Options and limitations of statistical modelling as a tool for understanding and predicting benthic functions in an area with high environmental variability

Dissertation

zur

Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von

Alexander Darr, geb. am 16.06.1975 in Melle

aus Rostock

Rostock, den 12.12.2013

urn:nbn:de:gbv:28-diss2015-0042-4

Gutachter:

1. Gutachter: Prof. Dr. Gerhard Graf,
Institut für Biowissenschaften, Universität Rostock

2. Gutachter: Prof. Dr. Heide Schulz-Vogt
Leibniz-Institut für Ostseeforschung Warnemünde

Datum der Einreichung: 13. Dezember 2013

Datum der Verteidigung: 04. Juli 2014

Summary (Zusammenfassung)

The increasing demand on assessment tools during the last decades was one of the main drivers for accelerating the research on understanding the consequences of changes in biodiversity for key ecosystem processes. The recently increasing interest in transferring this knowledge into function-based indicators can be attributed to both the awareness that functional changes are at least as important as changes in biodiversity and species composition and the unsatisfying performance of many species-based indicators, especially in areas with high natural stress such as brackish or transitional waters. Besides the limited knowledge on the autecology of many invertebrate species, missing statistical tools to both quantify the function on a spatial scale and detect shifts in functional diversity or functional composition still is the main limitation. Especially areas with natural gradients such as estuaries and transitional waters remained underrepresented in studies analysing benthic functional composition.

The objective of this thesis is to investigate the potential functional changes of macrobenthic communities along the natural gradients in transitional waters using the effect of salinity gradient in the south-western Baltic Sea on soft-bottom communities. The results should provide a basis for later state assessment using functional parameters. Two different approaches were chosen: (1) the appraisal of functional diversity and composition under consideration of the whole benthic community and (2) mapping species biomass as a basic tool for quantification of single key functions.

The approaches were implemented in four independent case studies. The 1st case study provides an example for the challenge to use macrobenthic species as indicators in state assessment of transitional waters. The consistency of the substrate preferences of selected "sensitive" and "tolerant" species along the salinity gradient was investigated. As a first attempt, Quantile regression splines were used to quantify species responds to environmental parameter. It was concluded that the environmental preferences of macrobenthic species might severely differ between regions due to acclimatisation, genetic adaptation (including cryptic species) and different interspecific interactions. The result of this study clearly evidenced the drawbacks in the concept of static "sensitivity" values.

Changes in functional diversity and functional composition along a salinity gradient were analysed in the second case study. Biological traits analysis in combination with fuzzy correspondence analysis was used to detect pattern in functional composition of the benthic community in the south-western Baltic Sea. Rao's quadratic entropy was used as a proxy for functional diversity. It was shown that abundance-based and biomass-based approach led to different results in detected changes of functional diversity and redundancy. Despite the fact that the results were influenced by high variability, the shift of the functional composition along the salinity gradient was distinct. The dominance of ubiquitous species at least mathematically buffered the loss of highly specialised species in ecosystem functioning. It was followed by a gain in functional redundancy, increasing the robustness of the benthic ecosystem to environmental changes. But the potential capacity of the dominant species in brackish water to change for example their feeding behaviour might also affect the capability to detect occurring functional changes along other natural or anthropogenic induced gradients.

The second aim of the thesis was the development of quantitative distribution maps and their linkage with single benthic functions. In the presented case study the biomass distribution of common bivalves was modelled and mapped for the south-western Baltic Sea using Random Forest. Bivalves were chosen as they are the most important suspension feeding taxonomic group in soft-substrates of the southern Baltic Sea and are due to their longevity potentially more robust against long sampling periods than short-living group such as opportunistic polychaetes. In a second step, the filtering capacity of *Mya arenaria* in the study area was calculated by coupling separate biomass- and abundance-based models with the function provided by Riisgård & Seerup (2004). Hotspots with a filter-feeding capacity of 3-4 $m^3*m^{-2}*d^{-1}$ were identified in the southern Pomeranian Bay close to the mouth of River Oder. It was concluded that the development of quantitative species-distribution models provides an opportunity to fundamentally increase the knowledge on benthic functions, but that its applicability is still limited due to the limited number of available equations linking basic functions with species density or biomass.

Finally, it was concluded that despite the discussed limitations of the actual procedures, both chosen approaches provide a sound basis for a wide variety of application in state assessment, nature conservation and ecosystem analysis.

Contents

Figures

page

Tables

List of frequently used abbreviations

afdm	Ash-free dry matter
AIC_c	Aikike Information Criterion, correction for small datasets
BNatSchG	Bundesnaturschutzgesetz
BQI	Benthic Quality Index
BTA	Biological Traits Analysis
CART	Classification and Regression trees
EEZ	Exclusive Economic Zone
FD	Functional diversity
FR	Functional redundancy
GAM	General additive models
GIS	Geographic Information System
GLM	Generalized linear models
HD	Habitats Directive (92/43/EWG)
HELCOM	Helsinki Commission
IOW	Leibniz Institute for Baltic Sea Research
LOI	Loss on ignition
MarBIT	Marine Biotic Index Tool
MarLIN	Marine Life Information Network
MSFD	Marine Strategy Framework Directive (2008/56/EG)
OM	Organic matter
psu	Practical salinity unit
QRS	Quantile Regression splines
RF	Random Forests
SDM	Species Distribution Model
UTM	Universal Transverse Mercator coordinate system
WFD	Water Framework Directive
WGS	World Geodetic System

1. General Introduction

Understanding the consequences of changes in biodiversity for key ecosystem processes recently became one of the most prominent objections in marine ecological research (Ieno et al. 2006). The increasing scientific interest on this task was illustrated by Miloslavich & Klein (2009) who analysed the title keywords of all submissions to the World Conference on Marine Biodiversity (WCMB) in Valencia 2008. Almost all of the most often used keywords referred to this research theme with a special focus on the "benthic" part of marine ecosystems (Figure 1).

One of the main external drivers for this development was the increasing pressure on the marine environment during the last decades. The increasing usage of marine resources, followed by a rising frequency and intensity of economically and ecologically negative consequences of human activities such as the break-down of fish stocks or harmful algae blooms, induced a raised interest of stakeholders in state assessments of the marine environment. While the assessment by species-based approaches was preferred for many years, the recently increasing interest on function-based indicators can be attributed to both (1) the awareness that functional changes are at least as important as changes in biodiversity and species composition and (2) the unsatisfying performance of many species-based indicators, especially in areas with high natural stress such as brackish or transitional waters (Mouillot et al. 2006, Osowiecki et al. 2008, Fleischer & Zettler 2009).

Despite the current rush for function-based indicators, the observation of functional composition of benthic communities and its changes along natural or anthropogenic gradients are not new. Pearson & Rosenberg (1978) already described functional changes in their concept of macrobenthic succession as an effect of organic pollution. Almost at the same time, Word (1979) developed one of the first function-based indicators, the Infaunal Trophic Index (ITI), relating the feeding type of endobenthic invertebrates to changes and degradation of environmental conditions. Nevertheless, the functional approach remained underrepresented in the assessment of marine waters for many years. Besides the limited knowledge on the autecology of many invertebrate species, missing statistical tools to both quantify the function on a spatial scale and detect shifts in functional diversity or functional composition was the main drawback for the usage of benthic functions in state assessment.

Although much effort was made during the last decade, some of the main problems are still unsolved. Especially areas with natural gradients such as estuaries and transitional waters remained underrepresented in studies analysing benthic functional composition. Within these areas anthropogenic stress may have the same characteristics as natural stress. Elliott & Quintino (2007) stated that this resemblance of natural and anthropogenic stress may lead to higher difficulties to detect human impact on the ecosystem. Therefore, understanding the effect of natural gradients on benthic functioning can be considered to be a fundamental requirement for a function-based state assessment of the benthic ecosystem in transitional waters.

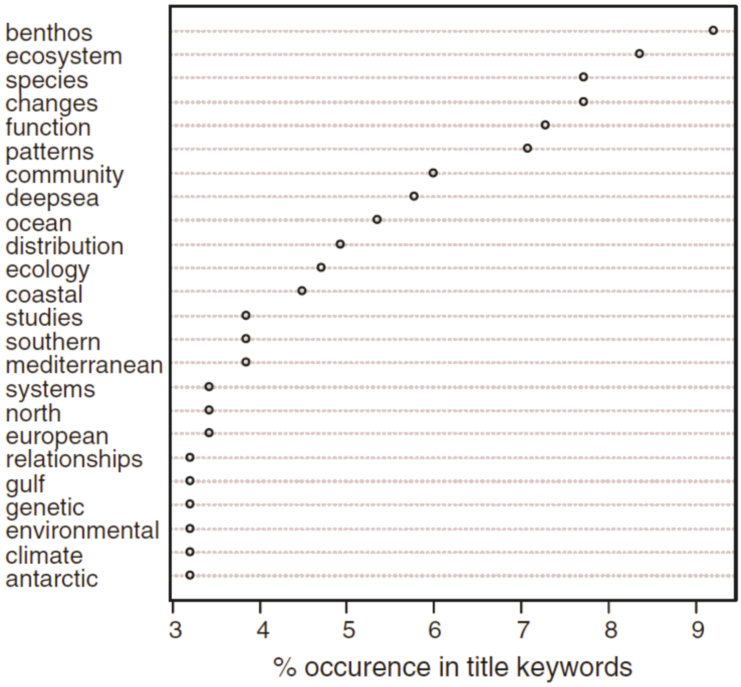


Figure 1: Occurrence of title keywords in the submissions to the 1st WCMB 2008 (taken from Miloslavich & Klein 2009)

1.1. The functional role of macrozoobenthos

The marine macrozoobenthos comprises almost all larger Phyla of the kingdom Animalia, such as the basic groups Porifera and Cnidaria and the higher systematic groups like Echinodermata and Chordata. It therefore provides a manifold variety of evolutionary strategies to cope with and to influence its surrounding environment. This environment is characterised by the water body and the sediment as both compartments play major roles within the life strategy of most benthic species. The water body is mainly used as food source and for active or passive dispersal in different life stages, e.g. eggs, larvae or adults. In contrast, the benthic compartment serves among others as static substrate for sedentary species, as shelter and also as food source.

Most activities of macrobenthic species happen in the interface between water body and substrate and influence the marine nutrient turnover by coupling the cycles of both compartments, known as the benthic-pelagic coupling (e.g. Graf 1992). It plays a major role in marine nutrient cycling especially in shallow coastal areas where the volume of the pelagic biotope is restricted. Besides the involvement of benthic invertebrates in the marine food web, the influence of macrobenthic species on benthic-pelagic coupling mainly consists of two activities: biodeposition and bioresuspension (Figure 2).

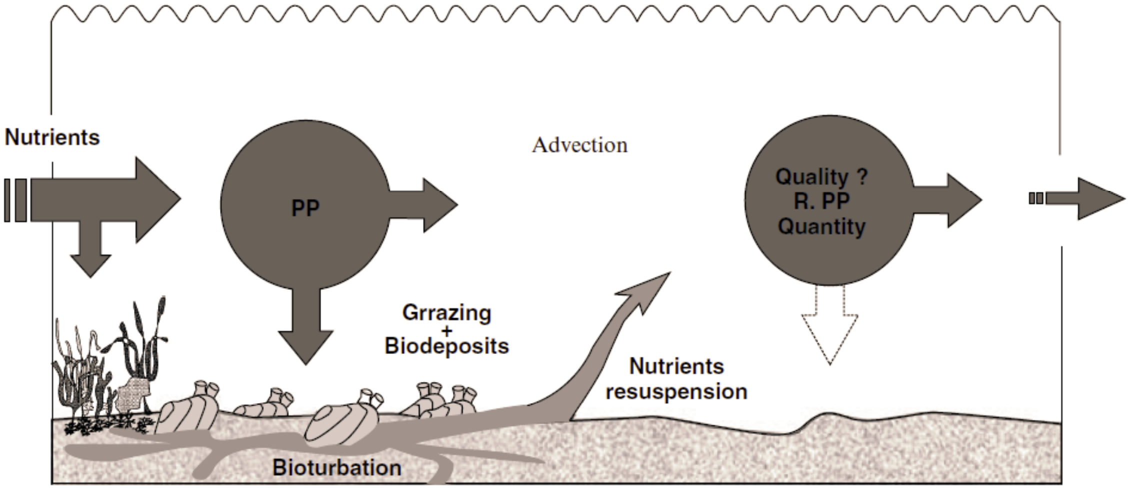


Figure 2: Biodeposition and bioresuspension are the most important functions of macrobenthic organisms in benthic-pelagic coupling. PP: primary production, R.PP: regenerated primary production (taken from Grall & Chauvaud 2002).

Resuspension of benthic nutrients can be facilitated by macrobenthic organisms either actively by excretion into the water column or passively by bioirrigation with the latter being by far more important. The construction of tubes and permanent burrows amplifies the surface area for nutrient exchange between the water column and the sediment. The active mixing of the sediment (bioturbation) has vast effects on the biogeochemistry of the sediment including oxygen, pH and redox gradients, substrate composition and sorting, pollutant and nutrient release, bacterial activity and composition and macrobenthic diversity (Queirós et al. 2013 and citations therein). The changes in the biological communities (microbial and macrobenthic) might multiply or counteract these effects (e.g. Bonaglia et al. 2013).

Biodeposition is mainly forced by suspension feeders, capturing food (plankton, detritus) from the water column and depositing the indigested organic material as faeces or pseudo-faeces on or in the sediment (Norkko et al. 2001). This might have significant effects both on the substrate characteristics and the food-web of the pelagic zone. For example Norkko et al. (2001) detected that large individuals of a suspension-feeding bivalve species of the genus *Atrina* significantly increased sedimentation rates, leading to an enrichment of the surrounding sediments with nitrogen and organic carbon. As the deposited organic material is an important food-source for deposit-feeders, the organic enrichment in the provided example was followed by an increase in abundance and diversity of the macrobenthic community. This correlation was pronounced especially in areas with little natural sedimentation (Norkko et al. 2001).

Although most bivalves species are considerably smaller than the members of the genus *Atrina* (length 30-40 cm), suspension-feeding species in general and bivalves in particular play a key role in soft-bottom communities throughout the oceans. Benthic bivalves modify their physical habitat and strongly influence the processes of nutrient flux between the near-bottom water and the sediment. Suspension feeders might even control the pelagic primary production by significantly reducing phytoplankton density (Grall & Chauvaud 2002 and citations therein). A suspension-feeder regulated system is defined by a high filtration capacity (large bivalve biomass), a long residence time (slow water exchange) and most-often a comparably small water volume (shallow areas, Figure 3).

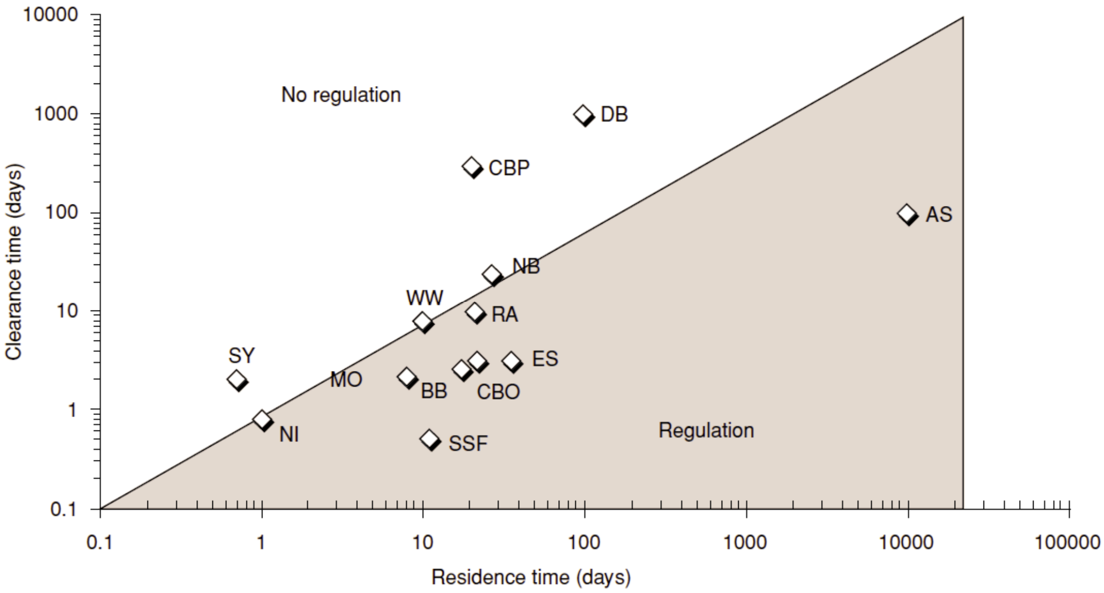


Figure 3: The ratio between residence time and filtration capacity can serve as a proxy for potentially suspension-feeder regulated coastal ecosystems (abbreviations: different coastal areas, e.g. SY: near Sylt, taken from Grall &Chauvaud 2002).

1.2. Objective and outline of the thesis

The objective of this thesis was to investigate the potential functional changes of macrobenthic communities along the natural gradients in transitional waters using the effect of salinity gradient in the south-western Baltic Sea on soft-bottom communities. The results should provide a basis for later state assessment using functional parameters.

Two different approaches were chosen:

(1) the appraisal of functional diversity and composition under consideration of the whole benthic community and

(2) mapping species biomass as a basic tool for quantification of single key functions.

The approaches were implemented in independent case studies. Each case study starts with an introductory section providing the background for the aim of the individual study. As the statistical approaches differed fundamentally among the individual studies, the used methods are described in detail within the case studies. An overview of the study area, the methods used for sampling and the laboratory work as well as a brief background for the statistical approaches are given in the sections foregoing the case studies.

The 1st case study provides an example of the challenge to use macrobenthic species as indicators in state assessment of transitional waters. The consistency of the substrate preferences of selected "sensitive" and "tolerant" species along the salinity gradient was investigated. As a first attempt, Quantile regression splines (QRS) were used to quantify species responds to environmental parameter.

The 2nd part represents the first approach to analyse the potential changes of functional composition and diversity along the salinity gradient. Functional composition was analysed using a fuzzy coding-approach by combining a biological traits table with fuzzy correspondence analysis. Functional diversity and associated measures were calculated using Rao's quadratic entropy.

The second target was to develop quantitative distribution maps and to link them with single benthic functions, which was approached in two separate case studies. In the 3rd case study the biomass distribution of common bivalve species was modelled and mapped for the south-western Baltic Sea using Random Forest algorithms. Bivalves were chosen as they (1) are the most important suspension feeding taxonomic group in soft-substrates of the southern Baltic Sea and (2) due to their longevity are potentially more robust against the varying sampling periods represented by the available data than short-living group such as opportunistic polychaetes.

In the last case study, the modelled biomass distribution of *Mya arenaria*, one of the most important soft-bottom bivalves, was linked with its filtering capacity as an example of the quantification of specific benthic functions.

The results and conclusions of the individual case studies were finally summarized and merged in the synthesis (chapter 7) including an outlook for the future challenges.

2. Material and Methods

2.1. Study area

Although the first case study also includes data from the North Sea, the analyses and conclusions of this thesis were mainly based on data from the south-western Baltic Sea between the Bay of Kiel and the Island of Usedom (Figure 4).

