Investigation of interrelations between sediment and near-bottom environmental parameters and macrozoobenthic distribution patterns for the Baltic Sea

Inauguraldissertation

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Ernst-Moritz-Arndt-Universität Greifswald

vorgelegt von

Mayya Gogina

geboren am 21.05.1982

in Moskau

Greifswald, 08 Januar 2010

Dekan:	Prof. Dr. Klaus Fesser
1. Gutachter :	Prof. Dr. Jan Harff
2. Gutachter:	Prof. Dr. Gerhard Graf
Tag der Promotion:	3 Juni 2010

A doctoral thesis at the Ernst Moritz Arndt University of Greifswald can be produced either as a monograph or, recently, as a collection of papers. In the latter case, the introductory part constitutes the formal thesis, which summarizes the accompanying papers. These have either been published or are manuscripts at various stages (in press, accepted, submitted).

Erklärung

Nach § 4 Abs. 1 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät der Ernst-Moritz-Arndt-Universität Greifswald vom 24. April 2007 (zuletzt geändert durch Änderungssatzung vom 26. Juni 2008):

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Ernst-Moritz-Arndt-Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.
Ferner erkläre ich, daß ich diese Arbeit selbständig verfasst und keine anderen als die darin angegebenen Hilfsmittel benutzt habe.

Greifswald, den

Abbreviations

AIC	Akaike's information criterion
BSH	Bundesamt für Schiffahrt und Hydrographie (Federal Maritime and Hydrographic Agency)
BTA	biological trait analysis
CCA	canonical correspondence analysis
DEM	digital elevation model
DYNAS	Dynamics of natural and anthropogenic sedimentation
GIS	geographic information system
GAM	generalized additive model
GLM	generalized linear model
HELCOM	Helsinki Commission
IOW	Institut für Ostseeforschung Warnemünde (The Leibniz Institute for Baltic Sea Research, Warnemünde)
LAEA	Lambert Azimuthal Equal Area projection
nMDS	non-metric multidimensional scaling
PCA	principal component analysis
UTM	Universal Transversal Mercator
WGS84	World Geographic System 84

Summary (Zusammenfassung)

The objectives of the present work are to relate the spatial distribution of benthic macrofauna in the Baltic Sea to patterns in environmental variables describing near-bottom hydrographical conditions and sediment characteristics, analyzing the data for two various spatial extents. It is mainly based on the data included in the Benthos Databank of the IOW. Other data considered originated from various available database and historical data on distribution of macrobenthic species (such as data provided by the Institut für Angewandte Ökologie, HELCOM monitoring data, Baltic Sea Alien Species Database). The external sources of abiotic data used included the data from BSH, Baltic Sea bathymetry datasets, modelled hydrographical data, sedimentological database of the IOW, and seabed sediments map produced by the EU-BALANCE project.

The first case study is devoted to an exploratory statistical description of the prevailing ecological structure within the limited area attached to the region of the Mecklenburg Bight. By defining the study area, we aimed to lessen the dominance of near-bottom salinity and oxygen concentration (known to be the dominating factors defining the Baltic Sea biodiversity) in the analysis to illuminate the impact of others. Detection of the induced spatial dependencies, examination of the environmental framework and isolation of abiotic predictors of species distribution were executed by means of various statistical methods (e.g. rank correlation, hierarchical clustering, nMDS, BIOENV, CCA). Thus, key environmental descriptors of spatial distribution of macrofaunal communities were disclosed. Within the area of investigation, these were: water depth, regarded as a proxy for other environmental factors (it determines food quality, food availability, light penetration; partial correlation analysis for the considered abiotic variables revealed depth to be the primary descriptor for total organic content, salinity and median grain size) and total organic content. Distinct benthic assemblages that are discriminated by particular species (*Hydrobia ulvae–Scoloplos armiger*, *Lagis koreni–Mysella bidentata* and *Capitella capitata–Halicryptus spinulosus*) were defined. Each assemblage is related to different spatial subarea and is characterized by a certain variability of environmental factors. This study represented the basis for the predictive modelling of species distribution in the selected investigation area, which constituted the next part of the investigation.

Species-specific models predicting the probability of occurrence relative to environmental and sedimentological characteristics were developed for 29 representative macrofaunal species using a logistic regression modelling approach. This way, for most species a good description of their occurrence along gradients of single environmental variables was obtained. Overall, the results showed moderate to high concordance (e.g., 64.1-93.5% for models considering the water depth as predictor, 57.4-94.3% for models predicting the probability of species occurrence relative to total organic content). Subsequently, the technique for a predictive

modelling of species distributions in response to abiotic parameters based on single-factor logistic regression models, utilizing Akaike's information criterion (AIC) and Akaike weights for multimodel inference, was used. Thus, probabilities of occurrence for selected exemplary species (*Arctica islandica, Hediste diversicolor, Pygospio elegans, Tubificoides benedii* and *Scoloplos armiger)* were modelled and mapped. The very similar approach was used to model the benthic species' response of their physical environment in the Pomeranian Bay (southern Baltic Sea). In the scheme of the dominance of strong salinity gradient over the brackish system, consistently small patches of comparatively higher or lower benthic diversity (the Shannon–Wiener diversity index ranges in various areas of the Pomeranian Bay approximately from 1 to 3.9) do emerge in areas where either environmental or anthropogenic impacts on the benthic habitat change drastically over short spatial distances. Hence, spatial diversity of ecological factors creates diversity among benthic colonization and community structures. The possibility to predict thereby induced benthic colonization areas and community structures inside the broad scheme of a brackish water habitat is shown through a logistic modelling approach. The study represents one of the first applications of this technique to benthic habitats of the Baltic Sea.

Finally, the investigation proceeded on a large spatial scale. The discriminating ability of such factors as salinity, bathymetry (as indirect variable replacing a combination of different recourses and direct gradients - a primary descriptor for other abiotic factors) and sediment characteristics (considered only generally due to the lack of more detailed data) to explain the occurrence of typical macrozoobenthic species on the Baltic Sea-wide extend was tested. Full coverage macrofauna distribution maps, though being increasingly demanded, are generally lacking, with information being merely restricted to point observations. In contrast to spatial interpolation, periled by presence of short distance changes in community structure and dependence of the result on density of the samples, predictive habitat suitability modelling allows to objectively produce distribution maps at a level of detail limited only by the availability and resolution of the environmental data. Various literature sources and available databases were analyzed in respect to the information on macrozoobenthos distribution in the Baltic Sea, resulting in the compilation of an extensive list of taxa and an inventory dataset on species distribution for the whole Baltic Sea.

The study demonstrates the need to analyze species' relationships in gradient systems such as the Baltic Sea and provides a basis for a tool to predict natural and anthropogenic forced changes in species distribution.

Table of contents

Part A: Scientific Framework

Part B: Scientific Papers Included

Declaration of the author's contribution

I. Gogina M., Glockzin M., Zettler M.L., 2010a. **Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 1. Causal analysis.** Journal of Marine Systems, 79: 112-123.

II. Meyer M., Harff J., Gogina M., Barthel A., 2008. **Coastline changes of the Darss-Zingst Peninsula - a modelling approach.** Journal of Marine Systems 74: 147-154.

III. Gogina M., Glockzin M., Zettler M.L., 2010b. **Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modelling and prediction.** Journal of Marine Systems, 80: 57-70

IV. Glockzin M., Gogina M., Zettler M.L., 2009. **Beyond salty reins – modelling benthic species' spatial response to their physical environment in the Pomeranian Bay (Southern Baltic Sea).** Baltic Coastal Zone, 13-2, 79-95.

V. Gogina M., Zettler M.L. **Diversity and distribution of benthic macrofauna in the Baltic Sea. Data inventory and its use for species distribution modelling and prediction.** Journal of Sea Research.

Appendix

1 Introduction

Analysis of species–environment relationship has always been a central issue in ecology, starting with the works of von Humboldt and Bonpland (1807) and de Candolle (1855) that recognized the importance of climate to explain animal and plant distribution (Guisan and Zimmermann, 2000). The climate in combination with other environmental factors has been much used to explain the distribution patterns of vegetation around the world in the course of the twentieth century. According to Guisan and Zimmermann (2000) the quantification of species–environment relationships represents the core of predictive geographical modelling in ecology, which is generally based on various hypotheses of how environmental factors control the species and communities distribution. These models have prime importance as a research tool in autecology, but besides that they recently gained importance as an instrument of land use impact assessment, biogeographic hypotheses test, mapping or conservation priorities set up. Most of the examples of statistical models currently in use to simulate the spatial distribution are developed for terrestrial plant species and vegetation units, some what less is done in the field for terrestrial animals, and the number of researches for aquatic system is hitherto considerably lower, but recently this field gains significance and increases rapidly (e.g. Franklin, 1995; Guisan and Zimmermann, 2000; Scott et al., 2002).

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

As consumers at intermediate trophic levels, macroinvertebrates are influenced by both bottom-up and top-down forces in the water system and serve as the conduits by which these effects are propagated. Benthic macrofauna can have an important influence on nutrient cycles, primary productivity, decomposition, and translocation of materials. Interactions among macroinvertebrates and their food resources vary among functional groups, macroinvertebrate species themselves constitute an important source of food for numerous fish. Macrobenthic communities also serve as important indicators of environmental status of the ecosystem, since composition of benthic communities integrates environmental conditions over longer periods of time owning to relative longevity of these organisms (year to decades). Thus, the many roles performed by macroinvertebrates underscore the importance of their study and conservation (Wallace and Webster, 1996; HELCOM, 2009).

1.1 Objectives and aims of the thesis

The objectives of the present work are to relate the spatial distribution of benthic macrofauna to patterns in environmental variables describing near-bottom hydrographical conditions and sediment characteristics, analyzing the data for two various spatial extents.

Regional scale study is focused on the area in the western Baltic Sea between 11.55° and 12.55° E. Depending on the scale and region, salinity together with near-bottom oxygen concentration is often regarded among the major factors affecting the species richness and composition of macrozoobenthic communities (e.g. Olenin, 1997; Zettler et al. 2000; Laine, 2003). By defining the study area, we aimed to lessen the dominance of these two factors in the analysis to illuminate the impact of others. The study area is attached to the region of Mecklenburg Bight close to the very beginning of the Baltic salinity gradient and that is why it is richer in macrofaunal species (with significant presence of both marine and brackish water species at the same time) than the inner, less saline, parts of the Baltic Sea. Geographically this investigation area is bounded by the eastern part of the Mecklenburg Bay and the southwestern Darss Sill area.

Paper I is devoted to an exploratory statistical description of the prevailing ecological structure within the region. Detection of the induced spatial dependencies, examination of the environmental framework and isolation of abiotic predictors of species distribution were executed by means of various statistical methods (e.g. rank correlation, hierarchical clustering, nMDS, BIOENV, CCA). Thus, key environmental descriptors of spatial distribution of macrofaunal communities were disclosed. Within the area of investigation, these were: water depth, regarded as a proxy for other environmental factors, and total organic content. Distinct benthic assemblages that are discriminated by particular species (Hydrobia ulvae–Scoloplos armiger, Lagis koreni–Mysella bidentata and Capitella capitata–Halicryptus spinulosus) were defined. Each assemblage is related to different spatial subarea and is characterised by a certain variability of environmental factors. This study represented the basis for the predictive modelling of species distribution in the selected investigation area.

Predictive estimates of species distribution for the whole investigation area required additional data sets to compile the grids of each abiotic descriptor. Paper I confirmed that bathymetry represents one of the most important proxies for spatial distribution of benthic species. Among numerous possible applications of high resolution digital elevation model (DEM), it is also a necessary input component for modelling the coastline changes scenarios. Paper II is focused on the modelling of the coastline changes of the Darss–Zingst Peninsula, with basic data used for the modelling constituted by the recent DEM, a map of sediment distribution, measured wave time series and a modelled eustatic scenario. Methodological details of DEM generation are also included in this publication. In Paper III the extended version of the same DEM design was used.

In Paper III species-specific models predicting the probability of occurrence relative to environmental and sedimentological characteristics were developed for 29 representative macrofaunal species using a logistic regression modelling approach. This way, a good description of the occurrence of species along gradients of single environmental variables was obtained. Subsequently, the technique for a predictive modelling of species distributions in response to abiotic parameters based on single-factor logistic regression models, utilizing AIC and Akaike weights for multimodel inference, was used. Thus, probabilities of occurrence for selected exemplary species (*Arctica islandica, Hediste diversicolor, Pygospio elegans, Tubificoides benedii* and *Scoloplos armiger)* were modelled and mapped.

In Paper IV the very similar approach was used to model the benthic species' response of their physical environment in the Pomeranian Bay (southern Baltic Sea). In the scheme of the dominance of strong salinity gradient over the brackish system, consistently small patches of comparatively higher or lower benthic diversity do emerge in areas where either environmental or anthropogenic impacts on the benthic habitat change drastically over short spatial distances. Hence, spatial diversity of ecological factors creates diversity among benthic colonization and community structures. Paper IV shows through a logistic modelling approach the possibility to predict thereby induced benthic colonization areas and community structures inside the broad scheme of a brackish water habitat. This study bases upon quantitative macrozoobenthic abundance data collected over a period of 4 years. It demonstrates the need to analyze species' relationships in gradient systems such as the Baltic Sea and proposes a tool to predict natural and anthropogenic forced changes in species distribution.

In the next working step the investigation proceeded on a large spatial scale (Paper V). The discriminating ability of such factors as salinity, bathymetry (as indirect variable replacing a combination of different recourses and direct gradients - a primary descriptor for other abiotic factors) and sediment characteristics (considered only generally due to the lack of more detailed data) to explain the occurrence of typical macrozoobenthic species on the Baltic Sea-wide extend was tested. Full coverage macrofauna distribution maps, though being increasingly demanded, are generally lacking, with information being merely restricted to point observations. In contrast to spatial interpolation, periled by presence of short distance changes in community structure and dependence of the result on density of the samples, predictive habitat suitability modelling allows to objectively produce distribution maps at a level of detail limited only by the availability and resolution of the environmental data. In continuation to regional case study, this part aimed to contribute to such essential ecological issue as estimations of potential response of biota to habitat changes (natural or anthropogenic) and generation of full coverage maps predicting the distribution of macrobenthic species within the whole Baltic Sea. Thus, its goals included the compilation of an extensive list of taxa and an inventory dataset on species distribution for the whole Baltic Sea; extraction of species distribution patterns regarding selected, that is to say available, environmental parameters; modelling and mapping the distribution of exemplary species in response to those factors, using logistic regression with model selection based on information theoretic approach and GIS tools.

A part of the discussion of this thesis is devoted to the review of knowledge on the role of benthic organisms in the near-bed sediment transport processes and the possible employment of the results of predictions of macrofaunal species distribution for related estimates of critical shear stress velocity and bottom roughness parameters. Numerical expressions of dependencies of bottom roughness parameters on marine zoobenthos communities are required for estimation of transport processes and evaluation of extraction caused changes. This addresses the following questions: Which species influence the sediment roughness (e.g. *Hydrobia ulvae*, *Lagis koreni*, *Pygospio elegans*)? How does the occurrence of certain species influence the sediment roughness, how does the influence vary among regions? What is the magnitude of its potential influence on sediment transport?

Outlining the aims of this thesis, they are: first of all - to examine the biotic/abiotic relations for the bottom habitats and to identify the crucial parameters for macrozoobenthos communities composition in the investigation areas. The next objective is to implement the released expressions of interrelations in the regionalization model for prediction of parameters of benthic macrofaunal assemblages and spatial distribution of macrobenthic species using sediment and environmental characteristics. The received consequences are to be reflected in the corresponding GIS model and used for the generation of habitat suitability maps for key species.

1.2 Causes of changes in benthic habitats and communities

Human disturbances which impact soft-sediment habitats act on various spatial and temporal scales (Fig. 1). For instance, bottom trawling has effect on spatial scales ranging from micrometers to many hundred kilometres (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 a warning signal for potential anthropogenic impact as well as baseline maps (Degraer et al., 2008).

Wiens (1989) declared the concept of *grain* and *extent* that describe the lower and upper limits of the resolution of a study. The issue of scales is central in ecology and has been frequently discussed in literature (e.g. Levin, 1992; Legendre and Legendre, 1998). Wiens's concept can be applied to various types of scales presented in ecology – with respect to size, time and complexity, respectively (Karle, 2006). The constraints of grain and extent imply that no extrapolations can be made beyond the extent, and similarly no patterns can be detected below the grain. Yet for logistical reasons expanding the extent of a study usually leads to enlarging the grain, and an increased resolution of the grain leads to a lesser scale extent, that emphasizes the importance of determining the ultimate scale of interest beforehand. Kemp et al. (2001) distinguished between three types of context when determining the grain and extent of a data set: observational scale for data observed in nature, experimental scale for data collected through manipulative experimental setup, and the natural or characteristic scale for data measured as intrinsic characteristics of a natural system, e.g. water depth.

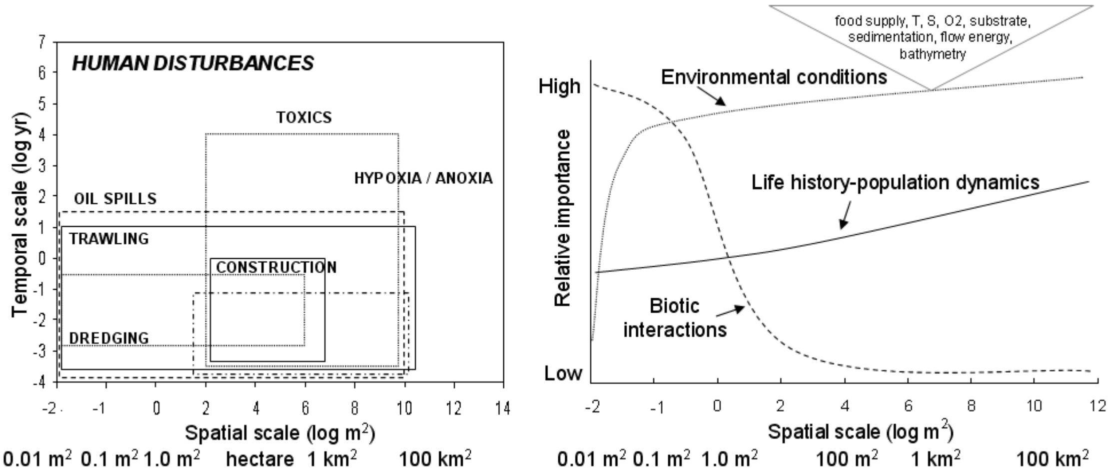


Fig. 1: The left diagram shows spatial and temporal scales of human disturbances which impact soft-sediment habitats (Zajac et al., 1998). On the right relative importance of factors influencing succession over spatial extents (modified after Zajac et al., 1998; Gray and Elliott, 2009) is visualized.

It was reported that relative importance of factors influencing succession in soft-sediment infaunal communities 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 abiotic 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). Statistical models are able to relate ecological features to environmental factors and, through validation and modification, to reveal the underlying mechanisms responsible for the structure and organization of communities (Austin, 1987; Glockzin and Zettler, 2008). Depending on the scale, different processes determine the spatial structure; therefore, inferences derived for one spatial scale cannot be extrapolated to another scale and studies on various spatial extends are important (Legendre and Legendre, 1998).

1.3 Successive steps of predictive geographical modelling

Consequently, there is a growing need to predict ecological responses to habitat changes. Guisan and Zimmermann (2000) summarized the successive steps of the model building process (Fig. 2). This conceptual scheme implies that particular establishments regarding the dynamics and structure of biotic/environmental interactions are required to evaluate natural and anthropogenic influences and effects on ecological systems (Glockzin and Zettler, 2008; Pavlikakis and Tsihrintzis, 2000). An exploratory statistical description of the prevailing ecological structure based on observations is always the indispensible first step (Bourget and Fortin, 1995).

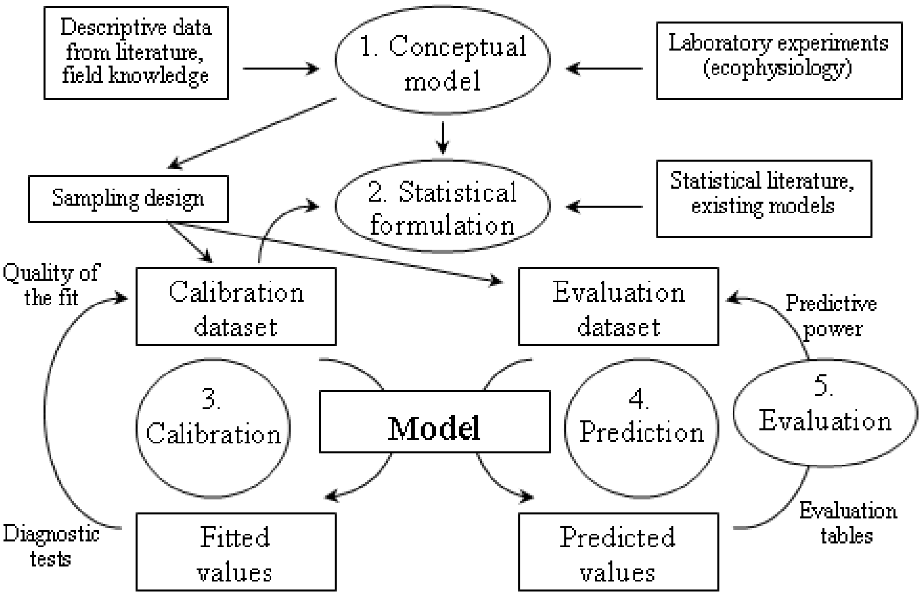


Fig. 2: Overview of the successive steps (1–5) of the model building process, when two data sets, one for fitting and one for evaluating the model, are available. Model evaluation can be made on the independent evaluation data set, by comparing predicted to observed values using preferentially a threshold-independent measure, such as the ROC-plot approach for presence/absence models (reproduced from Guisan and Zimmermann, 2000).

1.4 Exploring the prevailing ecological structure

Many studies have shown on varying scales that substrate characteristics and environmental factors determine the structure and composition of benthic communities (eg. Sanders, 1968; Rhoads, 1974; Gray, 1981; Fortin et al., 2005; O'Brien et al., 2003; Perus et al. 2004). Ekman

(1953) used the water temperature to define biogeographic partition of the continental shelf benthos (on warm-water, temperate, boreal, polar fauna); the second level partition was derived from coastal topography. Longhurst (2006) mentions Petersen's dictum in his classical 1918 paper: it is the nature of the physical biotope that determines the nature of the benthic biocoenosis. He points out and exploits the concept of regional benthic fauna characteristics to be largely dictated by the nature of the sediments, leading to the suggestion that structurally alike and recognizable assemblages of benthic macroorganisms are associated with each characteristic type of sediment on continental shelfs around the world at all latitudes. Continuing with this Thorson (1957) erects the structure of parallel benthic isocommunities (genera communities associated with deposits of certain geological grade and organic content: he recognizes Baltic as *Venus* – deeper shelly sands, Danish Baltic coast as *Macoma* – inshore soft, muddy deposits). Many examples are covered by Clarke and Warwick (2001) in their manual to the widely used PRIMER software package.

Only a few works were carried out establishing the assumed biotic/abiotic link particularly for Baltic Sea. For instance, Andersin et al. (1976) showed the decrease of Shannon index from the beginning of the Baltic Sea salinity gradient to the northern regions, a community study in the southern Baltic Sea was carried out by Warzocha (1995), and a large-scale investigation of soft-bottom macrofauna distribution related to environmental variability accomplished by Laine (2003). Glockzin and Zettler (2008) performed a most comprehensive study to identify potential environmental "key factors" causing spatial distribution of macrofaunal assemblages as a case study in the Pomeranian Bay (southern Baltic Sea). The observed distribution of selected species followed a strong gradient of depth and was explained best by the sediment parameters total organic carbon, median grain size and sorting.

The importance of spatial extent can be confirmed, for instance, considering the work of Zettler and Bick (1996) who concluded biologic interactions to be the critical factors for the fine-scale (ca. 5 by 5 m) variations in distribution of a polychaeta. Works of Olenin (1997), Laine (2003), Bonsdorff (2006), Zettler et al. (2008), Ojaver et al (in press), performed on a more global scale stressed that salinity together with oxygen concentrations define Baltic Sea diversity and best explained the patterns in community distribution. To illustrate the structuring function of salinity for benthic diversity of the Baltic Sea the salinity gradient was compared with the species richness in macrobenthos plotted for the sub-regions of the Baltic Sea (Fig. 3). While the number of marine species decreases rapidly from more saline southern areas to the brackish waters in the North, the fraction of freshwater species increases evidently. This is particularly notable for nearly freshwater inshore lagoons and very shallow offshore regions in the Northeastern part, where a large number of mostly freshwater insects, but also of oligochaetes and molluscs, is recorded.

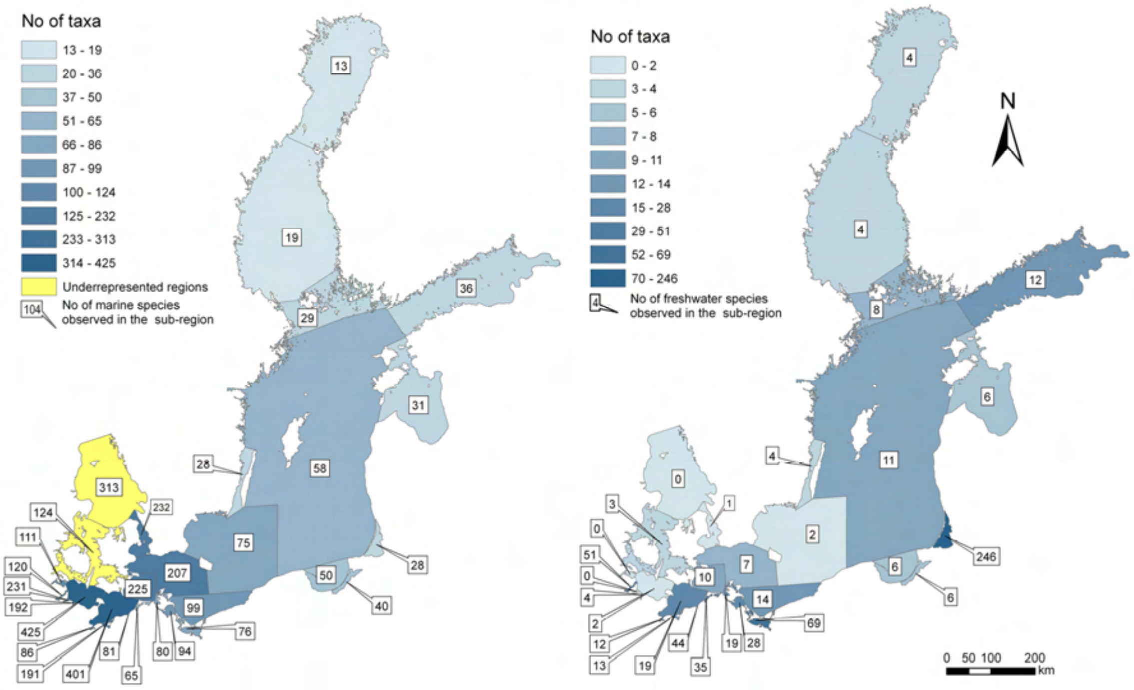


Fig. 3: Change of species richness in the macrozoobenthos across the Baltic Sea environmental gradient (left: marine species, right: freshwater species). Projection: ERTS89_LAEA CRS (Lambert Azimuthal Equal Area projection, ETRS89 datum), from Ojaveer et al., in press.

1.5 Quantifying species response

Recently, a number of studies have succeeded in the development of effective statistical models of benthic distribution. Ysebaert et al. (2002) successfully applied logistic regression to derive response surfaces of distributions for 20 common macrobenthic species found in the Netherland estuary related to salinity, depth, current velocity, and sediment characteristics. They concluded logistic regression modelling approach to be capable of predicting macrobenthic species distributions with a relatively high degree of success where patterns of distribution are strongly and directly coupled to physiochemical processes. Thrush et al. (2003) developed species-specific models for 13 benthic species of New Zealand estuaries that predicted probability of occurrence as well as maximum abundance relative to sediment mud content using logistic regression for distribution modelling and “factor ceiling” method (based on phenomenon described by Thomson et al., 1996) for maximum density modelling. The work highlighted the importance of heuristic approach in statistical modelling of marine ecosystems aiming to extend of our predictive ability for habitat change effects, and pointed out the need to test and iteratively develop the models across a range of spatial scales as well as to support them by the improved mechanistic understanding. Ellis et al. (2006) modelled the distribution of 13 representative macrobenthic species in New Zealand estuarine gradients using logistic regression and classification system based on ‘controlling factors’ with sediment characteristics, elevation, tidal currents, and wind-wave disturbance employed as predictors. They faced complications to fully test the latter approach due to differences in scales of collected benthic data and of higher level physical variables. Meissner et al. (2008) developed habitat models for *Nephtys* species in the German Bight (North Sea) with a median grain size, mud content, depth, and salinity as explanatory variables by application of multivariate adaptive regression splines techniques (MARS). Many researches indicated difficulties due to the complexity of identification of the underlying causal mechanisms controlling species distribution, further extended by the fact that animals modify their physical environment, and many physical parameters co-vary (Ellis et al., 2006).

Application of logistic regression methods in modelling species distribution is not new. This method was widely used in plant ecology (e.g. Guisan et al. 1999) and also in aquatic ecology, but to a lesser extend. Thrush et al. (2003) concentrated their investigation on a single environmental factor – sediment mud content, Ysebaert et al. (2002) performed a comprehensive study, using salinity, depth, flow parameters, median grain size and mud content as predictors. The present study represents one of the first applications of this technique to benthic habitats of the Baltic Sea.

1.6 Interactions between hydrography, sediments and benthic fauna

Benthic environments are at the receiving end of the accumulation and burial of organic material, healthy benthic communities play an important role in benthic-pelagic coupling and mineralization of organic matter settling on the seafloor (HELCOM, 2009). Hydrodynamics over mobile beds involves feedbacks between the sea floor and the overlying flows. The near bed flows as well as the bed forms can be affected by moving sediment and depend further on the properties of benthic communities (Fig. 4; Kuhrts et al., 2004). In an oxygenated environment, benthic animals populate the sediment. Their activity reduces the cohesive forces in the sediment and injects sedimentary material into the water column and, hence, contributes to re-suspension (Neumann and Schernewski, 2008).

The bottom roughness parameters are also dependent on marine zoobenthos communities. Numerical expressions of those dependencies as well as estimates of spatial distribution of macrobenthos are required for estimation of transport processes and evaluation of extraction caused changes. The roughness of the sea floor, used in calculations of bottom shear stress or friction velocities, varies over several orders of magnitude, from the diameter of a sediment grain to the size of large sand waves. The length scale, associated with the biogenic roughness, is difficult to predict, but it often dominates in muddy environments (Harris and Wiberg, 1997; Wheatcroft, 1994).

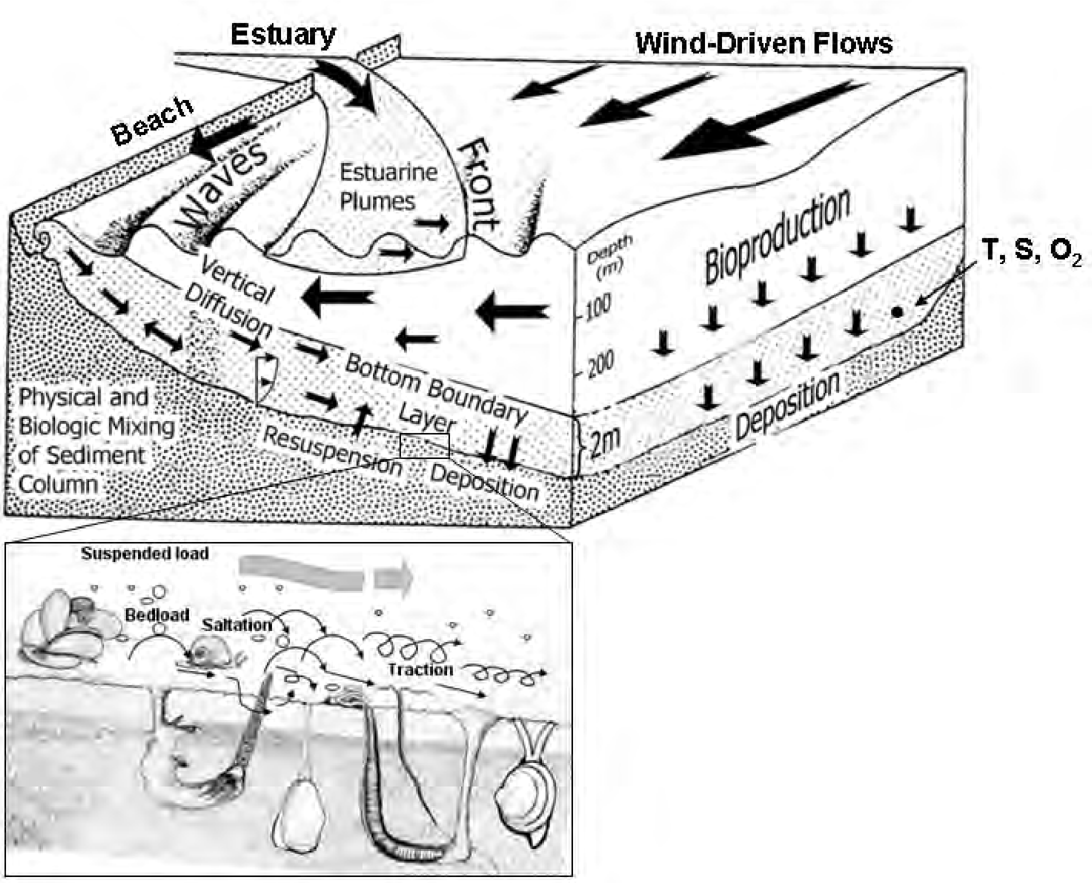


Fig. 4: Summary of sediment transport mechanisms in coastal seas (the upper box is modified after Nittrouer and Wright, 1994; Bobertz et al., 2009). The lower box is a sketch showing some exemplary the species responsible for biogenic structures in the western Baltic Sea - Mytilus edulis, Littorina littorea, Lagis koreni, Mya arenaria, Arenicola marina, Pygospio elegans and Arctica islandica (modified after Friedrichs, 2004).

2 Materials and methods

2.1 Study area

The morphological development of the Baltic Sea from its initial state as a freshwater lake, that formed after the retreat of the Pleistocene ice sheet, to its present state started about 10 000 years ago. It resulted in the formation of the semi-enclosed intracontinental marginal sea (see Fig. 3, Fig.5), one of the world's largest brackish water basins, that resembles a large estuary with steep horizontal and vertical environmental gradients (Bobertz and Harff, 2004; Bonsdorff and Pearson, 1999). Barotropic pressure gradients, river runoff and meteorological forces control the hydrodynamical system of the Baltic Sea (Fennel, 1995). 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. 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). Sea floor here is embedded by organic-reach silt-size sediments. Wave action and thus vertical mixing of the water body takes place in the shallower areas. Sands with coarse-grained lag sediments, formed by erosion in areas of outcropping glacial till moraines, cover the shallower regions. Intense salinity stratification and convective mixing during the cooling period is restricted to the upper 70 m. Large seasonal temperature variation at the surface in combination with the low surface salinity result in regular ice formation. Below 70 m the temperature is rather stable and can be approximated by 5.5°C found throughout the year in the deep basins (Tomczak and Godfrey, 2003).

The highest over all 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). Climatic variation from boreal to subarctic causes additional stress on the ecosystem. Generally, the number of species declines to the north resulting in the low-diversity communities, for example, in the Gulf of Bothnia (Ojaver et al. in press). The entire basin of the Baltic Sea is vulnerable to human influence due to densely inhabited drainage area compared to the water area of only about 412 000 km^2 and slow water exchange with residence time about 20 years. 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).

More comprehensive details on the morphological development and hydrological conditions of the whole Baltic Sea can be found in the book "State and evolution of the Baltic Sea" (recently published by Feistel et al., 2008).

The regional case study (Paper I, III) is focused on the limited area attached to the region of Mecklenburg Bight (Fig. 5) close to the very beginning of the Baltic salinity gradient. As was mentioned above, depending on the scale and region, salinity together with near-bottom oxygen concentration are often regarded among the major factors affecting the species richness and composition of macrozoobenthic communities. To lessen the dominance of those two factors and illuminate the effects of others, the area with reduced variability of both factors in distinct depth-zones, but other factors represented in wide ranges, was restricted. Whereas the range of near-bottom salinity in the whole area is relatively high (99.5% of sampling points lie within a salinity range of 8.3-21.8 psu), the variability in distinct depth-zones is lower (i.e. salinity varied from 9 to 16 PSU at approximately 75% of stations shallower than 18 m; and ranged from 15 to 22 PSU at the same percentage of stations below 18 m). The hypoxic events, connected with absence of oxygen-rich saline water inflows and formation of steady hydrographical stratification in the water column, run differently in various locations of the region and irregularly take place in the late summer – autumn period, e.g. in deepest parts of Mecklenburg Bight. However, they are not as frequent and their effect is not as dramatic (at least within the investigation time) and long-lasting as it is known for the deep basins of the Baltic Proper.

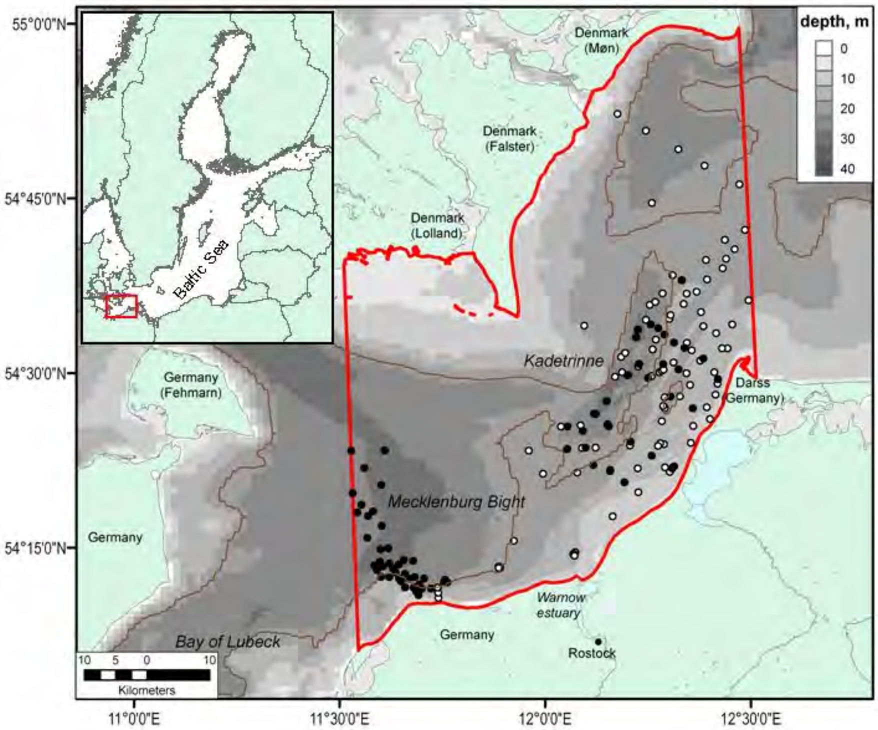


Fig. 5: Investigation area is bordered by the thick red line, distribution of 208 sampling stations is presented; dots are stations covered with quantitative data; filled dots indicate 72 stations with a full set of data available for all eight abiotic variables. Thin line is the 18 m isobath. Geographical data ESRI (2003); projection UTM on WGS84.

The study site ranges in the northwards direction approximately from 45 to 75 km and about 63 km in the eastwards direction. Within these boundaries the area amounts to about 4000 km^2. With an average depth of 16 m (depth at sampling points ranges from 7.5 to 29.9 m) its water volume approximates 46 km^3 (data origin: this study). The sea floor is characterised mainly by postglacial sand and gravel deposits, with organic-rich silt-size sediments observed in the deepest parts of Mecklenburg Bight (Bobertz and Harff, 2004).

The occurrence of brackish species in the area is, in particular, due to the desalinating effect of the Warnow-river that provides the most significant fresh water supply in the investigation area. With basin area of about 3.200 km² and length equal 148 km, its average water discharge in the site of Mühlendammwehr Rostock, before entering the brackish Unterwarmow, is 18,20 m³/s, and the high water discharge is 74,50 m³/s (StAUN, 2008). That is a rather small runoff volume of 0.57 km³/year. The average total river runoff to the Baltic Sea is 450 km³/year (data origin - SMHI, Norrköping, Sweden – provided by T. Seifert, personal conversation).

2.2 Data acquisition

For the regional case study in the Mecklenburg Bight, initially data for 217 stations in the investigation area was available. Station with no quantitative macrobenthic observations (e.g. only dredge sampling) have been eliminated from the analysis. Therefore, the core of data represents 208 stations, sampled simultaneously for benthic macrofauna and associated sediment and near-bottom environmental characteristics, in the depth range from 7.5 to 30 m, all collected from the millennium until the end of year 2007. Most of it is provided from the Benthos Databank of the IOW. 36 stations have the Institut für Angewandte Ökologie (IfAÖ) Database as a source, including only those stations from this data array located in the area, that were sampled with more then one sampling hole per station, with all the abiotc parameters measured parallel to macrozoobenthic sampling.

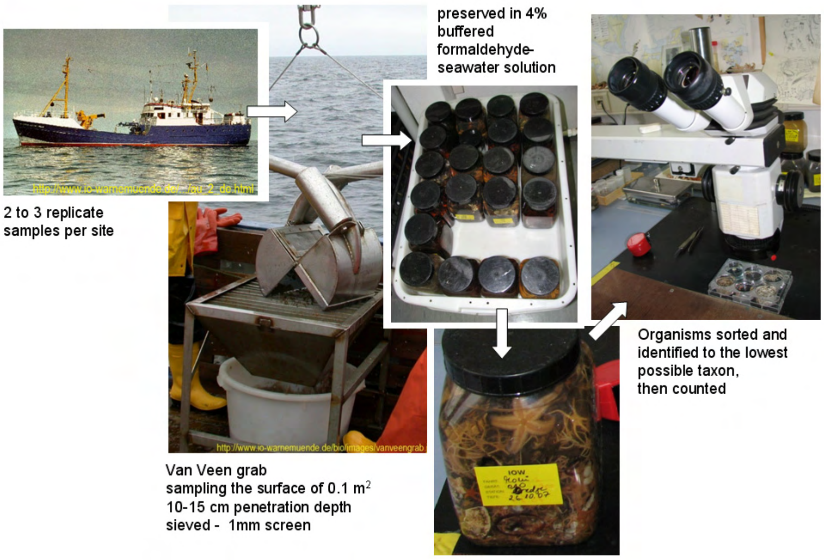
A noticeable part of samples (for 18 stations) were collected during the cruise of RV Prof. Albrecht Penck in May-June 2007 (in the Kadetrinne region) and were then fully analyzed particularly by the author (3 stations were sampled during the later survey due to technical problems, but sample analysis was performed as a set). A total of 208 sites (Fig. 5) include 72 stations with all considered near-bottom parameters (those are bathymetry, salinity, oxygen content, median grain size, total organic content, as well as calculated grain size distribution curve parameters sorting and skewness, and permeability). Most sampling locations (86%) were sampled once, few locations were measured 2 to 5 times during the sampling period considered. Some sampling sites are regular monitoring stations of the IOW, representing the offshore-water habitats of the Baltic Sea (Zettler et al. 2008). The highest density of distribution of sites with (as without) the whole set of parameters measured is observed in two areas, one near the Kühlungsborn and the other assigned to Kadetrinne region.

For the German part of the Pomeranian Bay (Paper IV) the core of the data was the environmental and quantitative macrozoobenthic abundance data collected at 191 stations over a period of 4 years (2003-2006).

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. For the study focused on the whole Baltic Sea a compatible dataset covering the whole Baltic was required to fulfill the aims. To compile it, 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), as well as 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 in respect to information on macrozoobenthos distribution in the Baltic Sea. The nomenclature for macrofauna species, identified to the lowest taxonomic level possible, was checked and revised following the World Register of Marine Species (SMEBD, 2009), and the BioLib taxonomic data (Zicha, 1999-2009) for the freshwater species. In this way an inventory list of about 11 hundred macrozoobenthic taxa grounded on valid taxonomy and synonymy was formed. It was published in Ojaver et al. (in press), and served as a base for developing the HELCOM Red List of the Baltic Sea macrofaunal species. The revised data on species occurrence within the defined Baltic Sea sub-regions was compiled together in GIS (software ArcGIS 9.1, ESRI, USA) including over 160 thousand entries (12200 stations, i.e. sampling events; see Paper V), sampled from 1839 to 2009.

2.2.1 Sampling macrofauna

For macrofauna, 2 to 3 replicate samples per site were taken with a Van Veen grab (0.1 m^2, 10-15 cm penetration depth, Fig. 6) and sieved on a 1.0-mm screen. Samples were preserved in 4% buffered formaldehyde-seawater solution. In the laboratory, the formalin was washed out of the samples prior to sorting. The organisms were sorted, identified to the lowest possible taxon, and counted. Sampling and preparation were conducted in accordance with institutional, national and international guidelines concerning the use of animals in research (HELCOM, 2008).



preserved in 4% buffered formaldehyde-seawater solution

2 to 3 replicate samples per site

Organisms sorted and identified to the lowest possible taxon, then counted

Van Veen grab
sampling the surface of 0.1 m^2
10-15 cm penetration depth
sieved - 1mm screen

Fig. 6: Steps of macrobenthic sampling and laboratory analysis.

2.2.2 Sampling sediments, analysis and calculation of sediment parameters

With regard to the sediment characteristics measurements (completed fully for 72 out of 208 stations), an additional grab sample was taken to extract the upper surface sediment layer (≤5cm) for the analyses. Grain size distribution was analysed using approximately 50 g of dried sediment, dry sieved on a RETSCH sieving machine (sieve set: 63 µm, 75 µm, 90 µm, 106 µm, 125 µm, 150 µm, 180 µm, 212 µm, 250 µm, 400 µm, 630 µm, 2000 µm) and laser sizing (CILAS 1180 Laser Particle Analyser). Total organic content was measured as loss on ignition (3 h at 500°C) of water-free material. In Penck cruise, sediment sample was extracted directly from the grab sample used for macrofauna analysis, and the extracted surface was subtracted from the further computations. On one hand, it reduced the amount of material used in quantitative macrozoobenthic analysis, on the other hand, this provided the sediment data directly corresponding with benthos observation (provides the availability to quantitative analysis of spatial variance). Parameters describing cumulative grain size distribution as median grain size, sorting, and skewness were then calculated by using a skewed s-shape function, fitted to the cumulative grain size data with the least sum of squares method, applying a special fitting algorithm (the description is given in Bobertz, 2000; Bobertz and Harff, 2004).

$$F(x)=\begin{cases}\left(1+e^{-1.7\frac{x-med}{so-sk\cdot\tanh(x-med)}}\right)^{-1}, & so>sk\cdot\tanh(z)\vee sk=0\\ 1, & so\le sk\cdot\tanh(z)\wedge sk>0\\ 0, & so\le sk\cdot\tanh(z)\wedge sk<0\end{cases}\tag{1}$$

where x is the grain size at the phi-scale; med is the median; so is the sorting, (so >0); sk is the skewness. med and so are generally comparable to the associated Folk and Ward (1957) parameters and in the following context approximate the d_{50} and σ, median grain size on phi-scale and the standard deviation of the frequency distribution. In general, skewness values stand for coarse-tailed distributions, characteristic for lag sediments. Negative values indicate a fine-tailed distribution, typical for the fine deposit sediments. Absolute values less than 0.1 should be considered as symmetrical (Bobertz and Harff, 2004).

Permeability, k, is the resistance of the bulk sediment to flow induced by pressure gradient, the notion proposed by Darcy (1856). A common permeability unit is the darcy (D) or m^2 (1 darcy $\approx 10^{-12}$ m^2). The equation to compute the related permeability using the data of grain size distribution (kKM) is derived by Krumbein and Monk (1942):

$$\mathrm{kKM}=7.5\times10^{-4}d_{50}e^{-1.31\sigma}\tag{2}$$

This relation is robust for sorting values less or equal 0.7, equivalent to moderately well sorted or better. Our data exceed this limit. However, the lack of porosity data did not allow the use of more unconfined Carman-Kozeny-relation (Carman, 1937; Forster at al. 2003). Foster et al. (2003) defined the critical threshold for k, a value of 2.5×10^{-12} m^2, below which the biochemical effects are negligible or could not be detected. In the analyzed data values below this limit roughly corresponds to those above the valid sorting magnitude. As there is no other alternative for deriving the permeability with the data available, but assumptions for sorting can not be fullfilled thoroughly, we consider the derived “permeability” as a sort of artificial variable.

Parameters of sediments, estimated from the grain size - sorting, skewness and permeability - can be assumed as more indirect, and, as is always the case, should be treated (e.g. in results interpretation) with response to the reliability of calculation methods.

2.2.3 Hydrographic measurements

For most data points hydrographic characteristics were measured simultaneously with macrozoobenthic sampling. At each station a water sample was taken by shipboard CTD (0.5 m above the sea floor). Oxygen content was determined by immediate potentiometric titration, near-bottom salinity was estimated by CTD as well and the depth at each station was measured and logged via shipboard sonar system.

All abiotic parameters considered in the regional case studies (Paper I, III, IV) are listed in Table 1, with the defined measurement and determination techniques.

Table 1: List of environmental parameters considered. Sampling, measurement and determination techniques are given

Parameter	Units	Method
Water depth	m	Determined and logged with shipboard sonar
Near-bottom oxygen	ml/l	Sampling by shipboard CTD system 0.5 m above the sea bottom, oxygen content determined by immediate potentiometric titration
Near-bottom salinity	PSU	Salinity estimated by shipboard CTD system 0.5 m above the sea bottom
Median grain size	µm	Sieving, laser sizing
Total organic content	%	The total organic content was measured as loss on ignition (3 h at 500°C) of water-free material
Sorting	phi	Derived from grain size distribution curve as described in (Bobertz and Harff, 2004)
Skewness	phi	Derived from grain size distribution curve as described in (Bobertz and Harff, 2004)
Permeability	darcy	Determined according to Krumbein and Monk (1942) using the data from median grain size analysis (see above)

2.2.4 Environmental data

Additional data sets where required to compile the grids of each abiotic descriptor, needed for predictive estimates of species distribution (probability of occurrence) for the whole investigation area. The distribution surfaces obtained for each of the environmental variables considered are presented in the corresponding papers.

For the regional scale digital elevation model with high resolution was designed using measurement data of BSH, and a regional grid dataset from Seifert et al. (2001) covering the Belt Sea region (Paper I-III). Grid datasets for near bottom oxygen content and salinity were based on the modelled hydrographical data, averaged for years 1960-2005 with the resolution 3x3 nm (Neumann and Schernewski, 2008).

Grids for sedimentological factors median grain size, sorting, skewnees and permeability are based on the sedimentological database of the IOW (Bobertz and Harff, 2004), integrating the data from about five decades of marine investigations, with average distance between adjacent sample sites less than 1 nm.

The external limited data on total organic content was also available from the sedimentological database of the IOW. This forced the merge of this data with the observed data used for model estimation. Nevertheless, only the part of the investigation area could be covered with the compiled grid of this parameter. Parameters were interpolated using ordinary kriging with spherical fitted models of semivariograms into a grid with the resolution of about 0.005 decimal degrees (approximately 0.5 km with respect to longitude).

Ysebaert et al. (2002) favoured the usage of modelled estimates of environmental variables over the data measured directly and simultaneously with benthic sampling, highliting its advantages of the available high spatial resolution and a sort of smoothing caused by simulation, e.g. elimination of outfits. However, taking into account the complexity of the functioning of ecosystems, the uncertainty of simulations may increase the complexity of the interpretation of derived empirical relationships. The preliminary explicit exploratory analysis of environmental framework should exclusively be based on direct in situ measurements. Therefore, to enable the investigation of autecological relationships the model calibration for the regional case study relied on directly observed data to the highest extent possible, applying minimum transformations to lessen the reduction of information contained in the data. Yet, the prediction was based partly on modelled data of sufficient resolution available for the study area, thus, allowing some sort of validation of modelling success. Utilisation of simulated data for the model estimation is forced merely by the necessity and absence of alternatives. This refers to the large scale case study, as only for the minor part of the inventory data set the in situ measured abiotic data was available. To maximize the applicability of the habitat suitability model, only widely available environmental variables were offered in the modelling exercise. Again bathymetry represented an indirect variable replacing a combination of different recourses and direct gradients - a primary descriptor for other abiotic factors (e.g. food quality and food availability, light penetration). The bathymetry dataset form Seifert et al. (2001) covering the whole sea was used to compile the DEM. The modelled near-bottom salinity averaged for 45 years to smoothen the variance was provided 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, with resolution of 200 m. The other environmental variables were subject to interpolation using nearest neighbor procedure in order not to lose the coarse, but valuable information on sediment class. Thus, values of 3 environmental variables, assumed to generally controlled species distribution, were obtained for each of the grid cells.

2.3 Statistical methods and data treatment

2.3.1 General basics

A co-called proxy-target concept is the procedure used to predict the parameters of interest from other parameters, either measured or modelled. It has been introduced in Harff et al.

(1992) and also described e.g. in Bobertz et al. (2005). Formally, it can be expressed as follows:

investigation area (or space)	R		
area with proxy known	$R^P \subset R$		
area with target known	$R^T \subset R$		(3)
proxy	$X^P(r)$	$r \in R^P$	
target	$X^T(r)$	$r \in R^P$	
relation	$X^T = f(X^P)$		

Generally, one can distinguish between a multidimentional proxy variable X^P for which the data are available in the whole investigation area and a target variable X^T to be predicted. A relation f between the proxy and the target variable can be given as a deterministic dependence (based on theoretical assumptions) or as regression function (based on the empirical data in case of a statistic dependency) and serve as the base for prediction.

Depending on the context, the proxy is also known as an "independent variable," "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable," or "input variable"; a target, dependent variable, can be referred to as "response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," "outcome variable," "experimental variable," or "output variable".

A variety of statistical techniques – ordination, GLM, GAM, etc. – have been proposed and used for statistic modelling of species distribution related to environment (Guisan et al. 1999). Widely applied in investigations of e.g. plant species distribution, their employment in benthic studies also gained importance in resent years. Application on the use of distinct approaches is best provided in review books (e.g. Legendre and Legedre, 1998). Good overview can be found, for example, in Clarke and Warwick (2001). In the latter work 4 stages of analyses are discussed each focusing on: graphical representation of communities, notions of response variables and predictors, collinearity of predictors, effect of various scales on results interpretation. Clustering and ordination techniques are able to generalize data variability and present it in format, convenient for further interpretation. Ordination is used, when gradients are sought, and clustering, when one is looking for a partitioning into subsets. Main objectives of the interpretation are discrimination of the structure of one or more descriptors, using descriptors at the origin of the structure, or others, potentially explaining it; forecasting one or more descriptors (response or dependant variables), using the number of other descriptors (explanatory or independent variables); prediction of one or more descriptors, using descriptors that exhibit natural variation or can be experimentally manipulated (e.g. Legendre and Legedre, 1998). Different possible tests setups and transformations accompany the analyses. When using the superimposition of clustering onto an ordination, as rooting highly recommended for marine ecological data sets, Legendre and Legedre (1998) advise single linkage clustering, voting for its fine relationships between closely similar objects. Clarke and Warwick (2001), in turn, suggest clustering with unweighted group average linking on sample similarities (e.g. Bray-Cuiris), appropriate for delineating groups of site of different species abundance composition. Modelling species responses to changes of abiotic factors using logistic regression approach on the binary (presence/absence) data seemed appropriate in terms of validity when the abundance density data can not be considered as homogeneous. The GLM tool served as workhorse in many studies aiming to predict macrobenthic response to environmental conditions (e.g. Ysebaert et al. 2002, Thrush et al., 2003, Ellis et al. 2006).

2.3.2 Steps of the causal analysis

Since spatial distribution patterns of species are often to a great extend influenced by spatially structured environment of biological processes, they can be spatially autocorrelated – the location of sampling points in space influences the values of random variables (Legendre, 1993). We have started with the calculation of spatial autocorrelation for 72 stations of abiotic parameters and 208 sampling sites of species abundance data via Morans's I index. A build-in tool of ArcGIS (ESRI) was used, that calculates the Moran's I statistic value and both a Z score and *p*-value evaluating the significance of that index (Index value near +1.0 indicates clustering, near -1.0 – dispersion). Generally, when the p-value is small and the absolute value of the Z score is large enough that it falls outside of the desired confidence level, the null hypothsis ("there is no spatial clustering of the values associated with the geographic features in the study area") can be rejected. The Morans's I statistic is given as follows:

$$I = \frac{n}{S_o} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \qquad (4)$$

where z_i is the deviation of an attribute for the feature *i* from its mean (*xi* - $\overline{X}$), $\omega_{i,j}$ is the spatial weight between feature *i* and *j*, n is equal to the total number of features, and S_o is the aggregate of all the spatial weights.

Continuing with univariate statistics, Spearman's rank correlation between abundance data and corresponding environmental data was the next examination. To analyze the coherencies among environmental variables normal and partial correlation (using Pearson correlation coefficients) were computed (SPSS Inc.). The primary environmental descriptors were disclosed and its effect on species distribution was analysed by means of testing cumulative frequency distributions of the primary predictor versus species abundances using the

Kolmogorov-Smirnov test of significance (Perry and Smith, 1994; Simpson and Walsh, 2004; Glockzin and Zettler, 2008).

For mapping purposes, methods of gradient analysis were complemented by numerical classification that allows to cut the continuum even if there are no distinctive boundaries (Leps and Smilauer, 2003). Ordination contributes to the detection of systematic patterns and can disclose transitional zones. Therefore, multivariate analysis was carried out by superimposing the results of group averaged hierarchical clustering based on Bray-Curtis similarities of 4th–root transformed abundance data on a non-metric multidimensional scaling (nMDS) surface. This way, distinct benthic communities were defined. The covariance-based principal component analysis (PCA) confirmed the results. Species responsible for classification were determined applying SIMPER exploratory analysis (disaggregates the average Bray-Cutis similarities between all pairs of inter-group samples by breaking it down into the separate contribution from each species) and visual re-examination of the modified data matrix whereupon benthic communities have been determined and described (PRIMER; Clarke and Warwick, 2001). To draw a boarder between the communities (Paper I) based on the dataset of (208) irregularly distributed points each assigned to one of 3 classes according to hierarchical clustering of 29 species abundance accomplished with non-metric MDS, the problem of interpolating the nominal data had to be faced. To solve it the posterior probabilities of each class membership were derived using discriminant analysis algorithms (options: *a priori* probabilities from sample and separate covariance matrix for each group) of the SPSS. Then they were subject to interpolation using kriging and the class of each node of a grid was defined as the one having the max probability (this method is common to the combined hierarchical and supervised classification, employed e.g. in Bobertz et al., 2004; Bobertz, personal communication).

Correlations between biological and environmental variables were examined via BIOENV procedure of PRIMER software (Clarke and Warwick, 2001; Clarke and Gorley, 2006) and canonical correspondence analysis (CCA) via CANOCO (ter Braak, 1986; ter Braak and Šmilauer, 2002), applying the necessary data transformations (Davis, 2002; Legendre and Gallagher, 2001). The first algorithm (BIOENV) selects environmental variables, or species "best explaining" community pattern, by maximizing a rank correlation between their respective resemblance matrices, trying all permutations of the trial variables. CCA is a method of ordination of community data matrix constrained by a multiple regression on its relationships to environmental variables (a so-called "direct gradient analysis", as opposed to those that perform an ordination on just the community data, and then secondarily relate the ordination to the environmental variables).

2.3.3 Steps of the regional-scale predictive modelling

Univariate logistic regression: The logistic regression model relates the probability of observing the species *p* to one or more predictor variables *x* (in our study separately to each of the environmental factors) using the logistic link function. The regression model can be written as

$$p(x) = e^z / (1 + e^z) \tag{5}$$

where *z* is the function of the explanatory variable(s). When this is a first-order polynomial, the response is a logistic, S-shaped function. In the special case of Gaussian logistic model when *z* is a second-order polynomial, the response will approximate a bell-shaped function. For this particular case Eq. (5) can be written as

$$p(x) = e^{(b_0 + b_1 x + b_2 x^2)} / (1 + e^{(b_0 + b_1 x + b_2 x^2)}) \tag{6}$$

where b_0, b_1, b_2 are the regression parameters. They are estimated by maximum likelihood, assuming a binomially distributed error term (Legendre and Legendre, 1998; Ysebaert et al. 2002; Wisz and Guisan, 2009). When the estimation of z term parameters is based on log transformed data this can be interpreted as a further extension of the method, aiming to produce an ecologically more plausible response for certain species. This idea was adopted from Thrush et al. (2003) and Thrush et al. (2005) who found such a transformation to produce the most realistic response to changes in sediment mud content for the occurrence of some of investigated species. Thus, for each factor and taxon combination, different functions were used (linear, Gaussian, polynomial), based on either raw or log transformed data. The Wald statistic was used to estimate the model's significance, with a significance level defined at 0.05. The final single-factor model used for each species was the function (of the particular factor) that explained the most variability. The evaluation of the model fit was performed by means of visual control of half-normal plots of residuals and plots of residuals versus predicted values and by considering the percentage of concordant pairs.

