Available online at www.sciencedirect.com
ScienceDirect
Marine Pollution Bulletin 53 (2006) 508–522
MARINE POLLUTION BULLETIN
www.elsevier.com/locate/marpolbul

Indices for the assessment of environmental pollution of the Baltic Sea coasts: Integrated assessment of a multi-biomarker approach

Katja Broeg [a,*], Kari K. Lehtonen [b]

[a] *Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany*
[b] *Finnish Institute of Marine Research, P.O. Box 2, FI-00561 Helsinki, Finland*

Abstract

Two mathematical methods to assess the "health status" of flounder (*Platichthys flesus*), eelpout (*Zoarces viviparus*) and blue mussel (*Mytilus* spp.) populations of the Baltic Sea were applied on selected biomarker data collected during the EU project "BEEP" (Biological Effects of Environmental Pollution on Marine Coastal Ecosystems). The Bioeffect Assessment Index (BAI) and the Integrated Biomarker Index (IBR) combine different biomarkers to single values, which can be used to describe the toxically-induced stress level of populations in different areas. Both indices determined here produced essentially similar results, which in most cases agreed with the known contamination levels in the different study areas. Advantages and limitations of index applications and interpretations are critically discussed. The use of indices provides comprehensive information about biological effects of pollution in marine organisms and may therefore serve as a useful tool for environmental management by ranking the pollution status of marine coastal areas.

Keywords: Baltic Sea; Biomarkers; Integrated indices; Monitoring; Pollution

1. Introduction

The assessment of the integrity of coastal and marine habitats is a complex task and of vital importance for regulation and management activities. Recently, specific emphasis has been placed on the elucidation of potential links between complex contaminant exposure and observable biological effects in aquatic organisms to provide information about their biological significance (Stentiford et al., 2003; Broeg et al., 2005). Validation and intercalibration of biomarkers that are prognostic for the health status of indicator organisms is one step towards the monitoring of toxically induced degradation—as well as improvement—of environmental quality.

To assess whether the health status of a population or a community is at risk, various data on adverse effects at different levels of organisation have to be combined to get a comprehensive and reliable picture. Earlier, Kelly and Harwell (1989) emphasised that damage to ecosystems involves loss of ecosystem structure and function. The fact that ecosystems are seldom in a stable state complicates the identification of disorders and the return to the norm. Ecosystems are complex *and* dynamic systems in which specific endpoints (e.g. species diversity and composition) are not guaranteed to return to the pre-disturbance status after remediation measures (Power, 1999). Therefore, interpretation of results of environmental monitoring based only upon measures at the ecosystem level is extremely difficult.

Biomarkers of exposure and early effects can bridge the gap between chemical analytics on one hand, and late or advanced biological effects monitored at the level of populations and/or communities on the other. Means to assess pollution stress in marine organisms by combining selected biomarkers of exposure and effects have been under recent development (Adams et al., 1993; Narbonne et al., 1999; Beliaeff and Burgeot, 2002; Chèvre et al., 2003a,b; Broeg et al., 2005; Sturve et al., 2005). Environmental deterioration and recovery are processes (Depledge, 1999). Adverse

* Corresponding author. Tel.: +49 4714 8311618; fax: +49 4714 8311425.
E-mail address: kbroeg@awi-bremerhaven.de (K. Broeg).

0025-326X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.marpolbul.2006.02.004

effects occurring at different levels of biological organisation are measured at different time scales (Broeg et al., 2005). Early and progressive responses of organisms can be detected by the application of a set of biomarkers for toxic effects; this kind of a multiple biomarker approach may serve as a central, integrative part in the general multidisciplinary effort for the assessment of environmental integrity.

To allow comparisons of the "pollution status" or "health status" of different areas or populations and the detection of temporal trends, the integration of multifactorial measurements is a basic requirement. It is widely accepted that in ecotoxicology one key challenge is to integrate individual biomarker responses into a set of tools and indices capable of detecting and monitoring the degradation of the health of organisms. Allen and Moore (2004) proposed the term "environmental prognostics" based on responses of organisms to both natural and anthropogenic stress. As a first step of this approach they suggest to relate biomarker responses to the health status of individual organisms, followed by the derivation of integrated frameworks.

During the pan-European EU-funded project "BEEP" (Biological Effects of Environmental Pollution in Marine Coastal Ecosystems, 2001–2004) a large set of biomarkers was tested for their applicability and reliability in reflecting pollution-induced alterations in marine organisms in different coastal areas of Europe. In the present paper, selected biomarker data from target organisms collected in the Baltic Sea component of the BEEP project were subjected to integrated response analyses using two methods, the Bioeffect Assessment Index (BAI) (Broeg et al., 2005) and the Integrated Biomarker Response (IBR) (Beliaeff and Burgeot, 2002).

The Health Assessment Index (HAI, Adams et al., 1993) for fish species has been applied, refined and modified by several authors (Coughlan et al., 1996; Steyermark et al., 1999; Kovacs et al., 2002; Schleiger, 2004). Based upon this concept the Bioeffect Assessment Index (BAI) has recently been developed by Broeg et al. (2005). The concept of BAI is based on observations of progressive deleterious effects in the liver of flounder from early responses to late effects (Broeg et al., 2005). The index has been applied during a long-term study on biological effects of pollution in the German Bight aiming at establishing links between effects observed at different levels of biological organisation. Chemical toxicity is often aggravated by the cumulative and potentially damaging effects of biotic and abiotic factors (e.g. food availability, oxygen deficiency, eutrophication) present in any given environment. Therefore, exclusively biomarkers of general toxicity are included in the BAI to enable the assessment of the combined effects of a complex mixture of contaminants and other types of anthropogenic stressors. The biomarkers selected reflect deleterious effects of various classes of contaminants such as heavy metals, organochlorines, pesticides and PAHs, reflecting general toxicity in an integrative manner. In fish, they included lysosomal perturbations (reduced lysosome membrane stability, LMS) and storage disorders (lipid accumulation, NL) as early markers of toxic effects observed in liver cells, acid phosphatase activity of macrophage aggregates (M-ACT) indicative for the modulation of non-specific immune response which represents longer time scale responses after chronic exposure, and the frequency of micronuclei (MN) as a marker for mutagenicity. In mussels, LMS, NL and MN were used for the calculation of the BAI.

In regard to the IBR, in their original publication Beliaeff and Burgeot (2002) included only biochemical biomarkers [glutathione-*S*-transferase (GST), acetylcholinesterase (AChE) and catalase enzyme activities, and DNA adducts] in the mussel. In the present study also histochemical biomarkers of toxic effects, or "general health" (GH), reflecting cytotoxicity (LMS) and immunotoxicity [M-ACT and macrophage aggregate size (M-AREA)] as well as mutagenic damage (MN) were applied. Moreover, in addition to GH biomarkers, "specific" (SP) biomarkers [metallothionein induction (MT), AChE, metabolites of polyaromatic hydrocarbons (PAH) in bile measured as fluorescent aromatic compounds (FAC), and ethoxyresorufin-*O*-deethylase activity (EROD)] were added (when available) to the GH biomarker sets to examine the effects of their inclusion on the IBR values.

The mathematical basis of both indices is described in Section 2 below. The results and index calculation methods are critically examined with regard to the information that they yield and its consistency, the sensitivity of the indices to detect differences between and within the study areas, and their correspondence with contaminant data or known pollution gradients. The aim of this analysis was to examine the strengths and weaknesses of using integrated indices in ranking or assessing the environmental integrity of different areas of the Baltic Sea.

2. Material and methods

Most of the original data obtained during the BEEP Baltic Sea component used here for the calculations of the indices are presented elsewhere in this volume of Marine Pollution Bulletin (Baršienė et al., 2006a; Kopecka et al., 2006; Schiedek et al., 2006; Vuorinen et al., 2006). Data from four study areas in the Baltic Sea were used: Wismar Bay (Germany), Gulf of Gdańsk (Poland), Klaipeda-Butinge area (Lithuania) and Kvädöfjärden (Sweden, "reference site").

Sampling was carried out at 3–4 stations from all areas (except at Kvädöfjärden, one station), the selection of sites based on previous information on local pollution patterns. Tissue analyses of selected contaminants (mainly organochlorines) in the target organisms were also performed to define the current level of contamination at the different sites. Detailed information on the study areas and tissue contaminant levels can be found in the publications listed above.

The species chosen as indicator organisms were flounder (*Platichthys flesus*), eelpout (*Zoarces viviparus*) and blue mussel (*Mytilus* spp.). Flounder and mussel were present at all study areas while eelpout could be studied only at Wismar Bay and Kvädöfjärden. Only quality-assured data were included in the index calculations.

2.1. Calculation of the BAI

The parameters used in the calculation of the BAI were graded by expertise according to observations from a biomarker study on more than 1000 individual flounders from the German Bight. Specific key events for the progress of toxically induced alterations could be defined based on this study; the grading of alterations was defined according to Table 1 (for further details see Broeg et al., 2005). In the present study the BAI was also calculated for eelpout and mussels for the first time.

The integration of different biomarkers is made by substituting each parameter value measured for each individual according to the progression of toxically induced alteration with a numerical value as follows: 10 = Stage 1, 20 = Stage 2, 30 = Stage 3, 40 = Stage 4 (Table 2). The BAI value for each location is obtained by summing up all the BAI values calculated for each individual fish and dividing them by the number of individuals analysed. A higher BAI value indicates a poorer health condition and vice versa.

