Application of the Benthic Ecosystem Quality Index 2 to benthos in Dutch transitional and coastal waters

W.M.G.M. van Loon [a,*], A.R. Boon [b], A. Gittenberger [c], D.J.J. Walvoort [d], M. Lavaleye [e], G.C.A. Duineveld [e], A.J. Verschoor [f]

[a] Rijkswaterstaat, Ministry of Infrastructure and the Environment, Zuiderwagenplein 2, 8224 AD Lelystad, The Netherlands
[b] Deltares, P.O. Box 177, 2600 MH Delft, The Netherlands
[c] Gimaris, BioScience Park Leiden, J.H. Oortweg 21, 2333 CH Leiden, The Netherlands
[d] Alterra, Wageningen University & Research Centre, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[e] NIOZ, P.O. Box 59, 1790 AB Den Burg, The Netherlands
[f] National Institute for Public Health and the Environment, P.O. Box 1, 3720 BA Bilthoven, The Netherlands

ARTICLE INFO

ABSTRACT

The Benthic Ecosystem Quality Index 2 (BEQI2) is the Dutch multi-metric index (MMI) for assessing the status and trend of benthic invertebrates in transitional and coastal waters for the Water Framework Directive (WFD). It contains the same indicators, i.e. species richness, Shannon index and AMBI, as in the multivariate m-AMBI. The latter MMI has been adopted by several European countries in the context of WFD implementation. In contrast to m-AMBI, the BEQI2 calculation procedure has been strongly simplified and consists of two steps, i.e. the separate indicator values are normalized using their long-term reference values resulting in three Ecological Quality Ratios (EQRs), which are subsequently averaged to give one BEQI2 value. Using this method only small numbers of samples need to be analysed by Dutch benthos laboratories annually, without the necessity to co-analyse a larger historical dataset. BEQI2 EQR values appeared to correlate quantitatively very well with m-AMBI EQR values. In addition, a data pooling procedure has been added to the BEQI2 tool which enables the pooling of small core samples (0.01–0.025 m^2) into larger standardized data pools of 0.1 m^2 in order to meet the data requirements of the AMBI indicator and to obtain comparable reference values. Furthermore, the BEQI2 tool automatically and efficiently converts species synonym names into standardized species names. The BEQI2 tool has been applied to all Dutch benthos data monitored by Rijkswaterstaat in the period of 1991–2010 in the transitional and coastal waters and salt lakes and these results are reported here for the first time. Reference values for species richness and Shannon index (99 percentile values) and AMBI reference values (1 percentile values) were estimated for all water body–ecotopes and are discussed. BEQI2 results for all these water bodies are discussed in view of natural and human pressures. The pressure sensitivity of the BEQI2 for sewage and dredging/dumping, via the state variables oxygen and suspended matter respectively, was demonstrated.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The Water Framework Directive (WFD) is the European legislation aimed to protect and improve the European aquatic habitats and ecosystems for fresh, brackish and coastal marine waters (EC 2000). The WFD constitutes a comprehensive framework for the monitoring, assessment and reporting of the status and trends of ecological, chemical and hydrological quality elements of all European water bodies with the exception of the marine waters outside the 1-mile coastal zone (Hering et al., 2010). The definition of ecological status takes into account

specific aspects of the biological quality elements, one of them being the composition and abundance of benthic invertebrate fauna.

In the WFD good ecological status for benthos in transitional and coastal waters is defined by the following normative definition: "The level of diversity and abundance of invertebrate taxa is slightly outside the range associated with the type-specific conditions. Most of the sensitive taxa of the type-specific communities are present". Evaluation of the status of benthos is frequently achieved by applying a multi-metric index (MMI) which combines several indicators each addressing different stressors or components of the biocoenosis (Hering et al., 2010). The MMIs for benthos which have been reviewed by Boon et al. (2011) typically contain three indicators: (a) diversity, (b) (relative) abundance and (c) a ratio of sensitive, tolerant and opportunistic

species. An important advantage of the use of a multi-metric index is that it clearly reduces the sensitivity of the individual indicators for natural variations (Borja et al., 2011; Kröncke and Reiss, 2010) which enhances the possibility to detect changes of benthic communities as a result of changing human pressures.

Dutch transitional (Westerschelde, Ems–Dollard) and coastal (Wadden sea, Delta coast, Dutch coast and Wadden coast) water bodies (Fig. 1) are subject to a variety of natural and human pressures and resulting changes in physical and chemical state variables (De Jonge et al., 2014; Janssen and Mulder, 2005; Meire et al., 2005; Van Buuren et al., 2010). Typical natural pressures in tidal estuaries are relatively high maximum flow rates, salinity variations, and periodic exposure on intertidal flats (Bouma et al., 2005). Occasional heavy rainfall lead to large variations in salinity (Lacroix et al., 2004) while seasonal temperature variations and particularly extremely cold winters have an impact on benthos especially on intertidal flats (Beukema and Essink, 1986; Mees et al., 1993). Gradual climate warming appears to lead to density shifts of certain species (Beukema et al., 2009). Typical human pressures in the Dutch estuaries are discharges of organic material and nutrients from agriculture and sewage causing eutrophication and hypoxia (Essink and Beukema, 1986), toxic substances, dredging and dumping, bottom trawling, and introduction of non-native species by shipping. The dominant human pressure in the Dutch coastal zone is most likely beam trawling by small vessels (<200 hp). In addition,

extraction of sand just outside the coastal zone and subsequent nourishment of the beach and shore-face impact benthos on many parts of the Dutch coast (Van Dalfsen and Essink, 2001). The complex mix of natural and human pressures especially in estuaries makes it difficult to discriminate the different components quantitatively (Schaffner et al., 2001; Hiscock et al., 2004 for a review).

In the North East Atlantic countries, a range of MMIs for benthos in transitional and coastal waters has been developed (Borja et al., 2007; Josefson et al., 2009; Muxika et al., 2007; Teixeira et al., 2009). In a review, Boon et al. (2011) suggested that the indicators which are used by the m-AMBI, viz. species richness, Shannon–Wiener index and AMBI, are probably also suited for the WFD assessment of benthos in Dutch waters. The sensitivity of the m-AMBI for human pressures has been demonstrated in a suite of circumstances, including impacts due to organic material and pollutants, dredging and dumping, oil platforms and aquaculture (Muxika et al., 2005, Borja et al., 2008; Borja et al., 2011). Furthermore, the m-AMBI and comparable Benthic Assessment Tool (BAT) are used by several countries in Europe (Spain, Portugal, Germany) which is favourable for intercalibration purposes.

Therefore, it was first attempted to adopt this commonly used Western European multi-metric tool for Dutch marine benthos. However, several methodological and practical drawbacks were encountered when testing the m-AMBI software (http://ambi.azti.es). First is the

Fig. 1. Global locations of the investigated water bodies. A. Westerschelde, B. Oosterschelde, C. Lake Grevelingen, D. Delta coast, E. Dutch coast, F. Wadden coast, G. Wadden sea and H. Ems–Dollard. Red lines indicate the outer border of the 12-mile coastal zone, borders between regions mentioned above, or borders between water shed areas of the Scheldt, Meuse, Rhine and Ems. Stippled red line is the outer border of the 1-mile coastal zone.

large number of samples (50) required for the Factor Analysis which is an integral part of m-AMBI (Borja et al., 2008; Hair et al., 1999). Second, the small core samples in the Dutch dataset do not meet the requirements for assignment of an AMBI classification (Borja and Muxica, 2005; Borja et al., 2008; N > 6; S > 3). An alternative MMI for benthos in Dutch waters, i.e. the Benthic Ecosystem Quality Index (BEQI) was developed in earlier years by Van Hoey et al. (2007). The BEQI assesses the status of benthos at three levels. At level 1, total benthic biomass is compared with primary production which are closely related in well mixed and shallow estuaries (Herman et al., 1999). At level 2, the areas of selected important eco-elements, such as intertidal flats and mussel banks, are compared with reference values. At level 3, the assessment is comparable to most North East Atlantic benthos MMIs. A European intercalibration exercise where benthos multi-metrics were compared and harmonized, showed that the BEQI results deviated from those produced by other European metrics (Carletti and Heiskanen, 2009). Furthermore, Borja et al. (2011) reported that the BEQI showed the lowest correlation with human pressures in a broad European study on pressure–impact correlations of MMIs. These findings were confirmed in a review study by Boon et al. (2011).

The results above have led us to design BEQI2 being an MMI for benthos based on the common indicators species richness, Shannon index and AMBI as in the well-established m-AMBI. However, BEQI2 contains a simplified and straightforward calculation method for Ecological Quality Ratios (EQR). In contrast, the m-AMBI uses factor analysis to integrate the indicators which requires a large number of samples and this method was shown to have no apparent benefits (Sigovini et al., 2013). Furthermore, in the BEQI2 software functionality in data processing has been added. These added functions include: (a) small core samples are pooled to a standardized sample area of 0.1 m^2 before data analysis; (b) a multi-metric calculation method is used in which the separate indicator values are first normalized using long-term reference values into an EQR which are then averaged to give one EQR value; (c) species synonym names are automatically converted to their standardized names and (d) optionally a genus is automatically converted to an already identified species within the same sample. The performance of the BEQI2 tool is examined here using benthos data from Rijkswaterstaat over the period 1991–2010 by comparing BEQI2 trends with those in

environmental conditions in different ecotopes, and by correlating BEQI2 with m-AMBI EQRs.

2. Methods

2.1. Data, sampling effort and taxonomic analysis

Around 1990 Rijkswaterstaat (Ministry of Infrastructure and the Environment) started systematic monitoring of the macrobenthos in the Dutch transitional and coastal water bodies. A geographical overview of the water bodies is given in Fig. 1. The principal characteristics of the water bodies and natural and human pressures on the local benthos communities have been listed in Table 1.

For the currently ongoing WFD intercalibration of North-East Atlantic MMIs for marine benthos a relatively simple ecotope classification was agreed discriminating only intertidal/subtidal and mesohaline (salinity 5–18)/polyhaline (salinity 18–30) ecotope classes (Anonymous, 1959) for data analysis. The simplification of the ecotope classification was necessary for the effective comparison in the intercalibration process and to increase the precision of the assessment per ecotope. For the present study we adopt the same classification as for the intercalibration. The sampling strategy and sampling effort in each water body and ecotope are summarized in Table 2.