Information-theoretic approach and utilization of the Akaike weights: To combine the results of single-factor models and to draw the multimodel inference we considered the information-theoretic philosophy described in Burnham and Anderson (2004). It relies on the calculation of the Akaike's information criterion (AIC) as a model selection tool. AIC is a measure of the relative Kullback-Leibler information loss when the candidate model *i* is used to approximate the truth *j*. Generally appropriate small-sample version (designed for $n/K > 40$) of criterion $\mathrm{AIC_c}$ is calculated as

$$\mathrm{AIC_c} = -2\ln(L) + 2K(K + 1)/(n - K - 1) \tag{7}$$

where L is a value of the maximised log-likelihood over the estimated parameters given the data and the model, K – number of parameters in model i, and n is the sample sizes. AIC (or $\mathrm{AIC_c}$) are calculated for each of the candidate models, then these values are rescaled to calculate delta AIC (Δ_i) so that the model possessing the lowest AIC value has a Δ_i value of 0:

$$\Delta_i = \mathrm{AIC}_i - \mathrm{AIC}_{min} \tag{8}$$

where AIC_{min} is the smallest AIC value in the model set. The model with the lowest AIC value is considered to be the best approximating model in the candidate set. The larger the value of Δ_i the less plausible the fitted model i is. Burnham and Anderson (2004) suggest a simple rule of thumb in assessing relative merits of models in a set: models with $\Delta_i \leq 2$ have strong support, models with Δ_i values between 2 and 10 have considerable support and those with $\Delta_i > 10$ have essentially no support. Akaike weights (ω_i) are now calculated for each of the r candidate models:

$$\omega_i = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_{r=1}^{R}\exp\left(-\frac{1}{2}\Delta_i\right)} \tag{9}$$

The ω_i are scaled between 0 and 1, and all Akaike weights sum to one – ω_i values represent the proportion of evidence for a particular model i in the total evidence supporting all of the models (Wisz and Guisan, 2009). A model that possesses the largest ω_i value is the most parsimonious and has most support among the specified candidate models given the data. When more then one model is supported by the data it is possible to calculate a global model that is a weighted average of all the candidate models in the *a priori* defined set. New parameter estimates for each term in the global model can be computed by weighting them by the Akaike weights

$$\hat{\hat{\theta}} = \sum_{i=1}^{R}\omega_i\hat{\theta}_i \tag{10}$$

Where $\hat{\hat{\theta}}$ is the model averaged parameter estimate based on all R models, $\hat{\theta}_i$ is the parameter estimate for a term in a candidate model i with the Akaike weight ω_i. For terms that do not feature in a candidate model but are present in the global model the parameter estimate is taken to be zero. Thus, if the goal is the prediction, the point inference can be based on the entire set of models using Akaike weights within the overlapping of single-factor models as weight factors to generally estimate the probability of species occurrence.

Habitat suitability mapping was carried out based on estimates derived for exemplary species. They were implemented in a geographical information system. As the grid data for total organic content was only available for a limited part of the investigation area (see section 2.2.4), two model sets were considered for the final prediction: one with and one without the model of species occurrence probability as function of total organic content included. The final value assumed for the overlapping areas of two model sets was the value from the model set with all parameters considered; the rest of the area was filled out with the results based on the model set lacking the total organic content factor. Concordance was calculated to provide the estimate of modelling success.

2.3.4 Technique used for the large-scale modelling

For deriving the habitat suitability models for the Baltic Sea-wide scale the biotic data was also 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 considered 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 correct identification. 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 also chosen as an appropriate technique here. This branch of GLM is classical for binary response. It makes no assumptions 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 form of the preliminary model can be given as follows: $p(x)=\gamma_b/(1+\gamma_b)$, where $\gamma_b = e^{(b_0+b_1S+b_2S^2+b_3D+b_4D^2+b_5Sed_1+b_6Sed_2+b_7Sed_3+b_8Sed_4)}$. Thus it included simple polynomial response for (S) salinity, (D) water depth and sediment classes (Sed) coded as 4-level categorical variable; b_i are the regression parameters. They were estimated by maximum likelihood, assuming a binomially distributed error term. Various combinations of predictors and the functional forms of the relationships (i.e. special cases of the linear predictor, specified above) were tested. 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 best the balance between bias and variance of model prediction and provide the least information loss when approximating the truth.

Models' discriminatory power was tested 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 is 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), PRIMER (PRIMER-E, Plymouth Marine Laboratory; Clarke and Warwick, 2001), MATLAB and ArcMap (ESRI Inc., Redlands, USA).

3 Results

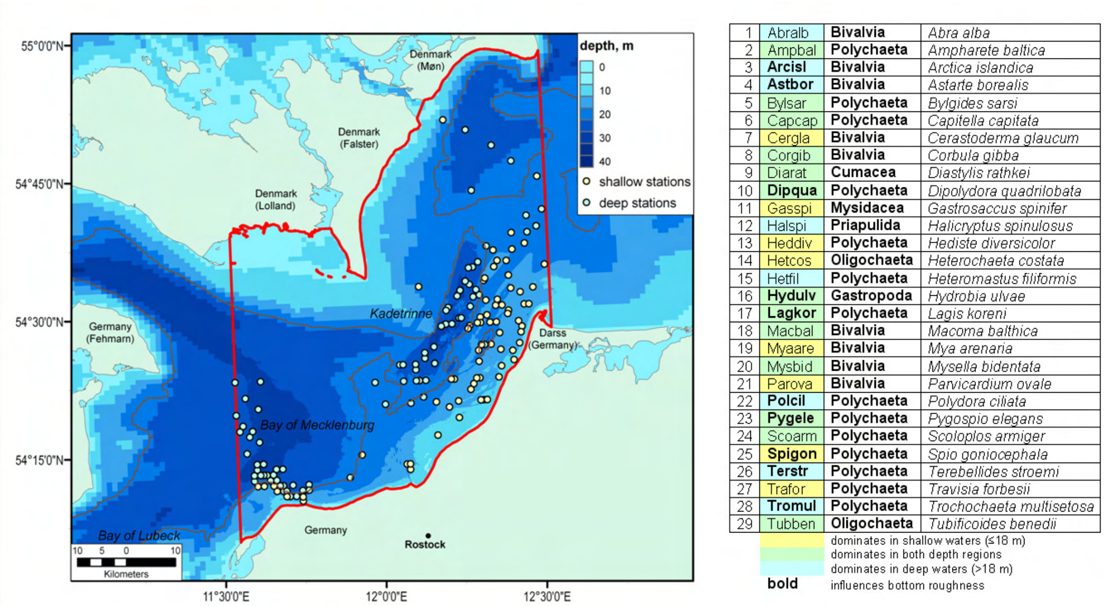
As the results for each part of the present research are presented in the corresponding paper, here the emphasis will be mainly given to some aspects that were not included in the corresponding manuscript (e.g. in terms of brevity).

3.1 Causal analysis

First, data reduction was necessary to allow the employment of various statistical methods. This meant the exclusion of all the uncountable species, remaining only endobenthic species, removal of species with high patchiness that causes unreliable sampling (this was the reason in our case for excluding *Mytilus edulis*), and finally, extraction of all species with a frequency of less than 6 % at all stations and species which account for small (we defined less than 0.3) % of total abundance over all stations. Using those criteria together with general knowledge of species habitats preferences and habits 20 most dominating species were extracted separately for each of two depth subarea of our region (bordered by the 18 m isobath). This formed a list of 29 species under interest (Fig. 7). The including of polychaeta *Bylgides sarsi* in the list represented a stumbling block. This marine species is characterized by relatively high mobility, indistinct substrate preferences and various feeding abilities (carnivore, detritus feeder, predator) (IfÄO, 2007). Its frequency is 58% at the shallower, and 97% at the deeper stations. Account of this taxon proved to bias the results of statistical analysis greatly (especially the ordination). A matrix of biotic data with abundance of those 29 species at 208 stations, and full abiotic data of 8 environmental parameters for 72 stations were created. Benthic sampling was carried out all around the year, 38% each were taken in spring (March, April, May) and summer (June, August), 11% and 13% in autumn (September, October) and winter (January, February) correspondingly. But as we are interested in general all-seasonal distribution patterns, no seasonal harmonization was done.

Next step was the detection of the induced spatial dependencies. There is a potential of strong gradients to monopolize much of a statistical analysis, which suggests their removal before the analysis unless it is the state aim to explore the influence of this particular gradient. We have tested the autocorrelation beforehand via Moran's Index and considered its possible bias in further investigation. We have attempted to reduce the effect of salinity gradient, characteristic for the Baltic Sea by limiting the study area. Smooth depth gradient is also present.

The partial correlation coefficient measures the intensity of this linear association of two variables while taking to the account their relationships with other variables. Correlation does not imply cause and effect. Path analysis may be used to assess the correspondence between the data and causal models, when causal ordering of the descriptors is stated by the prior information of ecological hypotheses. Pearson correlation coefficients were calculated for the z-transformed abiotic matrix using SPSS Inc. software (Tables 2, 3). Then path analysis was applied. All possible and meaningful three-variable combinations derived from zero-order correlation matrix and the matrix of partial correlation coefficients were considered and tested. Fig. 8 shows as an example the single correlation matrix for total organic content, median grain size and permeability, and the causal model supported by data. Schematically, all tested causal relations are illustrated by Fig. 9. Depth was proven to be the primary descriptor for total organic content, salinity and median grain size. Salinity shows weak positive association with organic content, even though its ranges are quite restricted in the particular data. Organic content directly controls median grain size, sorting, permeability (via interstice-blocking) and, to a smaller degree near-bottom oxygen concentrations. Sorting (been the second moment of the grain size distribution function and permeability (functionally derived from values of median grain size and sorting) are directly controlled by median grain size.



1	🟦 Abralb	**Bivalvia**	*Abra alba*
2	🟩 Ampbal	**Polychaeta**	*Ampharete baltica*
3	🟦 **Arcisl**	**Bivalvia**	*Arctica islandica*
4	🟦 **Astbor**	**Bivalvia**	*Astarte borealis*
5	🟩 Bylsar	**Polychaeta**	*Bylgides sarsi*
6	🟩 Capcap	**Polychaeta**	*Capitella capitata*
7	🟨 Cergla	**Bivalvia**	*Cerastoderma glaucum*
8	🟩 Corgib	**Bivalvia**	*Corbula gibba*
9	🟩 Diarat	**Cumacea**	*Diastylis rathkei*
10	🟩 **Dipqua**	**Polychaeta**	*Dipolydora quadrilobata*
11	🟨 Gasspi	**Mysidacea**	*Gastrosaccus spinifer*
12	🟦 Halspi	**Priapulida**	*Halicryptus spinulosus*
13	🟨 Heddiv	**Polychaeta**	*Hediste diversicolor*
14	🟨 Hetcos	**Oligochaeta**	*Heterochaeta costata*
15	🟦 Hetfil	**Polychaeta**	*Heteromastus filiformis*
16	🟩 **Hydulv**	**Gastropoda**	*Hydrobia ulvae*
17	🟩 **Lagkor**	**Polychaeta**	*Lagis koreni*
18	🟩 Macbal	**Bivalvia**	*Macoma balthica*
19	🟨 Myaare	**Bivalvia**	*Mya arenaria*
20	🟩 Mysbid	**Bivalvia**	*Mysella bidentata*
21	🟨 Parova	**Bivalvia**	*Parvicardium ovale*
22	🟦 **Polcil**	**Polychaeta**	*Polydora ciliata*
23	🟩 **Pygele**	**Polychaeta**	*Pygospio elegans*
24	🟩 Scoarm	**Polychaeta**	*Scoloplos armiger*
25	🟨 **Spigon**	**Polychaeta**	*Spio goniocephala*
26	🟦 **Terstr**	**Polychaeta**	*Terebellides stroemi*
27	🟨 Trafor	**Polychaeta**	*Travisia forbesii*
28	🟦 **Tromul**	**Polychaeta**	*Trochochaeta multisetosa*
29	🟩 Tubben	**Oligochaeta**	*Tubificoides benedii*

🟨 dominates in shallow waters (≤18 m)
🟩 dominates in both depth regions
🟦 dominates in deep waters (>18 m)
bold influences bottom roughness

Fig. 7: The division of stations into depth subareas (left) and the formed list of 29 species under interest (right; those marked with yellow dominating the shallower subarea, blue - dominant in deeper subarea, and green - typical for the whole region; in bold font are those species influencing the bottom roughness).

Apparently, sorting also affects permeability, whereas it is eliminated by the correlation matrix, for the relationship is not linear. Skewness, been the third statistical moment of the grain size distribution function, is defined by such a complex composition of physical processes, that its direct association with other environmental parameters considered is undisclosed.

Table 2: Pearson correlation coefficient calculated for the full set of z-transformed environmental data at 208 stations, coefficient values corresponding to $p<0.01$ (2-tailed) are in the bold font, n-values are given in brackets. The notation of environmental factor is as follows: Org – total organic content, Kg - median grain size, O2 – oxygen, S – salinity, d – depth, so – sorting, sk – skewness, kKM – permeability.

	Zsc(Org)	Zsc(KG)	Zsc(O2)	Zsc(S)	Zsc(d)	Zsc(so)	Zsc(sk)	Zsc(kKM)
Zsc(Org)								
	(109)							
Zsc(KG)	**-.363**							
	p<0.01							
	(109)	(110)						
Zsc(O2)	-0.187	**.262**						
	p=0.05	p<0.01						
	(109)	(110)	(181)					
Zsc(S)	**.544**	**-.338**	-0.082					
	p<0.01	p<0.01	p=0.27					
	(109)	(110)	(180)	(185)				
Zsc(d)	**.670**	**-.354**	-0.130	**.597**				
	p<0.01	p<0.01	p=0.08	p<0.01				
	(109)	(110)	(181)	(185)	(208)			
Zsc(so)	**.701**	**-.509**	-.269	**.500**	**.608**			
	p<0.01	p<0.01	p=0.02	p<0.01	p<0.01			
	(72)	(72)	(72)	(72)	(72)	(72)		
Zsc(sk)	-0.227	**.439**	.257	-0.226	**-.319**	**-.460**		
	p=0.05	p<0.01	p=0.03	p=0.06	p<0.01	p<0.01		
	(72)	(72)	(72)	(72)	(72)	(72)	(72)	
Zsc(kKM)	**-.549**	**.923**	**.305**	**-.574**	**-.747**	**-.521**	**.487**	
	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01	
	(72)	(72)	(72)	(72)	(72)	(72)	(72)	(72)
Factor unit	%	µm	ml/l	PSU	m	phi	phi	darcy
Factor range	0-28.44	7-1149	0.5-13	8.3-27.1	7.5-29.9	0.29-2.97	-1.46-0.37	0-49

Table 3: Pearson correlation coefficient separately calculated for z-transformed environmental data for two subareas: shallower (above the diagonal) and deeper (below) then the 18 m isobath, accounting only the station covered by full set of 8 abiotic parameters. The notation is as in Table 2.

	Zsc(Org)	Zsc(KG)	Zsc(O2)	Zsc(S)	Zsc(d)	Zsc(so)	Zsc(sk)	Zsc(kKM)
Zsc(Org)		-.357	.025	.233	.001	-.220	-.341	-.283
		p=0.11	p=0.91	p=0.31	p=0.99	p=0.34	p=0.13	p=0.21
		21	21	21	21	21	21	21
Zsc(KG)	**-.747**		.149	-.012	-.384	**.607**	**.745**	**.927**
	p<0.01		p=0.52	p=0.96	p=0.09	p<0.01	p<0.01	p<0.01
	51		21	21	21	21	21	21
Zsc(O2)	**-.452**	**.478**		.528	.154	-.046	.513	.107
	p<0.01	p<0.01		p=0.01	p=0.50	p=0.84	p=0.02	p=0.64
	51	51		21	21	21	21	21
Zsc(S)	**.574**	**-.576**	-.194		.214	-.218	.252	-.008
	p<0.01	p<0.01	p=0.17		p=0.35	p=0.34	p=0.27	p=0.97
	51	51	51		21	21	21	21
Zsc(d)	**.601**	**-.577**	-.338	.262		-.297	.089	-.456
	p<0.01	p<0.01	p=0.01	p=0.06		p=0.19	p=0.70	p=0.04
	51	51	51	51		21	21	21
Zsc(so)	**.620**	**-.432**	-.197	**.364**	**.507**		.320	.322
	p<0.01	p<0.01	p=0.17	p<0.01	p<0.01		p=0.16	p=0.15
	51	51	51	51	51		21	21
Zsc(sk)	-.031	.155	.164	.021	.013	-.353		**.649**
	p=0.83	p=0.28	p=0.25	p=0.88	p=0.93	p=0.01		p<0.01
	51	51	51	51	51	51		21
Zsc(kKM)	**-.649**	**.846**	**.433**	**-.544**	**-.596**	**-.649**	.319	
	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01	p<0.01	p=0.02	
	51	51	51	51	51	51	51	

	Perme-ability	Total organic content	Median grain size
Perme-ability	----	r = -.549***	r = .923***
Total organic content	Partial r = .337**	----	r = -.696***
Median grain size	Partial r = .901	Partial r = -.588***	----

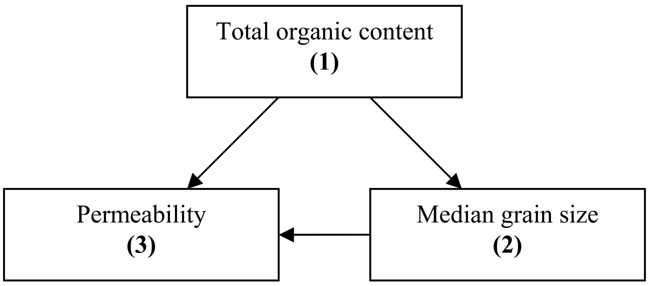


Fig. 8: Left: Pearson zero-order (above the diagonal) and partial (below) correlations among three variables: total organic content (1), median grain size (2) and permeability(3). Test of significance: **: 0.001<p<0.01; ***: p<0.001. Right: Causal model supported by data. Triangular model represents the combination of double effect, with (1) been the primary descriptor causing both (2) and (3), and intervening sequence, with secondary descriptor (2) intervening between (1) and (3).

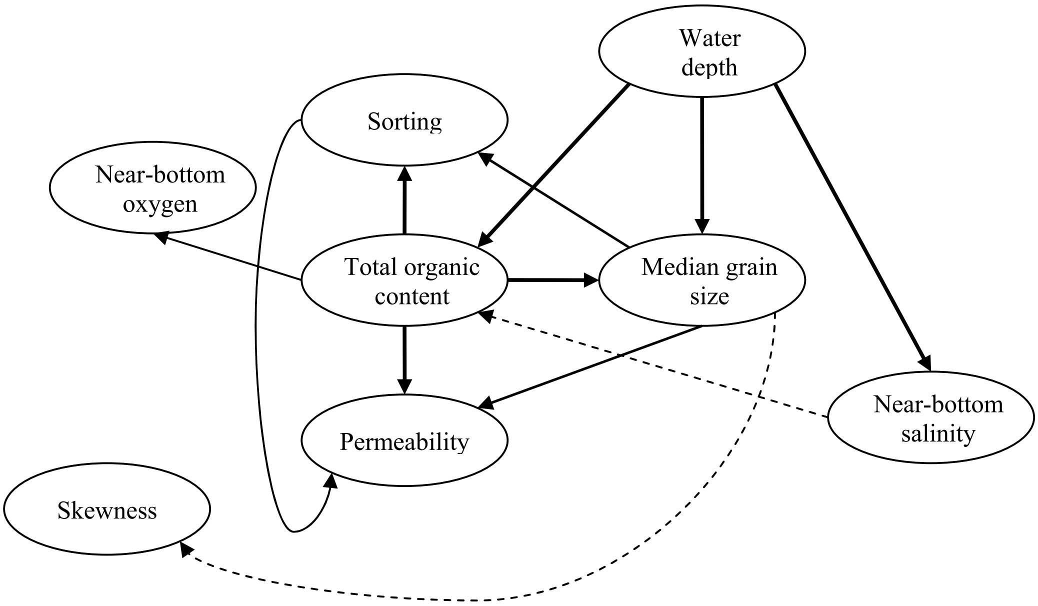
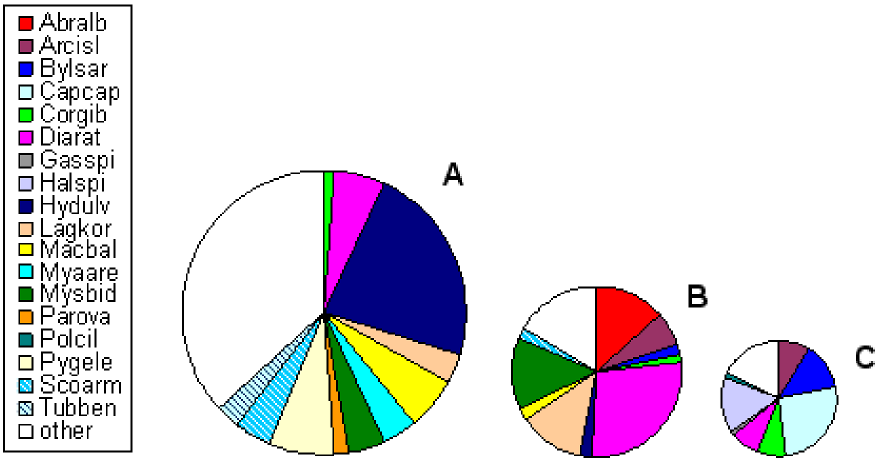


Fig. 9: Schema of causal relationships between environmental parameters, observed in the south western Baltic Sea, in accordance with tested ecological hypotheses. Linearity of the association is reflected by the shape of connectors, its straight is presented by the thickness of the line.

The complexity of abiotic interactions, reflected by the results of correlation analysis (Tables 3 and 4; Figs. 8, 9) is dominated by depth, whereas most species indicate strong response to it (see Paper I). This does not allow the complete removal of either gradient without the loss of essential information. The solution was fount in considering depth-detrended abiotic data as well as the initial environmental data in further analyses, suggested by previous work of Glockzin and Zettler (2008).



Total average abundance (ind/m^2)	7611	3219	736

Fig. 10: Generalized structure of the communities revealed in Paper I. The pie graphs illustrate the average species composition of the ‘classes’, pie diameter is relative to the average total abundance species. Only species from the selected 29 taxa forming more then 1 % of total abundance are included.

Prior to the identification of environmental factors responsible for benthic macrofauna distribution, the patterns contained in data set of species abundance are analyzed in order to detect differences in community structure. The valid zonation was received by means of hierarchal clustering combined with non-metric MDS technique and confirmed by other statistical methods (PCA, SIMPER, see Paper I). Three distinct benthic communities were defined, each related to different spatial subarea and characterized by certain variability of environmental parameters (Fig. 10; Table 4 in PaperI): *Hydrobia ulvae – Scoloplos armiger* (A) community inhabits the area of the shallow sandy bottom along the coast. Other species numerically dominating in the community are *Pygospio elegans*, *Macoma balthica, Mya arenaria, Diastylis rathkei*. *Lagis koreni – Mysella bidentata* (B) community with highest biodeversity occupy the intermediate-deep area stretched from south-west to north-east (middle of the Kadetrinne and the Darss Sill), including southern and northern parts of the Mecklenburg Bight. It is characterized by moderately sorted sediments of medium silt to fine sand and increased salinity. In terms of both occurrence and abundance the cumacean *Diastylis rathkei* and bivalve *Abra alba* are also characteristic for this community. *Capitella capitata – Halicryptus spinulosus* (C) community inhabits the deepest area in the Mecklenburg Bight, characterized by high organic content, fine to coarse poorly sorted silt sediments with low permeability. This community is characterized by low number of taxa. In terms of occurrence polychaeta *Bylgides sarsi* dominates in the community, but high frequency of this taxa is typical for the whole investigation area.

To draw a boarder between the communities (Paper I) the problem of interpolating the nominal data had to be faced. To solve it the posterior probabilities of each class membership were derived using discriminant analysis. Class distribution revealed two problems: the area is wide stretched - this has an influence on the maximum lag for semivariogram calculation, and there is one class (*Capitella capitata – Halicryptus spinulosus* community) clustered in the small south western area, that will result in small region with 1.0 probabilities and the large area with 0.0 (bad thing for calculating the semivariogram). The solution was found in a directional semivariogram, which implies, that the true semivariogram exsists in the direction. The contrary, that all other directions have a trend, must not be true. In fact, if there is a trend-free situation the omnidirectional semivariogram should be identical to each directional semivariogram. So, the directional semivariogram can be used instead of the omnidirectional one. As for the community concentrated in one limited area (*Capitella capitata – Halicryptus spinulosus*) the omnidirectional semivariogram was calculated for the data limited in the longitudinal direction (resulting in a small area with better distribution), that was then used for the entire area (Fig. 11). Then probabilities of each class membership were subject to interpolation using ordinary kriging based on derived semivariograms and the class of each node of a grid was defined as the one having the max probability. The method has its limits because one interpolates onto a grid and does a (re)classification of the grid nodes based on the interpolated probabilities - so the stations are not involved anymore; moreover, there can be stations with varying sampling time and conditions at close locations, leading to some miss assignmets (i.e. 5.8 % or 12 stations out of 207 obtain the max probabilities for the class they do not belong to).

The final step of exploratory analysis was the isolation of abiotic predictors of species distribution. Canonical correspondence analyses (Paper I, section 3.9) approximated and ranked the optima of individual species in multidimensional environmental space. Monte-Carlo permutation tests were preformed to test the effects of each environmental factor considered. The comparison of results for detrended and undetrended environmental data disclosed the predictors of benthic macrofauna composition. The influence of depth on all other factors was evidenced. Without the depth trend total organic content causes the best similarity between the abiotic and biotic data, followed by sorting and permeability. They differ in their rank in the results of BIOENV and CCA, but obviously remain dominating in the explanation of benthic variance.

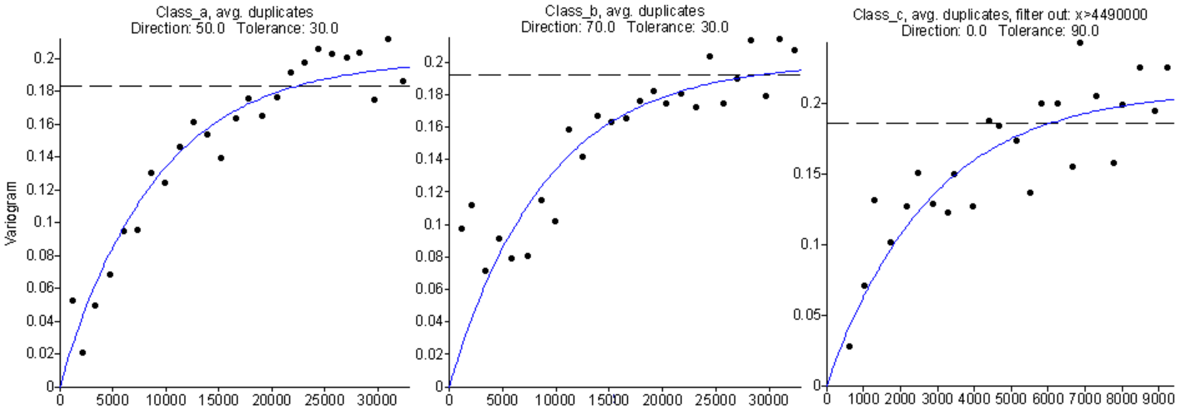


Fig. 11: Empirical directional semivariograms (for classes a and b, left and middle) and onmi-directional (for class c, right); fitted exponential models for the probabilities of class-membership.

3.2 Regional scale predictive modelling

The detailed analysis of patterns of benthic community distribution related to the selected environmental parameters provided basis for the predictive modelling of species distribution. Using the methodology defined in Section 2.3.3 the curves that approximate the response of species to changes of particular environmental factor were estimated using logistic regression with biotic data reduced to presence/absence. Such species-specific response curves were estimated for each factor-species combination using various mathematical functions. The final model was the function of the factor that explained most variability. Evaluation of model fit was preformed by means of visual control of half-normal plots of residuals and by considering the percentage of concordance. Thus – a set of single-factor models was compiled for each of the exemplary species. Logistic regression models predicting the probabilities of occurrence for 29 selected macrobenthic species relative to depth and total organic content are presented in Paper III, species response curves for median grain size and sorting are given in Fig. 12 and Fig. 13, respectively. Some species indicate strong response to the particular factor over the whole parameter range, whereas others do not seem to depend on the factor.

The multimodel inference serves to predict the distribution of species within the common limits of combined environmental factors. The information-theoretic approach is used to deal with model selection uncertainty, using derived Akaike weights (for each model of a set) within the overlapping of single-factor models as weight factors to receive the general estimate of probability of species occurrence. As total organic content data was limited: two model sets were considered for the final prediction - one with and one without the model of species occurrence probability as function of total organic content included.

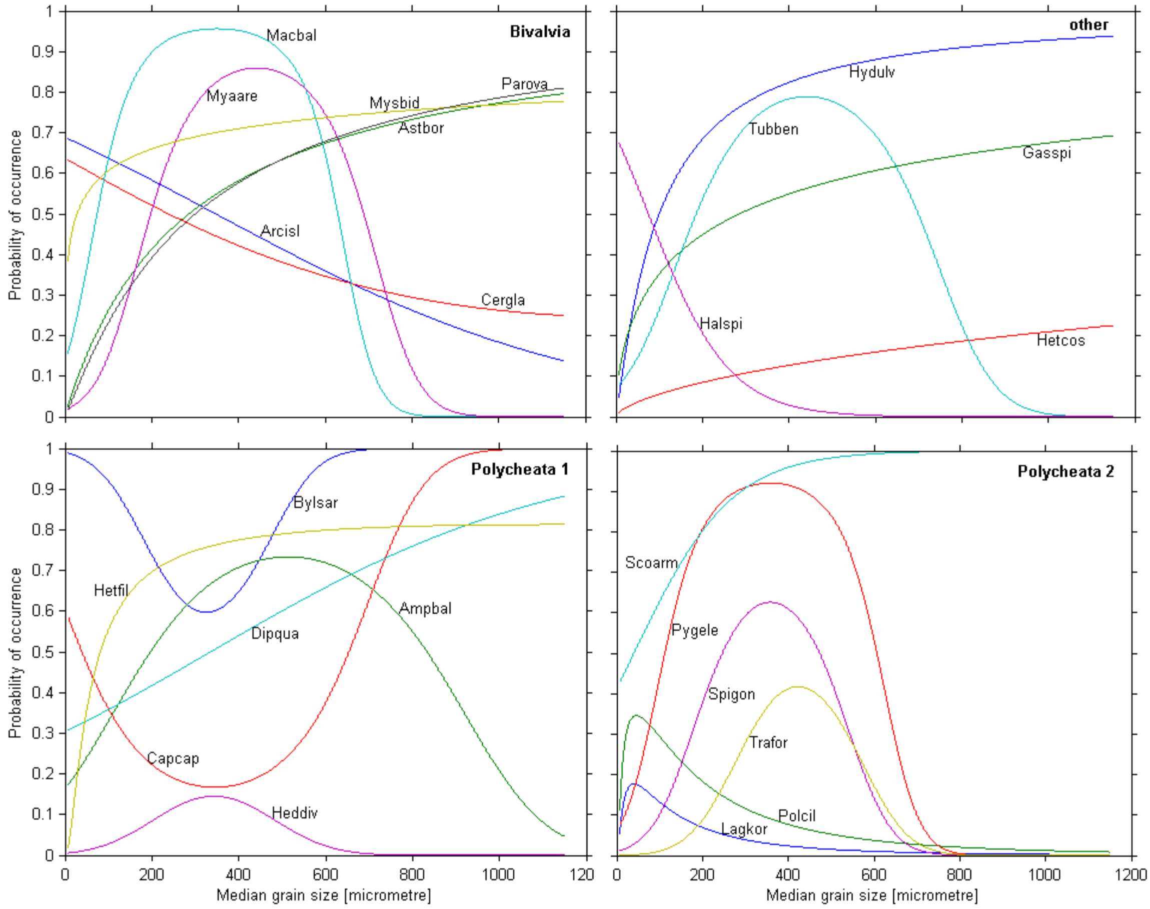


Fig. 12: Logistic regression models predicting the probabilities of occurrence for 29 selected macrobenthic species relative to median grain size. Species abbreviation: Abralb, Arcisl, Astbor, Bylsar, Capcap, Cergla, Corgib, Diarat, Dipqua, Gasspi, Halspi, Heddiv, Hetfil, Hydulv, Lagkor, Macbal, Myaare, Mysbid, Parova, Polcil, Pygele, Scoarm, Spigon, Terstr, Trafor, Tromul, Tubben stand for *Abra alba, Ampharete baltica, Arctica islandica, Astarte borealis, Bylgides sarsi, Capitella capitata, Cerastoderma glaucum, Corbula gibba, Diastylis rathkei, Dipolydora quadrilobata, Gastrosaccus spinifer, Halicryptus spinulosus, Hediste diversicolor, Heterochaeta costata, Heteromastus filiformis, Hydrobia ulvae, Lagis koreni, Macoma balthica, Mya arenaria, Mysella bidentata, Parvicardium ovale, Polydora ciliata, Pygospio elegans, Scoloplos armiger, Spio goniocephala, Terebellides stroemi, Travisia forbesii, Trochochaeta multisetosa,* and *Tubificoides benedii*, correspondingly.

Using the defined method, probabilities of occurrence for selected representative species (*Arctica islandica, Hediste diversicolor, Pygospio elegans, Tubificoides benedii* and *Scoloplos armiger)* within the investigation area are modelled and mapped (Paper III). The response surfaces obtained indicate relatively high degree of success. Water depth that represents a type of integral parameter remains the key player determining the species distribution among the parameters considered within the study scale, particularly for species that find their optima habitat here, such as *A. islandica*, but also for those as *H. diversicolor* that occur only locally and at comparatively low densities. In the results obtained for *T. benedii* and *S. armiger* total organic content, sorting and permeability also contribute significantly in defining the suitable habitat for the latter two.

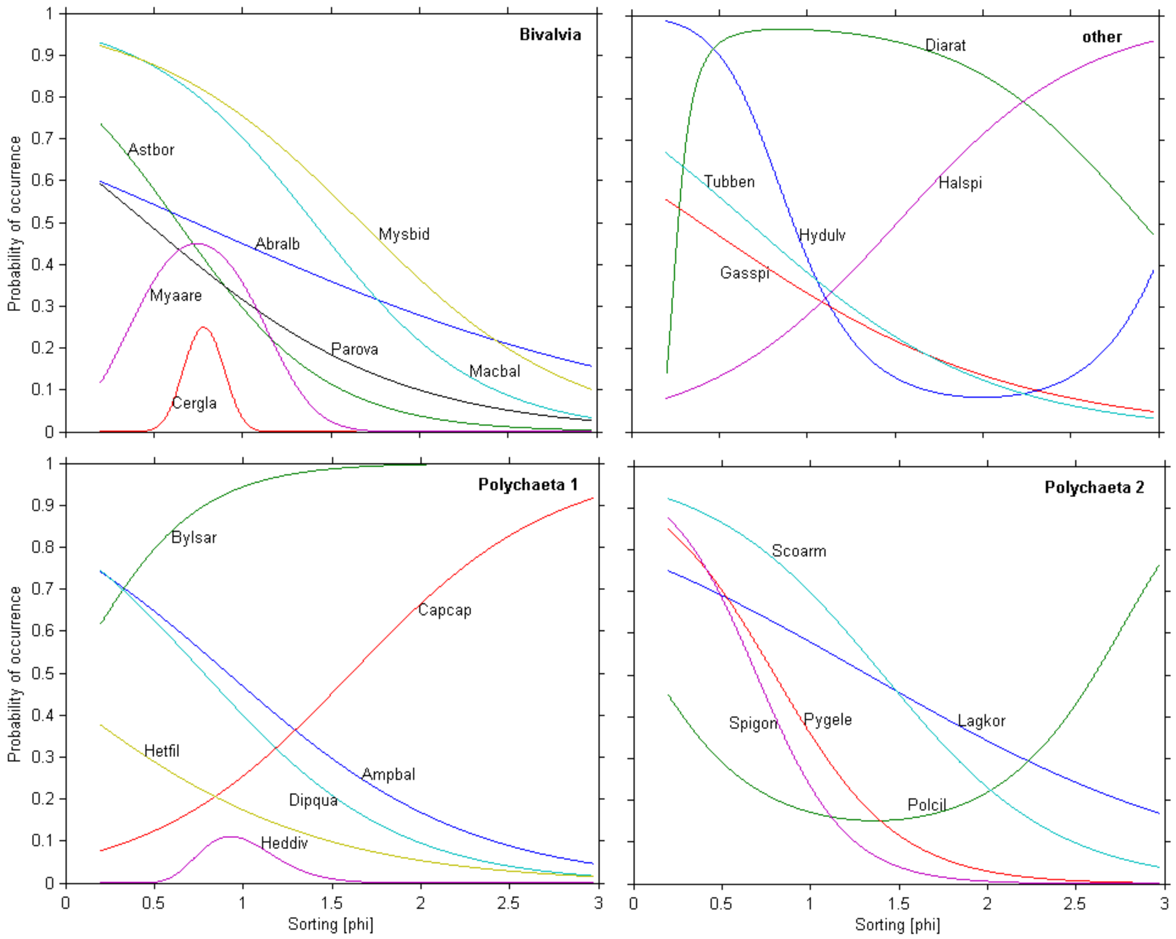


Fig. 13: Logistic regression models predicting the probabilities of occurrence for 29 selected macrobenthic species relative to sorting. Species abbreviation as in Fig. 12.

To gain more insight predicted probability of occurrence and measured abundance vs. abiotic factors for five selected species are plotted in Appendix A

3.3 Large scale predictive modelling

For this exercise 19 species from various functional groups representative for both marine and limnic communities, relatively abundant and well represented by the data were chosen. The frequency of occurrence for selected species ranged from 6 to 30 % within the dataset (Paper V). The observed distribution of the exemplary species along the ranges of considered environmental factors were investigated. Differences between species were observed in regard to their relative frequency of occurrence and density (Appendix B, Paper V). For instance, *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, 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*). Interestingly, the bivalve *A. islandica* shows a very clear bi-modal response to depth changes, with its highest densities found between 15 and 27 m depth, nearly entire absence of occurrence events at the interval of 30 to 40 m depth, and the appearance of the second essentially smaller peak of species abundance around the depth of 43-47 m. This latter sudden occurrence of species along the depth gradient corresponds to the deeper region of the Arkona Sea, the eastern frontier of species spread area. It can presumably be explained by the coupling effect of depth and salinity, and perhaps by the transport of larvae with the hydrodynamic flow, both some what sheltered conditions coinciding with the intermediate depth and distance to the shore, and still relatively high values of salinity.

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.

Using the method defined in Section 2.3.4, probabilities of species occurrence were modelled and mapped. The resulting maps with the observed species abundance and modelled probability of occurrence as function of considered environmental variables are presented in Paper V and in Appendix C. AUC value, estimated on the data assumed to be independent, that is not used for model calibration, generally indicated higher classification accuracy for specialist rather then for species with high tolerance to environmental gardients. For instance, for glacial liminic relict amphipoda *Pontoporeia affinis,* AUC value indicated very high degree of success. For bivalve *M. balthica* found almost allover the Baltic Sea and known to be very tolerant in response to environmental gradients, AUC value indicates the predictive power that can be regarded as only fair. Nevertheless, the consistency with data observations can be seen for this species as well. The considered environmental predictors are significant, but noticeable part of variance of species distribution is explained by some other factors not included in the analysis. For *Halicryptus spinulosus* model AUC values was relatively low.

Whereas, the combination of the factors considered seems to be suitable in the area north to the Aland 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. The probability of occurrence modelled for *A. islandica* known to prefer polyhaline environment corresponds with the observations very well. For this species it was also seen 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 analyzed.

4 Disscussion

Generally, this thesis contains studies on observational spatial scale that can be described as fine-grained (referring to the individual data units of grab samples), whereas spatial extend covered ranges from regional (Paper I-IV) to large (Paper V), though the subjectivity exists in the definition of the scale magnitude. As for temporal scales, they are not as simple to describe. As was pointed out by Praca et al. (2009) the use of temporally heterogeneous data confounds the effect of interannual variations in species occurrence and environmental conditions. The long-term distribution patterns were reflected in Paper V. However, our objectives were to attempt a general description of species habitats and to investigate the predictive abilities of the modelling techniques at the selected spatial scales. Models determining the distribution of exemplary macrofaunal species common for the particular regions and the whole Baltic Sea from changing environmental variables, such as depth, salinity and sediment characteristics, have been successfully developed. Those models are considered to be reasonably general, i.e. they are able to resolve the distribution of species over a comparatively large spatial scale as opposed to a metre-by-metre basis.

The static, comparative, models (those the presented thesis is mainly focused on) are opposed to more mechanistic models of ecosystem processes (Lischke et al., 1998). The use of and theoretical limitations of static models compared with dynamic approaches have been described e.g. in Lischke et al. (1998), who suggested that they do not provide a specific answer to the question of how change would affect future species' distributions, but rather attempt to explore the possible ranges of change and thus offer a basis for discussion about, for example, species' potential to reach new habitats. Moreover, they constitute the basic layers for future spatiotemporal distribution models.

Thus, the present study represents the first steps towards a benthic – abiotic interaction model using the data sets of varying complexity, based on recent investigations (Papers I-IV) and utilizing the historical data (Paper V) for the Baltic Sea regions. Such models allow making predictions on how and to what extent natural or anthropogenic influences affect benthic community assemblages. They can be a useful tool in marine ecosystem management and environment conservation planning (Pavlikakis and Tsihrintzis, 2000). However, the extraction of patterns of benthic community distributions and exploratory statistical description of the prevailing ecological structure based on in-situ observations (Paper I) is always an indispensible first step (Bourget and Fortin, 1995). Also, to assess the correctness of the calculated biotic – abiotic dependencies indispensably implies not only the geographical and ecological discrimination, but also the in-depth knowledge of autecology of the species. This expert knowledge (descended from study, utilization of physiological experiments on species, observation, habitat monitoring as well as experience) is the initial requirement that cannot be simply deduced from statistics (Sachs 1997). This is especially important since top-down correlations between the macrobenthos and the environmental variables represented by response models do not always reflect direct cause and effect relationships, since many environmental variables co-vary (van der Wal et al., 2008). Thus only based on the prior knowledge statistically modelled descriptions of species distribution patterns can be explained by causality as per ecological behavior of species.

It is essential to differentiate between the questions that the two types of model solutions presented serve to answer. Each single-factor model alone is able to answer the question of how certain species respond to changes of this separate factor, e.g. describe it as euryoecious or stenoecious organism. Thrush et al. (2005) acknowledge that simple models may fit well the purposes of response management, for instance by defining the sufficiently consistent general pattern of which species prefer muddy, intermediate or sandy sediment types across scales, so that a rank order of species can be developed. Such classification can be used to interpret changes in the distribution and abundance of monitored species or incorporated with other predictions of habitat change, and used in environmental risk assessment. The multimodel inference serve to predict the distribution of species within the common limits of combined environmental factors, e.g. to fill out the lack of information in the locations, where no observational data is available.

Thus, the results of statistical modelling of species distribution can be useful for: testing the hypotheses about the ranges of species distribution along environmental gradients and benthic stress impacts (for instance, the Pearson and Rosenberg (1978) model that reported a gradual loss of species as the degree of stress increased over space and/or time driven by the replacement pattern defined by different tolerance of species to stress); generation of habitat suitability maps that predict the specific ecological potential of a habitat (with limitations defined by the data analyzed) which can be considered in marine spatial planning and conservation management (Degraer et al., 2008); and predicting the possible consequences of habitat changes (either natural or antropogenic). Moreover, when accompanied by other relevant developments and investigations, 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). Within this context, the two particular approaches, the biological trait analysis (BTA) and the studies of impact of presence of macroinvertebrates on near-bed fluxes, are discussed in more details below.

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 the 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 (Mermillod-Blondin and Rosenberg, 2006; Crain and Bertness, 2006).

4.1 Ecosystem engineers, BTA

The presence or activities of invertebrate species in aquatic ecosystem 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 habitat modifying ecosystem engineers 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. In more benign environments engineers will tend to increase species coexistence and biodiversity or retain specific ecosystem functions, whereas in stressful environments (like most parts of the Baltic Sea), physical modifications can effectively create new habitat and enable establishment by organisms that would otherwise be unable to persist. In latter case, engineers are of critical importance as they are essential to any population of the habitat. Crain and Bertness (2006) therefore conclude, that across environmental stress gradients, most critical engineer species that retain community and ecosystem integrity and function can be identified and considered as a primary conservation target.

In their overview article Mermillod-Blondin and Rosenberg (2006) have linked invertebrate diversity to ecosystem functioning comparing functional groups of bioturbators (based on their mechanistic effects in reworking the sediments), and stressed the importance of accounting for physical habitat features (e.g. distinguish between diffusion-dominated benthic system with fine-grained sediments and advection-dominated system with coarse sediments) within a wider conceptual framework of functional approach. They highlighted that the impact of bioturbation on microbial activities in sediments decreases with increasing physically-induced interstitial flow rates within the system, supporting the statement that importance of ecosystem engineering is strongest when the engineered habitat provides a resource that is not available in the unengineered environment. One good example for that are the mussel shells that provide a substratum for the attachment of invertebrates and algae in soft-bottomed sediments, whereas it is not the case in rocky habitats where the structure of the unengineered environment provides habitats for the attachment of the organisms.

Alternatively, the ability of an ecological community to withstand disturbances (resilience) can be regarded as a function of functional (rather then species) diversity. The relevant questions stated by management and conservation of ecosystem functioning and resilience are then: what are the ecological commodities and services the system yields; which trophic levels are involved; what is the level of functional redundancy; is the system intrinsically prone to environmental extremes or subject to intense variability; to what extent is the system already (anthropogenically) disturbed (Ieno et al., 2006)?

The functional group-approach (BTA) to the analysis of benthic assemblages promises to provide more insights on ecosystem processes and ecosystem functioning as a whole and to enable the adaptation of a predictive approach with regard to assumed ecological disturbances, rather than a momentary snapshot of the benthic ecosystem quality (Bremner et al., 2006). The many effects on physico-chemical settings of the environment caused e.g. by dredging and construction have high potential amount of confounding factors, thus requiring a large number of rigorous analysis of effects in highly controlled environments (i.e. laboratory or mesocosm experiments) to enable any rigorous statements regarding species effects. According to Loreau (2002) the “functional groups” can be defined as sets of species that show either similar responses to the environment or similar effects on ecosystem processes. Ieno et al. (2006) states that in naturally occurring communities, the individual contribution of species effects to overall ecosystem functioning is a collective product of diversity effect (niche complementarily), species identity and species density, with differences in relative contribution of these components pronounced over the space and time. The fact that the effect of one species can masks other species effects supports the idea that functional diversity may be equally, or even more important in securing or delivering ecosystem processes. BTA uses a coded/scored information on series of life history, morphological and behavioural characteristics (trophic relations, dispersal methods, sensitivity etc.), as well as ecosystem engineering effects (potentially important traits can be defined subject to the ecosystem function(s) of interest) of species present in assemblages to indicate aspects of their ecological functioning (Padilla and Allen, 2000; Bremner, 2008). Changes in the patterns of trait expression within or between the assemblages, in terms of changes in the relative abundance/biomass of taxa exhibiting the traits, can be used to indicate the effects of human impacts on ecological functioning and compare the functioning of communities over large geographical ranges (as biological traits can be shared by organisms of different taxonomic identity).

The discriminating ability of the BTA rises with the number of traits analysed, which causes the potential problem related to the time and information required for preparing large trait databases to carry out the analysis, aggravated by the fact that information for each taxa of a pool might not be available and the trait information itself is not static. Thus Bremner (2008) notes, that especially systems with a relatively low species diversity and where there is a good knowledge of the environmental gradients, like the Baltic Sea, are deemed good areas to develop models of traits analysis.

4.2 Influence of benthic organisms on transport of sedimentary material

Sediment distribution patterns and its genesis are mainly driven by transport, erosion and accumulation of sedimentary material and refined insight into the corresponding processes is required by the decision-making authorities to evaluate impacts of dumping of dredged matter, ecological effects of submarine mining, and the transport of associated contaminants (Gray and Elliot, 2009; Harff et al., 2009). The effects of planned anthropogenic activities can be best estimated by investigating scenarios of the process under debate using numerical simulations and its results (Harff et al., 2009). Quantifying the effects of interactions between organisms, the sediments, and the flow field on the sediment transport is the subject of much current research. Organisms alter the flow field via the bottom roughness, modify sediment characteristics by breaking up aggregates or coating sediments with organic matter, or change the structure of the sediments during burrowing activities (e.g. Rhoads, 1974). Discussing in details the parameterization of sediment transport exceeds the framework of this study, thus, only the very general information will be included here in order to identify the link between sediment dynamics and the distribution of macroinvertebrate communities.

There can be distinguished the “physical roughness" (k_b) and the “hydrodynamic roughness", or roughness length, z_0. Both have units of length (m, cm, etc.). The “physical roughness" is a physically meaningful quantity - such as sediment grain size, ripple dimension, etc. Hydrodynamic roughness, z_0, is related to physical roughness, and to the flow structure, but it is really a mathematical construct. It is the height, z, at which a velocity profile, u(z), would go to zero if it were extrapolated towards the boundary. Abiotic as well as biotic factors should be taken into account within parameterization of bed roughness. Often, the biological effects are neglected (Kuhrts et al., 2004), in other cases the passive influence of marine benthic fauna are considered, without taking into account the influence of marine flora (Bobertz et al., 2009; Seifert et al., 2009). Usually, what effects and processes are considered in the modelling is the matter of questions addressed, knowledge on peculiarities of the investigation area, the appropriate parameterization defined and the data available. As was mentioned in the introduction, generally the processes involved in the sediment transport mechanisms can be summarized by the conceptual scheme (Fig. 4) from Nittrouer and Wright (1994).

Rhoads (1974) have published the pioneered systematic work devoted to the biological effects on the seafloor roughness and stability, where he defined the edges for both sedimentologists and ecologists, showing how organisms could control or otherwise modify the geological and geochemical properties of the sea floor, such as grain-size, water content, porosity, compaction, shear strength, biodeposition, bioadvection (bioturbation), and the critical threshold velocity for erosion. According to Jumars and Nowell (1984) benthic organisms can affect sediment transport via alteration of (1) fluid momentum impinging on the bed, e.g. caused by animal tubes in a flow, (2) particle exposure to the flow induced e.g. by burrowing invertebrates moving within the sediments, (3) adhesion between particles, e.g. in a conical mound of fecal pellets produced by burrowing bivalve and (4) particle momentum, e.g. via ejection of fecal pellets from the feeding pit by deposit feeding polychaete. They also concluded that consistent grouping of organisms as stabilizers vs. destabilizers, respectively decreasing or enhancing erodibility, is not possible. Back then the net effects of species or even individual on erosion and deposition thresholds seemed in general unpredictable. Authors highlighted the dependence of those effects upon the context of flow conditions, bed configuration, and community composition and claimed that separation of organism effects into the listed above four categories permits the specification of parameters for stochastic sediment dynamic models to enhance predictability of natural organisms-influenced marine settings. Unfortunately, even the present level of information hardly allows the robust definition of functional traits for most taxa of marine benthic communities (authors' own judgment, see the section on BTA), however a significant step forward was achieved supported by the results from a number of conducted experimental setups (e.g. Friedrichs, 2004; Peine, 2005; Rietschel, 2005; Schmid, 2009; Friedrichs et al. 2009). The conclusions of Jumars and Nowell (1984) suggested that if major sediment transport occurs during the infrequent large-magnitude events organism effects can be ignored, whereas benthic activity and bed forms gain significance in sediment transport only during low turbulence and at low velocity values and should only be considered if the variable of primary concern is not the total amount of sediment transported but rather the frequency of transport events or the spatial pattern of erosion and deposition.

In Dade et al. (2001) recent studies have been summarized of how biota influences the turbulence and velocity field near the sea floor. In general, biota greatly increases the physical roughness of the sea floor (k_b), and thereby the hydrodynamic roughness (z_0) compared to the roughness of a flat bed. The increased roughness comes from (1) pelletization of sediment into coarser aggregates by benthic organisms, (2) the presence of biogenic mounds or tracks, (3) protrusion of the organisms into the flow (worm tubes, grasses). Biological alterations of sediment properties or particle fluxes include both complex interactions and contradictory effects: for instance, bioturbation is assumed to facilitate erosion, whereas macrofauna mucus will consolidate the sediment; results of feeding tactics such as excretion or active particle expulsion by deposit feeders, or capture and biodeposition of suspended matter and refiltration of water by filter feeders (Orvain et al., 2004; Widdows et al., 2004; Peine, 2005) also can have a direct influence on sediment dynamics.

In 2000-2005 western part of the Baltic Sea served as an investigation area of the national German project on the dynamics of natural and anthropogenic sedimentation (DYNAS; Harff et al., 2009) that was based on multidisciplinary cooperation between sedimentologists, benthos- and microbiologists, sediment physicists and physical oceanographers. The project aimed to contribute to a profound understanding of sedimentation processes and to predict sediment transport in the region by means of numerical modelling.

Within the framework of the project Kuhrts et al. (2004) developed the numerical simulation tool linking the Baltic Sea Oceanographic Model (BOM; Fennel and Seifert, 1995; Schmidt et al., 1998) based on the Princeton Modular Ocean Model 3 (MOM-3; Pacanowski and Griffies, 2000) with a wave model, a bottom boundary and a sediment transport model, that was further improved in Seifert et al. (2009). The transport of sedimentary material in suspension was based on the co-called tracer equation $\frac{\partial c}{\partial t} + \vec{\nabla}(c \cdot \vec{u}) + \frac{\partial}{\partial z}(c \cdot w_{\sin k}) = \vec{\nabla}(v \vec{\nabla} c)$, where c is the concentration of the suspended sedimentary material, $\vec{u}$ is the full 3D current velocity, $w_{\sin k}$ is the settling velocity of the considered sediment type and v is the eddy diffusivity. The bottom boundary conditions determine the deposition or resuspension processes via the source term Q (the amount of sediment per unit of area to be resuspended or deposited at the bottom) that depends on the total skin friction velocity u_s^* and specific material constants:

$$Q = \begin{cases} (w_{\sin k} \cdot c)_{bottom,} & u_s^* < u_{cr}^* (deposition) \\ 0, & u_s^* = u_{cr}^* (equilibrium) \\ q_{r,} & u_s^* > u_{cr}^* (erosion) \end{cases}.$$

Here and q_r is the erosion rate, and u_{cr}^* is the critical shear velocity of considered sediment material which identifies the threshold, the point of incipient motion of the sediment particles. The latter can be approximated as function of mean grain size, density of sediment particles and water density, gravitational constant and water viscosity (e.g. Shields, 1936; Soulsby and Whitehouse, 1997 – for cohesive sediments; Hjulström, 1935; Zanke, 1977 – for cohesionless seiments). A brief summary on the widely used threshold approach can, for instance, be found in Bohling (2009) and Bobertz et al. (2009), though critics exist, e.g. due to the fact that no flow stage exists at which particles are suddenly placed in motion in a massive amount (e.g. Grass, 1970; Unsöld and Walger, 1987; Lavelle and Mofjeld, 1987). In the bottom boundary layer model developed by Kuhrts et al. (2004) and Seifert et al. (2009) the particle dynamics at the sea bottom are influenced by the shear stress velocity, in turn determined by waves and currents energy (meteorological forcing) and by the roughness of the sea floor. According to Bobertz et al. (2009) the roughness length, in the case of absence of macrozoobenthos and their structures, can be parameterized based on grain size too, employing Soulsby (1997), Nielsen (1983) and Yallin (1977). No equivalent simple parameterization for biological caused bed roughness exists. With some assumptions, findings of Friedrichs (2004) and results of the DYNAS project were employed to derive the roughness length from the abundance measurements of four selected key species (*Arenicola marina*, *Lagis koreni*, *Mytilus edulis*, *Pygospio elegans*). This was done using the estimated relationship between the roughness length and the roughness density, RD, which is the percent of the sediment surface covered by the population. In Bobertz et al. (2009) the proxy-target concept was used to extrapolated the roughness length (an input parameter for the modelling setup in Seifert et al., 2009) into the larger investigation area (bathymetry, median grain size, salinity and oxygen served as proxy variables); a combined hierarchical and supervised classification was applied. The basic assumption is that similar milieu conditions result in similar roughness densities. For this work, the macrozoobenthic data was the basis to derive the roughness density. Authors note that the appearance of macrozoobenthos may not be explained by the environmental parameters considered completely, thus stressing the importance of investigations of species-environmental relationships. The employment of methods developed in the present study (Papers III-V) can serve as a good supplement, aiming at utilization of more realistic distribution surfaces for the most important species. In order to include the results of flume experiment in the large-scale numerical simulation model for sediment transport the population density-dependant net sediment fluxes have to be coupled with faunal distribution data (Friedrichs et al., 2009). The results of data inventory can thus present a valuable contribution (Paper V). Peine et al. (2005) have visualized in ArcGIS the effect of the biogenic structures generated by blue mussel clumps (*Mytilus edulis*) on the roughness length (z_o). The GIS data base compiled in the present study can serve the common purposes.

Quantitative corroboration of effects of small-scale biological processes on the large-scale sediment transport and deposition was done in Borsje et al. (2008). Biological effects on cohesive sediment transport in the tidal Western Wadden Sea were quantitatively incorporated into the sediment transport module. The transport of fine suspended sediments was based on the advection-diffusion equation, with the source/sink terms that described the vertical fluxes between the bed and the water column as a result of erosion and deposition. The biological influence was parameterized through the empirical stabilizing and destabilizing factors (as functions of the chlorophyll-a concentrations and grazers biomass, respectively) for the critical bed shear for erosion and the erosion coefficient. For the modelling the grazers biomass found at two measurement sites was assigned to the whole study area based on classification of three different depth zones. One year setup was performed to capture the seasonal changes. Based on modelling results Borsje et al. (2008) concluded that seasonal variation in the sediment concentration is caused by the combining effect of the suspended sediments concentrations, wind and biological activity, with organisms assumed to be stabilizing (in this case diatoms) mainly responsible for seasonal variations in concentrations of suspended sediments, while the destabilizing ones (here the snail *Hydrobia ulvae* and the clam *Macoma balthica*) are mainly important for the spatial variation in the fine sediment on the bed. Subsequently, Borsje et al. (2009) have published the results of modelling of bio-geomorphological influences for offshore sandwaves, focusing on the effects of three benthic species characteristic for the North Sea: the sea urchin *Echinocardium cordatum*, the tube building worm *Lanice conchilega*, and the clam *Tellina fabula*. Mathematical inclusions to account for biological activity in idealized models suggested that biota is able to influence the wavelength of sandwave significantly, compared to the case of their absence. More importantly, the models indicated that biota is able to induce bed patterns while the physical parameters suggest a stable flat bed and vice versa. Authors stressed that future research should focus on improving the parameterization of biological activity influence on sediment dynamics and thereby on seabed patterns, but moreover, on gathering the site specific field data both on physical parameters and biological activity indispensable for model validation.

If the relationships between sediment fluxes and population densities are available, the data describing the distribution of macrofaunal population densities is required to integrate them into a numerical sediment transport model. Since one important feature of benthic assemblages is their intrinsic patchy distribution, the distribution of possible biogenic effects on sediment dynamics is also patchy. Works of Borsje et al. (2008), Bobertz et al. (2009), Seifert et al. (2009), Friedrichs et al. (2009) deal with this in their recent papers, although there were criticisms and skepticisms of technical nature (Seifert, Forster, pers. comm.). Nevertheless, the superimposition of biota-mediated sedimentation/erosion processes on the spatial extent should be feasible and should provide us with predictive information (of biotic effects) on sediment properties.

To complete the circle, in turn these sediment properties can then be used to make predictions on the ecological functioning and suitability of the sediment as a habitat (Diaz et al. 2004). Habitat suitability based on known gradients can serve as a forcing function for the modelling of population dynamics. Yet in turn, the result of the hydrodynamic and biogeochemical modelling can be translated in a potential for the existence of certain biological traits and used further for BTA.

5 Concluding remarks

The distribution of macrofaunal species of the Baltic Sea in response to selected environmental factors was investigated on varying scales using the exploratory statistical analysis approach and the developed predictive modelling techniques.

The first regional case study was focused on the limited area attached to the region of the Mecklenburg Bight. By defining the study area, we aimed to lessen the dominance of near-bottom salinity and oxygen concentration (known to be the dominating factors defining the Baltic Sea biodiversity) in the analysis to illuminate the impact of other abiotic factors considered. The exploratory statistical description illuminated the prevailing ecological structure. Distinct benthic assemblages discriminated by particular species (*Hydrobia ulvae–Scoloplos armiger*, *Lagis koreni–Mysella bidentata* and *Capitella capitata–Halicryptus spinulosus*) were related to different spatial subarea, characterised by a certain variability of environmental parameters. Water depth remained the key factor determining the species distribution among the parameters considered within the regional study scale in the region of the Mecklenburg Bight (Papers I-III). Total organic content and sorting also had noticeable effect. The method combining the parsimony of single-factor logistic regression models with an AIC solution of multimodel inference was applied to model and map the probabilities of occurrence for selected exemplary species (*Arctica islandica, Hediste diversicolor, Pygospio elegans, Tubificoides benedii* and *Scoloplos armiger*). It performed fairly well and can therefore be recommended on comparable spatial scales and environmental gradients. It is obvious, that only a fragment of control factors network responsible for macrozoobenthos distribution could be covered with the available data. The food web structure, complex predation interactions, flow energy and currents, temperature, turbidity, chemical substrate composition were not considered in this study. Additional variables need to be incorporated for more insightful analysis.

The similar approach was used to model the response of benthic species to their physical environment in the Pomeranian Bay (southern Baltic Sea). In the scheme of the dominance of strong salinity gradient over the brackish system, consistently small patches of comparatively higher or lower benthic diversity do emerge in areas where either environmental or anthropogenic impacts on the benthic habitat change drastically over short spatial distances. For instance, according to Zettler et al. (2007) the Shannon–Wiener diversity index ranges in various areas of the Pomeranian Bay approximately from 1 to 3.9. Hence, spatial diversity of ecological factors creates diversity among benthic colonization and community structures. Considering the fact that all species examined in the Pomeranian Bay (Glockzin and Zettler, 2008; Paper IV) are regarded as euryhaline and that in the central part of the region salinity varies only moderately, the apparent influence of salinity as the main predictor on the benthic assemblage becomes insignificant under certain scale, and depth was evidenced as a “master factor” mainly responsible for the spatial zonation of benthic species. The study clearly demonstrates the possibility to model species’ relationships in gradient systems such as the Baltic Sea where their patterns of distribution are strongly and directly coupled to abiotic processes.