For the classification of limits of the BAI scale a mean value of 25 was defined as the critical value since values above 30 are indicative for an advanced state of the pollution-induced deterioration of health condition reflecting a decline of environmental quality. In case of a BAI value of 25 or above at least one of the parameters is in Stage 4 (population risk), or two or more in Stage 3 (individual risk). The BAI categories are defined as follows: 10–15: good environmental condition; >15–25: tolerable environmental condition; >25–35: environmental risk; >35–40: bad environmental condition.

Since at least two parameters included in the BAI calculation show seasonal dependencies [macrophage aggregate activity (MA) (Broeg, 2003) and neutral lipid accumulation (NL), which is dependent on food availability] direct comparisons were made only between sampling campaigns carried out during the same season (spring or autumn).

Table 1
Characterisation of the single BAI stages on the basis of studies on lysosomal membrane stability and liver histopathology (Koehler et al., 2002; Broeg et al., 2005)

Stage 1	Destabilisation time 20 min and longer (reference)	*Healthy*
Stage 2	Destabilisation time >10 to <20 min	*Minor diseased*
Stage 3	Destabilisation time 5–10 min	*Reversible, progressively*
Stage 4	Destabilisation time <5 min	*Irreversible, degenerative*

2.2. Calculation of the IBR

The IBR is calculated by summing up triangular Star Plot areas calculated for each two neighbouring biomarkers in a given data set. The procedure described below was used:

For each biomarker: (1) Calculation of mean and SD for each station. (2) Standardisation of data for each station: $x_i' = (x_i - \text{mean } x)/s$, where x_i' = standardised value of the biomarker, x_i = mean value of a biomarker from each station, mean x = mean of the biomarker calculated for all the stations, and s = standard deviation calculated for the station-specific values of each biomarker. Result: variance = 1, mean = 0. (3) Using standardised data, addition of the value obtained for each station to the absolute (=non-negative) value of the minimum value in the data set: $B = x_i' + |x_{\min}|$. Result: adjusts the lowest value in the set to zero. For all the biomarkers treated this way: (4) calculation of Star Plot areas by multiplication of the obtained value of each biomarker (B_i) with the value of the next biomarker, arranged as a set, dividing each calculation by 2 and (5) summing-up of all values: $\{[(B_1 \times B_2)/2] + [(B_2 \times B_3)/2] + \cdots [(B_{n-1} \times B_n)/2]\}$. Result: IBR (average of different arrangements of biomarkers in the set).

In most of the cases, only 4–6 biomarkers could be used for each calculation. In some cases the data set consisted of more SP than GH biomarkers. Since the value of IBR is obtained by summing up the parameters derived from the actual biomarker values, i.e. after the calculation steps 1–4, it is directly dependent on the number of biomarkers in the set. Thus, in the results section the values of IBR are given divided by the number of biomarkers used in each

Table 2
Stages of toxically-induced alterations of the biomarkers and corresponding BAI values

	Stage 1	Stage 2	Stage 3	Stage 4
Lysosomal stability (destabilisation period in minutes)	⩾20	>10<20	5–10	<5
Numerical value BAI	***10***	***20***	***30***	***40***
Neutral lipid accumulation (mean absorbance)	1–2	3–4	5–6	>6
Numerical value BAI	***10***	***20***	***30***	***40***
Acid phosphatase activity in MA (mean absorbance)	0.5–0.6	0.4–0.49>0.6	0.3–0.39	<0.3
Numerical value BAI	***10***	***20***	***30***	***40***
Micronuclei	0	0.2–0.3	0.4–0.6	>0.6
Numerical value BAI	***10***	***20***	***30***	***40***

case and termed as IBR/n. It should also be noted that because the calculation method of IBR is based on relative differences between the biomarker responses in each given data set it is necessary to re-calculate all the index values each time when making new comparisons, e.g. adding new sites or comparing seasonal values (see also Section 4).

3. Results

3.1. BAI

3.1.1. Flounder

The BAI was calculated for flounder sampled in autumn 2001 in the Wismar Bay (station Walfisch) and the Swedish location Kvädöfjärden, and in autumn 2002 for populations from all Baltic Sea BEEP stations (Fig. 1). The highest mean BAI values were measured at Walfisch in 2001 (mean 31.5). One year later (autumn 2002) the BAI value had improved slightly (27) but was still above the critical value of 25. In contrast, flounder from Kvädöfjärden had relatively low values in 2001 (22) and showed a decline of health status in 2002 reflected by an increased BAI value (25). At the Lithuanian coast the BAI tolerance level was exceeded only at the Palanga location (25.5) in autumn 2002. Flounder at the two other study sites (Butinge and Nemirseta) showed similar values (around 22.5) in the tolerable range (Stage 2) of the BAI scale. In the Gulf of Gdańsk BAI values determined for flounder collected from Sopot and Mechelinki were also in the tolerable range while at Sobieszewo the local population exceeded the critical value (27.5) in 2002.

3.1.2. Eelpout

In autumn 2001, eelpout from Kvädöfjärden expressed BAI values slightly below the critical value (24.5) (Fig. 2). Individuals from the Wismar Bay stations had a mean BAI above the critical limit (26 in autumn 2001 and 28–29 in autumn 2002). The BAI value was higher at the Wismar Harbour station compared to Salzhaff.

3.1.3. Blue mussel

The BAI was calculated for mussels collected in spring and autumn 2001 in the Wismar Bay and in the Lithuanian coast, and in spring and autumn 2002 at all the Baltic Sea stations of the BEEP project (Fig. 3). The highest mean values of BAI (above 35) were measured at all three Wismar Bay stations in spring 2001. In autumn 2001 the BAI had slightly improved at the innermost station near the Wismar Harbour (25) but remained high at Eggers Wiek (30.5) outside the harbour. In the population of the Wismar Bay reference location Salzhaff the BAI indicated a marked recovery in autumn 2001 (15). In spring 2002, a deterioration of the BAI of mussels was observed again at the Wismar Harbour and Salzhaff study sites but the integrated response was not as pronounced as in spring 2001. In autumn 2002 the BAI was below 25 at all the study locations except the Gulf of Gdańsk, and in the Wismar Bay a clear gradient with decreasing values could be detected between Wismar Harbour and Salzhaff.

In the Gulf of Gdańsk, data on mussels from sampling campaigns of spring and autumn 2002 were used for BAI calculations. In spring, data from all the study stations

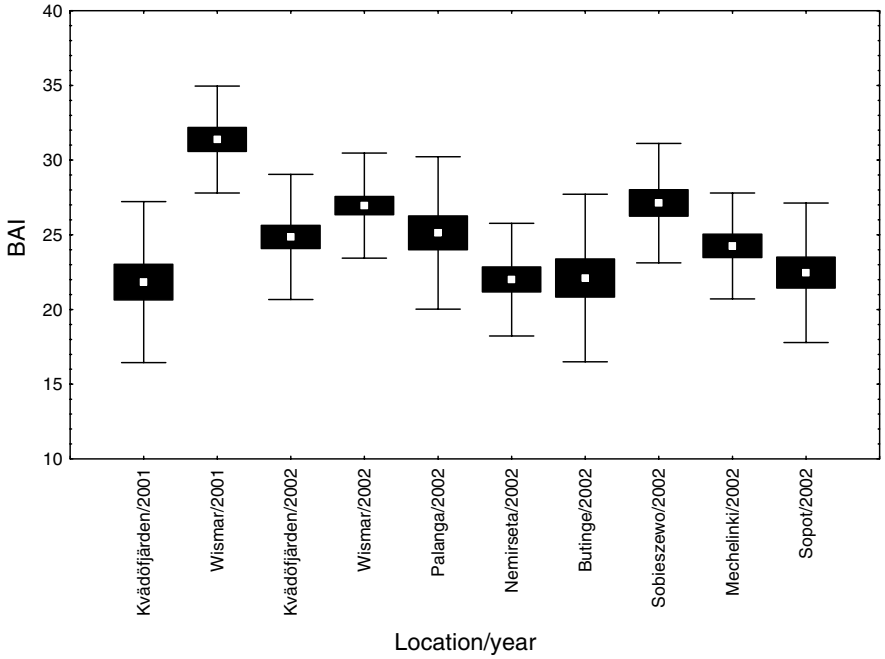

Fig. 1. BAI values (mean and STD), calculated for flounder (*Platichthys flesus*) from Kvädöfjärden and Walfisch (Wismar Bay) in autumn 2001 and 2002 and three locations at the Lithuanian coast (Palanga, Nemirseta and Butinge) and in the Gulf of Gdańsk (Sobieszewo, Mechelinki and Sopot) in autumn 2002.

