et al., 2002a). Oxygen concentrations declined with depth and time up to February 1998, and consequently the benthic communities were almost unaffected at 75 m depth, significantly reduced at 85 and 95 m, and periodically eliminated at 105 and 118 m (Alsbäck) depth. These investigations show both the decline and recovery of the number of species (Fig. 5B) and abundance (Fig. 5C) at the same sites at 15 sampling occasions. The information is particularly useful in this context as it includes both species tolerant and sensitive to increased

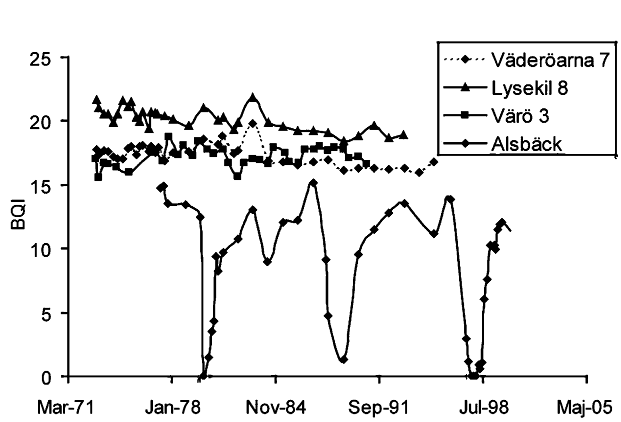

Fig. 4. Temporal changes in the benthic quality index (BQI) is shown for four stations in the Skagerrak sampled within the Swedish National Monitoring Programme. The benthic fauna at the station Alsbäck at 118 m in the Gullmarsfjord was severely impacted by hypoxia during three periods.

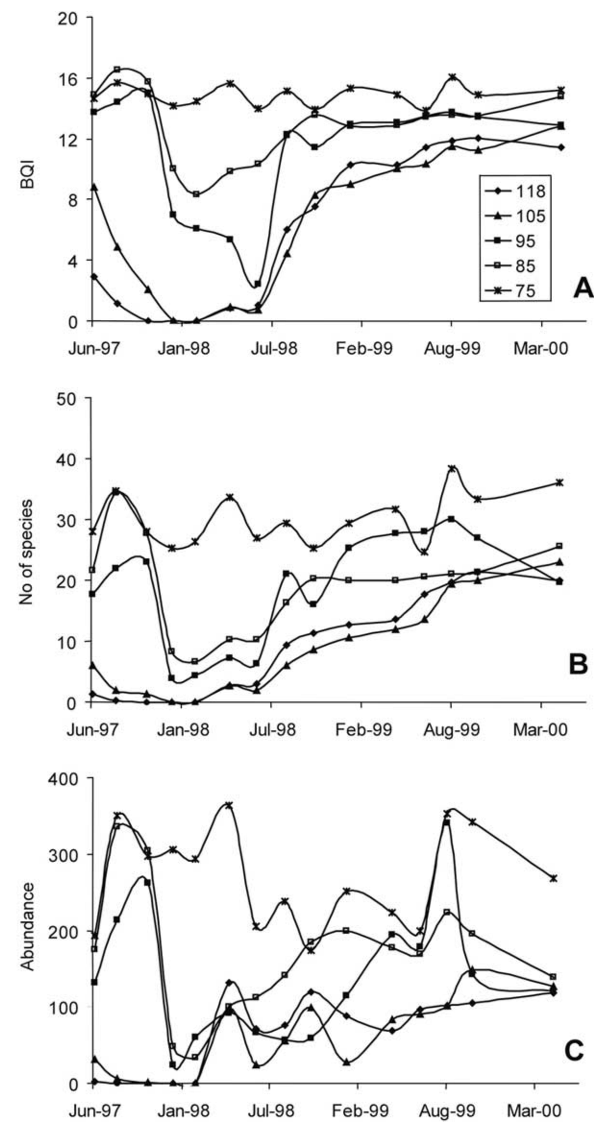

Fig. 5. Benthic quality index (BQI) (A), mean number of species (B) and abundance (C) per 0.1 m^2 are shown for five stations at depths between 75 (reference site) and 118 m in the Gullmars fjord during decreasing oxygen conditions from June 1997 up to February 1998 (from Nilsson and Rosenberg, 2000), and during the subsequent recovery during re-oxygenated conditions to April 2000 (Rosenberg et al., 2002). During hypoxia, the oxygen concentrations declined with depth. Error bars are omitted for simplification but shown in the original publications.

hypoxia and species that are rapid colonisers. The BQIs show a clear response to the changes in oxygen concentrations with clear declines related to increasing hypoxia and a successive increase following re-oxygenation of the bottom water. At the end of the sampling, the BQI at 85 m depth was close to that of the reference stations had 75 m, and the BQIs at the other stations has become close to that level. The succession of the BQIs was somewhat similar to that of number of species, whereas the abundance showed a greater temporal variation.