Salinity, bathymetry and sediment 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, bathymetry and rough sediment class information (Paper V). Models performed comparatively well in the whole sea, however their applicability outside the Baltic should be considered at least questionable. They require farther 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 presented exercise is only a first step. Implementation of other variables (e.g. characterizing temperature fluctuations, total organic content, nutrient supply) would obviously increase the model applicability. These 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 (Degraer et al. 2008). Presented habitat suitability maps predict the specific ecological potential of a habitat (the background) rather that a realized ecological structure, with limitations defined by the predictors considered, their range, and the specified scale.

On different spatial scales different factors most influence species distribution. The present thesis can be regarded as a baseline for creating a tool for predicting benthic community changes and its effects on the ecosystem functioning and last but not least for developing an instrument for conservation and management purposes. Bearing in mind the currently experienced impact of natural (e.g. via climate change) or anthropogenic forced changes (e.g. by pipeline building or dredge fishing) on the Baltic Sea ecosystem – never has the need for such a tool been more urgent than nowadays.

6 Future challenges

The presented study represents only the momentary state of benthic communities. But, for instance, temporal development of the south-western Baltic Sea has shown both high variability in species abundance and shifts in species composition (Andersin et al. 1978, Perus et al. 2004, Zettler et al. 2008), thus temporal aspects are to be incorporated in the future work.

Generally, model performance will benefit from inclusion of oxygen dynamics, sediment total organic content, nutrients etc. Until now only probability of occurrence based on presence/absence data was modelled, but abundance density of species is more informative (Thrush et al. 2003), and models describing it are required even more. Such methods as ‘factor ceiling’ or quantile regressions can be the solution.

Without accounting for distribution of common benthic invertebrates regarding the background variables in aquatic ecosystems, it is impossible to build a rigorous concept of conserving and restoring damaged ecosystems (Mermillod-Blondin and Rosenberg, 2006; Crain and Bertness, 2006). BTA represents a promising approach to gain the desired insights on ecosystem functioning and biodiversity-mediated processes. The results of faunal distribution modelling can be useful to generate an input data (for corresponding biologically-induced parameters) in the sediment transport numerical simulation model (e.g. Borsje et al. 2008, Friedrichs et al. 2009, Bobertz et al. 2009, Seifert et al. 2009)

Acknowledgments

Naturally, without colleagues, friends and family this thesis would not exist. There is no way to name all the people I am grateful to for their contribution to it. However, those who are not named here are not neglected, but are all secretly honoured in my thoughts. First I would like to acknowledge my official supervisor, Prof. Jan Harff, who found it possible to initiate this work, regardless of the many businesses he is always occupied with, and Prof. Dr. Detlef Schulz-Bull, who introduced us. Especially, I would also like to thank Dr. Michael L. Zettler for continues and stable guidance of my research, and Dr. Falk Pollehne for the advisement in all cases, including emergency. I gratefully acknowledge the work of all colleagues involved and responsible for field work and laboratory analysis, in particular Ines Glockzin, and others from the WG ‘Ecology of benthic organisms’: Dr. Ralf Bochert, Alexander Darr, Monique Utech, Nadine Keiser. I feel obliged to express great appreciation to Dr. Bernd Bobertz, Dr. Michael Meyer and Michael Glockzin for the valuable assistance on mathematical, statistical and GIS matters, but also dropping by or sitting behind the wall, and to Nicole Kowalski, for being permanently there. I also thank Dr. Torsten Seifert for his great help with modelled data on hydrographical parameters. Sincere thanks to the Hanse-Jobcenter Rostock and to Marion Madsack for helping to deal with it and to fight it. I thank all colleagues from the Leibniz Institute for Baltic Sea Research Warnemünde, the Shirshov Institute of Oceanology, RAS, and the M.V. Lomonosov Moscow State University, Faculty of Geography, Division of Mainland Hydrology. I am most grateful to all units of my family for the endless support, and my friends, who always stand behind me. Part of this work was funded by the EU-MARBEF. It would be impossible to perform it without the work of HELCOM and ICES database team, and the results of the BALANCE project.

Reference list

Al-Hamdani, Z., Reker, J., (Eds.), 2007. Towards marine landscapes in the Baltic Sea. BALANCE interim report #10. Available at 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 Marine Research 244, 137–144.

Andersin, A.-B., Lassig, J., Sandler, H. 1976. Community structures of soft-bottom macrofauna in different parts of the Baltic. in Keegan, B.F., Cei-Digh, P.O., Boaden P.J.S., (eds.): Biology of benthic organisms. Eleventh European Symposium on Marine Biology, Galway, October 1976. Pergamon Press, London.

Austin, M.P., 1987. Models for the analysis of species' response to environmental gradients. Vegetatio 69, 35–45.

Bobertz, B., 2000. Regionalisierung der sedimentären Fazies der südwestlichen Ostsee.- Dissertation, Ernst-Moritz-Arndt Universität, Greifswald, 121 pp.

Bobertz, B., Harff, J., 2004. Sediment facies and hydrodynamic setting: a study in the south western Baltic Sea. Ocean. Dynam. 54, 39-48.

Bobertz B., Harff J., Bohling B., 2009. Parameterisation of clastic sediments including benthic structures. J. Mar. Sys. 75, 371-381.

Bobertz, B., Kuhrts, C., Harff, J., Fennel, W., Seifert, T., Bohling, B., 2005. Sediment properties in the Western Baltic Sea for the use in sediment transport modelling. J. Coast. Res. 21, 288–597.

Bohling, B., 2009. Correlation of threshold values for incipient motion of sediment particles with two different erosion devices, J. Mar. Sys. 75, 330–335.

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.

Borsje, B.W., De Vries, M.B., Bouma, T.J., Besio, G., Hulscher, S.J.M.H., Herman, P.M.J., 2009. Modelling bio-geomorphological influences for offshore sand waves. Cont. Shelf Res. 29, 1289-1301.

Borsje, B.W., De Vries, M.B., Hulscher, S.J.M.H., De Boer, G.J, 2008. Modeling large-scale cohesive sediment transport with the inclusion of small-scale biological activity. Estuar. Coast. Shelf Sci. 78, 468-480.

Bourget, E., Fortin, M.-J., 1995. A commentary on current approaches in the aquatic sciences. Hydrobiologia 300/301, 1–16.

Bremner, J., 2008. Species' traits and ecological functioning in marine conservation and management. J. Exp. Mar. Biol. Ecol., 366(1-2), 37-47.

Bremner, J., Rogers, S., Frid, C., 2006. Methods for describing ecological functioning of marine benthic assemblages using biological traits analysis (BTA). Ecol. Indic. 6(3), 609-622.

Bromley, R.G., 1996. Trace Fossils: Biology, Taphonomy and Applications, second ed. 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.

de Candolle, A.I., 1855. Ge´ographique botanique raisonne´e. Masson, Paris.

Carman, P.C., 1937. Fluid flow through a granular bed. Trans. Inst. Chem. Engng. 15, 150–167.

Clarke, K.R., Gorley, R.N., 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E: Plymouth.

Clarke, K.R., Warwick, R.M., 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E: Plymouth.

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

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. Freshw. Res. 58, 293–305.

Dade, W.B., Hogg, A.J., Boudreau, B., 2001. Physics of Flow above the Sediment-Water Interface. In Boudreau, B., Jorgensen, B.B., (eds.) The benthic boundary layer: Transport processes and biogeochemistry. Oxford University Press, 404 pp.

Darcy, H., 1856. Dètermination des lois d'ècoulement de l'eau à travers le sable. In Les Fontaines Publiques de la Ville de Dijon. Victor Dalmont, Paris, pp. 590–594.

Davis, J.C., 2002. Statistics and Data Analysis in Geology, third ed. John Wiley and Sons, New York, 638 pp.

Degraer, S., Verfaillie, E., Willems, W., Adriaens, E., Van Lancker, V. and 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.

Diaz, R.J., Solan, M., Valente, R.M., 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. J. Environ. Manag. 73(3), 165-181.

Ekman, S., 1953. Zoogeography of the Sea. Sidgwick and Jackson, London.

Ellis, J., Ysebaert, T., Hume, T., Norkko, A., Bult, T., Herman, P., Thrush, S., Oldman, J., 2006. Predicting macrofaunal species distribution in estuarine gradients using logistic regression and classification systems. Mar. Ecol. Prog. Ser. 316, 69-83.

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

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

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

Fennel, W., Seifert, T., 1995. Kelvin wave controlled upwelling in the western Baltic. J. Mar. Sys. 6, 289–300.

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.

Franklin, J., 1995. Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geog. 19, 474–499.

Folk, R.L., Ward, W.C., 1957. Brazos river bar: a study in the significance of grain size parameters. J. Sed. Petrol. 27, 3-26.

Forster, S., Bobertz, B., Bohling, B., 2003. Permeability of sands in the coastal areas of the southern Baltic Sea: mapping a grain-size related sediment property. Aquat. Geochem. 9, 171–190.

Fortin, M.-J., Keitt, T.H., Maurer, B.A., Taper, M.L., Kaufman, D.M., Blackburn, T.M., 2005. Species' geographic ranges and distributional limits: pattern analysis and statistical issues. Oikos 108, 7-17.

Friedrichs, M., Leipe, T., Peine, F., Graf, G., 2009. Impact of macrozoobenthic structures on near-bed sediment fluxes, J. Mar. Sys. 75, 336-347.

Friedrichs, M., 2004. Flow-induced effects of macrozoobenthic structures on the near-bed sediment transport. Dissertation, Universität Rostock, 80 pp.

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.

Glockzin, M., Zettler, M.L., 2008. Spatial macrozoobenthic distribution patterns in relation to major environmental factors - a case study from the Pomeranian Bay (southern Baltic Sea). J. Sea Res. 59, 144-161.

Grass, A.J., 1970. Initial instability of fine bed sand. Journal of Hydraulic Research 96 (HY3), 619–632.

Gray, J.S., 1981. The Ecology of Marine Sediments.Cambridge, UK: Cambridge Univ. Press.

Gray, S., Elliott, M., 2009. Ecology of marine sediments: from science to management, second ed. Oxford University Press, New York.

Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modeling of plant species distribution. Plant Ecol. 143, 107–122.

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

Harff, J., Bobertz, B., Graf, G., 2009. Preface. Dynamics of Natural and Anthropogenic Sedimentation. J. Mar. Sys. 75, 315-316.

Harff, J., Davis J. C., Olea R. A., 1992. Quantitative Assessment of Mineral Resources with an Application to Petroleum Geology, Non-renewable Resources, pp. 74–84. Oxford University Press, London.

Harris, C. K., Wiberg, P. L., 1997. Approaches to quantifying long-term continental shelf sediment transport with an example from the northern California STRESS mid-shelf site. Cont. Shelf Res. 17(11), 1389-1418.

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. Eutrophication in the Baltic Sea – An integrated thematic assessment of the effects of nutrient enrichment and eutrophication in the Baltic Sea region: Executive Summary. Balt. Sea Environ. Proc. No. 115A.

von Humboldt, A., Bonpland, A., 1807. Essai sur la ge´ographie des plantes. Paris.

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

Hjulström, F., 1935. Studies in the morphological activity of rivers as illustrated by the river Fyris. Geological Institution of University of Uppsala: Bulletin 25, 221–528.

Hosmer, D., Lemeshov, S., 2000. Applied logistic regression, second ed. Wiley, New York.

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

Ieno, E.I., Solan, M., Batty, P., Pierce, G.J., 2006. How biodiversity affects ecosystem functioning: roles of infaunal species richness, identity and density in the marine benthos. Mar. Ecol. Prog. Ser. 311, 263–271.

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

Jumars, P.A., Nowell, A.R.M., 1984. Effects of benthos on sediment transport: difficulties with functional grouping. Cont. Shelf Res. 3 (2), 115–130.

Karle I.-M., 2006. On redox reactions and transport processes of solutes in coastal marine sediment. Dissertation, Analitical and Marine Chemistry, Göteborg University.

Kemp, W.M., Petersen, J.E., Gardner, R.H., 2001. Scale-dependence and the problem of extrapolation: implications for experimental and natural coastal ecosystems, in Gardner, R.H., Kemp, W.M., Kennedy, V., Petersen, J.E. (Eds.), Scaling relations in experimental ecology. Columbia University Press, New York. pp. 3-57.

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.

Krumbein, W.C., Monk, G.D., 1942. Permeability as a function of the size parameters of unconsolidated Sand. American Institute Mining and Metallurgy Engineering, Petroleum Technology, Technical Publication.

Kuhrts, C., Fennel, W., Seifert, T., 2004. Model studies of transport of sedimenary material in the western Baltic. J. Mar. Sys. 52, 167-190.

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.

Lavelle, J.W., Mofjeld, H.O., 1987. Do critical stresses for incipient motion and erosion really exist? J. Hydraul. Eng. 113, 370–385.

Legendre, P. 1993. Spatial autocorrelation: Trouble or new paradigm? Ecol. 74, 1659-1673.

Legendre, P., Gallagher, E.D., 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129, 271-280.

Legendre, P., Legendre, L., 1998. Numerical ecology, second ed. Developments in environmental modeling, 20. Elsevier, Amsterdam.

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

Leps, J., Smilauer, P., 2003. Multivariate Analysis of Ecological Data Using CANOCO. Cambridge University Press, Cambridge.

Levin, S.A., 1992. The problem of Pattern and Scale in Ecology. Ecol. 73, 1943-1967.

Lischke, H., Guisan, A., Fischlin, A., Bugmann, H., 1998. Vegetation responses to climate change in the Alps-Modeling studies. In: Cebon, P., Dahinden, U., Davies, H., Imboden, D., Jaeger, C. (Eds.), A View from the Alps: Regional Perspectives on Climate Change. MIT Press, Boston, pp. 309–350.

Longhurst, A.R., 2006. Ecological Geography of the Sea, second ed. Elsevier, 542 pp.

Loreau, M.E.A., 2002. Biodiversity and ecosystem functioning, Oxford University Press.

Meissner, K., Darr, A., Rachor, E., 2008. Development of habitat models for *Nephtys* species (Polychaeta: Nephtyidae) in the German Bight (North Sea). J. Sea Res. 60, 271–286.

Mermillod-Blondin, F., Rosenberg, R., 2006. Ecosystem engineering: the impact of bioturbation on biogeochemical processes in marine and freshwater benthic habitats. Aquatic Sciences, 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.

Nielsen, P., 1983. Analytical determination of nearshore wave height variation due to refraction shoaling and friction. Coast. Eng. 7, 233– 251.

Nittrouer, C.A., Wright, L.D., 1994. Transport of particles across continental shelves. Reviews of Geophysics 32 (1), 85–113.

O'Brien, K., Hänninen, J., Kanerva, T., Metsärinne, L., Vuorinen, I., 2003. Macrobenthic zonation in relation to major environmental factors across the Archipelago Sea, northern Baltic Sea. Boreal Env. Res. 8, 159-170.

Ojaveer, H., Jaanus, A., MacKenzie, B.R., Martin G., Olenin, S., Telesh, I, Zettler, M.L., Zaiko, A., (in press). 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. http://www.corpi.ku.lt/nemo/

Orvain, F., Sauriau, P.-G., Sygut, A., Joassard, L., Le Hir, P., 2004. Interacting effects of *Hydrobia ulvae* bioturbation and microphytobenthos on the erodibility of mudflat sediments. Mar. Ecol. Prog. Ser. 278, 205–223.

Pacanowski, R.C., Griffies, S.M., 2000. MOM 3.0 Manual. Technical report, Geophysical Fluid Dynamics Laboratory.

Padilla, D.K., Allen, B.J., 2000. Paradigm lost: reconsidering functional form and group hypotheses in marine ecology. J. Exp. Mar. Biol. Ecol. 250(1-2), 207-221.

Pavlikakis, G.E., Tsihrintzis, V.A., 2000. Ecosystem Management: A Review of a new concept and methodology. Water Resour. Manag. 14, 257–283.

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

Praca, E., Gannier, A., Das, K., Laran, S., 2009. Modelling the habitat suitability of cetaceans: example of the sperm whale in the northwestern Mediterranean Sea. Deep-Sea Res. I. 56, 648-657.

Peine, F., 2005. Influence of subtidalmacrofauna organisms generating biogenic structures on the near-bottom particle transport in the southwestern Baltic Sea. PhD. Dissertation, University of Rostock, Germany. 121 pp.

Perry, R.I., Smith, S.J., 1994. Identifying habitat of marine fishes using survey data: an application to the Northwest Atlantic. Can. J. Aquat. Sci. 51, 589-602.

Perus, J., Bäck, S., Lax, H.-G., Westberg, V., Kauppila, P., Bonsdorff, E., 2004. Coastal marine zoobenthos as an ecological quality element: a test of environmental typology and Europeaan Water Framework Directive, Coast. Rep. 4, 27–38.

Rhoads, D.C., 1974. Organism-sediment relations on the muddy sea floor. Oceanogr. Mar. Biol. Ann. Rev. 12, 263–300.

Sachs, L., 1997. Angewandte Statistik. Anwendung statistischer Methoden. Springer, Berlin.

Sanders, H.L., 1968. Marine benthic diversity: a comparative study. American Naturalist 102, 243-282.

Scott, J.M., Heglund, P.J., Samson, F. Haufler, J., Morrison, M., Raphael, M., Wall, B. 2002. Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press, Covelo, California.

Seifert, T., Fennel, W., Kuthrs, C., 2009. High resolution model studies of transport of sedimentary material in the south-western Baltic. Journal of Marine Systems 75, 382–396.

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.

Schmid, M., 2009. Der Einfluß des Wattwurmes *Arenicola marina* (Polychaeta: Capitellida; Linnaeus 1758) auf den bodennahen Partikeltransport. Diploma thesis, University of Rostock, Rostock.

Schmidt, M., Seifert, T., Lass, H.U., Fennel, W., 1998. Patterns of salt propagation in the Southwestern Baltic Sea. Deutsche Hydrographische Zeitschrift 50 (4), 345–364.

Shields, A., 1936. Anwendung der Ähnlichkeits-Mechanik und der Turbulenzforschung auf die Geschiebebewegung. Mitteilungen der Preussischen Versuchsanstalt für Wasserbau und Schiffahrt (26) 26 pp.

Soulsby, R.L., 1997. Dynamics of Marine Sands: a Manual for Practical Applications. Thomas Telford Publications, London. 249 pp.

Soulsby, R.L., Whitehouse, R., 1997. Threshold of sediment motion in coastal environments. Proc. Pacific Coasts and Ports '97 Conf. University of Canterbury, Christchurch, pp. 149–154.

Simpson, M.R., Walsh, S.J., 2004. Changes in the spatial structure of Grand Bank yellowtail flounder: testing MacCall's basin hypothesis. J. Sea Res. 51 (3–4), 199–210.

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

StAUN, 2008. Wasserbaumaßnahmen in der Warnow/Bützow 2004 – 2008, URL: http://www.staun-mv.de/cms2/StAUN_prod/StAUN/de/ro/Themen/Nachhaltigkeit_und_Umweltbildung/index.jsp

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

ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecol. 67, 1167–1179.

ter Braak, C.J.F., Šmilauer, P., 2002. CANOCOReferenceManual and CanoDrawforWindows User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca NY, USA.

Tomczak, M., Godfrey, J.S., 2003. Regional Oceanography: An Introduction, Second Edition, Daya Publishing House, Delhi.

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.

Thomson, J.D., Weiblen, G., Thompson, B.A., Alfaro, S., Legendre, P., 1996. Untangling multiple factors in spatial distributions: lilies, gophers and rocks. Ecol. 77, 1698–1715.

Thorson, G., 1957. Bottom communities (Sublittoral of shallow shelf), in: Hedgpeth, J.W. (Ed.), Treatise on marine ecology and paleoecology, vol. I, Ecology, Geol. Soc. Amer. Mem. 67, 461–534.

Unsöld, G., Walger, E., 1987. Critical entrainment conditions of sediment transport, in: Rumohr, J., Walger, E., Zeitschel, B. (Eds.), Seawater– Sediment Interactions in CoastalWaters— An Interdisciplinary Approach. Springer Verlag, Berlin, pp. 210–219.

van der Wal, D., Herman, P.M.J., Forster, R.M., Ysebaert, T., Rossi, F., Knaeps, E., Plancke, Y.M.G., Ides, S.J., 2008. Distribution and dynamics of intertidal macrobenthos predicted from remote sensing: response to microphytobenthos and environment. Mar. Ecol. Progr. Ser. 367, 57–72.

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

Warzocha, J., 1995. Classification and structure of macrofaunal communities in the southern Baltic. Arch. Fish. Mar. Res. 42, 225–237.

Wheatcroft, R.A., 1994. Temporal variation in bed configuration and one-dimensional bottom roughness at the mid-shelf STRESS site. Continental Shelf Research, 14, 1167-1190.

Wiens, J.A., 1989. Spatial scaling in ecology. Funct. Ecol. 3, 385-397.

Widdows, J., Blauw, A., Heip, C.H.R., Herman, P.M.J., Lucas, C.H.,Middelburg, J.J., Schmidt, S., Brinsley, M.D., Twisk, F., Verbeek, H., 2004. Role of physical and biological processes in sediment dynamics of a tidal flat in Westerschelde Estuary, SW Netherlands. Mar. Ecol. Prog. Ser. 274, 41–56.

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.

Yallin, M.S., 1977. Mechanics of Sediment Transport. Pergamon Press, New York. 298 pp.

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. Hydrobiologia 375-376, 227-240.

Zanke, U.C.E., 1977. Neuer Ansatz zur Berechnung des Transportbeginns von Sedimenten unter Strömungseinfluss. Mitteilungen des Franzius-Instituts der Universität Hannover 46, 156–178.

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., Bönsch, R. Gosselck, F., 2000. Das Makrozoobenthos der Mecklenburger Bucht— rezent und im historischem Vergleich. Meereswissenschaftliche Berichte, 42, 1–144.

Zettler, M.L., Schiedek, D., Bobertz, B., 2007. Benthic biodiversity indices versus salinity gradient in the southern Baltic Sea. Mar. Pollut. Bull. 55, 258-270.

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

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.

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

Curriculum vitae

PERSONAL INFORMATION

Name	Mayya Gogina
Gender	Female
Born / Citizenship	May 21, 1982, Moscow, Russia / Russian
Marital Status	Married, 1 child

PROFESSIONAL EXPERIENCE

Jan 2010-present	Leibniz Institute for Baltic Sea Research Warnemünde IOW. WG Ecology of benthic organisms, work under the project “Fehmarnbelt Fixed Link” (Marine Biology Services).
Oct 2006 - Jun 2009	PhD student, Leibniz Institute for Baltic Sea Research Warnemünde IOW. WG Ecology of benthic organisms, work under MARBEF project. Statistical analysis, GIS. Marine field experiments and laboratory analysis related to identification of potential environmental “key factors” causing spatial distribution of benthic macrofauna in the Baltic Sea. Habitat mapping and modelling of benthic species distribution. Nomenclature Checklist, data inventory. Actualization of the Database.
Oct 2004 - Oct 2005	Engineer at P.P.Shirshov Institute of Oceanology of the Russian Academy of Sciences (SIO RAS), Moscow, Russia. Laboratory of Biohydrochemistry. Field experiments: currents and turbulence structure of the Black Sea and its effect on the distribution of hydrochemical parameters and biogeochemical processes.
Aug 2003 - Jan 2004	“ENBAS” LLC (Environmental engineering), Moscow, Russia working in the field of environmental engineering. Primary data acquisition, preparation of pre-project environmental documentation, electronic subject maps design
Sep 1999 - Jun 2004	Division of Hydrology, Facutly of Geography of the Moscow State University, Moscow, Russia. Student assistant. Experimental data processing, numerical simulations, field experiments in rivers and artificial lakes of Russia. Project “Evaluation of Contribution of the River-Flow and Coastal Processes within the Russian Arctic Zone to the Global-Scale Climatic and Environmental Changes"

EDUCATION

Jun 2004	Dipl. Geogr. (Hydroecology) dergee (Lomonosov Moscow State University), thesis title: “Solute flow from the European part of Russia into the Arctic Seas“, mark – excellent
1999-2004	Studies of geography. M.V. Lomonosov Moscow State University, Faculty of Geography, Division of Mainland Hydrology
1995-1999	College under the Academy of Innovation Management Specialization in the field of economics, accounting and controlling, Moscow
1994-1999	Moscow Secondary School (UVK 1876) with in-depth study of English
1993-1994	Stone Mill Elementary School (USA, Atlanta, Georgia)

RESEARCH INTERESTS

Multivariate analysis, GIS, Geostatistics, Ecological modelling, Environmental modelling, Aquatic Ecosystem Functioning, Habitat mapping

PUBLICATIONS

Gogina M., Zettler M.L., 2010. Diversity and distribution of benthic macrofauna in the Baltic Sea. Data inventory and its use for species distribution modelling and prediction. Journal of Sea Research. doi:10.1016/j.seares.2010.04.005

Glockzin M., Gogina M., Zettler M.L., 2009. Beyond salty reins – modelling benthic species' spatial response to their physical environment in the Pomeranian Bay (Southern Baltic Sea). Baltic Coastal Zone, 13-2, 79-95.

Gogina M., Glockzin M., Zettler M.L., 2010b. Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modelling and prediction. Journal of Marine Systems 80, 57-70. doi:10.1016/j.jmarsys.2009.10.001

Gogina M., Glockzin M., Zettler M.L., 2010a. Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 1. Causal analysis. Journal of Marine Systems 79: 112-123. doi:10.1016/j.jmarsys.2009.07.006

Meyer M., Harff J., Gogina M., Barthel A., 2008. Coastline changes of the Darss-Zingst Peninsula - a modelling approach . Journal of Marine Systems 74: 147-154. doi:10.1016/j.jmarsys.2008.03.023

CONFERENCES

Gogina MA, Glockzin M, Zettler ML (2009): Predictive modelling of benthic macrofauna distribution: Data inventory and habitat suitability modelling for Baltic Sea macroinvertebrates in response to selected environmental factors. BSSC 2009, 17-21 August, Tallinn, Estonia. Abstract book, p. 91.

Gogina M, Zettler ML, Harff J (2009): Relationships of benthic coenoses to near-bottom environmental parameters in the south-western Baltic Sea and their predictive modelling. ASLO Meeting 2009, 25-30 January, Nice, France, Poster #82.

Gogina MA, Efimova LE, Zaslavskaja MB (2004): Solute flow from the European part of Russia into Arctic Seas // Seventh Workshop on Land Ocean Interactions in the Russian Arctic, LOIRA project. Moscow, Shirshov Institute of Oceanology of RAS. Abstracts of the Reports, pp. 31-32.2.

Gogina M, Kositskii A, Moreydo V, Popryaduhin A, Ramazanova S, Chalov S (2003): Hydrological and hydrochemical analysis of groundwater and surface water sources of the Tver Oblast // International scientific conference of graduate and post graduate students and young scientists "Lomonosov – 2003" Section of geography. MSU, Moscow. Abstracts of the Reports, p.168 (in Russian)

Gogina M, Kirilov A, Gorin S, Popryaduhin A, Chaly A (2002): Full-scale hydrological and hydrochemical analysis for the Mozhaisk artificial lake in winter conditions // International scientific conference of graduate and post graduate students and young scientists "Lomonosov – 2002". Section of geography. MSU, Moscow. Abstracts of the Reports (in Russian)

Balashova A, Bashinskii I, Belousova A, Bogdanova D, Gogina M, Gongalskii K et al. (2001): The environment of Abkhazia and its protection in conditions of the post-war crisis // International scientific conference of graduate and post graduate students and young scientists "Lomonosov – 2001". Section of geography. MSU, Moscow. Abstracts of the Reports, p.136 (in Russian)

Part B

Scientific Papers Included

Declaration of the author’s contribution

The thesis comprises five manuscripts, which are presented in Part B in their logical order. The author’s contribution to each of the manuscript is detailed below.

Paper I

Gogina M., Glockzin M., Zettler M.L., 2010a. **Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 1. Causal analysis**. Journal of Marine Systems, 79: 112-123.

The study addresses the question of species-environmental relationships focusing on the macrozoobenthic communities in the south-western Baltic Sea. The investigation was done using the multivariate statistical analysis. This study represents a basis for the predictive modelling of species distribution in the selected study area. Part of the sampling and laboratory analysis of the data presented in the manuscript, as well as the acquisition of data from the external sources were done by the author herself. The concept of the manuscript, all the analyses and interpretation of the results were developed by the author who also did the writing and editorial handling of the publication. Michael Glockzin and Michael L. Zettler contributed in editing the writing, expertise and scientific advice.

Paper II

Meyer M., Harff J., Gogina M., Barthel A., 2008. **Coastline changes of the Darss-Zingst Peninsula - a modelling approach**. Journal of Marine Systems 74: 147-154.

The study focuses on the modelling of the coastline changes of the Darss–Zingst Peninsula. The main objectives were to test the viability of sediment transport modelling software package SEDSIM as a tool to model the local scale coastline development. The concept of the manuscript was developed by Michel Meyer and Jan Harff. The basis for modelling was constituted by the recent digital elevation model (DEM), a map of sediment distribution, measured wave time series and a modelled eustatic scenario. The DEM generation, including development of methodology, was carried out by Mayya Gogina who also did the writing of the corresponding part. All modelling setup and results interpretation were done by Michel Meyer. Thus the manuscript was written by Michel Meyer with the help of all co-authors.

Paper III

Gogina M., Glockzin M., Zettler M.L., 2010b. **Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modelling and prediction**. Journal of Marine Systems, 80: 57-70.

The manuscript addresses the question of how certain macrozoobenthic species respond to changes of considered environmental factors within the particular spatial scale (the study is focused on a limited area in the south-western Baltic Sea). It contributes to the development of methods for quantification of such species-environmental relationships, proposing a multimodel inference approach to predict the distribution of species within the common limits of combined environmental factors. The concept and methodology was developed by the author and the co-author Michael Glockzin. All the analyses and interpretation of the results were carried out by the author who also did the writing and editorial handling of the publication. Michael Glockzin and Michael L. Zettler contributed in editing the writing, expertise and scientific advice.

Paper IV

Glockzin M., Gogina M., Zettler M.L., 2009. **Beyond salty reins – modelling benthic species' spatial response to their physical environment in the Pomeranian Bay (Southern Baltic Sea)**. Baltic Coastal Zone, 13-2, 79-95.

The study addresses the quantification of species- environmental relationships for the Pomeranian Bay in the southern Baltic Sea. Through a logistic modeling approach it shows the possibility to predict benthic colonization areas and community structures thereby induced by the distribution of ecological factors inside the broad scheme of a brackish water habitat. The concept and methodology was developed by Michael Glockzin and Mayya Gogina. All the analyses and interpretation of the results were carried out by Michael Glockzin who also did the writing. Mayya Gogina and Michael L. Zettler contributed in editing the writing, expertise and scientific advice.

Paper V

Gogina M., Zettler M.L., 2010. **Diversity and distribution of benthic macrofauna in the Baltic Sea. Data inventory and its use for species distribution modelling and prediction**. Journal of Sea Research.

The manuscript addresses the question of how certain macrozoobenthic species from various functional groups respond to changes of considered environmental factors within the large spatial scale (the whole Baltic Sea). Analyses were based on the inventory data set compiled by revising the numerous data sources on macrozoobenthos. The concept of the manuscript was develop by the author and Michael L. Zettler. Data compilation, all the analyses and interpretation of the results were carried out by the author who also did the writing and editorial handling of the publication. Michael L. Zettler contributed in editing the writing, expertise and scientific advice.

I

Journal of Marine Systems 79 (2010) 112–123

Contents lists available at ScienceDirect

Journal of Marine Systems

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

Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 1. Causal analysis

Mayya Gogina, Michael Glockzin, Michael L. Zettler *

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

ARTICLE INFO

Article history:
Received 20 January 2009
Received in revised form 20 July 2009
Accepted 22 July 2009
Available online 30 July 2009

Keywords:
Zoobenthos
Community composition
Environmental factors
Multivariate statistics
Zonation
Baltic Sea
Mecklenburg Bight
Between 11.55° to 12.55° E and 54.09° to 54.96° N

ABSTRACT

In this study we relate patterns in the spatial distribution of macrofaunal communities to patterns in near-bottom environmental parameters, analysing the data observed in a limited area in the western Baltic Sea. The data used represents 208 stations, sampled during the years 2000 to 2007 simultaneously for benthic macrofauna, associated sediment and near-bottom environmental characteristics, in a depth range from 7.5 to 30 m. Only one degree of longitude wide, the study area is geographically bounded by the eastern part of the Mecklenburg Bight and the southwestern Darss Sill Area. Spatial distribution of benthic macrofauna is related to near-bottom environmental patterns by means of various statistical methods (e.g. rank correlation, hierarchical clustering, nMDS, BIO-ENV, CCA). Thus, key environmental descriptors were disclosed. Within the area of investigation, these were: water depth, regarded as a proxy for other environmental factors, and total organic content. Distinct benthic assemblages are defined and discriminated by particular species (*Hydrobia ulvae–Scoloplos armiger*, *Lagis koreni–Mysella bidentata* and *Capitella capitata–Halicryptus spinulosus*). Each assemblage is related to different spatial subarea and characterised by a certain variability of environmental factors. This study represents a basis for the predictive modeling of species distribution in the selected study area.

1. Introduction

The relative roles of the environment, competition, predation and chance in determining the composition of communities have been largely debated in animal ecology (Austin, 1987). Some main factors usually named as responsible for benthos distribution are: food supply, water salinity, oxygen concentrations, current energy, temperature, turbidity, substrate composition, sedimentation rates and bathymetry (e.g. Wildish, 1977; Wilson, 1991; Bromley, 1996; Olenin, 1997; Laine, 2003; Coleman et al., 2007). It is important that a heuristic approach is taken for both validation and iterative development of ecological models across a range of spatial scales. An exploratory statistical description of the prevailing ecological structure of the observations made on site is always the indispensible first step (Bourget and Fortin, 1995); it should also be supported by an improving mechanistic understanding (Thrush et al., 2003).

The following study can be referred to a comparatively fine spatial scale. Depending on the scale, different processes determine the spatial structure; therefore, inferences derived for one spatial scale cannot be extrapolated to another scale (Legendre and Legendre, 1998). Our results comprise a detailed analysis of benthic community composition regarding near-bottom habitat characteristics of the area stretching through the Kadetrinne incision from the eastern Mecklenburg Bight to the southwestern Darss Sill area. The aim of the present investigation was to analyse the causal relations between benthic macrofauna and selected environment variables: water depth, near-bottom salinity and oxygen concentrations, total organic content, median grain size, as well as sorting, skewness, and permeability. An appropriate technique for predictive modeling of species distribution regarding the available abiotic data was defined and applied in Gogina et al. (2009-this volume). Additionally, the modeling results for selected species were mapped. Hence, the present work describes a causal analysis — a basis and a necessary first step towards the design of models able to predict species distribution.

2. Materials and methods

2.1. Study area

Depending on the scale and region, salinity together with near-bottom oxygen concentration is often regarded among the major factors affecting the species richness and composition of macrozoobenthic communities (e.g. Olenin, 1997; Zettler et al., 2000; Laine, 2003). By defining the study area, we aimed to lessen the dominance of these two factors in the analysis to illuminate the impact of others. The study area is attached to the region of Mecklenburg Bight (Fig. 1) close to the very beginning of the Baltic salinity gradient and that is

---

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

0924-7963/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmarsys.2009.07.006

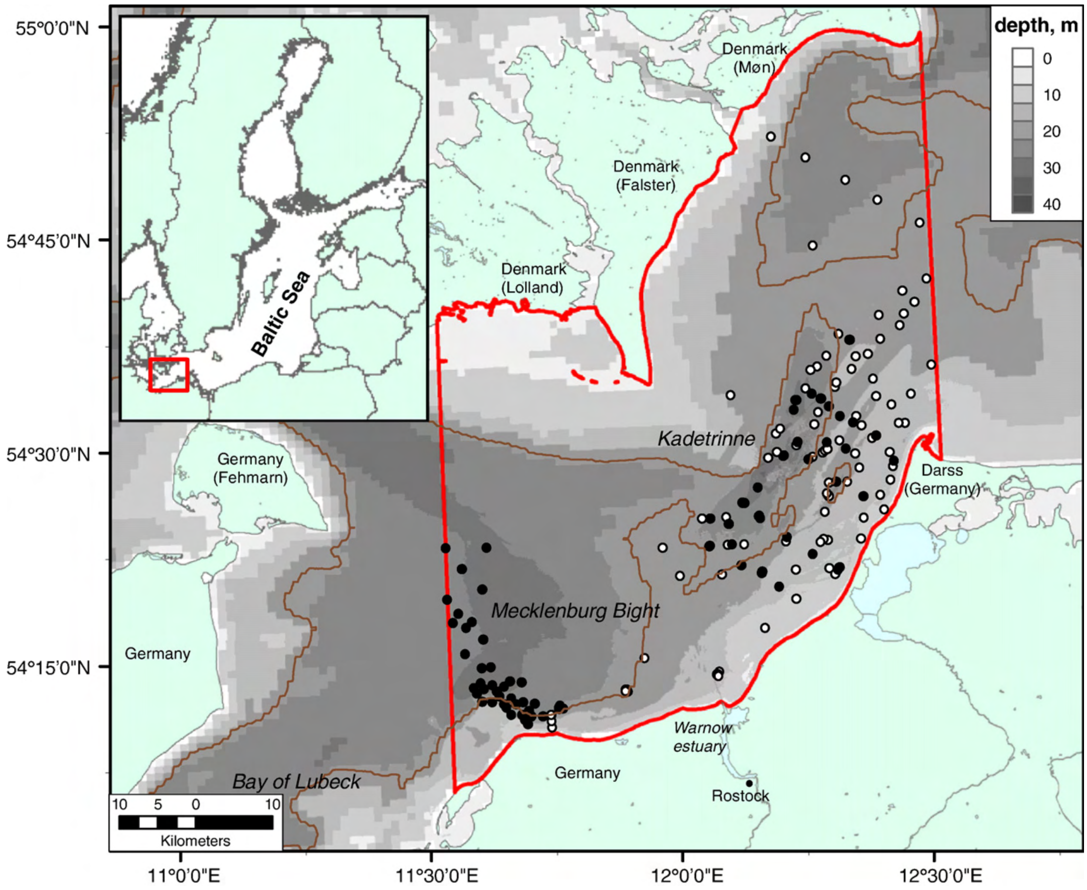


Fig. 1. Investigation area is bordered by the thick line, distribution of 208 sampling stations is presented; dots are stations covered with quantitative (abundance of benthic species) data; filled dots indicate 72 stations with a full set of data available for all eight abiotic variables. Thin line is the 18 m isobath which separates the two depth subareas of our region. Geographical data ESRI (1998); projection UTM on WGS84. The colour version of this figure is available online. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

why it is richer in macrofaunal species (with significant presence of both marine and brackish water species at the same time) than the inner, less saline, parts of the Baltic Sea.

The study area is assigned to the transition zone between the North Sea and the Baltic Sea. Barotropic and baroclinic gradients and local winds force currents that are controlled here by the bathymetry and stratification. Positive water balance caused by freshwater supply results in the outflow of brackish Baltic Sea water at the sea surface. Sea level differences forced by large scale winds drive the inflow of high saline Kattegat water into the transition zone. Density differences, mainly due to salinity, dominate the farther transport into the central parts of the Baltic. Local winds and thereby generated waves are important for vertical mixing in shallow waters. Particularly in summer, the thermal stratification stabilizes the water column and supports the salinity stratification, with halocline and thermocline occurring in the central part of the investigation area at around 12–16 m (Siegel et al., 2009). Generally, near-bottom salinity declines in the eastwards direction, dropping down in areas near the coastline due to the freshwater river runoff or the isolation from currents. Depending on the scale and region, salinity together with near-bottom oxygen concentration is often regarded among the major factors affecting the species richness and composition of macrozoobenthic communities. To lessen the dominance of those two factors and illuminate the effects of others we restricted the area to reduce the variability of both factors in distinct depth-zones, whereas other factors are represented in wide ranges. Whereas the range of near-bottom salinity in the whole area is relatively high (99.5% of sampling points lie within a salinity range of 8.3–21.8 PSU), the variability in distinct depth-zones is lower (i.e. salinity varied from 9 to 16 PSU at approximately 75% of stations shallower than 18 m; and ranged from 15 to 22 PSU at the same percentage of stations below 18 m). The hypoxic events, connected with absence of oxygen-rich saline water inflows and formation of steady hydrographical stratification in the water column, run differently in various locations of the region and irregularly take place in the late summer–autumn period, e.g. in deepest parts of Mecklenburg Bight. However, they are not as frequent and their effect is not as dramatic (at least within the investigation time) and long-lasting as it is known for the deep basins of the Baltic Proper.

The study site ranges in the northwards direction approximately from 45 to 75 km and about 63 km in the eastwards direction. Within these boundaries the area amounts to about 4000 km^2. With an average depth of 16 m (depth at sampling points ranges from 7.5 to 29.9 m) its water volume approximates 46 km^3 (data origin: this study). The sea floor is characterised mainly by postglacial sand and gravel deposits, with organic-rich silt-size sediments observed in the deepest parts of Mecklenburg Bight (Bobertz and Harff, 2004).

2.2. Macrobenthic sampling

Our data represents 208 stations, sampled over 8 years (2000–2007) simultaneously for benthic macrofauna, associated sediment and near-bottom environmental characteristics.

For 72 of the overall 208 stations a full set of abiotic variables considered is available. For macrofauna, 2 to 3 replicate samples per site were taken with a van Veen grab (0.1 m^2, 10–15 cm penetration depth) and sieved on a 1.0-mm screen. Samples were preserved in 4% buffered formaldehyde–seawater solution. In the laboratory, the formalin was washed out of the samples prior to sorting. The organisms were sorted, identified to the lowest possible taxon, and counted. Sampling and preparation were conducted in accordance with institutional, national and international guidelines concerning the use of animals in research (HELCOM, 2008).

2.3. Hydrographic measurements

Hydrographic characteristics were measured simultaneously with macrozoobenthic sampling for most data points. At each station a water sample was taken by shipboard CTD (0.5 m above the sea floor). Oxygen content was determined by immediate potentiometric titration, near-bottom salinity was estimated by CTD as well and the depth at each station was measured and logged via shipboard sonar system. All abiotic variables considered are listed in Table 1.

2.4. Estimation of sediment characteristics

With regard to the sediment characteristics measurements (completed fully for 72 out of 208 stations), an additional grab sample was taken to extract the upper surface sediment layer ($\leq$5 cm) for the analyses. Grain size distribution was analysed using approximately 50 g of dried sediment, dry sieved on a RETSCH sieving machine (sieve set: 63 µm, 75 µm, 90 µm, 106 µm, 125 µm, 150 µm, 180 µm, 212 µm, 250 µm, 400 µm, 630 µm, 2000 µm) and laser sizing (CILAS 1180 Laser Particle Analyser). Total organic content was measured as loss on ignition (3 h at 500 °C) of water-free material. Parameters describing cumulative grain size distribution as median grain size, sorting, and skewness were then calculated by using a skewed s-shape function, fitted to the cumulative grain size data with the least sum of squares method, applying a special fitting algorithm (the description is given in Bobertz and Harff, 2004). Permeability was estimated using the data of grain size distribution employing the relation derived by Krumbein and Monk (1942).

Table 1
List of environmental parameters considered in the study.

Parameter	Units	*N* stations	Method
Water depth	m	208	Determined and logged with shipboard sonar
Near-bottom oxygen	ml/l	181	Sampling by shipboard CTD system 0.5 m above the sea bottom, oxygen content determined by immediate potentiometric titration
Near-bottom salinity	PSU	185	Salinity estimated by shipboard CTD system 0.5 m above the sea bottom
Median grain size	µm	110	Sieving, laser sizing
Total organic content	%	109	The total organic content was measured as loss on ignition (3 h at 500 °C) of water-free material
Sorting	phi	72	Derived from grain size distribution curve as described in Bobertz and Harff (2004)
Skewness	phi	72	Derived from grain size distribution curve as described in Bobertz and Harff (2004)
Permeability	darcy	72	Determined according to Krumbein and Monk (1942) using the data from median grain size analysis (see above)

Sampling, measurement and determination techniques and the number of stations with each parameter observed are given.

2.5. Statistical methods and data treatment

To focus the investigation on biotic/environmental interactions rather than on other aspects (e.g. energetic criteria vs. species development) only species abundance as a biotic variable was subjected to statistical analyses (Young and Young, 1998).

Since spatial distribution patterns of species are influenced by biological processes structured by physiography, they can be spatially autocorrelated—the location of sampling points in space influences the values of random variables (Legendre, 1993). We calculated the amount of spatial autocorrelation for 72 stations (abiotic data) and 208 sampling sites (species abundance data) via Morans's I index using ArcView 9.1 software (ESRI). To examine correlations between species abundance and corresponding environmental data, Spearman's rank correlation coefficients were evaluated. To analyse the structure of coherencies among environmental variables normal and partial Pearson correlation factors were also computed (SPSS Inc.). This disclosed a primary environmental descriptor. Its effect on species distribution was analysed by means of testing cumulative frequency distributions of the primary predictor versus species abundances using the Kolmogorov–Smirnov test of significance (Perry and Smith, 1994; Simpson and Walsh, 2004; Glockzin and Zettler, 2008).

For mapping purposes, methods of gradient analysis were complimented by numerical classification that allows to cut the continuum even if there are no distinctive boundaries (Leps and Smilauer, 2003). Ordination contributes to the detection of systematic patterns and can disclose transitional zones. Therefore, multivariate analysis was carried out by superimposing the results of group averaged hierarchical clustering based on Bray–Curtis similarities of 4th–root transformed abundance data for 208 stations on a non-metric multidimensional scaling (nMDS) surface. In this way, distinct benthic communities could be defined. Species responsible for classification were determined applying SIMPER exploratory analysis and visual re-examination of the modified data matrix whereupon benthic communities have been determined and described (PRIMER; Clarke and Warwick, 2001). Correlations between biological and environmental variables were examined via BIO-ENV procedure of PRIMER software (Clarke and Ainsworth, 1993) and canonical correspondence analysis (CCA) via CANOCO (ter Braak, 1986), applying the necessary data transformations (Legendre and Gallagher, 2001; Davis, 2002).

3. Results

3.1. Data reduction

A total of 240 species were indentified at 217 sampling stations. For 9 of the sites no quantitative community assemblage data was available, reducing the number of analysed stations to 208. Polychaetes ranked first with 71 species, followed by the mollusca (56), crustacea (43), cnidaria (24), bryozoa (11), oligochaeta (9), nemertea (5), porifera (5), chordata (4), pycnogonida (3), echinodermata (2), kamptozoa (2), priapulida (2), arachnida (1), insecta (1), plathelminthes (1).

In order to employ various statistical methods, appropriate data sets were formed by applying the following steps: exclusion of all uncountable species, retention of strictly endobenthic species, removal of species with high patchiness that causes unreliable sampling (e.g. *Mytilus edulis*), and finally, extraction of all species with a frequency of less than 6% at all stations and species which account for less than 0.3% of total abundance over all stations (Legendre and Gallagher, 2001; Lozán and Kausch, 2004). Using those criteria together with a general knowledge of species' habitat preferences and habits, the most dominating species were extracted separately for each of the two depth subareas of our region. In the southwestern Baltic Sea at depths between 15 and 20 m, a transition zone with substantial shift in community composition from shallow- to deep-water species occurs. There is no certain value defined as a boundary, but we had to decide the fixed limit value to work with the data, so the 18 m isobath was assumed to separate two depth subareas of our region in a proper way (Fig. 1). This formed a list of 29 species of interest, cited in Table 2.

Matrices of biotic data with abundance values of 29 species at 208 stations and a full set of all available 8 environmental variables for 72 stations were created.

Benthic sampling was carried out throughout the year, 38% of the stations were sampled during spring (March, April, May) and summer (June, August) periods, each, whereas 11% and 13% of samples were taken during autumn (September, October) and winter (January, February) periods, respectively. We were interested in general all-seasonal distribution patterns; therefore no seasonal harmonisation of data was carried out.

3.2. Spatial autocorrelation via Moran's I Index

For all abiotic variables defined at the 72 sampling stations (see Fig. 1) a positive autocorrelation is assumed and the hypothesis of a random distribution of stations should be rejected. The Index values range between 0.06 and 0.4, with Z values by far exceeding the confidence interval (2.78–14.8). The index values calculated for most of the 29 species sampled at the 72 stations range from −0.04 to 0.04 with corresponding Z values between −1.14 and 2.15. Here, species abundance data appears to have a generally random pattern. Data point distribution of species *Cerastoderma glaucum*, *Gastrosaccus spinifer* and *Pygospio elegans* exhibit a slightly clustered pattern (index values 0.09 to 0.22) with Z scores slightly exceeding the confidence interval ($p = 0.05$); for some taxonomic units, the Z score noticeably exceeds the confidence interval (e.g. for *Abra alba*, *Arctica islandica*, *Bylgides sarsi*, *Diastylis rathkei*, *Dipolydora quadrilobata*, *Halicryptus spinulosus*, *Mya arenaria* and *Polydora ciliata*). Moran's I Index values calculated for the data set containing species abundance data sampled at 208 stations indicate nearly the same patterns for most species (−0.04 to 0.14) with Z values by far exceeding the confidence interval (−3.01 to 12.77). For this data set a positive autocorrelation is also assumed. However, the autocorrelation coefficients have possessed no values *near* ±1.0, that would indicate strong clustering or dispersion, for any variables of the datasets analysed. The obtained results rather indicate a slight tendency that has to be considered in further analyses and results interpretation.

Table 2
List of 29 species under interest.

N	Group	Species	6-letter code	Dominating in
1	Bivalvia	*Abra alba*	Abralb	Deeper waters (>18 m)
2	Polychaeta	*Ampharete baltica*	Ampbal	Both depth regions
3	Bivalvia	*Arctica islandica*	Arcisl	Deeper waters (>18 m)
4	Bivalvia	*Astarte borealis*	Astbor	Deeper waters (>18 m)
5	Polychaeta	*Bylgides sarsi*	Bylsar	Both depth regions
6	Polychaeta	*Capitella capitata*	Capcap	Both depth regions
7	Bivalvia	*Cerastoderma glaucum*	Cergla	Shallow waters (≤18 m)
8	Bivalvia	*Corbula gibba*	Corgib	Both depth regions
9	Cumacea	*Diastylis rathkei*	Diarat	Both depth regions
10	Polychaeta	*Dipolydora quadrilobata*	Dipqua	Both depth regions
11	Mysidacea	*Gastrosaccus spinifer*	Gasspi	Shallow waters (≤18 m)
12	Priapulida	*Halicryptus spinulosus*	Halspi	Deeper waters (>18 m)
13	Polychaeta	*Hediste diversicolor*	Heddiv	Shallow waters (≤18 m)
14	Oligochaeta	*Heterochaeta costata*	Hetcos	Shallow waters (≤18 m)
15	Polychaeta	*Heteromastus filiformis*	Hetfil	Deeper waters (>18 m)
16	Gastropoda	*Hydrobia ulvae*	Hydulv	Both depth regions
17	Polychaeta	*Lagis koreni*	Lagkor	Both depth regions
18	Bivalvia	*Macoma balthica*	Macbal	Both depth regions
19	Bivalvia	*Mya arenaria*	Myaare	Shallow waters (≤18 m)
20	Bivalvia	*Mysella bidentata*	Mysbid	Both depth regions
21	Bivalvia	*Parvicardium ovale*	Parova	Shallow waters (≤18 m)
22	Polychaeta	*Polydora ciliata*	Polcil	Deeper waters (>18 m)
23	Polychaeta	*Pygospio elegans*	Pygele	Both depth regions
24	Polychaeta	*Scoloplos armiger*	Scoarm	Both depth regions
25	Polychaeta	*Spio goniocephala*	Spigon	Shallow waters (≤18 m)
26	Polychaeta	*Terebellides stroemi*	Terstr	Deeper waters (>18 m)
27	Polychaeta	*Travisia forbesii*	Trafor	Shallow waters (≤18 m)
28	Polychaeta	*Trochochaeta multisetosa*	Tromul	Deeper waters (>18 m)
29	Oligochaeta	*Tubificoides benedii*	Tubben	Both depth regions

Group, species name, abbreviation used in further analysis, and depth regions where species dominate are indicated.

3.3. Spearman's rank correlation

To gain an insight on how the underlying processes of biotic/abiotic interactions cause species distributions and benthic zoning, Spearman's rank correlation factors were calculated for 29 species along with 8 environmental variables (Fig. 2). Because of the formerly described positive autocorrelation, no significance levels but descriptions of weak or strong correlations are given here.

The strongest Spearman's rank correlations between species abundance and environmental variables were found for water depth and salinity. A particularly strong and positive correlation with depth was indicated for *B. sarsi, Trochochaeta multisetosa* and *A. alba*. Beside depth and salinity, total organic content ranged on third position regarding the number of strong correlations calculated. The highest positive correlations between species abundance and sedimentological parameters median grain size, skewness and permeability were found for *Macoma balthica*.

3.4. Multiple causality–correlation among environmental variables

Pearson's correlation coefficients are calculated considering all the available abiotic data for 208 stations (Fig. 1). Prior to the analysis, the data matrix was z-transformed to make the data dimensionless, thus avoiding problems with the different measuring units used. Calculated values indicate a complex and relatively strong mutual association between environmental parameters prevailing in this region of the Baltic Sea (Table 3).

After computing Pearson's correlation coefficients for the z-transformed abiotic matrix, path analysis was applied in accordance with the model analysis described in (Legendre and Legendre, 1998). Therefore, all possible three-variable combinations derived from the matrix of partial correlation coefficients were considered and tested (not presented here for brevity). Path analysis revealed water depth as a primary descriptor for all other environmental factors in our study area.

3.5. Cumulative distribution curve analysis

To analyse the effect of water depth as the primary descriptor on all other environmental factors and its consequential influence on the spatial distribution of benthic species and communities, cumulative frequency distributions of depth and species abundances were tested versus each other (Fig. 3). Cumulative histograms represent 9 depth classes in a range from 7.5 to 30 m with a bounding step of 2.5 m. As a measure of dependency, the congruency between cumulative distribution curves of species and an environmental variable was used. The sharp increase of deviation, for instance for *A. islandica* between depth classes 4 and 5 (15 to 20 m), indicates the range of maximum response of the species against depth. The depth-dependence hypotheses can only be rejected for bivalve *M. balthica* with the significance level assumed. Some other species, bivalves *Mysella bidentata*, *Astarte borealis*, *Corbula gibba* and polychaeta *Ampharete baltica*, can be considered as only slightly depth-dependant species. Settling preferences are often reflected by the position of the species curve against the depth curve: for those species favoring shallower waters, the cumulative abundance curve lies above the depth curve.

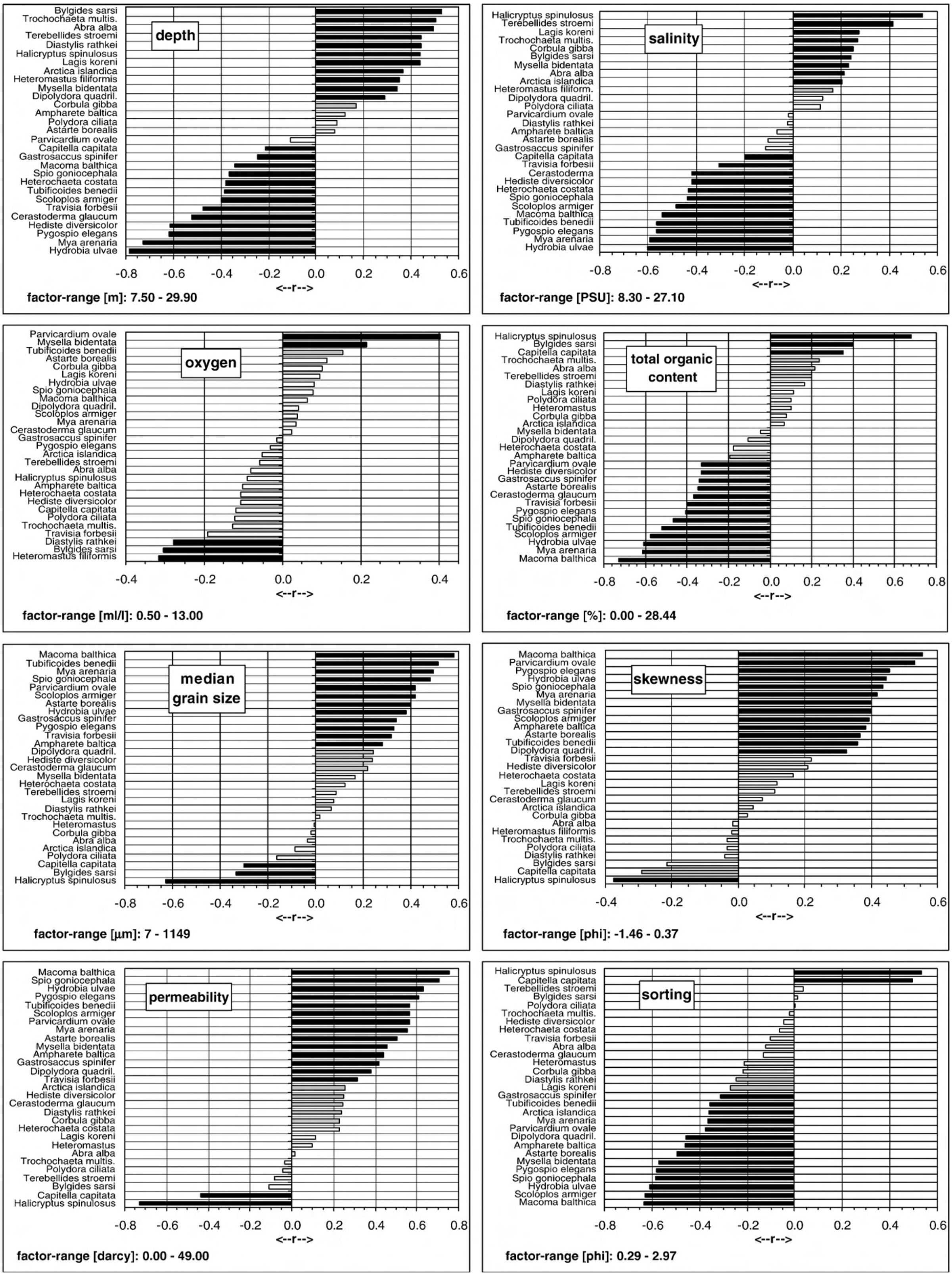


Fig. 2. Spearman's rank correlation factors calculated for 29 species along 8 environmental and sedimentological variables. Black and gray bars indicate significance levels of 0.01 and 0.05 (2-tailed), interpreted as strong and some correlation, correspondingly.

Table 3
Pearson correlation coefficients estimated for the full set of *z*-transformed environmental data sampled at 208 stations: sample size is indicated in brackets, coefficient values corresponding to $p<0.01$ are in bold font.

Environmental factors	Total organic content	Median grain size	Oxygen	Salinity	Depth	Sorting	Skewness	Permeability
Total organic content	1 (109)							
Median grain size	**−0.363** (109)	1 (110)						
Oxygen	−0.187 (109)	**0.262** (110)	1 (181)					
Salinity	**0.544** (109)	**−0.338** (110)	−0.082 (180)	1 (185)				
Depth	**0.670** (109)	**−0.354** (110)	−0.130 (181)	**0.597** (185)	1 (208)			
Sorting	**0.701** (72)	**−0.509** (72)	−.269 (72)	**0.500** (72)	**0.608** (72)	1 (72)		
Skewness	−0.227 (72)	**0.439** (72)	.257 (72)	−0.226 (72)	**−0.319** (72)	**−0.460** (72)	1 (72)	
Permeability	**−0.549** (72)	**0.923** (72)	**0.305** (72)	**−0.574** (72)	**−0.747** (72)	**−0.521** (72)	**0.487** (72)	1 (72)
Factor unit	%	µm	ml/l	PSU	m	phi	phi	darcy
Factor range	0–28.44	7–1149	0.5–13	8.3–27.1	7.5–29.9	0.29–2.97	−1.46–0.37	0–49

The results described above fully correlate with the results of Spearman's rank correlation analysis.

3.6. Macrobenthic communities and their discriminating species

Before computing the similarity matrices, the data for the selected 29 species at 208 stations was 4th-root transformed to reduce the impact of the species with the highest abundances on the assessment of the community similarities (Clarke and Warwick, 2001). Hierarchical clustering analysis based on Bray–Curtis similarity and unweighted group average linking was then carried out to test the similarity among the species.

At the 35% similarity level, a division into three groups of sampling stations and one single site emerged. Results of hierarchal clustering analysis were supplemented by the nMDS ordination (Fig. 4, left). The inverse analysis with the grouping of species corresponding to a 15% similarity level is presented in Fig. 4, right.

The SIMPER tool of the PRIMER software (Clarke and Corley, 2006) combined with visual re-examination of the modified display of the original data matrix (Clarke and Warwick, 2001) is able to identify the species responsible for the defined clustering pattern. Thus, three main benthic macrofaunal communities were distinguished and allocated in the investigation area (Fig. 5).

Group a: *Hydrobia ulvae–Scoloplos armiger* community inhabited the area of the shallow sandy bottom along the coast. The water depth of the stations ranged from 7.5 to 18.9 m. Median grain size varied from 108 to 527 µm, sorting and skewness ranged from 0.57 to 1.12 and from −0.31 to 0.33 phi respectively, organic content was low (0 to 1.2% with mean value ($\bar{x}$) of 0.4%). Hydrological conditions in the areas inhabited by this community were typical for shallow waters of the region: relatively low salinity (8.3 to 17.9 PSU), high oxygen content (3.0 to 12.9 ml/l, all values below 5 ml/l are dated by the August 2000). Other species numerically dominating in the community were the polychaete *P. elegans*, the bivalves *Macoma baltica* and *M. arenaria*, and the cumacean *D. rathkei*. The total number of species observed at the stations in the boundary of the community reached 111.