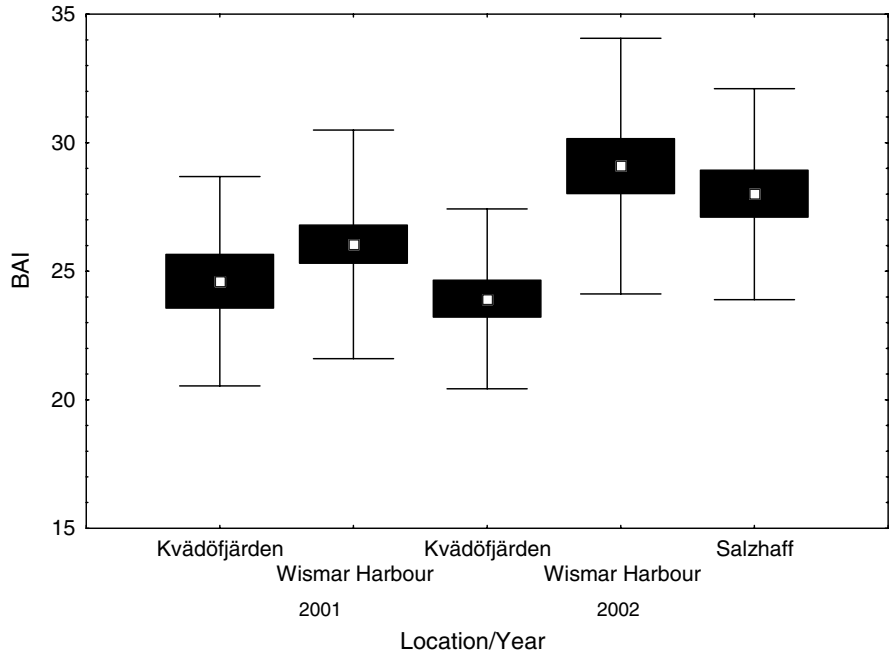

Fig. 2. BAI values (mean and STD), calculated for eelpout (*Zoarces viviparus*) from Kvädöfjärden and Wismar Bay in autumn 2001 and 2002.

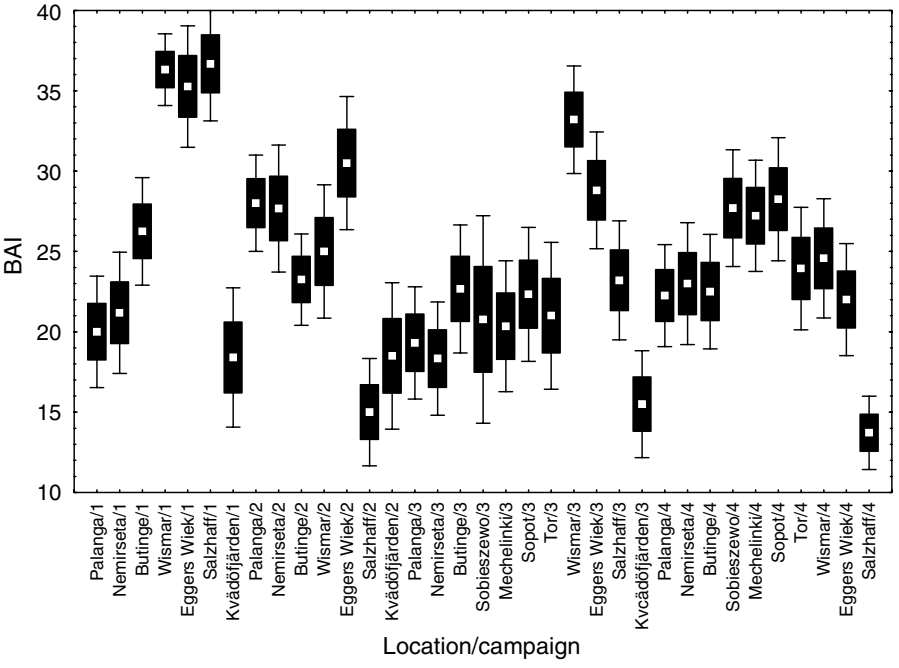

Fig. 3. BAI values (mean and STD), calculated for mussel (*Mytilus* spp.) from the whole BEEP Baltic Sea study area during four sampling campaigns: (1) spring 2001; (2) autumn 2001; (3) spring 2002; (4) autumn 2002.

produced similar BAI values (below 25). In autumn 2002 a decline in the health of populations was observed, indicated by mean BAI values between 28 and 29 at the Sobieszewo, Mechelinki and Sopot sampling sites, while mussels at the outermost location Tor showed a better condition (24).

In mussels from the Lithuanian coast the BAI tolerance level was exceeded at three occasions during the four sampling campaigns, in spring 2001 at station Butinge (26.5) and in autumn 2001 at stations Nemirseta and Palanga (28). In 2002 all BAI values measured in mussels from the Lithuanian coast were within the tolerable range (15–25) with mean values higher in autumn (23–24) compared to spring (18–19) at all locations.

Mussels from the reference location Kvädöfjärden showed BAI values below 20 during all sampling campaigns.

3.2. IBR

To illustrate the characteristics of the data sets used, standardised biomarker parameters (obtained from original data as described in Section 2) used for the calculation of the IBR are shown in Tables 3 and 4. Data used for between-season comparisons are not shown here. Since, for reasons explained above, the numerical values of IBR are comparable only within each data set they are not presented in the text.

3.2.1. Flounder

In spring 2002, a biomarker set consisting of all GH biomarkers LMS, NL, M-AREA, M-ACT and MN and the SP biomarker MT was used for the calculation of the IBR for flounder sampled from the whole area (Table 3). In the Wismar Bay, a marked difference was observed between the two sampling stations with the IBR/n values measured at station Offentief being clearly the lowest and those measured at station Walfisch the highest, also in regard to the whole study area (Fig. 4A). In the Gulf of Gdańsk, the flounder populations at stations Sobieszewo and Sopot showed a higher integrated response compared to Mechelinki. In the Lithuanian sites the index values were markedly higher at Palanga compared to Nemirseta (no data from station Butinge). The IBR/n determined for the flounder population of the Kvädöfjärden "reference" site was at the same low level as the values recorded at stations Nemirseta and Mechelinki. When MT was removed from the biomarker set, inter-station differences in IBR/n remained similar as described above.

Data obtained in autumn 2002 enabled the application of all GH indicators, and MT and FAC as SP biomarkers. Using only GH biomarkers the highest IBR/n values were obtained for flounder populations from Kvädöfjärden and Sobieszewo while low index values were recorded in Butinge and Sopot (Table 3, Fig. 4B). Quite opposite to spring 2002, flounder from Walfisch now showed a low IBR/n value compared to most other stations. The inclusion of MT and FAC to index calculations resulted in substantial elevations in IBR/n at stations Sobieszewo and Kvädöfjärden (already high using GH biomarkers alone), and also at Sopot (characterised by a low value using GH biomarkers).

To compare the seasonal levels of IBR, data from spring and autumn 2002 were used. The GH biomarkers LMS, M-AREA, M-ACT and MN produced a markedly higher IBR/n in spring compared to autumn at all Lithuanian stations and in the German site Walfisch (Fig. 4C). Inconsistent patterns of index values were obtained at the Gulf of Gdańsk stations. In Kvädöfjärden no seasonal variability could be observed. When MT was added to the biomarker set as a SP parameter a quite different seasonal pattern was

Table 3
Standardised biomarker parameters used in IBR calculations (see text body for details)

	General health biomarkers					Specific biomarkers			
	LMS	MN	M-AREA	M-ACT	NL	MT	AChE	FAC	EROD
Flounder—Spring 2002									
Kvädöfjärden	0.44	0.99	0.76	0.65	0.19	0.78	–	–	–
Nemirseta	1.04	*0.00*	0.67	1.29	0.06	0.58	–	–	–
Palanga	1.08	**1.35**	0.61	**1.54**	*0.00*	0.20	–	–	–
Mechelinki	*0.00*	0.87	0.62	0.49	0.83	*0.00*	–	–	–
Sobieszewo	**1.14**	0.81	0.63	0.70	0.84	0.42	–	–	–
Sopot	0.03	0.73	0.73	1.17	**1.32**	0.04	–	–	–
Walfisch	1.01	0.40	**0.91**	0.86	1.24	**1.29**	–	–	–
Offentief	0.41	0.26	*0.00*	*0.00*	0.03	0.06	–	–	–
Flounder—Autumn 2002									
Kvädöfjärden	0.24	1.02	1.17	0.86	0.93	**1.95**	–	0.92	–
Butinge	0.50	*0.00*	0.70	0.16	0.16	1.34	–	*0.00*	–
Nemirseta	0.53	0.31	**0.71**	0.72	0.91	1.28	–	0.16	–
Palanga	1.05	0.39	0.66	0.42	**1.36**	1.24	–	0.35	–
Mechelinki	0.53	**1.29**	0.02	**1.48**	0.45	*0.00*	–	1.48	–
Sobieszewo	**1.33**	1.08	0.05	0.98	0.84	1.43	–	**2.08**	–
Sopot	*0.00*	0.69	*0.00*	1.43	*0.00*	0.83	–	1.11	–
Walfisch	0.61	0.66	0.08	*0.00*	1.15	0.78	–	0.17	–
Eelpout—Spring 2001									
Wismar Harbour	–	0.05	–	–	–	0.47	**0.83**	**0.84**	*0.00*
Eggers Wiek	–	*0.00*	–	–	–	**0.86**	0.70	0.07	**0.84**
Salzhaff	–	**0.59**	–	–	–	*0.00*	*0.00*	*0.00*	0.44
Eelpout—Autumn 2002									
Kvädöfjärden	*0.00*	**0.58**	*0.00*	*0.00*	*0.00*	–	*0.00*	*0.00*	–
Wismar Harbour	**1.51**	0.07	**1.07**	**0.77**	0.67	–	**1.38**	1.11	–
Salzhaff	0.82	*0.00*	0.40	0.35	**1.25**	–	0.88	**2.03**	–

Zero values shown here in italics indicate the site of lowest response observed in each biomarker while highest values are bolded. For biomarker abbreviations, see text.