3.3. Reference value and classification of environmental status

According to the WFD, the coastal environmental status should be classified into 5 categories. This should be done separately for each coastal area based on the typology of that area. In the present paper, all coastal soft bottoms in the Skagerrak, Kattegat and Öresund have been used as one unit separated by depth only; i.e. bottoms ⩽20 m and bottoms >20 m, respectively, which is discussed below. Most sediments in this study are muddy with silt–clay as the main component; at some stations ⩽20 m depth the silt–clay could have been mixed with sand. Samples from sandy sediments were not used in this study as grab samples from such sediments have a poor vertical penetration and are not quantitative. Morover, sandy bottoms are erosion bottoms which are more or less continuously physically disturbed with no or minor accumulation of organic material and contaminants.

Frequency distributions of BQI from all coastal stations are presented in Fig. 3. The reference value of a defined bottom area should, according to the WFD, be selected as the greatest value; i.e. in this study BQI = 20. This was found at >20 m depth, and in the following we focus on this depth stratum. Based on this reference value the BQI classes were divided into five classes according to the WFD of equal size between 0 and 20, i.e. a class breadth of 4 units. Thus, the frequency distribution at BQI = 16.0 was defined as the border between "High" and "Good", i.e. values ⩾16.0 will classify such environments as having "High" status (Fig. 6). According to the WFD, this is when "all disturbance-sensitive taxa associated with undisturbed conditions are present". The border between "Good" and "Moderate" was set as 12.0, and "Good" was defined in the WFD to relate to "most of the sensitive taxa of the type-specific communities are present". The lower limit of "Moderate" was at 8.0, and here "Taxa indicative of pollution are present". The border between "Poor" and "Bad" conditions was set at 4.0. Classification of BQI for depths ⩽20 m was made in the same way and the result is presented in Fig. 6. Most of the stations at depths >20 m in this study were from areas considered as not particularly impacted and away from discharge points. That means that the peak is likely to occur within the range of "High" and "Good" environmental status. It is important for this assessment that the material is large. A small set of information from disturbed areas should not be used for the classification of reference value. It may happen in this type of analysis that a single or few BQIs are extremely high and outliers of the frequency distribution. If this is the case and these index classes contribute <1% of the total numbers, we suggest they are omitted from the classification.

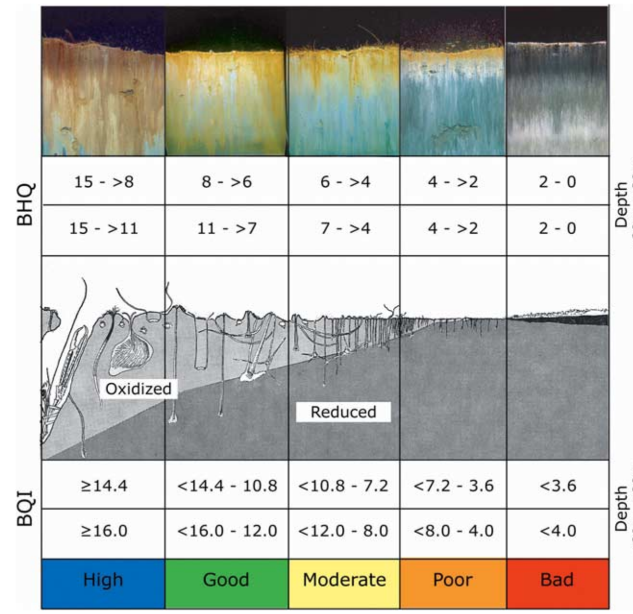

Fig. 6. Model of the faunal successional stages along a gradient of increasing disturbance from left to right (after Pearson and Rosenberg, 1978). Sediment profile images (colours enhanced) are shown on the top where brownish colour indicate oxidised conditions and black reduced conditions, and the benthic habitat quality (BHQ) indices (Nilsson and Rosenberg, 1997) are presented for depths >20 m and ⩽20 m. The benthic quality indices (BQIs) for the different environmental status according to the Water Framework Directive are presented for depths >20 m and ⩽20 m at the bottom of the figure.

3.4. MDS analysis of BQI

In the present study BQIs varied between 0 and 22 for depths >20 m (Fig. 4). MDS was used to analyse how different BQIs were distributed in relation to the composition of the benthic communities. MDS is a useful technique to analyse similarity between benthic communities (Gray et al., 1988). Similarities of benthic faunal compositions at 80 randomly selected stations at the Swedish west coast were analysed and presented in a MDS plot. Instead of presenting the station numbers in the plot, values of BQIs are shown (Fig. 7A). High numbers appear together in the lower central part of the plot, and low numbers are found at some distance from these. Thus, stations suggested to have "High" or "Good" environmental status are plotted together, whereas those with "Poor" and "Bad" status are spread out in the plot.

