Diversity and distribution of benthic macrofauna in the Baltic Sea
Data inventory and its use for species distribution modelling and prediction

Mayya Gogina, Michael L. Zettler *

Leibniz Institute for Baltic Sea Research, Seestrasse 15, D-18119 Rostock, Germany

ABSTRACT

This study attempts to model the probability of occurrence of some characteristic macro-invertebrate species of the Baltic Sea from different functional groups (i.e. grazers, deposit and suspension feeders, and predators) in response to major environmental forcing factors (salinity, water depth and seabed substrate type). Analyses were based on the inventory data set compiled by revising the data on macrobenthic species for over 12,000 sampling events in the Baltic Sea. In addition, data on environmental variables are retrieved from the results of modelling and large-scale mapping efforts. A simple logistic regression based modelling technique was applied and the candidate model with highest discriminatory power was selected for habitat suitability mapping. Habitat suitability models allowed to satisfactorily predict the potential distribution of macrofaunal species based solely on modelled salinity, bathymetry and rough sediment class information. Our results indicated that salinity, depth and substrate type are all important in determining the distribution of most characteristic macrobenthic species on the large-scale of the whole Baltic Sea. The present exercise is only a first step. Implementation of other variables (e.g. characterizing oxygen and temperature fluctuations, total organic content, and nutrient supply) would obviously increase the model applicability. Information on the ecological potential of habitat suitability can serve as the utmost important basis for scientifically sound marine spatial planning.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The Baltic Sea is amongst the world's largest seas isolated from the major oceans and therefore it contains unique communities of marine lifeforms, including benthic invertebrates (Leppäkoski et al., 2009; Zettler et al., 2008).

As consumers at intermediate trophic levels, macro-invertebrates are essential agents of both bottom-up and top-down forces in the water system (i.e. their populations are both regulated by resource limitation and predation, and they themselves represent resources and consumers for other levels of the food chain). Benthic macrofauna can have an important influence on nutrient cycles, primary productivity, decomposition, and translocation of materials (Wallace and Webster, 1996). The presence or activities of invertebrate species in aquatic ecosystems often alters the physical surrounding or the flow of recourses, thereby creating or modifying habitats, which then influence all other organisms in the community. This determines these organisms as ecosystem engineers. Crain and Bertness (2006) argued that most natural communities are hierarchically structured with ecosystem engineers able to modify the habitat providing the physical template of communities. Which engineers are important for maintaining the ecosystem functions of interest is dictated by the background environment and the limiting variables.

Human disturbances which impact soft-sediment habitats act on various spatial and temporal scales. For instance, bottom trawling has effect on spatial scales ranging from micrometers to many hundred kilometers (Gray and Elliott, 2009). Such marine activities as wind farms construction, stringing of pipelines and cables across the sea floor can be just as serious, and their spatial planning requires an assessment of the potential anthropogenic impact as well as baseline maps (Degraer et al. 2008). It was reported that relative importance of factors influencing succession varies over spatial extents. While biotic interactions are most important on the meter scope, environmental conditions dominate on a more global scale (Zajac et al., 1998). The environmental factors usually named as controlling for benthos distribution are food supply, water salinity, oxygen concentrations, currents, temperature, turbidity, substrate composition, sedimentation rates and bathymetry (e.g. Bromley, 1996; Olenin, 1997; Coleman et al., 2007). Since only very few species have been studied in detail in terms of their dynamic responses to environmental change, static distribution modelling often remains the only approach for studying the possible consequences (Woodward and Cramer, 1996; Guisan and Zimmermann, 2000).

A number of previous works have investigated the structuring factors for spatial distribution of benthic communities focusing on different spatial extends. For instance, Zettler and Bick (1996) in their study on small scale dispersion patterns concluded biological interactions to be the critical factors for the fine-scale (ca. 5 × 5 m)

* Corresponding author. Department of Biological Oceanography, Baltic Sea Research Institute (IOW), Seestrasse 15, D-18119 Rostock, Germany. Tel.: +49 381 5197 236; fax: +49 381 5197 440.
E-mail address: michael.zettler@io-warnemuende.de (M.L. Zettler).

variations in distribution of a polychaete. In the regional study in the Mecklenburg Bight Gogina et al. (2010) have found water depth to be the key factor determining the species distribution among 8 hydro-logical and sedimentological parameters considered. Works of Olenin (1997), Laine (2003), Bonsdorff (2006), Zettler et al. (2008) and Ojaveer et al. (submitted for publication), performed on a more global scale stressed that salinity together with oxygen concentrations define the Baltic Sea diversity and best explained the patterns in community distribution.

The present paper aims to contribute to such essential ecological issues as estimations of potential response of biota to habitat changes (natural or anthropogenic) and the generation of full coverage maps predicting the suitable areas for distribution of macrobenthic species within the Baltic Sea. Goals of the study included (1) the compilation of an extensive list of taxa and an inventory dataset on species distribution for the whole Baltic Sea, (2) extraction of patterns of species distribution in response to selected environmental parameters (salinity, depth, substrate type) and (3) modelling and mapping the probabilities of occurrence for exemplary species in response to those factors. Finally, we discuss our results as a baseline for creating a tool for modelling benthic community changes and its effects on the functioning of the ecosystem and last but not least for developing an instrument for conservation and management purposes.

2. Materials and methods

2.1. Study area

Barotropic pressure gradients, river runoff and meteorological forces control the hydrodynamical system of the Baltic Sea (Fennel, 1995; Feistel et al., 2008). It features salinity ranging from 25–30 psu in the Danish Sounds region to 1–3 psu or even lower in the northern Gulfs of Bothnia and the inshore lagoons. Climatic conditions range from temperate to boreal zone. Persistent pycnocline in the Baltic basin causes stagnation of bottom waters for long periods, with periodic hypoxia/anoxia as a consequence, over an area covering up to 100,000 km^2 (Bonsdorff and Pearson, 1999). Intense salinity stratifi-cation and convective mixing during the cooling period (generally winter and the beginning of spring) is restricted to the upper 60–80 m, below this level temperature is fairly stable and can be approximated to 5.5 °C found throughout the year in the deep basins. Large seasonal temperature variation at the surface in combination with the low surface salinity results in regular ice formation (e.g. Feistel et al., 2008; HELCOM, 2009).

The highest overall Baltic Sea diversity and number of benthic macroscopic species can be found in the south-western region, that is most influenced by marine conditions (Zettler et al., 2008). Generally, the number of species declines to the north resulting in the low-diversity communities, for example, in the Gulf of Bothnia (Ojaveer et al, (submitted for publication). In recent decades, eutrophication and pollution have significantly affected the biota of the Baltic Sea. The entire soft bottom infauna is defined as poor in terms of species composition, and functional complexity is considered to be low (Andersin et al, 1978; Laine, 2003; Bonsdorff and Pearson, 1999). The primary reason for the low-diversity is that very few species are endemic to brackish conditions, as both marine and limnic species meet their physiological limits (HELCOM, 2009).