In transitional and sheltered coastal waters autumn samples, if available, were preferred over spring samples because the organisms are usually slightly more mature in autumn and therefore easier to identify. In open coastal waters however, sampling by Rijkswaterstaat is scheduled in spring to avoid the massive settling of recruits, which can affect abundance dramatically. Consequently only spring benthos monitoring data for the open coast were available.

Benthos sampling for macrofauna was performed using manual cores at the intertidal flats, a flushing sampler in the shallow subtidal areas (0–5 m) and box core sampling in the deeper areas (<− 5 m). The sediment samples taken were sieved over 1 mm in the laboratory. Taxonomic analyses were performed at the species level as far as possible using the Dutch TWN taxa list (http://www.aquo.nl/tools/twn-lijst/), which is based on the WORMS species list (http://www.marinespecies.

Table 1
Dutch saline WFD water bodies studied and their major characteristics and human pressures.

Water body	Major characteristics	Major human pressures
Westerschelde	Estuary with large tidal action and strong salinity variations in the mesohaline part. Shipping channel for the harbour of Antwerpen (Belgium).	Sewage; has decreased strongly in the past two decades due to the implementation of sewage works in Belgium. Widening and deepening of shipping channel, possibly leading to increases of flow velocities and suspended matter in certain parts of this estuary.
Oosterschelde	A formerly open sea arm, which now has an open dam as a protective measure against flooding. Therefore the tidal action is reduced. Large banks of *Crassostrea gigas*.	Limited tidal action due to sluices.
Lake Veere	A salt lake, which was formerly in direct connection with the North Sea; was damned off as a safety measure against flooding.	No tidal action which led to deterioration of water quality in the period 1961–2004. In 2005 a small connection has been made with the Oosterschelde to provide a small tidal action and improve the oxygen and ecological status.
Lake Grevelingen	This formerly open sea arm has been closed for protection of the land against flooding. This has however led to insufficient water mixing and an increasing amount of stratification and areas with oxygen depletion and impaired benthic quality.	No tidal action. Consequently some stratification, especially in summer, which leads to local anoxic zones.
Dollard	Largely mesohaline-intertidal part of the Ems–Dollard estuary. In open connection with the Ems shipping channel.	Increased suspended matter concentrations due to regular dredging and dumping activities in the connected nearby Ems shipping channel
Wadden sea	Polyhaline sheltered (lagoonal) coastal area with tidal mudflats and subtidal gullies. A EU Habitat Directive protected area.	Mussel seed fisheries, shrimp fisheries, manual cockle collection.
Delta coast	Subtidal polyhaline shallow coastal area with gullies and tidal sandbanks. Largely a EU habitat directive protected area.	Small beam and shrimp trawling fisheries.
Dutch coast	Subtidal and polyhaline coastal area, exposed to wave action and highly dynamic.	Small beam and shrimp trawling fisheries.
Wadden coast	Subtidal and polyhaline coastal area, exposed to wave action and highly dynamic. Part of the so-called plaice box (Piet and Rijnsdorp, 1998), which prohibits large beam trawl fisheries.	Small beam and shrimp trawling fisheries.

Table 2
Benthos sampling effort per water body–ecotope in the period 1991–2010.

Water body	Ecotope	Strategy, sample number and size per ecotope	Total sampled area m²/year
Westerschelde	Mesohaline-intertidal	19–75 random samples (0.016 m²)	0.15–1.10
	Mesohaline-subtidal		0.45–0.66
	Polyhaline-intertidal		0.30–1.00
	Polyhaline-subtidal		0.45–0.90
Oosterschelde	Polyhaline-intertidal	30 random samples (0.016 m²)	0.45–1.30
		In 2009 78 samples (0.016 m²)	
		In 2010 80 samples (0.016 m²)	
	Polyhaline-subtidal	90 random samples (0.016 m²)	0.74–1.40
		In 2009 50 samples (0.016 m²)	
		In 2010 46 samples (0.016 m²)	
Lake Veere	Polyhaline-subtidal	40 random samples (0.016 m²) and	1.20–3.50
		20 random samples (0.020 m²).	
Lake Grevelingen	Polyhaline-subtidal	40 random samples (0.016 m²) and	1.00
		20 random samples (0.020 m²)	
Dollard	Mesohaline-intertidal	3 fixed transects each 20 stations with 5 replicate samples (0.009 m²)	1.40–1.60
Wadden Sea inner	Polyhaline-intertidal	6 fixed transects with in total 45 samples (0.06 m²)	1.60–3.00
Wadden Sea coast	Open coast	4–6 box cores (0.068 m² (1991–2000) or 0.078 m² (2001–2010))	0.27–0.31
Delta coast	Open coast		0.27–0.31
Dutch coast	Open coast		0.27–0.47

org/). Representatives of the Oligochaeta, Nemertea, Phoronida were not further identified and are reported at this higher taxonomic level. Following the AMBI rules epifaunal and hard substrate species were removed from the dataset (Borja and Muxika, 2005).

2.2. Data pooling

Because of the sample size requirements of the AMBI indicator, a pooling routine was written to attain a standard sample size of 0.1 m² (Walvoort and Van Loon, 2014; available as an R package from CRAN at http://cran.r-project.org/; R Core Team, 2014). Data pooling was performed with sample data from one specific water body–ecotope–year combination. The tool first randomly selected a sample from the pool of all available samples from a water body–ecotope–year combination, and determined the sample area. The tool then randomly selected a second sample, and calculated the combined sample area. This procedure was repeated until a combined sample area of 0.1 ± 0.01 m² was obtained; and this data pool was set apart. The user can specify other sample area pool ranges. Note that a specific sample was only used once in a single data pooling run. Then the procedure was repeated to form a second data pool etc., until no more data pools could be formed. The effect of random variations of the data pool compositions and the remaining unpooled samples was strongly reduced by repeating this pooling procedure 10 times. As the final step in the procedure, the arithmetic mean indicator values of 10 data pooling runs per water body–ecotope–year combination were stored as the final indicator results. Box core samples with a surface area of 0.068–0.078 m² from the Dutch coastal waters were analysed without pooling because these samples fulfilled the AMBI criteria.

The relationship between the detected species richness of box core samples, and pooled box core subsamples with the same total sample area, was investigated by a simulation experiment using a dataset of 98 box core samples of 0.078 m² taken in 2012 from five regions in the Dutch part of the North Sea: Coastal zone (19 samples), Offshore (28 samples), Frisian front (20 samples), Oyster grounds (24 samples) and Dogger bank (7 samples). For each of these box core samples (0.078 m²) the present individual benthic specimens were randomly split up into 6 smaller subsamples of 0.013 m², under the assumption that benthic specimens are randomly distributed within the box core sample. The resulting set of subsamples was data pooled and assessed within each region using the standard BEQI2 procedure.

2.3. BEQI2 components: species richness, Shannon diversity and AMBI

The BEQI2 contains three indicators: species richness (S), Shannon diversity index (H′) and AMBI. Species richness was calculated as the number of species counted in a single data pool of 0.1 m² or in a box core sample of 0.078 m². Species names are automatically converted by the BEQI2 tool to standardized species name (WORMS). In some cases, a species can only be identified at the genus level due to e.g. a juvenile life stage or damage. If in a specific sample a species has been reported from the reported genus, it is assumed that this genus is probably the already reported species; and the abundance of the genus is added (or proportionally distributed) to the already found species in the sample. By using both conversions an overestimation of the species richness and Shannon index is avoided. The Shannon index, with log2 basis, was calculated using the formula given by Shannon and Weaver (1963). The AZTI Marine Biotic Index (AMBI), which quantifies the sensitivity of benthic species for organic enrichment and associated pollutants using five classes (very sensitive, indifferent, tolerant, opportunistic and very opportunistic species), is calculated using the formula from Borja et al. (2000). The result is a value between 0 (good) and 6 (bad). The AMBI species list from AZTI (ambi.atzi.es version October 2013) was used supplemented with a Dutch list of 88 benthic species not present in the AMBI list or for which Dutch benthic specialists judged that a different AMBI classification was necessary; and their corresponding AMBI classifications (Gittenberger and Van Loon, 2013; Walvoort and Van Loon, 2014). Correlation coefficients between species richness, Shannon index and AMBI were calculated and the median, minimum and maximum values are presented. Since an increasing AMBI value indicates a poorer benthos quality, the signs of the correlations coefficients of species richness and Shannon, respectively, with AMBI are reversed for easier interpretation of these coefficients.

2.4. BEQI2 calculation

For calculation of the BEQI2 first the Ecological Quality ratios (EQR) for each of the three indicators (S, H′ and AMBI) has to be calculated. The EQRs represent normalized indicator values which are obtained by dividing each indicator value by its long-term reference value. The calculation of these reference values was performed for each water body–ecotope combination in view of observed differences of reference values between water body–ecotopes. In this study reference values were estimated using the 99 percentile value for species richness and Shannon index, and the 1 percentile value for AMBI on the basis of

the Rijkswaterstaat benthos dataset over the period 1992–2007. Inspection of the indicator values in this period shows that the 99 (for S and H′) and 1 percentile value (for AMBI) is approx. the highest value which is visibly not an outlier of the indicator value distribution (Borja and Tunberg, 2011). Bad values are usually simple to estimate, and are set for species richness and Shannon diversity at 0. For AMBI, the bad state is commonly defined as containing at least 1 very opportunistic species, corresponding to a bad value of 6 (Borja et al., 2000). Formulas for the EQR of species richness (S), Shannon index (H′) and AMBI are as follows:

$$EQR(S) = S(sample)/S(ref) \tag{1}$$

$$EQR(H') = H'(sample)/H'(ref) \tag{2}$$

$$EQR(AMBI) = (AMBI(bad)-AMBI(sample))/(AMBI(bad)-AMBI(ref)). \tag{3}$$

The EQRs of the indicators were calculated per water body–ecotope combination for every year in the period 1991–2010. We found that the value of the indicator EQR value exceeded 1 in about 1 percent of the cases. The BEQI2-tool does not correct for this, but because BEQI2 is based on the average of the 3 indicator EQR values the value was always in the range 0–1.

Each indicator in the BEQI2 represents an indicator class as demanded by the WFD, as already demonstrated for the m-AMBI (Muxika et al., 2007). Due to lack of complete and reliable indices for the different pressures in the water body–ecotopes, individual EQRs could not be weighted with respect to their sensitivity for pressures. Therefore indicator EQRs were assigned equal weights resulting in the following formula for BEQI2:

$$BEQI2 = (EQR(S) + EQR(H') + EQR(AMBI))/3 \tag{4}$$

For the BEQI2 calculations a tool was written as an R package (Walvoort and Van Loon, 2014, available from CRAN at http://cran.r-project.org/). The BEQI2 results are always calculated at the water body–ecotope–year level. For this purpose, the BEQI2 results of several data pools or box core samples within the respective water body–ecotope–year are averaged arithmetically.