In Fig. 7B, the BQIs in Fig. 7A were replaced by station depths. It is clear that shallow stations group on the right side and deeper stations on the left side. The line in the figure shows approximately stations above and below 20 m depth, but some station around this depth are mixed in the MDS plot. The lower vertical distribution limit of the halocline is at approximately 20 m depth. Water circulation along the Swedish west coast is dominated by a two-layer estuarine flow, driven by the

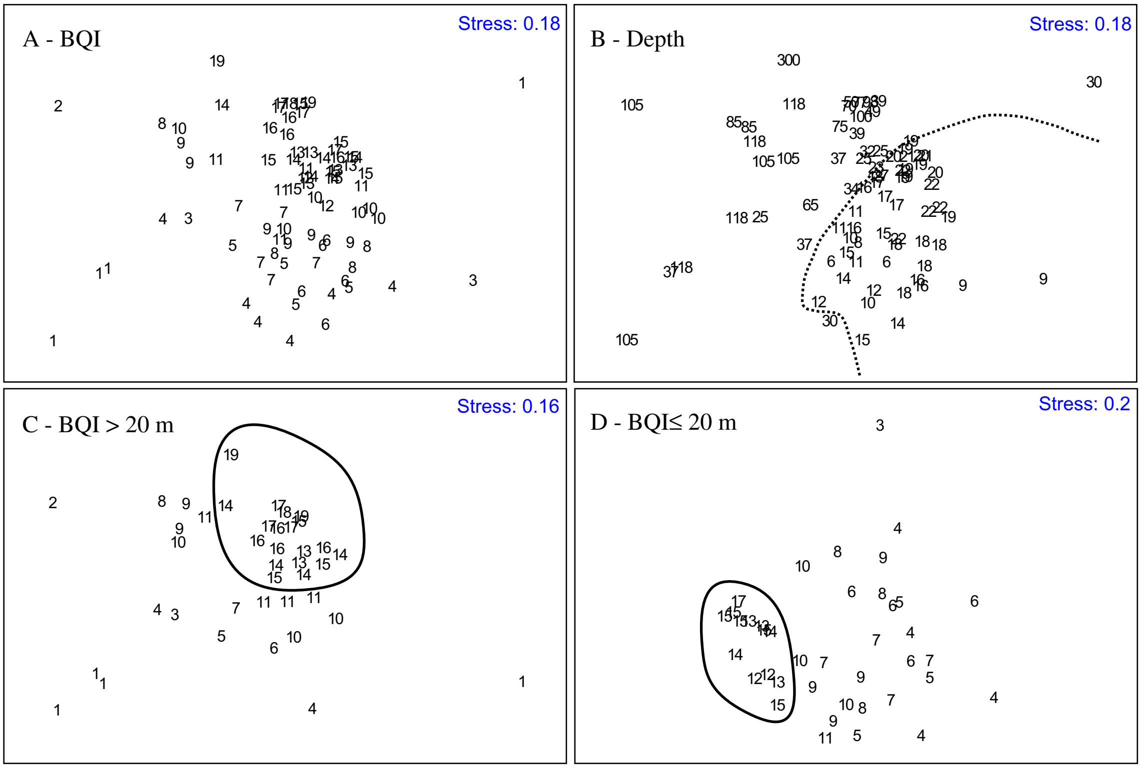

Fig. 7. MDS (multidimensional scaling) plots of benthic faunal similarities where the station numbers in the plot have been replaced by benthic quality indices (BQIs) from 80 randomly selected stations (A), and the same stations replaced by depths (B), where the demarcation line is approximately separating stations above and below 20 m depth. Below are MDS plots of the community composition from 40 randomly selected stations at depths >20 m (C) and ⩽20 m (D) where the station numbers are replaced by BQIs. Stations with high BQI are encircled.

outflow of low saline water from the Baltic. The surface water has a salinity of 15–30 psu down to the halocline at 12–20 m depth, and below this depth the salinity is between 32 and 34 psu (Rydberg et al., 1990). Greater variations in salinity (and temperature) above the halocline have been shown to create a more stressful environment than below the halocline. As a consequence the benthic faunal diversity, abundance and biomass are all generally lower in surface waters than below the halocline (Rosenberg and Möller, 1979; Rosenberg et al., 1992). Thus, the approximate lower distribution limit of the halocline is suggested to be the main reason why the BQIs are different. In other marine areas without salinity stratification such separation would not be necessary and the below-halocline conditions could probably apply also to shallow waters.