As the Baltic Sea is one of the largest brackish-water systems in the world, salinity is one of the major drivers for macroscopic benthic biodiversity (e.g. Remane 1934, Bleich et al. 2011, Zettler et al. 2013b). The microtidal Baltic Sea is comparably shallow, comprising a surface area of around 415 000 km² and a water volume of 21 760 km³ (HELCOM 2010). The water exchange with the North Sea through Skagerrak and Kattegat is limited by the narrow and shallow Danish Belts and the Sound. Also the underwater topography of the Baltic is characterised by sills hindering the inflow of saline waters into the Baltic basins. While the marine water enters the Baltic from the west, the discharge of the largest rivers (e.g. Neva, Vistula) and a net precipitation ratio cause a surplus of freshwater supply in the north-eastern part. This unbalanced water supply leads to a distinct horizontal salinity gradient with almost euhaline conditions in the Kattegat and almost limnic conditions in the Gulf of Bothnia and in the Gulf of Finland. The gradient is particularly pronounced in surface waters, whereas the bottom salinity, which is more relevant for benthic life, also depends on the topography. As the saline water possesses a higher density, inflows pass off in the bottom layer, replacing the stagnant water body. Therefore, salinity is in general higher in the deeper part of a particular basin.

The stratification of the water body due to thermohaline differences is stable in the deeper parts (i.e. in basins) where wind-induced mixing does not reach the sea bottom. By contrast, in shallower areas (e.g. Bay of Kiel, Mecklenburg Bay) stratification is maintained as long as strong winds hold off. In stagnation periods, the available oxygen is depleted by respiration and geochemical processes. As the exchange with oxygen-rich surface water is restricted, the near-bottom water layer first becomes hypoxic and at a later stage anoxic. Re-oxygenation may be driven in shallow areas by full mixing due to strong winds whereas the bottom layer in the deep basins has to be replaced by oxygen-rich saltwater inflows. Major saline water inflows happen irregularly and became even more infrequent since the end of the 1970s (Feistel et al. 2008). The last major inflow appeared in January 2003, ventilating large parts of the Baltic Sea including the Bornholm Basin, Gdansk Basin and the central Baltic Sea (Feistel et al. 2008 and citations therein).

Both salinity gradients (from west to east and from shallow to the deep) are pronounced within the south-western Baltic Sea (Figure 5A). Mean salinity of the near-bottom water body declines from 20-25 psu in the Bay of Kiel towards 7 psu in the Pomeranian Bay. The Darss Sill is acting as one of the major inhibitors of water exchange between the Danish straits and the southern Baltic. A strong seasonal stratification can be found in the western part. Mean salinities in the shallowest parts of Bay of Kiel and Fehmarnbelt are very similar to the salinity conditions in the Arkona Basin. But for benthic invertebrates, at least as important as a permanent reduced salinity for benthic invertebrates is the temporal variability (Atrill 2002), which is low in most parts of the study area in comparisons to tidal influenced estuaries. Salinity conditions may change dramatically within inflow events, but usually vary significantly only along the halocline and in the area between the Mecklenburg Bay and the Arkona Basin (Figure 5B). The strong seasonal stratification is frequently followed by hypoxic or anoxic conditions in the deeper parts of the Bay of Kiel, the Fehmarnbelt, the Mecklenburg Bay (including the Bay of Lübeck) and the Arkona Basin. Intensity and duration of hypoxia strongly depend on the productivity of the spring bloom, the frequency of storms and summer temperature, but in general lasts for several months especially in parts of the Bay of Lübeck (Figure 7).

Another fundamental driver for the distribution of macrobenthic invertebrates is substrate composition (Glockzin et al. 2009, Gogina et al. 2010a). The composition of surface sediments in the south-western Baltic mainly results from postglacial processes and the actual exposure to bottom currents. Shallow areas along the shore and on top of offshore glacial elevations are mainly abrasion areas characterised by a mosaic of rocks, till, gravel and coarser sand (Figure 6). Especially parts of the Fehmarnbelt area, the Kadet Trench/ Darss Sill area and the Adlergrund are characterised by a high small-scale variability including also fine substrates in sheltered areas. By contrast, the Pomeranian Bay is dominated by almost monotonous fine sand with an increasing portion of silt and organic material along the depth gradient towards the Sassnitz channel and the Arkona Basin. The basins and deeper parts of the bays are characterised by muddy sediments with accumulating organic matter.

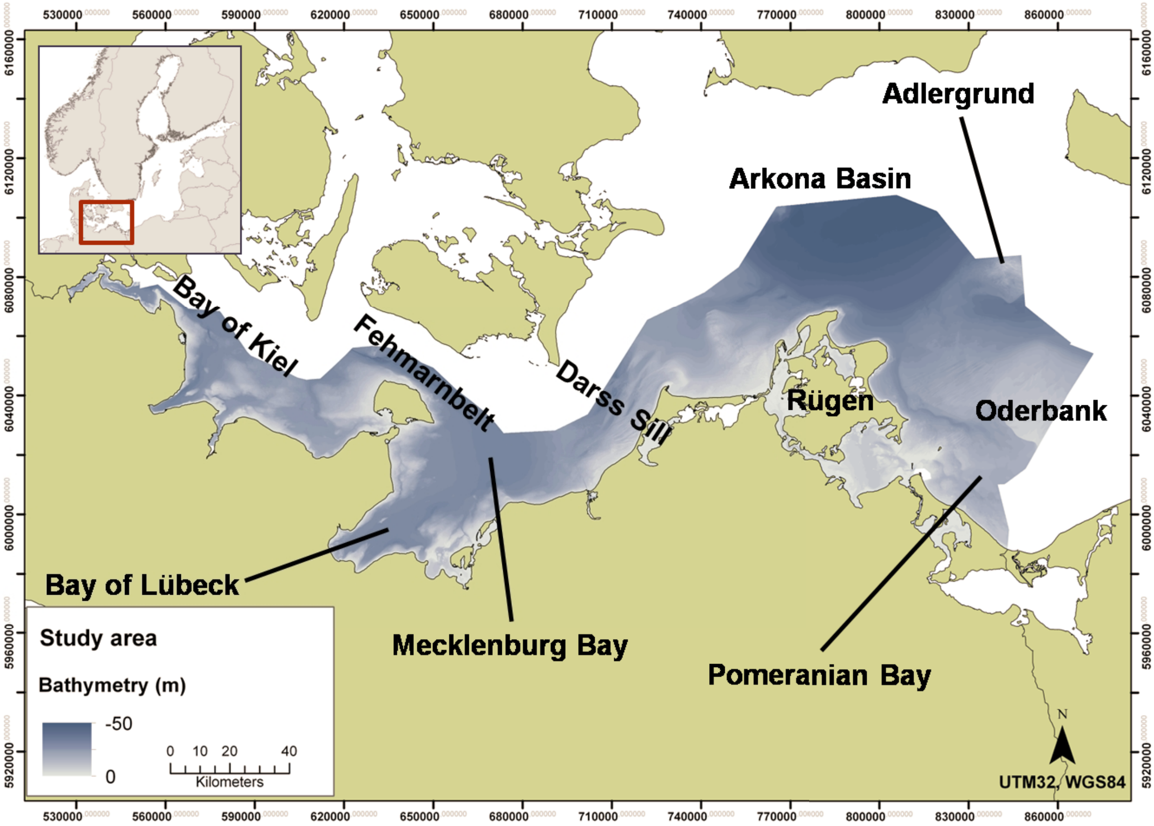


Figure 4: Study area displaying the location of sea areas frequently mentioned in the text.

[Bathymetric data provided by Tauber (2012)]

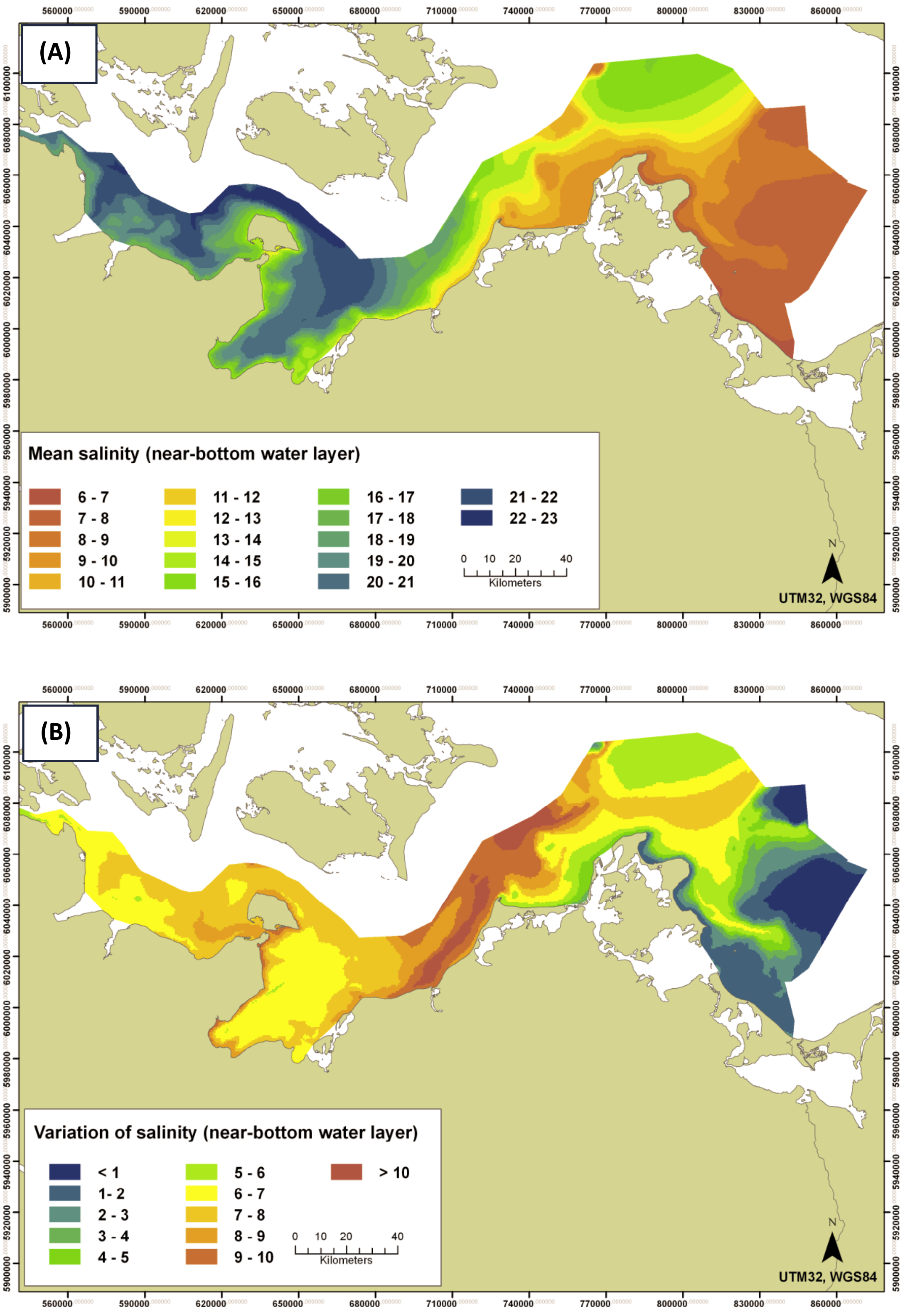


Figure 5: Mean (A) and variability (B) of salinity in the study area. Variability is expressed as the difference between 95% and 5%-percentile. Data were provided by Klingbeil et al. (2013).

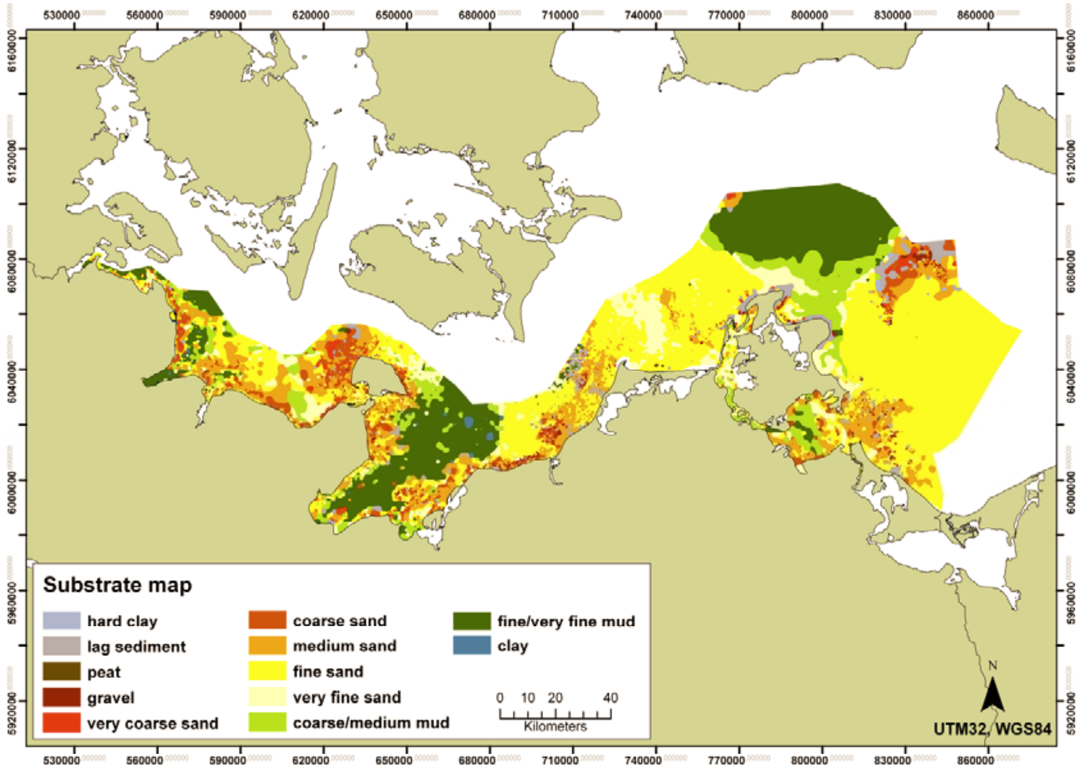


Figure 6: Simplified substrate map of the study area, based on Tauber (2012).

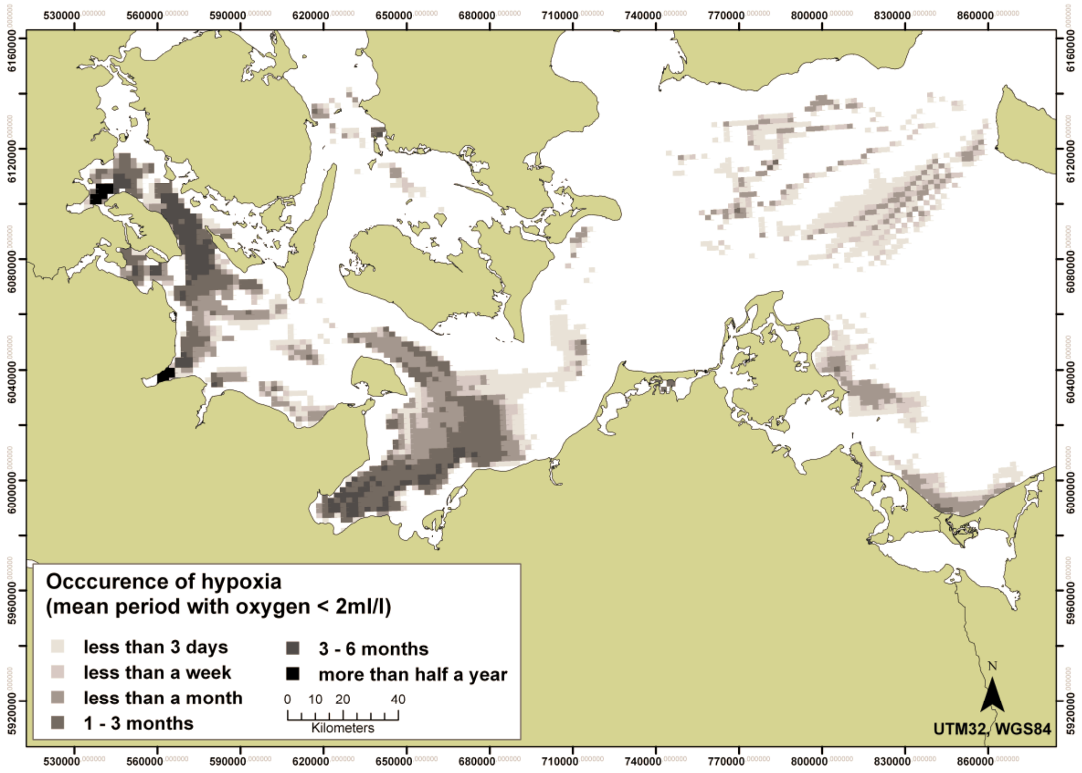


Figure 7: Mean period of hypoxia (oxygen < 2 ml*l^{-1}) modelled for the period 2004-2009 (provided by R. Friedland)

2.2. Acquisition of data: sampling and laboratory work

2.2.1. Generation of biotic data

For all intended analyses and modelling exercises, a large and consistent data base, including all necessary biotic and abiotic parameters, is a fundamental prerequisite. The data base of the working group “Ecology of benthic organisms” at the IOW comprehends more than 800 stations sampled between 2003 and 2012. The amount of available data was amended during seven cruises between 2009 and 2013, comprising more than 250 additional stations (Table 14 and Figure 44 in the appendix). The positioning of the stations depended on the objectives of the underlying project as follows:

- FEMA: Baseline study Fehmarnbelt fixed linked,
- Natura 2000: development of a monitoring concept for annex 2 habitats (habitats directive HD) in the German exclusive economic zone
- Cluster 4: implementation of an integrative monitoring and assessment concept for benthic biotopes considering the demands of the HD, the MSFD, BNatSchG and HELCOM
- Cluster 6: habitat mapping of the German EEZ with a focus on the Marine Protected Areas

Standard procedure included the sampling of three replicates at each station using a Van-Veen grab (70-75 kg; 0.1 m^2; 10-15 cm penetration depth, Figure 8A). Samples were washed through 1 mm mesh-size following HELCOM-guidelines (HELCOM 2013a, Figure 8B) and preserved in approximately 4 % buffered formaldehyde-seawater solution for quantitative appraisal in the laboratory (Figure 8C).

A small dredge (type: Kieler Kinderwagen, opening 100 cm, mesh-size: 10 mm, Figure 9A) was additionally used at each station to obtain an overview of epibenthic species. The dredge was trawled by wind-drift for one minute on mud and stony bottoms up to five minutes on clear sand. The samples were analysed qualitatively as they cannot be related to the sampled area. Furthermore, additional information on epibenthic species and habitat structures were obtained by underwater-videos (Figure 9B, C). However, only quantitative results from the grab sampling were used in the following analyses if not stated otherwise.

All quantitative samples were sorted in the laboratory using a stereomicroscope with 10-40 times magnification. The specimens were separated from the remaining substrate, identified to the lowest possible taxonomic level (see Table 15 in the appendix for taxonomic references of the species mentioned in this thesis), counted and weighted (fresh mass). Ash-free dry mass (afdm) was calculated from fresh mass using conversion factors generated from measurements conducted by the IOW-working group during the last years. Sorting, weighting and identification of the specimen was mainly carried out by the technical staff of the IOW. For all analyses, average abundance and biomass values were calculated for each station and extrapolated to 1 m².