2.2. Data inventory

Generally, there is a lack of data sets that would simultaneously fulfill the two important demands of rigorous assessment and analysis of the of Baltic Sea's macrozoobenthic biodiversity: the sufficiency of spatial cover and density of sampling points and its interior comparability, or at least homogeneity of taxonomic nomenclature. A compatible dataset covering the whole Baltic was required for our aims. Various literature sources, including the historical data from Knipowitsch (1909) and Hessle (1924), the data on distribution of species in the south-western Baltic Sea combined in Gerlach (2000) and Zettler and Röhner (2004), and available databases, including "Autecological Atlas" of the IfAÖ (2007), HELCOM monitoring data collected in the ICES-Database (ICES EcoSystemData), Baltic Sea Alien Species Database (Olenin et al., 2009), data obtained by the IOW monitoring and various research programmes, were analysed with respect to information on macrozoobenthos distribution in the Baltic Sea. Most latter sources provided the data collected by macrobenthic surveys following the guidelines of HELCOM (2008).

All macrofauna species were identified to the lowest taxonomic level possible. The nomenclature was checked and revised following the World Register of Marine Species (SMEBD, 2009), but also the BioLib taxonomic data (Zicha, 1999–2009) in case of freshwater species. This formed an inventory list of about 11 hundred macrozoobenthic taxa grounded on valid taxonomy and synonymy published by the co-author in Ojaveer et al, (submitted for publication).

Revised data on species occurrence within the defined Baltic Sea sub-regions were compiled together in GIS (software ArcGIS 9.1, ESRI, 2003) including over 160,000 entries (12,200 stations, i.e. sampling events; Fig. 1a), sampled from 1839 to 2009.

2.3. Environmental data

To maximize the applicability of the habitat suitability model, only widely available environmental variables were offered in the modelling exercise. Depth represents an indirect variable replacing a combination of different recourses and direct gradients — a primary descriptor of other environmental factors (e.g. food quality and food availability, light penetration). To compile the digital elevation model the bathymetry dataset from Seifert et al. (2001) was used (Fig. 1b). Data for near-bottom salinity (Fig. 1c) averaged for 45 years to smooth the variance are the result of dynamical 3D modelling and the methodology for producing it is documented by Neumann and Schernewski (2008). The only available data on seabed sediments covering the whole Baltic Sea region was produced by the EU-BALANCE project (Al-Hamdani and Reker, 2007). It is the categorical data represented by 5 classes (1 — bedrock, 2 — complex sediments, 3 — sand, 4 — hard clay, and 5 — mud and clay) with resolution of 200 m (Fig. 1d). The transformation of this categorical data into numerical data was not applied, as for instance was done in Meissner and Darr (2009). The reason for this is that no unambiguous correlation was found between the substrate classes and numerical characteristics of sediments (e.g. median grain size and total organic content analysed in Gogina et al., 2010). To retain the resolution of substrate data for the analysis, the other environmental data (salinity and bathymetry) were subject to resampling using nearest neighbor assignment that does not change any of the values of cells from the input layer. Thus, values of 3 environmental variables, assumed to generally controlled species distribution, were obtained for each of the grid cells 200 × 200 m.

2.4. Exemplary species

For this exercise 19 species from various functional groups representative of both marine and limnic communities, relatively abundant and well represented by the data, were chosen (Table 1). Here their feeding types, preferences in substrate and salinity and possible penetration depth according to the literature sources are listed. The frequency of occurrence ranging from 6 to 48% is indicated within the dataset.

2.5. Modelling technique

The process of construction of habitat suitability models included the following steps.
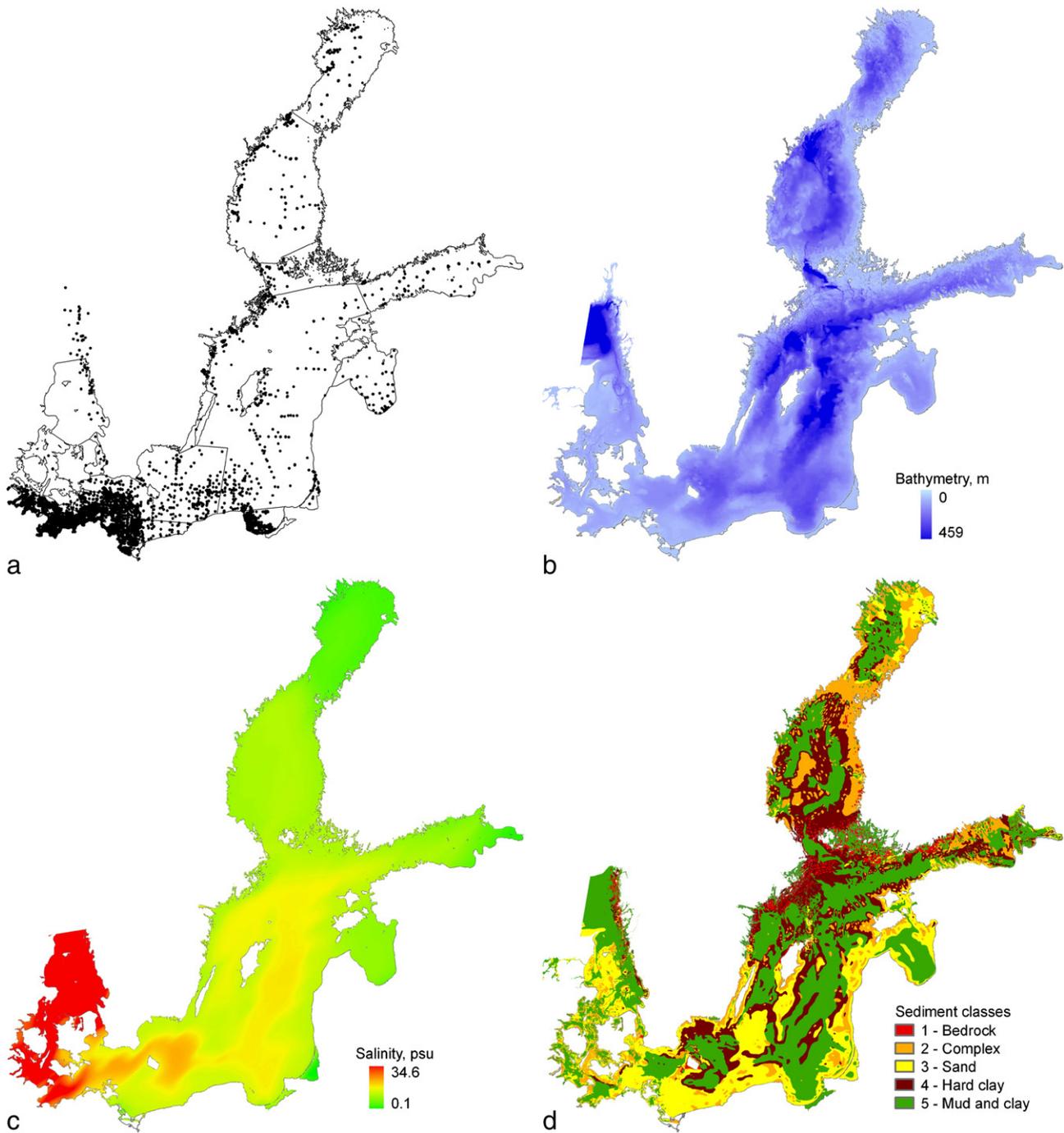
Fig. 1. Stations sampled for macrofauna within the Baltic Sea sub-regions are indicated by dots (a) and distribution surfaces of the environmental variables considered covering the Baltic Sea: bathymetry (b), near-bottom salinity (c), seabed sediment type (d). Geographical data ESRI (2003); projection UTM on WGS84.