In case the BEQI2 is calculated at the level of a water body and this contains more than one ecotope, the BEQI2 score is calculated as the average of the ecotope BEQI2 scores using the relative ecotope areas as weight factors.

2.5. State and trend assessment

State assessment was performed by calculating per water body–ecotope the arithmetic average of the BEQI2 scores of the last three years (2008–2010), in order to average out natural variations of the benthos composition. The advantage of using a three year period instead of using a WFD six year period is that more recent assessment values are obtained. The benthic status according to BEQI2 is determined by comparing this three-year-average with the Dutch WFD class boundaries for BEQI2. These class boundaries are in transitional waters, the salt lakes and the Wadden Sea: High/Good boundary 0.80, Good/Moderate boundary 0.60, Moderate/Insufficient 0.40, and Insufficient/Bad 0.20. In the Dutch open coastal waters, all these boundaries are set 0.02 lower (e.g. Good/Moderate boundary 0.58), because the AMBI reference value is lower in these water bodies, which leads to on average 0.02 lower BEQI2 scores.

Trend analysis was performed using Mann–Kendall non-parametric trend analysis and the associated Theil–Sen non-parametric slope estimation, using Mystat software (freely available from www.systat.com). A period of 12 years was used, because of the relatively slow

recovery time of benthos (see for more information Section 4.4) and because it corresponds to two WFD periods.

2.6. Pressure–impact validation of BEQI2

Pressure–impact validation for BEQI2 was performed in the Westerschelde mesohaline-intertidal ecotope and the Dollard mesohaline-intertidal ecotope. In the Westerschelde mesohaline-intertidal, oxygen concentrations, which are related to the human pressure sewage input into the Belgium part of the Schelde, were correlated with the BEQI2 values. In the Dollard, BEQI2 was correlated with suspended matter concentrations which are related to dredging and dumping activities in the connected Ems shipping channel. Oxygen data were collected for the subtidal location Schaar van Ouden Doel (Westerschelde, mesohaline part) and suspended matter concentrations were collected for the location Groote Gat Noord (Dollard) from the website http://live.waterbase.nl.

Additionally, the pressure–impact sensitivity of BEQI2 was evaluated by comparison of EQR values with the m-AMBI which has a well-demonstrated sensitivity for human pressure (Muxika et al. 2005, Borja et al., 2008; Borja et al., 2011). As a theoretical framework for pressure–impact validation the DPSIR (Driver Pressure State Impact Response) model was used (Borja et al., 2006a,b). In this model, a Driver (human activity, e.g. households) leads to a Pressure (untreated sewage), an impaired oxygen State, Impact on the benthos survival and a management Response (implementation of sewage works). The results of the BEQI2 and m-AMBI were compared using identical benthos data sets and reference values from the Westerschelde (450 data pools of 0.1 m², all ecotopes and years) and the open coastal water bodies Delta coast, Dutch coast, Wadden coast (20 annual averages per coastal body) and the Wadden Sea (polyhaline-intertidal ecotope; 20 annual averages). The m-AMBI tool was downloaded from http://ambi.azti.es.

2.7. BEQI2 validation through expert judgement

After calculation of the BEQI2 values for the different water bodies, a panel of macrobenthos experts was requested to give their expert opinion on the benthic ecological state (Good, Moderate etc.) of respectively the Wadden Sea (polyhaline-intertidal ecotope), the Dollard (mesohaline-intertidal ecotope), and the Westerschelde (averaged for all four ecotopes) for the period 2008–2010. The panel consisted of two regional experts for the northern region viz. Rob Dekker (NIOZ, NL) and Gabi Petri (NLWKN, DE) and an expert for the southern region (Dick de Jong-RWS, NL). Each of these experts has for a long period been involved in benthic monitoring and assessment of their respective areas (e.g. De Jong, 2000; Dekker and Beukema, 2014; Grage et al., 2013). Their expert opinion was inquired without them having prior knowledge of the BEQI2 scores obtained for these water bodies–ecotopes. For the Dollard, the expert judgement was based on the assessment report of Grage et al. (2013, in German). The expert opinion or judgement on the benthic ecological state of these selected water bodies was compared with the ecological state as determined using the BEQI2 scores and the appropriate class boundaries (see Section 2.5). Note that only expert opinion on the benthic ecological state was inquired, and not on the more precise BEQI2 scores.

3. Results

3.1. Effect of data pooling

The effect of the number of data pooling repetitions on the precision of the indicator results for species richness, Shannon index and AMBI is illustrated in Fig. 2. This figure shows that for all indicators, the relative deviation of the pooling average after 10 pool runs from the asymptotic value is less than 1% which is considered an adequately precise result.
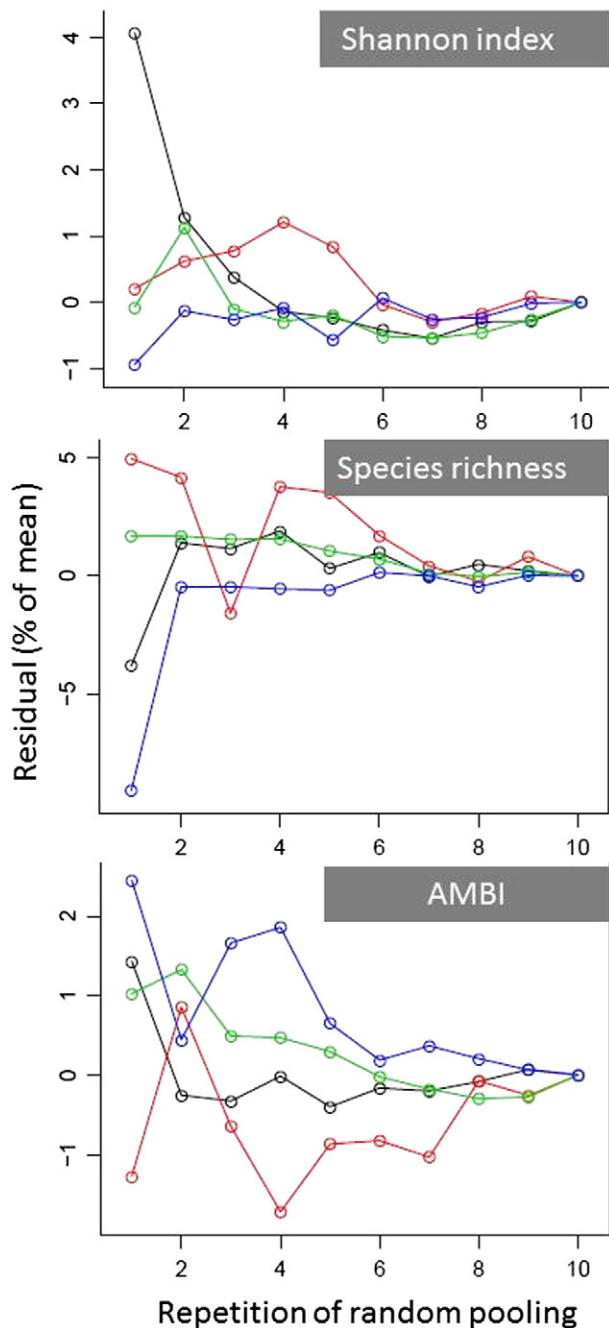
Fig. 2. Residual deviation (%) of reference values – from the reference value obtained using 10 repetitions – with increasing number of repetitions of the pooling procedure for the four Westerschelde ecotopes using data from the period 1991–2000. Black line is mesohaline-intertidal ecotope; red line is mesohaline-subtidal ecotope; green line is polyhaline-intertidal ecotope; blue line is polyhaline-subtidal ecotope.

Therefore, the data pooling procedure has been set at 10 repetitions for standardized BEQI2 data pooling. The variation of the BEQI2 values as a result of data pooling variation appeared to be on average <0.01 of the EQR.

The fraction of samples that meet the AMBI criteria in the Westerschelde ecotopes and the Dollard before and after data pooling to 0.1 m^2 are shown in Table 3. This shows that in the Westerschelde polyhaline ecotopes only 9–12% and in the Westerschelde mesohaline ecotopes 67–72% of the samples fulfilled all 3 AMBI criteria. After data pooling, 83–93% of the polyhaline ecotope samples and 85–92% of the mesohaline ecotope samples meet the AMBI criteria. Pooling improved

the applicability of these Westerschelde samples for AMBI assessment to an average of 88%. The Dollard samples, in which five replicate samples of 0.009 m^2 were first pooled, already met with the AMBI criteria. However, also in this case pooling data of two smaller samples to a standardized area of 0.1 m^2 is useful in order to obtain reference values comparable to other water bodies, and to make a correction of the Dollard reference values possible (see below).

The relationship between the average species richness detected and the sampling size and number used (1 box core of 0.078 m^2 versus 6 box core subsamples of 0.013 m^2 pooled) for several North Sea regions in 2012, as determined via a simulation experiment is presented, in Fig. 3. The use of real box core benthic data ensures that a fairly realistic simulation of the benthic variation is obtained. This figure strikingly shows that the use of 6 small box core subsamples detects a higher species richness compared to 1 box core of the same total area. Generally these results mean that the use of more samples with a smaller sampling area is better capable of detecting the species heterogeneity of a region compared to less larger samples of the same total sampling area. Therefore, data pooling has to be performed on samples of comparable small sampling area, in order to give data pools of 0.1 m^2 with a comparable species richness. In addition, reference values have to be determined for benthos data from the specific water body–ecotope monitored using comparable sample areas. Both these conditions are met in this study, because fixed sample areas were used for each water body–ecotope during the entire period studied. It is estimated that for the Shannon index and AMBI this effect will be much smaller, since these indicators are probably less sensitive for the sample area used.