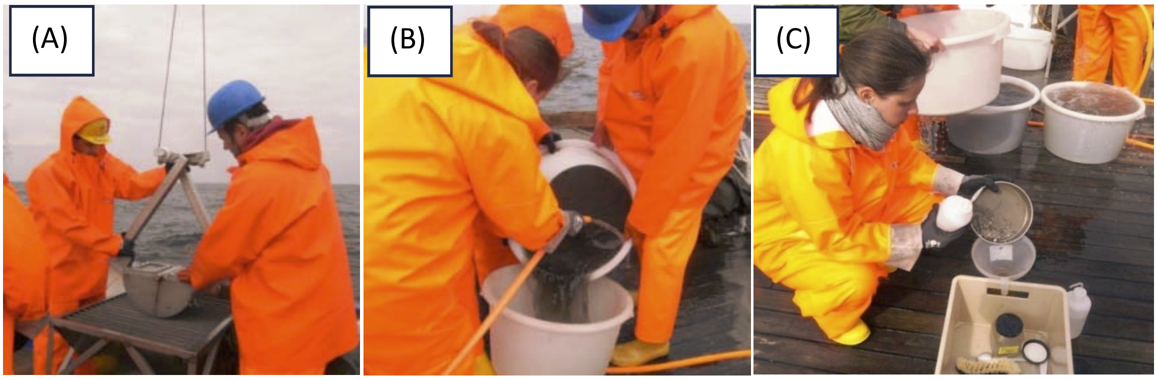


Figure 8: Process of benthos-sampling onboard: (A) sampling with standard gear (Van-Veen grab), (B) sieving over 1 mm, (C) storage into Kautex-bins

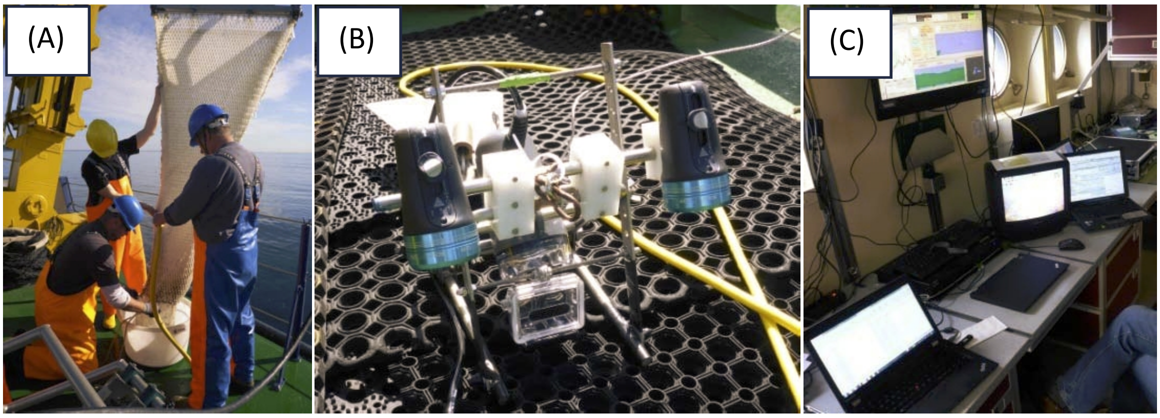


Figure 9: Additional information on epibenthic species and habitat structures were gathered by using a dredge (A) and a towed underwater video-system (B: camera, C: control unit)

2.2.2. Generation of abiotic data

For the determination of sediment characteristics, an additional sample was taken at each station. Grain size distribution was analysed following DIN 66165 using a RETSCH sieving machine (mesh scale: 63 µm, 90 µm, 125 µm, 180 µm, 250 µm, 355 µm, 500 µm, 710 µm, 1000 µm, 1400 µm, 2000 µm) for sandy and gravelly substrates and a CILAS 1180 L Laser Particle Analyzer for muddy samples. In general, median grain size (d50) was used as a proxy for substrate composition. The organic content of the sediment was analysed as loss on ignition (in % dry weight) by heating the material for 5 h at 500°C in a muffle furnace. Additionally, sediment characteristics of each biotic replicate sample were documented as differences between the samples at an individual station might occur especially in heterogeneous areas.

Water parameters such as salinity, water temperature and water depth were measured during the sampling event using ship-based CTD (Figure 10A, B). Oxygen content was measured by automatic Winkler-titration using a Metrohm TITRINO plus 848-System (Figure 10C).

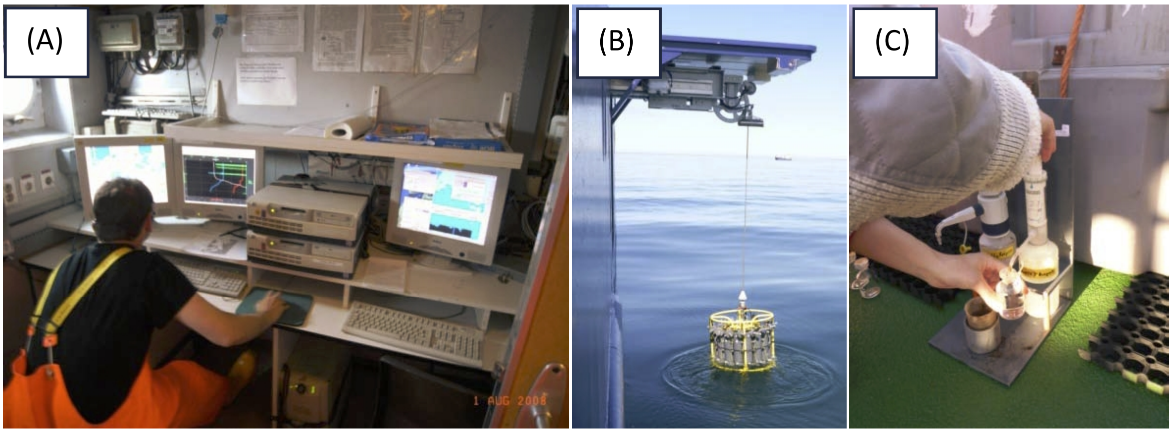


Figure 10: Salinity and water temperature were measured at each station using a CTD-unit (A: control unit, B: rosette). Oxygen concentration was measured by automatic Winkler-titration (C).

2.3. Analysis & Modelling techniques

As the objectives differed considerably between the case studies, different statistical methods were used. Therefore, the exact procedures were embedded in the relevant sub-chapters of the individual case studies. The following sections present a brief theoretical framework of the tools which were used to determine functional diversity in the 2nd case study and for the different modelling techniques analysing the relationship between species distribution and environmental parameters in the residual case studies.

Most analyses of this thesis were performed under R environment, an open-source computer language (R development core team 2012, http://www.r-project.org/). The required statistical methods and modelling techniques were either included in the basic program or provided by additional packages also named in the relevant sections of the individual case study. The scripts used for this thesis were self-coded, but based on the examples given in the manuals of the used packages and various textbooks (Zuur et al. 2007, Zuur et al. 2009, Zuur et al. 2012).

All maps in this thesis were drawn using ESRI ArcGIS 10®. The maps refer to the UTM-system, zone 32 north and are projected on the WGS 84 ellipsoid.

2.3.1. The concept of Functional Diversity

While species diversity is a measure of the number of species inhabiting an area (species richness) and of the dispersion of the number of individuals between the present species (evenness), functional diversity is considered to measure the present range of behavioural patterns and life strategies. Analogously to species diversity, it covers the aspects of functional richness (to which degree is the functional niche space utilized by the present community) and functional evenness (what is the proportion between the available traits, Schleuter et al. 2010). Although this concept sounds quite simple, its application turned out to be much more complicated than the concept of species diversity, leading to a variety of new methodical developments (summarized in Schleuter et al. 2010).

One of the most promising approaches is the usage of specific functional traits measured for each species (e.g. Bremner et al. 2003). Basic requisites for the compilation of a biological traits table are fundamental knowledge on the autecology of the inhabiting species and a tool to transform the available information into distinct numbers for the projected calculations. Various approaches for functional traits tables are described in the literature (e.g. Bremner et al. 2003 Tillin et al. 2006, Norling et al. 2007, Törnroos & Bonsdorff 2012, van der Linden et al. 2012). They typically combine life-history traits like reproduction techniques or life-span with those relevant for ecosystem processes like feeding type or position in and alteration of substrates. Some of the cited authors also include traits, which are directly linked with a specific environmental parameter. For example, van der Linden et al. (2012) included a trait called “salinity preference”. However, the usage of such traits might be improper if the aim of the study is to describe the impact of the corresponding parameter on the functional composition. The number of reflected traits and modalities (categories within a trait, see chapter 4.2.3) heavily varies between the studies. Therefore, Bremner et al. (2006) recommended limiting the number of traits and modalities as an excessive number of traits might increase stochastic variability and masks the underlying relationship.

These multiple traits tables can serve as basis to analyse the functional composition of the community using a special correspondence analysis or to calculate the indices for functional diversity as performed in the 2nd case study.

2.3.2. Usage of species distribution models in benthic ecology

Within the last decade the usage of species distribution models (SDM) became more and more popular in marine science as well after previously being used mainly in terrestrial research. The applications ranged from the description of the relationships between species distribution and environmental parameters to solely mapping purposes. Depending on the aim of the study, quite different modelling techniques are available (Elith & Graham 2009, Robinson et al. 2011, Reiss et al. *in prep*.). Especially the development of its usage for mapping purposes was accelerated by the increasing availability of spatial environmental data (Brown et al. 2011) and new modelling techniques.

Basic tools to link the biological response to environmental predictors are linear regression and its derivatives, e.g. generalized linear models (GLM, Zuur et al. 2007). A special case of GLM is logistic regression which uses presence/absence-data as response variable. For a long period, this method has predominantly been used by ecologists, for example by Gogina et al. (2010b) and Gogina & Zettler (2010), who predicted the probability of occurrence of selected benthic species in the south-western Baltic and in the entire Baltic Sea, respectively. The linear regression-family is a useful tool as it corresponds to both main purposes. It directly provides information on the species-environment relation and the results can easily be interpreted by the resulting equations. But the response of a species to an environmental gradient is affected by the simultaneous response of the species to multiple other parameters in natural systems. This increases heterogeneity and often masks the relationship between the considered parameter and the response variable in the ecological models that focus on single parameter gradients (Cade et al. 1999). Therefore, the modelling of an “outer envelope” *sensu* Anderson (2008) might be more appropriate than modelling the mean response, as the outer envelope is less susceptible to the influence of other parameters. A possible solution for this problem is to model the conditional percentiles of the response variable against the provided predictor. The mathematical possibility to use other parts of the distribution of the response variable than the mean was already described in the 1970s by Mosteller & Tukey (1977) and mathematically formulated by Koenker & Bassett (1978). But its usage in marine ecology remained comparably scarce, although the advantages in dealing with the large variability of the response variable are evident. A brief description of the method is given in chapter 3.2.2.

One fundamental drawback of linear regression techniques is the strict underlying mathematical assumptions, which are often violated by the available biotic data. Zuur et al. (2010) gave a comprehensive overview on possible violations and their impact on the reliability of the results and the ecological conclusions. They identified (1) zero-inflation (i.e. much larger proportion of zeros in the biological data than expected by normal distribution), (2) co-linearity among predictors, (3) missing independence of observations (which is in case of spatial data similar to spatial auto-correlation) and (4) overdispersion (the variation is much larger than the mean) as being the most influential violations. A variety of extensions of regression modelling techniques are currently under development to deal with all these problems, but the applicability is still limited as the risk for mathematical errors is still high (Zuur et al. 2009, 2012).

A different approach is pursued by data-driven machine learning techniques. Their algorithms incorporate explicit relationships with mechanistic components independently of current distribution (Guisan & Zimmermann 2000, Buckley et al. 2010, Reiss et al. *in prep*.). Among these techniques, artificial neural networks as well as tree-based classifiers and its extensions became the most popular.

Several studies have recently been published comparing the performance of different modelling techniques for marine benthic invertebrates and other functional groups (e.g. Reiss et al. 2011, Bučas et al. 2013). Both cited studies concordantly came to the conclusion that Random Forest (RF) algorithms are currently the technique with highest predictive accuracy and are used in the 3rd and 4th case study of this thesis.

3. First case study: Changes in substrate preferences of macrobenthic species along a salinity gradient

The results of the first case study form the basis for the manuscript entitled "On the myths of indicator species: issues and further consideration in the use of static concepts for ecological applications" written by M.L. Zettler, C.E. Proffitt, A. Darr, S. Degraer, L. Devriese, C. Greathead, J. Kotta, P. Magni, G. Martin, H. Reiss, J. Speybroeck, D. Tagliapietra, G. Van Hoey, T. Ysebaert and published with Open Access in PLOS one. The publication is available following the link

http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0078219

The authors of the publication acted jointly to accomplish the analysis and the text of the publication as stated in the link above (follow "About the Authors" -> "Author contributions") and as approved by all authors during the submission process. The actual case study presents the results of analysing solely IOW-owned data. Nevertheless, some tables and figures are included within this thesis to strengthen the argumentation.

3.1. Introduction

One of the most challenging aspects in the assessment of biotopes in brackish waters is the fact that the reduced salinity represents a permanent natural stress to the inhabiting macrobenthic communities. The impact of this natural stress on the benthic organism highly resembles the stress induced by human activities (Barbone et al. 2012). As a consequence, most of the species are rather tolerant against changes in their environment. This lack of so-called "sensitive" species often leads to a classification into "bad" or "deficient" categories in state assessment (Zettler et al. 2013c and citations therein).

Additionally, in most cases species are classified into a single sensitivity or tolerance class throughout their distributional range. Thus, it is assumed that their presence in the community always indicates comparable environmental conditions. But it has been noted that the apparent preferences of macrobenthic species might change along natural gradients (Fleischer & Zettler 2009, Grémare et al. 2009). Therefore, the interpretation has to be used with caution, as the real sensitivity of the species to environmental changes might be less pronounced than assumed. As a consequence, some indices such as the Benthic Quality Index (BQI) in the form adapted by Fleischer & Zettler (2009), incorporate the shift in sensitivity along natural gradients and use different sensitivity values for example for different depth and salinity classes.

The target of this case study was to test the general statement that the indicative value of benthic species in the assessment of the impact of human activities might change along the natural gradients. Changes in grain size composition, as a consequence of physical disturbance, and organic load of the substrate, as a consequence of eutrophication, were considered to be among the most severe impacts, therefore both parameters were reflected.

3.2. Material and methods

3.2.1. Available data and selection of species

In order to depict essential parts of the salinity gradient, the dataset from the study area was complemented by IOW-owned data from the German Bight (North Sea, Figure 11).

Sampling and sample treatment were conducted for all stations as described in chapter 2.2. The included stations had to contain information on required abiotic variables such as salinity and median grain size and/or organic content (loss on ignition, LOI). Overall, 1080 stations were available for the analyses using LOI and 909 stations using median grain-size (Table 1).

The data were arranged into salinity classes following the proposal by Reusser & Lee (2011) for an adapted Venice System with six levels:

➢ euhaline zone	>34 to 30
➢ α-polyhaline zone	25 to 30
➢ β-polyhaline zone	18 to 25
➢ α-mesohaline zone	18 to 10
➢ β-mesohaline zone	10 to 5
➢ oligohaline zone	5 to 0.5.

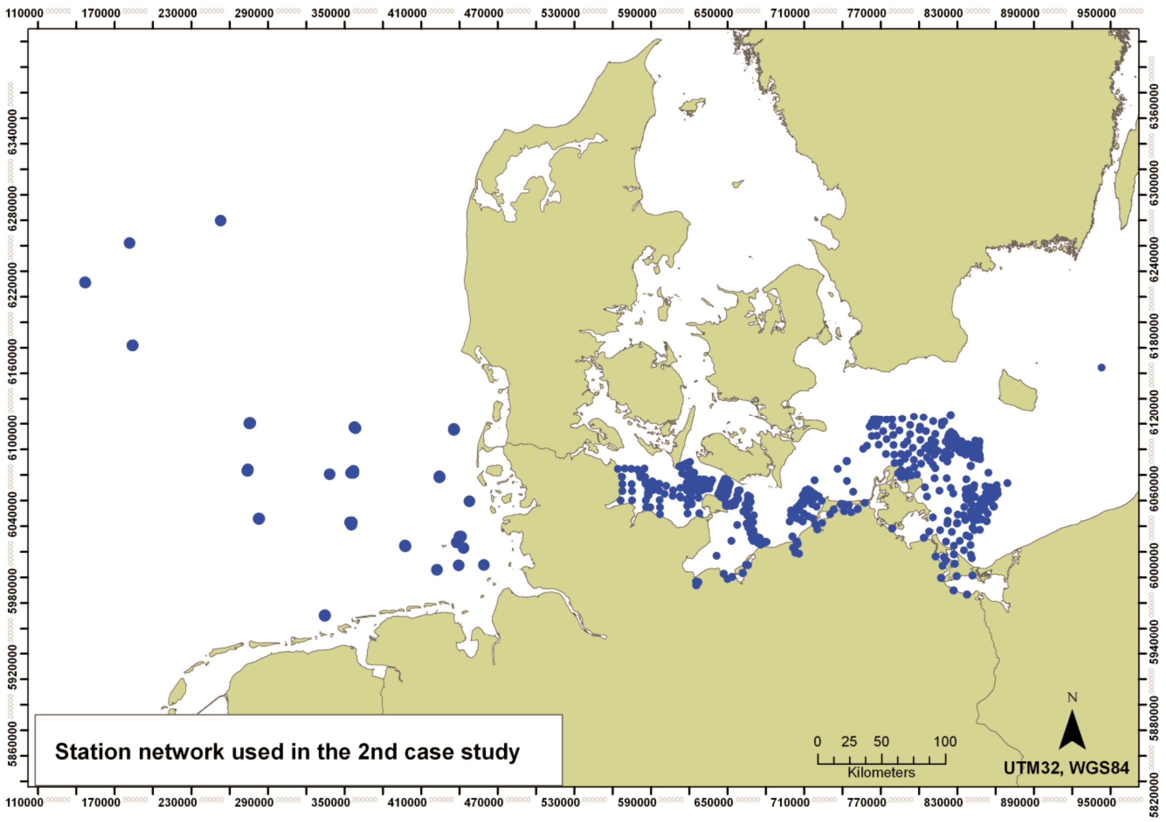


Figure 11: In the 2nd case study, data from the German Bight and the Baltic Sea (including lagoons) were used.

Table 1: Data availability considering loss on ignition and median grain size within the different salinity classes

Substrate parameter	Salinity class	German Bight	Baltic Sea
loss on ignition	oligohaline		72
	β-mesohaline		362
	α-mesohaline		209
	β –polyhaline		278
	α –polyhaline	3	75
	euhalin	81	
median grain size	oligohaline		72
	β-mesohaline		269
	α-mesohaline		184
	β –polyhaline		225
	α –polyhaline	3	75
	euhalin	81	

As it was intended to combine the present analyses with data from the wider North Sea and the Mediterranean (compare Zettler et al. 2013c), the targeted species had to fulfil three basic requirements:

- wide distribution also in other European areas
- presence along a wide range of the salinity gradient (preferably three or more of the classes mentioned above)
- inconsistent sensitivity classification among the major German or European assessment indices including AMBI (Borja et al. 2000), MarBIT (Meyer et al. 2008), BQI (Rosenberg et al. 2004) and MarLIN-data base (www.marlin.ac.uk)

Six species fulfilled these requirements and were selected for the analyses. Among them were the polychaete-species *Ampharete baltica*, *Hediste diversicolor* and *Travisia forbesii*, the amphipod-species *Bathyporeia pilosa* and the bivalve-species *Cerastoderma glaucum* and *Corbula gibba* (Table 2).

The reasoning for the selection process is provided by Zettler et al. (2013c) and can be summarized as follows:

Ampharete baltica: A considerable mud-content in the substrate seems to be mandatory for this tube-building species (Hartmann-Schröder 1996). While it is neither classified as particularly sensitive species nor as very tolerant species in MarBIT and AMBI, its sensitivity seems to change along the Baltic salinity gradient (BQI sensitivity-values).

Bathyporeia pilosa is classified as “very sensitive” in both the MarBIT and the AMBI-System, but is regarded as a more tolerant species in BQI. Although it prefers fine sand and obviously tolerates higher densities of finer particles, *B. pilosa* is frequently considered to be sensitive against changes in sediment characteristics including organic load (d’Udekem d'Acoz 2004 and references therein).