First, biotic data was reduced to presence/absence. We assumed that our data contains reliable "true absences" *sensu* Wisz and Guisan (2009), as species we are considering can be regarded as discriminatory, common and relatively well known for the Baltic Sea. The latter promotes its presence in the sample if the animal is present on the site as well as its identification.

Then, in order to be able to evaluate the model performance, data was randomly separated into the calibration and the evaluation datasets (each containing about a half of initial sites). This solution is suggested, for example by Guisan and Zimmermann (2000).

Binary logistic regression was chosen as an appropriate technique. This branch of GLM is classic for binary response. It makes no assump-

tions about the data distributions, including the possibility to use categorical predictors, and allows predicting the probability of observing the species (rather than predicting binary presence–absence).

The logistic regression model relates the probability of observing the species p to one or more predictor variables x (Legendre and Legendre, 1998; Ysebaert et al. 2002; Wisz and Guisan, 2009). The form of preliminary model can be given as follows: $p(x) = \gamma_b / (1 + \gamma_b)$, where $\gamma_b = e^{(b_0 + b_1 S + b_2 S^2 + b_3 D + b_4 D^2 + b_5 Sed_1 + b_6 Sed_2 + b_7 Sed_3 + b_8 Sed_4)}$. It includes simple polynomial response for (S) salinity, (D) water depth and substrate classes (Sed) coded as 4-level categorical variable; b_i are the regression parameters. They are estimated by maximum likelihood, assuming a binomially distributed error term.

Table 1

List of exemplary species, their habitat preferences according to the authors experience (some key references are also indicated in the footnote), median (P_{50}) and 99th percentile (P_{99}) of abundance density (ind/m^2) and frequency of occurrence (Freq., in %) in the analysed data set.

nn	Taxon and author	Feeding type	Substrate type	Penetration depth	Salinity (psu)	P_{50}	P_{99}	Freq.
Crustacea[*]								
1	*Bathyporeia pilosa* Lindström, 1855	Grazing	Fine sands	0–3 cm	7.3–14.7	11	5431	7.3
2	*Corophium volutator* (Pallas, 1766)	Deposit/suspension feeding	Muddy sands	2–5 cm	5–35	28	11,157	10.7
3	*Diastylis rathkei* (Kröyer, 1841)	Deposit feeding	Muddy sands and mud	1–5 cm	7.7–30.3	49	3970	34
4	*Pontoporeia affinis* Lindström, 1855	Deposit feeding	Mud to sand	0–5 cm	0–10	141	10,163	14.3
5	*Pontoporeia femorata* Krøyer, 1842	Deposit feeding	Mud to sand	0–5 cm	11.5–30.3	20	2157	12.3
6	*Saduria entomon* (Linnaeus, 1758)	Predation	Mud to sand, complex	0–10 cm	3–13	10	168	13.7
Mollusca[**]								
7	*Arctica islandica* (Linnaeus, 1767)	Suspension feeding	Mud to sand	0–14 cm	15–31	21	409	21.2
8	*Astarte borealis* (Schumacher, 1817)	Suspension feeding	Mud to sand	0–1 cm	15.8–40	27	969	13.1
9	*Hydrobia ulvae* (Pennant, 1777)	Grazing/deposit feeding	Mud to sand	0–1 cm	10–33	267	18,762	24.9
10	*Hydrobia ventrosa* (Montagu, 1803)	Grazing/deposit feeding	Mud to sand	0–1 cm	6–20	70	37,664	6.2
11	*Macoma balthica* (Linnaeus, 1758)	Deposit/suspension feeding	Mud to sand	5–6 cm	4.6–30.3	90	3438	48.4
12	*Mya arenaria* Linnaeus, 1758	Suspension feeding	Fine–medium sands	Up to 40 cm	7.3–30.3	40	5500	25.8
13	*Mytilus edulis* Linnaeus, 1758	Suspension feeding	Mud to boulders	0 cm	6.8–30.3	42	55,000	34.5
Polychaeta[***]								
14	*Heteromastus filiformis* (Claparède, 1864)	Deposit feeding	Mud to sand	Up to 30 cm	15–30.3	34	4040	18.4
15	*Lagis koreni* Malmgren, 1866	Deposit feeding	Muddy sands	0–10 cm	15–30.3	30	1751	16.4
16	*Pygospio elegans* Claparède, 1863	Deposit/suspension feeding	Fine–medium sands	4–6 cm	7.2–29.3	121	8640	28.9
17	*Scoloplos armiger* (Müller, 1776)	Deposit feeding	Muddy sands and mud	5–15 cm	11.5–32.4	84	1657	30.7
18	*Terebellides stroemii* Sars, 1835	Deposit feeding	Mud to gravel	0 cm	35–10	27	1351	16.7
Priapulida[****]								
19	*Halicryptus spinulosus* von Siebold, 1849	Deposit feeding, predation	Muddy sands and mud	1–6 cm	6.8–21.3	11	249	21.5

   [*] Schulz, 1969; Ankar, 1977; Fenchel et al., 1975.
  [**] Schulz, 1969; Ankar, 1977; Fenchel et al., 1975.
 [***] Schulz, 1969; Fauchald and Jumars, 1979.
[****] Schulz, 1969; Aarnio et al, 1998.

We have tested various combinations of predictors and the functional forms of the relationships (i.e. special cases of the linear predictor, specified above). The optimal model selection was based on information theoretic approach after (Burnham and Anderson, 2004) relying on the calculation of the Akaike Information Criteria. Models with the lowest AIC value within a set strike the best balance between bias and variance of model prediction and provide the least information loss when approximating the truth.

We tested the models' discriminatory power using calculations of the Area Under the ROC Curve (a so-called AUC) of a Receiver Operating Characteristic Plot on the evaluation data not used for models building. AUC is an appropriate metric for evaluating classification accuracy because it estimates the percentage of locations where the species is observed to be present that are expected to have a higher predicted probability of occurrence than places where the species are absent (Fielding and Bell, 1997). Moreover, it is a threshold independent metric, which means it assesses classification accuracy across the entire range of predicted probabilities, and not just for a specified probability threshold. Hosmer and Lemeshov (2000) suggest the following interpretation of AUC values when evaluating the model's discriminatory ability: AUC > 0.9 outstanding, 0.8 < AUC < 0.9 excellent, 0.7 < AUC < 0.8 acceptable, AUC ≤ 0.5 no discrimination. Assigning a threshold value for dichotomising the occurrence probabilities to presence–absence predictions is conditional to prior information about species prevalence (e.g. Strauss and Biedermann, 2007). The threshold-dependent metrics used to additionally assess model discrimination were sensitivity (true positive rate), specificity (true negative rate) and correct classification rate, evaluated for the test data (not used for model building), given for the defined threshold.

Using the defined method, probabilities of species occurrence were modelled and mapped. All analysis were carried out using SPSS (SPSS, Inc.), Statistica (StatSoft Inc., 2007), MATLAB and ArcMap (ESRI Inc., Redlands, USA).

3. Results

The 19 macrozoobenthic species selected for this exercise are all among the most dominant and representative benthic macro-invertebrates of the Baltic Sea, but are representatives of various functional groups (Table 1). Thus, their habitat preferences varied considerably (Fig. 2). Box-and-whisker plots were used to illustrate the observed distribution of the exemplary species along the ranges of depth and salinity. The number of sampled locations within each of the substrate classes varied drastically (96, 1675, 4288, 729 and 4694 stations within class 1, 2, 3, 4 and 5, respectively). Thus to visualize the patterns of occurrence of exemplary species over the substrate classes the fraction of stations where the particular species is found in the total number of stations sampled within the class was derived and plotted.