Table 4
Standardised biomarker parameters used in IBR calculations (see text body for details)

	General health biomarkers			Specific biomarkers	
	LMS	MN	NL	MT	AChE
Mytilus spp.—*Spring 2001*					
Butinge	0.69	0.72	–	2.42	–
Nemirseta	**1.61**	0.44	–	**2.43**	–
Palanga	1.02	*0.00*	–	2.30	–
Wismar Harbour	0.19	**1.83**	–	1.03	–
Eggers Wiek	*0.00*	1.42	–	0.99	–
Salzhaff	0.10	1.42	–	*0.00*	–
Mytilus spp.—*Autumn 2001*					
Kvädöfjärden	**1.33**	–	0.06	**2.08**	–
Butinge	*0.00*	–	1.45	1.81	–
Palanga	0.29	–	**1.92**	0.93	–
Wismar Harbour	1.14	–	0.76	0.33	–
Eggers Wiek	1.08	–	1.73	*0.00*	–
Salzhaff	0.18	–	*0.00*	0.14	–
Mytilus spp.—*Spring 2002*					
Kvädöfjärden	0.29	*0.00*	–	1.98	1.98
Butinge	*0.00*	1.11	–	2.67	1.53
Nemirseta	0.46	0.50	–	2.88	1.64
Palanga	0.80	0.75	–	**2.90**	1.61
Sobieszewo	0.91	0.80	–	1.50	1.89
Mechelinki	0.30	0.82	–	0.89	**2.41**
Sopot	0.65	0.78	–	0.66	1.60
Wismar Harbour	**1.51**	**1.87**	–	1.18	1.58
Eggers Wiek	1.18	1.47	–	0.52	1.01
Salzhaff	0.69	0.78	–	*0.00*	*0.00*
Mytilus spp.—*Autumn 2002*					
Kvädöfjärden	*0.00*	0.01	1.01	2.16	1.51
Butinge	0.16	1.07	0.44	0.85	2.17
Nemirseta	0.83	1.04	0.32	0.60	2.23
Palanga	1.28	0.73	*0.00*	0.15	1.98
Sobieszewo	1.86	1.60	0.57	1.50	2.71
Mechelinki	**2.07**	1.15	1.32	0.53	**2.79**
Sopot	1.35	**2.19**	0.63	2.09	2.57
Tor	1.44	1.17	0.95	**2.20**	2.60
Wismar Harbour	1.79	0.90	0.63	*0.00*	1.56
Eggers Wiek	1.15	0.42	**2.08**	0.09	2.20
Salzhaff	0.40	*0.00*	0.13	0.08	*0.00*

Zero values shown here in italics indicate the site of lowest response observed in each biomaker while highest values are bolded. For biomarker abbreviations, see text.

recorded with smaller between-site differences in IBR/n at the Lithuanian stations and markedly elevated values in autumn at all Gulf of Gdańsk sites (especially at Sobieszewo) and Kvädöfjärden (Fig. 4D).

3.2.2. Eelpout

In spring 2001, a data set consisting of GH biomarkers LMS and MN and SP biomarkers MT, AChE, FAC and EROD was used to examine differences in IBR between eelpout populations inhabiting the Wismar Bay sites. A gradient could be observed, with eelpout at station Salzhaff expressing the lowest and those from the Wismar Harbour site the highest IBR/n values (Table 3, Fig. 5A).

In autumn 2002, IBR calculations utilizing the full set of GH biomarkers in eelpout collected from Kvädöfjärden, Wismar Harbour and Salzhaff showed clear differences between the sites (Table 3, Fig. 5B). In Kvädöfjärden the IBR/n became zero and the values calculated for the Wismar Harbour population were markedly higher than at the Salzhaff site. The inclusion of AChE and FAC to index calculations did not change the difference between the two German sites and the IBR/n value for the eelpout population at Kvädöfjärden remained zero.

To examine seasonal differences the IBR was calculated from data covering spring and autumn 2002 GH biomarkers LMS and MN and SP biomarkers AChE and FAC. The results showed a higher integrated stress response in autumn compared to spring except at Kvädöfjärden where no difference between the seasons could be observed (Fig. 5C).

3.2.3. Blue mussel

In spring 2001, the IBR calculations consisting of GH biomarkers LMS and MN and SP biomarker MT yielded

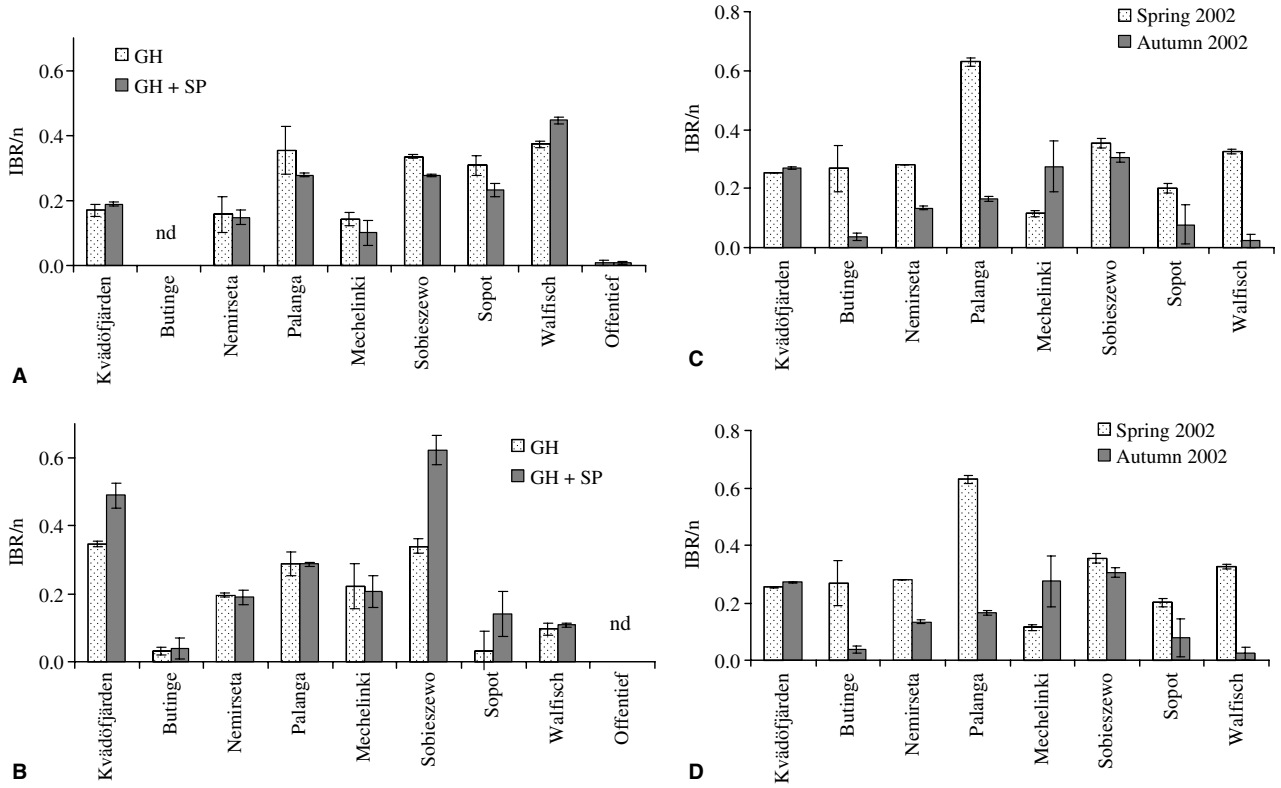

Fig. 4. IBR/n in *Platichthys flesus*. (A) Spring 2002: "general health" (GH) biomarkers LMS, NL, M-AREA, M-ACT and MN, and "specific" (SP) biomarker MT. (B) Autumn 2002: GH biomarkers LMS, NL, M-AREA, M-ACT and MN, and SP biomarkers MT and FAC. (C) A comparison of seasonal levels in IBR using spring and autumn 2002 data collected from the different study stations by using GH biomarkers LMS, M-AREA, M-ACT and MN, and (D) using the same GH biomarkers supplemented by SP biomarker MT. Mean and standard deviation calculated for IBR values obtained using different arrangements for biomarkers within each set. nd = no data.

higher values at the Lithuanian sites Butinge and Nemirseta compared to the local reference station Palanga and all German sites (Table 4, Fig. 6A). The latter showed a clear gradient in the index with the highest values at Wismar Harbour.

In autumn 2001, a set consisting of GH biomarkers LMS and NL and SP biomarker MT generated high index values also for the population at the reference station Kvädöfjärden, which was not studied in spring 2001 (Fig. 6B). Mussels of Salzhaff in the Wismar Bay showed the lowest IBR while no differences could be observed between the Lithuanian stations.

In spring 2002 the parameters used for IBR calculations were GH biomarkers LMS and MN and SP biomarkers AChE and MT. Compared to most other sites the IBR/n values of mussels at all the Lithuanian stations were high, especially at station Palanga (Table 4, Fig. 6C). In the Gulf of Gdańsk the IBR/n of mussels at Mechelinki and Sobieszewo was higher compared to Sopot, regarded as the local reference site. In mussel populations of the Wismar Bay a gradient in IBR/n was again distinct.