3.2. BEQI2 reference values for the period 1992–2007

The calculated 99 (for species richness, Shannon) or 1 percentile (for AMBI) reference values are listed in Table 4. These reference values are discussed in Section 4.2 per water body–ecotope. In general, the BEQI2 values calculated with these reference values give a good indication of the benthic quality status of the different Dutch transitional and coastal waters according to the consulted Dutch and German regional benthic experts. The only exception is the Dollard mesohaline-intertidal ecotope. In this particular case the initial BEQI2 value indicated a good benthic state (BEQI2 > 0.6). However, the regional benthic experts for the Dollard region indicated that the Dollard is in a moderate benthic state (Grage et al., 2013). This discrepancy is explained by the fact that reference values were computed over a period that the Dollard has been under continuous stress of elevated concentrations of suspended particulate matter (van Loon et al., 2011). Hence these references are too low and the resulting BEQI2 values are too high. Therefore, a procedure was set-up to correct the reference values in the northern Dollard. This was only possible by using reference values from comparable ecotopes and sample sizes in the southern Delta region and adapting these for regional differences.

The sample sizes of the small core samples in the Dollard, Wadden Sea and Delta region are not identical (see Table 2). However, due to the use of relatively small sample areas and a certain amount of spatial sample distribution in all three ecotopes, we estimate that the reference values for species richness obtained for these ecotopes are reasonably comparable. However, note that the reference values obtained for species richness for the large box core samples in the coastal zone and the data pooled small samples in the transitional waters and salt lakes are not comparable due to both a difference in sample area (0.078 m^2 box core versus 0.1 m^2 pooled) and sample numbers (1 versus 6–10). It is estimated that for Shannon index and AMBI the reference values will be comparable in all these cases, because the dependence on sample sizes and numbers is estimated to be much smaller. This difference can also be observed in the Regional Difference Factors for species richness and Shannon below.

Table 3

Fraction of samples that meet the AMBI criteria in the Westerschelde ecotopes (average sample area 0.017 m^2) and the Dollard (sample area 0.045 m^2), before and after data pooling to 0.1 m^2.

Criterion	n ≥ 6		S ≥ 3		AMBI$_{classified}$ ≥ 20%		Overall	
	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled	Unpooled	Pooled
Westerschelde								
Mesohaline-intertidal	0.94	0.98	0.91	0.96	0.74	0.94	0.72	0.92
Mesohaline-subtidal	0.83	0.92	0.87	0.91	0.68	0.88	0.67	0.85
Polyhaline-intertidal	0.41	1.00	0.43	1.00	0.14	0.93	0.12	0.93
Polyhaline-subtidal	0.28	0.99	0.38	0.99	0.10	0.83	0.09	0.83
Dollard								
Mesohaline-intertidal	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00

The correction procedure for species richness (S) and Shannon index involves three steps which are explained for S. In the first step the ratio was calculated between reference S values belonging to the polyhaline intertidal ecotopes in the Wadden Sea and those in the Oosterschelde. These ecotopes have similar salinity (29) which makes it possible to compare their reference values directly. Their ratio is termed the apparent Regional Difference Factor (RDF) of reference S values between the northern Wadden/Dollard region and the southern Delta region and represents differences due to all factors other than salinity and depth, e.g. productivity, grain size. In the second step, an estimate was made of a regional Delta reference S value for a salinity of 15 – as in the Dollard – by linear interpolation between the reference S values belonging to the mesohaline intertidal ecotope (average salinity 7.7) and to the polyhaline intertidal ecotope (average salinity 22.7) in the Westerschelde. Finally, this interpolated reference S value for salinity 15 in the Delta region was multiplied with the RDF obtained in step 1 to obtain an estimate of the reference S value for the mesohaline-intertidal ecotope (salinity = 15) in the Wadden/Dollard region. The same procedure was used to obtain a reference value of the Shannon index in the Dollard mesohaline-intertidal ecotope. The AMBI reference value of the Westerschelde mesohaline-intertidal ecotope, viz. 0.57, was directly applied to the Dollard, i.e. without an RDF, since it was assumed to be little dependent on the species richness and diversity differences between these two regions, and comparable for the similar ecotope in both regions.

The correction procedure for the reference values of the Dollard mesohaline-intertidal ecotope, as described above, resulted in a regional difference factor (North/South) of 0.64 for S and of 0.97 for H′ for the polyhaline ecotopes in the Wadden/Dollard region and Delta region, respectively. In Table 4, the corrected reference values for S and H′ are given, using both the salinity calibration model for the Delta region and the regional difference factors for S and H′ for the Wadden/Dollard and Delta region, respectively.

The median correlation coefficient between species richness and Shannon index is 0.59 (minimum −0.09, maximum 0.80) which shows as expected a fairly good correlation between these two diversity indicators. Note that Shannon index is used in BEQI2 to account for the WFD demand to incorporate (relative) abundances in the benthos multi-metric. The median correlation coefficient between species richness and AMBI is 0.1 (minimum −0.31, maximum 0.56). The median correlation coefficient between Shannon index and AMBI is 0.1 (minimum −0.28, maximum 0.51). These low correlation coefficients show that the AMBI provides a different and complementary type of benthic information compared to the diversity indicators species richness and Shannon index.

3.3. BEQI2 validation

The sensitivity of BEQI2 for human and natural stressors was explored by regression analysis of regional BEQI2 and time-series of measurements of dissolved oxygen (O$_2$) in the Westerschelde mesohaline-intertidal ecotope and of the suspended matter (SPM) concentrations in the Dollard mesohaline-intertidal ecotope. Fig. 4A shows that O$_2$ in the Westerschelde mesohaline-intertidal ecotope showed a strong increase in the period 1991–2010 from an annual average of ~3 mg/L to ~8 mg/L. The very significant relation
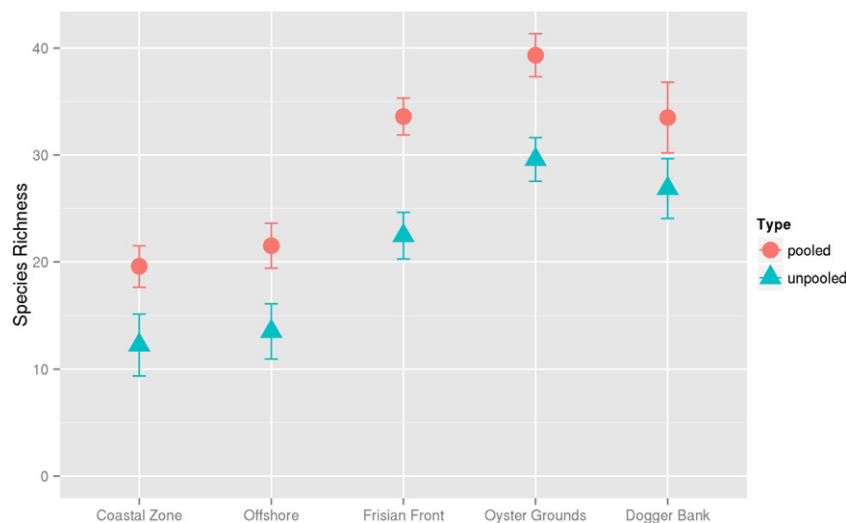
Fig. 3. Simulation results for the average species richness determined for the Dutch marine areas Coastal zone, Offshore, Frisian front, Oyster banks and Dogger bank using 98 original box core samples (0.078 m^2) from 2012; and simulated smaller samples (0.013 m^2) constructed from these box core samples. These smaller samples were pooled (6 samples per data pool) using the BEQI2 method, leading to an equal sample area of the data pools and the box cores. The whiskers indicate the 95% confidence intervals of the average indicator values per area.

Table 4
Estimated references values for S, H′ (99 percentile values) and AMBI (1 percentile values). These percentile values were calculated using MWTL data from 1992–2007. Salinity is also reported as an important state parameter which co-determines the reference values.

Water body	Ecotope	Salinity (average)	Ref (S) (99 perc.)	Ref (H) (99 perc.)	Ref. (AMBI) (1 perc.)
Westerschelde	Mesohaline-Intertidal	7.7	29	3.3	0.57
Westerschelde	Mesohaline-Subtidal	8.9	22	3.2	0.54
Westerschelde	Polyhaline-Intertidal	23	41	3.6	1.2
Westerschelde	Polyhaline-Subtidal	23	31	3.8	0.63
Oosterschelde	Polyhaline-Intertidal	29	45	3.7	0.54
Oosterschelde	Polyhaline-Subtidal	29	67	5.1	0.50
Lake Veere	Polyhaline-Subtidal	22	30	3.8	0.52
Lake Grevelingen	Polyhaline-Subtidal	29	44	4.2	0.54
Dollard	Mesohaline-Intertidal	15	22 (17)[a]	3.3 (2.5)[a]	0.54
Wadden sea	Polyhaline-Intertidal	24	29	3.6	0.54
Wadden sea	Polyhaline-Subtidal	24	23	3.5	0.50
Delta coast	Polyhaline-Subtidal	31	33	3.8	0.009
Dutch coast	Polyhaline-Subtidal	30	30	4.0	0.16
Wadden coast	Euhaline-Subtidal	31	30	3.7	0.12

[a] The uncorrected values are between brackets (Van Loon and Verschoor, 2012).

(p < 0.000) between O_2 and the BEQI2 score is shown in Fig. 4C. The slope of this relation is 0.014 [EQR/mg O_2], showing that an increase in oxygen concentration leads to a higher BEQI2 EQR. The development of the average SPM concentrations in the Dollard mesohaline-intertidal ecotope is shown in Fig. 4B illustrating a strong increase in the period 1991–2010 from an annual average of approx. 100 mg/L in 1991 to an annual average of approx. 200 mg/L in 2010. The very significant relation (p < 0.000) between the SPM concentration and the BEQI2 score is shown in Fig.4D. The slope of

this relation is −0.001 [EQR/mg SPM], showing that a higher SPM concentration leads to a lower BEQI2 EQR.