Cerastoderma glaucum is classified as being one of the very few sensitive species occurring in the coastal waters east of the Darss Sill in the German assessment tool for the Water Framework Directive, the MarBIT-Index. By contrast, the cockle is considered to be tolerant in AMBI and BQI. Own observations suggest different preferences between the population inhabiting inner (Estuaries and lagoons) and outer coastal waters.

Corbula gibba is widely accepted to be a species with a very high tolerance to environmental changes and can frequently be found in polluted or otherwise disturbed areas (Hrs-Brenko 2006). Nevertheless, the sensitivity value as used in BQI obviously varies along the salinity gradient of the Baltic Sea.

Hediste diversicolor has been included as counter-example. Also this species is known to be tolerant against environmental changes and pollution and is treated as such in all indicators.

Travisia forbesii is considered to be very sensitive against both nutrient load and changes in sediment composition and has therefore frequently been used as indicator species in the monitoring studies focussed on the effects of sediment extraction and dumping (Krause 2002 and citations therein). In contrast, sensitivity-values used in BQI reflect a higher tolerance.

Table 2: Sensitivity classification of selected species in European assessment systems and sensitivity against nutrient load (MarLIN-database) referring to the following references (Fleischer & Zettler 2009, Borja et al. 2000, Meyer et al. 2008, MarLIN). The table is taken from Zettler et al. (20013c).

Species	MARBIT (German WFD-tool)	AMBI/M-AMBI	BQI (adaptation southern Baltic) - range $ES50_{0.05}$	MarLIN
Ampharete baltica	indifferent (N)	indifferent (II)	tolerant to sensitive (6.6-10.5)	not included
Bathyporeia pilosa	very sensitive (O)	very sensitive (I)	tolerant (4.5-6.9)	not included
Cerastoderma glaucum	very sensitive (O)	tolerant (III)	tolerant (5.7-6.6)	nutrient load: intermediate intolerance, low recoverability, high sensitivity
Corbula gibba	indifferent (N)	very tolerant (IV)	tolerant to sensitive (5.9-9.4)	*C. gibba* are indicative of unstable environments such as ones with low oxygen levels and areas of eutrophication within the Mediterranean [see 87]
Hediste diversicolor	tolerant (T)	very tolerant (IV)	tolerant (5.2-6.9)	high tolerance against changes in nutrient level, may benefit
Travisia forbesii	very sensitive (O)	very sensitive (I)	tolerant (6.3-6.9)	not included

3.2.2. Statistics

Quantile regression spline (QRS) models were chosen as modelling technique. A comprehensive introduction into the usage of quantile regression is given by Cade (2003). The author remarked that the conditional quantiles are the inverse of the conditional cumulative distribution function of the response variable with the selected quantile τ always between 0 and 1 ($\tau \in [0, 1]$). That means, that the quantile $\tau = 0.95$ equals the 95th percentile of the distribution and that 95% of the values are less than or equal to the specified function. A given $\tau = 0.5$ equals the median of a distribution. Basically, the usage of all quantiles is possible, but upper percentiles like the maximum ($\tau = 1$, Thrush et al. 2005) or the 95th percentile ($\tau = 0.95$, Reusser & Lee 2011) seemed to be the most appropriate. The advantage of the 95th percentile over the maximum value is the lower sensitivity to outliers and is therefore recommended (Anderson 2008).

Simple regression models with the response variable y (e.g. species abundance) as a function of a predictor X (e.g. environmental variable) are generally formulated as $y = f(X)$ that is, the function is defined for the expected value of y $[E(y|X)]$ (adopted from Cade 2003). For quantile regression, the basic regression formula remains valid, but the expected value is replaced by the quantile, for example for $\tau = 0.95$ it is denoted by $Q_y(0.95|X)$.

In Quantile regression spline (QRS) models, this special type of response is combined with smoothing splines following the proposal of Koenker et al. (1994). Using smoothing splines is a common approach to fit curves to a set of data-points in different fields of science (Wang 2011). One of the classical approaches is to fit different polynomial functions to the data. Obviously, the fundamental difficulty is the designation of an objective criterion between smoothness of the curve and its closeness to the data points (Pollack 1993). On one hand, if one allows for highest smoothness, the result will be a straight line. On the other hand, if one allows for best fitting, the curve will pass through all data-points (Pollack 1993). As the divergence of a regression line to the data-points is basically expressed as least square error, an approach to deal with this problem is to add a penalty value as a trade-off for the complexity of the smoothing curve (Pollack 1993).

The QRS for the 95th-percentile were modelled for a set of four models per species and salinity class with a polynomial function of degree = 2, 3, 4 and 5. The quantile regression was modelled using the function ‘rq’ of the package quantreg (Koenker 2012) and a smoother was added using the function ‘bs’ (Hastie 1992) of the R statistical analysis software (R Development Core Team 2012).

Akaike’s information criterion corrected for small sample sizes AIC_c (Burnham & Anderson 2004) was used for model selection. The derivation of the AIC_C is described as:

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

where L is a value of the maximised log-likelihood over the estimated parameters, K is the number of parameters in the model, and n is the sample sizes.

AIC_c is calculated for all candidate models. The model with the lowest AIC_c-value is considered to be the best approximating model in the candidate set. This automatic choice of the best model can be modified if another model has a similar AIC_c value (within 2 units) and visually shows a better fitting to the data.

In the figures providing the results of the QRS-models (Figure 12, Figure 14 - Figure 19), vertical dashed lines indicate the highest predicted density of the species. The corresponding median grain size/ LOI-values were interpreted as an estimated optimum for the species within the particular salinity class.

3.3. Results

Ampharete baltica

Loss on ignition was chosen as predictor in the models for *A. baltica*. As only few records of this species were available for oligohaline and euhaline areas, QRS-models were only achieved for the four remaining salinity classes (Figure 12). Abundance declined with increasing organic load of the sediment in polyhaline waters and to a lesser extent in β-mesohaline waters, although in the latter case this pattern was not detected by the QRS-model. By contrast, in α-mesohaline waters, the abundance also declined primary with increasing organic load until LOI ≈ 5%, but rapidly increased again when LOI > 10%. However, this pattern also was only reflected in part by the QRS-model. Due to the reason that the results of the models poorly reflected this potential trend, further analyses were performed subdividing the dataset into three regions, namely the Baltic Sea east and west of Darss sill and the German Bight (North Sea, Figure 13). East of the Darss Sill (salinity 9-15 psu) *A. baltica* can be found in high densities both in substrates with high and low organic load, whereas west of the Darss Sill (salinity > 10 psu) it remains restricted to substrates with low organic load. The species was rarely found at the available North Sea stations.

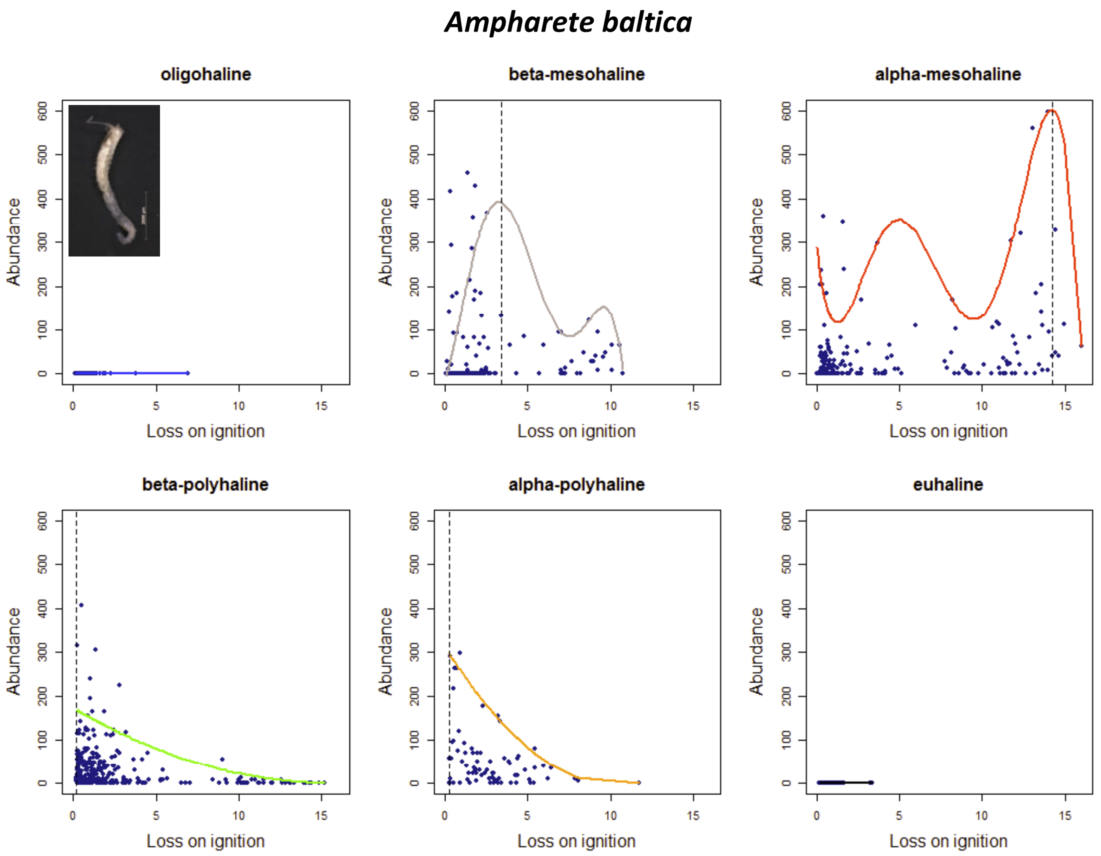


Figure 12: Abundance-response curves (95th percentile QRS) of *Ampharete baltica* against Loss on ignition (in % dry weight) in six different salinity classes in the Baltic and the North Sea (species photo: IOW).

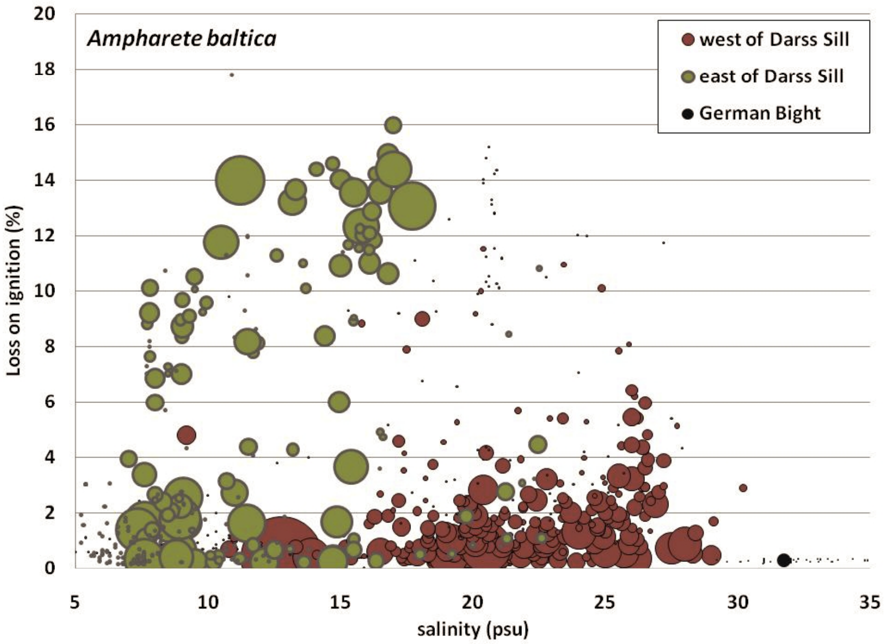


Figure 13: Abundance of *Ampharete baltica* in relation to bottom salinity and organic sediment content in the study area separate into North Sea and Baltic Sea east and west of the Darss sill respectively. The relative bubble size indicates the density of the species with smallest dots referring to species absence and largest bubble size to ca. 1700 ind. * m^{-2}.

Table 3: Summary table of the modelled "optima" for (a) loss on ignition and (b) median grain size

(a) Loss on ignition (%)

species	oligo-haline	β-mesohaline	α-mesohaline	β -polyhaline	α -polyhaline	eu-haline
Ampharete baltica	NA	3.38	14.2	min	min	NA
Cerastoderma glaucum	NA	min	1.98	NA	NA	NA
Corbula gibba	NA	NA	min	3.7	0.24	0.8
Hediste diversicolor	1.26	1.5	2.74	NA	NA	NA
Travisia forbesii	NA	NA	min	min	min	NA

(b) median grain size (µm)

species	oligo-haline	β -mesohaline	α -mesohaline	β -polyhaline	α -polyhaline	eu-haline
Bathyporeia pilosa	NA	198	626	NA	NA	NA
Travisia forbesii	NA	NA	556	(720)	411	NA

NA: no model achieved, min: the optimum equals the minimum of the available predictor value

Bathyporeia pilosa

The modelled optimum grain size for *B. pilosa* differed between the two salinity classes with sufficient data (Figure 14, Table 3). High densities were reached in fine sand at lower salinity, but the specificity to finer substrates seemed to be weaker at higher salinities. Sediments with high organic load (LOI > 2%) were avoided throughout the full salinity gradient (results not shown).

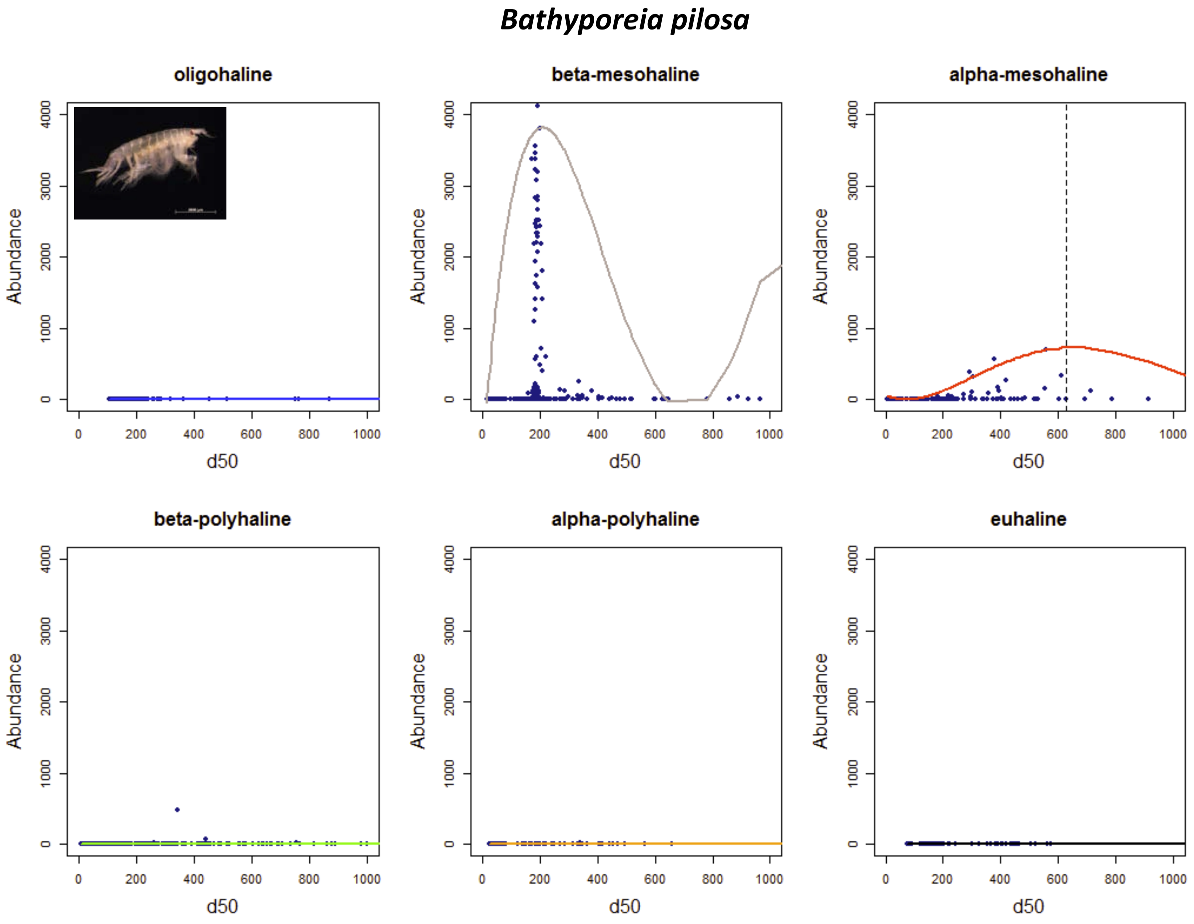


Figure 14: Abundance-response curves (95th percentile QRS) of *Bathyporeia pilosa* along a sediment grain size gradient in six different salinity classes in the North Sea and the Baltic Sea. Note that the y-axis had to be cut in some cases due to outliers (species photo: IOW).

Cerastoderma glaucum

Although the lagoon cockle was found throughout major parts of the salinity gradient, findings from both low and high salinity remained sporadic. High abundances of *C. glaucum* were restricted to mesohaline waters. The tolerance of higher organic load of the sediment did not fundamentally differ between β- and α-mesohaline salinities (Figure 15). In both classes, *C. glaucum* was restricted to substrates with LOI < 5%. The hypothesis that the tolerance differs between populations in sheltered lagoons and in the outer coastal areas, was only partly supported by the available data (not shown). The informative value was restricted by the gap of data for substrates with organic load from sheltered lagoons within the IOW data base.

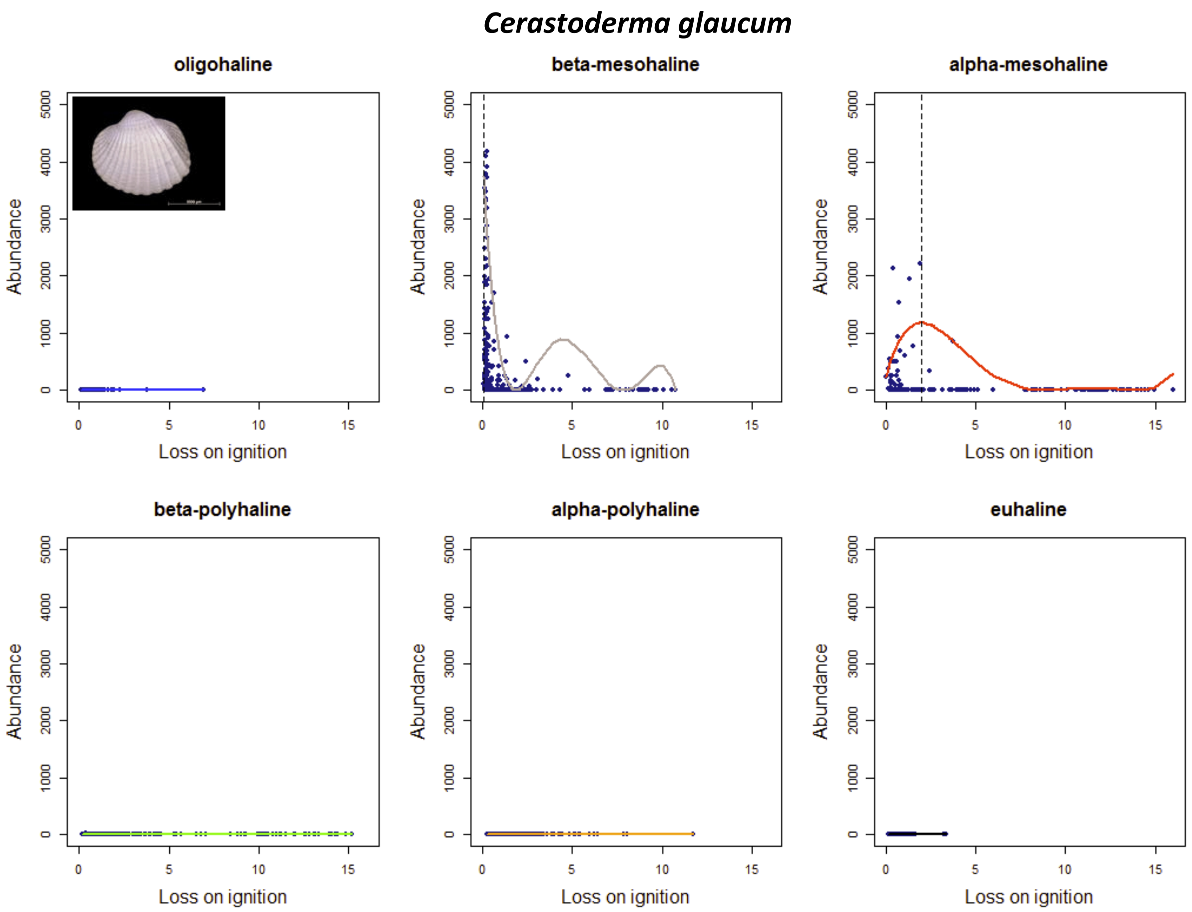


Figure 15: Abundance-response curves (95th percentile QRS) of *Cerastoderma glaucum* along the gradient of organic content (loss on ignition in %) in six different salinity classes in the North Sea and the Baltic Sea (species photo: IOW).