In autumn 2002, IBR calculations based on GH biomarkers LMS, MN and NL showed very low IBR for all mussels populations from the Lithuanian sites and markedly higher values at all sites from the Gulf of Gdańsk (Table 4, Fig. 6D). The addition of AChE and MT to calculations led to a similar pattern except for a marked elevation in IBR/n at some stations (mainly Sopot and Palanga). In the Wismar Bay, Salzhaff stood out as the least affected station regardless of the biomarker set employed. Also in Kvädöfjärden the IBR/n of the local mussel population was low.

To examine seasonal differences in the levels of integrated response of populations in the different areas, data from spring and autumn 2002 consisting of GH biomarkers LMS and MN and SP biomarkers MT and AChE were used for calculations. In Kvädöfjärden as well as at the Lithuanian and Wismar Bay stations the IBR/n of mussels was slightly higher in spring (except for one case) but quite the opposite was recorded at the Polish stations characterised by markedly higher values in autumn compared to the spring (Fig. 6E).

4. Discussion

4.1. General aspects

One of the main issues in performing realistic assessments of integrated pollution stress encountered by populations is the appropriate selection of biomarkers used for the calculation of the indices. The BAI is defined as a "general health" index (Broeg et al., 2005) and is therefore

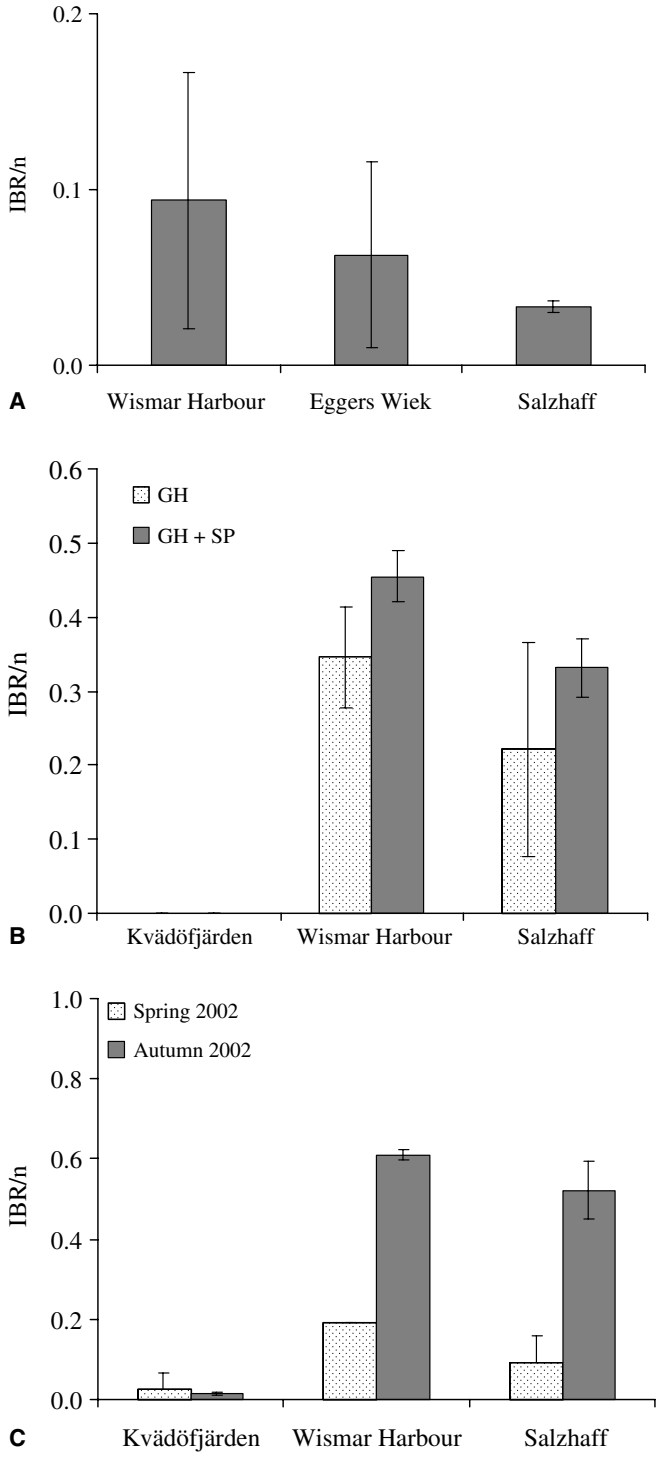

Fig. 5. IBR/n in *Zoarces viviparus*. (A) Spring 2001: "general health" (GH) biomarker MN and "specific" (SP) biomarkers MT, AChE, EROD and FAC. (B) Autumn 2002: GH biomarkers LMS, NL, M-AREA, M-ACT and MN, and SP biomarkers AChE and FAC. (C) A comparison of seasonal levels in IBR using spring and autumn 2002 data collected from the different study stations by using GH biomarkers LMS and MN, and SP biomarkers AChE and FAC. Mean and standard deviation calculated for IBR values obtained using different arrangements for biomarkers within each set.

based on parameters of non-specific stress while the IBR values were constructed either from GH biomarkers alone or complementing the set with SP biomarkers. This was done to enable direct comparisons of the results using the two mathematical methods on one hand as well as examining how the incorporation of SP biomarkers may affect the results.

When using SP biomarkers in any index calculations, e.g., here in case of IBR, the selection of several parameters that more-or-less respond to the same type of pollution (e.g. PAH metabolites in bile, EROD activity and DNA adducts as marker for PAH exposure) will inherently overemphasise the importance of the presence *or* absence of a certain group of compounds in the overall stress assessment of environmental quality. In regard to the present study, differences in the nature of chemical impact (different kinds and mixtures of contaminants present at the different locations) were observed. For example, at the Lithuanian coast an oil spill in autumn 2001 (Baršienė et al., 2006a,b) caused marked elevations in the IBR calculated including EROD and FAC in the biomarker set using data collected during the spring 2002 campaign whereas the responses of the BAI consisting of general stress indicators only were less distinct. Oppositely, in the Wismar Bay where a strong impact related to organochlorine compounds was observed in the target organisms (Schiedek et al., 2006) the general stress responses indicated by BAI were more pronounced compared to the IBR including also biomarkers of specific exposure; nevertheless, the observed trend was identical by using either BAI or IBR.

Furthermore, the selection of many biomarkers representing the same biological level or function (e.g. indicators of genotoxicity, immunotoxicity, metabolism of xenobiotics, endocrine disruption) causes this particular level of response becoming over-represented in the final index value at the expense of responses at some other, possibly more important or more sensitive levels. To summarize: various kinds of parameters may be used for the calculation of integrated responses of organisms to pollution but one must be fully aware of what the parameters represent and what the index derived from them is supposed to tell us.

For the use of integrated indices in biological effects monitoring the following aspects should be considered:

1. The battery of biomarkers monitored should be selected with flexibility to calculate different *relevant* indices that reflect general stress and/or imply the type of specific pollution or damage.
2. If the biomarker set is correctly selected and large enough, various indices can be calculated from the biomarker data obtained in a monitoring programme (e.g. using only a part of the biomarkers).
3. Continuous time series of the biomarkers selected for the calculation of the indices are essential to obtain temporal trends.

A major difference between classical toxicology and ecotoxicology is that in the latter context we cannot expect fully controlled exposure situations. A large variety of