The results of a quantitative comparison between BEQI2 and m-AMBI for five Dutch marine water bodies are shown in Fig. 5 and Table 5. The correlation coefficients between the BEQI2 and m-AMBI are very high (Pearson r = 0.997–0.999), and the average quantitative difference between the BEQI2 and m-AMBI scores is very low, i.e. in all cases less than 0.01. The match between the two MMIs holds for single box core samples as well as for pooled data.
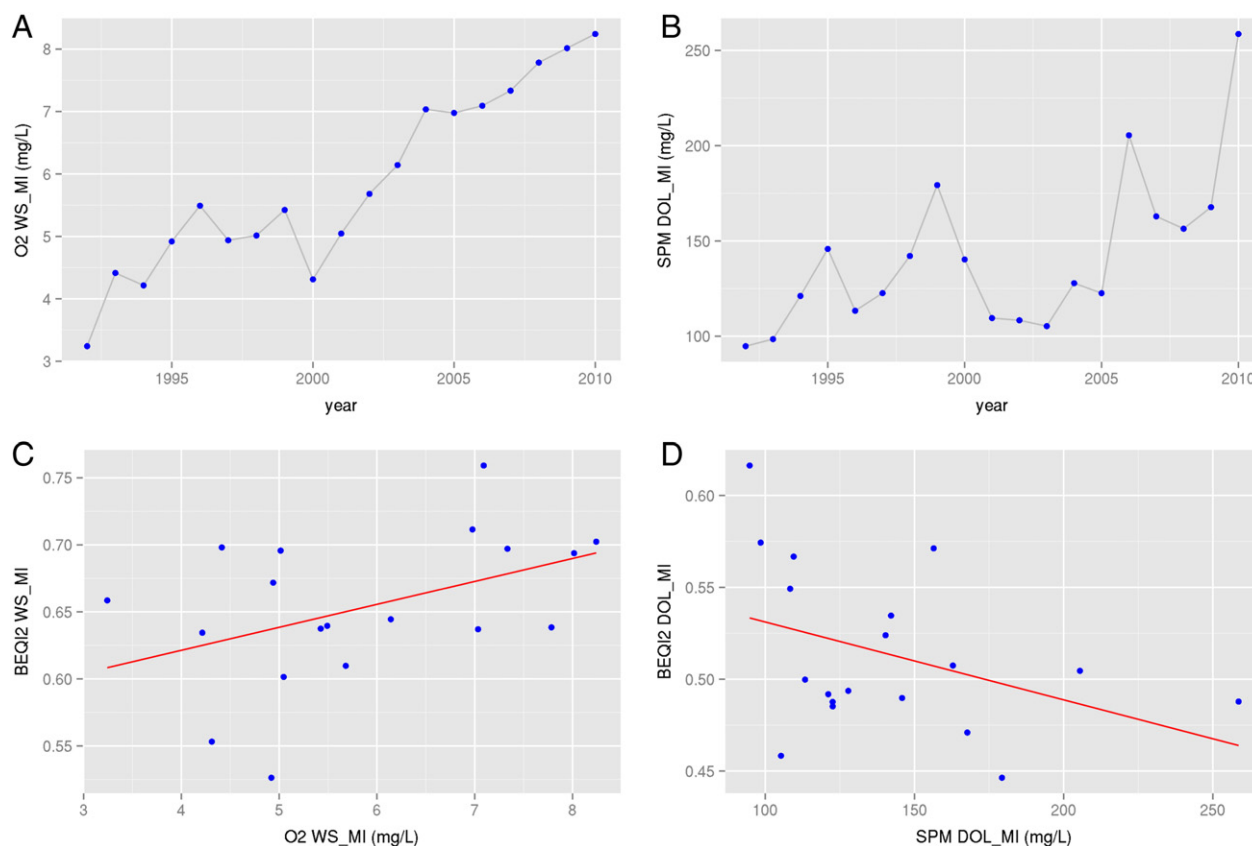
Fig. 4. A–B. Time trends of the state parameters oxygen (slope = 0.32; p < 0.000) and suspended matter (slope = 7.2; p < 0.000) in the water body–ecotopes Westerschelde mesohaline-intertidal (WS_MI) and Dollard mesohaline-intertidal (DOL_MI), respectively. C–D. State–impact correlations for oxygen concentrations (slope = 0.014; p < 0.000) and suspended matter (slope = −0.001; p < 0.000) with BEQI2 EQRs in the water body–ecotopes Westerschelde mesohaline-intertidal (WS_MI) and Dollard mesohaline-intertidal (DOL_MI), respectively.
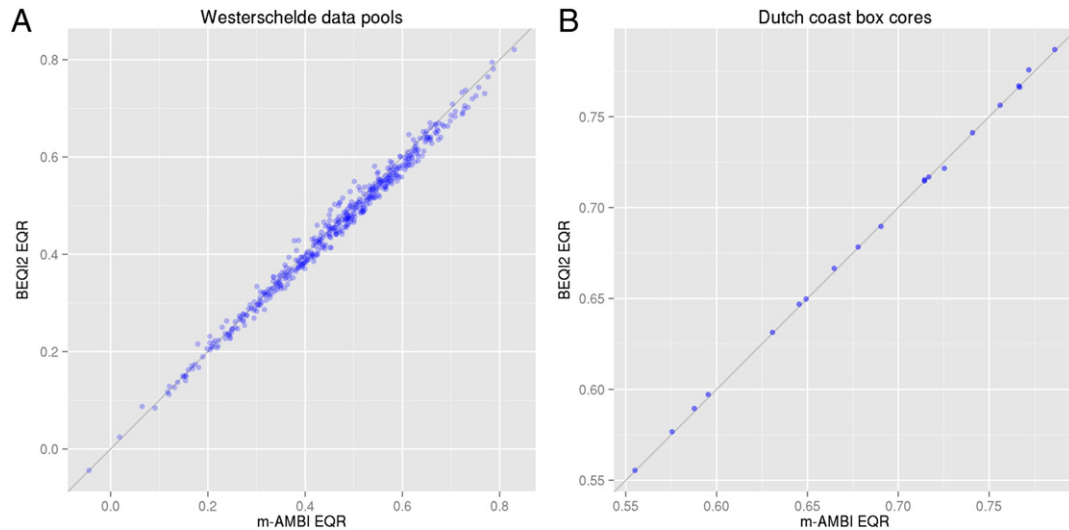
Fig. 5. Illustration of the quantitative correlations of BEQI2 and m-AMBI scores for data pools (Fig. 5A; water body Westerschelde) and box core samples (Fig. 5B; water body Dutch coast).

These results additionally show that the pressure sensitivity of BEQI2 is similar to that of m-AMBI.

3.4. State and trend assessments

The complete set of BEQI2 scores, state assessments and trend assessments are given in Table 6.

The scores show that nine water body–ecotopes have a good BEQI2 status and four water body–ecotopes have a moderate BEQI2 status. The latter are the Westerschelde polyhaline-subtidal, the Lake Grevelingen, the Dollard mesohaline-intertidal, and the Delta coast. The trends range from −0.019 to +0.018 EQR/y. The four significant BEQI2 trends (p < 0.05) are shown in Fig. 6A–D. Positive BEQI2 trends over the past 12 years were observed in the Westerschelde mesohaline-intertidal, Westerschelde polyhaline-intertidal and Oosterschelde polyhaline-intertidal, respectively. One negative BEQI2 trend is observed in the Dutch coast polyhaline-subtidal.

4. Discussion

4.1. BEQI2 calculation and reference values

Ecological quality under the WFD is expressed as a normalized ecological quality ratio (EQR), on a scale of 0 to 1. For normalization of these EQRs it is necessary to define the end points of this scale, i.e. the reference state value and bad state value per indicator. Clearly, the selection of reference and bad state values per indicator is crucial for obtaining realistic EQR values (Borja et al., 2011). There are in general four methods to obtain reference values (OSPAR, 2011): (a) the use of historical benthos data from periods without significant human pressures (e.g. around 1850) to determine reference state values; (b) the use of benthic data from comparable areas which are in a pristine reference state; (c) the use of the highest indicator value in a sufficiently large dataset (e.g. 10–15 years) which is not a statistical outlier (e.g. the 99 percentile

value) to represent the average benthic quality in the reference state (Borja and Tunberg, 2011; Borja et al., 2011); (d) the use of models or theoretical considerations to define reference and bad state values. The use of historical data is usually not possible as in the present study, because historical records are not available or if present, often of insufficient quality. Furthermore, benthos monitoring data from comparable areas in a pristine reference state are usually very hard to find. In this study reference state values were estimated using method c). The use of 99 to 95 percentile indicator values of a large benthos dataset to represent the average benthic quality in the reference state is a commonly used method in the WFD assessments of marine benthos (Borja and Tunberg, 2011; Borja et al., 2011).

Expert judgement is commonly used to check the plausibility of EQR and ecological state results obtained using specific reference and bad state values, and to check geographical distribution and consistency (e.g. Muxika et al., 2007). Weisberg et al. (2008) reported that expert judgement is a valid procedure to evaluate the benthic status of marine areas using benthos data, and may be used to evaluate the results of indicators and MMIs within WFD context. In the present study, the EQR outcomes showed good agreement with expert judgement or opinion in all but one case, viz. the mesohaline-intertidal ecotope in the Dollard. Hence a correction was applied to the Dollard references using data from another tidal estuary in the south of the Netherlands, the Westerschelde (Fig. 1).

In the Westerschelde, the reference values for species richness (S) for polyhaline ecotopes are higher than for mesohaline ecotopes. This well-known phenomenon (e.g. Remane, 1934) is explained by the more stable salinity conditions in polyhaline ecotopes compared to mesohaline ecotopes where strong salinity variations reduce the number of species. According to Remane (1934) the number of species in the mesohaline and polyhaline region is approximately linearly related to the salinity. This allowed us to estimate S for a salinity 16 in the Westerschelde. By interpolation and through multiplication with the RDF for S between the Wadden Sea and Westerschelde the corresponding

Table 5

Quantitative comparison of the BEQI2 and m-AMBI EQR values for a selection of Dutch marine water bodies. Data were annual averages of the sample values from the period 1991–2010.

Water body	Data aggregation	Correlation Pearson r	Average difference BEQI2 − m-AMBI
Westerschelde, all ecotopes	Pooling	0.997	− 0.008
Wadden Sea, polyhaline-intertidal	Pooling	0.997	+ 0.001
Dutch coast, polyhaline-subtidal	No pooling (box core)	0.999	+ 0.0009
Delta coast, polyhaline-subtidal	No pooling (box core)	0.998	− 0.006
Wadden coast, euhaline-subtidal	No pooling (box core)	0.998	− 0.002

Table 6
BEQI2 EQR values for all Dutch marine WFD water bodies for the period 1992–2010. Moderate status values (average of last 3 years) are printed fat. Significant trends (last 12 years) are underlined. The codes presented contain first a water body code and then an ecotope code (e.g. WS_MI). Water body codes used are: WS = Westerschelde, OS = Oosterschelde, LV = Lake Veere, LG = Lake Grevelingen, DOL = Dollard, WAD = Wadden sea, DEC = Delta coast, DUC = Dutch coast, WAC = Wadden coast. Ecotope codes used are: MI = mesohaline-intertidal, MS = mesohaline-subtidal, PI = polyhaline-intertidal, PS = polyhaline-subtidal. Oxygen and SPM concentrations (mg/L) are given for WS_MI and DOL_MI. TS-slope = Theil–Sen slope. MK-p = significance of Mann–Kendall trend test.