Group b: *Lagis koreni–M. bidentata* community occupied a deeper area, stretching from the south-west to the north-east (middle of the Kadetrinne and the Darss Sill), including the southern and northern parts of the Mecklenburg Bight. The shallowest stations were at 12 m, deepest—at 29.5 m. Sediments were characterised by median grain size of medium silt to fine sand, mainly moderately sorted, the average organic content was 2.5%, ranging from 0.2 to 17.4%. Environmental conditions were defined by higher salinity (9.9 to 27.1, $\bar{x}=16.3$ PSU) and variable oxygen conditions (0.5 to 13.0, $\bar{x}=6.5$ ml/l). In terms of both occurrence and abundance, the cumacean *D. rathkei* and the bivalve *A. alba* were also characteristic for this community. In total, 137 taxa were recorded.

Group c: *Capitella capitata–H. spinulosus* community. This community inhabited the deepest area in the Mecklenburg Bight (20.9 to 29.9, $\bar{x}=24.7$ m), characterised by high organic content (2.3–28.4, $\bar{x}=11.3$%), and poorly sorted fine to coarse silt. All measurements at stations assigned to the community were performed in February–March 2006, the observed salinity ranged from 17.4 to 21.1 PSU, oxygen varied from 4.6 to 7.0 ml/l. This community was characterised by a low number of taxa (23). *P. ciliata* was an additional characteristic species here. In terms of occurrence, the polychaete *B. sarsi* dominated in the community (high frequency of this taxon was typical for the whole investigation area). Other species under interest found in the community were the bivalves *C. gibba*, *A. islandica*, the cumacean *D. rathkei* and the polychaete *L. koreni*.

3.7. Distribution of benthic communities along environmental factors

To relate macrobenthic assemblages to environmental factors on the multivariate level, hierarchal clustering, Bray–Curtis similarity together with unweighted group average linking was applied to the 4th-root transformed abundance data of 29 species for 72 sites with the full set of abiotic variables observed and complemented by the nMDS. At the 43% similarity level, three classes representing different community structure and corresponding to different sea areas or depth-zones emerged. They matched (with few exceptions) the groups defined from the whole dataset of 208 stations (see Section 3.6). We have not adduced here the dendrogram and nMDS plots for this dataset, but abbreviated the groups distinguished based on the 72-stations data matrix as groups A, B and C, overlapping with groups (communities) a, b and c, respectively (see Fig. 5). Table 4 provides the average and standard deviation values of species abundance and environmental factors in three groups.

3.8. Linking macrobenthic community structure to environmental factors—BIO-ENV

The BIO-ENV procedure was employed on a species assemblage similarity matrix adjusted for 72 sites (as in Section 3.7) and the resemblance matrices generated using three various transformations of primary environmental 72-by-8 matrix (overall *z*-transformed; total organic content, median grain size, sorting and permeability are log-transformed prior to the overall *z*-transformation; standardised residuals of 7 environmental factors detrended from depth by

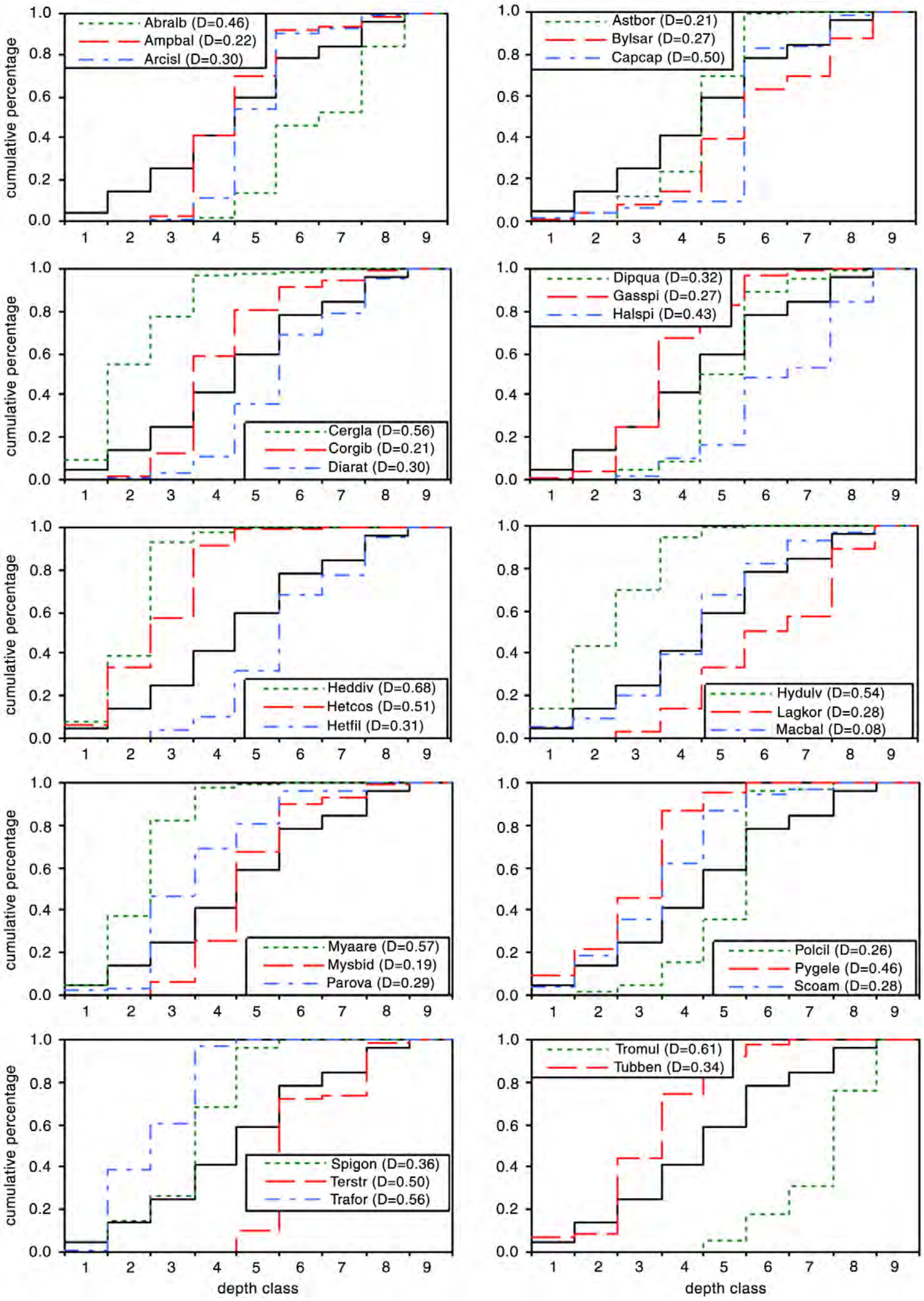


Fig. 3. Cumulative histograms of depth versus species abundances. Taxon names are abbreviated as in Table 2. Values of two-tailed two-sample Kolmogorov–Smirnov statistic (D) are given in brackets, for sample size $n = 208$ $D_{\mathrm{critical}} = 0.13$. The repeated black continuous line represents water depth. If the calculated D is below D_{critical}, the dependence can be rejected under the correspondent significance level $p = 0.05$ (Sachs, 1997), which is the case only for the bivalve *Macoma balthica* (result shown in bold). The color version of this figure is available online.

polynomial regression). The results are presented in Table 5. The Spearman correlation coefficient (r) was chosen as a rank correlation meaure. For the overall z-transformed environmental matrix, total organic content revealed the best association with the observed species distribution, ($r = 0.499$). It was followed by water depth ($r = 0.431$). Those two factors were responsible for most of the

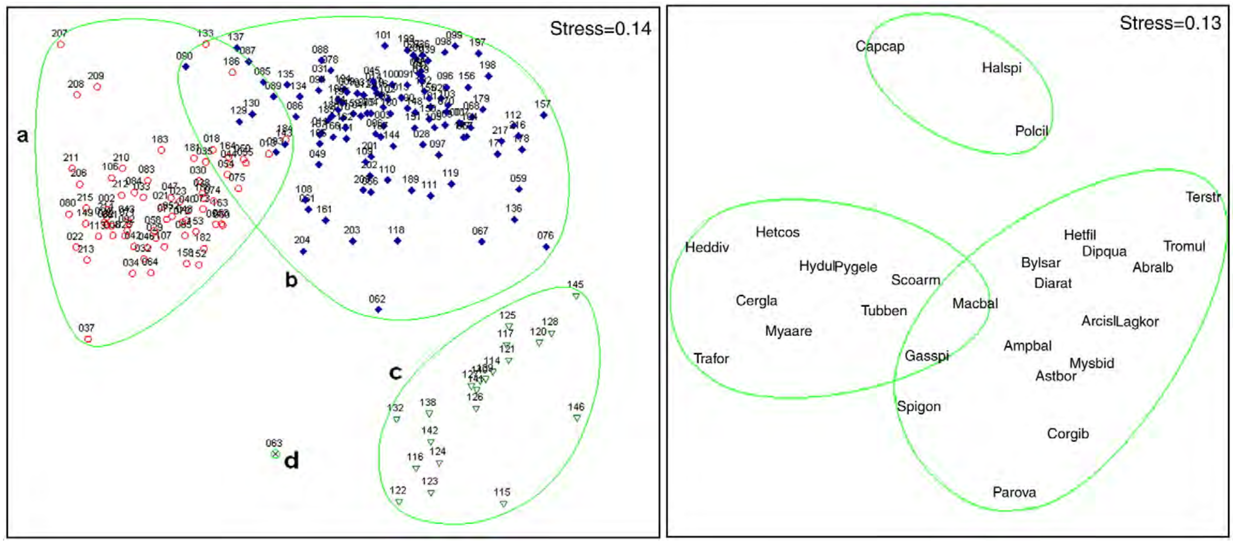


Fig. 4. Two-dimensional MDS-ordination of the 208 sampling sites regarding the Bray–Curtis dissimilarity based on 4th-root transformed abundance data for 29 selected species (left: normal analysis, grouping of sites correspond to 0.35 similarity level; right: independent inverse species analysis, grouping of species corresponds to 0.15 similarity level; dendrograms are not presented here for brevity). The color version of this figure is available online.

similarity between the biotic and abiotic matrices (total organic content alone accounted for over 80% of total similarity).

In the case of the prior log-transformation of selected abiotic variables, performed in order to reduce the heteroscedasticity of the descriptors, the highest overall correlation ($r = 0.605$) is found for a combination of 5 factors: median grain size, salinity, water depth, sorting and permeability.

For depth-detrended matrix of environmental descriptors, absolute correlations relating them to the community structure are very low for all possible combinations.

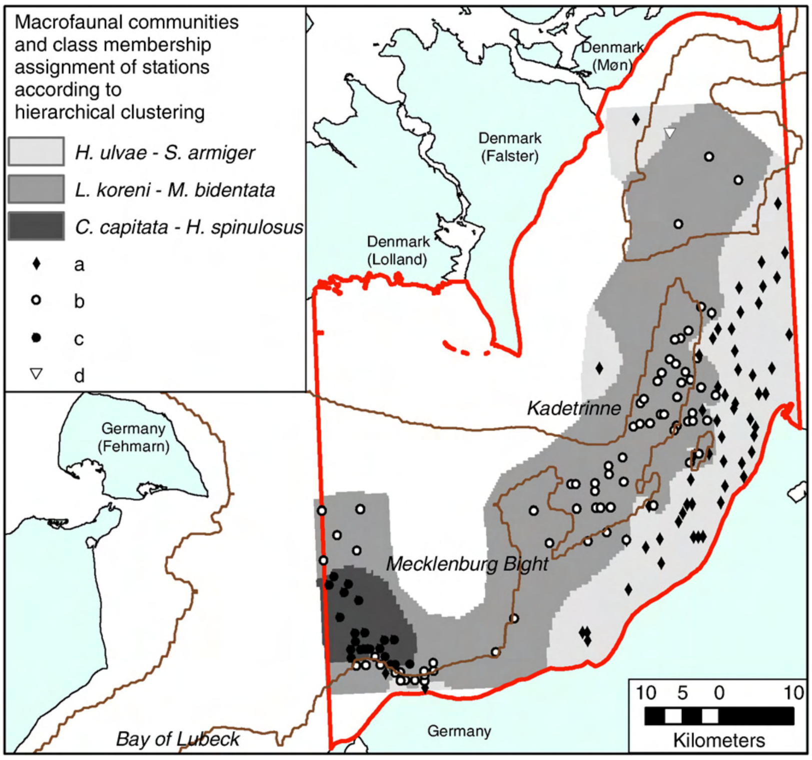


Fig. 5. Spatial distribution of macrobenthic coenoses. The zonation is based on the results of hierarchal clustering and nMDS analysis (class membership assignment of single stations is indicated as a, b, c and d according to Fig. 4, left). The colour version of this figure is available online.

Table 4
Average and standard deviation values of species abundance and environmental factors in three cluster groups based on the results of hierarchical clustering and non-metric MDS analysis for data set of 72 stations with all considered environmental data available.

Group	A	B	C
Number of sites	16	35	21
Polychaeta			
Ampharete baltica	15 ± 29	14 ± 27	
Bylgides sarsi	16 ± 15	54 ± 44	22 ± 16
Capitella capitata	6 ± 10	1 ± 3	535 ± 1403
Dipolydora quadrilobata	7 ± 12	12 ± 20	
Hediste diversicolor	3 ± 9		
Heteromastus filiformis	2 ± 9	4 ± 8	
Lagis koreni	256 ± 482	521 ± 1029	1 ± 3
Polydora ciliata	11 ± 21	2 ± 5	44 ± 125
Pygospio elegans	559 ± 815	7 ± 15	
Scoloplos armiger	162 ± 161	48 ± 49	1 ± 5
Spio goniocephala	20 ± 27	5 ± 12	
Terebellides stroemi		29 ± 130	
Travisia forbesii	21 ± 53		
Trochochaeta multisetosa	1 ± 3	29 ± 88	
Oligochaeta			
Heterochaeta costata	12 ± 40	0 ± 2	
Tubificoides benedii	72 ± 77	8 ± 20	
Cumacea			
Diastylis rathkei	114 ± 192	1009 ± 1258	14 ± 14
Mysidacea			
Gastrosaccus spinifer	18 ± 23	2 ± 4	1 ± 4
Gastropoda			
Hydrobia ulvae	2080 ± 3162	100 ± 397	
Bivalvia			
Abra alba	1 ± 4	446 ± 528	
Arctica islandica	14 ± 18	142 ± 144	13 ± 19
Astarte borealis	29 ± 69	21 ± 32	
Cerastoderma glaucum	7 ± 17		
Corbula gibba	51 ± 83	51 ± 135	14 ± 18
Macoma balthica	186 ± 183	62 ± 100	0 ± 2
Mya arenaria	91 ± 101	1 ± 6	
Mysella bidentata	197 ± 283	362 ± 551	2 ± 5
Parvicardium ovale	31 ± 48	16 ± 63	
Priapulida			
Halicryptus spinulosus	1 ± 3	2 ± 5	33 ± 37
Sum of average abundance of 29 selected species	3984 ± 3675	2947 ± 1816	681 ± 1536
Total abundance	7611 ± 9158	3219 ± 1837	736 ± 1605
Environmental factors			
Depth (m)	14.3 ± 2.1	21.8 ± 3.8	24.7 ± 2.4
Salinity (PSU)	13.6 ± 2.2	15.8 ± 3.1	20.4 ± 0.8
Total organic content (%)	0.6 ± 0.4	3.5 ± 3.8	11.3 ± 5.8
Dissolved oxygen (ml/l)	6.9 ± 0.7	6.2 ± 1.2	6.1 ± 0.7
Median grain size (µm)	295.4 ± 131.5	128.2 ± 77.5	41.2 ± 73.5
Sorting (phi)	0.7 ± 0.2	1 ± 0.6	1.9 ± 0.4
Skewness (phi)	0 ± 0.2	−0.3 ± 0.4	−0.5 ± 0.4
Permeability (darcy)	26.9 ± 16.6	6.8 ± 6.4	0.8 ± 3.2

3.9. The species' response to the environment—CCA

To support the results of BIO-ENV findings, the canonical correspondence analysis (CCA) was performed using CANOCO software (ter Braak and Šmilauer, 2002) on macrobenthic species abundance data (29 species) and corresponding environmental factors sampled at 72 stations. Two separate analyses were performed: the first one with 4th-root transformed abundance data and an untransformed matrix of 8 environmental factors and the second one with 4th-root transformed abundance data and a matrix of standardised depth-detrended residuals of abiotic factors. The Monte-Carlo permutation tests accompanied both analyses to test the effects of each environmental variable on the explanation of total variance in species data. Ordination diagrams of canonical correspondence analyses are shown in Fig. 6.

In the case of undetrended environmental data, the first two CCA axes together account for approximately 79% of the relations between species and environmental data (Fig. 6, left). With the matrix of standardised residuals of abiotic factors detrended from depth, this percentage is reduced to 70% (Fig. 6, right). Water depth reveals the highest positive correlation with the first canonical axis in case of undetrended environmental data followed by total organic content, whereas median grain size and permeability indicate the strongest negative correlation (Fig. 6, left). Sorting shows the highest correspondence to the second canonical axis, with organic content being on the second position. Near-bottom oxygen concentration shows the least correspondence to both first and second axes.

Table 5
BIO-ENV procedure results.

K	Factor/-s	r (ρ_{normal})	Factor/-s	r ($\rho_{\text{log-transf\_normal}}$)	Factor/-s	r ($\rho_{\text{depth-detrended}}$)
1	*1*	0.499	*8*	0.575	*1*	0.127
1	*5*	0.431	*2*	0.514	*8*	0.119
2	*1,5*	0.543	*5,8*	0.595	*1,8*	0.152
3	*4–6*	0.575	*4,5,8*	0.600	*1,6,8*	0.169
4	*1,4–6*	0.597	*2,4,5,8*	0.601	*1,2,6,8*	0.164
5	*1,2,4–6*	0.596	*2,4–6,8*	0.605	*1,2,4,6,8*	0.154
6	*1,2,4–7*	0.590	*2,4–8*	0.605	*1,2,4,6–8*	0.120
7	*1,2,4–8*	0.569	*1,2,4–8*	0.599	*1–4,6–8*	0.094
8	*1–8*	0.540	*1–8*	0.571		

K indicated the number of abiotic factors considered at a time. The highest Spearman rank correlation coefficients (ρ) evaluated between 4th-root transformed biotic similarity matrix (abundance data of 29 species, Bray–Curtis similarity) and three matrices of abiotic factor/-s (Euclidean distance) at 72 sampling sites are shown. Abiotic matrix 1 (ρ_{normal}): overall z-transformed; abiotic matrix 2 ($\rho_{\text{log-transf\_normal}}$): prior log-transformation of total organic content, median grain size, sorting and permeability and the following overall z-transformation; abiotic matrix 3 ($\rho_{\text{depth-detrended}}$): detrended from water depth component by means of polynomial regression of power 6 (calculated using the MATLAB as a tool), standardised regression residuals of 7 abiotic factors. The environmental factors associated to ρ are: *1* total organic content, *2* median grain size, *3* oxygen concentration, *4* salinity, *5* water depth, *6* sorting, *7* skewness, *8* permeability.

In the results of the CCA accounting for the depth-detrended residuals of environmental variables (Fig. 6, right), the highest positive correlation with the first canonical axis is assigned to total organic content, followed by sorting and salinity, while median grain size, skewness and permeability indicate negative correlation. Permeability and median grain size show strong negative correlation with the second CCA axis.

Comparison of both ordination diagrams highlights the importance of the water depth.

4. Discussion

In general, the highest overall Baltic Sea diversity and number of macrobenthic species can be found in the southwestern region that is most strongly influenced by marine conditions (Zettler et al., 2008). The number of species declines to the north resulting in the low-diversity communities, for example, in the Gulf of Bothnia (Bonsdorff, 2006). The rank of dominance of polychaeta, mollusca, and crustacea in the study area observed during the time period of this study is identical to patterns of community structure in the southwest Mecklenburg Bight between 1980 and 1994, described by Prena et al. (1997). They characterise a general temporal variation in the number of species observed by a slight increase from spring to summer whereas a decline in abundances towards autumn was most evidently seen at stations deeper than 20 m, with slight or no decline occurring in the shallower and outer zones. This is in agreement with findings of Zettler et al. (2000). This assemblage was also characterised in earlier papers (e.g. Schulz, 1969; Gosselck and Georgi, 1984). These works documented the retreat of populations of *A. islandica* and *Astarte* species from the deep areas of the Bight affected by irregular hypoxic events, followed by the alternate replacement of the returning communities and populations of several species that are able to survive the temporal oxygen deficiency or even colonize a so-called "bottom deserts". Oxygen-deficiency events, controlled by

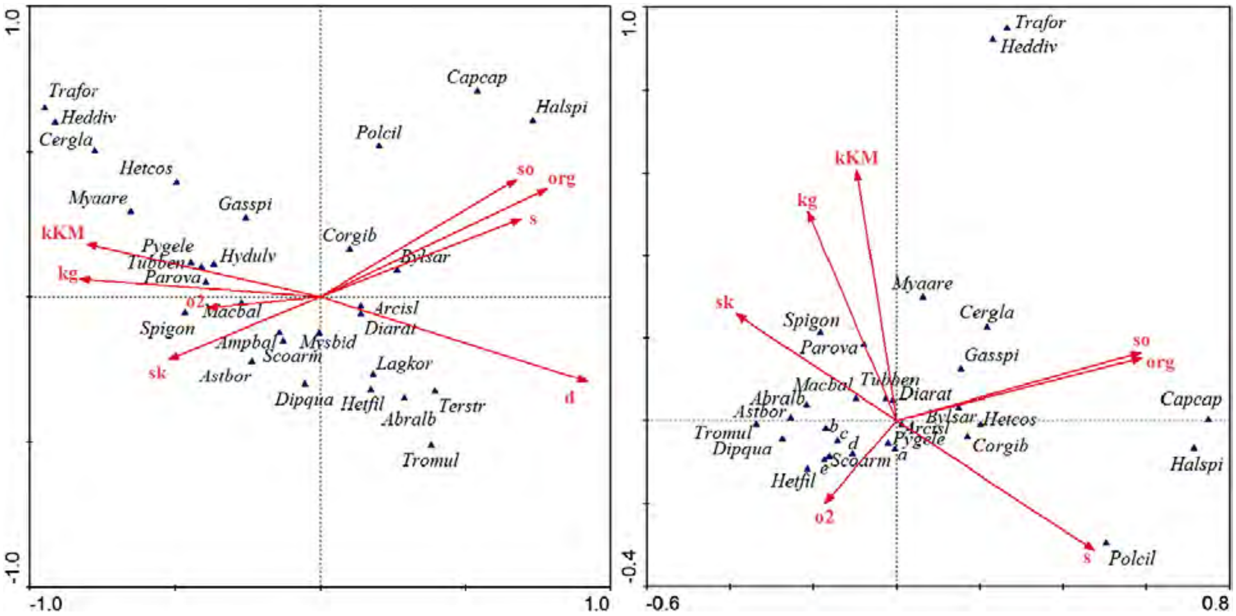


Fig. 6. Ordination diagrams of canonical correspondence analyses (CCA) of macrobenthic species abundance data (29 species) and corresponding environmental factors sampled at 72 stations (see Fig. 1). Left: CCA calculated for 4th-root transformed abundance data and untransformed abiotic matrix of 8 environmental factors. Right: CCA calculated for 4th-root transformed abundance data and matrix of standardised residuals of abiotic factors detrended from depth. Species notation: *a—H. ulvae*, *b—T. stroemi*, *c—A. baltica*, *d—M. bidentata*, *e—L. koreni*, others as in Table 2. Environmental factors notation: d—water depth, s—near-bottom salinity, org—total organic content, o2—ear-bottom oxygen concentration, kg—median grain size, so—sorting, sk—skewness, kKM—permeability. The colour version of this figure is available online.

bathymetry and hydrographical stratification, run differently in various locations of the region and usually take place in late summer–autumn, resulting in the retreat of species and low-diversity communities recorded during this season. However, those regions of poor bio-diversity or even deserted sea floor were always colonized within a relatively short time, leading to the general outcome: long-living species abandon the territory and are substituted by short life span polychaeta species (Zettler et al., 2000). This corresponds with the *C. capitata–H. spinulosus* community in our data, where *A. islandica* was absent or found in low abundances, and *A. borealis* was not observed at all. We deliberately avoided the analysis of temporal (seasonal or interannual) variability in our work. We had in the availability a relatively large dataset that covered both the distribution of macrobenthic species and simultaneously measured environmental variables. Analysing the data we aimed to identify ranges of abiotic parameters where certain species are in general likely to occur. Together with prior autecological knowledge this allows to recognize optimal, sub-optimal, or even pessimal conditions for them.

There is a potential of strong gradients to monopolize statistical analyses. Therefore, such gradients should be removed before unless it is the stated aim to explore the influence of these particular gradients (Legendre and Legendre, 1998; Bourget et al., 2003). We tried to reduce the effect of salinity gradient occurring in the Baltic Sea by restricting the study area and reducing the variability of salinity in distinct depth-zones. A smooth depth gradient could be allocated along the North–South axis of the area investigated. To examine the coherencies among environmental variables, Pearson's correlation coefficient was computed. However, correlation between parameters does not imply cause and effect. Path analysis may be used to assess the correspondence between the data and causal models, when causal ordering of the descriptors is stated by the prior information of ecological hypotheses (Legendre and Legendre, 1998). The complexity of abiotic interactions, reflected by the results of correlation analysis, does not allow the complete removal of either gradient without the loss of essential information. A solution was found in depth-detrending of abiotic data for certain statistical analyses (see also Glockzin and Zettler, 2008).

Three distinct benthic communities were defined, each related to different spatial subareas, characterised by a certain variability in environmental parameters (see Fig. 5). Previous studies explain the division of benthic species or communities over depth classes by food availability and/or certain food quality due to various sedimentation and resulting accumulation rates of organic material (Elmgren, 1978; Pearson and Rosenberg, 1978; Olenin, 1997; Laine, 2003).

The observed biotic structure was linked to environmental variability by examining correlations between the two multivariate patterns applying the BIO-ENV procedure (Clarke and Ainsworth, 1993; Clarke and Warwick, 2001; Lu, 2005). CCA was done to describe the full community variation with respect to a particular complex set of measured environmental variables. The attendant evaluating numerical analyses (Monte-Carlo permutation tests) was important to provide a fairly accurate picture (ter Braak, 1986; Legendre and Gallagher, 2001; ter Braak and Šmilauer, 2002). The comparison of results of both analyses accounting for detrended and undetrended environmental data points out the predictors of benthic macrofauna composition. The influence of depth on all other factors becomes clear. Like altitude through its physical forcing on other environmental factors plays an essential role in plant ecology (Guisan et al., 1999), depth can be regarded as a master factor in marine benthology. A tribute to the importance of this factor for benthic distribution is paid by the analysis of estimated cumulative curves for depth classes versus species abundances (Fig. 3). According to BIO-ENV analysis calculated with the undetrended dataset using no prior log-transformation, the total organic content causes the best similarity between abiotic and biotic data. The same implies to the results of the analysis with depth-detrended abiotic dataset. This is also confirmed by the Monte-Carlo simulation of the CCA performed on the same data. At the same time, in BIO-ENV results for depth-detrended environmental descriptors absolute correlations relating them to the community structure are very low for all possible combinations. This can be interpreted as an evidence of depth being the driving factor for community composition in the region. Considering the strong correlation between the total organic content and water depth, the first factor mentioned is likely to change drastically over relatively short distances. The same implies to both sorting and permeability. They differ in their rank in the results of BIO-ENV and CCA, but obviously remain dominating in explaining the benthic variance. The difference may be explained by the fact that BIO-ENV is based on similarity matrices, it might work best with sharp and short gradients (McGarigal et al., 2000) in contrast to CCA that most likely prefers distinct but uniformly continuous species response (Legendre and Legendre, 1998).

For *T. forbesii* and *H. diversicolor*, their habitat preferences seem to differ in correlation analysis and depth-detrended CCA. Such divergent results found for different statistical methods may account for the fact that our limited set of environmental variables did not fully reflect the complexity of natural species–environment relationships. This again does not only underline the importance of prior in-depth knowledge about species autecology for the interpretation of statistical results (Sachs, 1997), but furthermore illustrates the impossibility of deriving autecology from statistical results.

Our study defined dominating species and distinct benthic assemblages within the investigation area. It provides an insight on species that identify the strongest response to various factors and highlights environmental factors playing the major role for the distribution of benthic macrofauna. Hitherto, this is the most detailed study focused on benthic–abiotic interactions using a complex dataset of recent investigations performed for this region of the Baltic Sea. To that effect, it is an initial first step towards the design of models able to predict macrofaunal distribution regarding autecological species–environmental interactions (Glockzin and Zettler, 2008), that being the content of the second part of this work (Gogina et al., 2009-this volume).

It is obvious that only a fragment of the controlling factors network, responsible for macrozoobenthos distribution, could be covered with the available data. Additional factors, e.g. food web structure, complex predation interactions, currents, temperature, turbidity, chemical substrate composition etc., were not considered in this study. Additional variables incorporated would deliver a more insightful analysis (Guisan and Zimmermannm, 2000). On different scales various factors take the leading stand in influencing species distribution. Biotic factors are expected to be the dominating ones more on the local scale, responsible for fluctuations of abundance within the community. On a large scale, e.g. Baltic-wide, salinity is likely to be the primary or even the only descriptor in species–environmental relationships (e.g. Laine, 2003; Bonsdorff, 2006). Moreover, the analysis of such short-term data is able to represent only the momentary state of benthic communities, whereas the temporal development of the southwestern Baltic Sea has shown both high variability in species abundance and shifts in species composition (e.g. Andersin et al., 1978; Perus and Bonsdorff, 2004; Zettler et al., 2008).

However, for the area examined, we gained new insights into the distribution and habitat selection of dominating benthic species and analysed the impact of different environmental factors on the distribution of benthic communities. Discriminating species are found responsible for distinct benthic assemblages, namely *H. ulvae–S. armiger*; *L. koreni–M. bidentata* and *C. capitata–H. spinulosus.* It was highlighted that depth, salinity and certain sediment characteristics, above all total organic content, seem to determine the suitability of habitats for various taxa. Our findings suggest the interpretation of water depth as a complex factor, which integrates the effects of several other parameters, such as annual temperature variation, nutrient supply etc. This factor also controls sedimentological parameters such as organic content, sorting or permeability in the study area. Numerous authors already assumed a depth dependency of environmental factors and thereby caused macrobenthic invertebrates spatial distribution (e.g. Wildish, 1977; Kube et al., 1996; O'Brien et al., 2003; Perus and Bonsdorff, 2004; Kröncke et al., 2004; Bonsdorff, 2006; Zettler et al., 2006; Glockzin and Zettler, 2008). Our results show a significant response to change in water depth for 29 selected species and thus seem to approve the aforementioned hypothesis.

5. Conclusion and outlook

Predictions of the climate change models assume changes in the Baltic Sea ecosystem such as a drastic change in the food web structure, an increase of warm water species and growing benthic deserts on the sea floor (Philippart et al., 2007). Predictive modeling of species distribution can become an important tool in ecosystem/habitat management, supporting a sustainable development of the Baltic Sea ecosystem.

The analysis of causality performed in this study provides an essential basis and pre-work for the modeling of species distribution. In the second part of this work (Gogina et al., 2009-this volume), species response curves were estimated by means of logistic regression and used to model the spatial distribution of selected species utilising the Akaike's information criterion for multimodel inference (Ysebaert et al., 2002; Thrush et al., 2003; Burnham and Anderson, 2004; Guisan et al., 2006). On this basis, habitat suitability maps, representing the predicted probabilities of species occurrence, were created in GIS.

Acknowledgments

Part of this work was funded by the EU-MARBEF. We thank all colleagues, in particular Ines Glockzin, involved and responsible for field work and laboratory analysis, Dr. Bernd Bobertz and Dr. Michael Meyer for valuable assistance on mathematical, statistical and GIS matters. We express our thanks to Irina Steinberg and Anastasia Semenova for correcting the English. We also acknowledge the contribution of Prof. Jan Harff in initiating this work.

References

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 Marine Research 244, 137–144.

Austin, M.P., 1987. Models for the analysis of species' response to environmental gradients. Vegetatio 69, 35–45.

Bobertz, B., Harff, J., 2004. Sediment facies and hydrodynamic setting: a study in the south western Baltic Sea. Ocean Dynamics 54, 39–48.

Bonsdorff, E., 2006. Zoobenthic diversity-gradients in the Baltic Sea: continuous post-glacial succession in a stressed ecosystem. Journal of Experimental Marine Biology and Ecology 330, 383–391.

Bourget, E., Fortin, M.-J., 1995. A commentary on current approaches in the aquatic sciences. Hydrobiologia 300/301, 1–16.

Bourget, E., Ardisson, P.L., Lapointe, L., Daigle, G., 2003. Environmental factors as predictors of epibenthic assemblage biomass in the St Lawrence system. Estuarine, Coastal and Shelf Science 57, 641–652.

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

Burnham, K.P., Anderson, D.R., 2004. Multimodel Inference: understanding AIC and BIC in Model Selection. Sociological Methods & Research 33-2, 261–304.

Clarke, K.R., Ainsworth, M., 1993. A method of linking multivariate community structure to environmental variables. Marine Ecology Progress Series 92, 205–219.

Clarke, K.R., Corley, R.N., 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.

Clarke, K.R., Warwick, R.M., 2001. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation, second ed. PRIMER-E, Plymouth.

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. Marine and Freshwater Research 58, 293–305.

Davis, J.C., 2002. Statistics and Data Analysis in Geology, third ed. John Wiley and Sons, New York, p. 638.

Elmgren, R., 1978. Structure and dynamics of Baltic benthos communities, with particular reference to the relationship between macro- and meiofauna. Kieler Meeresforschungen 4, 1–22.

Glockzin, M., Zettler, M.L., 2008. Spatial macrozoobenthic distribution patterns in relation to major environmental factors—a case study from the Pomeranian Bay (southern Baltic Sea). Journal of Sea Research 59, 144–161.

Gogina, M.A., Glockzin, M., Zettler, M.L., 2009. Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modeling and prediction. J. Mar. Sys 79, 112–123 (this volume).

Gosselck, F., Georgi, F., 1984. Benthic recolonization of the Lübeck Bight (Westren Baltic) in 1980/81. Limnologica 15, 407–414.

Guisan, A., Zimmermannm, N.E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135 (2–3), 147–186.

Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modeling of plant species distribution. Plant Ecology 143 (1), 107–122.

Guisan, A., Lehman, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R., Hastie, T., 2006. Making better biogeographical predictions of species' distributions. Journal of Applied Ecology 43, 386–392.

ELCOM, 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.

Kröncke, I., Stoeck, T., Wieking, G., Palojarvi, A., 2004. Relationship between structural and functional aspects of microbial and macrofaunal communities in different areas of the North Sea. Marine Ecology Progress Series 282, 13–31.

Krumbein, W.C., Monk, G.D, 1942. Permeability as a function of the size parameters of unconsolidated sand. AIME Technical Publications (ZDB-ID 2589 114), p. 1492.

Kube, J., Powilleit, M., Warzocha, J., 1996. The importance of hydrodynamic processes and food availability for the structure of macrofauna assemblages in the Pomeranian Bay (Southern Baltic Sea). Archiv fuÉr Hydrobiologie 138, 213–228.

Laine, A.O., 2003. Distribution of soft-bottom macrofauna in the deep open Baltic Sea in relation to environmental variability. Estuarine, Coastal and Shelf Science 57, 87–97.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.

Legendre, P., Gallagher, E.D., 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129, 271–280.

Legendre, P., Legendre, L., 1998. Numerical Ecology. Second English edition. In: Developments in environmental modeling, vol. 20. Elsevier Science B.V, Amsterdam.

Leps, J., Smilauer, P., 2003. Multivariate Analysis of Ecological Data Using CANOCO. Cambridge University Press, Cambridge.

Lozán, J., Kausch, H., 2004. Angewandte Statistik für Naturwissenschaftler. Wissenschaftliche Auswertungen, Hamburg.

Lu, L., 2005. The relationship between soft-bottom macrobenthic communities and environmental variables in Singaporean waters. Marine Pollution Bulletin 51, 1034–1040.

McGarigal, S., Cushman, S., Stafford, S., 2000. Multivariate statistics for wildlife and ecology research. Springer, New York. 279 pp.

O'Brien, K., Hänninen, J., Kanerva, T., Metsärinne, L., Vuorinen, I., 2003. Macrobenthic zonation in relation to major environmental factors across the Archipelago Sea, northern Baltic Sea. Boreal Environment Research 8, 159–170.

Olenin, S., 1997. Benthic zonation of the Eastern Gotland Basin. Netherlands Journal of Aquatic Ecology 30 (4), 265–282.

Pearson, T.H., Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanography and Marine Biology. Annual Review 16, 229–311.

Perry, R.I., Smith, S.J., 1994. Identifying habitat associations of marine fishes using survey data: an application to the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences 51 (3), 589–602.

Perus, J., Bonsdorff, E., 2004. Long-term changes in macrozoobenthos in the Åland archipelago, northern Baltic Sea. Journal of Sea Research 52 (1), 45–56.

Philippart, C.J.M., Anadón, R., Danovaro, R., Dippner, J.W., Drinkwater, K.F., Hawkins, S.J., Oguz, T., O'Sullivan, G., Reid, P.C, 2007. Impacts of Climate Change on European Marine and Coastal Environment. ESF-Marine Board Position Paper 9.

Prena, J., Gosselck, F., Schroeren, V., Voss, J., 1997. Periodic and episodic benthos recruitment in southwest Mecklenburg Bay (western Baltic Sea). HelgolaÉnder Meeresuntersuchungen 51, 1–21.

Sachs, L., 1997. Angewandte Statistik. Anwendung statistischer Methoden. Springer, Berlin. 883 pp.

Schulz, S., 1969. Benthos und Sediment in der Mecklenburger Bucht. BeitraÉge zur Meereskunde 24 (25), 15–55.

Siegel, H., Gerth, M., Heene, T., Ohde, T., Rüß, D., Kraft, H., 2009. Hydrography, currents and distribution of suspended matter during a dumping experiment in the western Baltic Sea at a site near Warnemünde. Journal of Marine Systems 75, 397–408.

Simpson, M.R., Walsh, S.J., 2004. Changes in the spatial structure of Grand Bank yellowtail flounder: testing MacCall's basin hypothesis. Journal of Sea Research 51 (3–4), 199–210.

ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67, 1167–1179.

ter Braak, C.J.F., Šmilauer, P., 2002. CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca NY, USA.

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. Marine Ecology Progress Series 263, 101–112.

Wildish, D.J., 1977. Factors controlling marine and estuarine sublittoral macrofauna. HelgolaÉnder Meeresuntersuchungen 30, 445–454.

Wilson, W.H., 1991. Competition and predation in marine soft-sediment communities. Annual Review of Ecology and Systematics 21, 221–241.

Young, L.J., Young, J.H., 1998. Statistical Ecology, A Population Perspective. Kluwer Academic Publisher, Dordrecht, A population perspective.

Ysebaert, T., Meire, P., Herman, P.M.J., Verbeek, H., 2002. Macrobenthic species response surfaces along estuarine gradients: prediction by logistic regression. Marine Ecology Progress Series 225, 79–95.

Zettler, M.L., Bönsch, R., Gosselck, F., 2000. Das Makrozoobenthos der Mecklenburger Bucht - rezent und im historischem Vergleich. Meereswissenschaftliche Berichte 42, 1–144.

Zettler, M.L., Röhner, M., Frankowski, J., 2006. Long term changes of macrozoobenthos in the Arkona Basin (Baltic Sea). Boreal Environment Research 11, 247–260.

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

II

Journal of Marine Systems 74 (2008) S147–S154

Contents lists available at ScienceDirect

Journal of Marine Systems

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

Coastline changes of the Darss–Zingst Peninsula — A modelling approach

Michael Meyer [*], Jan Harff, Mayya Gogina, Anke Barthel

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

ARTICLE INFO

Article history:
Received 4 October 2007
Received in revised form 20 December 2007
Accepted 12 March 2008
Available online 15 October 2008

Keywords:
Coastline change
Baltic Sea
Darss–Zingst Peninsula
Sediment dynamic
SEDSIM

ABSTRACT

Coastline development is a function of eustasy and isostasy. On a regional scale, the precision of models based on these two components is sufficient for the investigation of coastline scenarios. However, on a local scale sediment dynamic is responsible for the evolution of coastal structures that alter the coastline in detail considerably. The Darss–Zingst Peninsula at the southern coast of the Baltic Sea serves as a example for a case study of such a structure. Longshore sediment transport driven by waves and responsible for the formation of this barrier during the Holocene is simulated for the next 840 years with the sediment transport modelling software package SEDSIM. The main objectives are to test if SEDSIM is a viable tool for the modelling of such structures and if the available data basis is suitable for this task. Basic data used for the modelling are a recent digital elevation model, measured wave time series, a map of sediment distribution, and a modelled eustatic scenario. In result, SEDSIM proved to be a stable and highly configurable program that can handle the requested simulations. Input data have to be improved and must include event data like storm surges.

1. Introduction

Coastline change is a function of endogenic and exogenic processes. First, movements of the earth's crust have an impact on morphology and therefore on coastline course. Besides long-term acting tectonics, triggered by deep-crustal processes, isostasy is a well investigated phenomenon (Watts, 2001). Especially in times of transitions between glacials and interglacials the change of balance between continental ice and marine water volume results in high magnitudes of glacioisostasy (Miettinen, 2004). Secondly, variations of the climate are also very important. Climate controls the eustatic regime by regulating the balance of marine and continent bound water masses, but also the marine water temperature. Warming climate leads to higher water temperatures, causing expansion and sea level rise (Cabanes et al., 2001). The investigation of these processes is subject to many research projects. Especially the anthropogenic influence to climate change is discussed intensely (Metz et al., 2007). Naturally, isostasy and eustasy are interacting. In order to quantify the impact of the effective sea level change to coastal evolution, records about palaeo sea level have to be combined with surface elevation data. This method was discussed in detail by Meyer and Harff (2005).

On a local scale, coastline change is also affected by eustasy and isostasy. But, on a closer look, coastal morphogenesis cannot be explained by these two components alone. Sediment dynamics like abrasion, transport, and accumulation, forced by wind and waves, play an important role for the evolution of coasts, too (Lehfeldt and Milbradt, 2000; Harff et al. 2009). Approaches used within regional scale modelling have to be downscaled and complemented by routines for calculating the physical behaviour of coastal sediment. The sediment transport modelling software package SEDSIM (Tetzlaff and Harbaugh, 1989; Martinez and Harbaugh, 1993) implements these components and can be applied for geological timescales. On the example of the Darss–Zingst peninsula, a structure, located at the southern Baltic Sea coast and shaped by wave-driven longshore sediment transport, the application will by verified based on a set of measured primary data. This evaluation is an essential precondition for palaeo-modelling on a millennial time scale based on secondary data derived from proxy variables only.

---

[*] Corresponding author. Present address: Leibniz Institute for Baltic Sea Research, University Rostock, Seestraße 15, D-18119 Rostock, Germany. Fax: +49 381 5197 502.
E-mail address: michael.meyer@io-warnemuende.de (M. Meyer).

0924-7963/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmarsys.2008.03.023

2. Area of investigation

The Baltic Sea area serves as a model region for investigations of coastal change processes, on regional and on local scales as well. All primary coastal types defined by Harff et al. (2001) are represented. In the northern part, the land rises with rates up to 9 mm/year whereas the south reports sinking values of about 1–2 mm/year (Rosentau et al., 2007). Today's overall eustatic influence is approximated by Harff et al. (in press) with about 1 mm/year sea level rise. Therefore it can be stated, that the north is strongly isostatic dominated and in the southern part eustasy and isostasy are competing for predominance.

Since 8000 years BP, especially in the south the Littorina transgression caused an enormous changes of the coastal landscape (Lemke, 1998; Björck, 1995). In the beginning, up to around 4000 years BP, the rapid sea level rise caused a simple drowning of the relief. Former landscape was kept intact, and besides the drowning only minor sediment relocations have occurred. This phase is well documented by relative sea level curves with a rapid sea level rise (Lampe et al., 2005). But after this initial phase, during the last 4000 years numerous new Holocene structures evolved along the southern Baltic Sea because of the slow down of the sea level rise and an increasing influence of sediment dynamics. Examples for such structures are the Hel peninsula or the Vistula Spit. Large parts of the German Baltic Sea coast were reshaped by sedimentary longshore transport too, such as Hiddensee Island, Darss–Zingst or Wustrow Peninsula (Fig. 1).

The Darss–Zingst Peninsula consists of Pleistocene cores and Holocene spits (Janke and Lampe, 1998). Most prominent features are the Fischland barrier in the West, Darsser Ort on the northernmost mainland, and the Zingst lowlands on the eastern side. Seaside, the coast is smoothed while towards the land side, a couple of lagoons form a highly diversified inner coast.

3. Methodology

3.1. Sedsim

SEDSIM is a three-dimensional stratigraphic forward modelling software. Originally developed by Tetzlaff and Harbaugh (1989), the version used here is maintained by The Predictive Geoscience Group, CSIRO Petroleum Australia. The general goal of the software package is to calculate changes in sediment budgets over time as a function of the depositional environment. Hydrodynamic conditions are calculated as an approximation of the Navier–Stokes equations (Chorin and Marsden, 2000). The program is written in Fortran and provides an interface in terms of formatted text files. Results can be presented by a separate, graphic driven module, called SedView.

SEDSIM is able to perform model calculations from geological (millions of years) down to decadal time scales.

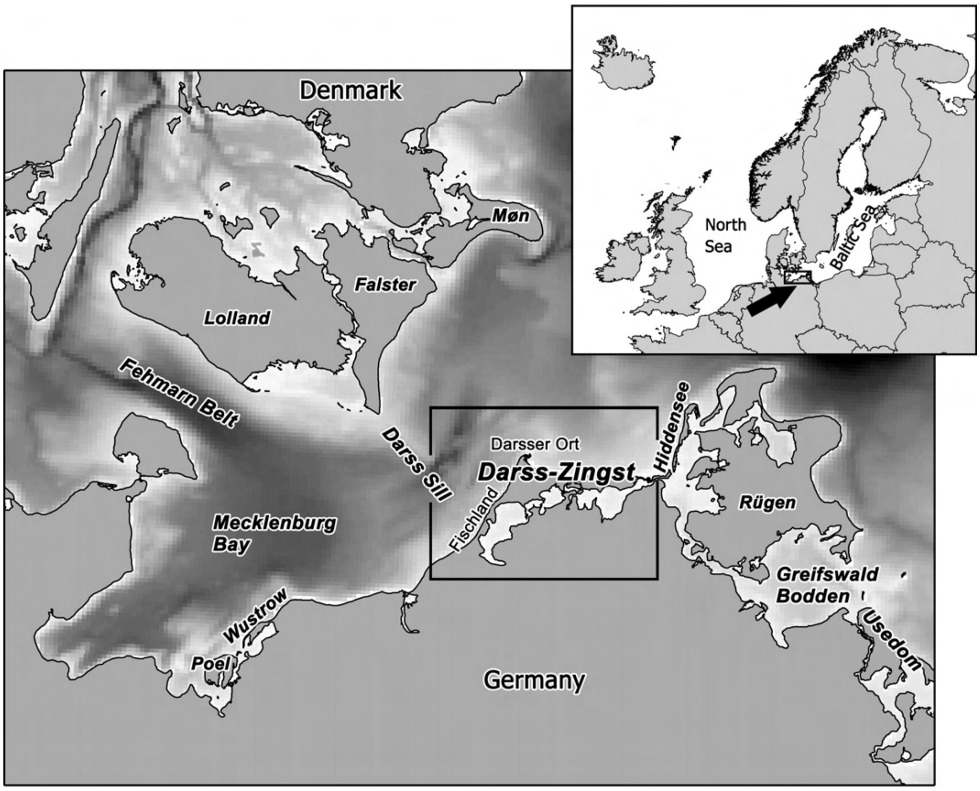


Fig. 1. Area of investigation. The Darss–Zingst Peninsula is located at the southern Baltic Sea Coast. The shape is a result of accelerated straightening after the initial phase of the Littorina transgression.

Table 1
Description of datasets used for the compilation of the DEM

Dataset name	Coordinate system reference	Spatial resolution
BSH	WGS 1984; in degrees, minutes, and seconds	Mostly below 50 m (point observations)
Seifert et al., 2001	Reference to any special earth ellipsoid is not defined; all data used were considered to refer to geographical longitudes and latitudes in decimal degrees	Approx. 1 km (grid)
DGM 25	Spatial reference is assigned in linear units (meters) referring to the Gauss–Krüger coordinate system	25 m (grid)

The software was used and verified by many investigations on coastal morphogenesis (e.g. Li et al., 2003; Tetzlaff and Schafmeister, 2007). The applied techniques and principles proved to by valid and able to model coastal dynamics. SEDSIM consists of required and optional modules. Here, special focus has the WAVE module. It is responsible for the calculations of sediment dynamics caused by waves and plays the most important role for long-term long-shore transport processes. Therefore, input parameter required are a digital elevation model, information about geology, relative sea level change, and data about heights and angles of waves.

3.2. Digital elevation model

In order to test the sensitivity of SEDSIM to the geological situation in the area of investigation, plausibility tests on the basis of the recent situation are required. For this, several input parameters are necessary. First, a digital elevation model (DEM) for the area of investigation was compiled.

The DEM represents both bathymetry and topography for the area. Its bathymetric part is based on measurements of water depths done by the Bundesamt für Seeschifffahrt und Hydrographie (2006, furthermore BSH). The resolution of these data is about +−30 cm in vertical and +−5 m in horizontal scale (Glockzin, 2006). The Nearest Neighbour Statistics (Davis, 2002) shows that the 75%-tile of the distances between the data points is below 50 m. Their spatial distribution is relatively homogeneous. The processing of the BSH data with standard software packages for spatial analysis was complicated because of the large amount of available data points (over 5 million). Therefore, first these data were separated into subsections 0.1° wide which overlap by 0.05°. After this, a Delaunay triangulation with Linear Interpolation was applied with Surfer (Golden Software, 2002) to each of these data portions. Then, the overlapping areas used to prevent edge errors by the interpolation where eliminated. Finally, the compilation of the resulting grid stripes was carried out in ArcGIS (Environmental Systems Research Institute, 2004).

The topographic part of the DEM was taken from the DGM 25 dataset (Amt für Geoinformation, Vermessungs- und Katasterwesen im Landesamt für innere Verwaltung Mecklenburg-Vorpommern, 2006). In addition, a regional dataset from Seifert et al. (2001) covering the Belt Sea region was employed to fill in the lack of data for a small marine section in the north-western corner of the area of investigation. Table 1 lists all used data sets with corresponding spatial resolutions.

The final DEM covers the area between 12.11° to 13.17° E and 54.07° to 54.76° N with a spatial resolution of 50 m, referring to the Gauss–Krüger coordinate system (Fig. 1).

3.3. Geological setting

The spatial distribution of surface sediments is displayed by Fig. 3. The map is a compilation from terrestrial (Heck et al., 1957) and marine (Tauber and Lemke, 1995; Tauber et al., 1999) mapping results. It was adjusted to the spatial resolution given by the digital elevation model. The reduction of the highly diversified classification given by the primary maps to three major sediment types is essential for the translation of the data grid into a format applicable by SEDSIM.

Within SEDSIM, main sediment physical properties are required as grain size and porosity. Table 2 shows the scheme applied to describe the three most common sediment types within the area of investigation: mud/very fine sand, sand, and glacial till. Most of the data are taken from Hoffmann et al. (2004) who described the Usedom peninsula, a geological analogue to the area of investigation.

Sediment transport was calculated for the whole area depicted by Figs. 2 and 3. For this selection the sediment transport boundary conditions were set to closed. Looking at the Darss–Zingst peninsula, valid boundary values can be assumed here because of the nesting.

As lower geological bound (basement) for the modelling with SEDSIM the surface of Pleistocene sediments (in this case glacial till resp. late Pleistocene basin sands) was used. In the area of recent Darss–Zingst Peninsula these sediments are partly situated on the surface (esp. Pleistocene core “Altdarss”) with heights up to 15 m NN but mostly they rest in the subsurface. Maximum depths of Pleistocene till are reached in the regions Darsser Ort (−12.4 m NN) and Commune Zingst (−17.3 m NN).

Additional to geological surface maps 71 borehole data were available for the estimation of geological basement. By interpolation between these sediment cores and map data it was possible to estimate the depth of the geological basement in the investigation area. Ordinary Kriging was used for spatial

Table 2
Generalized sediment classification for the area of investigation

	Coarse [%] Ø: 0.75 mm ρ: 2650 kg/m^3	Medium [%] Ø: 0.375 mm ρ: 2650 kg/m^3	Fine [%] Ø: 0.15 mm ρ: 2650 kg/m^3	Silt [%] Ø: 0.03 mm ρ: 2550 kg/m^3	Porosity [%] Ø: 0.75 mm ρ: 2650 kg/m^3
Mud/very fine sand	0	0	0	100	80
Sand	8	28	54	10	40
glacial till	12	22	35	31	20

Ø: diameter, ρ: density.

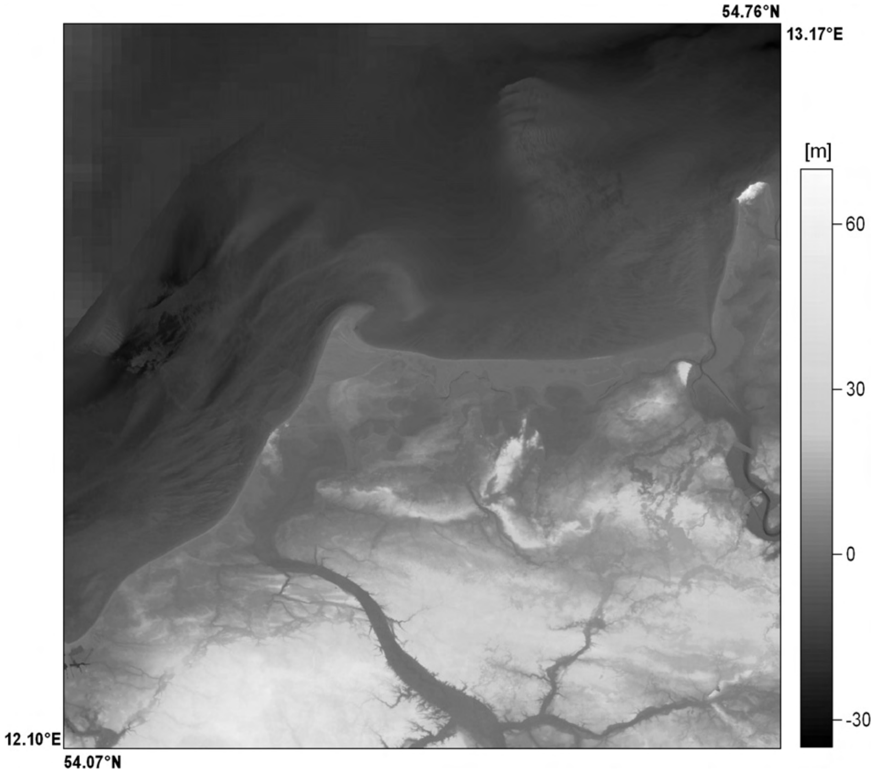


Fig. 2. Digital elevation model (elevation scale in m).

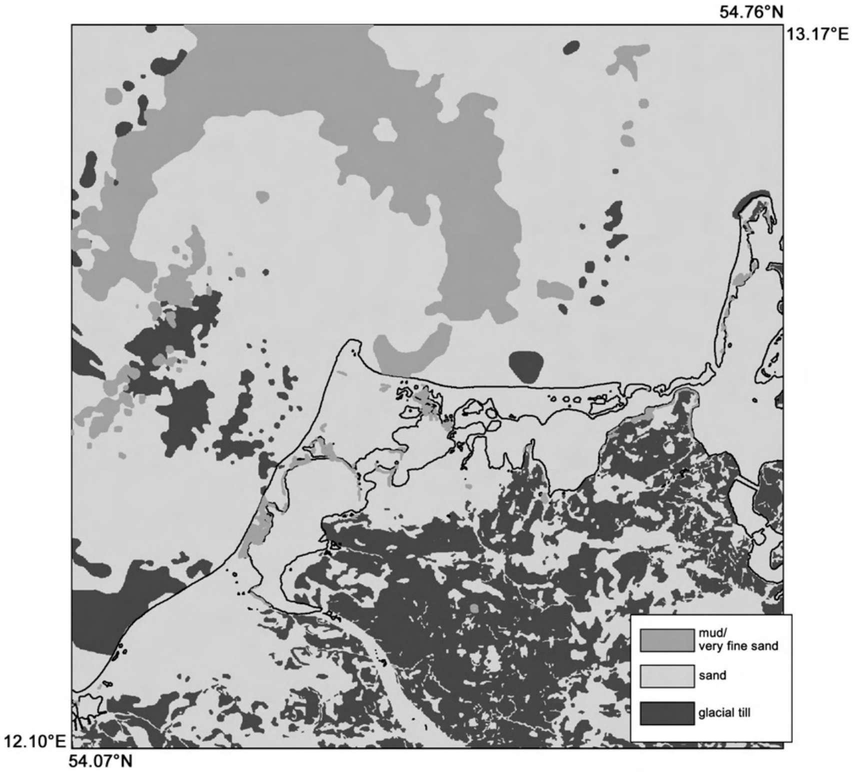


Fig. 3. Geological surface model.

interpolation. For the estimation a directional semivariogram (direction 30°, apex angle 30°) nested by a combination of Nugget, Spherical and Exponential models was applied to the data.

According to recent investigations about isostasy, vertical movements of the earth's crust within the area of investigation can be neglected. In Fig. 4 the map of the recent glacio-isostatic adjustment is given (from Harff et al., in press). Obviously, the area of investigation coincides with the transition zone between uplift in the north and subsidence in the south.

3.4. Wave regime

Sediment dynamics is a function of the depositional environment. Within the SEDSIM modelling, waves are considered the main driving force. For the area of investigation an experimental data set was provided by the GKSS-Research Centre Geesthacht (pers. comm. R. Weisse). It covers the time span between 1958 and 2002 and consists of wave heights and wave directions at a single location. This gauging station is located at 54.69° north, 12.69° east. Variations of waves according to deflection and reflection along the coast are calculated by the WAVE module of SEDSIM during the simulation. Original temporal resolution of the wave data is one hour, but this resulted in a very large amount of data. For SEDSIM parameterization the resolution was averaged into winter and summer season. The boundaries are defined by the beginning of October and April. To cover the time span between today and the next 840 years, the time span covered by the eustatic scenario (Voß et al., 1997), a linear continuation of the available 44 years function was applied.

The preferred directions of wave origin are aligned WSW and ENE (Fig. 5). This can be explained by the shape of the marine part of the area of investigation, that is more elongated in these than in it's perpendicular directions (see Fig. 1) providing a longer running way.

Wave heights differ depending on the season. During summer, generally low wave heights around 0.6 m are common, but during winter time, they may reach 0.9 m. It has to be noted, that these values are averaged heights. The spectrum of the original wave heights, shown in Fig. 5, has a median around 60–70 cm. Maximum height reaches 420 cm. Performing nonparametric statistical tests (One-Sample Kolmogorov–Smirnov, Chi-Square Test), as well as the analysis of simple descriptive statistics and histogram, suggest that a hypothesis of a log-normal distribution of the wave heights is not applicable.

3.5. Relative sea level change

It depends on the sea level height, what areas of sediment surface will be accessible for sediment dynamic processes. But, measured primary data of sea level change for the future are not available. For predictive scenarios it is necessary to use secondary data derived from climate modelling. An accepted and widely used data set is available from Voß et al. (1997). These data retrieved from by the ECHAM/LSG global atmosphere-ocean circulation model and considers the effect of global warming to the thermal expansion of ocean water and resulting sea-level rise. Voß et al. (1997) have simulated global sea-level rise for the next 800 years, based on IPCC-scenario A for the rise of atmospheric CO_2-concentration due to anthropogenic impacts (Houghton et al., 1990). Because of the rough spatial resolution of the data grid with 5.6°, information from the western North Atlantic and the North Sea were averaged for a regionalized sea level change function (Fig. 6). This function is available for case studies with a different development of the CO_2-budget. Here, the concentration is assumed to increase

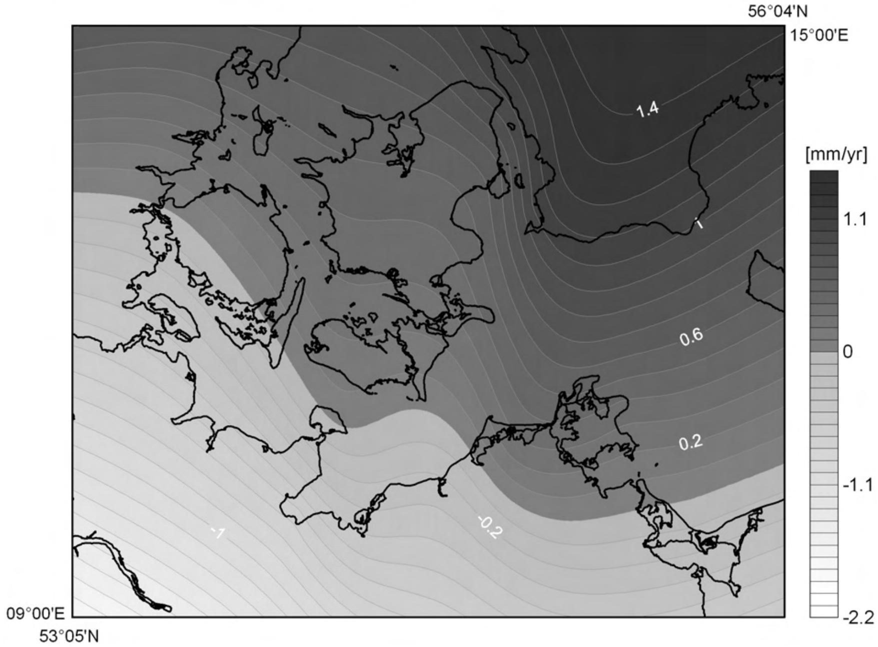


Fig. 4. Map of recent glacio-isostatic adjustment (from Harff et al., in press).