Differences between species were observed in regard to their relative frequency of occurrence and density. The bivalve *Macoma balthica* was the most common species. The gastropod *Hydrobia ventrosa* and the amphipods *Bathyporeia pilosa* and *Corophium volutator* occurred in the most moderate percentage of samples of the data set, still exceeded the overall 5% and thus cannot be assigned to the uncommon. In terms of density, *Hydrobia ulvae* indicated the highest median value of abundance at occupied locations, whereas the highest absolute value of density (P_{99} was considered to adjust the outlier observations) is featured by *Mytilus edulis*. Yet, as only a few species are considered here, no obvious general differences can be seen for these parameters between the various functional groups.

Additionally to Fig. 2, changes of species densities were plotted against the abiotic factors using scatter plots (not presented here for brevity). Both graphical outputs confirm that the crustaceans *Pontoporeia affinis*, *Pontoporeia femorata* and *Saduria entomon* are the species that are most tolerant to the increase of water depth, with their highest frequencies and associated densities observed between 40 and 60 m. Oppositely, distribution of *B. pilosa* and *C. volutator* and,
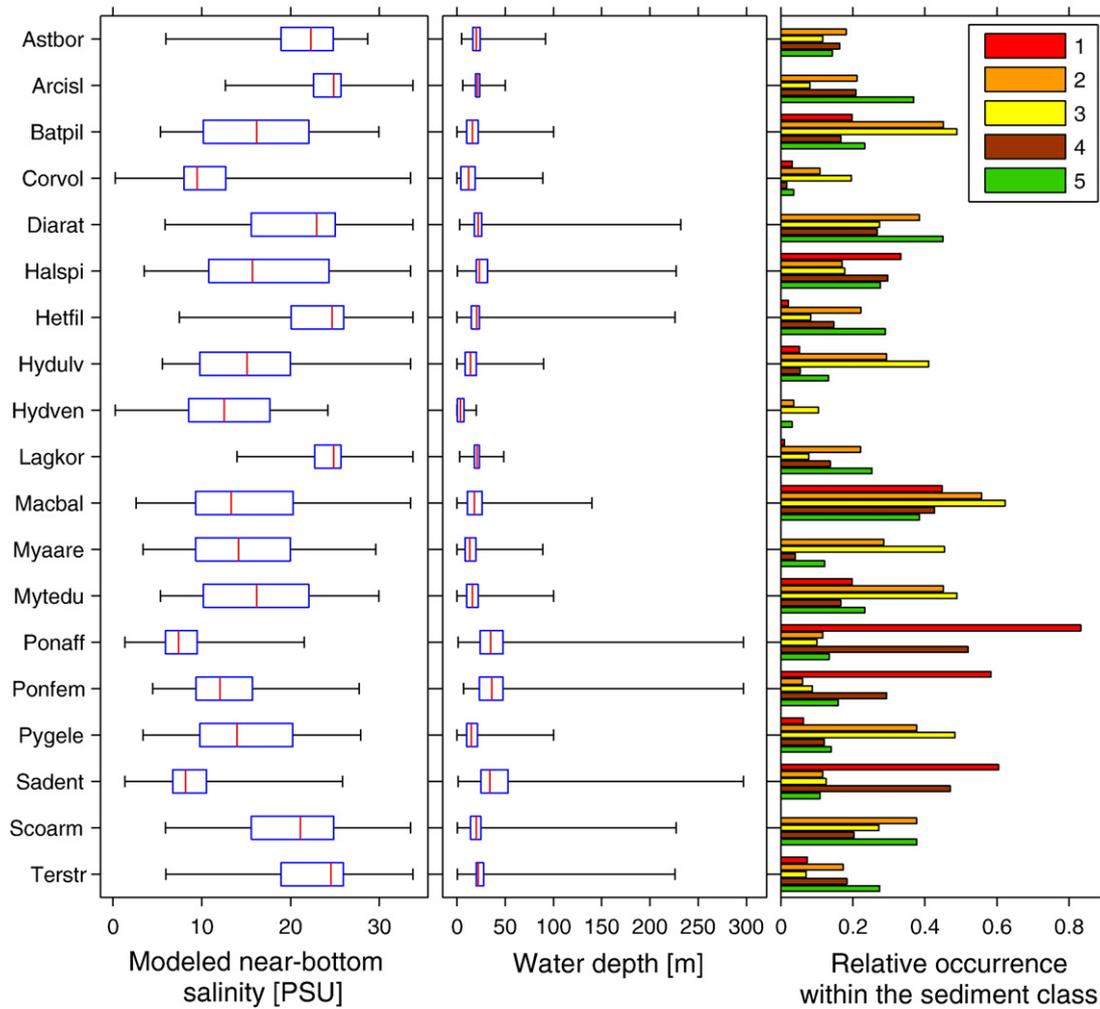
Fig. 2. Occurrence of 19 exemplary species along the ranges of investigated environmental parameters. Species are ordered alphabetically, 6-letter codes from top to the bottom referring to *Astarte borealis, Arctica islandica, Bathyporeia pilosa, Corophium volutator, Diastylis rathkei, Halicryptus spinulosus, Heteromastus filiformis, Hydrobia ulvae, Hydrobia ventrosa, Lagis koreni, Macoma balthica, Mya arenaria, Mytilus edulis, Pontoporeia affinis, Pontoporeia femorata, Pygospio elegans, Saduria entomon, Scoloplos armiger,* and *Terebellides stroemi.* The tops and bottoms of each "box" are the 25th and 75th percentiles of the samples, respectively. The line in the middle of each box is the sample median. The "whiskers" extending to the left and the right of each box represent minimum and maximum value of the abiotic parameter corresponding to the occurrence of species. Sediment classes are numbered as in Fig. 1d.

especially, *H. ventrosa* is most strongly constrained by the factor, with highest densities observed in regions shallower then 10 m and animals hardly ever present in samples from below 30 m isobaths (15 m for *H. ventrosa*).

In terms of near-bottom salinity, *P. affinis* was the most evident limnic species, with highest densities observed between 5 and 9 psu, and only rear events of its occurrence at stations with salinity values above 10 psu. *C. volutator* and *S. entomon* also favour the common salinity range with highest abundances recorded at this part of the gradient; however, their occurrence at higher values of salinity is not unlikely. Species as *Terebellides stroemi* and *Diastylis rathkei*, indicate the opposite behavior with wide range of suitable salinities (approximately 10 to 30 psu) and are most abundant in the most saline regions. Other species as *M. balthica* and *H. spinulosus*, though also recorded at station with salinity values from ca. 4 to 27 psu, show the peaks of abundance density near the lower end of this range.

As for substrate preferences of the exemplary species, most of them seem to avoid the areas where seabed substrate is represented by the bedrock class, with the evident exceptions of *P. affinis*, *P. femorata* and *S. entomon*. Generally, this type of substrate is the most infrequent for the Baltic Sea (Fig. 1d), particularly, it is exposed in many areas of the Åland Sea and Archipelago Sea, where it shows traces of the last ice age; all of the three species are glacial relicts

that in terms of both abundance and occurrence favour very fine sediments. *M. balthica* and the priapulid *Halicryptus spinulosus* also indicate their occurrence on bedrock substrates. *M. balthica* seem generally to have a high tolerance for variations of substrate types. As for *H. spinulosus* it is important to note that is rather a data artefact. This species is known to be a mud/clay dweller and its tolerance to rocky beds is highly unlikely. Apparently, the rough substrate map available for this study does not reflect some isolated muddy patches that can occur within the areas where bedrock is prevailing, especially in the regions of Stockholm archipelago and the Archipelago Sea.