Corbula gibba

Corbula gibba is considered to be rather tolerant against pollution including organic load. The euhaline species was absent in oligohaline water and only infrequently found in β-mesohaline areas, but QRS-models were achieved for the remaining four salinity classes (Figure 16). The modelled optimum was considerably below LOI = 5% in three salinity classes (Table 3) and close to LOI = 5% for the β-polyhaline salinity class. In all four cases abundance distinctly declines with organic load LOI > 5%.

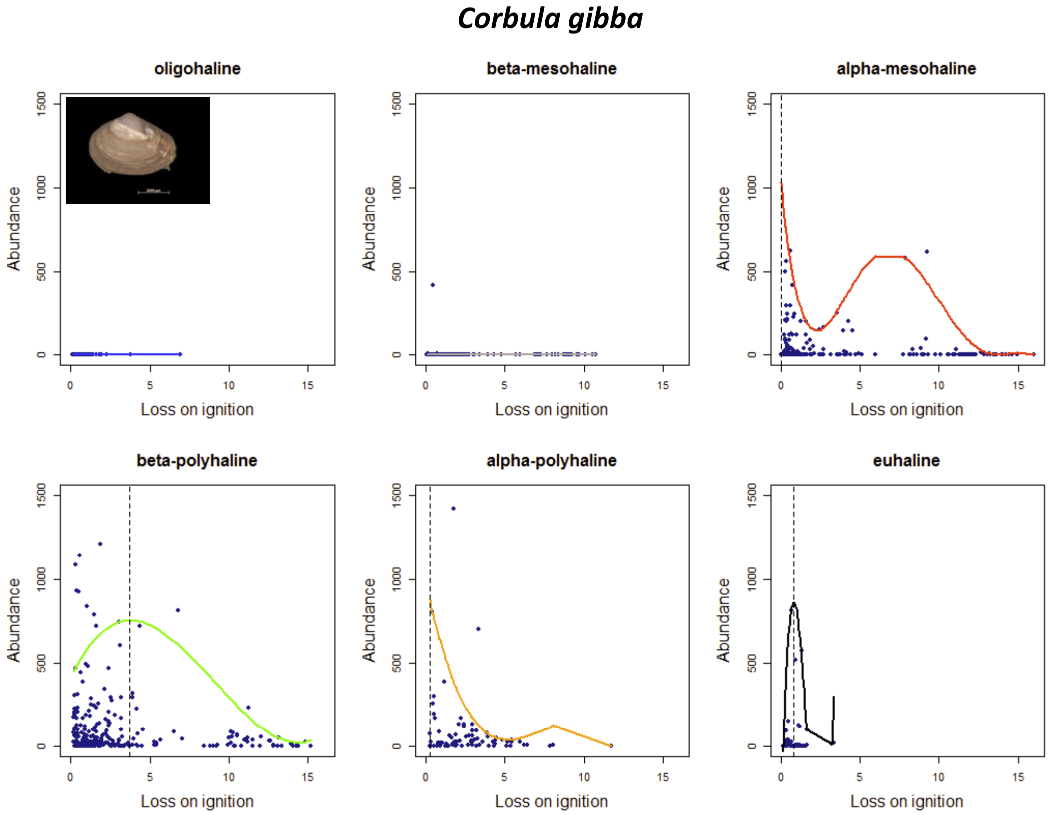


Figure 16: Abundance-response curves (95th percentile QRS) of *Corbula gibba* along the gradient of organic content (loss on ignition in %) in six different salinity classes in the North Sea and the Baltic Sea (species photo: IOW).

Hediste diversicolor

The QRS-models for *H. diversicolor* showed patterns similar to those for *C. gibba* with the exception that the polychaete-species remained restricted to lower salinity classes (oligo- and mesohaline). Although high organic load was tolerated, highest abundances and therefore the modelled optimum were found at low or very low organic content in all three classes (Figure 17, Table 3).

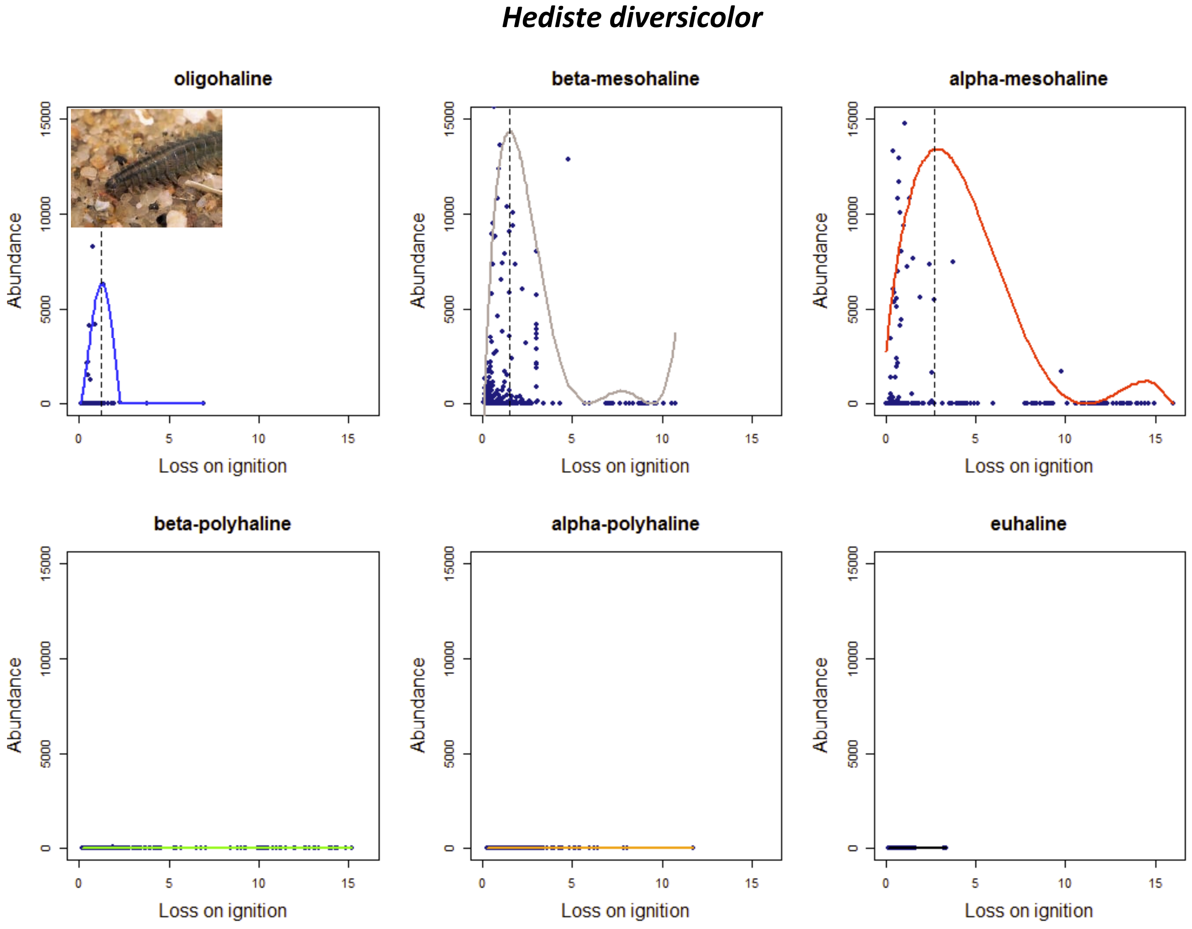


Figure 17: Abundance-response curves (95th percentile QRS) of *Hediste diversicolor* along the gradient of organic content (loss on ignition) in six different salinity classes in the North Sea and the Baltic Sea. Note that the y-axis had to be cut in some cases due to outliers (species photo: IOW).

Travisia forbesii

The QRS-models underlined the assumption that *T. forbesii* is sensitive to organic load. All compiled models resulted in an optimum in substrates almost without organic load (Figure 18). The species was not found in oligohaline and rarely found in euhaline waters. No model was achieved for β-mesohaline salinity class as only few observations were available.

No changes were also found in the modelled optima of sediment grain size for *T. forbesii* along the salinity gradient (Figure 19). Predicted optima varied between 500 and 800 µm in α-mesohaline, polyhaline and euhaline waters. As occurrence of the species was too low, no model was achieved for β-mesohaline waters.

The models obtained for other salinity classes did not seem to represent the underlying data in a satisfying way, as the cloud of dots representing high abundances seems to shift towards smaller median grain sizes with decreasing salinity. This pattern was more pronounced in a descriptive bubble plot. It indicated a tendency of *T. forbesii* populations at lower salinity to accept smaller median grain sizes (Figure 20).

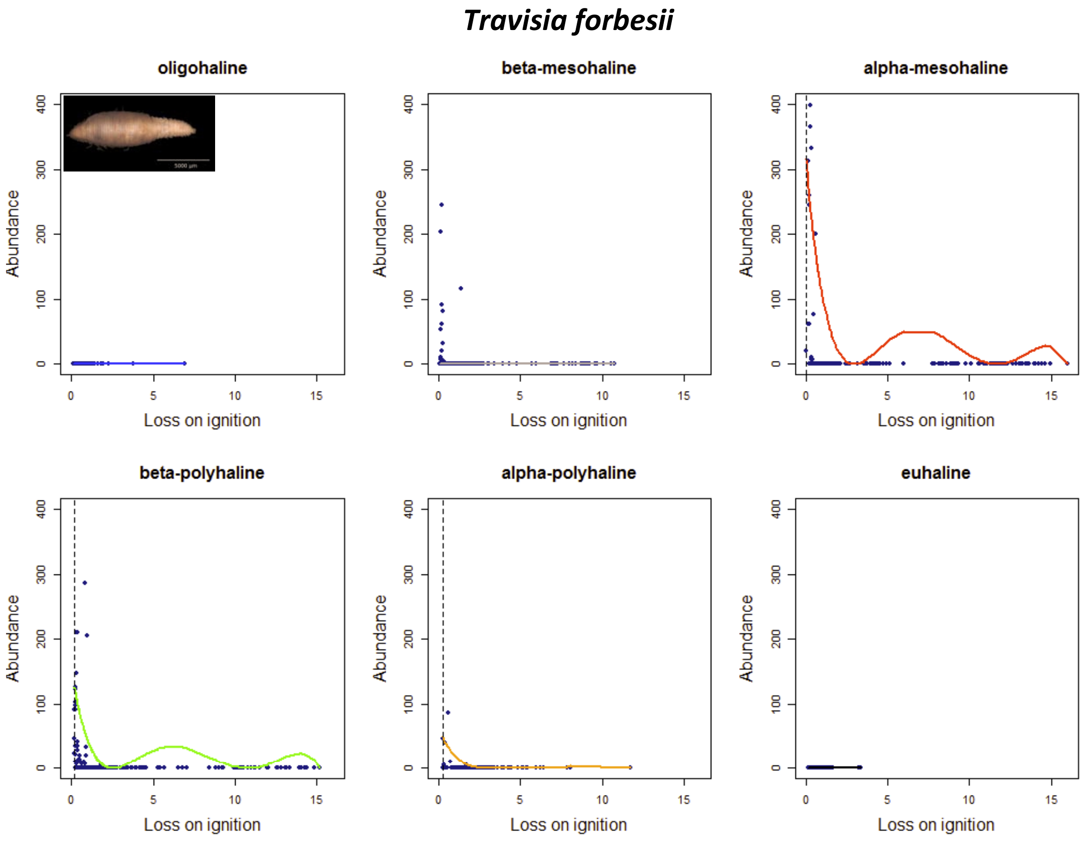


Figure 18: Abundance-response curves (95th percentile QRS) of *Travisia forbesii* along gradient of organic content (loss on ignition in %) in six different salinity classes in the North Sea and the Baltic Sea (species photo: IOW).

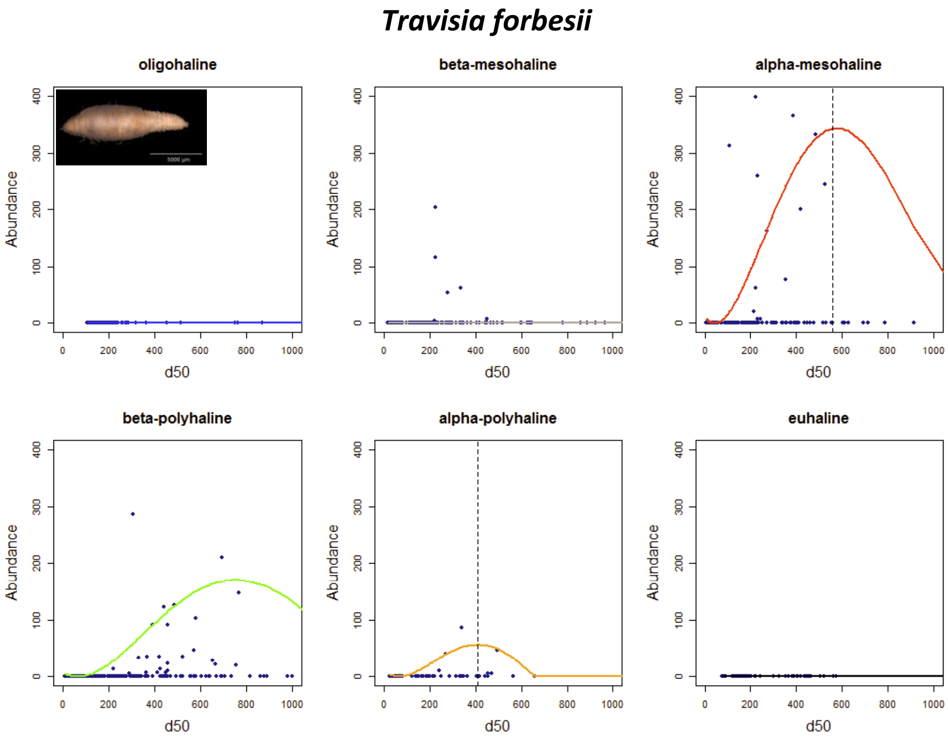


Figure 19: Abundance-response curves (95th percentile QRS) of *Travisia forbesii* along sediment grain size gradient in six different salinity classes in the North Sea and the Baltic Sea. Note that the y-axis had to be cut in some cases due to outliers (species photo: IOW).

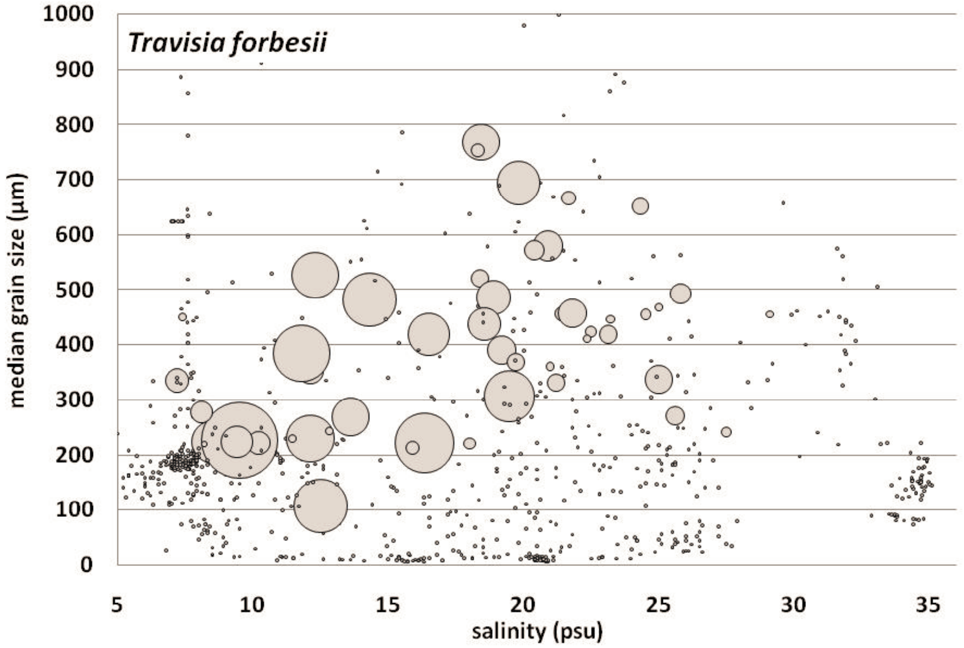


Figure 20: Abundance of *Travisia forbesii* along the full salinity gradient combining the North and Baltic Sea data set and along median grain size. Relative bubble size: species density (absent - 670 ind. * m^{-2})

3.4. Discussion and conclusions

3.4.1. Suitability of the modelling technique

Quantile regression splines were chosen as modelling technique as they had been proven to be a suitable tool in comparable kinds of studies due to their ability to deal with data variability in a much better way than ordinary linear models and GAM (e.g. Cade 1999, Anderson 2008, Keely et al. 2012). Additionally, they provide much more information for the interpretation of the underlying relationship between response variable and individual predictor than for example machine learning techniques.

The usage of QRS-models was only partly successful in this study. The reasons for the results might be manifold. One important drawback might be the variability within the available dataset. In contrast to most examples given in literature, the present data had not been sampled during a single campaign, but over a period of almost ten years and during different seasons. This might increase the heterogeneity and potentially covers the relationship between the species abundance and selected predictor. However, the intended analyses required a large dataset, which could only have been achieved over several years. Additionally, the interannual variability of species' distribution is another important parameter that has to be considered in state assessment.

3.4.2. Relevance of the results for the usage of static indicative values

The examples provided for the German sector of the Baltic Sea and the German Bight clearly documented the challenge to classify benthic species according to their "sensitivity" (Table 4). For instance, the "sensitive" polychaete-species *T. forbesii* showed a distinct shift in the tolerance of finer substrates along the salinity gradient. By contrast, the modelled optima of organic content did not differ distinctly between *C. glaucum* and *H. diversicolor* although the first species is categorized as sensitive and the latter species as tolerant in the German WFD assessment-tool MarBIT (Meyer et al. 2008).

The evaluation of species preferences becomes even more complex, comparing the response of the species from different sea areas. Zettler et al. (2013c) combined and compared the IOW-dataset with data from the eastern Baltic Sea, the North Sea and the

Mediterranean. For instance the sediment preferences of *B. pilosa* differed obviously between the sea areas, thus distinct responses to the median grain size solely became visible in salinity classes including data from a unique sea area. The optimum median grain size appeared to decrease with increasing salinity, but a preference of fine substrates as stated in d'Udekem d'Acoz (2004) could not be verified (Figure 21).