Year	Oxygen WS_MI	WS_MI	WS_MS	WS_PI	WS_PS	OS_PI	OS_PS	LV_PS	LG_PS	Susp.matter DOL_MI	DOL_MI	WAD_PI	DEC_PS	DUC_PS	WAC_PS
1991	3.55									102	0.60	0.69	0.60	0.58	0.65
1992	3.24	0.66	0.59	0.68	0.65	0.65	0.69	0.63	0.67	95	0.62	0.71	0.58	NA	0.68
1993	4.41	0.70	0.63	0.69	0.64	0.61	0.79	0.76	0.60	98	0.57	0.72	0.61	0.64	0.75
1994	4.21	0.63	0.53	0.69	0.49	0.61	0.68	0.69	0.60	121	0.49	0.64	0.60	0.70	0.73
1995	4.92	0.53	0.46	0.56	0.47	0.54	0.67	0.62	0.62	146	0.49	0.62	0.57	0.64	0.74
1996	5.49	0.64	0.51	0.65	0.61	0.56	0.68	0.60	0.66	113	0.50	0.57	0.42	0.55	0.45
1997	4.94	0.67	0.63	0.64	0.57	0.53	0.69	0.68	0.63	123	0.49	0.51	0.55	0.54	0.55
1998	5.01	0.70	0.68	0.65	0.59	0.56	0.67	0.72	0.62	142	0.53	0.57	0.55	0.68	0.76
1999	5.42	0.64	0.52	0.62	0.54	0.63	0.76	0.65	0.63	179	0.45	0.62	0.50	0.69	0.71
2000	4.31	0.55	0.54	0.65	0.65	0.65	0.74	0.66	0.57	140	0.52	0.63	0.61	0.63	0.75
2001	5.05	0.60	0.71	0.55	0.52	0.47	0.71	0.67	0.55	110	0.57	0.60	0.60	0.75	0.74
2002	5.68	0.61	0.48	0.59	0.61	0.47	0.68	0.72	0.54	108	0.55	0.59	0.59	0.73	0.69
2003	6.14	0.64	0.41	0.70	0.59	0.52	0.65	0.60	0.59	105	0.46	0.56	0.61	0.71	0.63
2004	7.03	0.64	0.42	0.68	0.48	0.58	0.67	0.61	0.68	128	0.49	0.60	0.58	0.73	0.64
2005	6.98	0.71	0.57	0.71	0.61	0.61	0.65	0.74	0.51	123	0.49	0.62	0.63	0.69	0.64
2006	7.09	0.76	0.63	0.67	0.54	0.51	0.64	0.61	0.51	205	0.50	0.50	0.61	0.65	0.67
2007	7.33	0.70	0.82	0.78	0.63	0.61	0.79	0.75	0.65	163	0.51	0.64	0.55	0.61	0.70
2008	7.78	0.64	0.67	0.60	0.47	0.65	0.74	0.58	0.55	156	0.57	0.59	0.61	0.60	0.65
2009	8.02	0.69	0.64	0.73	0.59	0.71	0.68	0.73	0.53	168	0.47	0.63	0.56	0.56	0.70
2010	8.24	0.70	0.53	0.69	0.51	0.68	0.62	0.61	0.60	259	0.49	0.59	0.54	0.58	0.67
3 year average	8.01	0.68	0.61	0.67	**0.52**	0.68	0.68	0.64	**0.56**	194	**0.51**	0.60	**0.57**	0.58	0.67
Trend period	90–10	99–10	99–10	99–10	99–10	99–10	99–10	99–10	99–10	90–10	99–10	99–10	99–10	99–10	99–10
TS-slope	0.320	0.011	0.014	0.010	−0.003	0.018	−0.008	0.001	−0.003	7.240	0.000	0.000	−0.002	−0.019	−0.004
MK-p	0.000	0.010	0.146	0.040	0.179	0.054	0.071	0.509	0.353	0.071	0.456	0.509	0.353	0.001	0.179

reference S value in the Dollard was obtained. The RDF of 0.64 that we used for this purpose denotes that reference S values in the Wadden Sea are lower than in the comparable ecotopes in the Westerschelde. These regional differences may be caused by (a) small systematic differences in temperature (slightly colder Wadden Sea); and (b) a potentially larger input of exotic species in the Westerschelde due to shipping traffic and/or proximity to the Atlantic fauna in the Channel.
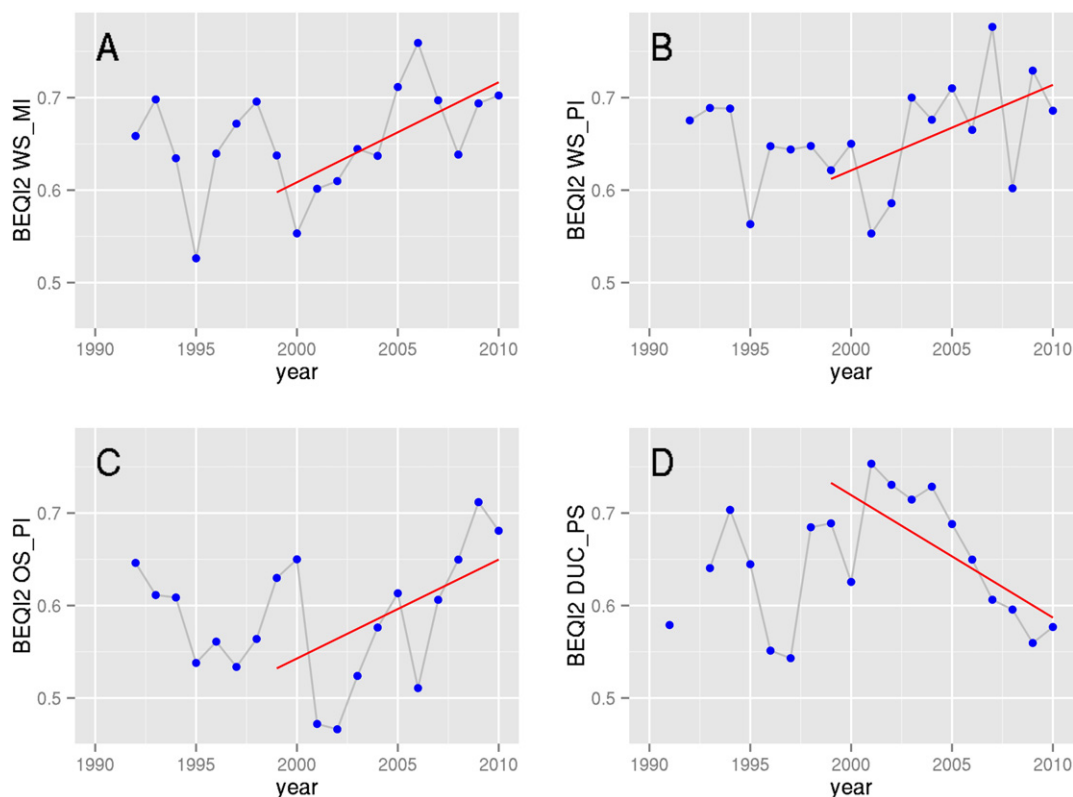
Fig. 6. A–D. BEQI2 time trends in the water body–ecotopes which show significant BEQI2 trends in the period 1999–2010: A. Westerschelde mesohaline-intertidal (WS_MI); B. Westerschelde polyhaline-intertidal (WS_PI); C. Oosterschelde polyhaline-intertidal (OS_PI); and D. Dutch coast polyhaline-subtidal (DUC_PS).

Similarly as for S, Shannon index (H′) reference values in the Westerschelde polyhaline ecotopes are also higher than for the mesohaline zone, but this effect is numerically less pronounced probably because other environmental factors than salinity also play a role in determining H′ such as enrichment, disturbance frequency (Widdicombe and Austen, 2001), heterogeneity and steepness of gradients (e.g. Reizopoulou et al., 2014). Moreover, Shannon index is particularly sensitive to rare species (Peet, 1974). The calculated 99 percentile reference values for H′ in the Westerschelde and Wadden Sea appear to be fairly well comparable for the corresponding intertidal and subtidal ecotopes with the same salinity, resulting in an RDF for H′ of 0.97.

For AMBI the 1 percentile values ranged from 0.50 to 1.2 for the Westerschelde ecotopes, from 0.50 to 0.54 for the Wadden Sea ecotopes, and from 0.009 to 0.16 for the open coastal waters. Overall, the dynamic saline and/or wet/dry conditions in transitional water and sheltered coastal waters, which pose a certain amount of natural pressure on the benthos, limit the AMBI reference value at a minimum value of 0.5. The very low AMBI reference conditions in the open coastal waters can be explained by the relatively stable physico-chemical conditions, with constant submersion and a stable salinity.

We found a mismatch between calculated state of a benthic ecosystem and expert judgement only with respect to the Dollard. This estuary has an almost century long history of negative human pressures starting with discharge of waste water from the pulp industry to recent dredging and deepening for shipping causing enhanced turbidity (De Jonge et al., 2014). Reference values for the Dollard calculated over this period were therefore near or at the lower end of their potential range, and hence the BEQI2 values become too high. In conclusion, reference values should be taken from a period in which the benthic quality of a water body is near the Good/Moderate class boundary, or even higher, in order to obtain realistic estimates of the reference state with the 99 percentile method.

4.2. BEQI2 pressure–impact validation

BEQI2 showed a significant relation with the concentrations of dissolved oxygen in the Westerschelde and those of SPM in the Dollard over the past 19 years (Fig. 4C,D). The strong improvement of the oxygenation in the Westerschelde mesohaline-intertidal is attributed to the implementation of sewage works in the Belgian part of the Schelde basin in the past decades (Maes et al., 2007). The removal of organic matter by these sewage works has led to strong reduction of the organic load of the river water flowing into this ecotope, and consequently to an increase of the oxygen concentrations. Gray et al. (2002) reported that growth of benthic organisms is affected when oxygen concentrations drop below 6 mg/L, that metabolism is affected below 4 mg/L O_2, and that mortality occurs below 2 mg/LO_2. Although Fig. 4C displays a statistically significant relation there is noticeable scatter. Scatter could arise from different factors such as replacement of hypoxia sensitive by more tolerant species (Levin et al., 2009) thereby compensating losses of species and diversity. Furthermore body size and life traits (bioirrigation, Van Colen et al., 2012) present in the community determine the state of sediment oxygenation. Also, wind plays an important role for instance by intermittent oxygenation and benthos survival despite overall hypoxic conditions. Nevertheless considering published limits and ample studies on hypoxia, in combination with the increase in oxygen concentration over the 19 year period from 3 to 8 mg/L (Fig. 4A), the trend in BEQI2 in the Westerschelde mesohaline intertidal (Fig. 6A) and the significant relation with oxygen (Fig. 4C) are explicable.