Habitat models were developed for all the investigated species. For most species the final model with greatest discriminatory power included all three environmental factors considered, with the polynomial response for salinity and depth and the 4-level categorical assignment to one of the substrate classes. The exceptions were the habitat models for *P. affinis* and *C. volutator*, where the quadratic term for salinity was eliminated, as well as the final models for the cumacea *D. rathkei* and the pectinariid *Lagis koreni* that did not account for the substrate type factor. Both latter marine species are known to prefer muddy sands. The habitat suitability models constructed for them indicated relatively high AUC values (0.83 and 0.9, respectively).

The examples of produced maps with probabilities of occurrence derived using the logistic regression models based on three

environmental factors are shown in Figs. 3 and 4. Symbols indicate the observed species abundance with their size corresponding to the value of the abundance density. Color fields show the modelled probability of occurrence.

On the Fig. 3, left the results are shown for the glacial limnic relict amphipod *P. affinis*. According to Table 2, AUC value estimated on the data assumed to be independent, that is not used for model calibration, indicates very high degree of success. On the right the results are mapped for the bivalve *M. balthica* found almost all over the Baltic Sea and known to be very tolerant in response to environmental gradients. In this case AUC value indicates the predictive power that can be regarded as only fair. Nevertheless, the consistency with data observations can be seen even here. The considered environmental predictors are significant, but noticeable part of the variance of species distribution is explained by some other factors not included in the analysis.

Fig. 4 shows the results of prediction for the priapulid *H. spinulosus* and the bivalve *Arctica islandica* both species are not as abundant and frequent as the previous ones. For *H. spinulosus* model AUC value was relatively low. Whereas, the combination of the factors considered seems to be suitable in the area north to the land Sea, the observational data evidences the barrier of modelling here, and the existence of some other limiting factor preventing the species from inhabiting the territory. This can supposedly be variations in ice cover, temperature or nutrient supply. Also, the locally bad performance of the model for this species is most probably caused by the data artefact (the roughness of substrate type data results in the indication of the occurrence of species on bedrock substrates, what in turn fully contradicts with the autecological preferences known for the species). The probability of occurrence modelled for *A. islandica* known to prefer polyhaline environment corresponds with the observations very well. It can be also seen here that areas with higher predicted values match with the locations of higher observed abundance densities.

Thus, habitat suitability maps predict the specific ecological potential of a habitat rather that a realized ecological structure, with limitations defined by the data analysed.

4. Discussion and conclusions

Physical limits of species distribution are caused by environmental and physiological constrains, e.g. many zoobenthic species exploit the physical characteristics of the environment to obtain their foods, survival of larvae is conditional on hydrographical variables, etc. (Guisan and Zimmermann, 2000; Gray and Elliott, 2009).

On different scales various factors take the leading stand in influencing species distribution. Biotic factors are expected to be the dominating ones on the more local scale. The spatial scale of this exercise can be described as fine-grained (referring to the individual data units of grab samples), whereas spatial extent covered can be defined as large (377,000 km²). On a large-scale, e.g. Baltic-wide, salinity is likely to be the primary (or even the only) descriptor in species–environmental relationships (Laine, 2003; Bonsdorff, 2006; Ojaveer et al, (submitted for publication)).

The core of predictive geographical modelling is represented by the quantification of species–environment relationships (Guisan and Zimmermann, 2000). In the present study we have tested the discriminating ability of such factors as salinity, depth (as indirect factor replacing a combination of different recourses and direct gradients — a primary descriptor for other abiotic factors) and substrate characteristics (considered only generally due to the absence of more detailed data available to the authors at the time of carrying out the analysis) on a Baltic Sea-wide scale to explain the occurrence of typical macrozoobenthic species. Presented models are based on the hypothesis of species distribution being generally controlled by the environmental factors analysed. The environmental variables analysed in this study represent both direct and indirect ecological factors, yet, we expect the measured processes to be a constraint on but not the sole determinant of the benthic organisms' responses (Cade and Dong, 2008). For instance, numerous factors, including the existence of a circulatory system, diffusion distances, temperature, degree of locomotor activity, effect of dissolved oxygen levels on the persistence and bioavailability of some chemicals, ability to regulate external respiration and the existence of respiratory pigments determine the dependency of
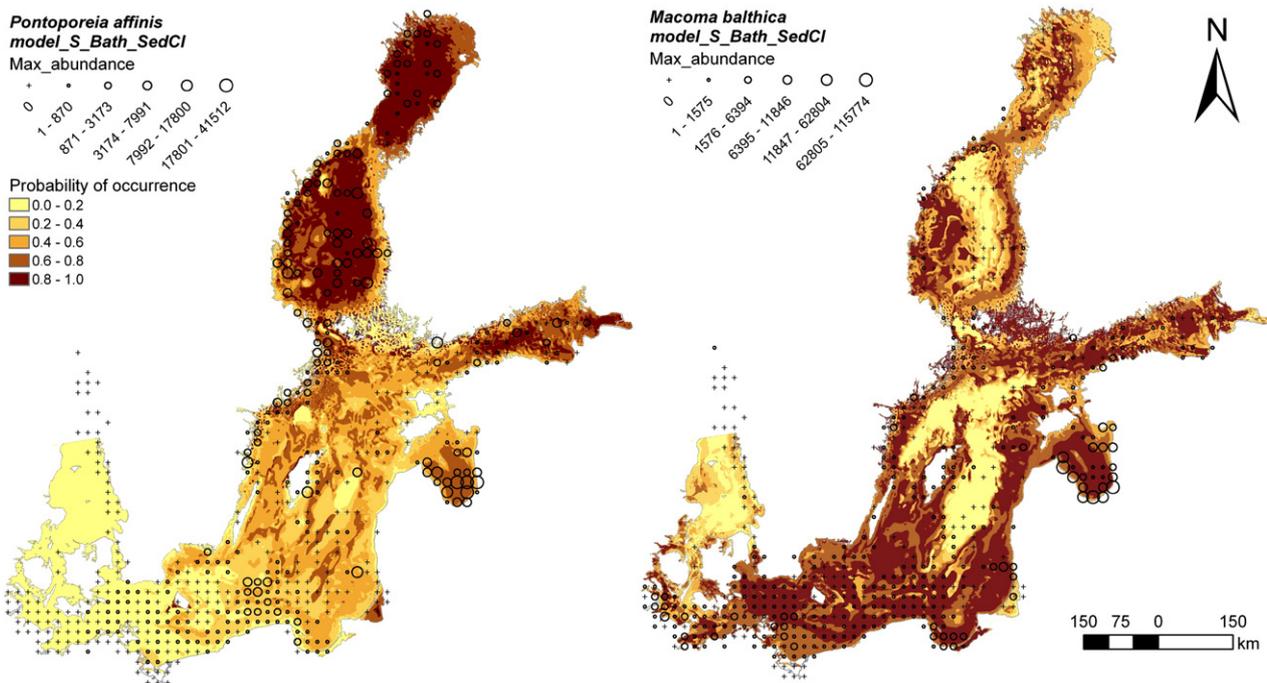