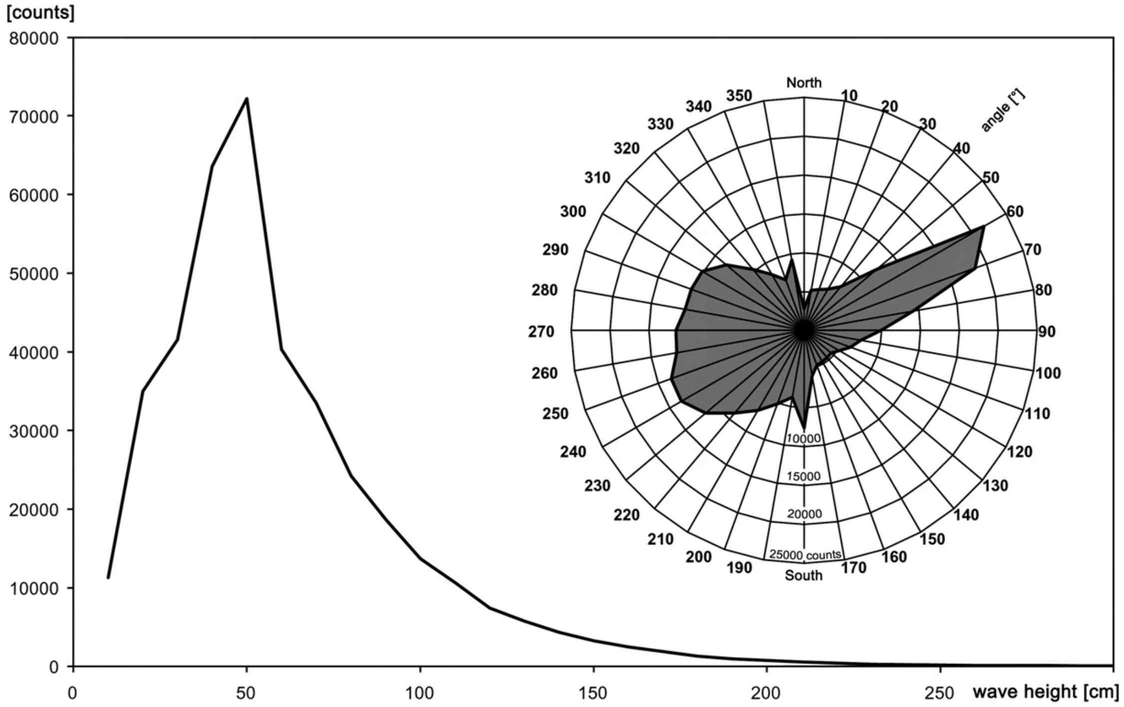


Fig. 5. Semi-diurnal distribution of wave heights and directions for a dataset 1958–2002. Wave direction is defined by origin.

linearly during the next 120 years to four times the present concentration and remains constant thereafter. Interestingly, the derived sea level is continuously rising even after the CO_2-level has reached its constant level. This phenomenon is due to the time lag of heat transfer from the atmosphere to the ocean by the global circulation system.

4. Results and discussion

According to the available data sets, the time span to be investigated was defined by the years 2000 (recent time) and 2800. The initial distribution of land and see is shown in Fig. 7a. Already 20 years after starting the SEDSIM modelling with parameters described in the chapter before, changes are visible (Fig. 7b). Especially along the part of the peninsula aligned toward the mainland, sediment seems to be eroded from the coast and deposited in the inner basin. This filling continues at least for the next 100 years (Fig. 7c and d). During this time, the offshore coast is affected very slightly only. At the top of the Darsser Ort a small amount of sediment seems to be transported to the north-eastern direction. Later on, in Fig. 7e and f the seaside coastline has changed dramatically, mostly

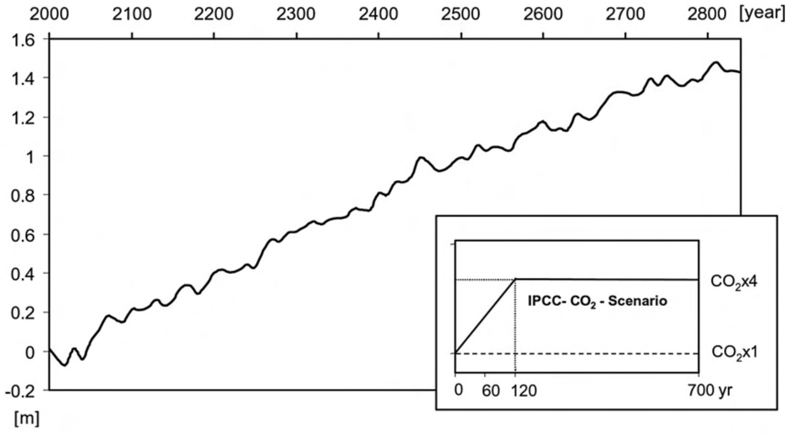


Fig. 6. Relative sea level change model.

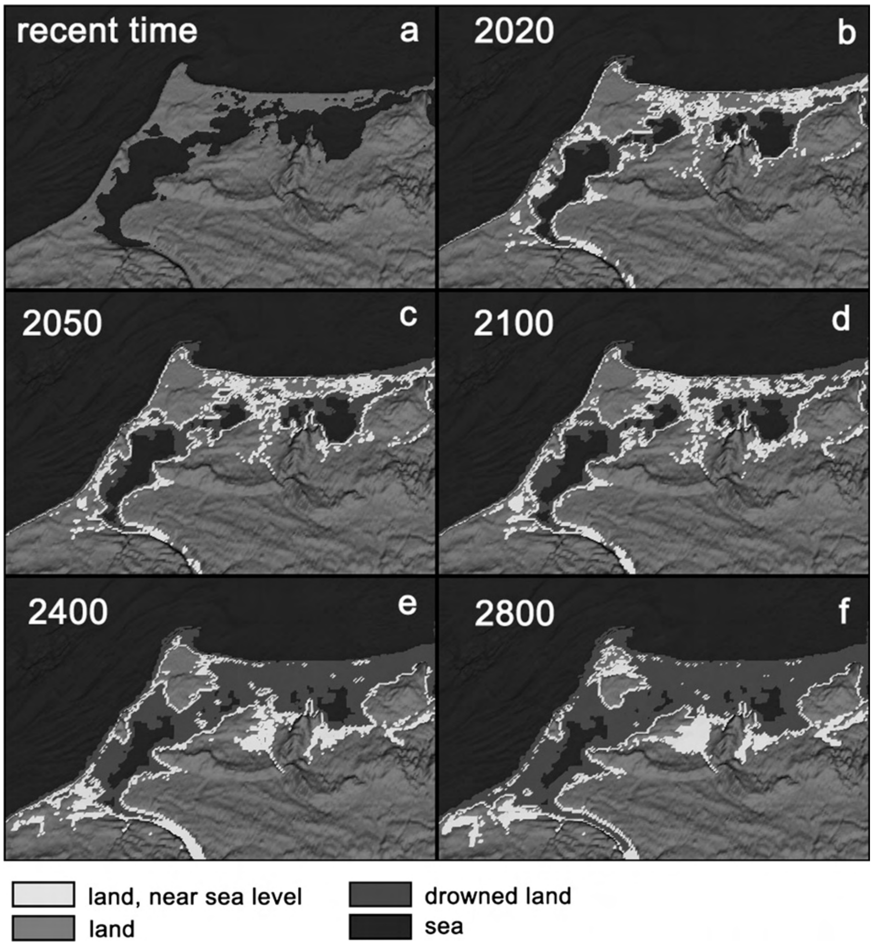


Fig. 7. Modelling results.

because of the sea level rise. Fischland is broken into a small island chain. To the north, the lowlands forming a protective barrier for the inner lagoons, are gone hence exposing the former inner coast to the marine environment.

Still, the morphology itself doesn't have changed remarkably. This indicates two results. First, the sediment itself has not been moved in an considerable amount by the model. And secondly, sea level rise is also for this high resolution modelling of major significance. Especially the rapid sea level rise proposed by the IPCC scenario superimposes the sediment transport component by far. Longshore sediment transport is not able to keep the coastal pattern intact but rising water leads to a breakup into small islands. In order to confirm the outcome of this experiment, events like storm surges should by included into the list of original data. They may alter the results significantly, but are not considered in the modelling here.

5. Summary

On a regional scale, coastline change is controlled mainly by eustasy and isostasy. But, the southern coast of the Baltic Sea today is also a result of longshore sediment transport. Therefore, especially for modelling of coastal morphogenesis on a local scale sediment dynamics has to be taken into account. The Darss–Zingst Peninsula serves as a case study for a modelling of these processes. The main target, the modelling of the evolution of the coastline on a millennial time scale, requires a validation of the sediment transport model that is to be applied. In a first approach, SEDSIM was chosen in order to model near future coastline scenarios for the Darss–Zingst area. The calculations are based on a recent digital elevation model, measured wave time series, a detailed map of sediment distribution, and a eustatic sea level curve proposed by IPCC. The time span covered ends in 2840 AD.

Results show a major impact of eustasy on the coastline change for the next centuries, but only a subordinated influence of sediment dynamic. However, the input parameters used within these case scenarios does not include event data. The effect of events like storm surges may vary the results significantly. But, the implementation of such processes is difficult because of their unpredictability. This has to be accomplished as a next step in order to improve the reliability of model results.

Acknowledgements

The authors would like to thank Cedric Griffith and Chris Dyt from CSIRO for support and assistance with SEDSIM. Ralf Weisse from the GKSS Geesthacht deserves thanks for the provision of wave data. We also like to thank the referees for helpful comments. The study was funded by the German Research Foundation for which we are grateful.

References

Amt für Geoinformation, Vermessungs- und Katasterwesen im Landesamt für innere Verwaltung Mecklenburg-Vorpommern, 2006. Digitales Geländemodell 25 — DGM25. Digital Elevation Model, Grid Size 25 m. Landesamt für innere Verwaltung, Schwerin.

Björck, S., 1995. Late Weichselian to early Holocene development of the Baltic Sea — with implications for coastal settlements in the southern Baltic region. In: Fischer, A. (Ed.), Man and Sea in the Mesolithic. Coastal Settlements Above and Below Present Sea Level, vol. 53. Oxbow Monographs, Oxford, pp. 23–34.

Bundesamt für Seeschifffahrt und Hydrographie, 2006. Topograpische Karten des Seegrundes für Nord- und Ostsee. Digital Data of Water Depth Measurements. Bundesamt für Seeschifffahrt und Hydrographie, Rostock.

Cabanes, C., Cazenave, A., Le Provost, C., 2001. Sea level rise during past 40 years determined from satellite and in situ observations. Science 294, 840–842.

Chorin, A.J., Marsden, J.E., 2000. A mathematical introduction to fluid mechanics. Texts in Applied Mathematics 4, 169 pp.

Davis, J.C., 2002. Statistics and Data Analysis in Geology. Wiley and Sons, New York. 638 pp.

Environmental Systems Research Institute, 2004. ArcGIS 9. Environmental Systems Research Institute (ESRI) Inc, Redlands, CA.

Glockzin, M., 2006: Die Abhängigkeit der benthischen Besiedlung ausgewählter Bereiche der Pommernbucht von ausgewählten Umweltfaktoren. Diplomarbeit, Universität Rostock, 143 pp.

Golden Software, 2002. Surfer 8 User's Guide. Golden Software Inc, Golden, CO.

Harff, J., Frischbutter, A., Lampe, R., Meyer, M., 2001. Sea level change in the Baltic Sea: interrelation of climatic and geological processes. In: Gerhard, J., Harrison, W.E., Hanson, B.M. (Eds.), Geological Perspectives of Climate Change, vol. 47. AAPG Studies in Geology, Oklahoma (Tulsa), pp. 231–250.

Harff, J., Graf, G., Bobertz, B., 2009. Dynamics of Natural and Anthropogenic Sedimentation (DYNAS). Journal of Marine Systems. doi:10.1016/j.jmarsys.2008.04.008.

Harff, J., Meyer, M., Lampe, R., in press. Changing Holocene coastal zones of the Baltic Sea – a modelling approach. In: Harff, J., Lüth, F. (Eds.) (in press), Sinking Coasts – Geosphere Ecosphere and Anthroposphere of the Holocene Southern Baltic Sea. Berichte der Römische Germanischen Kommission (RGK), Publikationsjahr (2009).

Heck, H.-L., Breitbach, J., Büttner, K., Groba, E., König, G., Stahff, U., Tattenberg, P., Vollbrecht, K., 1957. Geologische Karte des Norddeutschen Flachlandes, 1:1000000. - Geologische Karte, Einheitsblatt, 10: 2 sheets; Berlin.

Hoffmann, G., Musolff, A., Meyer, T., Schafmeister, M.-Th., 2004. Der geologische Aufbau des oberflächennahen Grundwasserstockwerkes im Nordosten der Halbinsel Gnitz (Usedom/Mecklenburg-Vorpommern). Rostocker Meeresbiologische Beiträge 12, 9–21.

Houghton, J.T., Jenkins, G.J., Ephraums, J.J. (Eds.), 1990. Climate change: the IPCC scientific assessment. Cambridge University Press, Cambridge.

Janke, W., Lampe, R., 1998. Die Entwicklung der Nehrung Fischland - Darß - Zingst und ihres Umlandes seit der Litorina-Transgression und die Rekonstruktion ihrer subrezenten Dynamik mittels historischer Karten. Zeitschrift für Geomorphologie, N.F., Suppl. 112, 177–194.

Lampe, R., Endtmann, E., Janke, W., Meyer, H., Lübke, H., Harff, J., Lemke, W., 2005. A new relative sea-level curve for the Wismar Bay, N-German Baltic coast (Eine neue relative Meeresspiegelkurve für die Wismarbucht, norddeutsche Ostseeküste). Meyniana 57, 5–35.

Lehfeldt, R., Milbradt, P., 2000. Longshore sediment transport modeling in 1 and 2 dimensions. — Advances in hydro-science and -engineering. Proceedings of the 4th International Conference on Hydro-Science and -Engineering, Seoul. Abstract, vol. 262. Korea Water Resources Association, Seoul.

Lemke, W., 1998. Sedimentation und paläogeographische Entwicklung im westlichen Ostseeraum (Mecklenburger Bucht bis Arkonabecken) vom Ende der Weichselvereisung bis zur Litorinatransgression. Meereswissenschaftliche Berichte 31, 156 pp.

Li, F., Dyt, C., Griffiths, C., 2003. A coastal morphodynamic model for cross-shore sediment transport. Coastal Engineering 2003, 335–344.

Martinez, P.A., Harbaugh, J.W., 1993. Simulating Nearshore Environments. 265 pp.; New York.

Metz, B., Davidson, O., Bosch, P., Dave, R., Meyer, L. (Eds.), 2007. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge. 851 pp.

Meyer, M., Harff, J., 2005. Modelling palaeo coastline changes of the Baltic Sea. Journal of Coastal Research 21 (3), 598–609.

Miettinen, A., 2004. Holocene sea-level changes and glacio-isostasy in the Gulf of Finland, Baltic Sea. Quaternary International 120, 91–104.

Rosentau, A., Meyer, M., Harff, J., Dietrich, R., Richter, A., 2007. Relative sea level change in the Baltic Sea since the Littorina Transgression. Zeitschrift für Geologische Wissenschaften 35 (1/2), 3–16.

Seifert, T., Tauber, F., Kayser, B., 2001. A High Resolution Spherical Grid Topography of the Baltic Sea — 2nd Edition. Baltic Sea Science Congress, Stockholm. Poster #147.

Tauber, F., Lemke, W., 1995. Map of sediment distribution in the Western Baltic Sea. Deutsche Hydrographische Zeitschrift 47 (3), 171–178.

Tauber, F., Lemke, W., Endler, R., 1999. Map of sediment distribution in the Western Baltic Sea (1: 100 000), sheet Falster - Møn. Deutsche Hydrographische Zeitschrift 51 (1), 5–32.

Tetzlaff, D.M., Harbaugh, J.W., 1989. Simulating Clastic Sedimentation, Computer Methods in the Geosciences. Van Nostrand Reinhold, New York. 196 pp.

Tetzlaff, D.M., Schafmeister, M.-T., 2007. Interaction among sedimentation, compaction, and groundwater flow in coastal settings. In: Harff, J., Hay, W.W., Tetzlaff, D.M. (Eds.), Coastline Changes: Interrelation of Climate and Geological Processes. Special Paper, vol. 426. The Geological Society of America, pp. 65–87.

Watts, A.B., 2001. Isostasy and Flexure of the Lithosphere. 458 pp.; Cambridge.

Voß, R., Mikolajewicz, U., Cubasch, U., 1997. Langfristige Klimaänderungen durch den Anstieg der CO2-Konzentration in einem gekoppelten Atmosphäre-Ozean-Modell. Annalen der Meteorologie 34, 3–4.

III

Journal of Marine Systems xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Journal of Marine Systems

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

Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modelling and prediction

Mayya Gogina, Michael Glockzin, Michael L. Zettler *

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

ARTICLE INFO

Article history:
Received 3 March 2009
Received in revised form 29 September 2009
Accepted 10 October 2009
Available online xxxx

Keywords:
Benthic macrofauna
Environmental factors
Ecological niche modelling
Logistic regression
Statistical modelling
Multimodel inference
AIC
Baltic Sea
Mecklenburg Bight
Between 11.55° to 12.55° E and 54.09° to 54.96° N

ABSTRACT

The detailed analysis of patterns of benthic community distribution related to selected environmental parameters provides a basis for predictive modelling of species distribution. Species-specific models predicting the probability of occurrence relative to environmental and sedimentological characteristics were developed in this study for 29 macrofaunal species common for our study area using a logistic regression modelling approach. This way, a good description of the occurrence of species along gradients of single environmental variables was obtained. Subsequently, we used a technique for a predictive modelling of species distributions in response to abiotic parameters based on single-factor logistic regression models, utilizing AIC and Akaike weights for multimodel inference. Thus, probabilities of occurrence for selected exemplary species (*Arctica islandica*, *Hediste diversicolor*, *Pygospio elegans*, *Tubificoides benedii* and *Scoloplos armiger*) were modelled and mapped. For all species the use of this newly available combination of methods provided fairly accurate results of a distribution prediction. Water depth that represents a type of integral parameter remained the key factor determining the species distribution among the parameters considered within the study scale. This is particularly relevant for species that find their optima habitat here, but also for those as *H. diversicolor* that occur only locally and in comparatively low densities. Total organic content, sorting and, for *S. armiger*, salinity also had noticeable effect in the determination of suitable habitats for benthic macrofauna. The employed technique proved to be appropriate for modelling of the benthic species habitat suitability, at least within comparable spatial scales and variability of environmental factors.

1. Introduction

Climate change models assume a drastic change in the food web structure, a shift in species composition towards warm water species and growing benthic deserts on the sea floor as up-coming changes for the Baltic Sea ecosystem (Philippart et al., 2007). Predictive modelling of species distribution can be a valuable tool in management directed towards the sustainable development of the Baltic Sea. Studies on various scales are required to extend our knowledge of habitat change effects.

Response of macrofaunal assemblages to substrate composition, hydrographic parameters and their variation is declared by many different studies (e.g. Sanders, 1968; Rhoads, 1974; O'Brien et al., 2003; Laine, 2003; Perus and Bonsdorff, 2004; Ellis et al., 2006). Particular establishments regarding the dynamics and structure of biotic/environmental interactions are required to evaluate natural and anthropogenic influences and effects on the ecological systems (Pavlikakis and Tsihrintzis, 2000; Glockzin and Zettler, 2008a). An exploratory statistical description of the prevailing ecological structure based on observations is always the indispensible first step (Bourget and Fortin, 1995).

Recently, a number of studies have succeeded in the development of effective statistical models of benthic distribution. Ysebaert et al. (2002) successfully applied logistic regression to derive response surfaces of distributions for 20 common macrobenthic species found in the Schelde estuary in the Netherlands related to salinity, depth, current velocity, and sediment characteristics. Thrush et al. (2003) developed species-specific models for 13 benthic species of New Zealand estuaries that predicted probability of occurrence as well as maximum abundance relative to sediment mud content using logistic regression for distribution modelling and 'factor ceiling' method (Blackburn et al., 1992) for maximum density modelling. Ellis et al. (2006) modelled the distribution of 13 representative macrobenthic species in New Zealand estuarine gradients using logistic regression and classification system based on 'controlling factors' with sediment characteristics, elevation, tidal currents, and wind-wave disturbance employed as predictors. They faced complications to fully test the latter approach due to differences in scales of collected benthic data and of higher level physical variables. Meissner et al. (2008) developed habitat models for *Nephtys* species in the German Bight (North Sea) with median grain size, mud content, depth, and salinity as explanatory variables by application of

---

* Corresponding author. Leibniz Institute for Baltic Sea Research, Seestrasse 15, D-18119 Rostock, Germany. Fax: +49 381 5197 440.
E-mail address: michael.zettler@io-warnemuende.de (M.L. Zettler).

0924-7963/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmarsys.2009.10.001

multivariate adaptive regression spline techniques (MARS). Many researches indicated difficulties due to the complexity of identification of the underlying causal mechanisms controlling species distribution, further extended by the fact that animals modify their physical environment, and many physical parameters co-vary (Ellis et al., 2006).

The present study contributes to the development of statistical models that are able to predict the distribution of benthic macrofaunal species as a function of environmental variables. Models that forecast the behaviour of species distribution versus changes in environmental factors (sensu Legendre and Legendre, 1998) provide an insight into chronic habitat change (regarding these parameters), though they do not provide insight into the acute effects associated with disturbance events (Thrush et al., 2003; Ellis et al., 2006). Model estimations based on a data set consisting of the response variables (e.g. species occurrence or abundance) and on a set of predictor variables (e.g. environmental parameters) can be used to predict the spatial distribution of species in a habitat with known or defined environmental settings (Ysebaert et al., 2002; Ellis et al., 2006; Meissner et al., 2008). Statistical models are able to relate ecological features to environmental factors and, through validation and modification, are able to reveal details in the underlying mechanisms responsible for structure and organization of communities (Austin, 1987; Glockzin and Zettler, 2008a).

Our investigation is focused on a limited area in the south-western Baltic Sea. The pre-work of an exploratory statistical description of the prevailing ecological structure is an essential first step towards modelling, and it was completed for the area of interest beforehand (Gogina et al., 2009—this volume). The identified distinct benthic assemblages have proved to associate with certain spatial regions and specific limits of environmental parameters. The reduction of macrozoobenthic data to presence/absence is forced by the absence of temporal homogeneity and is used here to eliminate the influence of patchiness in spatial distribution that macrofauna exhibits (McArdle and Blackwell, 1989; Legendre et al., 1997; Thrush et al., 2003).

2. Materials and methods

2.1. Study area

The study area is located in the south-western Baltic Sea, between 11.55° to 12.55° E and 54.09° to 54.96° N (Fig. 1). It is bounded by the eastern part of the Mecklenburg Bight and the western region of the Kadetrinne, with its northern and southern limits defined by Danish and German land boundaries. Some geographical details about the area, which is characterised by a relatively high biodiversity of both saline and brackish water species, as well as the analysis of benthic community structure, can be found in Gogina et al. (2009—this volume).

2.2. Data used for model estimation

The study is based upon the data of benthic macrofauna and associated sediment and near-bottom environmental characteristics, sampled at 208 stations (Fig. 1a). For 72 of these stations a full set of the abiotic parameters considered is available. For modelling purposes the species abundance data was reduced to presence/absence.

The description of methods for benthic macrofauna sampling and abiotic factor determination, as well as the selection process for extraction of 29 representative macrobenthic species modelled here, can be found in Gogina et al. (2009—this volume).

2.3. Additional environmental data for predictive modelling

Additional data sets were required to compile the grids of each abiotic descriptor, needed for predictive estimates of species distribution (probability of occurrence) for the whole investigation area. The distribution surfaces obtained for each of the environmental variables considered are presented in Fig. 1.

For the bathymetry a high-resolution digital elevation model (DEM) was created using measured data provided by the Federal Maritime and Hydrographic Agency (BSH) and a regional grid data set from Seifert et al. (2001), covering the Belt Sea region. For more details see the description of DEM design in Meyer et al. (2008). Grid data sets for near-bottom oxygen content and salinity were based on the modelled hydrographical data, averaged for years 1960–2005 with the resolution of 3 nautical miles (Neumann and Schernewski, 2008), covering the whole western Baltic sea area. Grids for sediment parameters like median grain size, sorting, skewness and permeability are derived from the internal database of the Leibniz Institute for Baltic Sea Research Warnemuende (IOW; Bobertz and Harff, 2004), integrating the data of about five decades of marine investigations. The average distance between adjacent sample sites is less than 1 nautical mile. From the IOW database external data on total organic content was also available, however, only for a limited area. Hence, this data was agglomerated together with the observed data used for model estimation to increase the area covered and the density of data points. Nevertheless, only a part of the investigation area could be covered with the compiled grid of this parameter (Fig. 1d). Parameters were interpolated using ordinary kriging with spherical fitted models of semivariograms into a grid with the resolution of about 0.005 decimal degrees (approximately 0.5 km with respect to longitude).

Ysebaert et al. (2002) favoured the usage of modelled estimates of environmental variables over the data measured directly and simultaneously with benthic sampling. The argumentation included the available high spatial resolution and a sort of smoothing caused by simulation, e.g. elimination of outfits. However, taking into account the complexity of the functioning of ecosystems, the uncertainty of simulations may increase the complexity of the interpretation of derived empirical relationships. Utilization of simulated data for the model estimation is forced merely by the necessity and absence of alternatives. The preliminary explicit exploratory analysis of environmental framework should exclusively be based on direct in situ measurements. Therefore, to enable the investigation of autecological relationships we rely our model calibration on directly observed data to the highest extent possible, applying minimum transformations to lessen the reduction of information contained in the data (Gogina et al. 2009—this volume). Yet, the prediction is based partly on modelled data of sufficient resolution available for the study area, thus, allowing the validation of modelling success.

2.4. Statistical analysis and data treatment

2.4.1. Univariate logistic regression

Logistic regression of biotic data reduced to presence/absence was employed to model the probability of occurrence of 29 discriminating species, using the considered environmental factors (water depth, salinity, oxygen concentrations, total organic content, median grain size, sorting, skewness and permeability of sediments) as explanatory variables. These factors are generally assumed to have direct or indirect impact on distribution of macrobenthic species. The logit function in a logistic regression is the special case of a link in a generalised linear model, known as canonical for the binomial distribution. Application of logistic regression methods in modelling species distribution is not new. This method was widely used in plant ecology (e.g. Guisan et al. 1999) and also in aquatic ecology, but to a lesser extent. Thrush et al. (2003) concentrated their investigation on a single environmental factor—sediment mud content, Ysebaert et al. (2002) performed a comprehensive study, using salinity, depth, flow parameters, median grain size and mud content as predictors. The present study represents one of the first applications of this technique to benthic habitats of the Baltic Sea.

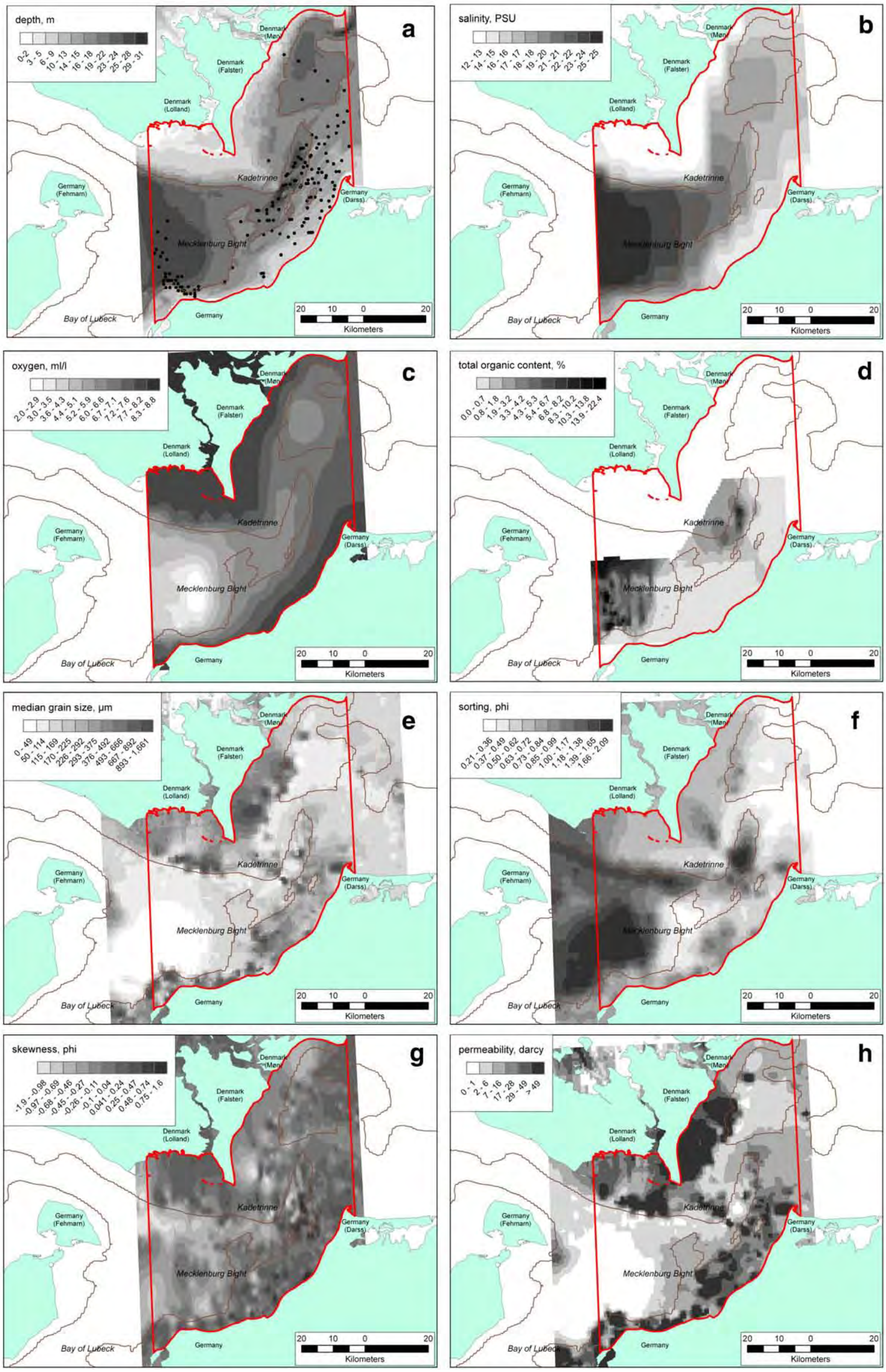


Fig. 1. Distribution surfaces covering the study area (limited by the thick line) generated for each of the environmental variables considered using ordinary kriging. Dots indicate stations sampled for macrofauna. Geographical data ESRI (2003); projection UTM on WGS84.

The logistic regression model relates the probability of observing the species p to one or more predictor variables x (in our study separately to each of the environmental factors) using the logistic link function. The regression model can be written as

$$p(x) = e^{z} / (1 + e^{z}) \tag{1}$$

where z is the function of the explanatory variable(s). When this is a first-order polynomial, the response is a logistic, S-shaped function. In the special case of Gaussian logistic model when z is a second-order polynomial, the response will approximate a bell-shaped function. For this particular case Eq. (1) can be written as

$$p(x) = e^{(b_0 + b_1 x + b_2 x^2)} / (1 + e^{(b_0 + b_1 x + b_2 x^2)}) \tag{2}$$

where b_0, b_1, and b_2 are regression parameters. They are estimated by maximum likelihood, assuming a binomially distributed error term (Legendre and Legendre, 1998; Ysebaert et al. 2002; Wisz and Guisan, 2009). When the estimation of z term parameters is based on log-transformed data this can be interpreted as a further extension of the method, aiming to produce an ecologically more plausible response for certain species. This idea was adopted from Thrush et al. (2003, 2005) who found such a transformation to produce the most realistic response to changes in sediment mud content for the occurrence of some of the investigated species.

Thus, for each factor and taxon combination, different functions were used (linear, Gaussian, polynomial) and were based on either raw or log-transformed data. The Wald statistic was used to estimate the model's significance, with a significance level defined at 0.05. The final single-factor model used for each species was the function (of the particular factor) that explained the most variability. The evaluation of the model fit was performed by means of visual control of half-normal plots of residuals and plots of residuals versus predicted values and by considering the percentage of concordant pairs.

2.4.2. Collinearity among predictors

The complexity of prediction of species distribution is caused by the complexity of interactions of various factors. To give some insight in the extent and direction of collinearity among the considered factors, relationships between pairwise combinations of individual predictors are portrayed in Fig. 2 using scatterplot matrix. Additionally, analyses of correlations among environmental parameters and values of Pearson correlation coefficients can be found in Gogina et al. (2009—this volume).

2.4.3. Information-theoretic approach and utilization of the Akaike weights

To combine the results of single-factor models and to draw the multimodel inference we considered the information-theoretic philosophy described in Burnham and Anderson (2004). It relies on the calculation of the Akaike's information criterion (AIC) as a model

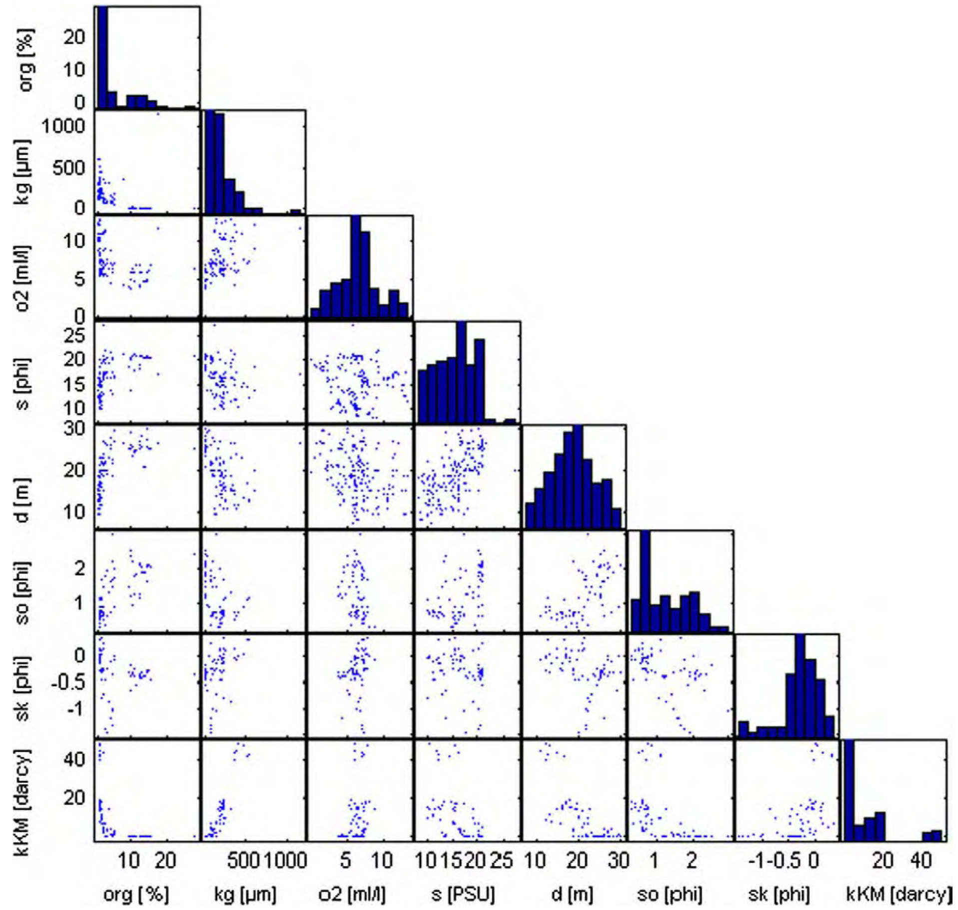


Fig. 2. Scatterplot matrix showing relationships between environmental factors over the data set. Environmental factors notation: d—water depth, s—near-bottom salinity, org—total organic content, o2—near-bottom oxygen concentration, kg—median grain size, so—sorting, sk—skewness, kKM—permeability.

selection tool. AIC is a measure of the relative Kullback–Leibler information loss when the candidate model *i* is used to approximate the truth *j*. Generally appropriate small-sample version (designed for $n/K>40$) of criterion $\mathrm{AIC_c}$ is calculated as

$$\mathrm{AIC_c} = -2\ln(L) + 2K(K+1)/(n\text{–}K\text{–}1) \tag{3}$$

where *L* is a value of the maximised log-likelihood over the estimated parameters given the data and the model, *K*—number of parameters in model *i*, and *n* is the sample sizes. AIC (or $\mathrm{AIC_c}$) are calculated for each of the candidate models, then these values are rescaled to calculate delta AIC (Δ_i) so that the model possessing the lowest AIC value has a Δ_i value of 0:

$$\Delta_i = \mathrm{AIC}_i\text{–}\mathrm{AIC_{min}} \tag{4}$$

where $\mathrm{AIC_{min}}$ is the smallest AIC value in the model set. The model with the lowest AIC value is considered to be the best approximating model in the candidate set. The larger the value of Δ_i the less plausible the fitted model *i* is. Burnham and Anderson (2004) suggest a simple rule of thumb in assessing relative merits of models in a set: models with $\Delta_i \le 2$ have strong support, models with Δ_i values between 2 and 10 have considerable support and those with $\Delta_i > 10$ have essentially no support. Akaike weights (ω_i) are now calculated for each of the *r* candidate models:

$$\omega_i = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum\limits_{r=1}^{R} \exp\left(-\frac{1}{2}\Delta_i\right).} \tag{5}$$

The ω_i are scaled between 0 and 1, and all Akaike weights sum to one—ω_i values represent the proportion of evidence for a particular model *i* in the total evidence supporting all of the models (Wisz and Guisan, 2009). A model that possesses the largest ω_i value is the most parsimonious and has most support among the specified candidate models given the data. When more than one model is supported by the data it is possible to calculate a global model that is a weighted average of all the candidate models in the *a priori* defined set. New parameter estimates for each term in the global model can be computed by weighting them by the Akaike weights

$$\hat{\hat{\theta}} = \sum_{i=1}^{R} \omega_i \hat{\theta}_i \tag{6}$$

where is $\hat{\hat{\theta}}$ the model averaged parameter estimate based on all *R* models, and $\hat{\theta}_i$ is the parameter estimate for a term in a candidate model *i* with the Akaike weight ω_i. For terms that do not feature in a candidate model but are present in the global model the parameter estimate is taken to be zero. Thus, if the goal is prediction, the point inference can be based on the entire set of models using Akaike weights within the overlapping of single-factor models as weight factors to generally estimate the probability of species occurrence.

2.4.4. Habitat suitability mapping

The method applicability was tested on five selected species (*Arctica islandica*, *Hediste diversicolor*, *Pygospio elegans*, *Tubificoides benedii* and *S. armiger*). Finally, estimates derived for these species were implemented in a geographical information system. As the grid data for total organic content was only available for a limited part of the investigation area (see Section 2.3), two model sets were considered for the final prediction: one with and one without the model of species occurrence probability as a function of total organic content included. The final value assumed for the overlapping areas of two model sets was the value from the model set with all parameters considered; the rest of the area was filled out with the results based on the model set lacking the total organic content factor. Concordance was calculated to provide the estimate of modelling success.

All analysis were carried out using SPSS (SPSS, Inc.), Statistica (StatSoft, Inc., 2007), PRIMER (PRIMER-E, Plymouth Marine Laboratory;

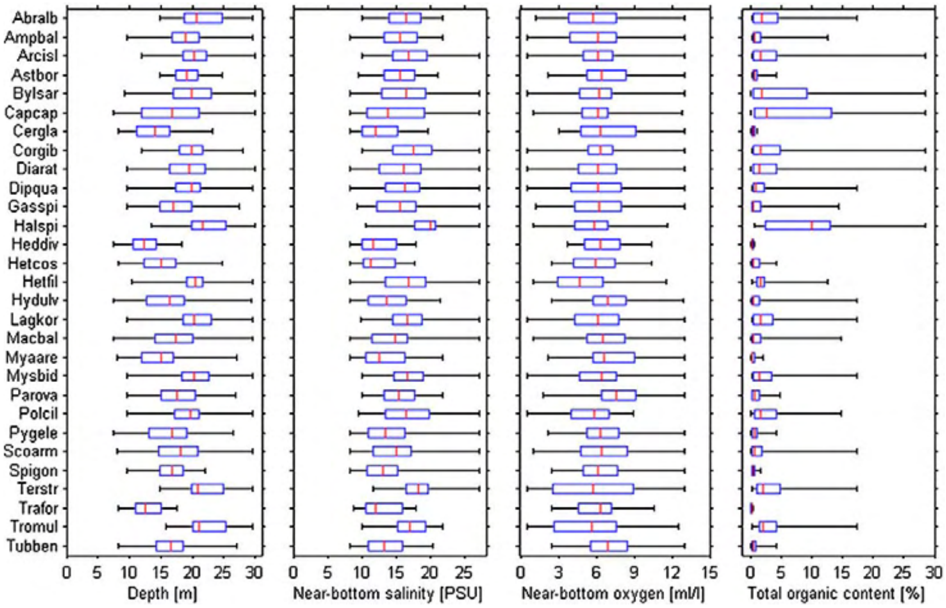


Fig. 3. Box-and-whisker plots for 29 macrobenthic species with respect to depth, near-bottom salinity and oxygen, total organic content. Species are ordered alphabetically, 6-letter codes from top to the bottom referring to *Abra alba*, *Ampharete baltica*, *Arctica islandica*, *Astarte borealis*, *Bylgides sarsi*, *Capitella capitata*, *Cerastoderma glaucum*, *Corbula gibba*, *Diastylis rathkei*, *Dipolydora quadrilobata*, *Gastrosaccus spinifer*, *Halicryptus spinulosus*, *Hediste diversicolor*, *Heterochaeta costata*, *Heteromastus filiformis*, *Hydrobia ulvae*, *Lagis koreni*, *Macoma balthica*, *Mya arenaria*, *Mysella bidentata*, *Parvicardium ovale*, *Polydora ciliata*, *Pygospio elegans*, *Scoloplos armiger*, *Spio goniocephala*, *Terebellides stroemi*, *Travisia forbesii*, *Trochochaeta multisetosa*, and *Tubificoides benedii* . 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 of the observations.

Clarke and Warwick, 1994), MATLAB and ArcMap (ESRI Inc., Redlands, USA).

3. Results

3.1. Distribution of the observed species occurrence along considered environmental variables

The observed frequency distribution of 29 macrobenthic species (selected as representative/dominating within the study area as described in Gogina et al., 2009—this volume) along considered environmental variables is presented as box-and-whisker plots in Figs. 3 and 4.

With respect to depth, while the range of species occurrence represented by “whiskers” generally corresponds to the limits defined by cumulative abundance distribution curves (see Gogina et al., 2009—this volume), analysis of “boxes” suggests differences between distributions of cumulative abundance and occurrence curves. E.g. 25th and 75th occurrence percentiles for *Capitella capitata* correspond to ca.12 and 21 m depth, however, the cumulative percentage of its abundance increases between depth values of 20 and 22.5 m from less than 0.2 to more than 0.8, indicating the range of maximum response against depth within the data limits. This is noteworthy, because this sudden increase in cumulative abundance is due to very few locations where the density of these polychaetes is up to two orders of magnitude higher than the average abundance in the region.

Though generally all the dominating species are mesohaline, and the salinity gradient is limited within the study area, the graphs clearly distinguish characteristic species tending to polyhalinity (*Halicryptus spinulosus*, *Terebellides stroemi* and *Trochochaeta multisetosa*) from species preferring the lower salinity within the considered factor range (*Cerastoderma glaucum*, *H. diversicolor*, *Heterochaeta costata*, *Hydrobia ulvae*, *Mya arenaria*, *Spio goniocephala*, *Travisia forbesii*, *T. benedii*). Other species take intermediate positions regarding salinity, with widest range of occurrence observed in our data, e.g. for *Bylgides sarsi* and *Diastylis rathkei*.

The influence of total organic content on the occurrence of species seems to be most evident. Some species indicate an exceptionally narrow range of occurrence at low values of organic content, e.g. *C. glaucum*, *H. diversicolor*, *T. forbesii*, and, to a smaller degree, *Astarte borealis*, *H. costata*, *M. arenaria*, and *S. goniocephala*. Others, such as *B. sarsi*, *C. capitata*, *H. spinulosus*, occur along the whole range of organic content.

3.2. Single-factor response curves

Logistic regression models predicting the probability of species occurrence as a function of each of the abiotic parameters considered were generated for each of the 29 species. Single-factor models are defined here following Eqs. (1) and (2). Table 1, Figs. 5 and 6 give an example of obtained response curves for a single environmental variable, showing the fitted logit curves for the 29 macrobenthic species in relation to water depth and to the total organic content. These are the factors which reveal the most distinct patterns in explaining the variability of benthic fauna distribution. Water depth actually represents a type of integral parameter (‘proxy’) that combines the effects of various habitat features that are either described by the available data, or not measured (or even not measurable). The response curves obtained were in general agreement with the observed distribution from Figs. 3 and 4, which supports the accuracy of the logistic modelling approach employed. Models predicting the probability of species occurrence relative to water depth show a concordance between 64.1 and 93.5%, relative to the total organic content—between 57.4 and 94.3% (Table 1). The variety of functional forms among the species response to water depth indicates that the occurrence of species is driven by species-specific sensitivity to that factor, with a non-constant rate of change characteristic for all of the species. The probability of occurrence of species such as *C. glaucum*, *M. arenaria*, *H. costata*, *H. diversicolor* and *T. forbesii* was higher at shallow depths. The sharp drop of the curve with increasing depth indicates that some species, e.g. *H. ulvae* and *P. elegans*, are highly sensitive to this parameter and avoid deep regions. Bell-shaped curves with an optimum at intermediate depths, as seen for *S. goniocephala*, *T. benedii* and *Gastrosaccus spinifer*, indicate the decrease of probability of occurrence both at the lower and upper ends of the factor range. Some species, e.g. *Mysella bidentata*,

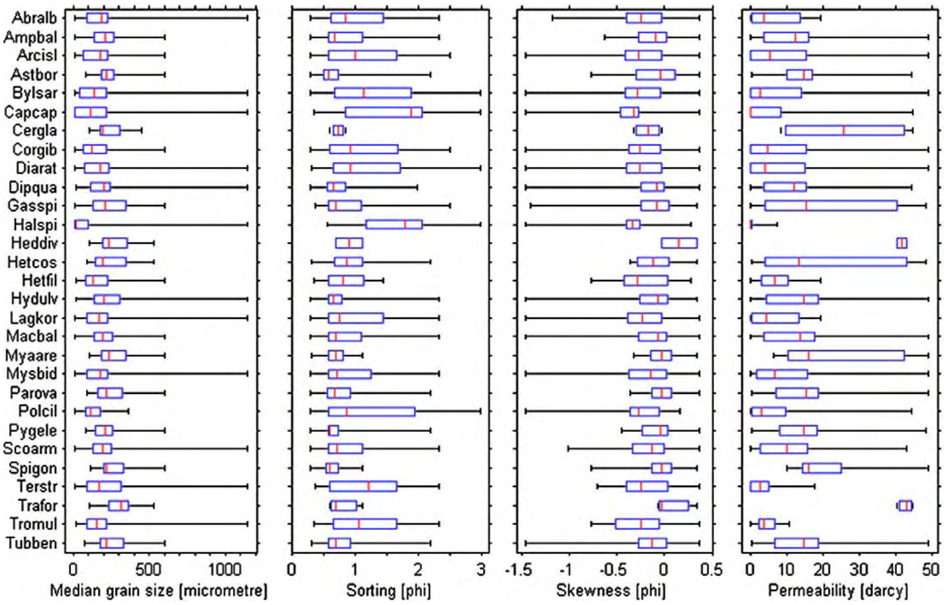


Fig. 4. Box-and-whisker plots for 29 macrobenthic species with respect to median grain size, grain-size parameters sorting and skewness, and permeability of surface sediments. For all the explanation see caption of Fig. 3.

Table 1
Logistic regression models of species occurrence.

Species	Model_d (x — water depth, m)	p(model)	Con %	Model_org (x — total organic content, %)	p(model)	Con %
Abralb	$p(x)=\exp(-4.43+0.21^*x)/(1+\exp(-4.43+0.21^*x))$	<0.0001	76.0	$p(x)=\exp(-0.84+0.48x-0.04x^2)/(1+\exp(-0.84+0.48x-0.04x^2))$	0.0046	71.6
Arcisl	$p(x)=\exp(-20.2+1.97x-0.04x^2)/(1+\exp(-20.2+1.97x-0.04x^2))$	<0.0001	82.1	$p(x)=\exp(-0.62+2.32\log(x+1)-0.75(\log(x+1))^2)/(1+\exp(-0.62+2.32\log(x+1)-0.75(\log(x+1))^2))$	0.0161	65.5
Astbor	$p(x)=\exp(-42.3+4.55x-0.12x^2)/(1+\exp(-42.3+4.55x-0.12x^2))$	<0.0001	86.7	$p(x)=\exp(0.35-0.52x)/(1+\exp(0.35-0.52x))$	<0.0001	71.0
Cergla	$p(x)=\exp(4.02-0.30x)/(1+\exp(4.02-0.30x))$	<0.0001	84.0	$p(x)=\exp(-0.22-1.5x)/(1+\exp(-0.22-1.5x))$	<0.0001	59.9
Corgib	$p(x)=\exp(-16.0+1.59x-0.04x^2)/(1+\exp(-16.0+1.59x-0.04x^2))$	<0.0001	73.9	$p(x)=\exp(-0.84+2.1\log(x+1)-0.64(\log(x+1))^2)/(1+\exp(-0.84+2.1\log(x+1)-0.64(\log(x+1))^2))$	0.0353	63.4
Macbal	$p(x)=\exp(73.9-6.0x+0.12x^2)/(1+\exp(73.9-6.0x+0.12x^2))$	<0.0001	92.2	$p(x)=\exp(2.62-0.47x)/(1+\exp(2.62-0.47x))$	<0.0001	94.3
Myaare	$p(x)=\exp(8.04-0.45x)/(1+\exp(8.04-0.45x))$	<0.0001	89.3	$p(x)=\exp(1.25-1.5x)/(1+\exp(1.25-1.5x))$	<0.0001	87.3
Mysbid	$p(x)=\exp(-12.7+1.24x-0.03x^2)/(1+\exp(-12.7+1.24x-0.03x^2))$	<0.0001	75.7	$p(x)=\exp(-0.91+4.14\log(x+1)-1.53(\log(x+1))^2)/(1+\exp(-0.91+4.14\log(x+1)-1.53(\log(x+1))^2))$	<0.0001	76.3
Parova	$p(x)=\exp(-5.56+0.55x-0.02x^2)/(1+\exp(-5.56+0.55x-0.02x^2))$	0.0114	64.1	$p(x)=\exp(0.01-0.32x)/(1+\exp(0.01-0.32x))$	<0.0001	68.6
Diarat	$p(x)=\exp(-9.06+1.11x-0.03x^2)/(1+\exp(-9.06+1.11x-0.03x^2))$	0.0003	80.8	$p(x)=\exp(0.13+4.54\log(x+1)-1.5(\log(x+1))^2)/(1+\exp(0.13+4.54\log(x+1)-1.5(\log(x+1))^2))$	0.0008	77.8
Hydulv	$p(x)=\exp(8.58-0.41x)/(1+\exp(8.58-0.41x))$	<0.0001	89.0	$p(x)=\exp(1.08-0.28x)/(1+\exp(1.08-0.28x))$	<0.0001	82.9
Gasspi	$p(x)=\exp(-4.22+0.59x-0.02x^2)/(1+\exp(-4.22+0.59x-0.02x^2))$	0.0004	69.0	$p(x)=\exp(0.09-0.18x)/(1+\exp(0.09-0.18x))$	0.0002	69.5
Hetcos	$p(x)=\exp(2.01-0.2x)/(1+\exp(2.01-0.2x))$	<0.0001	77.0	$p(x)=\exp(-1.96-0.29x)/(1+\exp(-1.96-0.29x))$	0.0448	69.8
Tubben	$p(x)=\exp(-6.02+0.93x-0.03x^2)/(1+\exp(-6.02+0.93x-0.03x^2))$	<0.0001	78.5	$p(x)=\exp(1.23-0.78x)/(1+\exp(1.23-0.78x))$	<0.0001	81.7
Ampbal	$p(x)=\exp(-11.8+1.25x-0.03x^2)/(1+\exp(-11.8+1.25x-0.03x^2))$	<0.0001	70.3	$p(x)=\exp(0.19-0.17x)/(1+\exp(0.19-0.17x))$	0.0003	61.1
Bylsar	$p(x)=\exp(-3.75+0.31x)/(1+\exp(-3.75+0.31x))$	<0.0001	84.1	$p(x)=\exp(0.15+1.31x)/(1+\exp(0.15+1.31x))$	<0.0001	83.0
Capcap	$p(x)=\exp(7.05-0.76x+0.02x^2)/(1+\exp(7.05-0.76x+0.02x^2))$	0.0003	66.2	$p(x)=\exp(-1.41+0.16x)/(1+\exp(-1.41+0.16x))$	<0.0001	68.6
Dipqua	$p(x)=\exp(-12.4+1.22x-0.03x^2)/(1+\exp(-12.4+1.22x-0.03x^2))$	<0.0001	71.2	$p(x)=\exp(0.12-0.14x)/(1+\exp(0.12-0.14x))$	0.0013	57.4
Heddiv	$p(x)=\exp(7.5-0.56x)/(1+\exp(7.5-0.56x))$	<0.0001	93.5	$p(x)=\exp(-0.38-3.62x)/(1+\exp(-0.38-3.62x))$	<0.0001	86.6
Hetfil	$p(x)=\exp(-20.6+1.91x-0.04x^2)/(1+\exp(-20.6+1.91x-0.04x^2))$	<0.0001	79.4	$p(x)=\exp(1.94-0.49\log(x+1)+0.04(\log(x+1))^2)/(1+\exp(1.94-0.49\log(x+1)+0.04(\log(x+1))^2))$	0.0005	73.6
Lagkor	$p(x)=\exp(-3.21+0.17x)/(1+\exp(-3.21+0.17x))$	<0.0001	72.7	$p(x)=\exp(-1.2+4.07\log(x+1)-1.48(\log(x+1))^2)/(1+\exp(-1.2+4.07\log(x+1)-1.48(\log(x+1))^2))$	<0.0001	73.7
Polcil	$p(x)=\exp(-8.07+0.76x-0.02x^2)/(1+\exp(-8.07+0.76x-0.02x^2))$	0.0020	68.5	–	–	–
Pygele	$p(x)=\exp(5.99-0.3x)/(1+\exp(5.99-0.3x))$	<0.0001	82.2	$p(x)=\exp(0.69-0.5x)/(1+\exp(0.69-0.5x))$	<0.0001	74.4
Scoarm	$p(x)=\exp(3.69-0.12x)/(1+\exp(3.69-0.12x))$	0.0003	68.3	$p(x)=\exp(2.18-0.28x)/(1+\exp(2.18-0.28x))$	<0.0001	81.8
Spigon	$p(x)=\exp(-13.1+1.82x-0.06x^2)/(1+\exp(-13.1+1.82x-0.06x^2))$	<0.0001	79.7	$p(x)=\exp(0.91-2.49x)/(1+\exp(0.91-2.49x))$	<0.0001	80.1
Terstr	$p(x)=\exp(-5.28+0.22x)/(1+\exp(-5.28+0.22x))$	<0.0001	78.1	$p(x)=\exp(-3.14+4.06\log(x+1)-1.29(\log(x+1))^2)/(1+\exp(-3.14+4.06\log(x+1)-1.29(\log(x+1))^2))$	0.0006	71.7
Trafor	$p(x)=\exp(3.89-0.34x)/(1+\exp(3.89-0.34x))$	<0.0001	87.4	$p(x)=\exp(0.97-7.78x)/(1+\exp(0.97-7.78x))$	<0.0001	92.3
Tromul	$p(x)=\exp(-5.77+0.24x)/(1+\exp(-5.77+0.24x))$	<0.0001	80.5	$p(x)=\exp(-4.49+6.32\log(x+1)-2.06(\log(x+1))^2)/(1+\exp(-4.49+6.32\log(x+1)-2.06(\log(x+1))^2))$	<0.0001	80.1
Halspi	$p(x)=\exp(-5.79+0.23x)/(1+\exp(-5.79+0.23x))$	<0.0001	78.0	$p(x)=\exp(-2.15+0.35x)/(1+\exp(-2.15+0.35x))$	<0.0001	91.2

$p(x)$ is the estimated probability that the species occurs. Species abbreviation is as in Fig. 3. p(model) indicates the model significance level, tested using Wald test based on the χ^2-distribution ($p<0.05$), and Con % is the percentage of concordant pairs.

Dipolydora quadrilobata and *D. rathkei*, showed a clear optimum towards the higher end of the depth range, differing in the position of their optimum and in their tolerance towards the lower end of the factor range. *Abra alba*, *H. spinulosus*, *T. multisetosa*, *T. stroemi* and, to a lesser extent, *B. sarsi* had the highest probability of occurrence at the deepest zones. A broad tolerance for water depth was indicated, for instance for *S. armiger*. The inverse form of the Gaussian response towards that factor shown by *Macoma balthica* and *C. capitata* can be interpreted owning to their cosmopolite behaviour in the context of depth within the studied ranges and dependency on other variables. It may also be regarded as a realistic form of the response function, when considered as a slice of the bimodal response explained by the competition exclusion in the middle of broad tolerance to an environmental gradient.

Most macrobenthic species showed a high probability of occurrence at the lower end of the total organic content range (e.g. *M. balthica*, *M. arenaria*, *D. quadrilobata*) with only two species, *C. capitata* and *H. spinulosus*, having an increasing probability of occurrence with the increase of total organic content. Species such as *M. bidentata*, *A. alba*, *Lagis koreni*, *T. stroemi* showed skewed unimodal curve forms with an optimum tending towards the lower end of the factor range. The models developed for *Heteromastus filiformis*, and to a lesser extent *A. islandica* and *Corbula gibba*, have indicated that the occurrence of these species is not sensitive to a wide range of sediment total organic content.

3.3. Modelling—predicting the probability of species occurrence

The multimodel inference technique was applied to five selected species (*A. islandica*, *H. diversicolor*, *P. elegans*, *T. benedii* and *S. armiger*), exemplarily chosen as representatives for soft and sandy-bottom regions of the investigation area. Maximum densities and frequencies of occurrence for selected species are given in Table 2.

Table 3 provides maximum-likelihood estimates of logistic regression parameters for species response surfaces to each of the abiotic factors. Single-factor models are defined as described in Section 3.2. For *H. diversicolor* models of the factors sorting, skewness and permeability were not supported by a sufficient value of occurrence events (the species was recorded only at 2 stations out of 78 covered with data on corresponding abiotic variables), and thus

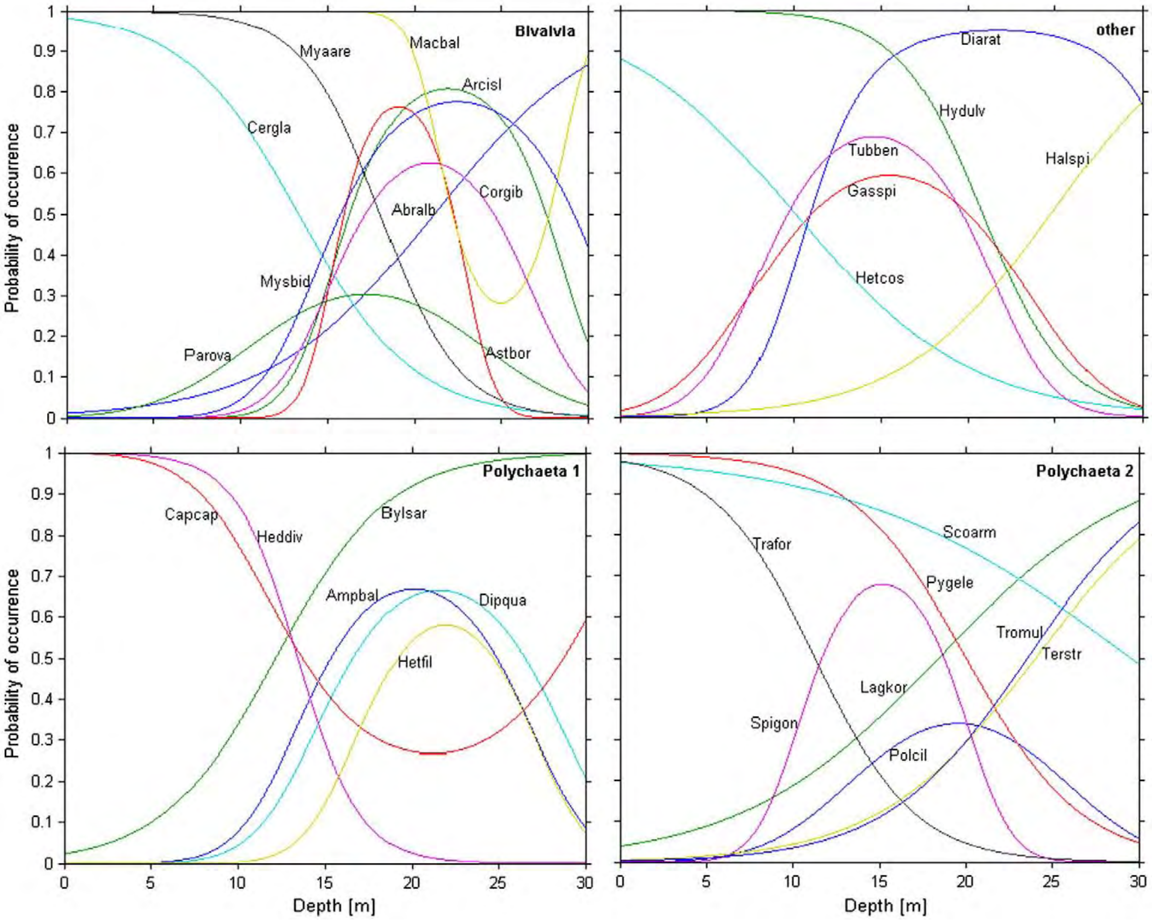


Fig. 5. Logistic regression models predicting probability of occurrence for 29 macrobenthic species relative to water depth. Species abbreviation as in Fig. 3.

models for these factors were not included in the set for model averaging defined for this species.