The 20 m depth level was consequently used for a depth-stratified analysis of the benthic faunal composition of another 40 randomly selected stations at each depth interval (Fig. 7C and D). The groups of stations with the highest BQIs were together and encircled, whereas the other indices were spread out. Encircled BQIs at depths >20 m were 12 and 19, whereas at depths ⩽20 m they were between 11 and 17. This discrepancy shows that the classification was depth related. The encircled numbers are suggested to represent relatively undisturbed benthic environments of "High" and "Good" status according to the WFD.

4. General discussion

4.1. Comparison of different methods to classify tolerant and sensitive species

The first attempt to classify tolerant and sensitive marine species to various degrees of disturbance was based on a literature review of scientific data from different marine areas (Pearson and Rosenberg, 1978). The authors classified some common benthic species appearing in four zones along a gradient of disturbance: normal, transitory, polluted and grossly polluted. Rhoads and Germano (1986) made a similar classification for a gradient of disturbance in the USA. The pioneer species appearing in the early succession following improved conditions are often called opportunists or r-strategists (Gray, 1979) and are characterised by small size, short life cycles and being rapid colonisers of defaunated areas (e.g. *Capitella capitata*). Larvae of these species may be tolerant to colonise enriched and sulphidic sediments, but adults may not necessarily be tolerant towards oxygen deficiency or physical disturbance.

Thus, during increased stress, such as increased oxygen deficiency, other species may be more tolerant than the opportunists (Diaz and Rosenberg, 1995). One example is the bivalve *Arctica islandica*, which is tolerant to hypoxia for weeks (Rosenberg and Loo, 1988); it may live up to 100 yr, and is a typical K-strategist. Thus, both r-strategists and K-strategists may be classified as tolerant species, where the former invade areas when conditions improve and the latter category may be among the last survivors when conditions worsen. It is not the purpose of this work to further analyse these discrepancies of tolerance. Another factor affecting the distribution of tolerant species is salinity. Pearson and Rosenberg (1978) demonstrated that the general dominance of *C. capitata* in disturbed marine areas was replaced by the bivalve *Macoma balthica* and oligochaetes in low salinities in the Gulf of Bothnia (Northern Baltic).

Grall and Glémarec (1997) classified species into five groups of sensitivity. Most of the species fell into the category of sensitive species, and among the most tolerant species were *Capitella capitata* and *Scolelepis fuliginosa*, the same as in the categorisation by Pearson and Rosenberg (1978). Based on these publications, Borja et al. (2000) and Borja et al. (2003) classified more than 2000 taxa into five different groups of sensitivity. Data were obtained from six European areas with different pollution sources. Simboura and Zenetos (2002) similarly classified species from a number of Mediterranean areas, particularly Greek waters, but they grouped the species into three categories only. The classifications of taxa were used in different mathematical formulas to assess the degree of disturbance; Borja et al. (2003) used eight levels in a "Biotic Index", and Simboura and Zenetos (2002) presented a "Benthic Index" (BENTIX) ranging between 0 and 6.

As these validations were used on literature data and personal experience of the authors, they seem to be useful in analysing the degree of disturbance for the areas presented. The classifications of the different species or taxa were, however, subjectively made and may vary between scientists and geographical areas (see Table 1 for summary of indices). An objective way to rank species was described by Gray (1981) and analysed for larger datasets by Pearson et al. (1983). By plotting log-normal frequency distributions of individuals among species, the authors were able to identify sensitive species at a break point of the curve. This method is useful in identifying species that may either increase or decrease in relation to a change in disturbance. The technique was not used for the assessment of the degree of disturbance in different areas. Methods to do such spatial disturbance analyses were presented by Reish (1955) for Los Angeles Harbour, by Leppäkoski (1975) for Swedish and Finnish coastal areas, and by Weisberg et al. (1997) for the Chesapeake Bay estuary. Although the first mentioned two methods were useful in the particular cases they were used, they are not applicable in a wider sense. For example, Leppäkoski's "Benthic Pollution Index" is based on disturbed bottom areas and not immediately useful for the identification of tolerant species. The evaluation of benthic habitat quality in the Chesapeake estuary by Weisberg et al. (1997) is objective, and they used among their 17 candidate metrics the depth distribution of benthic species in the sediment. Admittedly, this may be a most useful variable for a quality assessment as deep burrowing species commonly represent mature successional stages. In practice, however, many of the variables used in their "Index of Biotic Integrity" (B-IBI) are not available for such calculations.