The increase of suspended matter (SPM) concentrations in the Dollard in the past two decades, especially between 2005–2010 (Fig. 4B), is most likely caused by increasing dredging and dumping activities in the Ems–Dollard in order to maintain and deepen the Ems shipping channel (De Jonge et al., 2014). This increased SPM constitutes a pressure on filter- and deposit-feeding benthos because of smothering of benthic species and impairment of their food

quality (Essink, 1999, Gittenberger and Van Loon, 2013) but apparently did not negatively affect the annual average oxygen concentration in the Dollard (8.4–9.2 mg/L).

4.3. BEQI-2 correlation with m-AMBI

The BEQI2 method has been introduced in the NEAGIG intercalibration process in 2011 in a report format and was published on the internet for intercalibration use (Van Loon et al., 2011). More recently, Sigovini et al. (2013) independently developed and published a similar simplified m-AMBI calculation method using normalized indicator values. They also reported excellent correlation coefficients between the so-called normalized m-AMBI and original m-AMBI using several benthos datasets. Sigovini et al. (2013) emphasize the importance of transparency and simplicity in the calculation method of MMIs. This supports the simplified BEQI2 normalization method presented in this paper.

The high quantitative comparability between the BEQI2 and m-AMBI multi-metrics depicted in Fig. 5, shows that the difference between the two calibration methods, when using identical data sets and reference values, may be considered negligible. In view of this result, the well-established human pressure sensitivity of the m-AMBI is most likely also a characteristic of the BEQI2. This result supports the pressure sensitivity of BEQI2 demonstrated in Section 4.2.

4.4. State and trend assessments based on BEQI2

State assessment of all WFD quality elements is in the Netherlands standardly performed using the three-year-average of the most recent annual average BEQI2 scores, in order to average out year-to-year variations, and the results are presented in Table 6.

For the Westerschelde polyhaline-subtidal ecotope, the observed moderate state (BEQI2 = 0.52) can presumably be assigned to the increased penetration of the tidal wave causing increased flow and height due to widening and deepening of the shipping channel to the harbour of Antwerp (De Vriend et al., 2011). In the last century the Mean High Water Level (MHW) of the station Bath (near the Belgian–Dutch border) in the Westerschelde has increased by some 0.50 m. Although much effort is invested by the regional water authorities to maintain and reduce flow pressures, it remains a point of water management attention. The moderate benthic quality of Lake Grevelingen (BEQI2 = 0.56) is an acknowledged water quality problem which is caused by the gradually increasing stratification of this lake, especially in summer, since it has been closed of from the sea to protect the region from flooding by the sea. Due to this increasing stratification, oxygen depletion occurs in certain parts of this lake in summer (Seitaj et al., 2013). A recent Dutch report on the benthic status of Lake Grevelingen (Wanink, 2014) using the BEQI2 tool shows that the BEQI2 scores in the depth stratum <− 10 m are moderate and a significantly decreasing BEQI2 time trend is observed in this zone. The currently moderate benthic quality of the Dollard originates from the large organic pollution which started in 1969 caused by large discharges of unpurified beetsugar mill waste water directly into the Dollard (Essink and Beukema, 1986). These discharges diminished considerably from 1974 onwards but a new human pressure on the Dollard benthos appeared in the form of gradually increasing amounts of suspended matter caused by dredging of the Ems shipping channel and dumping of sludge near or in the Dollard (De Jonge et al., 2014). This new pressure, possibly combined with insufficient restoration time from the organic load pressure in the sixties/seventies (Van Colen et al., 2012), seems to have prevented the restoration of the benthos in the Dollard to a Good state (Table 6). The benthic states of the Delta and Dutch coast are just moderate and on the border of a Good and Moderate, i.e. BEQI2 score of 0.57 and 0.58, respectively. Most likely the relatively high intensity of small beam trawling and shrimp fisheries is the cause of this situation (ICES, 2014, see below).

Trend analysis is an important second step in benthic quality assessment to determine if a Good or Moderate state is developing i.e. improving or deteriorating. It is well known that damage to benthic communities may occur rapidly (within one or a few years), especially if sand extraction/dredging or sand nourishment/dumping has occurred (Van Dalfsen and Essink, 2001). In contrast, recovery of benthic communities may take longer than 8 years for fragile and sensitive species (Kaiser et al., 2006) and even >15 years as reported by Hering et al. (2010). Therefore, a trend analysis period of 12 years (corresponding to the duration of 2 WFD assessment periods) was regarded to be the most suitable to detect possible recovery of benthic communities. In four regions in the Dutch WFD marine water bodies–ecotopes significant trends over a 12 year period were found (Table 6, Fig. 6). The positive trend of BEQI2 scores in the Westerschelde mesohaline-intertidal and polyhaline-intertidal ecotopes shows that there are also positive developments in the benthic quality of the Westerschelde estuary. These intertidal areas are probably less affected by hydrodynamic pressures and fisheries, and benefit from the increasing oxygen status of the Westerschelde, especially in the mesohaline part. The BEQI2 scores in the Oosterschelde are also improving, which matches with the general view of regional water managers that the ecological quality of the Oosterschelde is good. This is for a large part caused by the controlled open connection of the Oosterschelde with the sea, which gives a very good oxygen status and water mixing but without large flow velocities resulting in lower turbidity.

For the Dutch coast a significantly decreasing BEQI2 trend was found suggesting that this water body shifts into a moderate benthic status. This negative trend is possibly caused by the intensive small beam trawling and shrimp fisheries in the coastal zone which has been categorized as being "ploughed and raked" habitat by Lindeboom et al. (2008). Chronic trawling has been shown to depress species richness as well as biomass and productivity of benthic communities (e.g. Hiddink et al., 2006; Tillin et al., 2006). These fundamental impacts and the intensity of trawling in Dutch waters, call for better insight in the pressure–impact relations of different types of fisheries (e.g. shrimp, sumwing, electric) on indicators and MMIs of local benthic communities. The growing VMS (Vessel Monitoring System) dataset on the distribution of trawling (Hintzen et al., 2012; ICES, 2014; Piet and Hintzen, 2012) will form one important contribution to this goal. Further closure of part of the Dutch coastal zone for fisheries in the Habitat Directive framework will provide opportunity to test the effectiveness of separate indicators and BEQI2 with respect to fishing impact.

References

Anonymous, 1959. The Venice System for the Classification of Marine Waters according to Salinity. The final resolution of the symposium on the classification of brackish waters. Arch. Oceanogr. Limnol. 11, 243–248 (Suppl.).

Beukema, J.J., Essink, K., 1986. Common patterns in the fluctuations of macrozoobenthic species living at different places on tidal flats in the Wadden Sea. Hydrobiologia 142, 199–207.

Beukema, J.J., Dekker, R., Jansen, J.M., 2009. Some like it cold: populations of the tellinid bivalve Macoma balthica (L.) suffer in various ways from a warming climate. Mar. Ecol. Prog. Ser. 384, 135–145.

Boon, A.R., Gittenberger, A., van Loon, W.M.G.M., 2011. Review of marine benthic indicators and metrics for the WFD and design of an optimized BEQI, Deltares, report nr 1203801-000. http://publicaties.minienm.nl/documenten/review-of-marine-benthic-indicators-and-metrics-for-the-wfd-and.

Borja, A., Muxika, I., 2005. Guidelines for the use of AMBI (AZTI's Marine Biotic Index) in the assessment of the benthic ecological quality. Mar. Pollut. Bull. 50, 787–789.

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

Borja, A., Franco, J., Perez, V., 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Mar. Pollut. Bull. 40, 1100–1114.

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

Borja, A., Muxika, I., Franco, J., 2006b. Long-term recovery of soft-bottom benthos following urban and industrial sewage treatment in the Nervión estuary (southern Bay of Biscay). Mar. Ecol. Prog. Ser. 313, 43–55.

Borja, A., Josefson, A.B., Miles, A., Muxika, I., Olsgard, F., Philips, G., Rodriguez, J.G., Rygg, B., 2007. An approach to the intercalibration of benthic ecological status assessment in the North Atlantic ecoregion, according to the European Water Framework Directive. Mar. Pollut. Bull. 55, 42–52.

Borja, A., Mader, J., Muxika, I., Rodrıguez, J.G., Bald, J., 2008. Using M-AMBI in assessing benthic quality within the Water Framework Directive: some remarks and recommendations. Mar. Pollut. Bull. 56, 1377–1379.

Borja, A., Barbone, E., Basset, A., et al., 2011. Response of single benthic metrics and multimetric methods to anthropogenic pressure gradients, in five distinct European coastal and transitional ecosystems. Mar. Pollut. Bull. 62, 499–513.

Bouma, H., de Jong, D.J., Twisk, F., Wolfstein, K., 2005. A Dutch Ecotope System for Coastal Waters (ZES.1). To map the potential occurrence of ecological communities in Dutch coastal and transitional waters. Report RIKZ/2005.024 http://publicaties.minienm.nl/documenten/zoute-wateren-ecotopenstelsel-zes-1-a-dutch-ecotope-system-for-c.

Carletti, A., Heiskanen, A.S., 2009. Water Framework Directive intercalibration technical report part 3: coastal and transitional waters. JRC report http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/10473/1/3010_08-volumecoast.pdf.

De Jong, D.J., 2000. Ecotopes in the Dutch marine tidal waters: a proposal for a classification of ecotopes and a method to map them. CM Documents — ICES, CM 2000(T:05). ICES, p. 22 (sol.).

De Jonge, V.N., Schuttelaars, H.M., Van Beusekom, J.E.E., Talke, S.A., De Swart, H.E., 2014. The influence of channel deepening on estuarine turbidity levels and dynamics, as exemplified by the Ems estuary. Estuar. Coast. Shelf Sci. 139, 46–59.