Akaike weights derived for two model sets for each species—one set including the single-factor model of total organic content used as predictor and one without it—are presented in the lower part of the Table 3. To assess the accuracy of prediction the values of concordance were estimated.

Regarding Akaike weight values, the factor dominating in the explanation of variability of *A. islandica* was depth. Some significant part of distribution variability was also explained by permeability, while all other environmental parameters considered explain a small part of uncertainty, though models with oxygen, median grain size and skewness did not differ significantly from the random chance. The water depth model also provided the greatest strength of evidence for *H. diversicolor*, followed by the model of total organic content, which possessed a considerably lower weight in the model averaging. Factors dominating in the explanation of variability in distribution of *P. elegans* were total organic content and depth. Some significant part of distribution variability was also explained by sorting. When it was not possible to take the total organic content into account salinity also explained a noticeable part of variance. *S. armiger* indicated a well-defined response to salinity, to total organic content, (when it was possible to use this factor as predictor) and to sorting. Among the considered abiotic factors, the variability of distribution of *T. benedii* was almost entirely explained by depth, with some information hidden in the gradient of total organic content.

In Fig. 7 the results of the prediction with application of the described technique are mapped and compared to the observed species abundance data. The visual inspection reveals that higher values of species abundance prevail where higher probabilities of species occurrence are modelled.

For *A. islandica* the preferable conditions for the settlement within an area and considered ranges of environmental factors included high values of depth (approximately >18 m) and salinity as well as wide ranges of total organic content and sedimentological parameters. On the contrary, *H. diversicolor* preferred low-saline regions shallower than 18 m. Therefore, these species seem to present the biological antipodes in our study area. *P. elegans* revealed disinclination to regions with high total organic content and preferred more permeable substrate (see Fig. 7).

4. Discussion

As pointed out by Praca et al. (2008) the use of temporally heterogeneous data confounds the effect of interannual variations in species occurrence and environmental conditions. However, our objectives were to attempt a general description of species habitats and to investigate the predictive abilities of the modelling technique at the selected spatial scale. Models determining the distribution of exemplary macrofaunal species common for the south-western Baltic Sea from changing environmental variables, such as depth, salinity and sediment characteristics, have been successfully developed. These

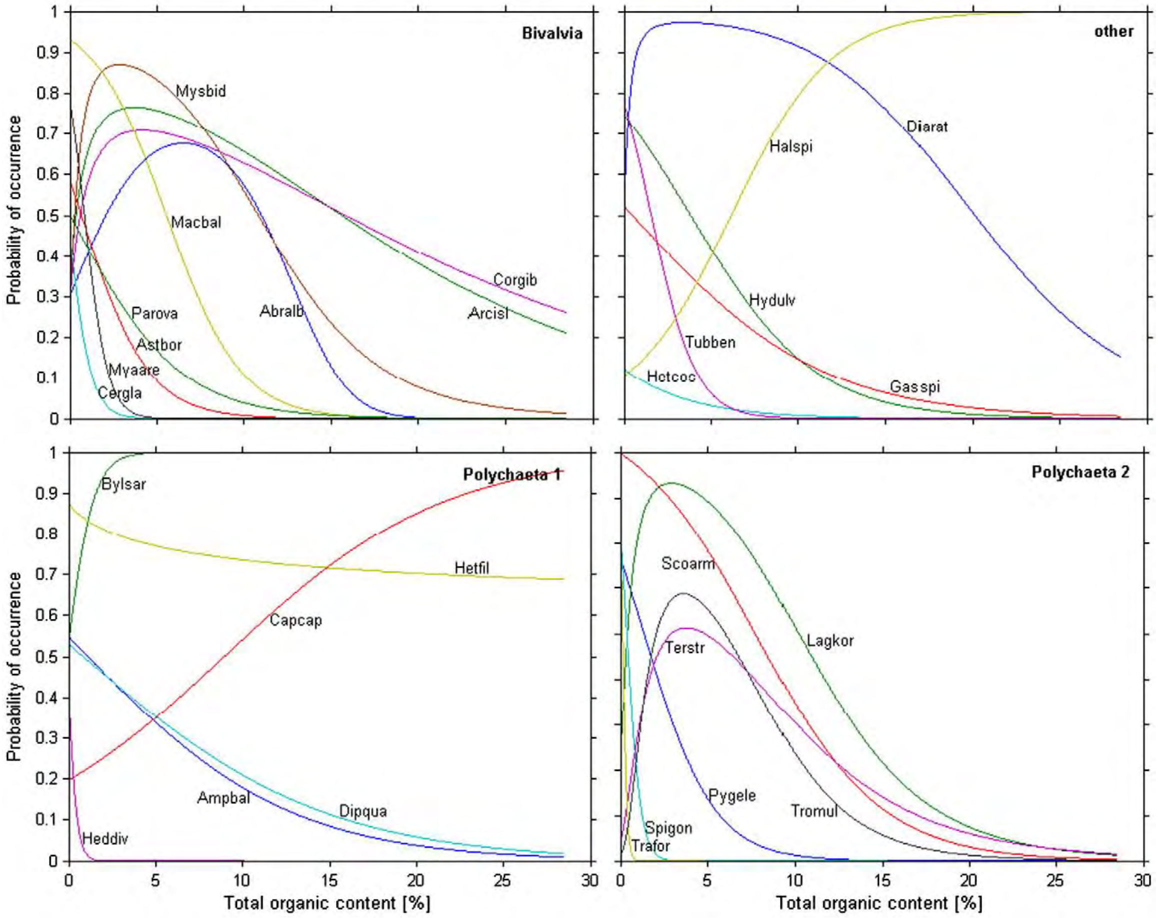


Fig. 6. Logistic regression models predicting probability of occurrence for 29 macrobenthic species relative to sediment total organic content. Species abbreviation as in Fig. 3.

models are considered to be reasonably general, i.e. they are able to resolve the distribution of species over a comparatively large spatial scale (ca. >100 m) as opposed to a metre-by-metre basis. Confirming the postulates of Ysebaert et al. (2002), Thrush et al. (2003), and Ellis et al. (2006), we have found that logistic regression is a useful and relatively transparent approach to predict the response of species occurrence as a function of various environmental conditions.

Beside the widely used generalised linear models (GLM; Guisan et al., 1998; Wisz and Guisan, 2009), other techniques employed in order to unravel the complexity of interactions between distribution and environmental factors are generalised additive models (GAM; Yee and Mitchell, 1991), classification and regression trees (CART; Moore et al. 1991), artificial neural networks (ANN; Fitzgerald and Lees, 1992) and multivariate adaptive regression splines (MARS; Friedman, 1991). The generation of 'potential habitat distribution maps' is stated to be among the predictive modelling goals (Munoz and Felicisimo, 2004), therefore the convenience of cartographic implementation is crucial. Such properties are mainly assigned to generalised methods as well as to the MARS method, that builds complex regression models by fitting piecewise linear regressions. A successful development of habitat models includes both geographical and ecological discrimination of species. Since only very few species have been studied in detail in terms of their dynamic responses to environmental changes, static distribution modelling often remains the only approach for studying the possible consequences (Woodward and Cramer, 1996). For the regions where the distribution is strongly and directly coupled to physiochemical processes statistical models are also capable of satisfactorily predicting the species distribution (Ellis et al., 2006). This is the case for our study area as was concluded in Gogina et al. (2009—this volume). Among the objectives we had for selecting the relatively simple and intuitive GLM approach was its ability to construct a parsimonious model that strikes a balance between bias and variance; identify the relative importance of the predictor variables; explore and interpret the response of the species to each predictor; estimate the uncertainty associated with parameter estimates; predict the probability of observing the species (rather than predicting binary presence–absence) and explore spatially explicit patterns of uncertainty in predictions. Decisive was its availability for direct interpretation of the results, accompanied by its reputation for providing a competitively high accuracy (e.g. Praca et al., 2008; Wisz and Guisan, 2009). The comparable method that employed logistic regressions together with weighted overlay was also successfully applied and

Table 2
Maximum densities and frequencies of occurrences of selected species within the data set.

Species	Maximum density, ind/m^2	Frequency of occurence, %
A. islandica	622	52.4
H. diversicolor	667	20.7
P. elegans	11459	57.2
S. armiger	1684	77.9
T. benedii	1469	44.7

Table 3
The maximum-likelihood estimates of logistic regression parameters for response surfaces of selected species to each of abiotic factors and normalized Akaike's weights derived for the model set including single-factor model of total organic content as predictor and the model set without it.

Parameter estimates

		A. islandica	*H. diversicolor*	*P. elegans*	*S. armiger*	*T. benedii*
d	b_0	−20.1829	7.5037	5.9871	3.6877	−6.0225
	b_1	1.9694	−0.5584	−0.2993	−0.1247	0.9314
	b_2	−0.0449	–	–	–	−0.03179
	p(model)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
s	b_0	−2.8957	3.3046	5.0447	6.6534	5.5642
	b_1	0.1868	−0.3310	−0.3101	−0.3215	−0.3635
	b_2	–	–	–	–	–
	p(model)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
o2	b_0	0.5838	−3.9298	0.0453	4.1259	−1.1231
	b_1	−0.0904	0.7660	0.0183	−1.1573	0.1690
	b_2	–	−0.0520	–	0.1029	–
	p(model)	0.1149**	0.0946**	0.7473**	0.0192	0.0085
org	b_0	−0.6234*	−0.3851	0.6887	2.1783	1.2297
	b_1	2.3222*	−3.6250	−0.5053	−0.2814	−0.7844
	b_2	−0.7467*	–	–	–	–
	p(model)	0.0161	<0.0001	<0.0001	<0.0001	<0.0001
kg	b_0	0.7970	−5.3292	−2.7251	−0.3493	−2.5964
	b_1	−0.0023	0.0207	0.0288	0.0087	0.0177
	b_2	–	−0.00003	−0.00004	–	−0.00002
	p(model)	0.0647**	0.0215	<0.0001	<0.0001	<0.0001
so	b_0	1.4313	–	2.2991	2.8781	1.0022
	b_1	-0.4871	–	-2.8793	-2.0383	-1.4740
	b_2	–	–	–	–	–
	p(model)	0.2075	–	<0.0001	<0.0001	0.0008
sk	b_0	0.7319	–	0.1435	0.9031	−0.2341
	b_1	−0.3262	–	3.7708	1.9679	1.6366
	b_2	–	–	–	–	–
	p(model)	0.6379**	–	<0.0001	0.0043	0.0285
kKM	b_0	0.5043	–	−2.7986	−1.0221	−2.0592
	b_1	0.1260	–	0.3755	0.5344	0.2287
	b_2	-0.0031	–	-0.0068	-0.0118	-0.0036
	p(model)	0.0300	–	<0.0001	<0.0001	0.0128

Akaike's weights estimated for two model sets

	A. islandica		*H. diversicolor*		*P. elegans*		*S. armiger*		*T. benedeni*	
Set	Full	No org	Full	No org	Full	No org	Full	No org	Full	No org
d	0.7591	0.7669	0.8646	0.9939	0.4431	0.9885	0.0000	0.0000	0.9902	0.9999
s	0.0188	0.0190	0.0003	0.0003	0.0000	0.0001	0.5861	0.9450	0.0001	0.0001
o2	0.0517	0.0522	0.0046	0.0053	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
kg	0.0197	0.0199	0.0004	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
so	0.0411	0.0415			0.0025	0.0056	0.0341	0.0549	0.0000	0.0000
sk	0.0208	0.0210			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
kKM	0.0786	0.0794			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
org	0.0101		0.1301		0.5543		0.3798		0.0097	

Concordance of merged model results, %

	A. islandica	*H. diversicolor*	*P. elegans*	*S. armiger*	*T. benedeni*
	82.86	94.59	81.24	75.78	77.75

The environmental factors notation is as follows: org—for total organic content, kg—for median grain size, o2—for oxygen concentration, s—for salinity, d—for water depth, so—for sorting, sk—for skewness, and kKM—for permeability.

* Estimated for log-transformed total organic content variable, that indicated a better model fit.

** No significant relation of species occurrence probability to the factor could be derived.

tested on a similar spatial scale by Glockzin and Zettler (2008b), who modelled habitat suitability maps for exemplary species of the Pomeranian Bay. In their work they emphasised not only the importance of spatial scale, but in-depth knowledge of species autecology used in such studies and models. This is especially important since top-down correlations between the macrobenthos and the environmental variables represented by response models do not always reflect direct cause and effect relationships, since many environmental variables co-vary (van der Wal et al., 2008).

It is essential to differentiate between the questions that the two types of model solutions performed serve to answer. Each single-factor model alone is able to answer the question of how certain species respond to changes of this separate factor, e.g. describe it as euryoecious or stenoecious organism. Thrush et al. (2005) acknowledge that simple models may fit well the purposes of response management, for instance by defining the sufficiently consistent general pattern of which species prefer muddy, intermediate or sandy sediment types across scales, so that a rank order of species can be developed. Such classification can be used to interpret changes in the distribution and abundance of monitored species or incorporated with other predictions of habitat change, and used in environmental risk assessment. The multimodel inference serves to predict the distribution of species within the common limits of combined environmental factors, e.g. to fill out the lack of information in the locations, where no observational data is available. Thus among the applications of statistical modelling exercises such as the one presented in this paper are: testing the hypotheses about the ranges of species distribution along environmental gradients and benthic stress impacts (for instance, the Pearson and Rosenberg (1978) model that reported a gradual loss of species as the degree of stress increased over space and/or time driven by the replacement pattern

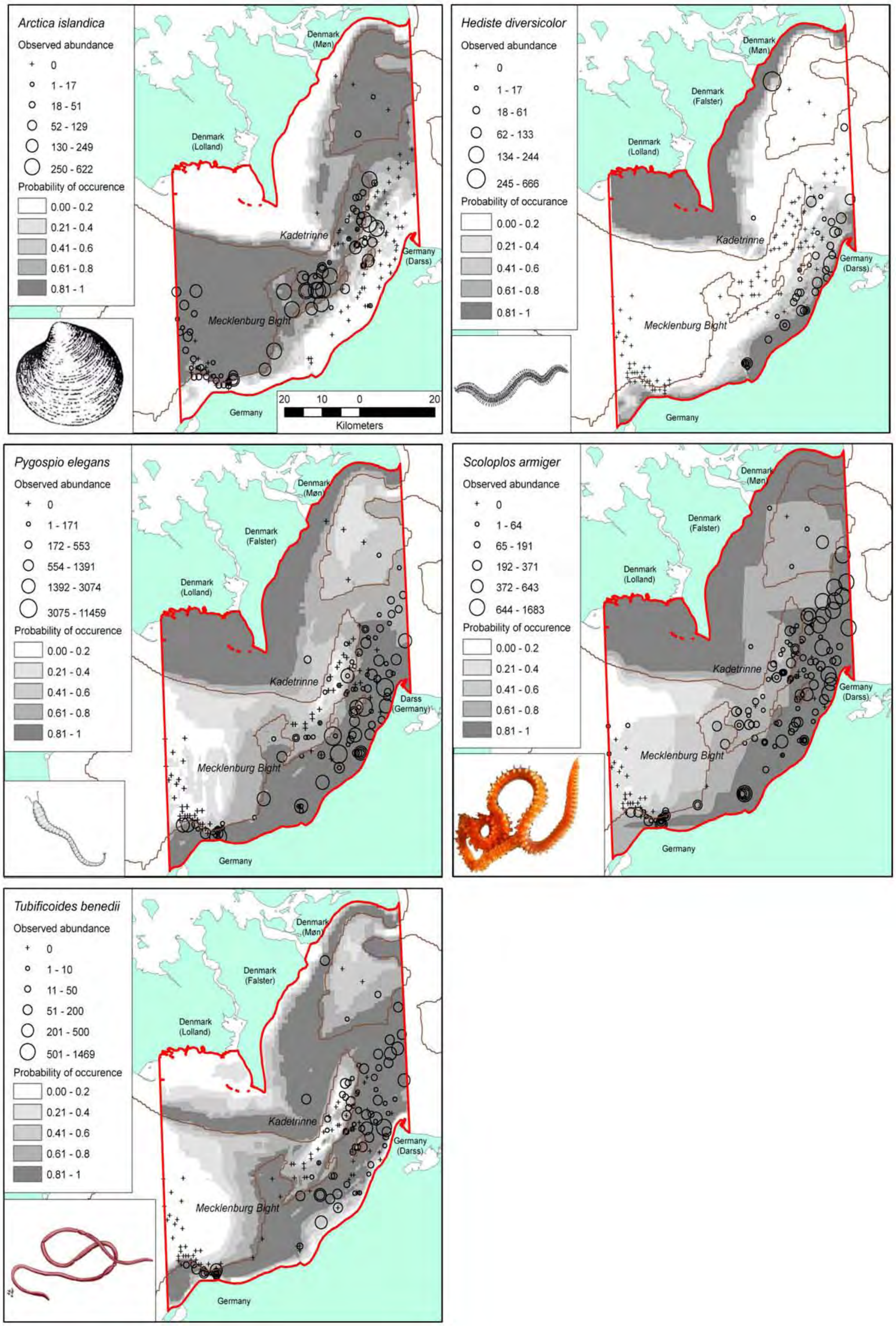


Fig. 7. Modelled probability of species occurrence determined using Akaike's weghts for multimodel inference. Circles indicate the observed distribution with their size corresponding to the value of abundance density (abundance, ind/m^2). The intensity of shading corresponds to increasing probability of occurrence.

defined by different tolerance of species to stress); generalization of habitat suitability maps that predict the specific ecological potential of a habitat (with limitations defined by the data analyzed) which can be considered in marine spatial planning and conservation management (Degraer et al., 2008); and predicting the possible consequences of habitat changes (either natural or antropogenic). When accompanied by other relevant developments and investigations possible use can be found in the comparison the of species' spatial distribution at different scales (e.g. Thrush et al., 2005). 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).

Considerable differences of the importance of various environmental conditions were highlighted by the prediction maps for selected species (Fig. 7). All obtained response surfaces indicated a relatively high percentage of concordance, though a more robust validation of derived models is still to be executed and will unambiguously require external data. The essential autecological features found in literature sources and evidences derived from modelling results are summed up in Table 4.

The discussion above illustrates, that to distinguish between colonization types, prior biological knowledge must support the statistical examination. Bonsdorff (2006) states the "ecological age" of the present Baltic Sea ecosystem to be only about 8000 years, resulting in still ongoing primary succession processes and numerous ecological niches remaining available for immigration. This seems to be evidenced by the fact, that

Table 4
Essential autecological features of exemplary species and its evidences based on modelling results (Fig. 7).

Species	Essential autoecological features	Evidence from the modelling results
A. islandica[a]	Habitat: arctic-boreal bivalve, found in high concentrations at 25–61 m (recorded at 8 to 256 m), in the Baltic Sea found in high saline areas (at depths of 16 to 30 m, eastern distribution limit—Arkona basin). Substrate: firm sediments, medium to fine grain sand, sandy mud, silty sand. Oxygen: resistant to oxygen deficiency; can burrow into substrate and respire anaerobically for up to week. Salinity: found at oceanic salinities, in the lab resists salinities as low as 22 psu. Feeding mode: suspension feeder Notes: occur in dense beds over level bottoms. Extremely long life-span (ca. 80 years in the Baltic and about 405 in the Atlantic). Strong recruitment of the species in the Mecklenburg Bight during the past decades and probably the displacement of recruitment zone from below 20 to 15–20 m depth.	Confirm: wide ranges of organic content and sediment parameters, high values of depth (>18 m) and salinity define suitable habitat. Very unlikely to occur in the shallow coastal zone, influenced by freshwater runoff. Highest densities found at depth between 15 and 20 m, whereas highest probabilities of occurrence coincide with the regions slightly below 20 m. As densities of species are known to correlate negatively with size/age this confirms the displacement of recruitment zone. Not recorded in the region at salinities below 10 psu, highest densities found between 14 and 16 psu (much lower than oceanic salinity), but occurrence increases with salinity. The preference of dense beds reflected in low likelihood of occurrence on coarse sediments and at high permeability values. Oxygen concentration had no strong effect.
H. diversicolor[b]	Habitat: inhabits shallow marine and brackish waters. In the Baltic Sea prefers enclosed bays and lagoons with smooth slopes and absence of strong currents. Substrate: sandy mud but also gravels, clays, even turf. Oxygen: able to survive drastic conditions of hypoxia. Salinity: euryhaline, lowest limit of salinity (determined through its larvae development) of about 5 psu. Feeding mode: carnivore, scavenger, filter feeder and a surface deposit feeder, however also having the ability to live as a suspension feeder Notes: generally endobethic species, able to penetrate the substrate up to depths of about 30 cm.	Confirm: does not find its optima within the study area (greatly higher densities are recorded in the enclosed coastal estuarine regions of the Baltic Sea). Preference of low-saline regions shallower than 18 m within the investigated spatial and environmental limits. Among considered abiotic factors the largest effect size featured water depth, followed by total organic content. Species also responded here noticeably to oxygen, grain size and salinity, but with relatively low weights in the model averaging inference. Its opportunistic abilities seem to allow species to avoid locations where it may be disturbed by other competitive species.
P. elegans[c]	Habitat: mainly lower superlittoral down to depth of 100 m, in the Baltic Sea from ca 5 m down to 20 m. Substrate: fine to medium sands Oxygen: highly sensitive to hypoxia, hardly ever found in areas affected by the phenomena Salinity: from 2 psu to hypersaline pools Feeding mode: deposit and filter feeder Notes: features the penetration depth of 4–6 cm. Evolves opportunistic life strategies: the species is able to rapidly re-colonize defaunated substrates. Regarded as a semi-sedentary species. Avoids strong currents. Negative response to organic enrichment.	Confirm: the model with highest weight in the model averaging was the function of total organic content. Response indicates almost no chance of finding the species in sediments with organic content exceeding 5%. Water depth (containing the effect of currents) indicated slightly smaller but comparable influence on species distribution. Typically favors shallower waters, but the response against increasing water depth is not as rapid as in case of *H. diversicolor*. Surprisingly, only sorting had non-zero (yet low) model averaging weight among sediment factors, perhaps due to the limiting power of covarying organic content.
S. armiger[d]	Habitat: cosmopolitan species with intertidal as well as in the subtidal occurrence encountered in all zoogeographic regions. In the Baltic Sea, e.g. in the Mecklenburg Bight, species habitat is limited from 5 to ca. 30 m, with highest abundance values found around 10 m depth and absence in purely silt regions. Substrate: muddy sands and mud Oxygen: resistant to hypoxia down to 0.5 ml/l Salinity: cannot survive at salinities lower than 10.5 psu Feeding mode: deposit feeder Notes: mobile non-selective species burrowing freely through sediments, that builds non-persisting tubes in the sediment down to 15 cm.	Confirm: among the considered abiotic factors only salinity, total organic content and, to small extend, sorting explained the distribution pattern of *S. armiger* in multimodel inference. Bleidorn et al. (2006) revealed that *S. armiger* represents a species complex and is not a cosmopolitan species that explains the inconsistency between general environmental setting found in literature and its observed and predicted distribution limits within the study area. Thus, there is a negative response along increase of both salinity and organic content in terms of both occurrence and abundance within the region (though observed salinities have the lower limit of 8 psu).
T. benedii[e]	Habitat: ubiquitous marine oligochaete that dominates in coastal areas. It is often typified as 'opportunist' that is adapted to the rapid environmental fluctuations and harsh conditions in estuaries. Substrate: fine organic- enriched sediments Oxygen: resistant to hypoxia Salinity: found in meso- to euhaline waters. Feeding mode: deposit feeder Notes: very successful adaptive strategies in sulfidic benthic environments. Able to penetrate into the substrate up to 10 cm.	Confirm: the prediction map for this species is derived nearly exclusively from water depth as a proxy, with highest probability of its occurrence around 10 to 20 m, and solitary occurrence events below 23 m. In the investigation area this species never seems to settle on the truly mud substrate and does not favor high organic contents as such conditions here coincide with undesirable water depth (that probably comprises a range of other limiting factors for this animal).

[a] Cargnelli et al., 1999; Zetter and Röhner, 2004; Wanamaker et al., 2008.

[b] Nithart et al., 1998; Scaps, 2002; Kristensen, 2001; Henning et al., 2004; Zetter and Röhner, 2004; van der Wal et al. 2008.

[c] Fauchald and Jumars, 1979; Morgan et al., 1999; van der Wal et al. 2008.

[d] Fauchald and Jumars, 1979; Zettler et al., 2000; Bleidorn et al. 2006.

[e] Dubilier et al., 1994; Giere, 2006.

most of the representative species selected for predictive modelling in the study area are named as opportunists.

The presented models should be best applicable for species that find their optima within the observed ranges of environmental factors. *Inter alia*, such models can also reflect the tendencies of an opportunistic species distribution with satisfactory result, e.g. *H. diversicolor* that reaches only low abundances within the study area compared to its average density in more favourable in-shore estuarine conditions. For such species the modelling results reveal the best of available conditions in contrast to their optima. However, it should be mentioned that, as for all statistical methods, the inference about the model selection uncertainty is conditional on both data and the full set of a priori models considered (Burnham and Anderson, 2004). Coudun and Gégout (2006) suggest a general minimum value of 50 occurrences for species to derive acceptable ecological response curves with logistic regression. The available data for *H. diversicolor* slightly drops out of this condition with 43 cases of occurrence within total 208 observations.

A high variance in the species distribution that usually causes an impediment for modelling, is often not a sampling error or random “noise”, but rather the mechanistic consequence of shifts between limiting resources or other effects and factors (e.g. intra- and inter-species competition, predation, mortality or dispersal). The abundance of species may be very low under favourable conditions if, for some reason, the number of propagule is very low or species never even reach a given area. This natural phenomena of a species failing to colonize all areas where it could potentially thrive (Huston, 2002) may partly explain some inconsistencies between the predicted high probability of species occurrence and its factual absence according to observations. Under optimal conditions species can reach maximal reproduction rates and maximal abundances, but macrobenthic surveys often show an entirely different reality: species and communities are distributed rather patchily and often the relatively smooth structure of abiotic gradients and other characteristics can increase statistical uncertainty and blur the picture.

5. Conclusions and outlook

For all exemplary species the used technique provided the results of distribution prediction based on environmental data with a fairly satisfactory accuracy. The method combining the parsimony of single-factor logistic regression models with an AIC solution of multimodel inference can therefore be recommended for modelling of the habitat suitability for benthic species, at least on comparable spatial scales and environmental gradients.

It is important to point out that only a fragment of the whole factor network responsible for the macrozoobenthos distribution was covered in this study. Consideration of additional variables may reveal more insightful analysis. Results of the present study may only be interpreted within the context of the momentary state of benthic communities.

In the present study only the probability of benthic species occurrence based on presence/absence data was modelled, allowing the prediction of species distribution derived from environmental data. However, the density of species is a considerably more informative figure. Modelling of benthic species abundance as a function of abiotic parameters is likely to be performed applying ‘factor ceiling’ (Thrush et al., 2003) or quantile regression (Cade and Noon, 2003) methods.

Acknowledgments

Part of this work was funded by the EU-MARBEF. We gratefully acknowledge the work of all colleagues involved and responsible for field work and laboratory analysis, in particular Ines Glockzin, the valuable assistance on mathematical, statistical and GIS matters of Dr. Bernd Bobertz and Dr. Michael Meyer. We also thank Dr. Torsten Seifert for his great help with modelled data on hydrographic parameters. The generation of this data was partly supported by BMBF within the framework of Project IKZM-Oder (03F0403A) and by a DFG-grant NE G17/3-1. We are very grateful to Prof. Peter Herman, for his critical, essential and helpful review. We express our thanks to Irina Steinberg, Anastasia Semenova and Gunda Wieczorek for correcting the English.

References

Austin, M.P., 1987. Models for the analysis of species' response to environmental gradients. Vegetatio 69, 35–45.

Blackburn, T., Lawton, J.H., Perry, J.N., 1992. A method of estimating the slope of upper bounds of plots on body size and abundance in natural animal assemblages. Oikos 65, 107–112.

Bleidorn, C., Kruse, I., Albrecht, S., Bartolomaeus, T., 2006. Mitochondrial sequence data expose the putative cosmopolitan polychaete *Scoloplos armiger* (Annelida, Orbiniidae) as a species complex. BMC Evol. Biol. 6, 47.

Bobertz, B., Harff, J., 2004. Sediment facies and hydrodynamic setting: a study in the south western Baltic Sea. Ocean. Dyn. 54, 39–48.

Bobertz, B., Harff, J., Bohling, B., 2009. Parameterisation of clastic sediments including benthic structures. J. Mar. Syst. 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.

Bourget, E., Fortin, M.-J., 1995. A commentary on current approaches in the aquatic sciences. Hydrobiologia 300 (301), 1–16.

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., Noon, B.R., 2003. A gentle introduction to quantile regression for ecologists. Front. Ecol. Environ. 1, 412–420.

Cargnelli, L.M., Griesbach, S.J., Packer, D.B., Weissberger, E., 1999. Ocean quahog Arctica islandica, life history and habitat characteristics. NOAA Tech Memo NMFS-NE-148.

Clarke, K.R., Warwick, R.M., 1994. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. Plymouth Marine Laboratory, Plymouth, UK.

Coudun, Ch., Gégout, J.-C., 2006. The derivation of species response curves with Gaussian logistic regression is sensitive to sampling intensity and curve characteristics. Ecol. Model. 199, 164–175.

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.

Dubilier, N., Giere, O., Grieshaber, M.K., 1994. Concomitant effects of sulfide and hypoxia on the aerobic metabolism of 241 the marine oligochaete *Tubificoides benedii*. J. Exp. Zool. 269, 287–297.

Ellis, J., Ysebaert, T., Hume, T., Norkko, A., Bult, T., Herman, P., Thrush, S., Oldman, J., 2006. Predicting macrofaunal species distribution in estuarine gradients using logistic regression and classification systems. Mar. Ecol. Prog. Ser. 316, 69–83.

ESRI, 2003. ESRI Data & 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.

Fitzgerald, R.W., Lees, B.G., 1992. The application of neural networks to the floristic classification of remote sensing and GIS data in complex terrain, in: American Society of Photogrametry and Remote Sensing (Eds.), Proceedings of the XVII Congress ASPRS, Bethesda, MD, pp. 570–573.

Friedman, J.H., 1991. Multivariate adaptive regression splines. Ann. Stat. 19, 1–141.

Giere, O., 2006. Ecology and biology of marine Oligochaeta—an inventory rather than another review. Hydrobiologia 564, 103–116.

Glockzin, M., Zettler, M.L., 2008a. Spatial macrozoobenthic distribution patterns in relation to major environmental factors—a case study from the Pomeranian Bay (southern Baltic Sea). J. Sea Res. 59, 144–161.

Glockzin, M., Zettler, M.L., 2008b. Which factors cause benthic diversity in a low salinity environment? An approach to species response. Book of Abstracts, MARBEF World Congress on Biodiversity, Valencia, Spain, Session 1.10, p. 18.

Gogina, M.A., Zettler, M.L., Glockzin, M., 2009-this volume. Distribution of benthic macrofaunal coenoses in the western Baltic Sea with regard to near-bottom environmental parameters. Part 1. Casual analysis. J. Mar. Syst. doi:10.1016/j.jmarsys.2009.07.006.

Guisan, A., Theurillat, J.-P., Kienast, F., 1998. Predicting the potential distribution of plant species in an alpine environment. J. Veg. Sci. 9, 65–74.

Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modelling of plant species distribution. Plant Ecol. 143, 107–122.

Henning, D., Zauke, G.-P., Stief, P., Vareschi, E., 2004. Integrated effects of marine benthic photosynthetically active microorganisms and a bioturbating polychaete, *Hediste diversicolor* (O. W. Müller), on oxygen distribution at the sediment–water interface: a geostatistical analysis. Proceedings of the Second International Symposium on GIS/Spatial Analyses in Fishery and Aquatic Sciences, Fishery-Aquatic GIS Research Group, Kawagoe-City, Japan.

Huston, M.A., 2002. Critical issues for improving predictions. In: Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B. (Eds.), Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press, Washington, pp. 7–21.

Kristensen, E., 2001. Impact of polychaetes (*Nereis* spp. and *Arenicola marina*) on carbon biogeochemistry in coastal marine sediments. Geochem. Trans. 12.

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, Second English edition. Developments in Environmental Modelling, vol. 20. Elsevier Science B.V., Amsterdam.

Legendre, P., Thrush, S.F., Cummings, V.J., Dayton, P.K., Grant, J., Hewitt, J.E., Hines, A.H., McArdle, B.H, Pridmore, R.D., Schneider, D.C., Turner, S.J., Whitlatch, R.B., Wilkinson, M.R., 1997. Spatial structure of bivalves in a sandflat: scale and generating processes. J. Exp. Mar. Biol. Ecol. 216, 99–128.

McArdle, B.H., Blackwell, R.G., 1989. Measurement of density variability in the bivalve *Chione stutchburyi* using spatial autocorrelation. Mar. Ecol. Prog. Ser. 52, 245–252.

Meissner, K., Darr, A., Rachor, E., 2008. Development of habitat models for *Nephtys* species (Polychaeta: Nephtyidae) in the German Bight (North Sea). J. Sea Res. 60, 271–286.

Meyer, M., Harff, J., Gogina, M., Barthel, A., 2008. Coastline changes of the Darss-Zingst Peninsula—a modelling approach. J. Mar. Syst. 74, 147–154.

Moore, D.M., Lee, B.G., Davey, S.M., 1991. A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. Environ. Manag. 15, 59–71.

Morgan, T.S., Rogers, A.D., Paterson, G.L.J., Hawkins, L.E., Sheader, M., 1999. Evidence for poecilogony in *Pygospio elegans* (Polychaeta: Spionidae). Mar. Ecol. Prog. Ser. 178, 121–132.

Munoz, J., Felicisimo, A., 2004. A comparison between some statistical methods commonly used in predictive modelling. J. Veg. Sci. 15, 285–292.

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.

Nithart, M., Alliot, E., Salen-Picard, C., 1998. Population dynamics and secondary production of Nereis diversicolor in a north Norfolk Saltmarsh (UK). J. Mar. Biol. Assoc. U. K. 79, 1029–1037.

O'Brien, K., Hänninen, J., Kanerva, T., Metsärinne, L., Vuorinen, I., 2003. Macrobenthic zonation in relation to major environmental factors across the Archipelago Sea, northern Baltic Sea. Boreal Environ. Res. 8, 159–170.

Pavlikakis, G.E., Tsihrintzis, V.A., 2000. Ecosystem management: a review of a new concept and methodology. Water Resour. Manag. 14, 257–283.

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

Perus, J., Bonsdorff, E., 2004. Long-term changes in macrozoobenthos in the Åland archipelago, northern Baltic Sea. J. Sea Res. 52, 45–56.

Philippart, C.J.M., Anadón, R., Danovaro, R., Dippner, J.W., Drinkwater, K.F., Hawkins, S.J., Oguz, T., O'Sullivan, G., Reid, P.C., 2007. Impacts of climate change on European marine and coastal environment. ESF Marine Board Position Paper 9.

Praca, E., Gannier, A., Das, K., Laran, S., 2008. Modelling the habitat suitability of cetaceans: example of the sperm whale in the northwestern Mediterranean Sea. Deep-Sea Res. I. doi:10.1016/j.dsr.2008.11.001.

Rhoads, D.C., 1974. Organism sediment relations on the muddy sea floor. Oceanogr. Mar. Biol. Annu. Rev. 12, 263–300.

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

Scaps, P., 2002. A review of the biology, ecology and potential use of the common ragworm *Hediste diversicolor* (O.F. Müller) (Annelida: Polychaeta). Hydrobiologia V470, 203–218.

Seifert, T., Tauber, F., Kayser, B., 2001. A high resolution spherical grid topography of the Baltic Sea, Baltic Sea Science Congress, Stockholm, Poster #147, 2nd edition.

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

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.

van der Wal, D., Herman, P.M.J., Forster, R.M., Ysebaert, T., Rossi, F., Knaeps, E., Plancke, Y.M.G., Ides, S.J., 2008. Distribution and dynamics of intertidal macrobenthos predicted from remote sensing: response to microphytobenthos and environment. Mar. Ecol. Progr. Ser. 367, 57–72.

Wanamaker Jr., A.D., Heinemeier, J., Scourse, J.D., Richardson, C.A., Butler, P.G., Eiríksson, J., Knudsen, K.L., 2008. Very-long lived molluscs confirm 17th century AD tephra-based radiocarbon reservoir ages for north Icelandic shelf waters. Radiocarbon 50, 3.

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.

Yee, T.W., Mitchell, N.D., 1991. Generalized additive models in plant ecology. J. Veg. Sci. 2, 587–602.

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.

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

Zettler, M.L., Bönsch, R., Gosselck, F., 2000. Das Makrozoobenthos der Mecklenburger Bucht—rezent und im historischem Vergleich. Meereswiss. Ber. 42, 1–144.

IV

Baltic Coastal Zone No. 13 part II	
(79-95) 2009	Institute of Biology and Environmental Protection Pomeranian University in Słupsk

BEYOND SALTY REINS – MODELING BENTHIC SPECIES' SPATIAL RESPONSE TO THEIR PHYSICAL ENVIRONMENT IN THE POMERANIAN BAY (SOUTHERN BALTIC SEA)

Michael Glockzin, Mayya Gogina, Michael Lothar Zettler*

** Corresponding author*
Leibniz Institute for Baltic Sea Research (IOW),
Department of Biology, Seestr. 15, D-18119 Warnemuende, Germany,
e-mail: michael.zettler@io-warnemuende.de

Abstract

The brackish water environment of the Baltic Sea is dominated by a strong gradient of salinity and along with salinity the benthic diversity decreases – salinity is regarded as the master factor regulating benthic diversity in brackish habitats. In this scheme, consistently small patches of comparatively higher or lower benthic diversity do emerge in areas where either environmental or anthropogenic impacts on the benthic habitat change drastically over short spatial distances. Hence, spatial diversity of ecological factors creates diversity among benthic colonization and community structures. We show through a logistic modeling approach the possibility to predict thereby induced benthic colonization areas and community structures inside the broad scheme of a brackish water habitat. This study bases upon quantitative macrozoobenthic abundance data collected over a period of 4 years. It clearly demonstrates the need to analyze species' relationships in gradient systems such as the Baltic Sea and provides a tool to predict natural and anthropogenic forced changes in species distribution.

Key words: Baltic Sea, Pomeranian Bay, macrozoobenthos, ecological modeling, environmental factors, multivariate analysis, species' response curves

INTRODUCTION

Across the Baltic Sea, a long and strong gradient of salinity stretches over a distance of about 1100 km, steadily decreasing from the Skagerrak and Kattegat in the West (about 25-30 psu) over the Belt Sea (about 10-25 psu) and Baltic Proper (about 5-10 psu) up to the Baltic's most northerly and north-easterly parts, the Finnish and Bothnian Gulfs (about 0-5 psu) – and with salinity as dispersion barrier for marine species, the benthic diversity as well decreases. Hence, the factor salinity is regarded as the master factor regulating benthic diversity in brackish habitats (Zettler et al. 2007). On the whole, it seems to explain on a large spatial scale benthic diversity patterns examined in the Baltic Sea as well as in other brackish habitats all over the world. However, inside this general picture of decline in diversity, consistently small scale patches of comparatively higher or lower benthic diversity do emerge – namely in areas where either the environmental "texture" (e.g. the sea bottom structure), other environmental factors (e.g. sediment type) or anthropogenic impacts (e.g. man-made pollution) influencing the benthic habitat change drastically over short spatial or temporal distances. Thus, spatial and temporal diversity in ecological factors can create diversity among benthic colonization and community structures. Due to the variability of species in terms of habitat selection, reciprocal effects between species' distribution and environmental factors manifest in patterns, visible in their abundances or assemblages (Keitt et al. 2002, Fortin et al. 2005). Always, benthologists have dreamt of possibilities to forecast such patterns and structures – modern comforts now at hand via computer-aided statistical model development, providing us with a useful tool to relate ecological features to environmental factors. Through validation and modification, it can even reveal the underlying mechanisms responsible for the structure and organization of communities (Austin 1987). However, an exploratory statistical description of the prevailing ecological structure based on observations always is a necessity – not only to examine and understand the structure and dynamics of benthic biotic/environmental interactions and processes but to evaluate and model natural and anthropogenic influences and effects on ecological systems (Bourget and Fortin 1995, Legendre and Legendre 1998). Furthermore, in-depth knowledge is indispensible regarding the autecology of the species for the interpretation of the found relationships (Sachs 1997). On this account, we described patterns in the distribution of benthic organisms and determined the parameters causing such patterns via a multivariate statistical approach, thus establishing a statistical link between benthic infauna distribution and environmental factor patterns for the German part of the Pomeranian Bay in a preliminary study (Glockzin and Zettler 2008). There, we already announced the possibility of "[…] modeling species response curves on the basis of this study, using an appropriate model […] and using the found equations in a GIS-based approach shall enable us to create a two-dimensional ecological model of the Pomeranian Bay and to predict species assemblages [...] by two-dimensional morphological, geological or hydrological data sets". On this account, we utilize in this present study species' response towards responsible environmental key factors to model spatial distributions for selected benthic species via a binomial logistic regression approach (Trush et al. 2003) in a GIS based environment (ArcGIS, ESRI Inc.).

STUDY AREA, MATERIALS AND METHODS

Proper choice of a spatial scale for our investigation was important – to bring into focus the prevailing biotic-abiotic interactions and to diminish or even eliminate the influence of an all-dominating factor like salinity in brackish habitats (Fig. 1). Therefore, we chose the Pomeranian Bay as an adequate "test site" for the methods used in this study. This work bases upon environmental and quantitative macrozoo-

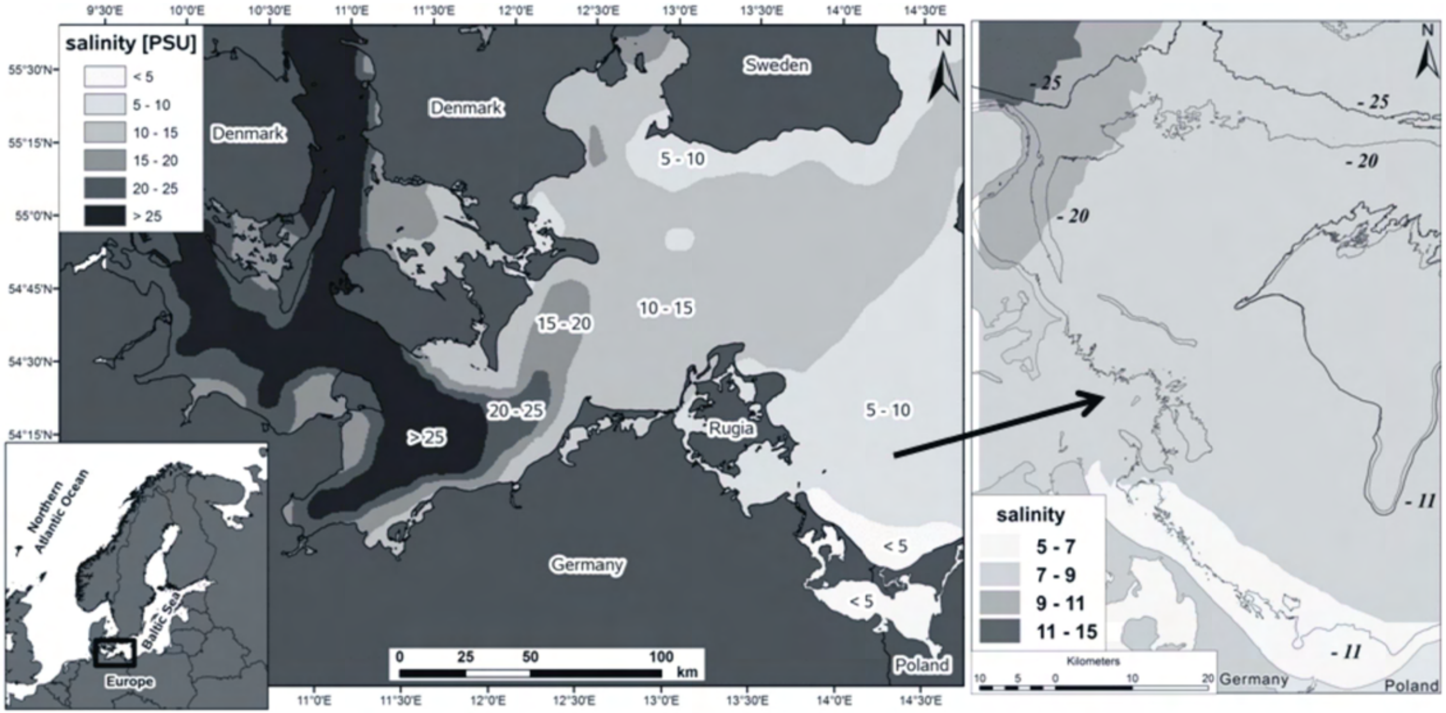


Fig. 1. Left: spatial distribution of the environmental factor near-bottom salinity in the Baltic Sea. Right: spatial distribution of the environmental factor near-bottom salinity in the Pomeranian Bight (study area). Numbers and isolines indicate the water depth (e.g. -20-20 m water depth)

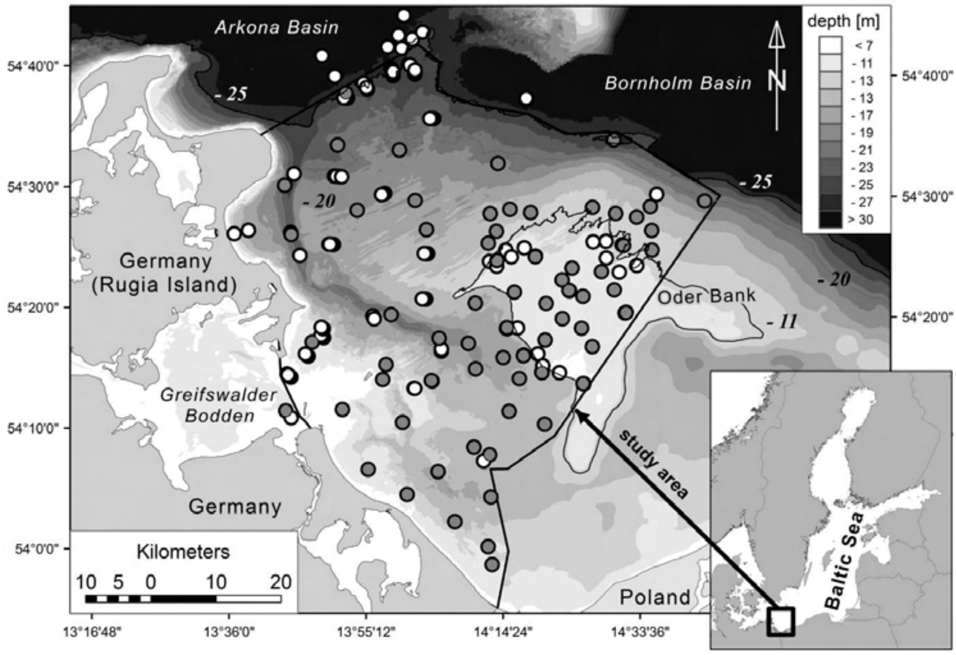


Fig. 2. Study area and distribution of the 191 sampling stations, filled circles indicate stations with a full set of data available for all six abiotic variables. Sampled stations per year: 2003 (53 stations), 2004 (78 stations), 2005 (30 stations), 2006 (30 stations)

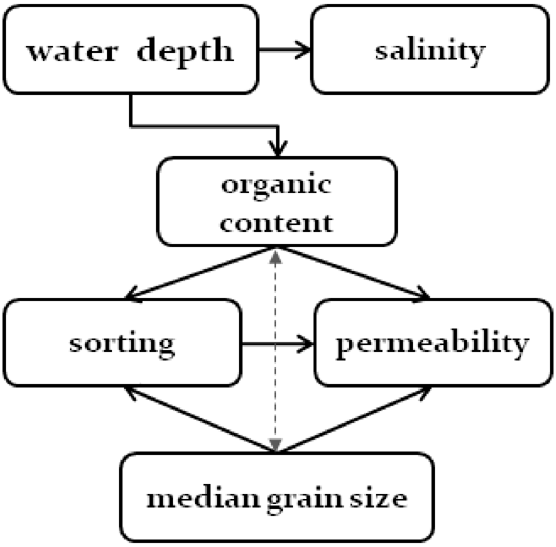


Fig. 3. Environmental framework prevailing in the study area and inter-factor relationships derived from path model analysis with partial correlations (Legendre and Legendre 1998). The dotted line indicates a two-way connection added due to causal interpretation

benthic abundance data collected at 191 stations in the German part of the Pomeranian Bay (Southern Baltic Sea) over a period of 4 years (2003-2006, Fig. 2). A detailed description of the study area, sampling procedures, sample analysis as well as factor and species exclusion/choice would go beyond the scope of this paper and can be found in Glockzin and Zettler (2008). In order to evaluate existing coherencies among environmental parameters, and to assess the primary descriptor predicting all other environmental variables in the Pomeranian Bay, a partial correlation analysis (path analysis) according to the methods described in (Legendre and Legendre 1998) was already conducted in Glockzin and Zettler (2008). Figure 3 illustrates the underlying interactions among environmental factors prevailing in the Pomeranian Bight. Taking a closer look on the prevailing environmental conditions in the study area is indispensible to understand the physical framework of the Bight and to illuminate the later described selection of environmental factors for the modeling approach. Due to prevailing wind, wave conditions and saline water inflows from deeper areas of the Bay and fresh water from the coastal river runoff no homogeneous near bottom salinity zonation develops. Deeper areas of the southern Baltic as Sassnitzrinne, Arkona and Bornholm Basin are the main accumulation areas for fine particles with input of high organic material from the Pomeranian Bay (Kuhrts et al. 2006). This explains the apparent correlation between salinity and organic content, i.e. with increasing water depths and equally rising salinity, the content of organic carbon in the sediment also increases. With increasing amounts of fine particles at the sea bottom, permeability of the sediment in turn decreases rapidly because of the blocking of the interstice (Forster et al. 2003). Due to the glacifluviatile and glaciaeolian genesis of the Bay, median grain size and sorting are merely modified by hydrography, but not created. To highlight the physical and physiographical control of water depth on most of the environmental conditions prevailing in the Pomeranian Bay, a detailed bathymetric structure for the study area is also given in Figure 2. The species' response towards this environmental framework was modeled in terms of probability of occurrence via a generalized linear modeling approach (GLM), using species binomial data (abundance data reduced to presence/absence) in logistic regression with corresponding (logit) link functions (Guisan et al. 1999, Thrush et al. 2003, Gogina et al. 2010) available in STATISTICA 6 (StatSoft Inc.). Therefore, a biotic matrix of three exemplary species (*Bathyporeia pilosa, Tubificoides benedii* and *Hediste diversicolor*) and an abiotic matrix of six environmental parameters (water depth, salinity, median grain size, organic content, permeability and sorting), estimated for 78 out of 191 sampling stations, were used. For the six environmental factors, measuring units together with lower and upper boundaries are given (Tab. 1). Water depth influences through its physical and physiographical forcing almost all other environmental parameters in the study area, it can act as a "proxy", a type of integral parameter combining the effects of various available (measured) or unavailable (not measured or incapable of measurement) habitat features. But its influence can be spatially inhomogeneous or simply superimposed by other factors and therefore warp the outcome of the analysis. With almost all other factors being strongly dependant on it and keeping in mind a certain susceptibility of models to such redundancy; we only examined it in the model building procedure but excluded it as

Table 1

Measuring units together with lower and upper boundaries for the six environmental factors used in the modelling approach

Factor	Factor range	Unit
Water depth	4.4-35	[m]
Salinity	5.7-15.4	[psu]
Median grain size	80-348	[µm]
Organic content	0.12-9.31	[%]
Sorting	0.29-1.40	[no unit]
Permeability	0.4-4.5	[10^{-5} cm s^{-1}]

a factor from further analysis. Regarding the aforementioned physical behavior of salinity, this factor was also excluded. Thus, for model and habitat map building we regarded species probabilities of occurrence calculated as functions of only four factors (organic content, median grain size, sorting and permeability). To find the best combination from the set of single-factor models we considered the philosophy described in Burnham and Anderson (2004). Here, consideration of maximum likelihood as a possible approach to model selection uncertainty is discussed. Hence, the predictor set with the highest amount of likelihood by means of maximum likelihood scores was selected as the set of factors for further analysis. For each factor and taxon combination, estimates of significance from Wald statistic (STATISTICA) were used to estimate the single model significance, with a significance level defined at 0.05. Then, the single-factor model for each species that explained most of the observed variability and fit best to the data was selected. The same maximum likelihood scores, but calculated for the single predictors, were used in further analysis as scaled weights in weighted overlay of probability maps to create habitat suitability maps. The habitat suitability maps were generated with the ArcView spatial analyst (raster calculator), implemented in the Geographic Information System ArcGIS (ESRI Inc.). The factor weights for each species and predictor were scaled, assigning a value of “1” to the factor with the highest maximum likelihood score (Tab. 2). The underlying mechanism of the whole modeling approach is given in Figure 4.

Table 2

Scaled weights used in GIS for a weighted overlay for three species and four factors

Species	*Bathyporeia pilosa*	*Tubificoides benedii*	*Hediste diversicolor*
Organic content	1.00	1.00	1.00
Median grain size	0.78	0.39	0.46
Sorting	0.77	0.65	0.20
Permeability	0.70	0.72	0.62

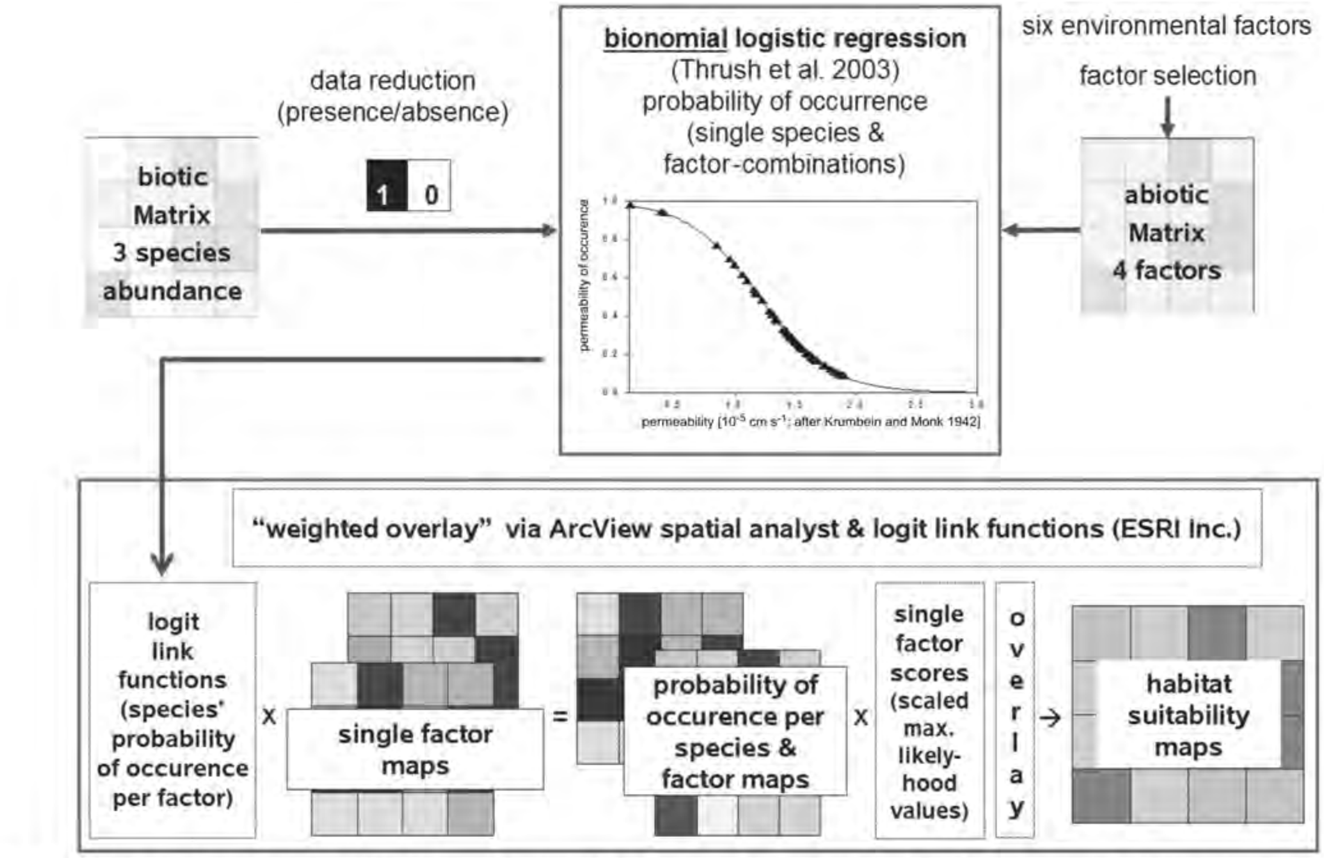


Fig. 4. Flowchart of the modeling approach to model spatial distributions for three selected benthic species and six environmental factors via species' response towards responsible environmental key factors a binomial logistic regression approach in a GIS based environment (ArcGIS, ESRI Inc.)

RESULTS

Figure 5 shows the probability of species occurrence as well as data points and fitted logit curves for three species and each of the four environmental factors used for the generation of probability maps. They explained best the variability of the benthic fauna distribution and were used in the last step of weighted overlay procedure. Here, the organic content of the sediment shows the most dramatic effect on *Bathyporeia pilosa*. The response curve estimated for this species indicates almost no chance of finding it in sediments with organic content exceeding 0.5%. In contrast, *Tubificoides benedii* seems to benefit strongly from organic enriched sediments. For this species, sediments with too low organic content ($< 0.5\%$) seem to be neglected or even avoided. Yet, *Hediste diversicolor* does not discriminate too much between sediments with different organic content. Though its probability of occurrence decreases with a rising organic load, it can occur over the whole range of this factor. Good sorted sediments (i.e. with low values of sorting) appear to favor the occurrence of *B. pilosa* to the point of vanishing where sediments are poorly sorted, whereas such sediments seem to advantage *T. benedii* and *H. diversicolor*. Occurrence of these species can be suppressed strongly but not entirely by this factor for they can most probably exist over its entire range. A different picture can be seen for the probability of occurrence regarding the factor permeability. Here, a rising probability of occurrence of *B. pilosa* and *H. diversicolor* seems to be associated strongly with a rising permeability of sediments, whereas the likelihood of occurrence of *T. benedii* falls close to zero at sediments more permeable than $2.5\text{x}10^{-5}$ cm s^{-1}. In any case, permeability below $0.3\text{x}10^{-5}$ cm s^{-1} seems to be an exclusion criterion for *H. diversicolor*. *B. pilosa* and *T. benedii* are both affected by median grain size – but with opposing trends. While the probability of occurrence of *B. pilosa* rises with decreasing grain size it is almost exactly the other way round with *T. benedii*. The occurrence of *H. diversicolor* seems to be almost unaffected by sediments median grain size. Though probability of occurrence drops to low values all three species can occur with the utmost probability over the whole factor range in the study area.