Sanders (1968) diversity index, the rarefaction technique, was initially used to compare the diversity between different latitudes and depth gradients in the sea. Rygg (2002) successfully used the improved version by Hurlbert (1971) of this index for an objective identification of species and taxa in relation to various degrees of disturbance. Rygg used 1080 grab samples from different areas along the Norwegian coast and presented sensitive index values for 200 commonly found taxa. The index used was a calculation of the expected number of species among 100 individuals (ES100) from which Rygg calculated the average of the five lowest index values (ES100 $\min_5$) to obtain an "Indicator Species Index" (ISI) for each species represented in at least 50 samples. Based on this, benthic communities were classified into five degrees of disturbance by calculating the mean ES100 $\min_5$ for all species at a station.

By calculating ES100 $\min_5$ only information from five samples is used. There is then a possibility that a sensitive species is represented by one or few individuals in some samples where otherwise tolerant species dominate. The index for such a species will than be low. Instead, it is here suggested to use the abundance values calculated from the 5% lowest abundance of a particular species ($ES50_{0.05}$). This species tolerance value is assumed to be representative for the greatest tolerance level for that species along an increasing gradient of disturbance, i.e. if the stress would be increased a little more that species will not be present any more. This method is similar to that proposed by Gray and Pearson (1982) to assess tolerant species at a breaking point in the log-normal frequency distribution. This will reduce the weight of out-layers in calculating the index.

Based on $ES50_{0.05}$, the benthic environmental quality was assessed by calculating a BQI. The MDS plot of stations with known disturbance was used as a reference for this index (Fig. 7). The mature communities all had high index values, whereas communities in pioneering or declining successional stages were much lower but rather comparable. As the methods for obtaining the species tolerance value and the index for a community at a particular station are objective, the same method could be used also for other marine areas in Europe and elsewhere. The calculation of ES50 is based on criteria that the size of the sample area is the same, and that distribution of individuals among species is random. The first criterion is generally fulfilled as the same grab (0.1 m^2 Smith–McIntyre grab) was used, and means were calculated from similar number of replicates ($n = 3$–5). The latter criterion of random distribution may not always be the case, particularly not when some species appear as strong dominants. For practical purposes, however, the sampling technique seems to work in an accurate way. It is, however, advisable to use many stations and replicates for the quality assessment of an area. One factor that may cause misinterpretations is, however, taxonomy. Different taxonomists have a various degree of skilfulness and the quality of keys for the identification varies between geographic areas and animal groups. Juvenile specimens are often particularly difficult to identify to species level. It is not possible to evaluate how this will affect the techniques used for environmental assessment, but it should have an impact on all methods where number of species or taxa are involved. In the present study, species had to be found at $\geqslant 20$ occasions to be included, which will exclude rare and some falsely identified species from this analysis.

4.2. Other methods for benthic quality assessment

Identification of species is not only a matter of skilful taxonomists, it is laborious and tedious; thus it is expensive. Another or complementary method for the analysis of benthic habitat quality is to analyse sediment profile images (SPIs). This is comparatively more cost-efficient and many more samples than grab samples can be obtained over the same time, and the analysis is rapid (Rhoads and Germano, 1986). The SPI technique has been successfully used to assess the impact of, e.g. pollution (Valente et al., 1992), mariculture (O'Connor et al., 1989; Karakassis et al., 2002), impact of demersal trawling (Nilsson and Rosenberg, 2003), drilling (Rumohr and Schomann, 1992), and oxygen deficiency (Nilsson and Rosenberg, 1997; Nilsson and Rosenberg, 2000; Rosenberg et al., 2002a).

Rhoads and Germano (1986) suggested than an organism-sediment index (OSI) could be used for assessing the sediment habitat quality. That index is based on the depth distribution of the apparent redox potential discontinuity (aRPD) in the sediment, the presence/absence of methane in the sediment, measurements of oxygen in the near-bottom water,and the subjective determination of the faunal successional stage (Table 1). Thus index varies between −10 and +11. The OSI has been used to assess the sedimentary habitat quality in Rhode Island Sound (Valente et al., 1992) and from mariculture in Ireland (O'Connor et al., 1989) and in Greece (Makra et al., 2001). Nilsson and Rosenberg (1997) developed a benthic habitat quality (BHQ) index where structures on the sediment surface, structures in the sediment, and the aRPD were parameterized. This index varied between 0 and 15 where high numbers were associated with mature benthic faunal successional stages and low numbers with pioneering stages or azoic bottoms (Fig. 6). The assignment of images to successional stages when using the OSI index could be made in an objective way from the BHQ index (Solan and Kennedy, 2002). Thus, use of SPI technique may be considered an alternative or a complementary objective method to traditional analysis of benthic faunal composition. Analysis of SPIs does, however, not include the identification of species or quantification of abundance and biomass.