De Vriend, H.J., Wang, Z.B., Ysebaert, T., Herman, P.M., Ding, P., 2011. Eco-morphological problems in the Yangtze estuary and the Western Scheldt. Wetlands 31, 1033–1042.

Dekker, R., Beukema, J.J., 2014. Phenology of abundance of bivalve spat and of their epibenthic predators: limited evidence for mismatches after cold winters. Mar. Ecol. Prog. Ser. 513, 17–27.

Essink, 1999. Ecological effects of dumping of dredged sediments; options for management. J. Coast. Conserv. 5, 69–80.

Essink, K., Beukema, J.J., 1986. Long-term changes in intertidal flat macrozoobenthos as an indicator of stress by organic pollution. Hydrobiologia 142, 209–215.

Gittenberger, A., van Loon, W., 2013. Sensitivities of marine macrozoobenthos to environmental pressures in the Netherlands. Ned. Faun. Meded. 41, 79–112.

Grage, A., Grotjahn, M., Heiber, W., Herleyn, M., Knaackj, Kolbe K., Petri, G., Witt, J., 2013. Gewässerüberwachungssystem Niedersachsen, GütemessnetzÜbergangs- und Küstengewässer — 2013. Küstengewässer und Ästuare Band 6 (50p.).

Gray, J.S., Wu, R.S.S., Or, Y.Y., 2002. Effects of hypoxia and organic enrichment on the coastal marine environment. Mar. Ecol. Prog. Ser. 238, 249–279.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1999. Analisis Multivariate. 5th ed. Prentice Hall Iberia, Madrid (832 pp).

Hering, D., Borja, A., Carstensen, J., et al., 2010. The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. Sci. Total Environ. 408, 4007–4019.

Herman, P.M.J., Middelburg, J.J., van de Koppel, J., Heip, C.H.R., 1999. Ecology of estuarine macrobenthos. Adv. Ecol. Res. 29, 195–240.

Hiddink, J.G., Jennings, S., Kaiser, M.J., Queiro's, A.M., Duplisea, D.E., Piet, G.J., 2006. Cumulative impacts of seabed trawl disturbance on benthic biomass, production and species richness in different habitats. Can. J. Fish. Aquat. Sci. 63, 721–736.

Hintzen, N.T., Bastardie, F., Beare, D., Piet, G.J., Ulrich, C., Deporte, N., Egekvist, J., Degel, H., 2012. VMS tools: open-source software for the processing, analysis and visualisation of fisheries logbook and VMS data. Fish. Res. 115–116, 31–43.

Hiscock, K., Langmead, O., Warwick, R., Smith, A., 2004. Identification of seabed indicator species to support implementation of the EU Habitats and Water Framework Directives. Report to the Joint Nature Conservation Committee and the Environment Agency from the Marine Biological Association http://www.marlin.ac.uk/PDF/CDEP84_5_244FinalRpt_screen.pdf.

ICES, 2014. General advice. OSPAR request on mapping of bottom fishing intensity using VMS data Version 2, 29-09-2014. http://ices.dk/sites/pub/Publication%20Reports/Advice/2014/Special%20Requests/OSPAR_mapping_bottom_fishing_intensity.pdf.

Janssen, G., Mulder, S., 2005. Zonation of macrofauna across sandy beaches and surf zones along the Dutch coast. Oceanol. Acta 47, 265–282.

Josefson, B., Blomqvist, A.M., Hansen, J.L.S., et al., 2009. Assessment of marine benthic quality change in gradients of disturbance: comparison of different Scandinavian multi-metric indices. Mar. Pollut. Bull. 58, 1263–1277.

Kaiser, M.J., Clarke, K.R., Hinz, H., Austen, M.C.V., Somerfield, P.J., Karakassis, I., 2006. Global analysis of response and recovery of benthic biota to fishing. Mar. Ecol. Prog. Ser. 311, 1–14.

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

Lacroix, G., Ruddicka, K., Ozera, J., Lancelot, C., 2004. Modelling the impact of the Scheldt and Rhine/Meuse plumes on the salinity distribution in Belgian waters (southern North Sea). J. Sea Res. 52, 149–163.

Levin, L.A., Gooday, A.J., Jorissen, F., Middelburg, J.J., Naqvi, S.W.A., Neira, C., Rabalais, N.N., Zhang, J., 2009. Effects of natural and human-induced hypoxia on coastal benthos. Biogeosciences 6, 2063–2098.

Lindeboom, H.J., Dijkman, E.M., Bos, O.G., Meesters, H.W.G., Cremer, J.S.M., Raad, I. de, Bosma, A., 2008. Ecologische Atlas Noordzee ten behoeve van gebiedsbescherming. Wageningen IMARES 9789074549127, p. 289.

Maes, J., Stevens, M., Breine, J., 2007. Modelling the migration opportunities of diadromous fish species along a gradient of dissolved oxygen concentration in a European tidal watershed. Estuar. Coast. Shelf Sci. 75, 151–162.

Mees, J., Dewicke, A., Hamerlynck, O., 1993. Seasonal composition and spatial distribution of hyperbenthic communities along estuarine gradients in the Westerschelde. Neth. J. Aquat. Ecol. 27, 359–376.

Meire, P., Ysebaert, T., Van Damme, S., Van den Bergh, E., Maris, T., Struyf, E., 2005. The Scheldt estuary: a description of a changing ecosystem. Hydrobiologia 540, 1–11.

Muxika, I., Borja, A., Bald, J., 2007. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. Mar. Pollut. Bull. 55, 16–29.

Muxika, I., Borja, A., Bonne, W., 2005. The suitability of the marine biotic index (AMBI) to new impact sources along European coasts. Ecol. Indic. 5 (1), 19–31.

OSPAR, 2011. OSPAR's MSFD Advice Manual on Biodiversity, version 3.2, 5 March 2012. http://www.ospar.org/v_publications/download.asp?v1=p00581.

Peet, R.K., 1974. The measurements of species diversity. Annu. Rev. Ecol. Syst. 5, 285–307.

Piet, G.J., Hintzen, N.T., 2012. Indicators of fishing pressure and seafloor integrity. ICES J. Mar. Sci. 69, 1850–1858.

Piet, G.J., Rijnsdorp, A.D., 1998. Changes in the demersal fish assemblage in the southeastern North Sea following the establishment of a protected area ('plaice box'). ICES J. Mar. Sci. 55, 420–429.

R Core Team, 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (URL http://www.R-project.org/).

Reizopoulou, S., Simboura, N., Barbone, E., et al., 2014. Biodiversity in transitional waters: steeper ecotone, lower diversity. Mar. Ecol. 35, 78–84.

Remane, A., 1934. Die Brackwasserfauna. Verh. Deutsch. Zool. Ges. 36, 34–74.

Schaffner, L.C., Dellapenna, T.M., Hinchey, E.K., et al., 2001. Physical energy regimes, seabed dynamics, and organism–sediment interactions along an estuarine gradient. In: Aller, J.Y., Woodin, S.A., Aller, R.C. (Eds.), Organism–sediment Interactions. The Belle W. Baruch Library in Marine Science, University of South Carolina Press, pp. 159–179.

Seitaj, D., Malkin, S.Y., Schauer, R., Meysman, F.J.R., 2013. Competition for sulfide in marine sediments: electrogenic filamentous bacteria versus Beggiatoa. Mineral. Mag. 77, 2174.

Shannon, C.E., Weaver, W., 1963. The Mathematical Theory of Communication. University of Illinois Press, Urbana (117 pp.).

Sigovini, M., Keppel, E., Tagliapietra, D., 2013. M-AMBI revisited: looking inside a widely-used benthic index. Hydrobiologia 717, 41–50.

Teixeira, H., Neto, J.M., Patricio, J., et al., 2009. Quality assessment of benthic macroinvertebrates under the scope of WFD using BAT, the Benthic Assessment Tool. Mar. Pollut. Bull. 58, 1477–1486.

Tillin, H.M., Hiddink, J.G., Jennings, S., Kaiser, M.J., 2006. Chronic bottom trawling alters the functional composition of benthic invertebrate communities on a sea-basin scale. Mar. Ecol. Prog. Ser. 318, 31–45.

Van Buuren, A., Gerrits, L., Teisman, G.R., 2010. Understanding and managing the Westerschelde — synchronizing the physical system and the management system of a complex estuary. Hydrol. Earth Syst. Sci. 14, 2243–2257.

Van Colen, C., Rossi, F., Montserrat, F., Andersson, M.G., et al., 2012. Organism–sediment interactions govern post-hypoxia recovery of ecosystem functioning. PLoS ONE 7, 1–10.

Van Dalfsen, J.A., Essink, K., 2001. Benthic community response to sand dredging and shoreface nourishment in Dutch coastal waters. Senckenb. Marit. 31, 329–332.

Van Hoey, G., Drent, J., Ysebaert, T., Herman, P., 2007. The Benthic Ecosystem Quality Index (BEQI), intercalibration and assessment of Dutch coastal and transitional waters for the Water Framework Directive. Final report (http://www.vliz.be/en/imis?module=ref&refid=119203. Accessed 21 July 2014).

van Loon, W.M.G.M., Verschoor, A.J., 2012. Benthic Ecosystem Quality Index 2: Application to Dutch Marine Benthos Data from the period 1990-2010. Report http://publicaties.minienm.nl/documenten/benthic-ecosystem-quality-index-2-application-to-dutch-marine-be.

Van Loon, W.M.G.M., Verschoor, A.J., Gittenberg, A., 2011. Benthic Ecosystem Quality Index 2: design and calibration of the BEQI-2 WFD metric for marine benthos in transitional waters Report, 8 December 2011, http://publicaties.minienm.nl/documenten/benthic-ecosystem-quality-index-2-design-and-calibration-of-the-.

Walvoort, D.J.J., Van Loon, W.M.G.M., 2014. BEQI2: Benthic Ecosystem Quality Index 2. R package version 2.0–0. http://cran.r-project.org/BEQI2.

Wanink, J.H., 2014. Toepassing van het BEQI2 tool op diepte-ecotopen van het Grevelingenmeer. Koeman en Bijkerk, Rapport 2014-011.

Weisberg, S.B., Thompson, B., Ranasinghe, J.A., et al., 2008. The level of agreement among experts applying best professional judgment to assess the condition of benthic infaunal communities. Ecol. Indic. 8, 389–394.

Widdicombe, S., Austen, M.C., 2001. The interaction between physical disturbance and organic enrichment: An important element in structuring benthic communities. Limnol. Oceanogr. 7, 1720–1733.