Four habitat suitability maps, derived from the weighted overlay of probability maps calculated via species probability of occurrence, corresponding to four single environmental factors (Fig. 5), are shown in Figure 6. Additionally to modeled probability surfaces, sampled species abundance is given for evaluation of modeling success in each map. For *B. pilosa*, two different maps were created: Figure 6 (upper left) shows the modeling result for all four environmental factors; Figure 6 (upper right) illustrates the spatial probability for this species when only organic content of sediments was used as predictor. In both cases, the high probability of species occurrence covers the shallow areas of the German Part of the Bay to the extend of certain water depths, the Oder Bank, and a small area between Greifswalder Bodden Sill and Oder Bank, and it is very unlikely to observe *B. pilosa* in the Greifswalder Bodden Sill itself. Figure 6 (lower left) shows the modeled probability-of-occurrence-surface of *T. benedii*. As for this map, this species mostly occurs in the deeper parts of the Pomeranian Bay along the North-South axis of the Pomeranian Bay and in areas around the Greifswalder Bodden Sill. The probability-of-occurrence surface of

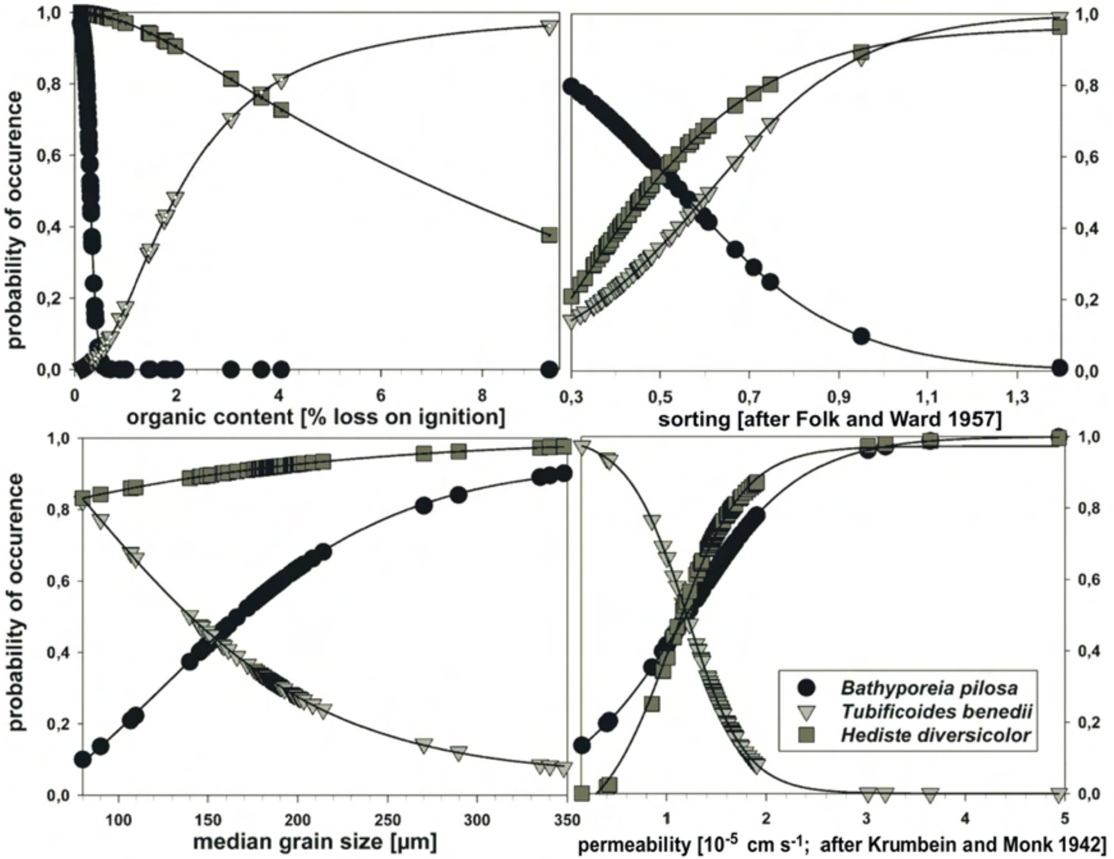


Fig. 5. Probability of species occurrence with data points and fitted logistic regression curves calculated from presence/absence data via logit-function for three species and each of the four environmental factors used for the generation of probability maps

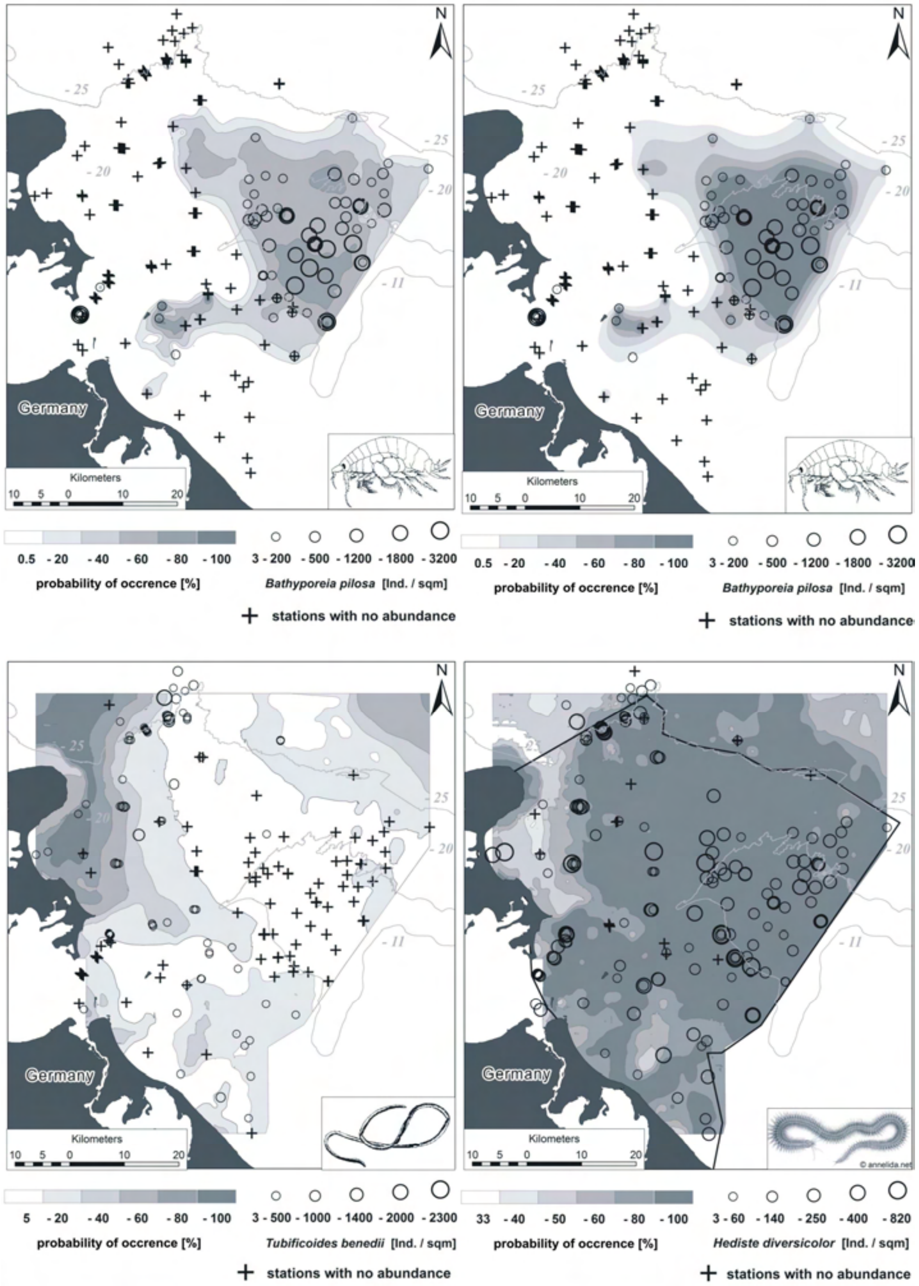


Fig. 6. Four maps for three species, derived from the weighted overlay of probability maps calculated via species probability of occurrence, corresponding to four single environmental factors (Fig. 5). Habitat suitability map for: *B. pilosa* vs. four environmental factors (upper left), *B. pilosa* vs. organic content of sediment only (upper right), *T. benedii* vs. four environmental factors (lower left), *H. diversicolor* vs. four environmental factors (lower right)

the third species, *H. diversicolor*, is shown in Figure 6 (lower right). Almost all areas of the Bay can be inhabited by *H. diversicolor* with high probability, except for the deeper parts along the Sassnitzrinne up to the Arkona Basin.

DISCUSSION

The present study was the first attempt towards a benthic – abiotic interaction model using a complex data set of recent investigations in the German part of the Pomeranian Bay. Such model allows making predictions on how and to what extent natural or anthropogenic influences affect benthic community assemblages not only in the Pomeranian Bay but in other areas of the Baltic Sea (e.g. Gogina et al. 2010). It can be a useful tool in marine ecosystem management (habitat mapping, e.g. Pavlikakis and Tsihrintzis 2000) and environment conservation planning (e.g. to plan the length of a pipeline in marine habitats etc.). However, an exploratory statistical description of the prevailing ecological structure based on in-situ observations is always an indispensible first step (Bourget and Fortin 1995). As in this study, the extraction of patterns of benthic community distributions using large-scale studies is characterized by a large number of data points randomly sampled over long distances and in irregular spatial intervals.

Modeling species spatial distribution via logistic regression and logit function from such data can be regarded as a legacy from plant ecology (e.g. Guisan et al. 1999). Few applications of the method are known in marine science, especially for the Baltic Sea, to date. For example, though not in the Baltic itself, Thrush et al. (2003) investigated response of species to a single environmental factor (sediment mud content) and Ysebaert et al. (2002) performed a comprehensive study, using salinity, depth, flow parameters, median grain size and mud content as predictors. For the regions where distribution is strongly and directly coupled to physiochemical processes statistical models are capable of satisfactorily predicting species distributions (Ellis et al. 2006). To assess the correctness of such calculated biotic – abiotic dependencies, not only geographical and ecological discrimination of species is necessary (Meissner et al. 2008), but also the in-depth knowledge of autecology of the species is indispensible. This expert knowledge stems from different sources: study, utilization of physiological experiments on (benthic) species, observation, (benthic) habitat monitoring as well as experience; it has to be at hand in the first place and cannot be simply deduced from statistics (Sachs 1997). Not till then, statistically modeled descriptions of species distribution patterns can be explained by causality as per ecological behavior of species, e.g. their habitat selection and colonization or the "response type" towards their physical environment (e.g. stenoecious, euryoecious etc.). The question that arises here is whether the modeled results for the three species used in our study mirror their autecology and thus legitimate the use of our modeling approach? *B. pilosa*, a coast-dwelling, sandlicking amphipod which grazes diatoms off the surface of sand grains (Nicolaisen and Kanneworff 1969, Sundbäck and Persson 1981) inhabits all sea-bottom elevations of post-glacial origin in the Pomeranian Bay (see Fig. 6, upper left and right). They consist of well sorted fine sand, deposited and sorted here by postglacial aeolian processes. Wind and wave energy induce currents strong enough to vent these fine grained sediments and to keep them organic-poor through abrasion. This scheme fits perfectly well the described autecology of this species as an inhabitant of good sorted, well oxygenated and organic-depleted sediments (e.g. Speybroeck et al. 2008). In sharp contrast, the euryoecious and meso- to euryhaline deposit feeder *T. benedii* is most abundant in sediments rich in nutrients and organic carbon which form a trail along the North-South axis of the Pomeranian Bay, formed by the deposition of sedimented fine material, discharged by the Oder River in great amounts (ca. 39 kt year-1 total nitrogen, ca. 3.1 kt year-1 total phosphorus, Pastuszak et al. 2005). For this opportunistic oligochaete, mass reproductions in areas with high accumulation rates of organic carbon in sediments are known (Diaz 1984). Such sediments are distributed from the Oder River mouth up to the northwest and along the submerged ancient riverbed of the Oder, the Sassnitzrinne, all the way to the Arkona Basin. The preference of *T. benedii* for impermeable silty sediments seems to originate from his apparent tolerance of hydrogen sulphide occurrence and oxygen deficiency. It can penetrate sediments up to a depth of about 10 cm. It is often typified as well adapted to rapid environmental fluctuations and harsh conditions in estuaries. This advantage makes *T. benedii* one of the most successful inhabitants of ecologically stressed benthic environments with the aforementioned appropriate environmental conditions (Giere et al. 1999, Giere 2006). This oligochaete can also dominate coastal areas with fine sediments highly enriched in organic matter (Dubilier et al. 1994). The environmental preferences of *B. pilosa* and *T. benedii* are very special and this is clearly illustrated by the fact that both habitat suitability maps appear like blanked out by one another. This could be caused by the fact that *B. pilosa* can be regarded as stenoecious to organic content of sediment, causing a "yes or no" response of this species towards this environmental parameter, limiting the possible habitat that can be colonized by this species (Fig. 5, upper left and Fig. 6, upper left). In this respect, *B. pilosa* and *T. benedii* seem to represent some kind of "ecological antagonists" whereas *H. diversicolor* can be regarded as a kind of "cosmopolite". This endobenthic euryoecious and oligohaline polychaete settles as well in lotic as in lenitic sediments with salinity minima given as 5 psu. Of all the species inhabiting the Pomeranian Bay, it has the widest diet range. *H. diversicolor* can survive as a predatory carnivore as well as a scavenger, grazer and suspension feeder. *H. diversicolor* can penetrate the sediment up to depths of about 30 cm (Zettler et al. 1994). This species colonizes the Pomeranian Bay apparently homogenously and, vice versa to *B. pilosa* and *T. benedii*, completely independent from depth. Food quality or at least food availability due to high sedimentation and resulting accumulation rates of organic material can lead to a discrimination of benthic species over depth (O'Brien et al. 2003, Pearson and Rosenberg 1978). Such accumulation of organic material occurs throughout the year in the Pomeranian Bay and causes shifts in macrobenthic community assemblages, and thus is not necessarily a result of eutrophication (Zettler et al. 2006). It seems more likely that such community shifts are typical for estuarine ecosystems, caused by physical abrasion, accumulation or biological decomposition of aforementioned organic enrichments. This might be, together with other major environmental parameters, most likely the cause of a periodically occurring "bouncing" of benthic zonation borders (Pazdro et al. 2001, O'Brien et al. 2003). The colonization success of such opportunistic species as *T. benedii* or *H. diversicolor* can locally be explained by their sheer surviving and reproduction abilities in the extreme environmental conditions prevailing in some areas of the Pomeranian Bight, e.g. sulphidic sediments or temporarily existing oxygen deficiency zones through upwelling from deeper areas (Kube et al. 1996). On a larger (time)scale it might be due to the relatively young "ecological age" of the present Baltic Sea ecosystem (about 8 000 years), resulting in still ongoing primary succession processes, leaving numerous ecological niches available for such species (Bonsdorff 2006). Modeling species response works best with species showing a clear and recognizable optimum in the factor range examined. Nevertheless, such models can even mirror tendencies of opportunistic species without the full factor range of the ecological niche of a species. This applies for instance to *H. diversicolor*: a clear and sharply defined optimum was never found in our modeled response curves. For such species, modeling results reveal the best of available conditions in contrast to their optima. The best results are obtained for species with a response towards a factor that equals a "jump function", as is the case for the response of *B. pilosa* to organic content (Fig. 5, upper left). Here the "speed of response" is the all-dominant term regarding the modeling results. All other factor responses are simply too slow and therefore is ruled out. With such a reaction to a factor change, modeling the probability of occurrence for this factor alone can give satisfactorily results (Fig. 6, upper left and right). A factor combination with "slower" terms of the weighted overlay equation can at best compliment the picture but in the last resort can even lead to diffused results. A further inaccuracy might be the use of data sampled temporally differently. But that is not necessarily a bad thing. As was pointed out by Praca et al. (2008), the use of temporally heterogeneous data can confound the effect of interannual variation in species occurrence and environmental conditions – though one has to accept a noise floor increase in the analysis (own investigations). However, it should be mentioned that, as for all statistical methods, the inference about model selection uncertainty is conditional on both data and the full set of a priori models considered (Burnham and Anderson 2004). To derive acceptable ecological species response curves with logistic regression, Coudun and Gégout (2006) suggest a general minimum value of 50 occurrences per species and factor. This assumption was met for every species – factor combination used in our modeling approach. High variance in species distribution is not always due to sampling errors or random "noise" but rather the mechanistic consequence of shifts between limiting resources or other effects and factors (e.g. intra- and inter-species competition, predation, mortality or dispersal). The abundance of species may be very low, even under favorable conditions if, for some reason, the number of propagule is very low or species never even reach a given area. This natural phenomenon of species failing to colonize all areas where it could potentially thrive may explain some inconsistencies between the predicted high probability of species occurrence and its factual absence according to observations (Huston 2002). Under optimal conditions, species might reach maximal reproduction rates and maximal abundances, but macrobenthic survey often shows an entirely different reality: species and communities are distributed rather patchily and often the relatively smooth structure of abiotic gradients and other characteristics can increase statistical uncertainty and blur the picture. For a detailed overview and reflection *inter alia* on our method and alternatives used in marine modeling of species response see also Guisan and Zimmermann (2000), Muñoz and Felicisimo (2004), Guisan and Thuiller (2005), Guisan et al. (2006), Wisz and Guisan (2009) and Gogina et al. (2010). However, our objectives were to attempt a general description of species habitats and to investigate the predictive ability of the modeling technique at the selected spatial scale. Confirming the postulates of Ysebaert et al. (2002), Thrush et al. (2003), and Ellis et al. (2006), we have found that logistic regression in combination with a weighted overlay approach is a useful and relatively transparent approach to predict the response of species occurrence as a function of various environmental conditions.

CONCLUSIONS

Often, salinity is regarded as the one and only primary descriptor in brackish estuarine ecosystems. Considering the fact that all species examined (this study, Glockzin and Zettler 2008) are regarded as euryhaline and that the salinity does not vary strongly in the center of the Pomeranian Bay, the apparent influence of salinity as the main predictor on the benthic assemblage becomes insignificant under certain scale. As already pointed out, the prevailing environmental parameters in the study area, such as salinity, organic content, sorting or permeability, seem to be controlled primarily by water depth; it could even be described as a “master factor” and most likely, through its impact on all other factors, it is mainly responsible for the spatial zonation of benthic species, not only in the Pomeranian Bay but other marine habitats as well (e.g. Kube et al. 1996, Bonsdorff et al. 2003, O’Brien et al. 2003, Kröncke et al. 2004, Warzocha 1995, Zettler et al. 2006, Glockzin and Zettler 2008, Gogina et al. 2010). Our study clearly demonstrates the possibility to model species’ relationships in gradient systems such as the Baltic Sea where their patterns of distribution are strongly and directly coupled to abiotic processes. By using species’ response towards responsible environmental key factors to model spatial distributions for selected benthic species via a binomial logistic regression approach, we propagate an easy-to-use tool to assess natural and anthropogenic forced changes in benthic species distribution patterns. Thinking through the consequences and possibilities of these methods and doing the necessary enhancements, e.g. by coupling species ecological functions (e.g. filtration rates, bioturbation modes, e.g. Forster and Zettler 2004) to our model, we might even be able to assess an ecosystems functioning – and the loss of it. Bearing in mind the currently experienced impact of natural (e.g. via climate change) or anthropogenic forced changes (e.g. by pipeline building or dredge fishing) on the Baltic Sea ecosystem – never has the need for such a tool been more urgent than nowadays.

Acknowledgements

Sampling and preparation in this study were conducted in accordance with institutional, national and international guidelines concerning the use of animals in research (HELCOM 1988). Sincere thanks are given to Dr. Thorsten Seifert for helpful assistance regarding the treatment of bathymetrical data and to Dr. Bernd Bobertz, Dr. Björn Bohling and Dr. Michael Meyer for helpful assistance on granulometry and GIS. Not to forget the work of Ines Glockzin and her technical assistance. This project was funded by the German Federal Agency for Nature Conservation (Federal Authority of Federal Ministry for the Environment, Nature Conservation and Nuclear Safety) with the grant number 80 285 210.

REFERENCES

Austin M.P., 1987. Models for the analysis of species'response to environmental gradients. *Plant. Ecol.* 69, 35-45.

Bonsdorff E., Laine A.O., Hänninen J., Vuorinen I., Norkko A., 2003. Zoobenthos of the outer archipelago waters (N. Baltic Sea) – the importance of local conditions for spatial distribution patterns. *Boreal. Env. Res.*, 8, 135-145.

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.

Bourget E., Fortin M.-J., 1995. A commentary on current approaches in the aquatic sciences. *Hydrobiologia*, 300/301, 1-16.

Burnham K. P., Anderson D. R., 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33-2, 261-304.

Coudun Ch., Gégout J.-C., 2006. The derivation of species response curves with Gaussian logistic regression is sensitive to sampling intensity and curve characteristics. *Ecol. Model.*, 199, 164-175.

Diaz R.J., 1984. Short term dynamics of the dominant annelids in a polyhaline temperate estuary. *Hydrobiologia*, 115, 153-158.

Dubilier N., Giere O., Grieshaber M.K., 1994. Concomitant effects of sulfide and hypoxia on the aerobic metabolism of 241 the marine oligochaete Tubificoides benedii. *J. Exp. Zool.*, 269, 287-297.

Ellis J., Ysebaert T., Hume T., Norkko A., Bult T., Herman P., Thrush S., Oldman J., 2006. Predicting macrofaunal species distribution in estuarine gradients using logistic regression and classification systems. *Mar. Ecol. Prog. Ser.*, 316, 69-83.

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

Folk R.L., Ward W.C., 1957. Brazos river bar, a study in the significance of grain size parameters. *J. Sed. Petrol.*, 27, 3-26.

Forster S., Bobertz B., Bohling B., 2003. Permeability of sands in the coastal areas of the southern Balic Sea, Mapping a grain-size related sediment property. *Aquat. Geochem.*, 9, 171-190.

Forster S., Zettler M.L., 2004. The capacity of the filter-feeding bivalve *Mya arenaria* L. to affect water transport in sandy beds. *Marine Biology*, V144(6), 1183-1189.

Fortin M.-J., Keitt T.H., Maurer B.A., Taper M.L., Kaufman D.M., Blackburn T.M., 2005. Species' geographic ranges and distributional limits, pattern analysis and statistical issues. *Oikos*, 108, 7-17.

Giere O., Preusse J.-H., Dubilier N., 1999. *Tubificoides benedii* (Tubificidae, Oligochaeta) – a pioneer in hypoxic and sulfidic environments. An overview of adaptive pathways. *Hydrobiologia*, 406, 235-241.

Giere O., 2006. Ecology and biology of marine Oligochaeta - an inventory rather than another review. *Hydrobiologia*, 564, 103-116.

Glockzin M., Zettler M.L., 2008 Spatial macrozoobenthic distribution patterns in relation to major environmental factors – a case study from the Pomeranian Bay (southern Baltic Sea). *J. Sea Res.*, 59, 144-161.

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

Guisan A., Weiss S.B., Weiss A.D., 1999. GLM versus CCA spatial modelling of plant species distribution. *Plant Ecol.*, 143, 107-122.

Guisan A., Zimmermann N.E., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2-3),147-186.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8, 993-1009

Guisan A., Lehman A., Ferrier S., Austin M., Overton J.M.C.C., Aspinall R., Hastie T., 2006. Making better biogeographical predictions of species' distributions. *J. Appl. Ecol.*, 43, 386-392.

HELCOM, 1988. Guidelines for the Baltic Monitoring Programme for the Third Stage, Part D. Biological Determinants. Baltic Sea Environmental Proceedings, 27.

Huston M.A., 2002. Critical issues for improving predictions. In: Predicting Species Occurrences: Issues of Accuracy and Scale. (Eds.), J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall, F.B. Samson, Island Press, Washington, 7-21.

Keitt T.H., Bjørnstad O.N., Dixon P.M., Citron-Pousty S., 2002. Accounting for spatial pattern when modelling organism-environment interactions. *Ecography*, 25, 616-625.

Kröncke I., Stoeck T., Wieking G., Palojärvi A., 2004. Relationship between structural and functional aspects of microbial and macrofaunal communities in different areas of the North Sea. *Mar. Ecol. Prog. Ser.*, 282, 13-31.

Krumbein W.C., Monk G.D., 1942. Permeability as a function of the size parameters of unconsolidated sand. American Institute Mining and Metallurgy Engineering, Petroleum Technology, Technical Publikation no. 1942.

Kube J., Powilleit M., Warzocha J., 1996. The importance of hydrodynamic processes and food availability for the structure of macrofauna assemblages in the Pomeranian Bay (Southern Baltic Sea). *Arch. Hydrobiol.*, 138, 213-228.

Kuhrts C., Seifert T., Fennel W., 2006. Modelling transport of fluff layer material in the Baltic Sea. *Hydrobiologia*, 554, 25-30.

Legendre P., Legendre L., 1998. Numerical Ecology. Amsterdam, Elsevier Science B.V.

Meissner K., Darr A., Rachor E., 2008. Development of habitat models for *Nephtys* species (Polychaeta: Nephtyidae) in the German Bight (North Sea). *J. Sea Res.*, 60, 271-286.

Muñoz J., Felicisimo A., 2004. A comparison between some statistical methods commonly used in predictive modelling. *J. Veg. Sci.*, 15, 285-292.

Nicolaisen W., Kanneworff E., 1969. On the burrowing and feeding habits of the amphipods *Bathyporeia pilosa* Lindstrom and *Bathyporeia sarsi* Watkin. *Ophelia*, 6, 231-250.

O'Brien K., Hänninen J., Kanerva T., Metsärinne L., Vuorinen I., 2003. Macrobenthic zonation in relation to major environmental factors across the Archipelago Sea, northern Baltic Sea. *Boreal Env. Res.*, 8, 159-170.

Pastuszak M., Witek Z., Nakel K., Wielgat M., Grelowski A., 2005. Role of the Oder estuary (southern Baltic) in transformation of the riverine nutrient loads. *J. Mar. Syst.*, 57, 30-54.

Pavlikakis G.E., Tsihrintzis V.A., 2000. Ecosystem Management: A Review of a New Concept and Methodology. *Water Resour. Manag.*, 14, 257-283.

Pazdro K., Staniszewski A., Bełdowski J., Emeis K.-Ch., Leipe T., Pempkowiak J., 2001. Variations in organic matter bound in fluffy layer suspended matter from the Pomeranian Bay (Baltic Sea). *Oceanologia*, 43, 405-420.

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

Praca E., Gannier A., Das K., Laran S., 2008. Modelling the habitat suitability of cetaceans: Example of the sperm whale in the northwestern Mediterranean Sea. *Deep-Sea Research*, I, 648-657.

Sachs L., 1997. Angewandte Statistik. Anwendung statistischer Methoden. (Applied Statistics). Springer, Berlin, (in German).

Speybroeck J., Van Tomme J., Vincx M., Degraer S., 2008. In situ study of the autecology of the closely related, co-occurring sandy beach amphipods *Bathyporeia pilosa* and *Bathyporeia sarsi*. *Helgol. Mar. Res.*, 62, 257-268.

Sundbäck K., Persson L.-E., 1981. The effect of microbenthic grazing by an amphipod, *Bathyporeia pilosa*, Lindström. *Kieler Meeresforsch.*, 5, 573-575.

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.

Warzocha J., 1995. Classification and structure of macrofaunal communities in the southern Baltic. *Arch. Fish. Mar. Res.*, 42, 225-237.

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).

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.

Zettler M.L., Bochert R., Bick A., 1994. Röhrenbau und Vertikalverteilung von *Marenzelleria viridis* (Polychaeta: Spionidae) in einem inneren Küstengewässer der südlichen Ostsee. (Tube-building and vertical distribution of *Marenzelleria viridis* (Polychaeta: Spionidae) in an inner coastal water of the Southern Baltic Sea). *Rostocker Meeresbiol. Beitr.*, 2, 215-225, (in German).

Zettler M.L., Röhner M., Frankowski J., 2006. Long term changes of macrozoobenthos in the Arkona Basin (Baltic Sea). *Boreal Env. Res.*, 11, 247-260.

Zettler M.L., Schiedek D., Bobertz B. 2007. Benthic biodiversity indices versus salinity gradient in the southern Baltic Sea. *Marine Pollution Bulletin*, 55, 258-270.

V

Journal of Sea Research xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Journal of Sea Research

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

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

A R T I C L E   I N F O

Article history:
Received 16 November 2009
Received in revised form 20 April 2010
Accepted 21 April 2010
Available online xxxx

Keywords:
Zoobenthos Distribution
Environmental Factors
Ecological Modelling
Logistic Regression
Large-scale Statistical Modelling
Habitat Suitability Mapping
Baltic Sea

A B S T R A C T

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.

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).

1385-1101/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.seares.2010.04.005

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 hydrological 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 stratification 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.

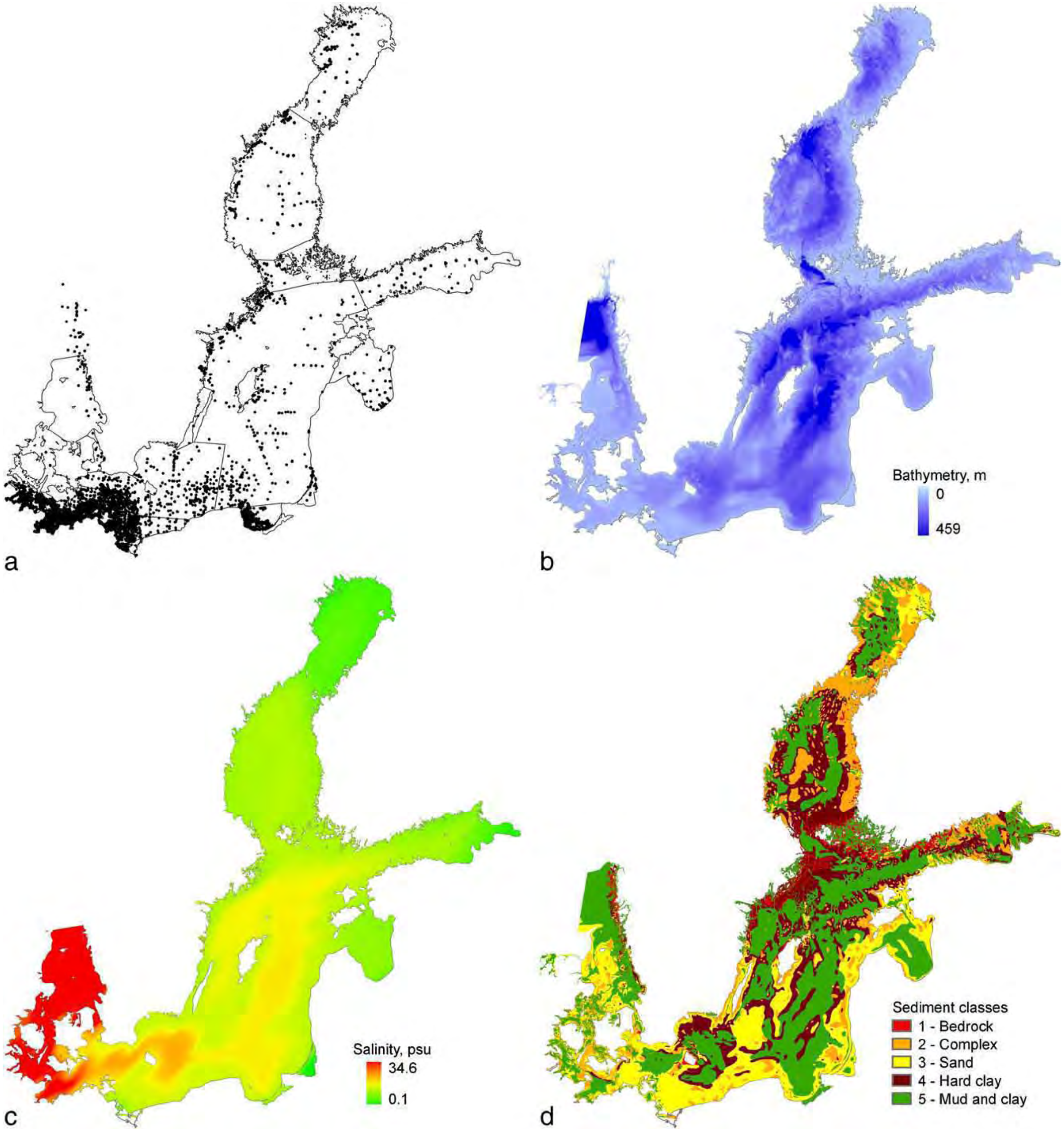


Fig. 1. Stations sampled for macrofauna within the Baltic Sea sub-regions are indicated by dots (a) and distribution surfaces of the environmental variables considered covering the Baltic Sea: bathymetry (b), near-bottom salinity (c), seabed sediment type (d). Geographical data ESRI (2003); projection 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 assumptions 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 $\gamma_b = e^{\left(b_0 + b_1S + b_2S^2 + b_3D + b_4D^2 + b_5\mathrm{Sed}_1 + b_6\mathrm{Sed}_2 + b_7\mathrm{Sed}_3 + b_8\mathrm{Sed}_4\right)}$. 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, $AUC\leq0.5$ 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 macroinvertebrates 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,

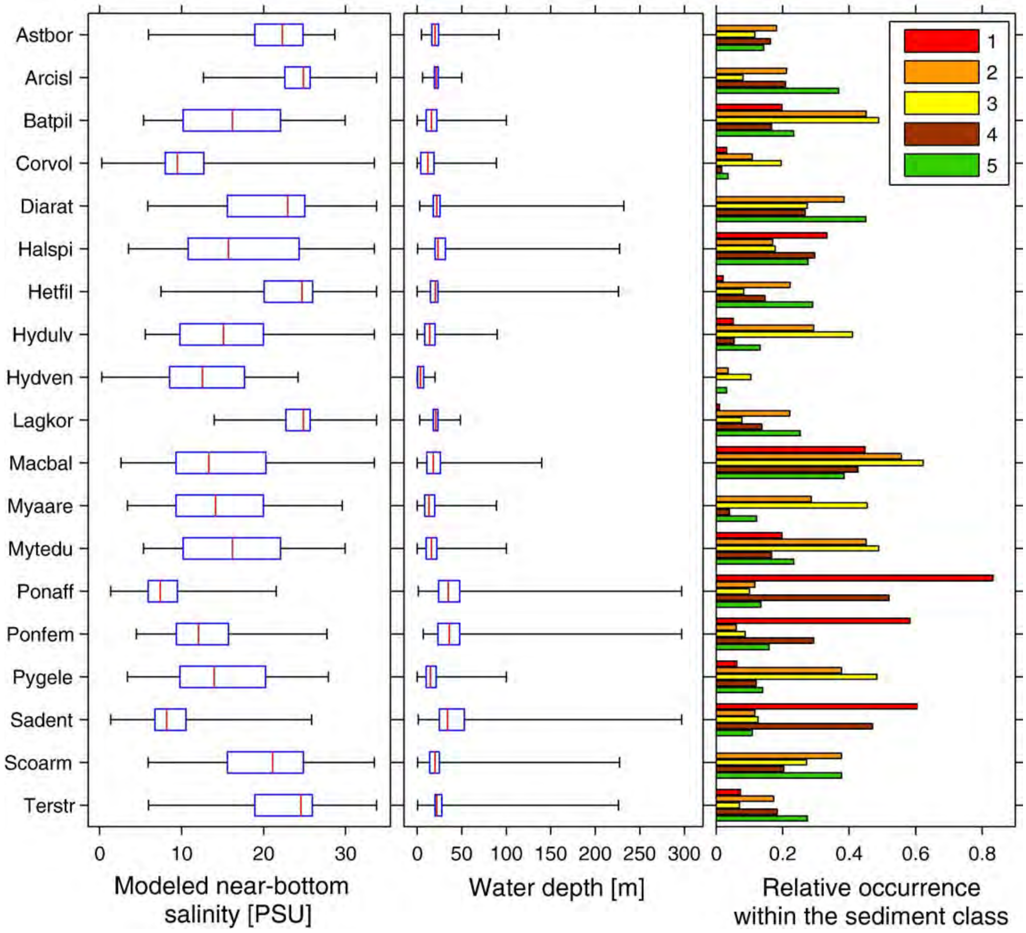


Fig. 2. Occurrence of 19 exemplary species along the ranges of investigated environmental parameters. Species are ordered alphabetically, 6-letter codes from top to the bottom referring to *Astarte borealis*, *Arctica islandica*, *Bathyporeia pilosa*, *Corophium volutator*, *Diastylis rathkei*, *Halicryptus 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 Åland 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

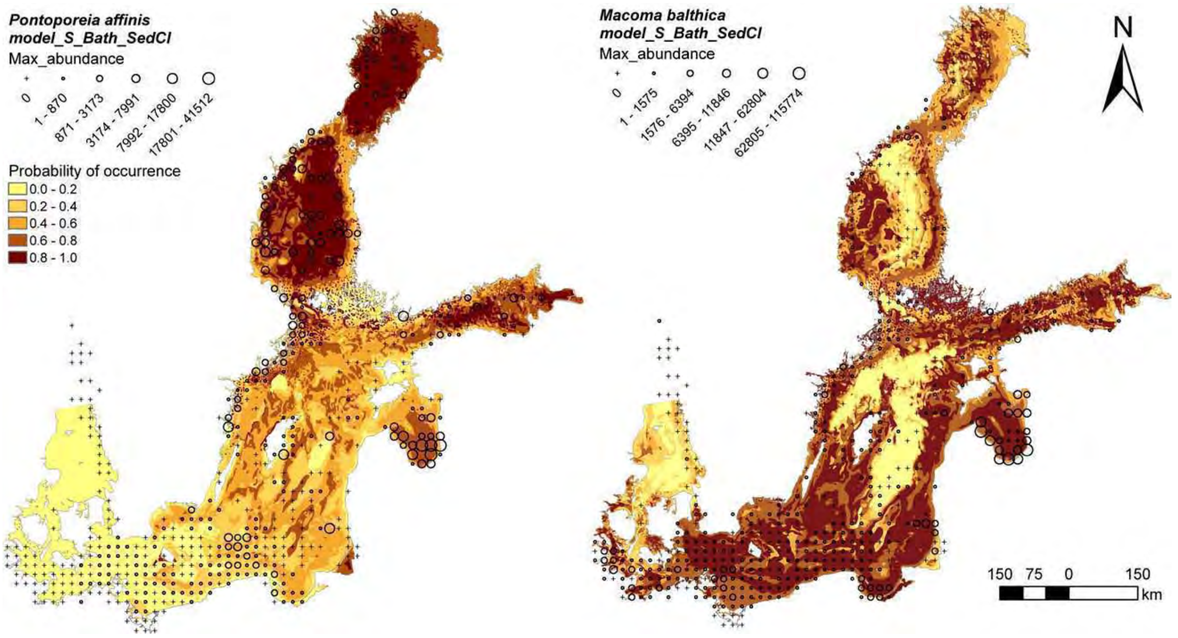


Fig. 3. Modelling results compared with the observed data for the amphipod *Pontoporeia affinis* (left) and the bivalve *Macoma balthica* (right). Symbols indicate the observed species abundance (ind/m^2) with their size corresponding to the value of abundance density (the maximal value observed for each cell of a regular 20 km grid is plotted in order to simplify visualization). Color fields show the modelled probability of occurrence. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

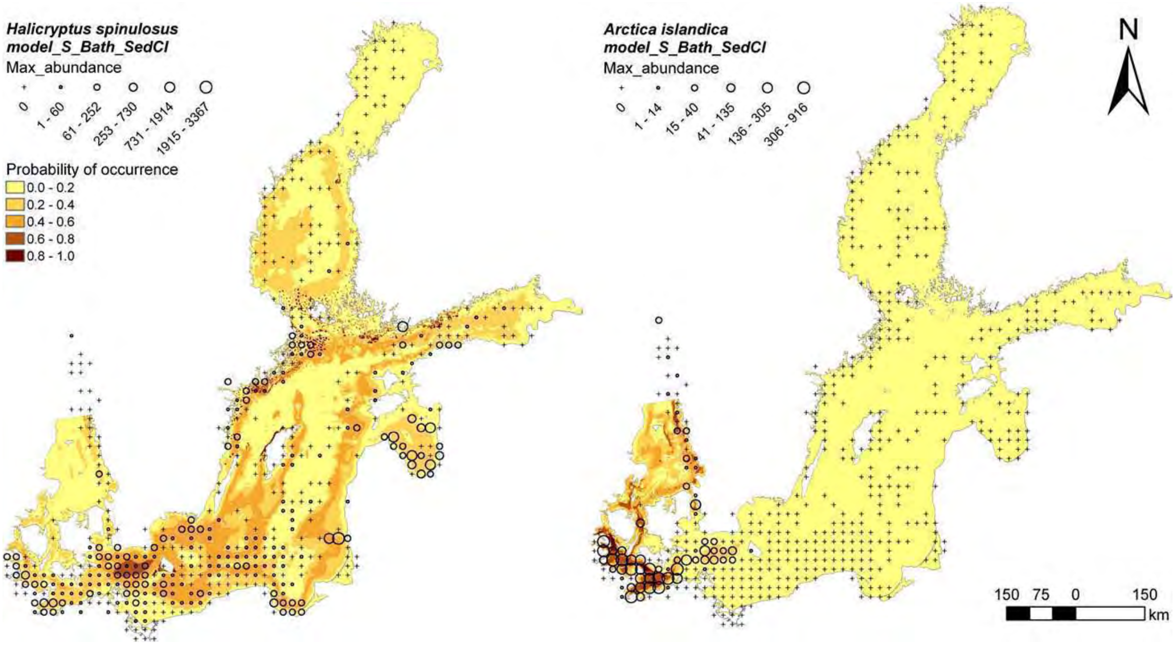


Fig. 4. Modelling results compared with the observed data for the priapulid *Halicryptus spinulosus* (left) and the bivalve *Arctica islandica* (right). Symbols indicate the observed species abundance (ind/m^2) with their size corresponding to the value of abundance density (the maximal value observed for each cell of a regular 20 km grid is plotted in order to simplify visualization). Color fields show the modelled probability of occurrence. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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), 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 fulfill 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

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

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

Part of this work was funded by the EU-MARBEF project. We gratefully acknowledge the work of all colleagues involved and responsible for field work and laboratory analysis, in particular Ines Glockzin, the valuable assistance on mathematical, statistical and GIS matters of Dr. Bernd Bobertz, Dr. Michael Meyer and Michael Glockzin. We also thank Dr. Torsten Seifert for his great help with modelled data on hydrographic parameters. The generation of this data was partly supported by BMBF within the framework of Project IKZM-Oder (03F0403A) and by DFG-grant NE G17/3-1. It would be impossible to perform this work without the work of HELCOM and ICES database team, and the results of the BALANCE project. We thank the two anonymous reviewers for their critical, essential comments on the manuscript that were very valuable to achieve this final version. We are indebted to Anastasia Semenova for correcting the English.

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), 57–7010.1016/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. Hydrobiologia 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.

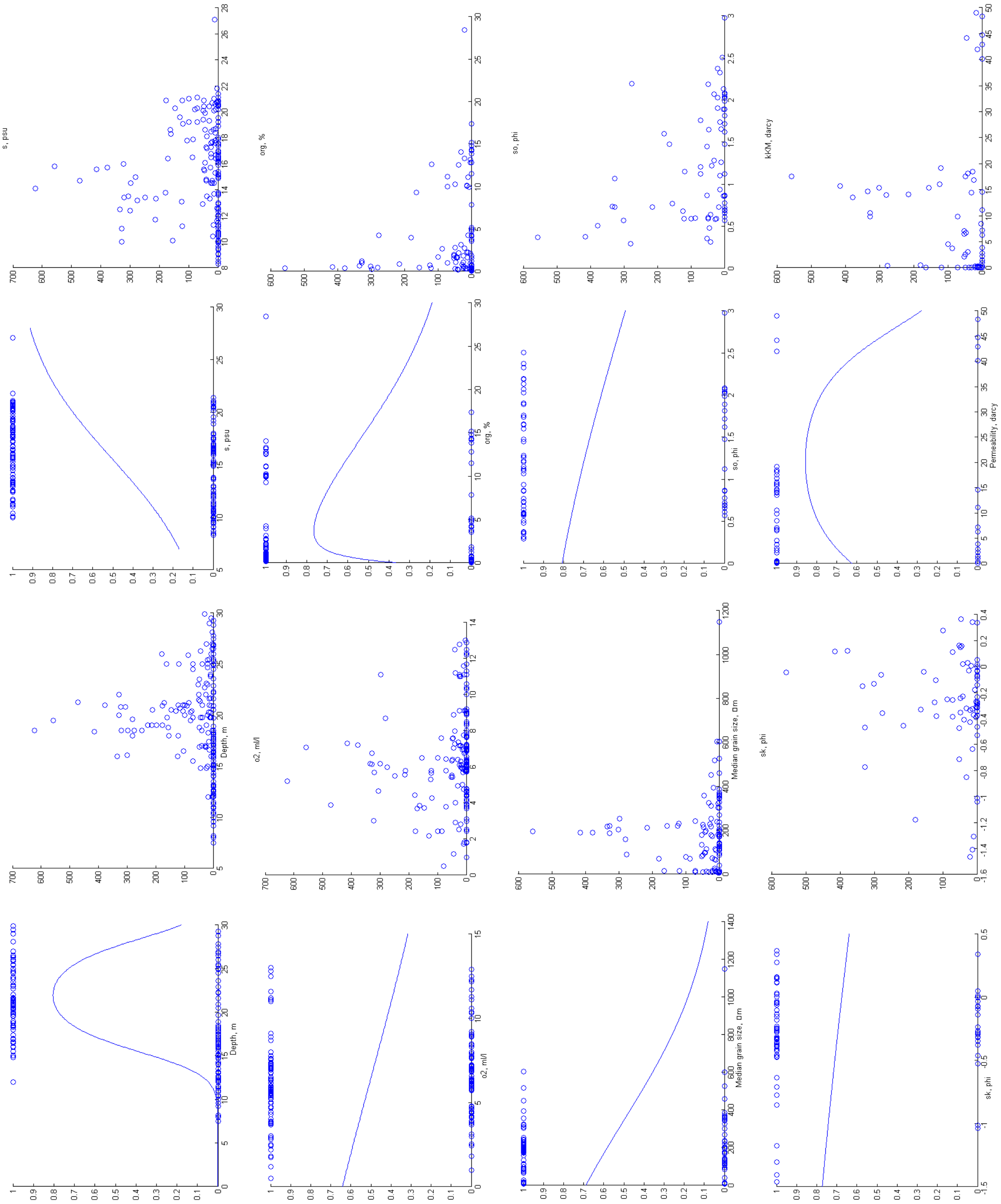
Zettler, M.L., Schiedek, D., Glockzin, M., 2008. Zoobenthos. In: Feistel, R., Nausch, G., Wasmund, N. (Eds.), State and Evolution of the Baltic Sea. John Wiley & Sons, Inc, pp. 517–541. 1952–2005.

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

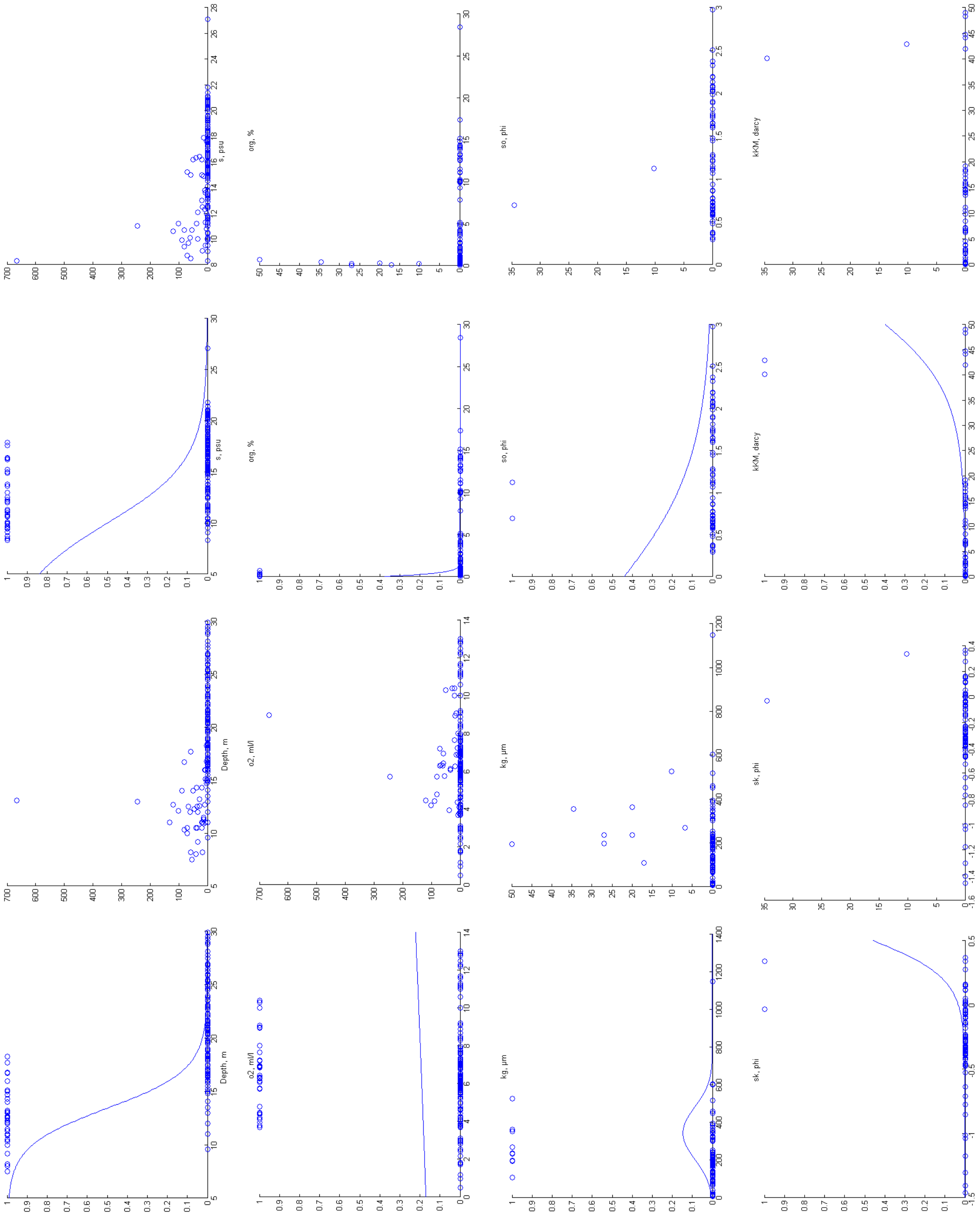
Appendix A

Predicted probability of occurrence and measured abundance vs. abiotic factors for five selected species: *Arctica islandica, Hediste diversicolor, Pygospio elegans, Tubificoides benedii and Scoloplos armiger*

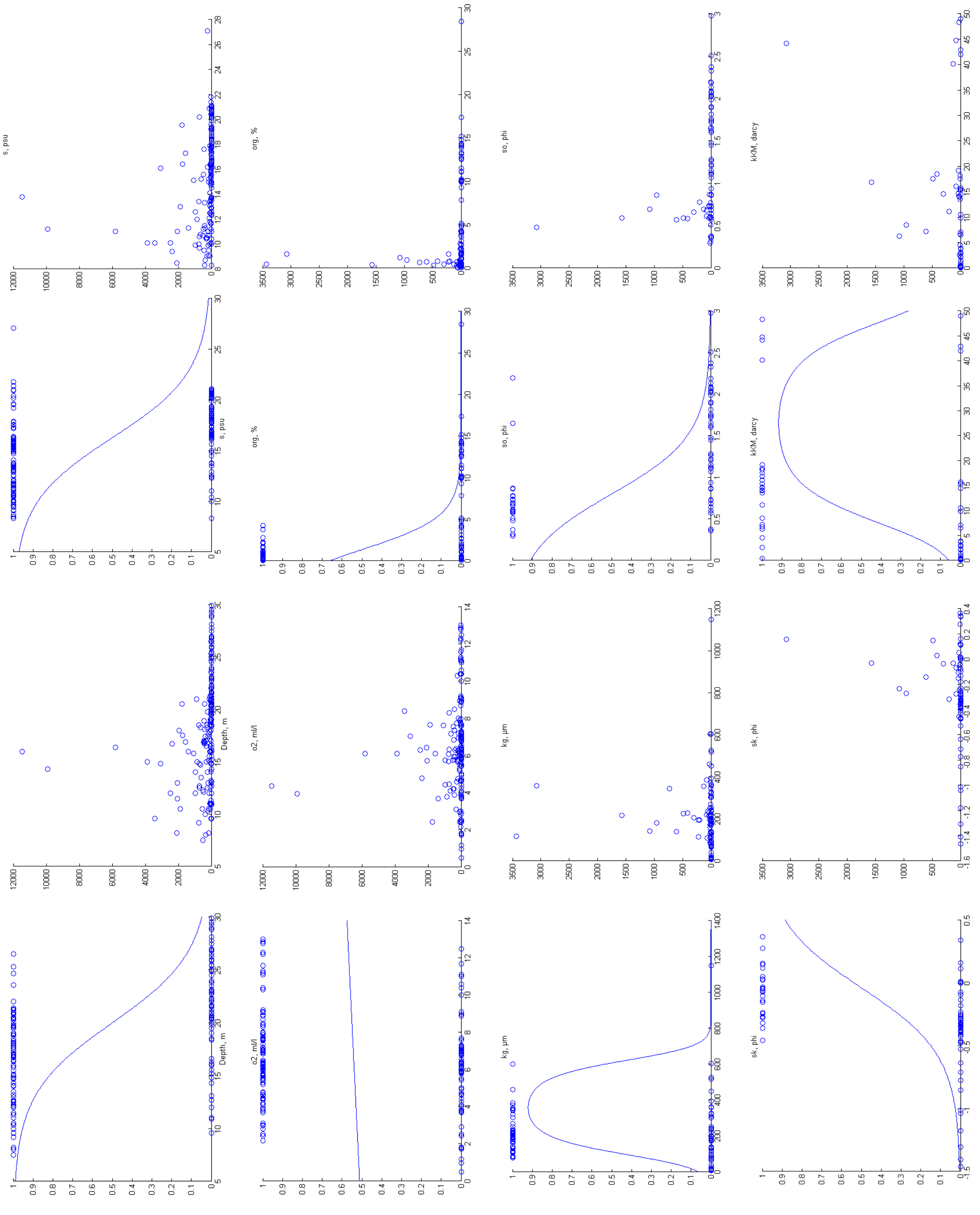
A. islandica, predicted probability of occurrence and measured abundance vs. abiotic factors



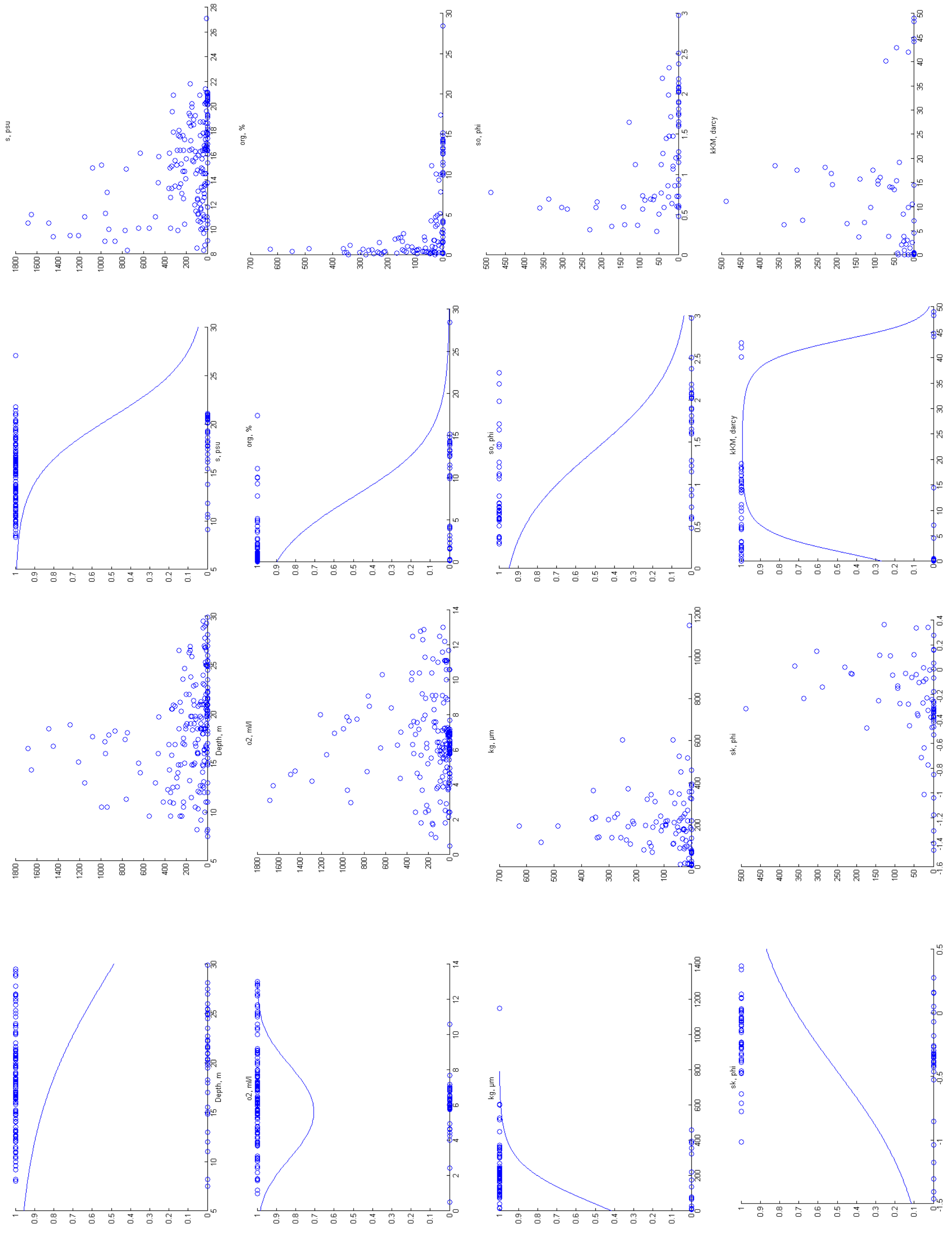
H. diversicolor, predicted probability of occurrence and measured abundance vs. abiotic factors



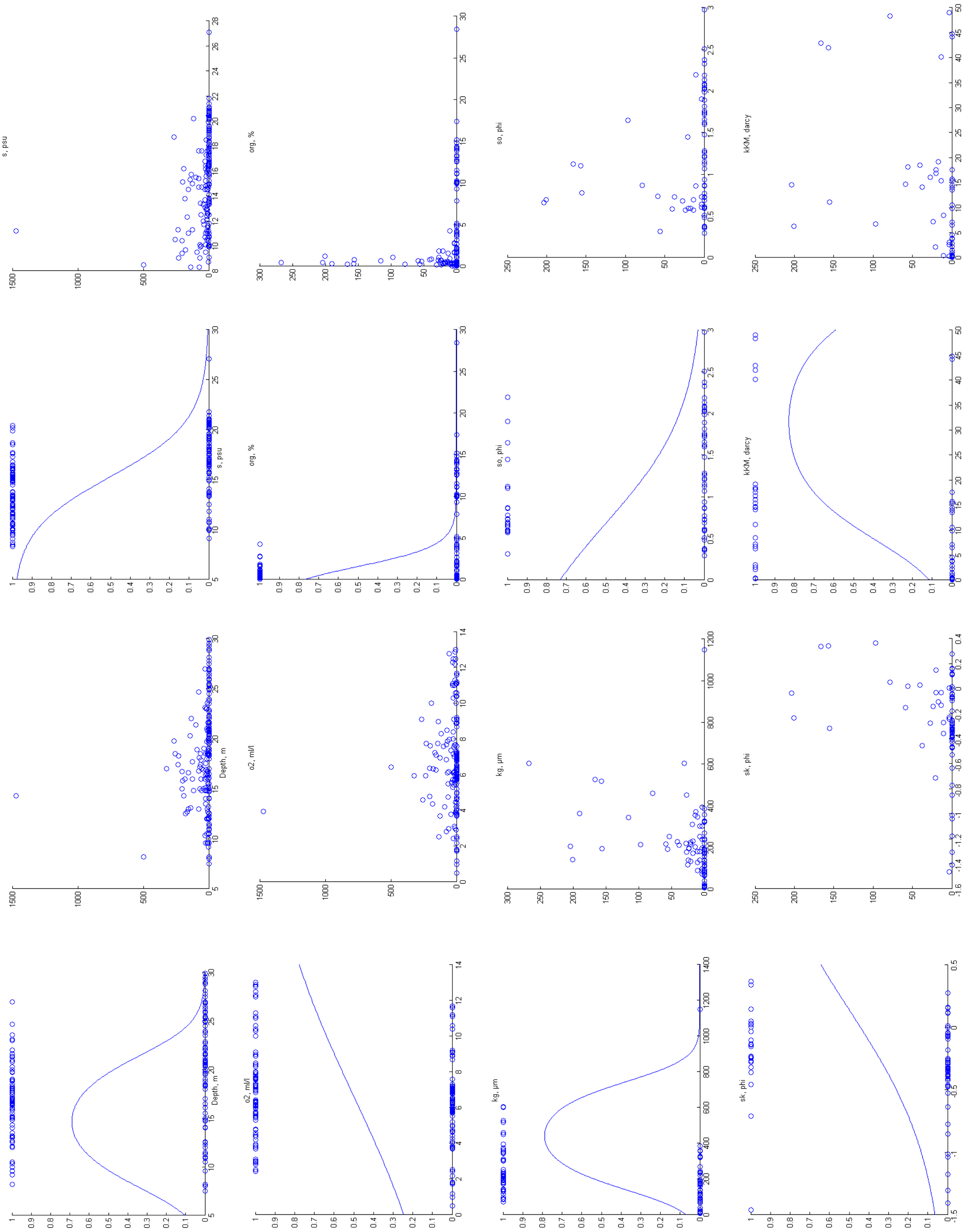
P. elegans, predicted probability of occurrence and measured abundance vs. abiotic factors