The assumed differences in the tolerance of organic load between populations of *C. glaucum* inhabiting open coastal waters and lagoons in the Baltic Sea could not be proofed as the number of stations from lagoons with muddy substrates was insufficient in the available dataset. Considering the full dataset provided by Zettler et al. (2013c), the lagoon cockle clearly tolerates much higher organic load (and higher salinities) in Mediterranean lagoons than in the Baltic Sea (Figure 22). Consequently, *C. glaucum* was classified as tolerant species in the AMBI which was primary developed for Spanish coastal waters (Borja et al. 2000) and as sensitive species in the MarBIT, developed for the German part of the Baltic Sea.

The few chosen examples evidenced the limitations of state indicators that based on static sensitivity values provided for the whole distributional range of species, especially if applied in transitional waters. Tolerating changes in environmental conditions is a fundamental requirement for species inhabiting transitional waters. As the effects of anthropogenic activities in general and the organic pollution in particular on the macrobenthic community resemble those of natural gradients (Elliott & Quintino 2007), it is a paradox to base estuarine quality assessment on the sensitivity of macrobenthic species (Dauvin et al. 2010, Dauvin et al. 2012, Zettler et al. 2013c).

The direct interpretation of the occurrence of a species and the concluded apparent preferences as sensitivity or tolerance of the species against environmental changes bases on the assumption that the apparent preference equals the fundamental niche of the species. But this assumption is hardly ever true in natural systems as the range of occurrence is always limited by competitive exclusion (Pulliam 2000). The realised niche of a species may differ between different parts of the distributional range either due to different competitive pressures or behavioural discrepancies between the populations. The tolerance of a broader substrate range of species like *T. forbesii* and *A. baltica* at lower salinity in the Baltic Sea might be a consequence of lower competition as species number significantly declines from polyhaline to mesohaline waters (e.g. Bleich et al. 2011, Zettler et al. 2013a, b).

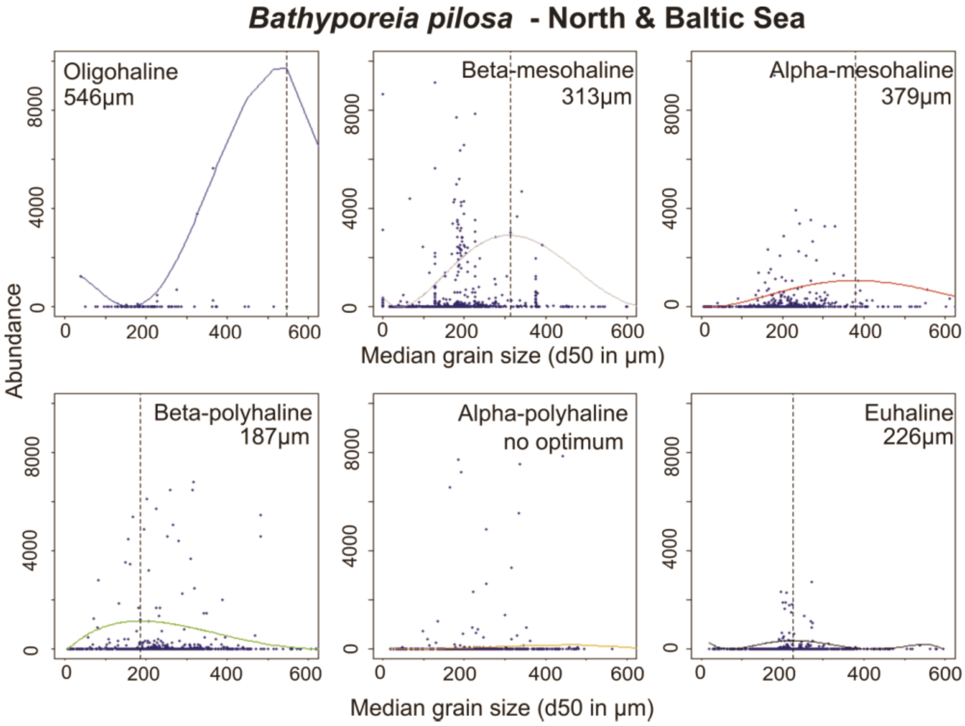


Figure 21: Abundance-response curves (95th percentile QRS) of *Bathyporeia pilosa* along a sediment grain size gradient in six different salinity classes in the North Sea and the Baltic Sea including data from Great Britain, the Netherlands, Belgium, Germany and Estonia. The dashed line marks the modelled optimum. Note that the y-axis had to be cut in some cases due to outliers (taken from Zettler et al. 2013c).

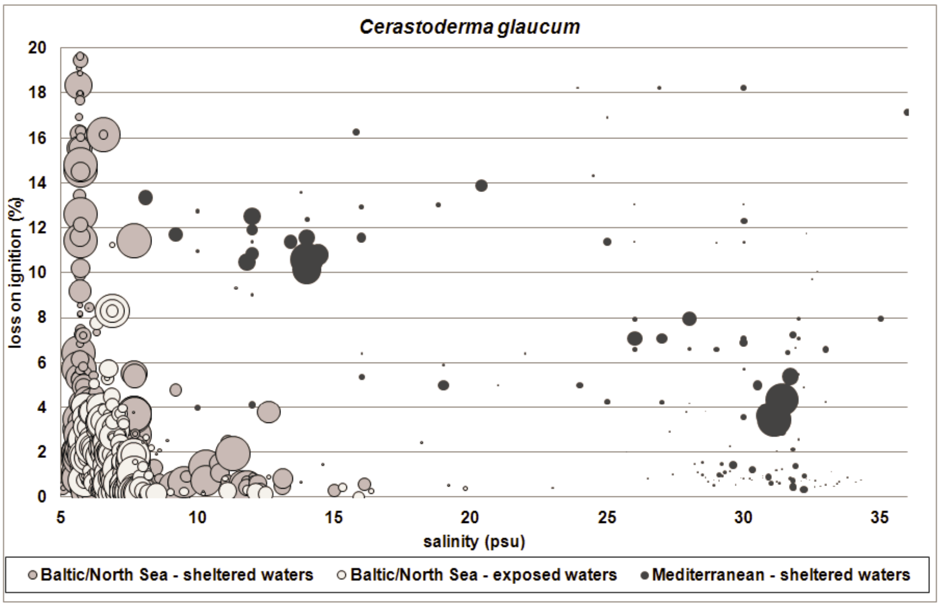


Figure 22: Abundance of *Cerastoderma glaucum* with relation to salinity and LOI both in the North/Baltic Sea and in the Mediterranean. Relative bubble size: density (absent - ca. 10 000 ind. * m^{-2}, taken from Zettler et al. 2013c)

Besides the competitive effect, the apparent differences in occurrence or tolerance of a species might be due to acclimatisation of the local populations to the particular environmental conditions or due to genetic adaptation. Nikula et al. (2008) described the co-existence of two populations with different genetic origin for *Macoma balthica* in the Baltic Sea. Within the western Baltic Sea, the two gene pools were principally separated along the depth gradient indicating a potential adaptation to the different environmental conditions in shallow and deep waters, respectively. Genetic analysis evidenced the potential presence of cryptic species throughout the taxonomic range of marine macrozoobenthos, for example in *Ampharete balthica* (Parapar et al. 2012), *Hediste diversicolor* (Röhner et al. 1997, Virgilio et al. 2009), *Scoloplos armiger* (Kruse et al. 2003) and *Cerastoderma glaucum* (Nikula & Väinölä 2003). As stated in Zettler et al. (2013c), cryptic species may occupy different niches and therefore may show different habitat preferences (Kesaniemi et al. 2012, Gamenick et al. 1998, Luttikhuizen et al. 2011). This might lead to a misclassification of the full species-complex as a “tolerant” species. Therefore, genetic phylogeographic studies and the transformation of genetic information to morphological distinguishable species may improve the accuracy in species identification (Pilgrim & Darling 2010) and in the concept of species sensitivity (Zettler et al. 2013c).

Besides the limitation of the static “sensitivity”-approach due to inter- and intra-specific interactions, it is evident that species show different sensitivities against different pressures (compare e.g. MarLIN database). The soft-bottom bivalve-species of the south-western Baltic Sea developed different responses against organic load or oxygen deficiency, but most of them show similar sensitivities against physical disturbance, for example by bottom-trawling, as they are hardly able to escape. This similarity in the response/sensitivity of life-history traits to particular disturbances might lead to a more accurate reflection of the communities’ response to particular disturbances in the functional approach.

Table 4: Summary table of ecological requirements of selected species, given as assumptions, results and its possible explanations (taken from Zettler et al. 2013c)

Species	*Ampharete baltica*	*Bathyporeia pilosa*	*Cerastoderma glaucum*	*Corbula gibba*	*Hediste diversicolor*	*Travisia forbesii*
assumption	indifferent to organic load	sensitive to changes in grain size distribution	sensitive or tolerant to organic load	tolerant to organic load	tolerant to organic load	sensitive to changes in grain size distribution
Result	avoidance of organic load at high salinity, higher tolerance at low salinities	large differences in preferred grain size detected between different salinity classes and ecoregions	no general sensitivity to organic load detected	found on various substrates, but absent on pure mud with high organic load	found on various substrates also with very high organic load, but highest abundances were found in substrates with medium organic load	higher sediment specificity at higher salinities, widening in mesohaline waters
Rationale	different adaptation of different populations, cryptic species	apparent substrate preferences result of strong/weak interspecific competition with grain size not being the limiting parameter	different adaptation of different populations	apparent substrate preferences result of strong/weak interspecific competition	cryptic species	apparent substrate preferences result of strong/weak interspecific competition with grain size not being the limiting parameter

4. Second case study: Changes in functional diversity and composition along a salinity gradient

The results of this case study have been published by A. Darr, M. Gogina & M.L. Zettler in the Journal of Sea Research entitled "Functional changes in benthic communities along a salinity gradient– a western Baltic case study" and is available following the link

http://dx.doi.org/10.1016/j.seares.2013.06.003.

4.1. Introduction

The results of the previous section have explicitly illustrated the potential pitfalls for species-based indicators in transitional waters and therefore stressed the importance to promote function-based approaches. Analogous to taxonomic-based indicators, shifts in the function within a community can be detected by analysing either the functional diversity or the functional composition. With respect to the latter, special attention is paid to the approach of the Biological Traits Analysis (BTA) and its usage in the context of nature conservation (e.g. Bremner et al. 2003, Bremner 2008, Frid et al. 2008). A few years ago, most studies that used the BTA were performed in full marine areas (e.g. Bremner et al. 2006, Bremner 2008, Hewitt et al. 2008, Papageorgiou et al. 2009, Barrio Froján et al. 2011, Dimitriadis & Koutsoubas 2011), whereas the number of case studies from areas with reduced salinity only increased more recently (e.g. Boström et al. 2010, Aarnio et al. 2011, Villnäs et al. 2011, van der Linden et al. 2012, Törnroos & Bonsdorf 2012). Nevertheless, studies on the effect of the salinity gradient on the functional composition of macrobenthic communities are still rare.

The concept of functional diversity is at least as complex as the concept of functional composition and is interpreted differently in different fields of ecology (e.g. Diaz & Cabido 2001, Petchey & Gaston 2006). Mason et al. (2005) identify evenness, richness and divergence as the primary components of functional diversity and propose different ways to calculate these components. However, in our study we followed the proposal of van der Linden et al. (2012) to calculate Rao's Quadratic Entropy (FD_Q, Rao 1982), in accordance with the specification by Champely & Chessel (2002) and Botta-Dukat (2005).

Van der Linden et al. (2012) also propose to analyse functional redundancy of benthic communities in transitional waters by calculating the ratio between functional diversity (FD_Q) and species diversity (H’). Functional redundancy is defined as the portion of functional traits shared by the species of a community and in general considered to be a proxy for the resilience of an ecosystem/community versus environmental changes (Naeem 1998).

The target of this case study was to identify potential functional changes in the benthic communities along a salinity gradient by analysing functional composition, functional diversity and functional redundancy.

4.2. Material and Methods

4.2.1. Study area

Within this case study, three distinct sub-areas were selected for comparison of functional composition and diversity, each representing a different part of the salinity gradient (Figure 23).

The Fehmarnbelt and adjacent waters were included as the westernmost area is directly supplied with polyhaline waters through the Belt Sea with salinity mainly ranging between 18 – 25 psu within the dataset (Figure 24). The Darss Sill represents the border between the western Baltic and the Arkona Basin, where salt water inflows are restricted not only by the low sill, but also by the relatively narrow Kadet Trench. This area is characterised by salinity ranging mostly from 9 to 15 psu. In contrast, the salinity in the easternmost area, the Pomeranian Bay, is rather stable ranging with few exceptions between 7 and 8.5 psu.

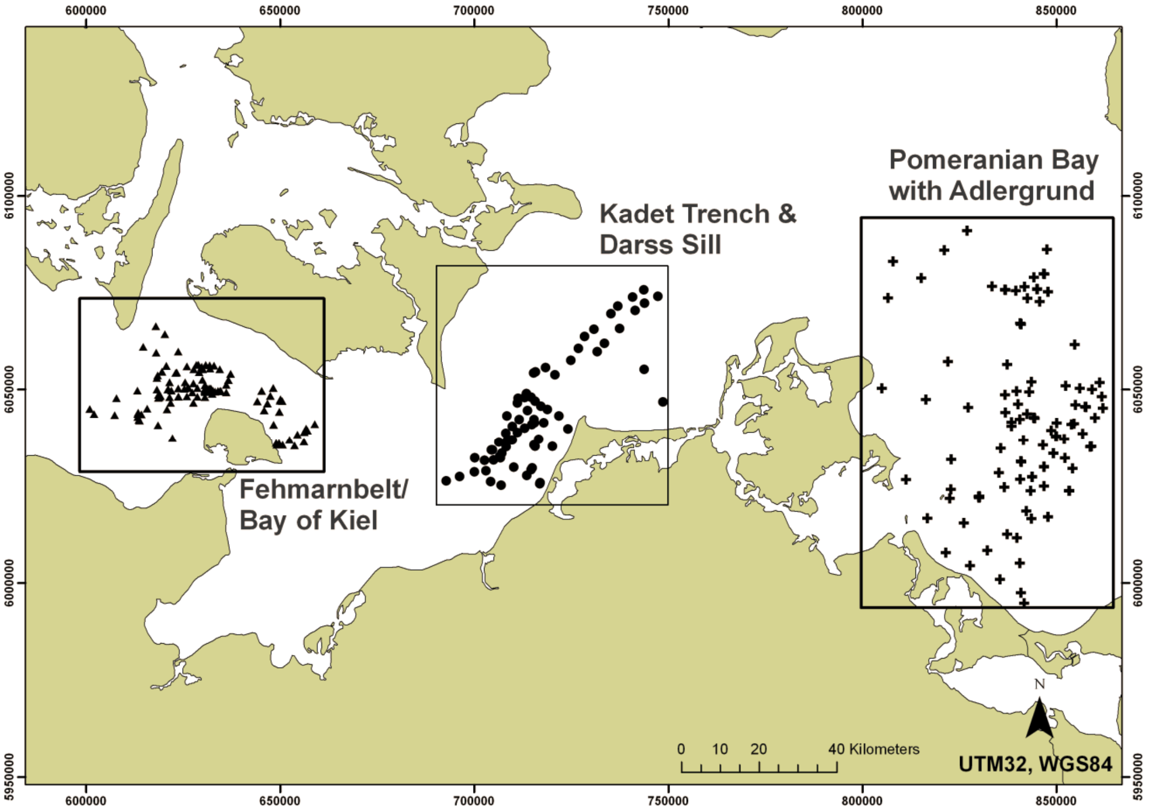


Figure 23: Map of the south-western Baltic Sea showing the position of the selected stations of the third case study in the areas Fehmarnbelt, Kadet Trench and Pomeranian Bay.

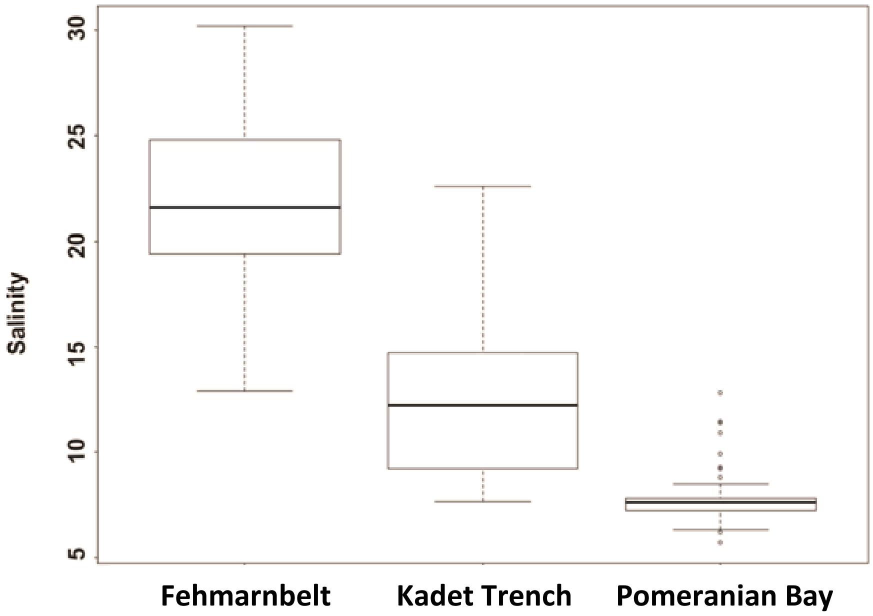


Figure 24: Boxplots showing the different salinity ranges in the three areas

4.2.2. Generation and selection of data sets:

The stations were sampled between 2002 and 2012 in the context of several projects (Table 5) following the procedure described in chapter 2.2.

The study focuses on the functioning of soft-bottom communities, although natural hard substrates (e.g. boulders, hard clay) are common in parts of the study area. Therefore all stations with abiotic (sediment description, parallel underwater-video) or biotic (very high dominance of epibenthic species, e.g. *Mytilus edulis*) indicators of the presence of macrophytes, stones or coarse sediments were excluded from the data set. Nevertheless, countable not colony-forming epibenthic species were included, if present at the sampled stations. The covered substrates ranged in all areas from organic mud to medium sand.

Colony-forming taxa like sponges and bryozoans were generally excluded, as they were not captured quantitatively. To ensure the comparability between the taxonomy-based data set and the biological traits set, all species without full biological traits information (see below) were also excluded, reducing the number of included taxa from 252 overall identified taxa to 208. To avoid a potential bias or a loss of information in the functional data set, only stations with a contribution of the reduced species set of at least 95 % of the original abundance and/or biomass were selected for further analysis. For the vast majority of stations more than 99 % of the original abundance and/or biomass could be translated into the functional information.

Most of the disregarded countable species were very rare, occurring three times or less in the data set with only few specimens. The only exception was the polychaete species *Neoamphitrite figulus* achieving high dominance values at five stations in the Kadet Trench (up to 20 %). Therefore, these stations were also excluded from further analysis, resulting in a dataset covering 370 stations (Fehmarnbelt n = 160, Kadet Trench N = 76, Pomeranian Bay n = 134, Table 5).