Fig. 3. Modelling results compared with the observed data for the amphipod *Pontoporeia affinis* (left) and the bivalve *Macoma balthica* (right). Symbols indicate the observed species abundance (ind/m²) with their size corresponding to the value of abundance density (the maximal value observed for each cell of a regular 20 km grid is plotted in order to simplify visualization). Color fields show the modelled probability of occurrence.
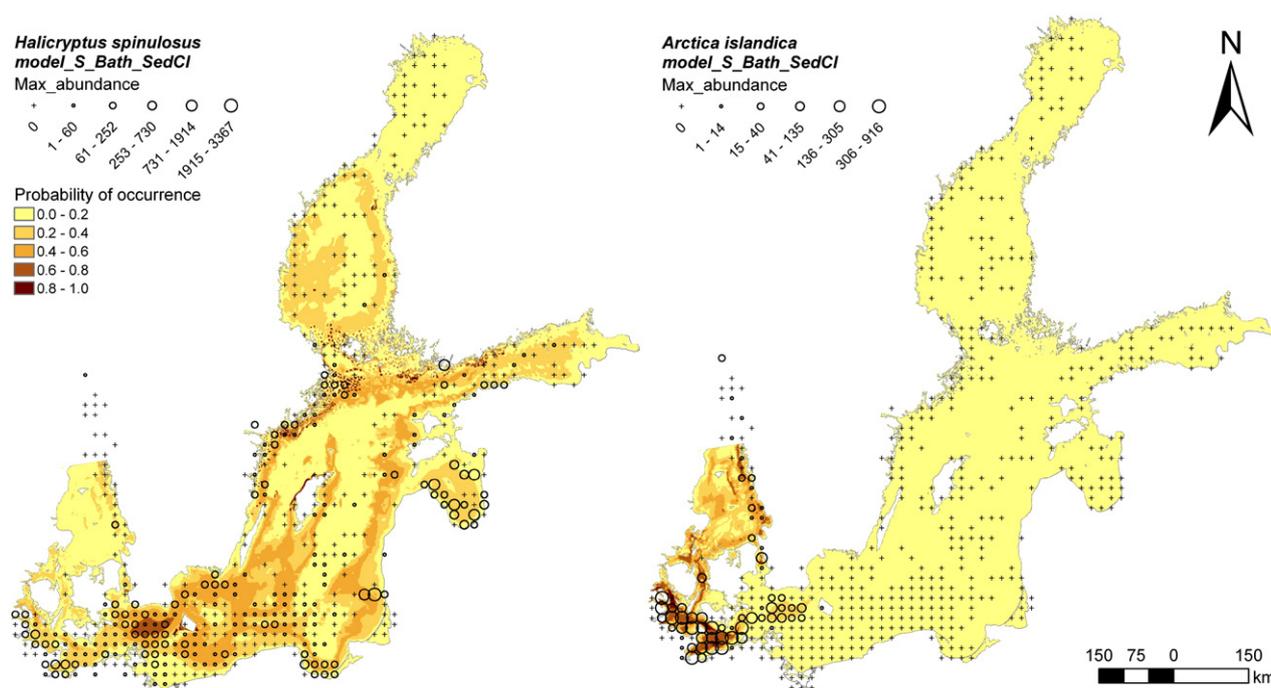
Fig. 4. Modelling results compared with the observed data for the priapulid *Halicryptus spinulosus* (left) and the bivalve *Arctica islandica* (right). Symbols indicate the observed species abundance (ind/m^2) with their size corresponding to the value of abundance density (the maximal value observed for each cell of a regular 20 km grid is plotted in order to simplify visualization). Color fields show the modelled probability of occurrence.

marine and estuarine invertebrates on oxygen; there are dramatic differences in need for this recourse between macrobenthic species (e.g. Davis, 1975). Often water depth is regarded as a proxy (indirect and integrated) for near-bottom oxygen conditions and distribution of oxygen depletion events. Oxygen regime in the Baltic Sea is depth-dependant, yet oxygenation of sub-halocline depends also and mainly on the advection of Kattegat waters (e.g. Fleischer and Zettler, 2009; Olenin, 1997). Prolonged changes in oxygen conditions result in modification of local community structure, with intolerant of depressed oxygen species abandoning the environment (mobile) or die (sessile),

Table 2

Comparison of modelling results concerning discrimination and predictive accuracy of developed models for the 19 macrobenthic species. AUC = area under the receiver operating curve, CI = confidence interval, calculated for the evaluation data not used for models building; CCR = correct classification rate in %, Sens. = sensitivity (true positive rate) in %, Spec. = specificity (true negative rate) in %, given for the arbitrary defined threshold value (P). Species notation as in Fig. 2.

	AUC (95% CI)	P	CCR	Sens.	Spec.
Arcisl	0.917 (0.910–0.924)	0.50	85.6	76.6	88.5
Astbor	0.826 (0.815–0.838)	0.20	73.2	80.6	71.9
Batpil	0.816 (0.799–0.833)	0.15	78.0	63.4	79.2
Corvol	0.810 (0.803–0.836)	0.15	78.7	72.0	79.4
Diarat	0.829 (0.818–0.840)	0.50	77.3	73.8	79.4
Halspi	0.747 (0.733–0.761)	0.30	71.7	58.4	75.9
Hetfil	0.848 (0.837–0.859)	0.30	76.3	70.1	77.8
Hydulv	0.811 (0.799–0.823)	0.40	75.9	62.4	80.5
Hydven	0.902 (0.891–0.914)	0.15	87.9	73.8	88.7
Lagkor	0.900 (0.893–0.908)	0.40	84.4	81.7	85.0
Macbal	0.746 (0.734–0.759)	0.50	67.4	72.2	62.4
Myaare	0.809 (0.798–0.821)	0.40	75.2	59.1	81.1
Mytedu	0.783 (0.771–0.795)	0.40	71.2	71.4	71.0
Ponaff	0.942 (0.935–0.949)	0.16	87.4	92.1	82.8
Ponfem	0.772 (0.754–0.790)	0.15	79.7	63.1	82.4
Pygele	0.790 (0.778–0.802)	0.50	72.5	62.5	76.9
Sadent	0.857 (0.841–0.874)	0.20	77.8	79.6	76.0
Scoarm	0.786 (0.774–0.798)	0.40	72.3	78.6	65.7
Terstr	0.836 (0.825–0.847)	0.30	77.7	65.4	80.0

inhabiting the territory tolerant species survive, or diversity drop to zero before the area is recolonized by the species able to stand the present conditions. Thus, the prior duration, frequency and regularity of hypoxia/anoxia are of matter for the current state of the macrobenthic community (Karlson et al., 2002). Yet, the complexity and variability of oxygen dynamics in the Baltic Sea causes difficulties for inclusion of this factor in the model. In further work authors hope to overcome the challenge of defining an appropriate variable (or a set of variables) to cover the inter-annual and seasonal variations and fulfil the frameworks of physics, autecology and mathematics involved in the generation of habitat suitability models. Also, benthic environments are at the receiving end of the accumulation and burial of organic matter, and models for many species would benefit from incorporating the corresponding variable; however the sufficient data was not yet available for this exercise.

Thus our results indicate the habitat suitability defined by the factors analysed under the conditions, when hypoxia does not play a limiting role. We believe that if such "oxygen-rich conditions" would last for sufficiently long period in the deepest regions of the sea (e.g. regions where long-lasting hypoxia plays a significant role in determining the spread of benthic species), the species under interest would be able to colonise the areas that are suitable for them in terms of the factors accounted by our study.