S. armiger, predicted probability of occurrence and measured abundance vs. abiotic factors

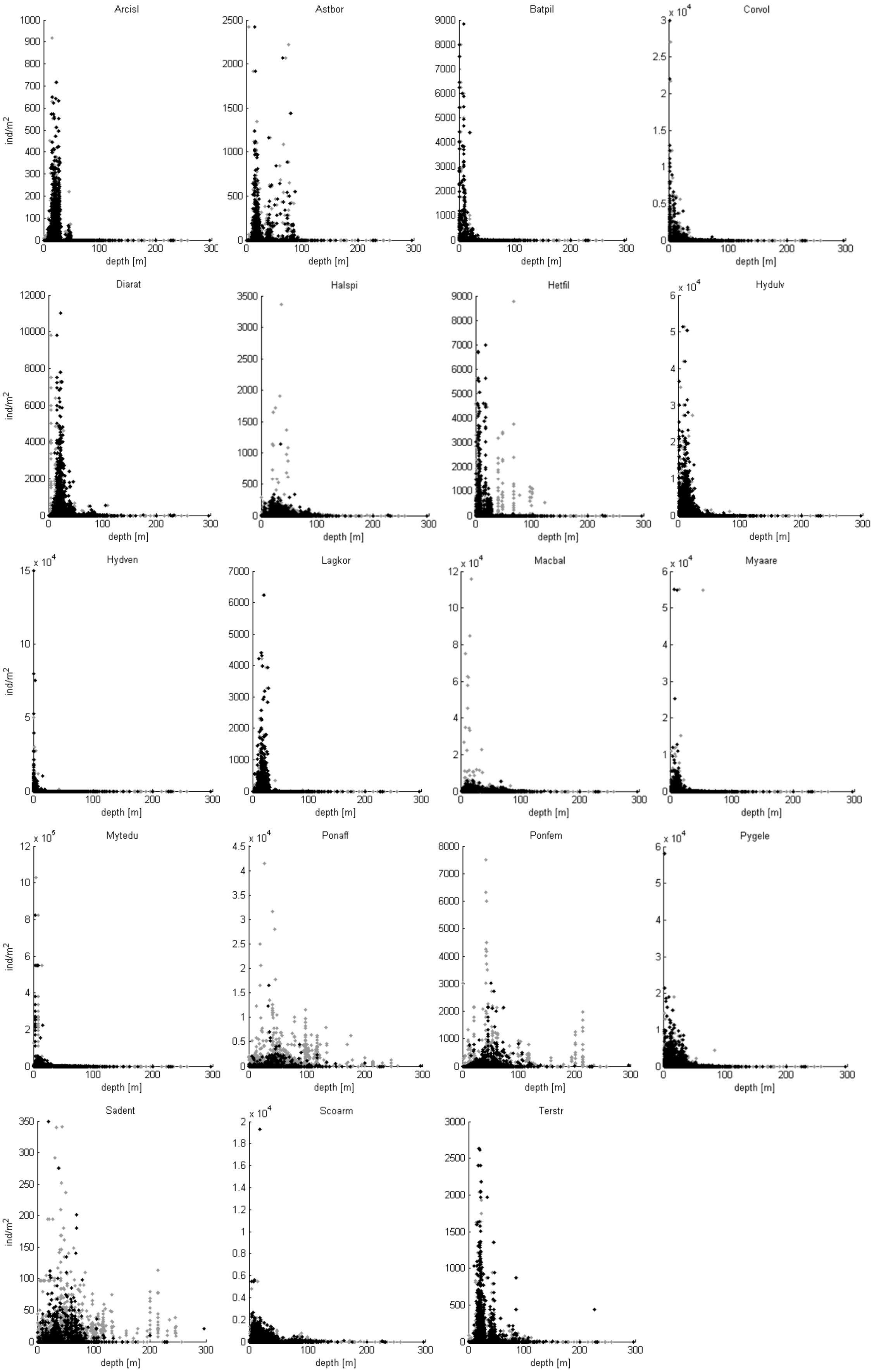


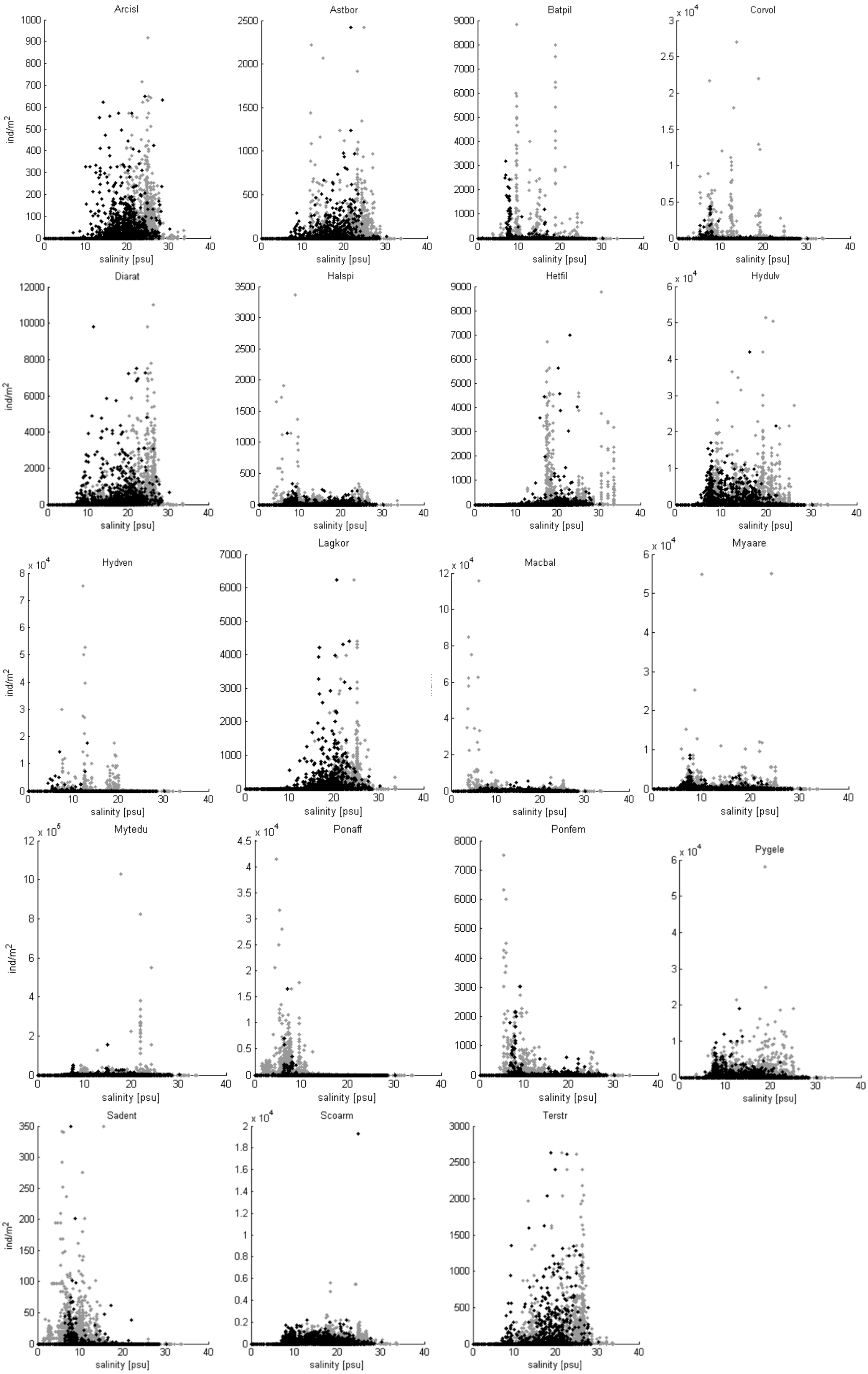
T. benedeni, predicted probability of occurrence and measured abundance vs. abiotic factors

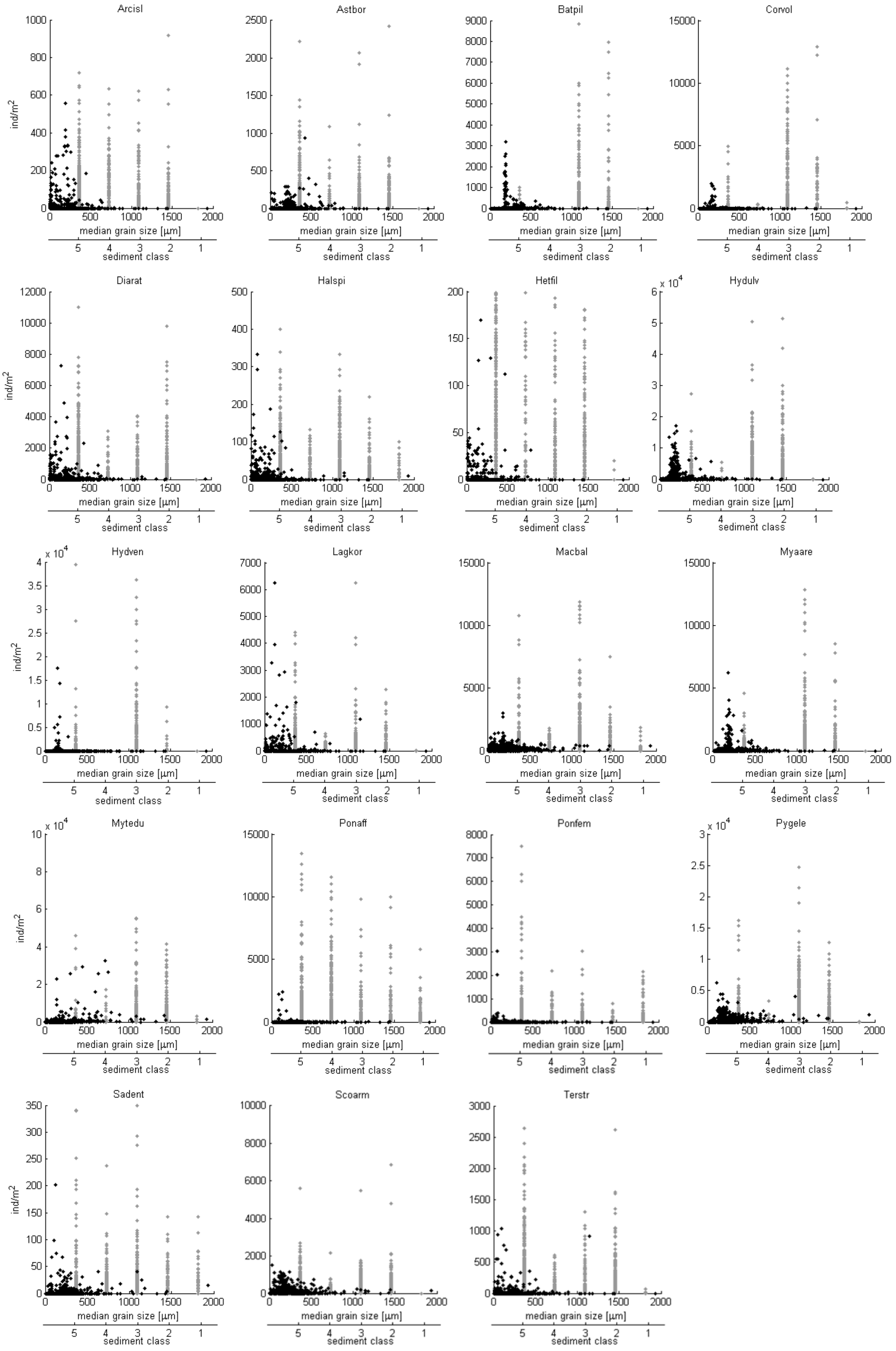


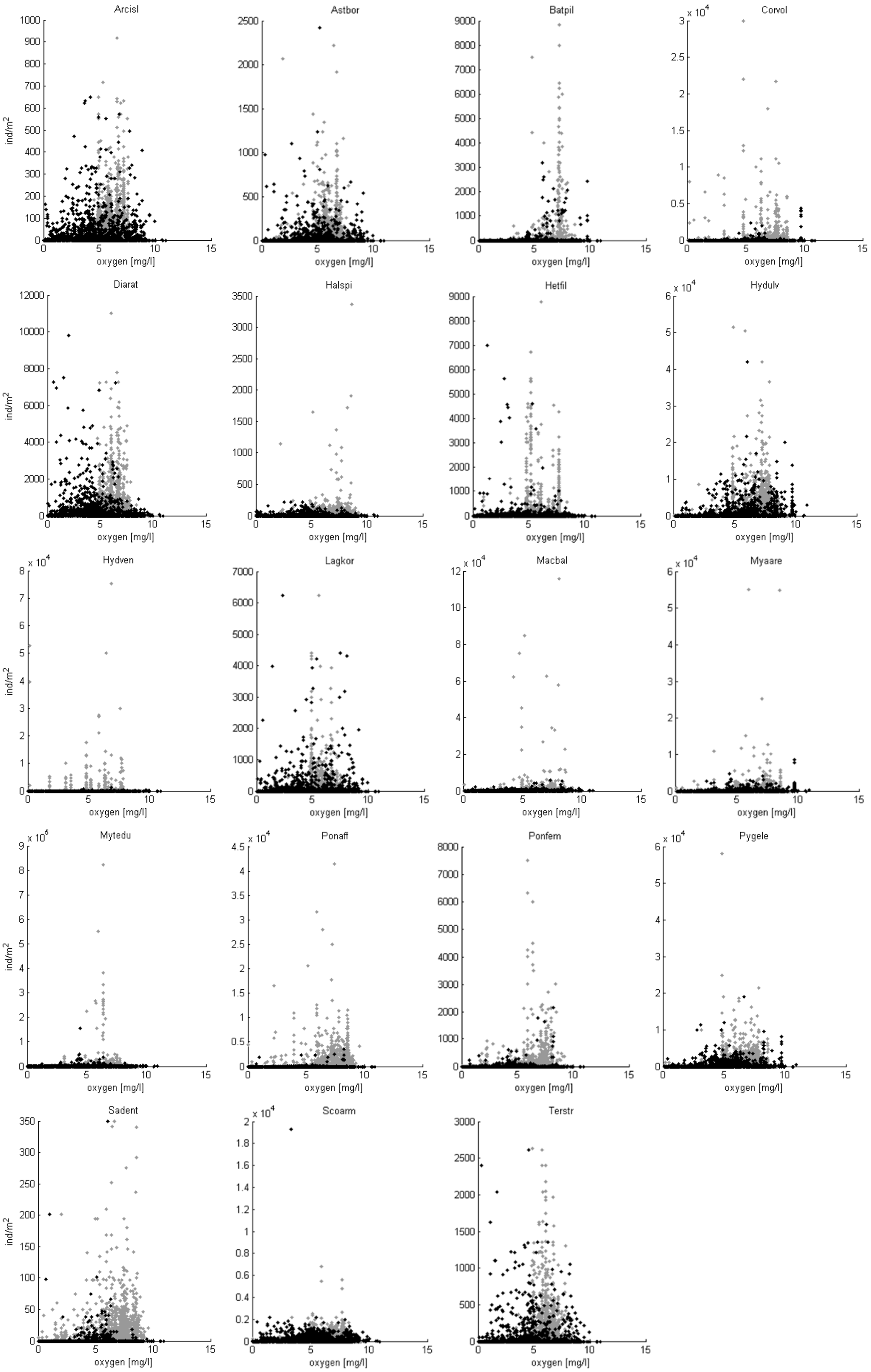
Appendix B

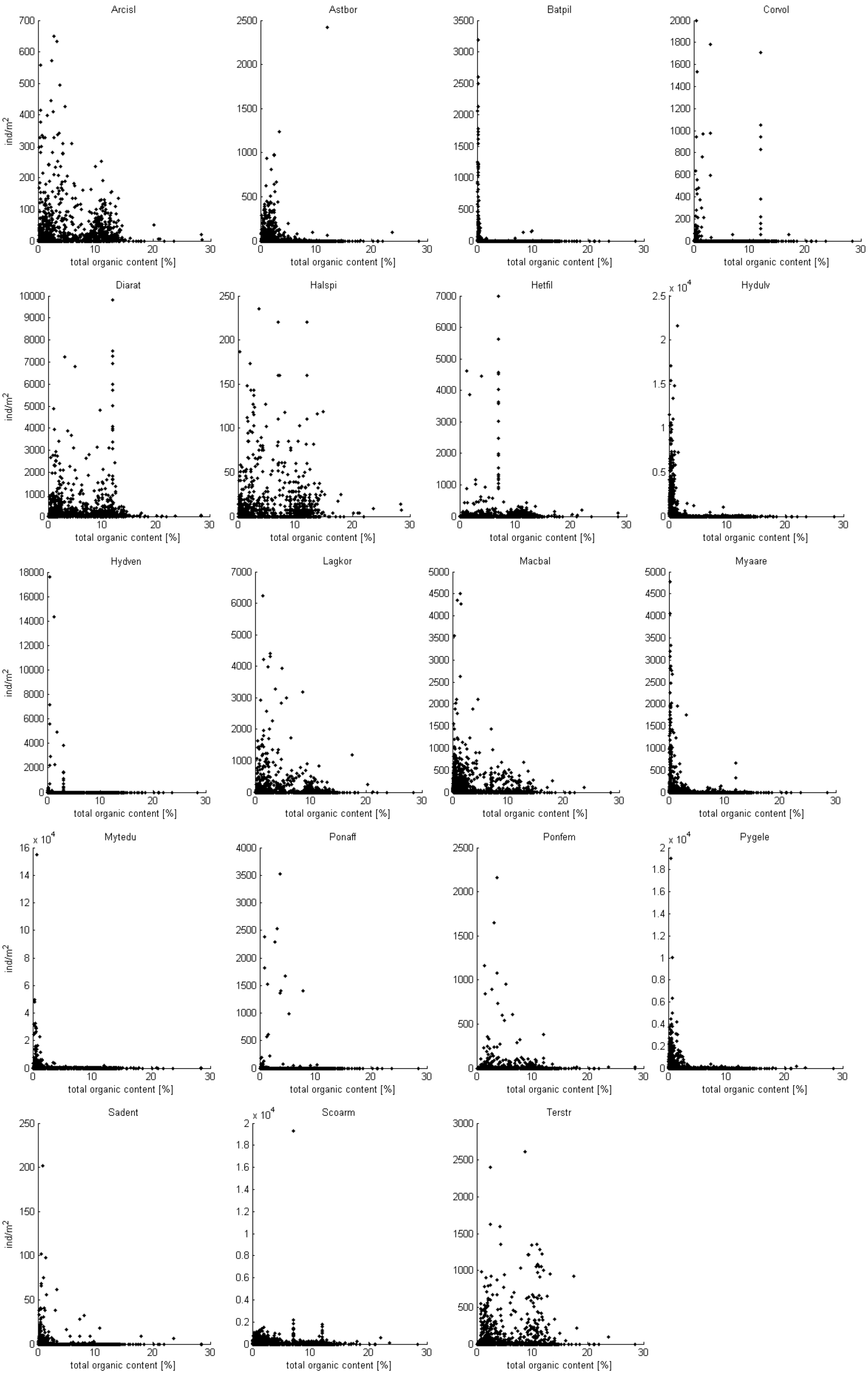
Scatter plots of abundance density values for 19 macrobenthic species representative for the Baltic Sea vs. environmental factors (depth, salinity, sediment class/median grain size, near-bottom oxygen concentrations, total organic content in sediments: measured data for environmental variables – black dots, modelled – gray dots). Species abbreviation: Astbor, Arcisl, Batpil, Corvol, Diarat, Halspi, Hetfil, Hydulv, Hydven, Lagkor, Macbal, Myaare, Mytedu, Ponaff, Ponfem, Pygele, Sadent, Scoarm, Terstr stand for *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*, *Terebellides stroemi*


Arcisl
Astbor
Batpil
Corvol
Diarat
Halspi
Hetfil
Hydulv
Hydven
Lagkor
Macbal
Myaare
Mytedu
Ponaff
Ponfem
Pygele
Sadent
Scoarm
Terstr
ind/m^2
depth [m]
0
100
200
300
x 10^4
x 10^5


Arcisl
Astbor
Batpil
Corvol
Diarat
Halspi
Hetfil
Hydulv
Hydven
Lagkor
Macbal
Myaare
Mytedu
Ponaff
Ponfem
Pygele
Sadent
Scoarm
Terstr
ind/m^2
salinity [psu]
x 10^4
x 10^5
0
10
20
30
40

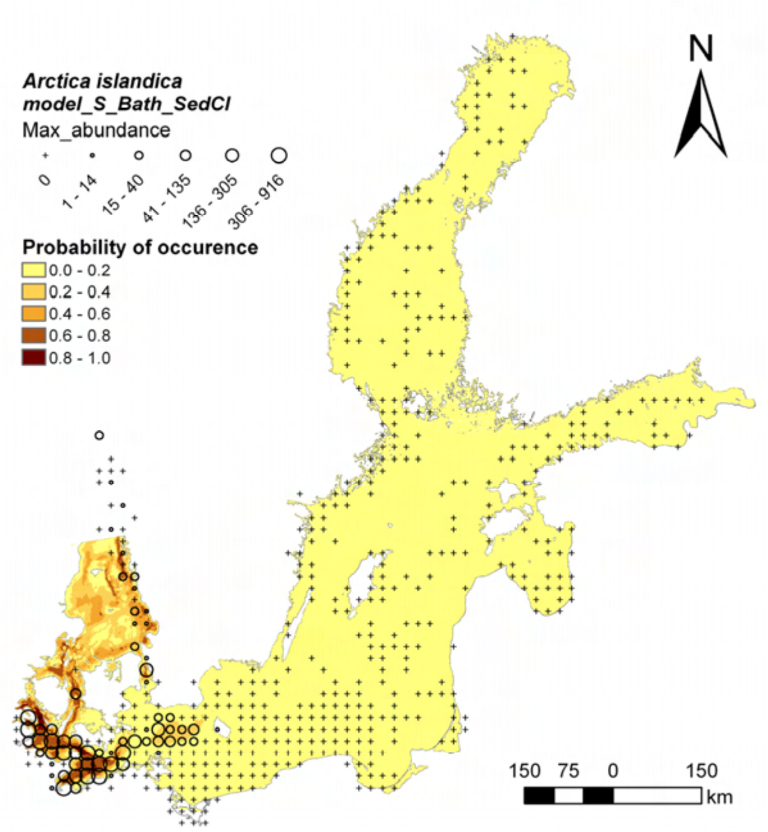

Arcisl
Astbor
Batpil
Corvol
Diarat
Halspi
Hetfil
Hydulv
Hydven
Lagkor
Macbal
Myaare
Mytedu
Ponaff
Ponfem
Pygele
Sadent
Scoarm
Terstr
ind/m^2
x 10^4
median grain size [µm]
0
500
1000
1500
2000
sediment class
5
4
3
2
1

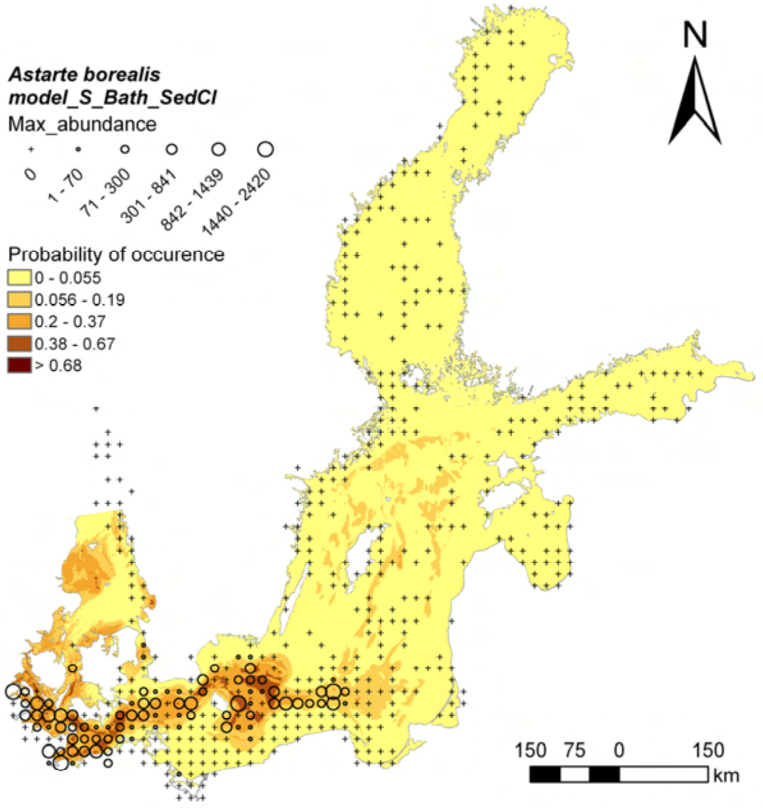

Arcisl
Astbor
Batpil
Corvol
x 10^4
Diarat
Halspi
Hetfil
Hydulv
x 10^4
Hydven
x 10^4
Lagkor
Macbal
x 10^4
Myaare
x 10^4
Mytedu
x 10^5
Ponaff
x 10^4
Ponfem
Pygele
x 10^4
Sadent
Scoarm
x 10^4
Terstr
ind/m2
oxygen [mg/l]
0
5
10
15

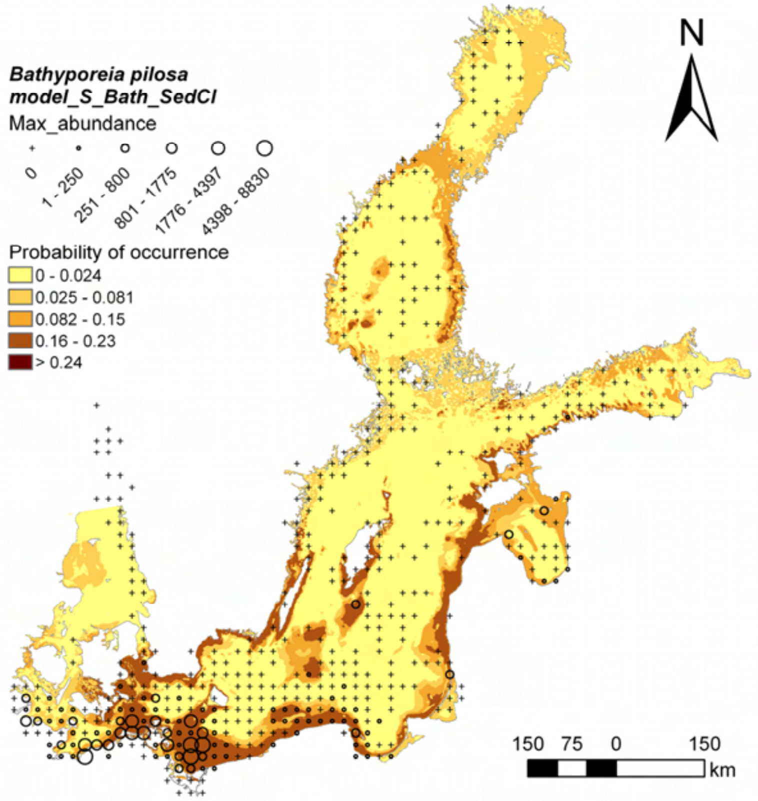

Arcisl
Astbor
Batpil
Corvol
Diarat
Halspi
Hetfil
Hydulv
Hydven
Lagkor
Macbal
Myaare
Mytedu
Ponaff
Ponfem
Pygele
Sadent
Scoarm
Terstr
ind/m^2
total organic content [%]
x 10^4
0
10
20
30

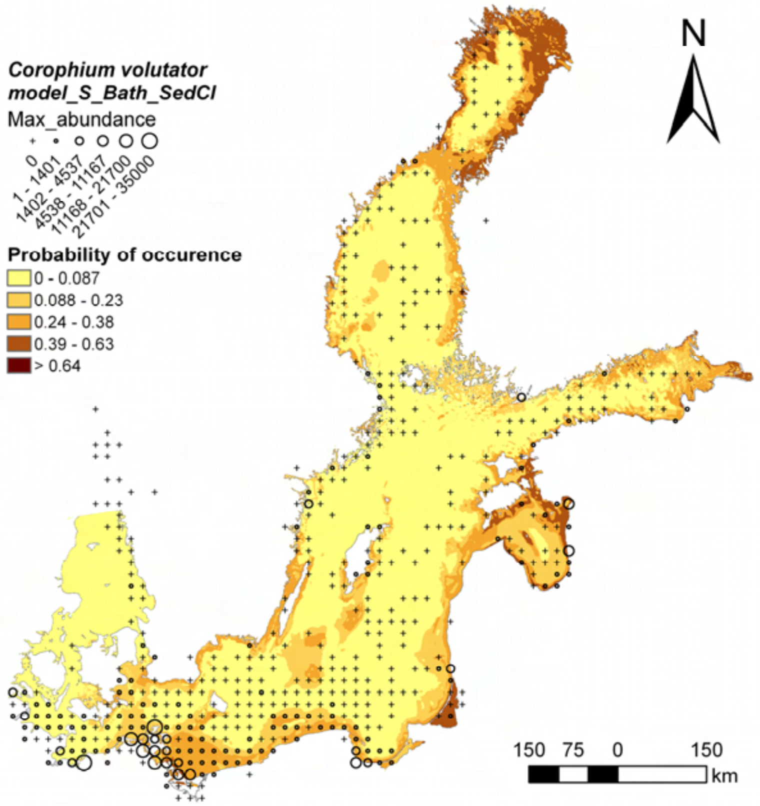
Appendix C

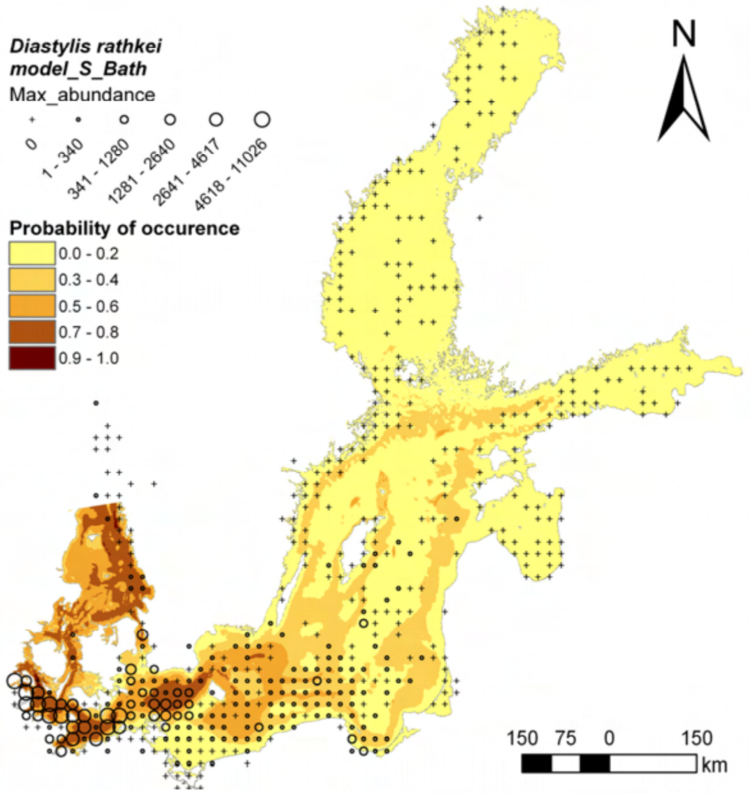
Modelling results compared with the observed data for the 19 macrobenthic species representative for the Baltic Sea. Symbols indicate the observed species abundance (ind/m^2) with their size corresponding to the value of abundance density (the maximal value observed for each cell of a regular 20 km grid is plotted in order to simplify visualization). Colour fields show the modelled probability of occurrence.

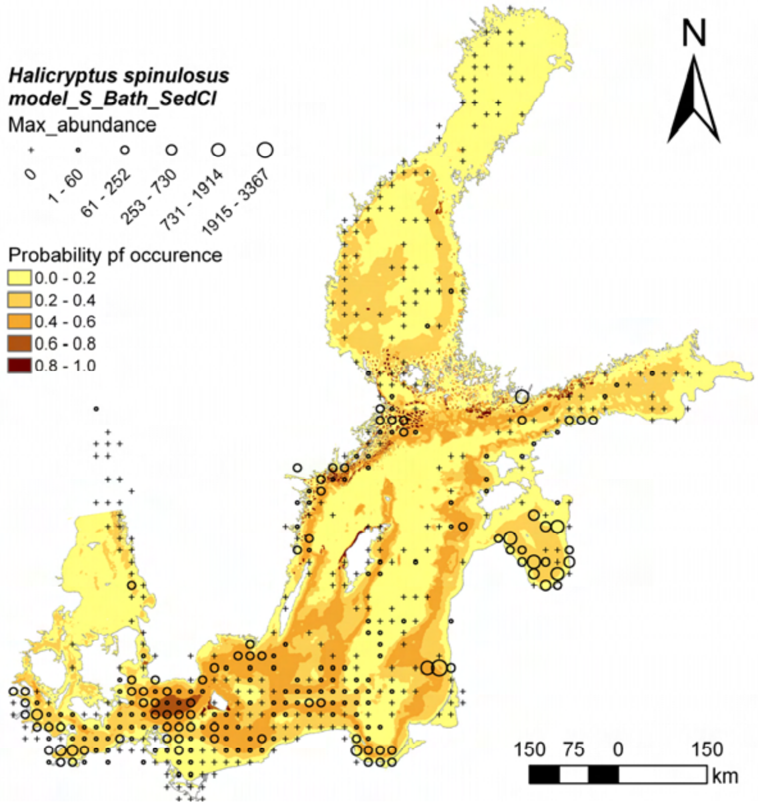

Arctica islandica
model_S_Bath_SedCl
Max_abundance
0
1 - 14
15 - 40
41 - 135
136 - 305
306 - 916
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150
75
0
150
km

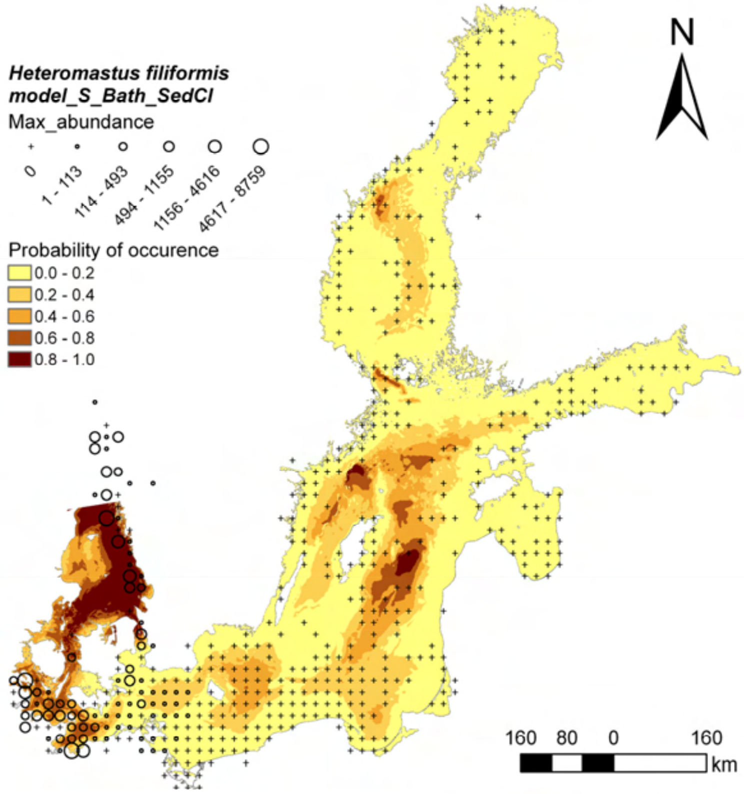

Astarte borealis
model_S_Bath_SedCl
Max_abundance
0
1 - 70
71 - 300
301 - 841
842 - 1439
1440 - 2420
Probability of occurence
0 - 0.055
0.056 - 0.19
0.2 - 0.37
0.38 - 0.67
> 0.68
N
150
75
0
150
km

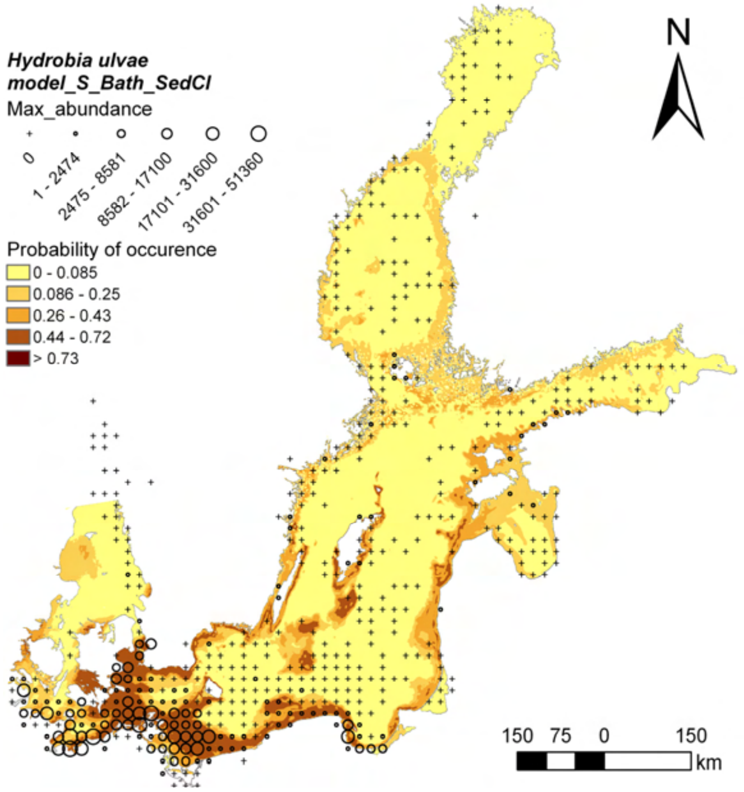

Bathyporeia pilosa
model_S_Bath_SedCl
Max_abundance
0
1 - 250
251 - 800
801 - 1775
1776 - 4397
4398 - 8830
Probability of occurrence
0 - 0.024
0.025 - 0.081
0.082 - 0.15
0.16 - 0.23
> 0.24
N
150
75
0
150
km

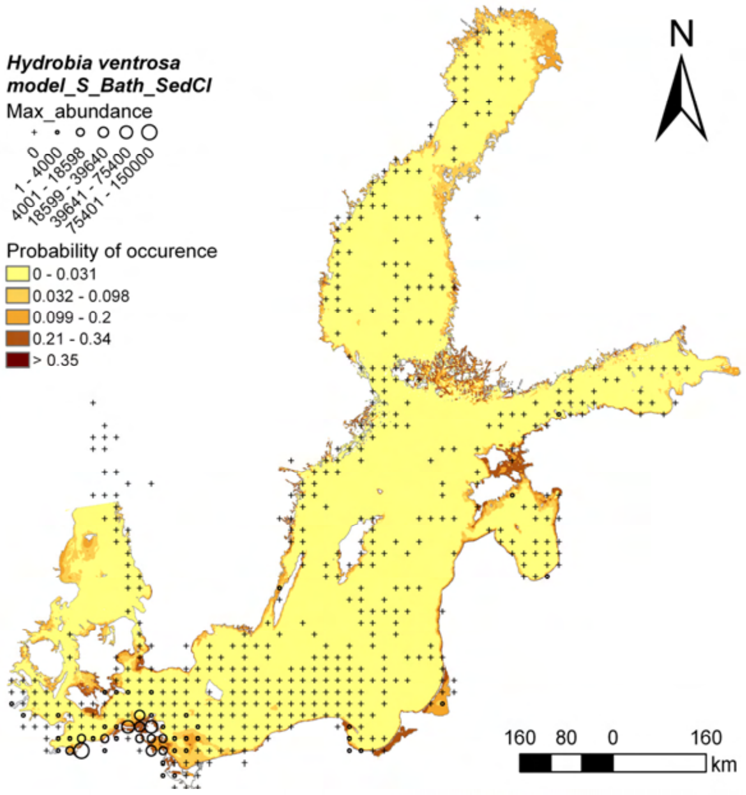

Corophium volutator
model_S_Bath_SedCl
Max_abundance
0
1 - 1401
1402 - 4537
4538 - 11167
11168 - 21700
21701 - 35000
Probability of occurence
0 - 0.087
0.088 - 0.23
0.24 - 0.38
0.39 - 0.63
> 0.64
N
150
75
0
150
km

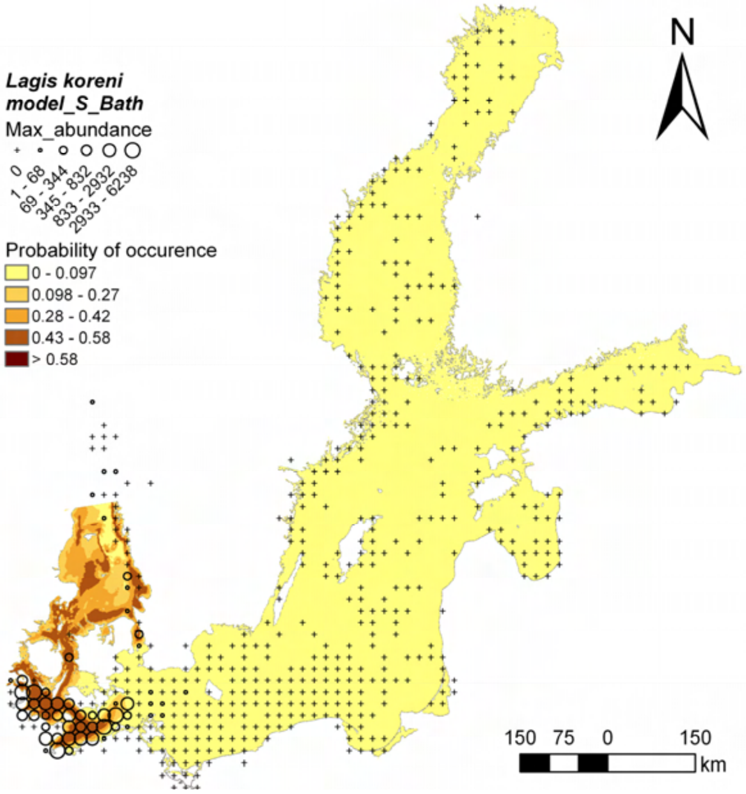

Diastylis rathkei
model_S_Bath
Max_abundance
0
1 - 340
341 - 1280
1281 - 2640
2641 - 4617
4618 - 11026
Probability of occurence
0.0 - 0.2
0.3 - 0.4
0.5 - 0.6
0.7 - 0.8
0.9 - 1.0
N
150 75 0 150
km

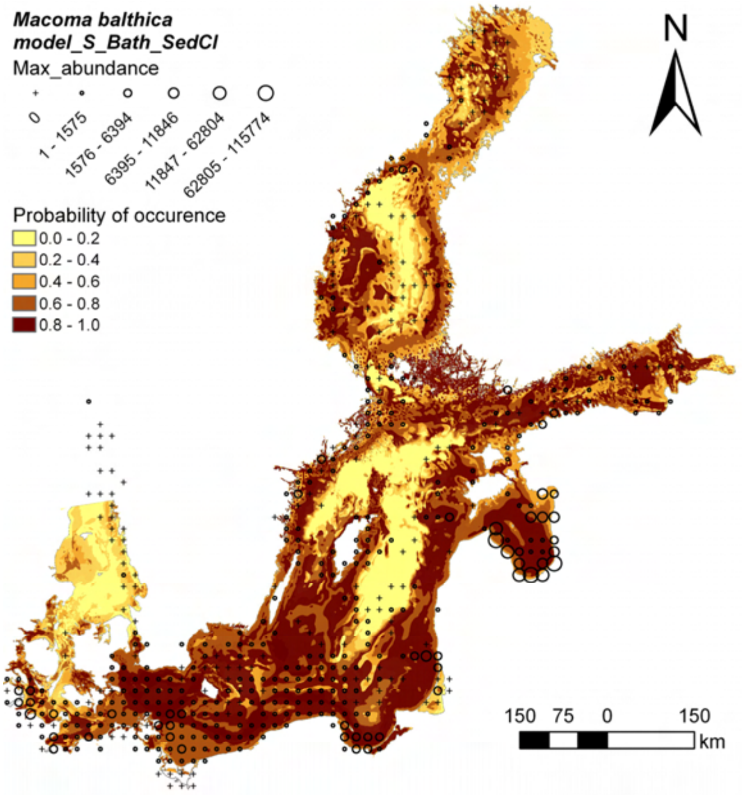

Halicryptus spinulosus
model_S_Bath_SedCl
Max_abundance
0
1 - 60
61 - 252
253 - 730
731 - 1914
1915 - 3367
Probability pf occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150 75 0 150
km

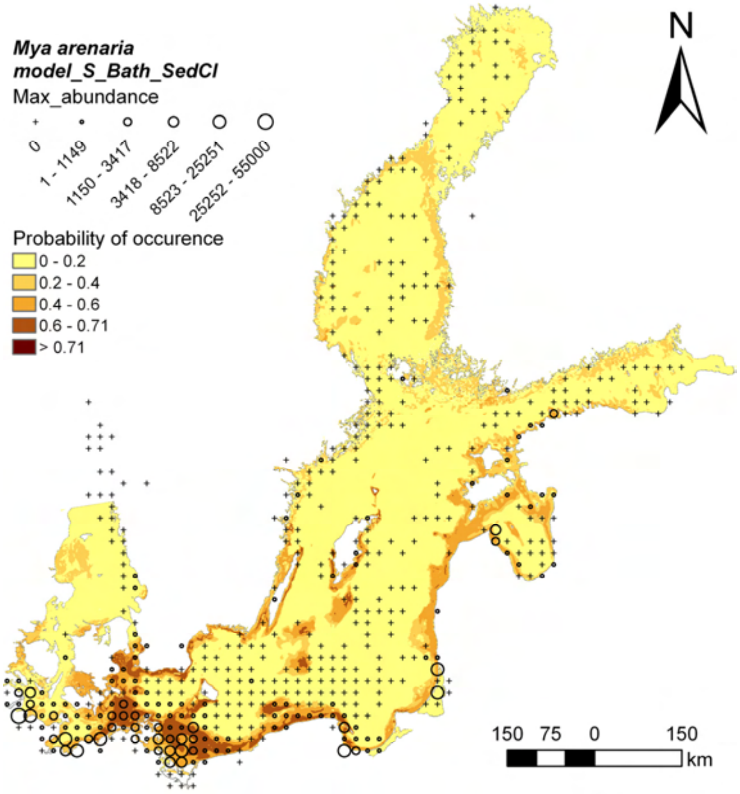

Heteromastus filiformis
model_S_Bath_SedCl
Max_abundance
0
1 - 113
114 - 493
494 - 1155
1156 - 4616
4617 - 8759
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
160 80 0 160
km

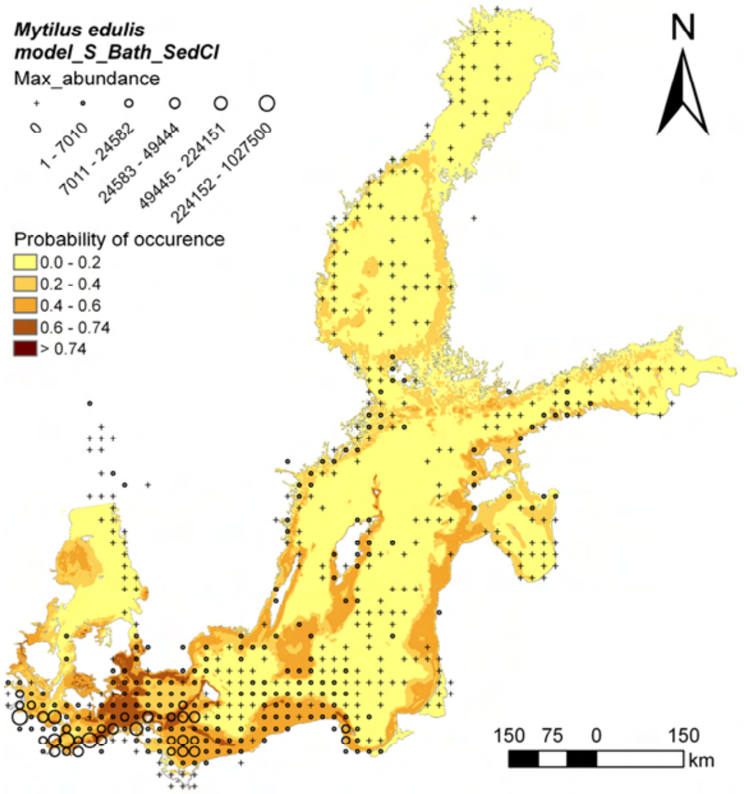

Hydrobia ulvae
model_S_Bath_SedCl
Max_abundance
0
1 - 2474
2475 - 8581
8582 - 17100
17101 - 31600
31601 - 51360
Probability of occurence
0 - 0.085
0.086 - 0.25
0.26 - 0.43
0.44 - 0.72
> 0.73
N
150 75 0 150
km

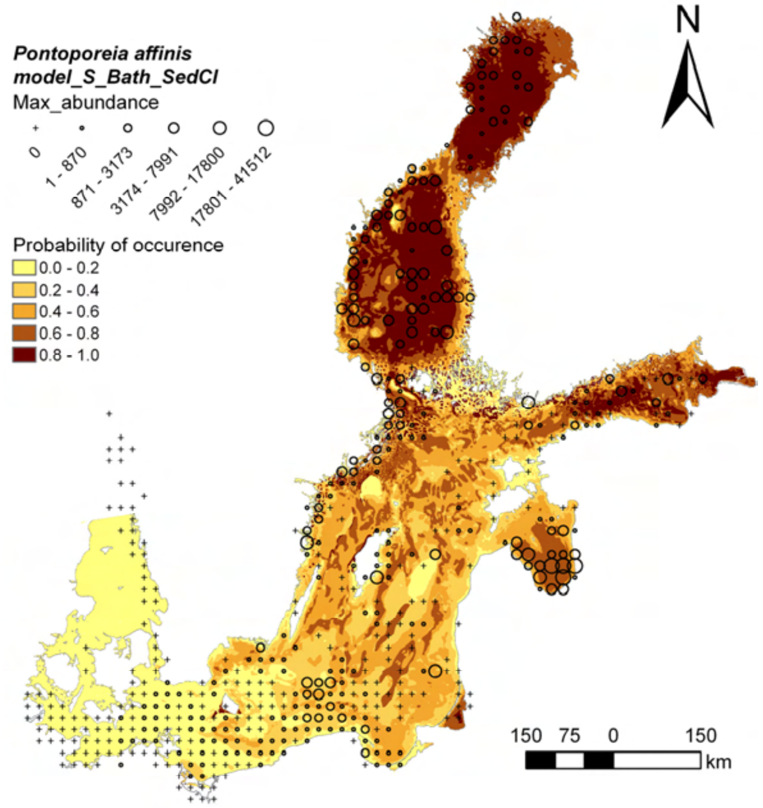

Hydrobia ventrosa
model_S_Bath_SedCl
Max_abundance
0
1 - 4000
4001 - 18598
18599 - 39640
39641 - 75400
75401 - 150000
Probability of occurence
0 - 0.031
0.032 - 0.098
0.099 - 0.2
0.21 - 0.34
> 0.35
N
160
80
0
160
km

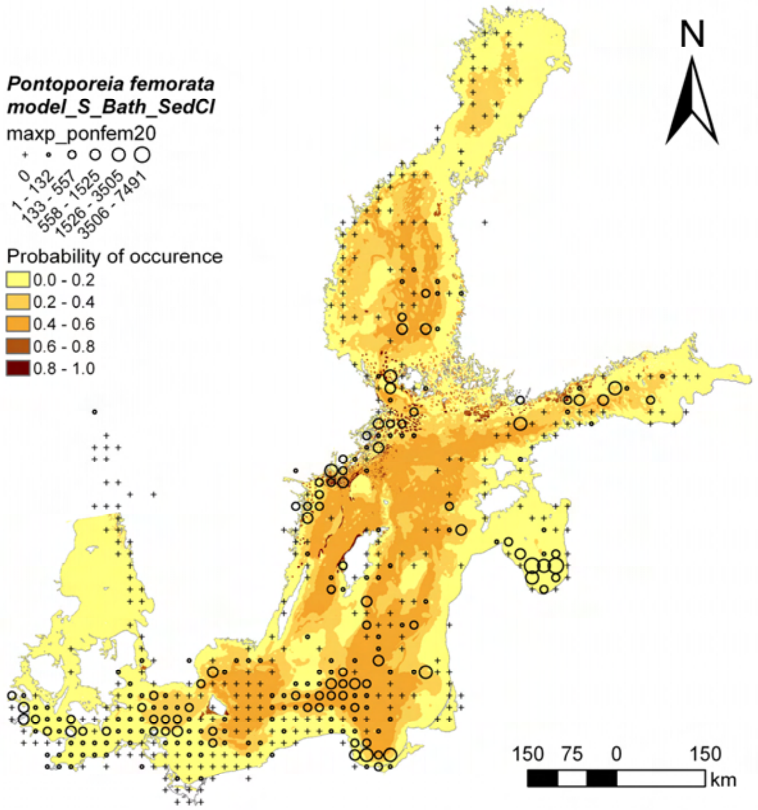

Lagis koreni
model_S_Bath
Max_abundance
0
1 - 68
69 - 344
345 - 832
833 - 2932
2933 - 6238
Probability of occurence
0 - 0.097
0.098 - 0.27
0.28 - 0.42
0.43 - 0.58
> 0.58
N
150
75
0
150
km

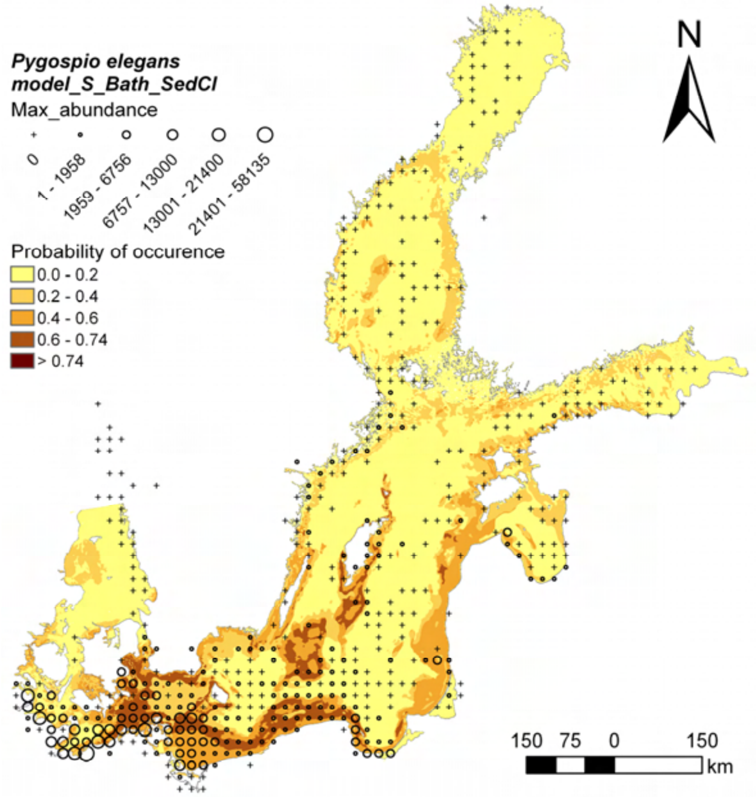

Macoma balthica
model_S_Bath_SedCl
Max_abundance
0
1 - 1575
1576 - 6394
6395 - 11846
11847 - 62804
62805 - 115774
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150 75 0 150
km

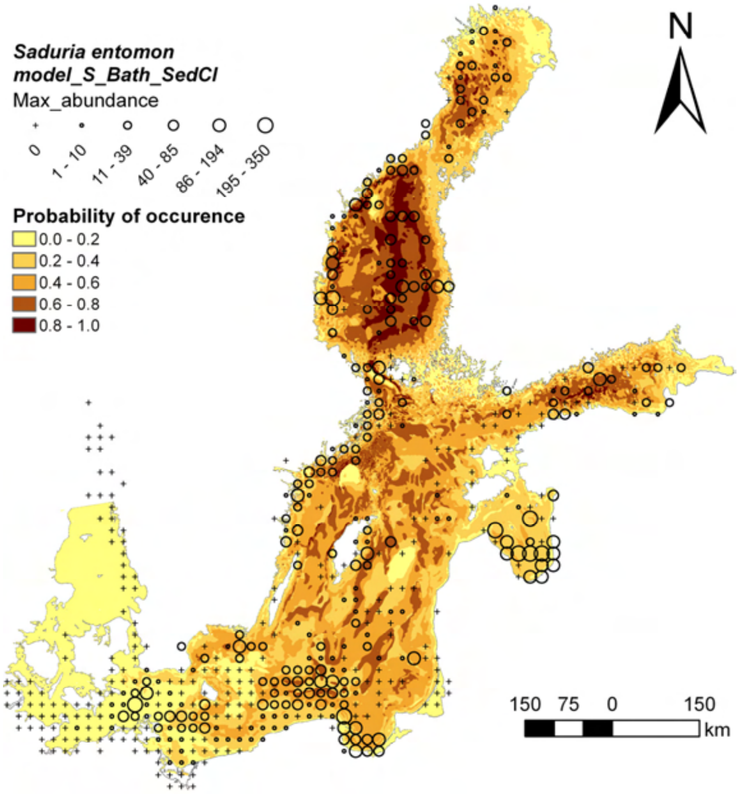

Mya arenaria
model_S_Bath_SedCl
Max_abundance
0
1 - 1149
1150 - 3417
3418 - 8522
8523 - 25251
25252 - 55000
Probability of occurence
0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.71
> 0.71
N
150 75 0 150
km

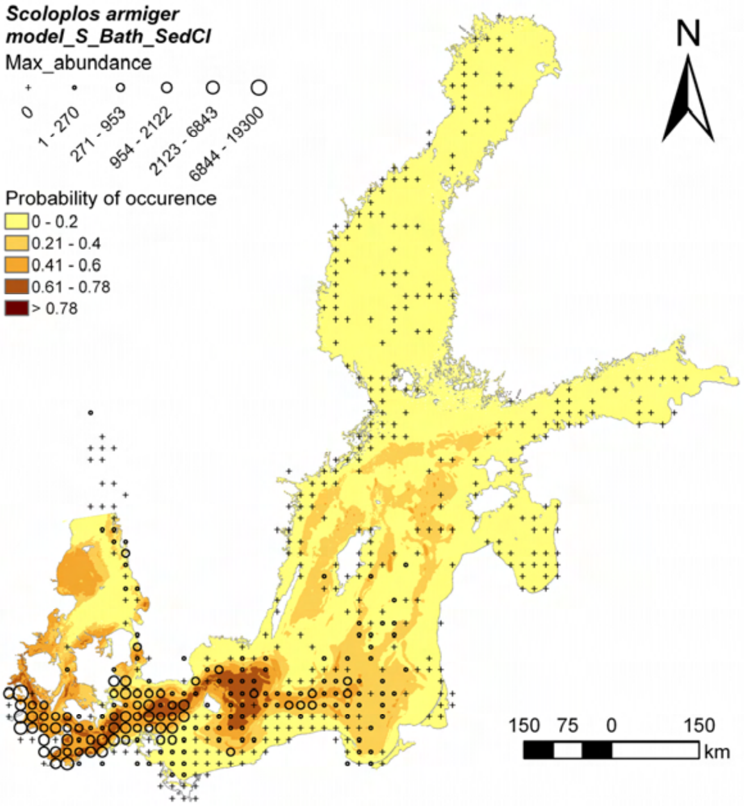

Mytilus edulis
model_S_Bath_SedCl
Max_abundance
0
1 - 7010
7011 - 24582
24583 - 49444
49445 - 224151
224152 - 1027500
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.74
> 0.74
N
150 75 0 150
km

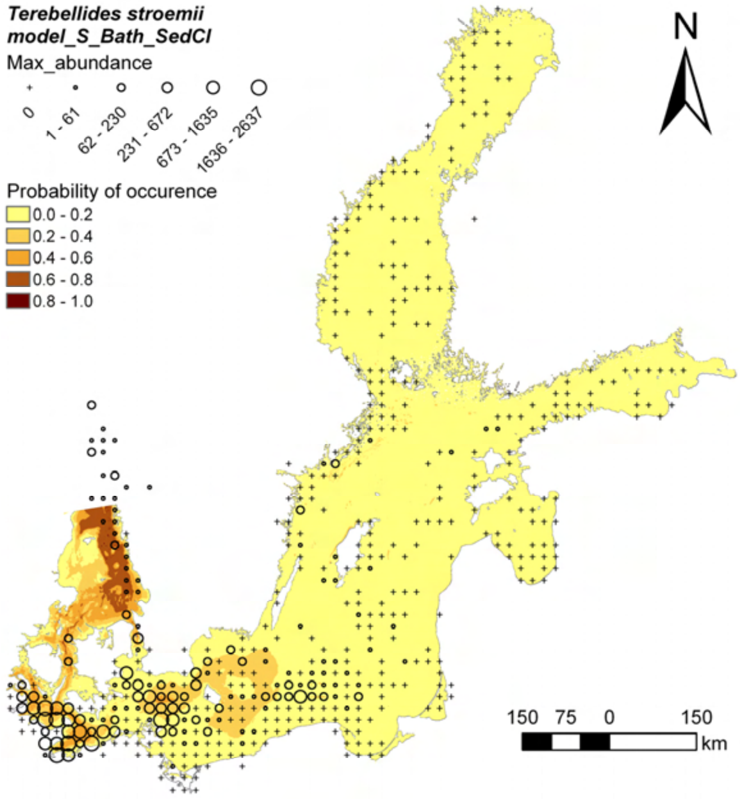

Pontoporeia affinis
model_S_Bath_SedCl
Max_abundance
0
1 - 870
871 - 3173
3174 - 7991
7992 - 17800
17801 - 41512
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150 75 0 150
km


Pontoporeia femorata
model_S_Bath_SedCl
maxp_ponfem20
0
1 - 132
133 - 557
558 - 1525
1526 - 3505
3506 - 7491
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150 75 0 150
km


Pygospio elegans
model_S_Bath_SedCl
Max_abundance
0
1 - 1958
1959 - 6756
6757 - 13000
13001 - 21400
21401 - 58135
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.74
> 0.74
N
150 75 0 150
km


Saduria entomon
model_S_Bath_SedCl
Max_abundance
0
1 - 10
11 - 39
40 - 85
86 - 194
195 - 350
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150
75
0
150
km


Scoloplos armiger
model_S_Bath_SedCl
Max_abundance
0
1 - 270
271 - 953
954 - 2122
2123 - 6843
6844 - 19300
Probability of occurence
0 - 0.2
0.21 - 0.4
0.41 - 0.6
0.61 - 0.78
> 0.78
N
150
75
0
150
km


Terebellides stroemii
model_S_Bath_SedCl
Max_abundance
0
1 - 61
62 - 230
231 - 672
673 - 1635
1636 - 2637
Probability of occurence
0.0 - 0.2
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1.0
N
150
75
0
150
km