Table 5: Number of stations sampled at each of the three survey areas between 2002 and 2012

Year	Fehmarnbelt	Kadet Trench	Pomeranian Bay
2002		15	
2003	25	10	
2004	9	10	28
2005	17	8	23
2006	1		30
2007	13	17	7
2008	1		2
2009	21		8
2010	48	1	6
2011	10	1	20
2012	15	14	10
Overall	**160**	**76**	**134**

4.2.3. Autecological data

The functional traits table used in this thesis had been developed by several scientists within the working group “Ecology of benthic organisms” at IOW, including the author of this thesis. The largest portion of literature review was provided by Dr. Mayya Gogina and Christiane Volkmann. Scientific literature and online information systems (e.g. MarLIN) were browsed for relevant ecological information. If information from literature was missing or vague, it was substituted by the known features of related species or complemented by expert knowledge. A focus was set to literature explicitly describing the peculiar properties of the Baltic Sea populations. Literature provided for populations from other sea areas was used with caution.

Overall, information was gathered for about 400 macrobenthic species including all common species of the study area and adjacent waters. Ecological information was categorized into 12 biological traits. The chosen traits covered information on mobility, feeding, habitat structuring, sediment transport mode, position in sediment, sensitivity to hypoxia, size and form of the individuals, reproduction season and technique as well as features of larvae type, life-stage and longevity. Each trait comprises several modalities summarizing the overall number of categories to 46 (Table 6).

Table 6: Categories of biological traits used to assess functional composition

Trait	Category	description	Label
1. Motility	1.1. Sedentary	non-moving	mot.sed
	1.2. Limited free movement	e.g., withdrawal into sediment	mot.lim
	1.3. Freely motile in/on sediment	free-moving, roaming	mot.fre
	1.4. Semi-pelagic	more water-associated, hyperbenthos, demersal nekton	mot.spel
2. Feeding type	2.1. Suspension	relying on water currents to deliver food particles/dissolved nutrients to body or actively pumping water through body to obtain oxygen and food	feed.sus
	2.2. Deposit feeder	feeding on refractory detrital material on or in the sediment	feed.dep
	2.3. Predator	actively hunting for live animals	feed.pre
	2.4. Scavenger	consuming dead animals	feed.sca
	2.5. Grazer	consuming algal or plant-like material	feed.gra
	2.6. Parasite or commensal	parasite/commensal	feed.com
3. Habitat structure	3.1. Permanent burrow	Permanent burrow that is used at least a part of its life	struc.bur
	3.2. Hole, pit or non-permanent burrow	hole, pit or non-permanent burrow	struc.pit
	3.3. Tube	Tube	struc.tub
	3.4. Mound	Mound	struc.mou
	3.5. Creating troughs, trampling on sediment	Creating troughs or trampling across sediment surface	struc.tra
	3.6. Forming biogenic epibenthic structures	Species body represents a habitat for other species	struc.epi
4. Sediment transport	4.1. Surface-to-deep	e.g. head-up conveyor belt feeder (defecate at depth)	sed.sd
	4.2. Deep-to-surface	e.g. head-down conveyor belt feeders (defecate at surface)	sed.ds
	4.3. Surface mixing	surface biodiffusive	sed.sm
	4.4. Deep mixing	deep biodiffusive	sed.dm

Trait	**Category**	**description**	**Label**
5. Position in sediment	5.1. Protruding surface	through and above sediment surface	pos.pro
	5.2. Attached to hard substrate, other structures	to other animals or small hard surfaces	pos.att
	5.3. Oxygenated zone	yellow or brown zone in surface sediment	pos.oxy
	5.4. Below oxygenated zone	dark grey or black and usually sulfidic zone deeper in sediment	pos.sox
6. Body size	6.1. Small	0.5 - 5 mm longest dimension, exclusive	size.s
	6.2. Medium	5-20 mm longest dimension, exclusive	size.m
	6.3. Large	>20 mm longest dimension	size.l
7. Body form	7.1. Vermiform	length >> width	form.ven
	7.2. Globulose	length ~ width	form.glo
	7.3. Others	e.g. multi-structured	form.oth
8. Reproductive season	8.1. Winter	(self-explanatory)	reps.win
	8.2. Spring		reps.spr
	8.3. Summer		reps.sum
	8.4. Autumn		reps.aut
9. Reproductive technique	9.1. Asexual	budding, parthenogenesis	sex.asex
	9.2. Spawn	gonochoristic, both sexes spawn gametes in water column	sex.spa
	9.3. Attached eggs	gonochoristic, the eggs are attached to substrate	sex.egg
	9.4. Brood	gonochoristic, the eggs are kept in or on the body or coelom	sex.bro
10. Larvae	10.1. Planktonic	planktonic larvae life-stage	lar.pla
	10.2. Benthic	benthic larvae life-stage	lar.ben
11. Hypoxia sensitivity	11.1. Low	high or long tolerance (≥ 21 days)	hyp.low
	11.2. High	low or short tolerance (0-2 days)	hyp.hig
12. Longevity (years)	12.1. ≤ 2	< 2 years	long.2
	12.2. 2-5	2-5 years	long.5
	12.3. 5-10	6-10 years	long.10
	12.4. 10+	> 10 years	long.10p

The ecological data was compiled using a "fuzzy coding" approach (Chevenet et al. 1994). For each modality, a score within the range between 0 and 1 was assigned. It was mandatory that the sum of all modalities equals 1 to avoid bias between the different traits in further analysis. The allocated score depended on the affinity of the species to the individual modality. This implied that each species was either prescribed to just one modality per trait (e.g. *Mytilus edulis* scored to 1 as "suspension feeder") or to a combination of different categories reflecting the differentiation suggested by available information (e.g. *Macoma balthica* scored to 0.6 as "deposit feeder" and 0.4 as "suspension feeder"). All zero scores were assigned in rare cases when some trait categories were irrelevant for certain species groups, such as "sediment transport" for epiphytic species grazing on plants and algae (e.g. *Theodoxus fluviatilis*) or for commensals (e.g. *Odostomia scalaris*).

The abundance of each species at each station was multiplied by its fuzzy coding score for each trait category and afterwards summed up across all species present at the particular station (Bremner et al. 2003, Hewitt et al. 2008).

Biomass (afdm) and abundance were chosen as response variables for the linkage of the species-by-stations table to the traits-by-species-table in independent analyses. To down-weight the influence of dominant taxa, biomass ($\log_{10}$+1) and abundance√)(were transformed before the analyses.

4.2.4. Statistical methods

Common indices describe species diversity as a function of the distribution of species density to species richness. The most widespread index is the Shannon-Index (Shannon 1948):

$$H' = -\sum_{i=1}^{Z} p_i \; x \; log_2 p_i \tag{2}$$

where p_i is the proportion of individuals belonging to the *i*-th species. The index was originally developed to quantify the information content, the so-called entropy, of single letters under the framework of information theory.

Hurlbert (1971) introduced the concept of rarefaction as an alternative approach to species diversity by calculating the expected number of species in a sample of *n* individuals selected at random from a collection containing *N* individuals, *S* species, and N_i individuals in the *i*-th species:

$$E(S_n) = \sum_i \left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right] \tag{3}$$

A different approach was given by Rao (1982), defining the quadratic entropy as

$$FD_Q = \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} d_{ij} \, x \, p_i^2 \tag{4}$$

where d_{ij} is the difference between the *i*-th and *j*-th species (d_{ij} = d_{ji} and d_{ii} = 0). This index has been tested and approved by Botta-Dukat (2005) as the most appropriate available index to measure functional diversity.

Functional diversity FD_Q was derived as proposed by van der Linden et al. (2012) by (1) calculating Rao's quadratic entropy (RQE) separately for each of the 12 traits and (2) summing up these single RQE values for each station. Van der Linden et al. (2012) suggested calculating functional redundancy (FR) by the ratio between functional diversity and species diversity (FD_Q/H'). This proposal is followed in this study.

Differences between areas were tested for significance using a Kruskal-Wallis-Test (Kruskal & Wallis 1952). In case of significant differences, pairwise comparisons were performed by applying Mann-Whitney U tests (Mann & Whitney 1947).

Fuzzy Correspondence analysis

Differences in functional composition between sites and areas were identified using a Fuzzy Correspondence Analysis (FCA, Chevenet et al. 1994). The FCA is an extension of the Correspondence Analysis (CA) and is able to deal with the specific demands of a fuzzy dataset. The purpose of these ordination techniques is the arrangement or 'ordering' of species/stations along the studied gradients. The simple Correspondence analysis (CA) has the same statistical background as principal component analysis (PCA), but it is used for categorical instead of continuous data.

Results of the multidimensional FCA are shown in two two-dimensional plots. In the first plot, the different modalities of all included traits along the gradient of the first two axes are plotted. Variability explained by these axes and correlation between the twelve traits and the axes are obtained from additional tables. The second plot shows the position of the stations and the centroid of the three areas in relation to the same axes.

All analyses were performed under the frame of the R environment (Version 2.13.2, R development Core Team 2012) using the packages vegan (Version 1.17-11, Oksanen et al. 2011) and ade4 (Version 1.4-17, Chessel et al. 2004).

4.3. Results

4.3.1. General description of communities, species richness and taxonomic diversity

Overall, 208 taxa were included into the data set, of which 190 taxa were detected in the Fehmarnbelt area, 124 species were identified in the Kadet Trench and 62 were present in the Pomeranian Bay dataset.

The community in the Fehmarnbelt area was numerically dominated by several species, but none of them achieved more than 10 % off the overall abundance (Table 7). By contrast, the ocean quahog *Arctica islandica* (68.9% of overall biomass) and *Astarte borealis* (10.8%) dominated the biomass in this area.

The ocean quahog also dominated the biomass within the Kadet Trench reaching a comparable dominance value (60.7% of overall biomass). The species with the second highest biomass in this area was the soft-shell clam *Mya arenaria* (16.0%). *Peringia ulvae* (18.0%) and *Diastylis rathkei* (14.1%) were the most widespread species reaching highest mean abundance.

In the Pomeranian Bay, the dominance value of the most abundant species (*Peringia ulvae*) was higher than in the other two areas (36.8% of overall abundance). In contrast the dominance of the species with the highest biomass (*Macoma balthica* 29.4% of the overall biomass, *Mya arenaria* 28.7%) was lower than in the western areas.

Table 7: The ten most dominant species regarding both abundance (left) and biomass (right) are listed for the three areas to characterise the communities.

Fehmarnbelt

Species	Density $[m^{-2}]$	Dominance	Species	Biomass $[g\ afdm\ m^{-2}]$	Dominance
Kurtiella bidentata	331	9.4%	*Arctica islandica*	32.51	68.9%
Scoloplos armiger	237	6.7%	*Astarte borealis*	5.12	10.8%
Abra alba	227	6.4%	*Astarte elliptica*	3.15	6.7%
Terebellides stroemii	212	6.0%	*Terebellides stroemii*	0.69	1.5%
Lagis koreni	195	5.5%	*Alitta virens*	0.47	1.0%
Dendrodoa grossularia	179	5.1%	*Nephtys ciliata*	0.42	0.9%
Diastylis rathkei	133	3.8%	*Mya truncata*	0.39	0.8%
Corbula gibba	125	3.5%	*Amauropsis islandica*	0.36	0.8%
Parvicardium pinnulatum	115	3.3%	*Macoma calcarea*	0.33	0.7%
Edwardsia danica	96	2.7%	*Nephtys caeca*	0.31	0.7%
Overall	**3538**		**Overall**	**47.2**	

Kadet Trench / Darss Sill

Species	Density $[m^{-2}]$	Dominance	Species	Biomass $[g\ afdm\ m^{-2}]$	Dominance
Peringia ulvae	577	18.0%	*Arctica islandica*	16.79	60.7%
Diastylis rathkei	449	14.1%	*Mya arenaria*	4.41	16.0%
Mytilus edulis	282	8.8%	*Mytilus edulis*	1.69	6.1%
Scoloplos armiger	278	8.7%	*Macoma balthica*	1.54	5.6%
Pygospio elegans	265	8.3%	*Lagis koreni*	0.45	1.6%
Macoma balthica	226	7.1%	*Astarte borealis*	0.39	1.4%
Lagis koreni	195	6.1%	*Nephtys caeca*	0.33	1.2%
Kurtiella bidentata	146	4.6%	*Scoloplos armiger*	0.32	1.2%
Abra alba	123	3.8%	*Abra alba*	0.23	0.8%
Arctica islandica	71	2.2%	*Peringia ulvae*	0.18	0.7%
Overall	**3197**		**Overall**	**27.7**	

Pomeranian Bay

Species	Density [m^{-2}]	Dominance	Species	Biomass [g afdm m^{-2}]	Dominance
Peringia ulvae	1989	36.8%	*Macoma balthica*	2.18	29.4%
Pygospio elegans	660	12.2%	*Mya arenaria*	2.13	28.7%
Bathyporeia pilosa	602	11.1%	*Cerastoderma glaucum*	0.78	10.6%
Mya arenaria	584	10.8%	*Marenzelleria viridis*	0.58	7.8%
Marenzelleria viridis	347	6.4%	*Mytilus edulis*	0.56	7.6%
Cerastoderma glaucum	251	4.6%	*Peringia ulvae*	0.48	6.5%
Macoma balthica	213	3.9%	*Hediste diversicolor*	0.33	4.4%
Mytilus edulis	174	3.2%	*Bathyporeia pilosa*	0.11	1.5%
Heterochaeta costata	131	2.4%	*Pygospio elegans*	0.06	0.9%
Hediste diversicolor	118	2.2%	*Corophium volutator*	0.05	0.6%
Overall	**5400**		**Overall**	**7.4**	

Highest abundance was reached in the Pomeranian Bay (mean 5400 ind.*m^2, Figure 25A), being significantly higher than in the other two areas (Kruskal-Wallis $p < 0.001$ for both comparisons). No differences in abundance were found between Fehmarnbelt and Kadet Trench.

Biomass, species richness and species diversity showed a significant decrease from west to east ($p < 0.001$ for all pairwise comparisons in KW-test). The mean biomass decreased by factor five from 47.4 g afdm*m^2 in the Fehmarnbelt area to 7.4 g afdm*m^2 in the Pomeranian Bay (Figure 25B). The mean species richness (ES_{50}) declined from 16.1 in the Fehmarnbelt area to 7.8 in the Pomeranian Bay (Figure 25C). Also species diversity (Shannon-index) was lowest in the Pomeranian Bay (H' = 2.3, Figure 25D) and highest in Fehmarnbelt (H' = 3.8). The Kadet Trench reached intermediate values for all parameters.

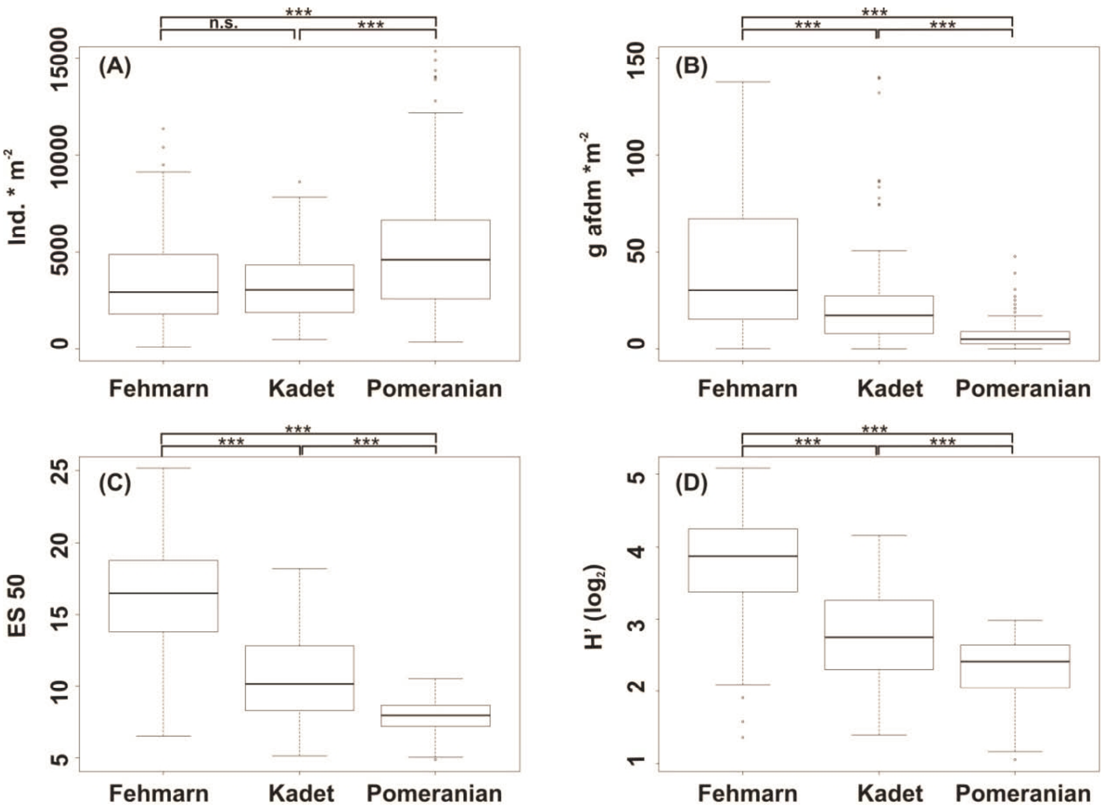


Figure 25: Boxplots showing the mean abundance (A), mean biomass (B), species richness (C) and biodiversity (D) in the three areas. Results of pairwise comparisons (KW-test) are shown above the plots with significance levels as follows: ***: $p < 0.001$, ** $p < 0.01$, *$p < 0.05$, n.s.: not significant.

4.3.2. Functional diversity and redundancy

Functional diversity was higher in the Fehmarnbelt area than in the Kadet Trench ($p < 0.05$) and in the Pomeranian Bay ($p < 0.05$) based on abundance data (Figure 26A). In contrast, functional diversity increased significantly from west to east when using transformed biomass (KW-test, $p < 0.01$ for all pairwise comparisons, Figure 26C).

Differences in functional redundancy ($FR = FD_Q/H'$) followed the same trend in both approaches (Figure 26B, D) with FR increasing significantly ($p < 0.001$ for all pairwise comparisons in both approaches) from Fehmarnbelt towards Pomeranian Bay.

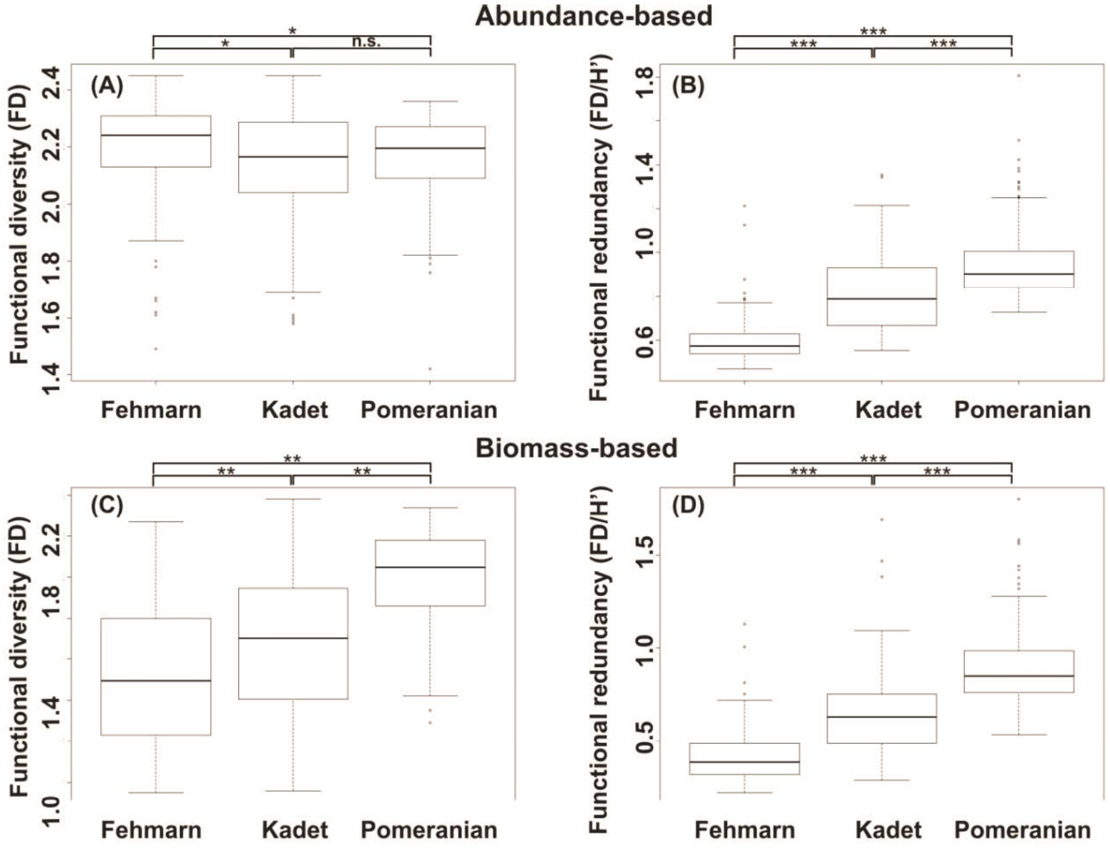


Figure 26: Functional diversity (A) and functional redundancy (B) based on the abundance data set and the biomass data set (C, D) respectively. Results of pairwise comparisons (KW-test) are shown above the plots with significance levels as follows: ***: $p < 0.001$, ** $p < 0.01$, *$p < 0.05$, n.s.: not significant.