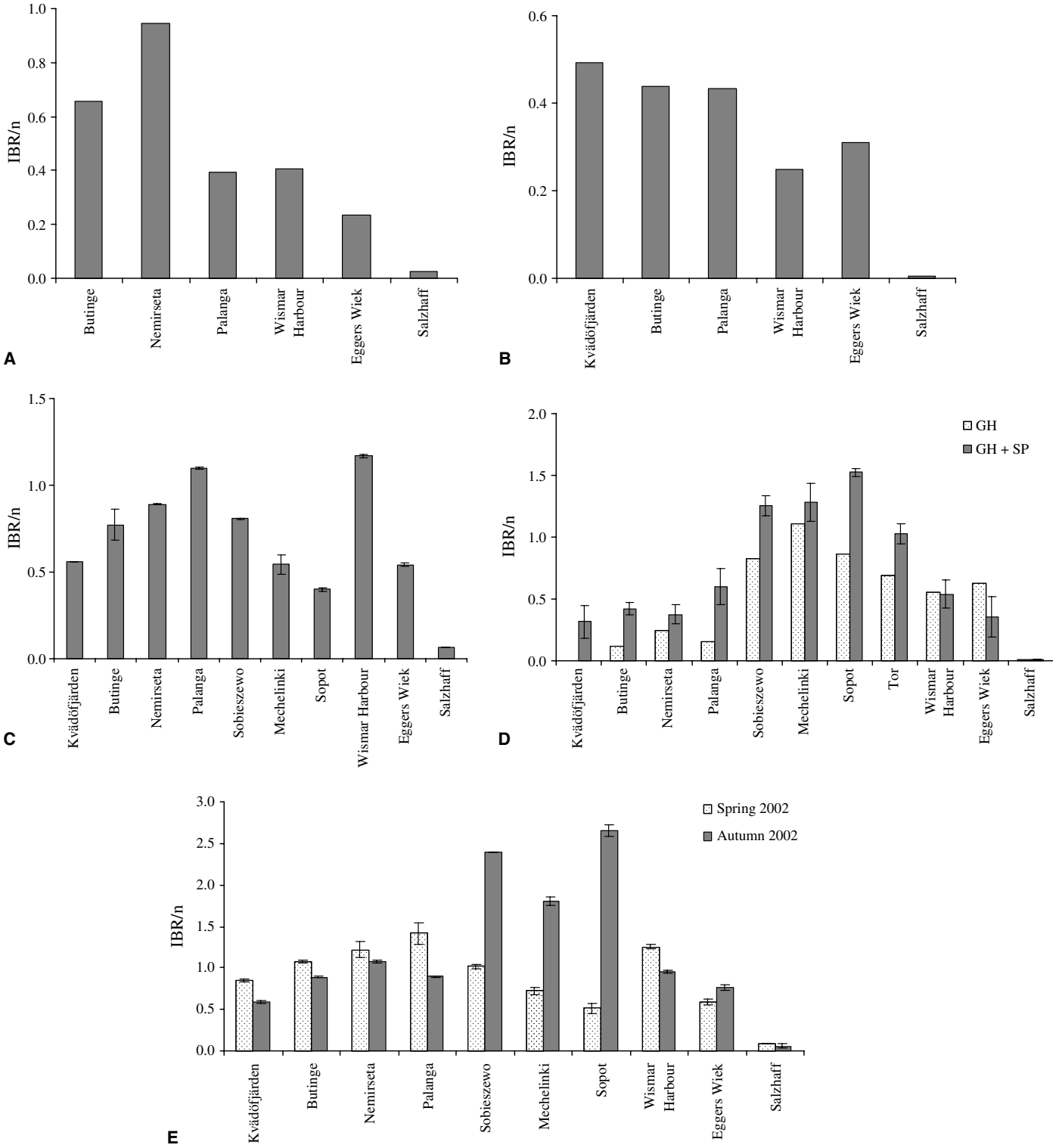

Fig. 6. IBR/n in *Mytilus* spp. (A) Spring 2001: "general health" (GH) biomarkers LMS and MN, "specific" (SP) biomarker MT. (B) Autumn 2001: GH biomarkers LMS and NL, and SP biomarker MT. (C) Spring 2002: GH biomarkers LMS and MN, and SP biomarker MT and AChE. (D) Autumn 2002: GH biomarkers LMS, NL and MN, and SP biomarkers AChE and MT. (E) A comparison of seasonal levels in IBR using spring and autumn 2002 data collected from the different study stations by using GH biomarkers LMS and MN, and SP biomarkers MT and AChE. Mean and standard deviation calculated for IBR values obtained using different arrangements for biomarkers within each set (when feasible).

uncontrolled endogenic and abiotic factors present in the field may modify the toxically induced responses. In the field, equal concentrations of toxicants may cause different intensities of effects dependent on season, temperature, nutritional status, reproductive stage, oxygen availability, etc. Significant effects may not be recorded at one time of

the year but during another season they might become apparent because the magnitude of the stress responses changes. These factors are also reflected in the integrated measures of pollution stress; therefore it should be well understood which factors affect which of the components used in index calculations, and what the direction and magnitude of the effect is.

4.2. Specific considerations concerning the BAI

The BAI is based on biomarkers that have been demonstrated to link with alterations at the ecosystem level, the parasite diversity in the individual host (Broeg et al., 1999; Schmidt et al., 2003; Broeg et al., 2005). Routine implementation of the BAI in monitoring programmes enables the detection of toxically induced disorders of organisms' health, and recoveries resulting from remediation measures.

Considering the comprehensive knowledge available on the single biomarkers included in the BAI the following aspects have to be taken into account for its successful application:

1. The activity of the non-specific immune response in fish is temperature and maturity dependent (Bly and Clem, 1992; Alcorn et al., 2002; Broeg, 2003). Thus, the BAI should only be implemented during one defined season.
2. The accumulation of neutral lipids as an indicator for toxically induced alteration of liver fat metabolism is most pronounced in autumn, after a long period of intensive feeding. Direct comparisons of fish BAI values obtained from measurements during different seasons are therefore not feasible.
3. Since the non-specific immune response shows age and gender specific differences (Tatner and Manning, 1985; Broeg, 2003) the definition of the age/size class and sex of the organisms included for the biomarker assessments is needed.

Furthermore, the following of standard operation procedures is essential during the analyses of the biomarkers included in the BAI.

The ecotoxicological relevance of the critical BAI value of 25 established for flounder during the initial study appears to be valid also for eelpout. This value was initially placed between Stages 2 and 3 because the validation study on North Sea flounder demonstrated a risk for the individual and population health between those stages, reflected by a suppression of the non-specific immune response and a decrease of parasite diversity in individual hosts (Broeg et al., 2005). The eelpout populations from the Wismar Bay exceeded the BAI value of 25, coinciding with distinct reproductive disorders that indicate adverse effects at the population level (Gercken et al., 2006). In eelpout from Kvädöfjärden with a BAI value in the range of 25 and below, no reproductive disorders were reported (Gercken et al., 2006). Thus, the definition of the critical BAI value of 25 for the estimation of a potential risk for the North Sea flounder population (Broeg et al., 2005) seems applicable also for eelpout from the Baltic Sea. In case of mussels, more detailed information about the effects of the alterations of the single biomarkers on the population health is needed to characterize a critical BAI value.

4.3. Specific considerations concerning the IBR

The IBR was able to distinguish inter- and intra-regional as well as seasonal differences of responses in all the three species studied. The IBR emerges as a feasible tool to examine differences in responses between different populations and sampling times by using variable biomarker combinations. The majority of the results of the IBR were remarkably consistent, reflecting the pollution levels measured at the different study locations regardless of the considerable variability in the biomarker sets used for the index calculations. However, various points have to be carefully considered when using this index.

A practical problem related to the mathematical basis of the IBR is that it does not allow for a direct comparison of the numerical values obtained earlier with new data from the site (temporal monitoring from the same site) or data from other sites (spatial comparisons). Thus, new data must be incorporated and processed together with the previous ones, resulting in new IBR values, which then are comparable. For example, in the present study the additional examination of seasonal differences of IBR levels was done this way. Thus, the IBR can be described as a "dynamic" index that is not able to give fixed, directly comparable numerical values to describe the "health status" of a population at any given time, e.g. in a time series, without recalculations. Nevertheless, in this study the IBR succeeded well in identifying temporal and spatial differences between populations and their magnitude, and can in this sense be regarded as a very useful tool.

Due to its mathematical basis the IBR becomes more robust when the number of biomarkers increases. In the original version of Beliaeff and Burgeot (2002) the order of biomarkers within the data sets used in the calculation of Star Plot areas was fixed, randomly or intentionally, ignoring the obvious problem of obtaining different values with different biomarker arrangements. In the modification applied here (Leiniö and Lehtonen, 2005; Lehtonen et al., in press) several IBRs were calculated from the same data changing the order of the biomarkers and using the mean $\pm$ SD of all the index values as the final index value. As one might expect the SD of mean becomes smaller as the number of biomarkers increases, illustrating that the order of the biomarkers in the calculation process does not play any notable role when their number is sufficiently large. In addition, the number of biomarkers included in the calculation of the IBR plays an important role affecting the "relative weight" of each biomarker in the final index value. When the set of biomarkers is relatively large, e.g. 6–8, the weight of one factor is markedly reduced compared to cases when 3–4 biomarkers are used.

One consequence of the calculation procedure of the IBR, based on multiplication of standardised biomarker values with each other, is that in cases when there are no zero values (i.e., the lowest standardised mean response level within all study stations) in none of the biomarkers at some of the sites and all biomarker levels at these sites are intermediate or low, the IBR may, as a result, be high compared to those populations where zero values exist but the levels of some biomarkers are comparatively high. In the present study this situation was observed e.g. in spring 2002 flounder data from the Sopot sampling site (Table 3, Fig. 4) and obviously calls for carefulness in interpretation.

In cases when only one biomarker is notably elevated it may still have a marked effect on the IBR value. However, in a less likely situation where one biomarker shows high values but all the others are the lowest of those among the sites investigated, the mechanism of the IBR calculation always produces a zero value for that site. In practice this scenario is, however, unlikely since many hazardous chemicals usually have effects on multiple mechanisms at different biological levels. Provided that the biomarker set applied is large enough and able to detect responses at various levels, more than one biomarker response is usually observed in contamination situations. In any case, when using the IBR, spatial differences in the levels of single biomarkers that constitute the index should also be examined.

In some cases the IBR was remarkably similar regardless of using only GH biomarkers or complementing them with SP biomarkers. Calculations using the autumn 2002 data on flounder showed marked elevations in the IBR only at stations where the levels of the SP biomarkers MT and FAC were in the high end of the range measured for all the stations (Table 3, Fig. 4). This implies that many of the responses observed are caused by the same impact factors or are even physiologically coupled, and that the effects of specific pollutants can also be detected using non-specific biomarkers.

4.4. Integrated responses observed in the different study areas

In eelpout and mussels both BAI and IBR showed good accordance with tissue levels of organochlorine compounds established for populations from the Wismar Bay and the reference station Kvädöfjärden (Schiedek et al., 2006). Along the Lithuanian coast and in the Gulf of Gdańsk where chemical contamination gradients did not exist or were not clear (Baršienė et al., 2006a,b; Kopecka et al., 2006) and the nature and distribution of pollutants between the sites was more heterogeneous, IBR and BAI exhibited no obvious spatial patterns. In flounder and mussels the high IBR values measured at the Lithuanian "reference" station Palanga in spring 2002 are most likely related to the oil spill affecting the site in November 2001 (Baršienė et al., 2006a,b).