The BHQ index could also be a useful tool for the WFD in assessing the benthic habitat quality. Nilsson and Rosenberg (2000) showed how the BHQ index could be assigned to different successional stages in the Pearson and Rosenberg (1978) model, and Nilsson and Rosenberg (2000) and Rosenberg et al. (2002a) showed that each of the variables: number of species, abundance and biomass strongly correlated with the BHQ index under changing oxygen concentrations. Karakassis et al. (2002) similarly showed that multivariate patterns obtained through SPI analysis were highly correlated to those obtained from standard multivariate analysis of macrofauna. Instead of the earlier separation of the BHQ index into four successional stages, we suggest that the BHQ index is dived into five classes (see Fig. 6) to be used according to the WFD. This index is, as the B-IBI index used by Weisberg et al. (1997), giving high scores to deep burrowing fauna in association with a deep distribution of the RPD, attributes that are associated with mature benthic communities.

4.3. New model adjusted to the WFD classification

We suggest that the Pearson and Rosenberg (1978) model is adjusted from four to five successional stages of benthic communities to support the environmental quality assessment as required from the WFD. These five stages are related to the different ecological status proposed in the WFD: "High", "Good", "Moderate", "Poor" and "Bad". The ecological status of a particular station or habitat could be assessed in an objective way by calculating the BQI. The suggested classification into the five groups is shown in Fig. 6. The delimitations between the different benthic communities are set for practical purposes to be a useful tool in environmental quality assessment. In nature, the changes occur as a continuum without any clear breakpoints. The WFD states that the environmental quality assessment should be related to an ecological quality ratio (EQR), which varies between 0 and 1. The EQR for the five classifications in Fig. 6 can each be calculated by dividing the BQIs in each class with the reference value, i.e. by 20 for depths >20 m and by 18 for depths $\leqslant$20 m.

We suggest that the new model may be useful on sublittoral soft bottoms in most temperate and boreal areas, as the original Pearson and Rosenberg (1978) model has proven to be useful for assessing the degree of disturbance in these waters. The model may also be useful in enclosed seas as the Mediterranean and the Baltic. Karakassis et al. (2002) and Rosenberg et al. (2002b) have shown that the sedimentary habitat quality in the Mediterranean can be evaluated from both faunal data and SPI, and that these variables correlate. Bonsdorff et al. (1996) showed that benthic faunal variables and SPI were both useful for habitat quality assessment in the Åland Archipelago in the northern Baltic proper. However, the benthic faunal composition in the brackish waters of the Baltic is indeed different from true marine areas, and the applicability of the BQI formula for such low saline areas has to be evaluated.

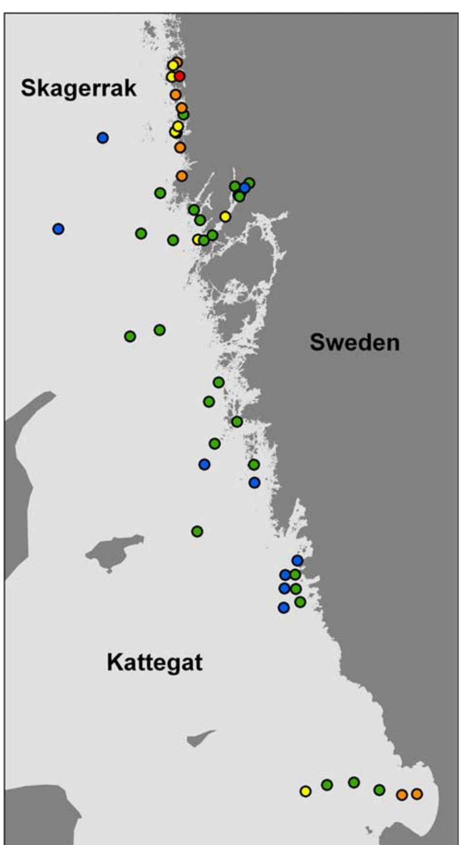

Fig. 8. Classification according to the Water Framework Directive of the benthic stations in the Kattegat and the Skagerrak sampled in 1990 based on information in Fig. 6.

4.4. Environmental quality status along the west coast of Sweden

Based on the BQI classifications in Fig. 6, we have as an example assessed the environmental quality status according to the WFD for the Kattegat and the Skagerrak in 1990 (Fig. 8). Stations in the northern Kattegat and offshore stations were classified as having a "High" or "Good" quality. A "Poor" quality was found at two stations in the inner part of the southeast Kattegat. These stations at 16 and 18 m depth were affected by hypoxia (Rosenberg et al., 1992). Other stations (unpublished personal records) with "Poor" conditions were found at 12–30 m depth along the Swedish Skagerrak coast. There, the number of species and abundance were exceptionally low at several stations. The station with "Bad" quality was from a deep depression in an enclosed fjord at 30 m depth. The reason for this degraded benthic quality needs further evaluation. The example shows that areas where the benthic environmental conditions are acceptable ("High" or "Good") can be separated from those that show clear signs of disturbance. Thus, the method presented here seems to be a useful tool within the WFD for assessing benthic habitat quality.