As a potential field of application for similar models Davies et al. (2000) suggests creating a model for prediction of local habitat features that are expected to occur at a site in the absence of the effects of human activities using large-scale variables. A list of habitat features that are expected at a site can farther provide targets for habitat restoration or enhancement. Moreover, when accompanied by other relevant developments and investigations the possible use can be found in the comparison of species' spatial distribution at different scales (e.g. Thrush et al., 2005; Gogina et al., 2010). Possibly, the coupling of species ecological functions (filtration rates, bioturbation modes, etc.) with the results of such modelling exercises via biochemical or sediment transport models may help to assess the ecosystem functioning (e.g. Bobertz et al., 2009). The analytical overviews on ecosystem

engineers and their role in ecosystem functioning (Mermillod-Blondin and Rosenberg, 2006; Crain and Bertness, 2006) suggest that without taking into account the distribution of common benthic invertebrates regarding the background environment and limiting variables in aquatic ecosystems, it is impossible to build a rigorous, predictive concept of conserving and restoring damaged ecosystems, to meet the challenge of answering such questions as why do species occur in or avoid a particular region; why are they frequent or rare; which species are responsible for particular ecosystem functions across environmental gradients; how can the consequences of changing habitats be qualitatively and quantitatively assessed, and whether these consequences will be alike in various environments.

The present exercise has confirmed, that salinity, depth and substrate type are all important in determining the distribution of most characteristic macrobenthic species on a large-scale of the whole Baltic Sea. Simple empirical (logistic regression based) habitat suitability models allow to satisfactorily predict the distribution of macrofaunal species even based solely on modelled salinity, depth and rough substrate class information.

Models performed comparatively well in the whole Sea, however their applicability outside the Baltic should be considered at least questionable. They require further development, calibration and validation, and adjustment to environmental patterns known for the region to be applied to, e.g. inclusion of other abiotic variables.

The present exercise is only a first step. Implementation of other variables (e.g. characterizing oxygen and temperature fluctuations, total organic content, and nutrient supply) would obviously increase the model applicability.

Information on the ecological potential of a habitat suitability is utmost important for scientifically sound marine spatial planning (for instance, accounting for precautionary principal, high potential areas should be avoided when planning new marine constructions, as suggested in Degraer et al. 2008).

The roles played by many macrofaunal species are influenced by density, not just occurrence (Thrush et al. 2003). "Factor ceiling" or quantile regressions can be sufficient for modelling of this parameter. Huston (1994) concluded that the effects of competition, predation, and general physical disturbance are alike in that individuals were removed from the assemblage. However, the presented models are aiming to reveal only the general patters. In absence of major anthropogenic impacts habitat suitability and thus ecological potential are far more temporally stable compared to fluctuating macrobenthic community structure. Presented habitat suitability maps predict the specific ecological potential of a habitat for a species (the background) rather that a realized ecological structure, with limitations defined by the predictors considered, their range, and the specified scale, and can be considered as a complementary to observations (Degraer et al. 2008).

Further development, rejection of such general averaging, assimilation of more detailed data and accounting for temporal sequences are indispensable to be able to provide a basis for more particular inferences directed towards management.

Acknowledgments

References

Aarnio, K., Bonsdorff, E., Norkko, A., 1998. Role of *Halicryptus spinulosus* (Priapulida) in structuring meiofauna and settling macrofauna. Mar. Ecol. Prog. Ser. 163, 145–153.

Al-Hamdani, Z., Reker, J. (Eds.), 2007. Towards Marine Landscapes in the Baltic Sea. BALANCE interim report #10. Available at 2007. http://balance-eu.org/.

Andersin, A.B., Lassig, J., Parkkonen, L., Sandler, H., 1978. Long-term fluctuations of the soft bottom macrofauna in the deep areas of the Gulf of Bothnia 1954–1974; with special reference to *Pontoporeia affinis* Lindström (Amphipoda). Finnish Mar. Res. 244, 137–144.

Ankar, S., 1977. The soft bottom ecosystem of the northern Baltic proper with special reference to the macrofauna. Contr. Askö Lab. 19, 1–62.

Bobertz, B., Harff, J., Bohling, B., 2009. Parameterisation of clastic sediments including 656 benthic structures. J. Mar. Sys. 75 (3–4), 371–381.

Bonsdorff, E., 2006. Zoobenthic diversity-gradients in the Baltic Sea: continuous post-glacial succession in a stressed ecosystem. J. Exp. Mar. Biol. Ecol. 330, 383–391.

Bonsdorff, E., Pearson, T.H., 1999. Variation in the sublittoral macrozoobenthos of the Baltic Sea along environmental gradients: a functional-group approach. Aust. J. Ecol. 24, 312–326.

Bromley, R.G., 1996. Trace Fossils : Biology, Taphonomy and Applications, second edition. Champman and Hall, London.

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociol. Methods Res. 33-2, 261–304.

Cade, B.S., Dong, Q., 2008. A quantile count model of water depth constraints on Cape Sable seaside sparrows. J. of Anim. Ecol. 77, 47–56.

Coleman, N., Cuff, W., Moverley, J., Gason, A.S.H., Heislers, S., 2007. Depth, sediment type, biogeography and high species richness in shallow-water benthos. Mar. Freshwater Res. 58, 293–305.

Crain, C.M., Bertness, M.D., 2006. Ecosystem engineering across environmental stress gradients: implications for conservation and management. Bioscience 56, 211–216.

Davies, N.M., Norris, R.H., Thoms, M.C., 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. Freshw. Biol. 45, 343–369.

Davis, J.C., 1975. Waterborne dissolved oxygen requirements and criteria with particular emphasis on the Canadian environment. National Research Council of Canada, Associate Committee on Scientific Criteria for Environmental Quality, Ottawa. NRCC No. 14100.

Degraer, S., Verfaillie, E., Willems, W., Adriaens, E., Van Lancker, V., Vincx, M., 2008. Habitat suitability as a mapping tool for macrobenthic communities: an example from the Belgian part of the North Sea. Cont. Shelf Res. 28 (3), 369–379.

ESRI, 2003. ESRI Data and Maps 2003. Media Kit. ESRI, Redlands, California, USA.

Fauchald, K., Jumars, P.A., 1979. The diet of worms: a study of polychaete feeding guilds. Oceanogr. Mar. Biol. Annu. Rev. 17, 193–284.

Feistel, R., Nausch, G., Wasmund, N. (Eds.), 2008. State and Evolution of the Baltic Sea, 1952–2005. John Wiley & Sons, Inc, Hoboken, New Jersey.

Fenchel, T., Kofoed, L.H., Lappalainen, A., 1975. Particle size-selection of two deposit feeders: the amphipod *Corophium volutator* and the prosobranch *Hydrobia ulvae*. Mar. Biol. 30, 119–128.

Fennel, W., 1995. Wasserhaushalt und Strömung. In: Reinheimer, G. (Ed.), Meereskunde der Ostsee. Springer Berlin Heidelberg, New York, pp. 56–68.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24 (1), 38–49.

Fleischer, D., Zettler, M.L., 2009. An adjustment of benthic ecological quality assessment to effects of salinity. Mar. Pol. Bull. 58, 351–357.

Gerlach, S.A., 2000. Checkliste der Fauna der Kieler Bucht und eine Bibliographie zur Biologie und Ökologie der Kieler Bucht. In: Bundesanstalt für Gewässerkunde (Hrsg.), Die Biodiversität in der deutschen Nord- und Ostsee, Band 3. Bericht BfG-1247, Koblenz.