In spring 2002, when the BAI value for mussels at station Wismar Harbour indicated strong toxically induced biological effects, the tissue concentrations of DDTs in mussels at this site (360 ng g lipid wt^{-1}) were almost twice higher than in autumn 2001 (Schiedek et al., 2006) and about three times higher than in mussels from the Lithuanian sampling stations in spring 2002 (Baršienė et al., 2006a). Mussel populations at all Lithuanian stations showed BAI values within the tolerance range (15–25), which agrees with the lower tissue levels of organochlorines measured compared to the Wismar Bay.

The health status of flounder populations of the Baltic Sea, reflected by BAI, did not differ significantly from that recorded in North Sea flounder in the initial validation study of the index (Broeg et al., 2005). However, in contrast to the North Sea study, no such heavily polluted estuary comparable to the Elbe estuary was sampled in the Baltic. On the other hand, it became obvious during the BEEP project that there probably is no true "reference" location in the Baltic Sea where "pristine" environmental conditions would allow the sampling of "healthy" fish (BAI < 15) (for further details see Baršienė et al., 2006a; Gercken et al., 2006; Kopecka et al., 2006; Lang et al., in preparation; Schiedek et al., 2006). As a result, most of the fish collected at the different sites did not differ significantly in regard to their BAI values. These findings are well in line with the results of histopathological surveys carried out during the BEEP project (Lang et al., in preparation). Nevertheless, the time course of improvement (Wismar Bay, 2001–2002) and deterioration (Kvädöfjärden, 2001–2002) of the health of flounder populations at single locations could be detected using the BAI. Thus, the BAI serves as a promising tool for routine screening or monitoring of general toxicity and subsequent effects on the health status of bioindicator organisms.

For the first time, the BAI was also applied on mussels. In mussels from the Wismar Bay, strongest deterioration of health condition was observed in spring 2001 followed by a stepwise recovery until autumn 2002.

In regard to IBR, analysis of flounder data from 2002 for seasonal differences demonstrated the strong effect of biomarkers characterized by distinct seasonal variability on the IBR (Fig. 4C and D). During the BEEP project, MT in flounder showed mean levels 80–96% higher in autumn compared to spring at most of the study stations except for Mechelinki (51%) and Walfisch (38%) (Kopecka et al., 2006; Baršienė et al., 2006a; unpublished BEEP data). This was clearly reflected in the IBR, which generally showed higher values in autumn. In this specific example the IBR calculated by using only GH biomarkers indicated that the condition of the flounder populations in most of the areas was better in autumn (lower IBR). However, by introducing the seasonally fluctuating MT into the IBR calculations the seasonal differences were either evened out, or, in cases where no seasonal differences existed, the index now showed a more stressed state in autumn. It is possible that MT levels in flounder were elevated in all study regions because of temporal differences of exposure to (metal) pollution between spring and autumn. However, it is more likely that the observed elevation of MT levels is linked

to endogenous regulation mechanisms and the reproductive stage (George and Olsson, 1994). Temporal variability in the reproductive status of populations from different parts of the Baltic can also be reflected by the MT levels of flounder regardless of actual stress originating from metal levels present in the local environment (Baršienė et al., 2006a; Kopecka et al., 2006). Conclusively, seasonal comparisons are only possible when the biomarkers used in the index calculations are known to be unaffected by season; in other cases direct comparisons between different seasons should be avoided (see also BAI).

In eelpout, data from 2002 showed a higher IBR in autumn compared to spring both in the Wismar Harbour and Salzhaff populations (Fig. 5C). In contrast, no seasonal differences were observed at the "reference" site Kvädöfjärden. As discussed above, considerably higher frequencies of reproductive disturbances were recorded in the German eelpout populations compared to Kvädöfjärden (Gercken et al., 2006), which suggests that the observed seasonal variability of IBR in the eelpout was related to local contamination rather than being a natural phenomenon.

In mussels from Kvädöfjärden and the Lithuanian sites the IBR was slightly elevated in spring compared to autumn 2002, while, in contrast, the values recorded in the Gulf of Gdańsk were markedly higher in autumn (Fig. 6E). Observations at the latter area indicate the effects of changes in the pollution situation since the majority of the biomarkers (LMS, MN and AChE) used for the calculation of the IBR in autumn 2002 showed the highest responses at the Gulf of Gdańsk sites compared to the other areas (Baršienė et al., 2006a; Kopecka et al., 2006; Schiedek et al., 2006).

4.5. Biotic and abiotic factors affecting the integrated responses

The nature and quantity of pollutants as well as their accumulation change during the year. The flounder accumulates contaminants mainly during the intensive feeding period in spring and summer when preparing for reproduction. Therefore, autumn is potentially the season when the strongest toxic effects occur. In dab (*Limanda limanda*), another flatfish species, lipophilic contaminants are partly shifted from the liver into the eggs during gonad maturation (Söffker, 2000), leading to reduced levels of contaminant residues in individuals after reproduction.

Factors related to reproduction and temperature are probably in key position in explaining the seasonal variability observed in several biomarkers and, subsequently, the seasonal differences observed in the integrated stress level indicated by biomarker indices. Therefore, comparisons of integrated stress levels between different areas should always be made using data collected at the same stage of reproduction of each species and during the same season. Eelpout and flounder reproduce in late autumn–winter and, in most cases, an elevated stress level indicated by a higher IBR was assessed in autumn compared to spring in both species. In the Baltic Sea the blue mussel usually reproduces between late spring and mid-summer, sometimes showing a second spawning peak in late summer (Kautsky, 1982). In the mussel and the soft-bottom clam *Macoma balthica* elevated levels of biomarkers (AChE, GST, catalase and MT) and the resulting IBR have been recorded during the reproductive period in the northern Baltic Sea (Leiniö and Lehtonen, 2005), corresponding with the results obtained in the present study.

Natural fluctuations of the hydrological regime in the form of spring floods, storm events, temperature changes, salinity and oxygen conditions may also influence the effects of contamination. According to the BEEP Baltic Sea data (Baršienė et al., 2006a; Kopecka et al., 2006; Schiedek et al., 2006) and data from the more northern part of the Baltic Sea (Gulf of Finland, Archipelago Sea: Leiniö and Lehtonen, 2005; Lehtonen et al., in press) it appears that levels of MT in *Mytilus* spp. are regulated by salinity, with higher mean levels of MT occurring at lower salinity. This is most likely related to the higher bioavailability of metals at low salinities. In this study, concerning some IBR calculations for *Mytilus* spp. made by using a very small number of biomarkers (three), biasing effects of salinity dependent MT on the IBR potentially occur (see Fig. 6D). Within a geographical area characterized by a relatively homogeneous salinity range the problem does not exist; in the Wismar Bay the mussel population at the least polluted site Salzhaff was always clearly distinguished from the two others by having the lowest IBR values. Conclusively, the use of biomarkers known to behave differently under specific local abiotic conditions (such as MT at different salinities) can be included in biomarker index calculations only when comparing areas with similar abiotic conditions.

Finally, the particular lifestyle of the target organisms must be taken into account when interpreting the results. Mussels were always collected from the same sites and also eelpout is considered to be a sedentary fish species (Ojaveer et al., 2004). Oppositely, in the southern Baltic Sea part of the flounder populations perform extensive migrations to reproduce in deeper waters in winter while others stay closer to the coast (Aro, 1989). The different migration behaviour of populations may lead to differences in exposure to contaminants depending on the area that the flounder have recently occupied. Since migration starts in late autumn, sampling in early autumn provides that flounder have stayed at least for more than half a year in the sampling area.

5. Conclusions

Based on the results obtained here, the use of integrated indices describing contaminant-induced stress as management and research tools is considered a useful approach. However, their development into a more "sophisticated" direction, including a scientifically or empirically based weighting of the different biomarkers used in the calcula-

tion of an index as well as the overall selection of biomarkers needs careful attention. Most importantly, it must be well understood that, in all cases, indices like these are oversimplifications of very complex exposure situations prevailing in the field, including combined effects of pollutants and abiotic factors, and also of the multiple physiological responses in the target organisms. Therefore, the results obtained by using integrated indices should never been taken as "face value" but rather as tools to direct further actions in the attempt to resolve causes of the differences observed, whether working on basic research or in environmental management.

Acknowledgements

The authors wish to thank the BEEP Baltic Sea partners (Janina Baršienė, Lars Förlin, Justyna Kopecka, Doris Schiedek, Pekka J. Vuorinen) for the permission to use the original data for the calculation of the biomarker indices. The BEEP Project (Biological Effects of Environmental Pollution in Marine Coastal Ecosystems) received support from the EU (EVK3-CT-2000-00025).

References

Adams, S.M., Brown, A.M., Goede, R.W., 1993. A quantitative health assessment index for rapid evaluation of fish condition in the field. Transactions of the American Fisheries Society 122, 63–73.

Alcorn, S.W., Murra, A.L., Pascho, R.J., 2002. Effects of rearing temperature on immune functions in sockeye salmon (*Oncorhynchus nerka*). Fish and Shellfish Immunology 12, 303–334.