Acknowledgements

Financial support was given by the Swedish Environment Protection Agency, and "Stiftelsen San Michele" (to R.R.). We appreciate the use of benthic data sets from the Swedish National Monitoring Programme, several regional monitoring programmes and scientific data collected and analysed of many persons. Alf Norkko gave valuable comments on the manuscript.

References

Bellan, G., 1985. Effects of pollution and man-made modifications on marine benthic communities in the mediterranean: A review. In: Moraitou-Apostopoulou, M., Kiortis, V. (Eds.), Mediterranean Marine Ecosystems. Plenum Publishing Corporation, New York, pp. 163–194.

Boesch, D.F., Rosenberg, R., 1981. Response to stress in marine benthic communities. In: Barrett, G.-W., Rosenberg, R. (Eds.), Stress Effects on Natural Ecosystems. Wiley, New York, pp. 179–200.

Bonsdorff, E., Diaz, R.J., Rosenberg, R., Norkko, A., Cutter, G.R.J., 1996. Characterization of soft-bottom benthic habitats of the Åland Islands, northern Baltic Sea. Mar. Ecol. Prog. Ser. 142, 235–245.

Borja, A., Franco, J., Pérez, V., 2000. A marine biotic index to establish the ecological quality of soft-bottom benthis within European estuarine and coastal environments. Mar. Pollut. Bull. 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. Mar. Pollut. Bull. 46, 835–845.

Clark, K.R., Warwick, R.M., (1994). Change in marine communities: An approach to statistical analysis and interpretation. Plymouth Marine Laboratory.

Diaz, R.J., Rosenberg, R., 1995. Marine benthic hypoxia: a review of its ecological effects and behavioural responses of marine macrofauna. Oceanogr. Mar. Biol. Ann. Rev. 33, 245–303.

Grall, J., Glémarec, M., 1997. Using biotic indices to estimate macrobenthic community perturbations in the Bay of Brest. Estuar. Coast. Shelf Sci. 44, 43–53.

Gray, J.S., 1979. Pollution-induced changes in populations. Philos. Trans. R. Soc. 286, 545–561.

Gray, J.S., 1981. Detecting pollution induced changes in communities using the log-normal distribution of individuals among species. Mar. Pollut. Bull. 12, 173–176.

Gray, J.S., Mirza, F.B., 1979. A possible method for detecting pollution induced disturbance on marine benthic communities. Mar. Pollut. Bull. 10, 142–146.

Gray, J.S., Pearson, T.H., 1982. Objective selection of sensitive species indicative of pollution-induced change in benthic communities. I comparative methodology. Mar. Ecol. Prog. Ser. 9, 111–119.

Gray, J.S., Aschan, M., Carr, M.R., Clarke, K.R., Green, R.H., Pearson, T.H., Rosenberg, R., Warwick, R.M., 1988. Analysis of community attributes of the benthic macrofauna of Frierfjord/Langesundfjord and in a mesocosm experimental. Mar. Ecol. Prog. Ser. 46, 151–165.

Heip, C., 1995. Eutrophication and zoobenthos dynamics. Ophelia 41, 113–136.

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

Josefson, A.B., Widbom, B., 1988. Differential response of benthic macrofauna and meiofauna to hypoxia in the Gullmar Fjord basin. Mar. Biol. 100, 31–40.

Karakassis, I., Tsapakis, M., Smith, C.J., Rumohr, H., 2002. Fish farming impacts in the Mediterranean studied through sediment profiling imagery (SPI). Mar. Ecol. Prog. Ser. 227, 125–133.

Kolwitz, R., Marsson, M., 1909. Ökologie der tierischen Saprobien. Internat. Rev. Ges. Hydrobiol. 2, 126–152.

Leppäkoski, E., 1975. Assessment of degree of pollution on the basis of macrozoobenthos in marine and brackish-water environments. Acta Acad. Abo. Ser. B 35, 1–90.

Makra, A., Thessalou-Legaki, M., Costelloe, J., Nicolaidou, A., Keegan, B.F., 2001. Mapping the pollution gradient of the Saronikos Gulf benthos prior to the operation of the Athens sewage treatment plant, Greece. Mar. Pollut. Bull. 42, 1417–1419.

Nilsson, H.C., Rosenberg, R., 1997. Benthic habitat quality assessment of an oxygen stressed fjord by surface and sediment profile images. J. Mar. Syst. 11, 249–264.