Gogina, M.A., Glockzin, M., Zettler, M.L., 2010. Distribution of benthic macrofaunal coenoses in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modelling and prediction. J. Mar. Syst. 80 (1–2), 57–70 10.1016/j.jmarsys.2009.10.001.

Gray, S., Elliott, M., 2009. Ecology of Marine Sediments: From Science to Management, second edition. Oxford University Press, New York.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135 (2–3), 147–186.

HELCOM, 2008. Manual for Marine Monitoring in the COMBINE Programme of HELCOM. Annex C-8 Soft Bottom Macrozoobenthos. . http://www.helcom.fi/groups/monas/CombineManual/AnnexesC/en_GB/annex8/?u4.highlight=c8.

HELCOM, 2009. Biodiversity in the Baltic Sea — an integrated thematic assessment on biodiversity and nature conservation in the Baltic Sea. Balt. Sea. Env. Proc. 116B.

Hessle, C., 1924. Bottenbonitoringar i inre Östersjön. Meddelanden fran Kungliga Lantbruksstyrelsen 250, 1–52.

Hosmer, D., Lemeshov, S., 2000. Applied Logistic Regression, Second edition. Wiley, New York.

Huston, M.A., 1994. Biological Diversity: The Coexistence of Species on Changing Landscapes. Cambridge University Press, Cambridge.

ICES EcoSystemData http://ecosystemdata.ices.dk/index.aspx.

IfAÖ, 2007. Autökologischer Atlas benthischer Wirbelloser der deutschen Seegebiete in Nord- und Ostsee. CD.

Karlson, K., Rosenberg, R., Bonsdorff, E., 2002. Temporal and spatial large-scale effects of eutrophication and oxygen deficiency on benthic fauna in Scandinavian and Baltic waters—a review. Oceanogr. Mar. Biol. Annu. Rev. 40, 427–489.

Knipowitsch, N.M., 1909. Otcetch o rabotach v Baltijskom morje N.M. Knipowitscha i S. A. Pawlowitscha ljetom 1908 g. po sobiraniju morskoi fauny dlja Zoologiceskago Museja Imperatorskoi Akademii Nauk. Annuaire du Musee Zoologique de l'Academie Imperiale des Sciences de St.-Petersbourg 14, 131–245.

Laine, A.O., 2003. Distribution of soft-bottom macrofauna in the deep open Baltic Sea in relation to environmental variability. Estuar. Coast. Shelf Sci. 57, 87–97.

Legendre, P., Legendre, L., 1998. Numerical ecology, Developments in Environmental Modeling, Second English Edition. Elsevier Science B.V, Amsterdam, p. 20.

Leppäkoski, E., Shiganova, T., Alexandrov, B., 2009. European enclosed and semi-enclosed seas. In: Rilov, G., Crooks, J.A. (Eds.), Biological Invasions in Marine Ecosystems: Ecological, Management, and Geographic Perspective. Springer, pp. 529–547.

Meissner, K., Darr, 2009. Distribution of *Magelona* species (Polychaeta: Magelonidae) in the German Bight (North Sea): a modeling approach. Zoosymposia 2, 567–586.

Mermillod-Blondin, F., Rosenberg, R., 2006. Ecosystem engineering: the impact of bioturbation on biogeochemical processes in marine and freshwater benthic habitats. Aquat. Sci. 68, 434–442.

Neumann, T., Schernewski, G., 2008. Eutrophication in the Baltic Sea and shifts in nitrogen fixation analyzed with a 3D ecosystem model. J. Mar. Syst. 74, 592–602.

Ojaveer, H., Jaanus, A., MacKenzie, B.R., Martin G., Olenin, S., Telesh, I, Zettler, M.L., Zaiko, A., (submitted for publication). Status and change of biodiversity in the Baltic Sea. PLoS ONE.

Olenin, S., 1997. Benthic zonation of the Eastern Gotland Basin. Neth. J. Aquat. Ecol. 30 (4), 265–282.

Olenin, S., Daunys, D., Dauniene, E., 2009. Baltic Sea Alien Species Database. WWW Page. . http://www.corpi.ku.lt/nemo/.

Schulz, S., 1969. Das Makrobenthos der südlichen Beltsee (Mecklenburger Bucht und angrenzende Seegebiete). Beiträge zur Meereskunde 26, 21–46.

Seifert, T., Tauber, F., Kayser, B., 2001. A high resolution spherical grid topography of the Baltic Sea, Baltic Sea Science Congress, Stockholm, Poster #147, second ed.

SMEBD, 2009. World Register of Marine Species. Online checklist, http://www.marinespecies.org.

StatSoft Inc., 2007. Electronic Statistics Textbook. Tulsa, OK: StatSoft. http://www.statsoft.com/textbook/stathome.html.

Strauss, B., Biedermann, R., 2007. Evaluating temporal and spatial generality: how valid are species–habitat relationship models? Ecol. Model. 204, 104–114.

Thrush, S.F., Hewitt, J.E., Norkko, A., Nicholls, P.E., Funnell, G.A., Ellis, J.I., 2003. Habitat change in estuaries: predicting broad-scale responses of intertidal macrofauna to sediment mud content. Mar. Ecol. Prog. Ser. 263, 101–112.

Thrush, S.F., Hewitt, J.E., Herman, P.M.J., Ysebaert, T., 2005. Multi-scale analysis of species–environment relationships. Mar. Ecol. Prog. Ser. 302, 13–26.

Wallace, J.B., Webster, J.R., 1996. The role of macroinvertebrates in stream ecosystem function. Annu. Rev. Entomol. 41, 115–139.

Wisz, M.S., Guisan, A., 2009. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecol. 9, 8.

Woodward, F.I., Cramer, W., 1996. Plant functional types and climatic changes: introduction. J. Veg. Sci. 7, 306–308.

Ysebaert, T., Meire, P., Herman, P.M.J., Verbeek, H., 2002. Macrobenthic species response surfaces along estuarine gradients: prediction by logistic regression. Mar. Ecol. Prog. Ser. 225, 79–95.

Zajac, R.N., Whitlach, R.B., Thrush, S.F., 1998. Recolonization and succession in soft-sediment infaunal communities: the spatial scale of controlling factors. Hydro-biologia 375–376, 227–240.

Zettler, M.L., Bick, A., 1996. The analysis of small- and mesoscale dispersion patterns of *Marenzelleria viridis* (Polychaeta: Spionida) in a coastal water area of the southern Baltic. Helgolaender Meeresuntersuchungen 50, 265–286.

Zettler, M.L., Röhner, M., 2004. Verbreitung und Entwicklung des Makrozoobenthos der Ostsee zwischen Fehmarnbelt und Usedom - Daten von 1839 bis 2001. In: Bundesanstalt für Gewässerkunde (Hrsg.), Die Biodiversität in der deutschen Nord- und Ostsee, Band 3. Bericht BfG-1421, Koblenz.

Zettler, M.L., Schiedek, D., Glockzin, M., 2008. Zoobenthos. In: Feistel, R., Nausch, G., Wasmund, N. (Eds.), State and Evolution of the Baltic Sea. John Wiley & Sons, Inc, pp. 517–541. 1952–2005.

Zicha, O. (Ed.), 1999–2009. BioLib. http://www.biolib.cz/en.