Allen, J.I., Moore, M., 2004. Environmental prognostics: is the current use of biomarkers appropriate for environmental risk evaluation? Marine Environmental Research 58, 227–232.

Aro, E., 1989. A review of fish migration patterns in the Baltic. Rapports et procés-verbal de la réunion—Conseil International pour l'Exploration de la Mer 190, 72–96.

Baršienė, J., Lehtonen, K.K., Broeg, K., Vuorinen, P.J., Lang, T., Pempkowiak, J., Balk, L., Šyvokienė, J., Dedonyte, V., Rybakovas, A., Repečka, R., Köhler, A., Vuontisjärvi, H., Kopecka, J., 2006a. Biomarker responses in mussel (*Mytilus edulis)* and flounder *(Platichthys flesus)* in the Klaipėda-Būtingė area (Baltic Sea). Marine Pollution Bulletin (this volume).

Baršienė, J., Syvokienė, J., Rybakovas, A., Schiedek, D., Kopecka, J., Förlin, L., 2006b. The assessment of cytogenetic and cytotoxic effects in gill cells of the blue mussel *Mytilus edulis* from different zones of the Baltic Sea. Marine Pollution Bulletin (this volume).

Beliaeff, B., Burgeot, T., 2002. Integrated biomarker response (IBR): a useful graphical tool for ecological risk assessment. Environmental Toxicology and Chemistry 21, 1316–1322.

Bly, J.E., Clem, L.W., 1992. Temperature and teleost immune functions. Fish and Shellfish Immunology 2, 159–171.

Broeg, K., 2003. Acid phosphatase activity in liver macrophage aggregates as a marker for pollution induced immunomodulation of the non-specific immune response in fish. Helgoland Marine Research 57, 166–175.

Broeg, K., Zander, S., Diamant, A., Körting, W., Krüner, G., Paperna, I., Westernhagen, H.v., 1999. The use of fish metabolic, pathological and parasitological indices in pollution monitoring I. North Sea. Helgoland Marine Research 53, 171–194.

Broeg, K., Westernhagen, H.V., Zander, S., Körting, W., Koehler, A., 2005. The Bioeffect Assessment Index—a concept for the quantification of effects of marine pollution by an integrated biomarker approach. Marine Pollution Bulletin 50 (5), 495–503.

Chèvre, N., Gagné, F., Blaise, C., 2003a. Development of a biomarker-based index for assessing the ecotoxic potential of aquatic sites. Biomarkers 8, 287–298.

Chèvre, N., Gagné, F., Gagnon, P., Blaise, C., 2003b. Application of rough sets analysis to identify polluted aquatic sites based on a battery of biomarkers: a comparison with classical methods. Chemosphere 51, 13–23.

Coughlan, D.J., Baker, B.K., Cloutman, D.G., Rash, W.M., 1996. Application and modification of the fish health assessment index used for largemouth bass in the Catawba River, North Carolina–South Carolina. American Fisheries Society Symposium 16, 73–84.

Depledge, M.H., 1999. Recovery of ecosystems and their components following exposure to pollution. Journal of Aquatic Ecosystem Stress and Recovery 6, 199–206.

George, S.G., Olsson, P.E., 1994. Metallothioneins as indicators of trace metal pollution. In: Kramer, K.J.M. (Ed.), Biomonitoring of Coastal Waters and Estuaries. CRC Press, Boca Raton, pp. 151–178.

Gercken, J., Förlin, L., Andersson, J., 2006. Developmental disorders in larvae of eelpout (*Zoarces viviparus*) from German and Swedish Baltic coastal waters. Marine Pollution Bulletin (this volume).

Kautsky, N., 1982. Quantitative studies on gonad cycle, fecundity, reproductive output, and recruitment in a Baltic *Mytilus edulis* population. Marine Biology 68, 143–160.

Kelly, J.R., Harwell, M.A., 1989. Indicators of ecosystem response and recovery. In: Levin, S.A., Harwell, M.A., Kelly, J.R., Kimball, K.D. (Eds.), Ecotoxicology: Problems and Approaches. Springer-Verlag, New York, pp. 9–39.

Koehler, A., Wahl, E., Söffker, K., 2002. Functional and morphological changes of lysosomes as prognostic biomarkers of toxic liver injury in a marine flatfish (*Platichthys flesus* L.). Environmental Toxicology and Chemistry 21 (11), 2434–2444.

Kopecka, J., Lehtonen, K.K., Baršienė, J., Broeg, K., Vuorinen, P.J., Gercken, J., Pempkowiak, J., 2006. Measurements of biomarker levels in flounder (*Platichthys flesus*) and blue mussel (*Mytilus trossulus*) from the Gulf of Gdańsk (southern Baltic). Marine Pollution Bulletin (this volume).

Kovacs, T.G., Martel, P.H., Voss, R.H., 2002. Assessing the biological status of fish in a river receiving pulp and paper mill effluents. Environmental Pollution 118 (1), 123–140.

Lang, T., Wosniok, W., Baršienė, J., Broeg, K., Förlin, L., Köhler, A., Kopecka, J., Lehtonen, K.K., Pempkowiak, J., Schiedek, D., Schneider, R., Vuorinen, P.J., in preparation. Multi-biomarker measurements in offshore flounder (*Platichthys flesus*) from the Baltic Sea.

Lehtonen, K.K., Leiniö, S., Schneider, R., Leivuori, M., in press. Biomarkers of pollution effects in the bivalves *Mytilus edulis* and *Macoma balthica* collected from two areas in the southern coast of Finland (Baltic Sea). Marine Ecology Progress Series.

Leiniö, S., Lehtonen, K.K., 2005. Seasonal variability in biomarkers in the bivalves *Mytilus edulis* and *Macoma balthica* from the northern Baltic Sea. Comparative Biochemistry and Physiology C 140, 408–421.

Narbonne, J.F., Daubèze, M., Clérandeau, C., Garrigues, P., 1999. Scale of classification based on biochemical markers in mussels: application to pollution monitoring in European coasts. Biomarkers 4, 415–424.

Ojaveer, H., Eero, M., Lankov, A., 2004. Microevolution of eelpout, *Zoarces viviparus*, in the Baltic Sea. Proceedings of Estonian Academy of Science Biology Ecology 53 (4), 292–305.

Power, M., 1999. Recovery in aquatic ecosystems: an overview of knowledge and needs. Journal of Aquatic Ecosystems Stress and Recovery 6 (3), 253–257.

Schiedek, D., Broeg, K., Baršienė, J., Lehtonen, K.K., Gercken, J., Pfeifer, S., Vuontisjärvi, H., Vuorinen, P.J., Dedonyte, V., Koehler, A., Balk, L., Schneider, R., 2006. Biomarker responses as indication of contaminant effects in blue mussel (*Mytilus edulis*) and female eelpout (*Zoarces viviparus*) from the southwestern Baltic Sea. Marine Pollution Bulletin (this volume).

Schleiger, S.L., 2004. Fish Health Assessment Index study of four reservoirs in West-Central Georgia. North American Journal of Fisheries Management 24 (4), 1173–1188.

Schmidt, V., Zander, S., Körting, W., Broeg, K.v., Westernhagen, H.v., Dizer, H., Hansen, P., Skouras, A., Steinhagen, D., 2003. Parasites of flounder (*Platichthys flesus* L.) from the German Bight, North Sea, and their potential use in biological effects monitoring C. Pollution effects on the parasite community and a comparison to biomarker responses. Helgoland Marine Research 57, 236–251.

Söffker, K., 2000. Saisonalität der Kontamination von Klieschen (*Limanda limanda* (Linnaeus, 1758), Osteichthyes) mit Organochlorverbindungen Bedeutung der Lipide für die Kinetik der Umlagerung von Kontaminanten zwischen verschiedenen Organen. Shaker Verlag, Aachen, Germany, 352 pp.

Stentiford, G.D., Longshaw, M., Lyons, B.P., Jones, G., Green, M., Feist, S.W., 2003. Histopathological biomarkers in estuarine fish species for the assessment of biological effects of contaminants. Marine Environmental Research 55, 137–159.

Steyermark, A.C., Spotila, J.R., Gillette, D., Isseroff, H., 1999. Biomarkers indicate health problems in brown bullheads from the industrialized Schuylkill River, Philadelphia. Transactions of the American Fisheries Society 128 (2), 328–338.

Sturve, J., Berglund, A., Balk, L., Broeg, K., Böhmert, B., Massey, S., Savva, D., Parkkonen, J., Stephensen, E., Koehler, A., Förlin, L., 2005. Effects of dredging in Göteborg harbour, Sweden, assessed by biomarkers in eelpout (*Zoarces viviparus*). Environmental Toxicology and Chemistry 24 (8), 1951–1961.

Tatner, M.F., Manning, M.J., 1985. The ontogenic development of the reticulo-endothelial system in the rainbow trout, *Salmo gairdneri* Richardson. Journal of Fish Diseases 8 (2), 189–195.

Vuorinen, P.J., Keinänen, M., Vuontisjärvi, H., Baršienė, J., Broeg, K., Förlin, L., Gercken, J., Kopecka, J., Köhler, A., Parkkonen, J., Pempkowiak, J., Schiedek, D., 2006. Use of biliary PAH metabolites as a biomarker of pollution in fish from the Baltic Sea. Marine Pollution Bulletin (this volume).