Nilsson, H.C., Rosenberg, R., 2000. Succession in marine benthic habitats and fauna in response to oxygen deficiency: analysed by sediment profile imaging and by grab samples. Mar. Ecol. Prog. Ser. 197, 139–149.

Nilsson, H.C., Rosenberg, R., 2003. Effects on marine sedimentary habitats of experimental trawling analysed by sediment profile imagery. J. Exp. Mar. Biol. Ecol. 285–286, 453–463.

O'Connor, B.D.S., Costolloe, J., Keegan, B.F., Rhoads, D.C., 1989. The use of REMOTS technology in monitoring coastal enrichment resulting from mariculture. Mar. Pollut. Bull. 20, 384–390.

Pearson, T.H., Gray, J.S., Johannessen, P.J., 1983. Objective selection of sensitive species indicative of pollution-induced change in benthic communities. 2. Data analysis. Mar. Ecol. Prog. Ser. 12, 237–255.

Pearson, T.H., Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanogr. Mar. Biol. Ann. Rev. 16, 229–311.

Pielou, E.C., 1969. An Introduction to Mathematical Ecology. Wiley, New York.

Reish, D.J., 1955. The relation of polychaetous annelids to harbor pollution. Pub. Healt. Rep. Wash. 70, 1168–1174.

Rhoads, D.C., Germano, J.D., 1986. Interpreting long-term changes in benthic community structure: a new protocol. Hydrobiologia 142, 291–308.

Rosenberg, R., Loo, L.-O., 1988. Marine eutrophication induced oxygen deficiency: effects on soft bottom fauna, western Sweden. Ophelia 29, 213–225.

Rosenberg, R., Möller, P., 1979. Salinity stratified benthic macrofaunal communities and long-term monitoring along the west coast of Sweden. J. Exp. Mar. Biol. Ecol. 37, 175–203.

Rosenberg, R., Nilsson, K., Landner, L., 1975. Effects of a sulphate pulp mill on the benthic macrofauna in a firth of the Bothnian Sea. Merentutkimuslait. Julk./Havsforskningsinst. Skr. No. 239, pp. 289–300.

Rosenberg, R., Loo, L.-O., Möller, P., 1992. Hypoxia, salinity and temperature as structuring factors for marine benthic communities in an eutrophic area. Neth. J. Sea Res. 30, 121–129.

Rosenberg, R., Agrenius, S., Hellman, B., Nilsson, H.C., Norling, K., 2002a. Recovery of benthic habitats and fauna in a Swedish fjord following improved oxygen conditions. Mar. Ecol. Prog. Ser. 234, 43–53.

Rosenberg, R., Grémare, A., Amouroux, J.-M., Nilsson, H.C., 2002b. Benthic habitats in the northwest Mediterranean characterised by sedimentary organics, benthic macrofauana and sediment profile images. Estuar. Coast. Shelf Sci. 57, 297–311.

Rumohr, H., Schomann, H., 1992. REMOTS sediment profiles around an exploratory drilling rig in the southern North Sea. Mar. Ecol. Prog. Ser. 91, 303–311.

Rydberg, L., Edler, L., Floderus, S., Granéli, W., 1990. Interaction between supply of nutrients, primary production, sedimentation and oxygen consumption in SE Kattegat. Ambio 19, 134–141.

Rygg, B., 2002. Indicator species index for assessing benthic ecological quality in marine waters of Norway. Norwegian Institute for Water Research, Oslo, Report SNO 4548-2002, 32p.

Sanders, H.L., 1968. Marine benthic diversity: a comparitive study. Am. Nat. 102, 243–282.

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

Solan, M., Kennedy, R., 2002. Observation and quantification of in situ animal–sediment relations using time-lapse sediment profile imagery (t-SPI). Mar. Ecol. Prog. Ser. 228, 179–191.

Swartz, R., Schults, D., Ditsworth, G., DeBen, W., Cole, F., 1985. Sediment toxicity, contamination, and macrobenthic communities near a large sewage outfall. Special Technical Testing Publication 865, American Society of Testing Materials Philadelphia, USA, pp. 152–175.

Valente, R.M., Rhoads, D.C., Germano, J.D., 1992. Mapping of benthic enrichment patterns in Narragansett Bay, Rhode Island. Estuaries 15, 1–17.

Warwick, R.M., 1986. A new method for detecting pollution effects on marine macrobenthic communities. Mar. Biol. 92, 557–562.

Warwick, R.M., 1988. The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar. Pollut. Bull. 19, 259–268.

Weisberg, S., Ranasinghe, J.A., Dauer, D.M., Schaffer, L.C., Diaz, R.J., Frithsen, J.B., 1997. An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. Estuaries 20, 149–158.