4.3.3. Functional composition

Two separate Fuzzy Correspondence Analyses (FCA) based on abundance and biomass, respectively, were performed to analyse the functional composition of the communities.

The separation of the traits modalities on the factorial map based on the abundance data was weak (Figure 27A). The body size modality “small”, the reproduction technique “eggs”, position attached, feeding type “grazer” and semi-pelagic motility were sited most left-handed of the first axis. The most right-handed coordinates were attained by the feeding type “commensal”, the reproduction season “winter”, unspecific other body forms and moderate longevity (long.5). The two latter and the modality “epibenthic” within the trait habitat structure were separated from the other traits along the ordinate. The correlation of the traits with the two first axis of the FCA remained weak (< 0.1 for all traits, Table 2). The weakness of the signal was stressed by the low part of the variation explained by the first two axes (42.5%). Also the station map showed that the separation of the areas was incomplete (Figure 27B). The centroids were ordered along the first axis with the area Pomeranian Bay, attaining the highest differentiation along the second axis.

Separation of the traits-modalities along the first two axes and also their correlation were stronger in the factorial map using the biomass data than on the one using the abundance data Figure 28A). This outcome was reflected in the higher proportion of variation explained by the first axis (32.5%). Two clusters of traits were separated along the first axis including the modality “deep to surface mixing” within the trait sediment transport, feeding type “grazer”, semi-pelagic motility, small size-class, planktonic larvae and short-lived species (long.2) in the left-hand group and “mound building” (trait habitat structures), long-lived species (long.10p), benthic larvae and feeding type “commensals” on the right hand side. The second axis mainly separates “forming epibenthic structures” (trait habitat structures), position attached (trait position in sediment) and unspecific other body forms from the predominant parts of traits. The traits larvae type (0.21), longevity (0.19) and habitat structures (0.15) were moderately correlated with the first axis (Table 8). The correlation with the ordinate was highest for the traits habitat structures (0.2) and position in sediment (0.16). The factorial map of the traits per station matrix showed an overlap of the investigated areas, but the Pomeranian Bay and the Fehmarnbelt were mostly separated (Figure 28B). The centroids of the three areas were almost lined up along the first axis analogous to their order from east to west, whereas the Kadet Trench area showed the strongest deviance along the second axis.

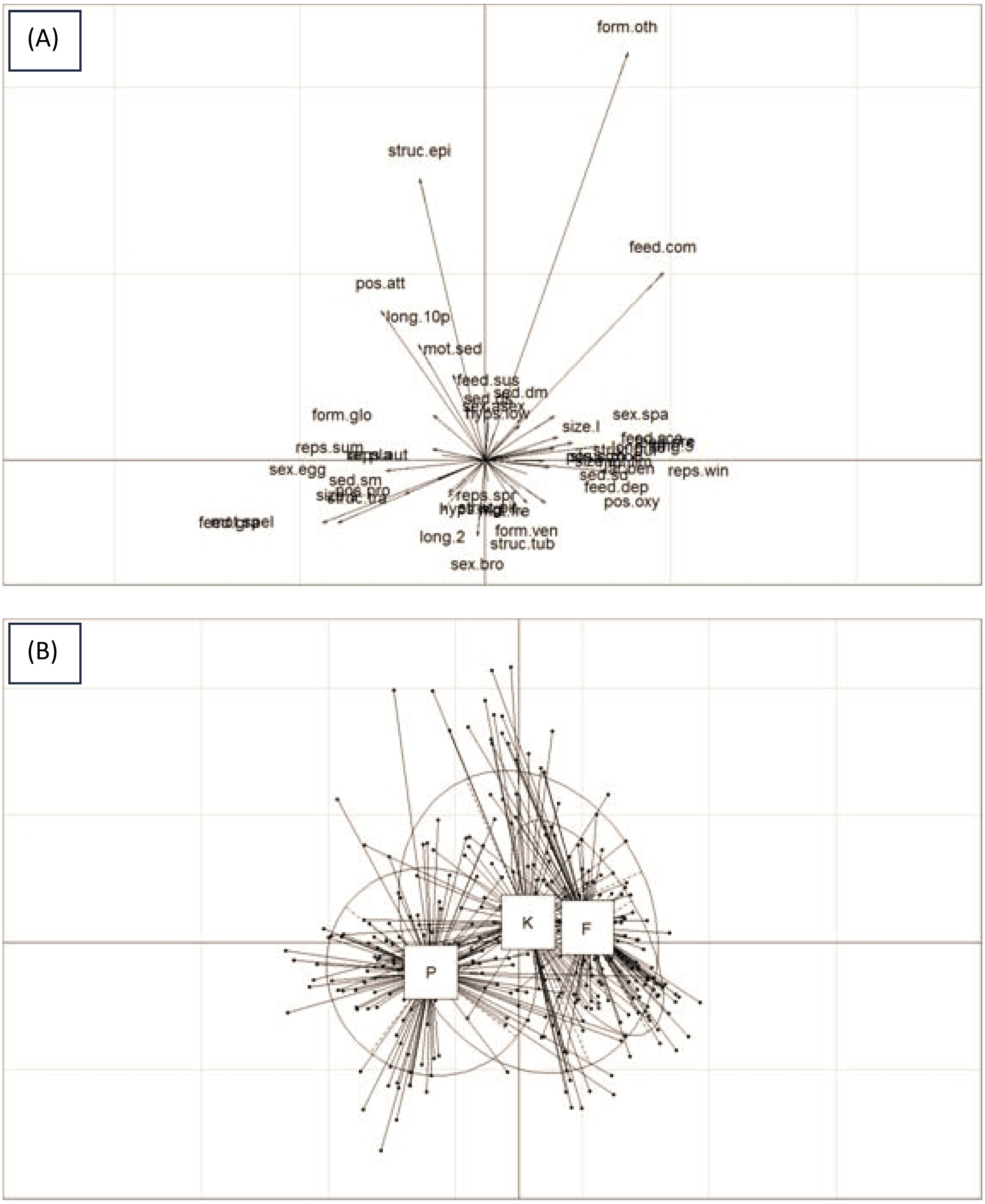


Figure 27: Factorial maps of the FCA based on abundance showing the distribution of the traits modalities (A) and the stations (B) along the first two factorial axes. Letters in plot B indicating the centroids of the three areas with labels F = Fehmarnbelt, K: Kadet Trench, P: Pomeranian Bay.

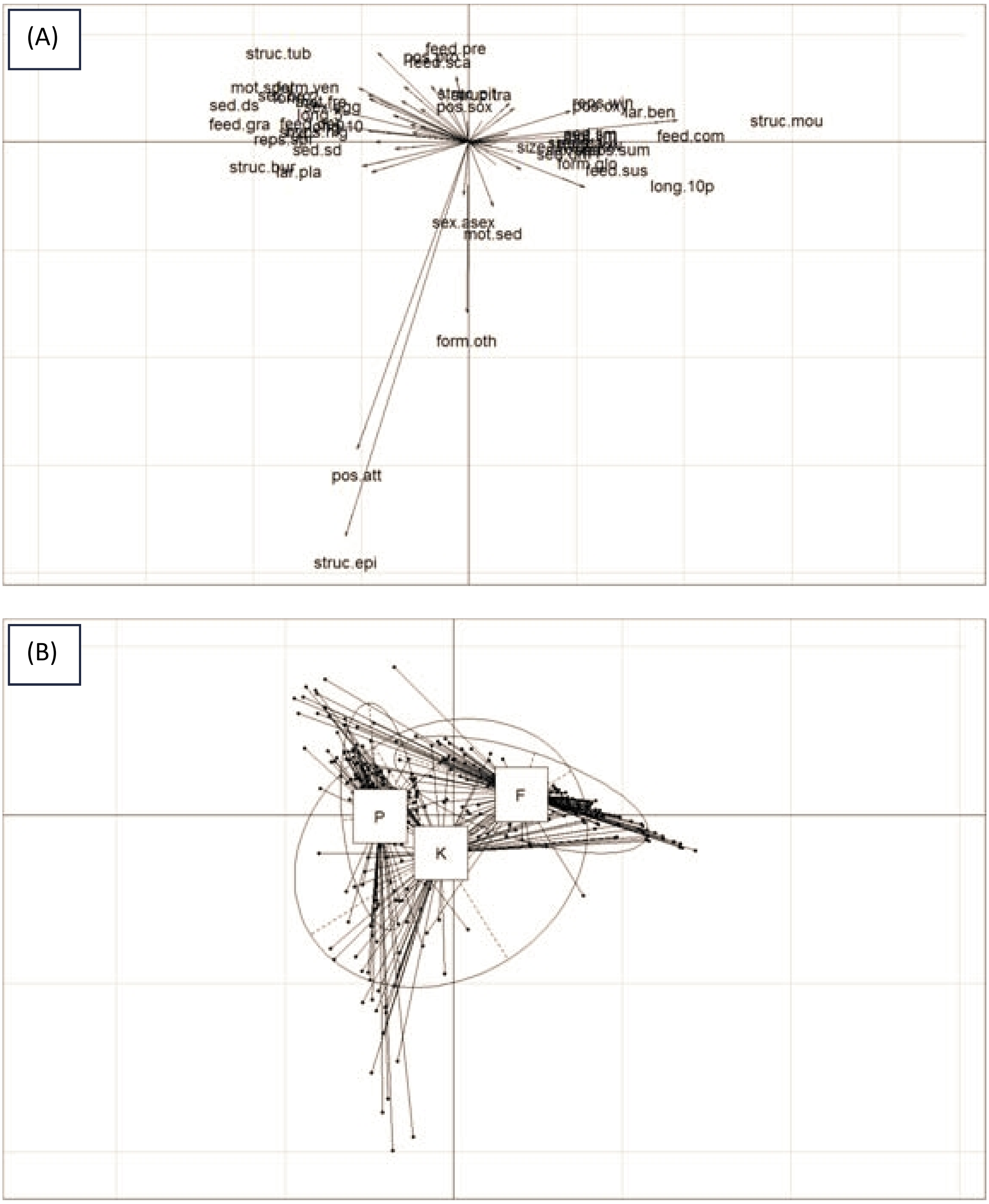


Figure 28: Factorial maps of the FCA based on biomass (afdm) showing the distribution of the traits modalities (A) and the stations (B) along the first two factorial axes. Letters in plot B indicate the centroids of the three areas with labels F = Fehmarnbelt, K: Kadet Trench, P: Pomeranian Bay.

Table 8: Correlation matrix between the first two axes of the FCA and the biological traits for the analyses basing on biomass (left) and abundance (right). Figures in brackets: Variation explained by the corresponding axis).

	Abundance		Biomass	
Trait	Axis 1 (25.0%)	Axis 2 (17.5%)	Axis 1 (32.5%)	Axis 2 (17.0%)
Feeding type	0.05	0.02	0.07	0.02
Body form	0.02	0.02	0.04	0.03
Hypoxia sensitivity	0.00	0.00	0.03	0.00
Larvae type	0.02	0.00	0.21	0.02
Longevity	0.03	0.03	0.19	0.03
Motility	0.03	0.02	0.06	0.04
Position in sediment	0.02	0.03	0.04	0.16
Reproduction season	0.02	0.00	0.07	0.00
Sediment transport	0.01	0.00	0.05	0.00
Reproduction technique	0.03	0.02	0.03	0.01
Body size	0.02	0.01	0.04	0.00
Habitat structures	0.01	0.06	0.15	0.20

4.4. Discussion and conclusions

4.4.1. Comparability of data sets

The three selected areas showed a significant decrease in salinity from the westernmost area in the Fehmarnbelt towards the easternmost area, the Pomeranian Bay. Other environmental gradients such as substrate composition, water depth or organic load of the sediment were present within each of the three areas. In addition, the sampling period of around ten years might result in temporal changes in the community structure, e.g. due to temporal oxygen depletion in the deeper parts of the Kadet Trench and the Fehmarnbelt. It was tried to work against a potential bias in the results by choosing the same sampling period and the same substrate types (medium sand to mud) for each of the areas. A reduction to a single gradient system would have diminished the dataset to an unrepresentative small number of samples. However, the presence of multiple gradients is common in natural systems and always needs to be taken into account when interpreting the results.

4.4.2. Functional composition

The heterogeneity of the dataset was clearly visible in the FCA-plots as the variation explained by the first axes in both approaches (abundance-based and biomass-based) was comparably low. Nevertheless, the separation of the three areas along the first axis reflected the influence of the salinity gradient. The decrease in salinity was followed by a shift from large, long-lived species and species with benthic larvae towards small, short-living species and planktonic larvae. This finding is in agreement with the expectations by Odum (1985, summarized in Elliott & Quintino 2007) that the proportion of k-strategists, mean organisms size and lifespan decrease in stressed ecosystems. Although the mean salinity was rather stable in the Pomeranian Bay and the present community might be called an actual "climax status", the reduced salinity implied a permanent physiological stress for the inhabiting species (e.g. Groth & Theede 1989).

The second axis of the factorial traits-maps of both FCAs basically sub-divided the communities into an epibenthic and an endobenthic part. Although all stations with indications of hard substrates had been excluded from the dataset, usually attached epibenthic species obviously heavily influenced species and functional composition (see chapter 4.3.3). In the Pomeranian Bay and parts of the Kadet Trench/ Darss Sill-area, basically the blue mussel *Mytilus edulis* and the brackish-water barnacle *Amphibalanus improvisus* contributed to this part of the community. An increasing density of floating cohorts of the blue mussel with the epizoic barnacles was a regional consequence of eutrophication (Kube et al. 1997). They are nowadays a dominating part of the soft-bottom community not only in the Pomeranian Bay and cannot be neglected when analysing community functioning. In the western area other species like sea anemones and the tunicates *Dendrodoa grossularia* and *Ciona intestinalis* were common on shell fragments of *Arctica islandica*.

4.4.3. Functional diversity and redundancy

Van der Linden et al. (2012) published a first case study from the estuary of the River Mondego in Portugal. No significant change in functional diversity was found along the analysed environmental gradient. Also in our study, the functional diversity remained stable along the salinity gradient when the analysis was based on abundance, but increased with decreasing salinity in analysis based on biomass. This contradiction was intriguing as species richness and species diversity significantly decreased along the same gradient. This result is directly linked to the shift in composition of the community. While an undisturbed community in fully marine areas is mainly characterised by specialists, the functioning of brackish water communities is determined by the capability of many species to switch between different modes of behaviour if necessary. Many species score on various modalities within most of the traits, resulting in a full functional coverage by including only few species and within relatively small areas (Törnroos & Bonsdorff 2012). This ubiquitous capability is followed by a high functional redundancy in brackish water communities as most modalities are shared by most of the dominant species. Consequently, the communities in brackish waters can be considered to be more robust against a loss of function by the disappearance of single species (Diaz & Cabido 2001).

4.4.4. Differences between abundance and biomass

Biomass is generally regarded as being the more suitable parameter to describe community function. One advantage of the parameter biomass is the lower temporal variability. Additionally, the key functions of a community are much more linked to the biomass than to abundance. The filter feeding capacity of a single large bivalve specimen (e.g. *Arctica islandica*) for example exceeds that of many small polychaetes (e.g. *Euchone papillosa*) tremendously. Thus, the composition of main functional traits of a community is better reflected by the biomass.

The opposite might be true for functional diversity. The diversity-index partly depends on the dominance structure – it is in principle lower if few species dominate the community. With respect to the abundance, this is true for the community in the reduced salinity of the Pomeranian Bay where few small species (especially *Peringia ulvae*) reach high abundance. The result shows the opposite regarding the biomass: the community in the Pomeranian Bay is dominated by several medium-sized bivalves, whereas the biomass in the western areas is dominated by large specimen of the ocean quahog *Arctica islandica*. Consequently, the functional diversity decreases with increasing dominance of large and long-lived species if biomass is used.

4.4.5. Final conclusions

Despite the fact that the results are influenced by high variability due to different gradients, the shift of the functional composition along the salinity gradient is dominant. The dominance of ubiquitous species at least mathematically buffers the loss of highly specialised species in ecosystem functioning. It is followed by a gain in functional redundancy, increasing the robustness of the benthic ecosystem to environmental changes. But the potential capacity of the dominant species in brackish water to change e.g. feeding behaviour might also affect the capability to detect the occurring functional changes along the other natural or anthropogenic induced gradients. This disguise of functional changes by “multifunctional” species has to be considered in the development of indicators based on functional diversity or composition of benthic communities in brackish waters.

5. Third case study: Predicting the biomass of selected bivalve species

The results of this case study have been submitted for publication by A. Darr, M. Gogina & M.L. Zettler to the Journal of Marine Systems entitled "Detecting hot-spots of bivalve biomass in the south-western Baltic Sea".

5.1. Introduction

Bivalves are regarded as an essential part of the benthic community in marine and brackish water systems (Gosling 2003). Especially in brackish water systems, where several important tribes of marine invertebrates do not occur due to reduced salinity, bivalves become more relevant. For instance in the south-western Baltic Sea bivalves often provide more than 80% of the benthic macrofauna biomass in soft-bottom communities (Kube et al. 1996, own observations). Bivalves are an important food source for benthophagous fish species (Brey et al. 1990, Siaulys et al. 2012), sea-birds (Lewis et al. 2007) and other higher trophic levels of the food web. Especially in soft-bottoms they play a predominant role in bentho-pelagic coupling by filtering the water column for nourishment and deposing pseudofaeces on or into the sediment (Graf 1992, Norkko et al. 2001).

Additionally, bivalve species combine several autecological features with potential value for assessment and management purposes. Most adult bivalves are, once settled, more or less sessile and therefore reflect the environmental conditions in the area where they were found. The ocean quahog *Arctica islandica* is among the most long-lived invertebrate species worldwide (Ridgway & Richardson 2010). But also the lifespan of other species like *Astarte elliptica* may exceed 20 years (Trutschler & Samtleben 1988). Therefore, these species do not only provide information on recent environmental conditions, but the state of their population structure may give information on the conditions during the last decades, as well.

However, the calculation of the different functions of benthic bivalves and the application of this information is up to now limited by the punctual knowledge of the distribution of benthic invertebrate species. Within the last decade, habitat suitability modelling became a common tool in benthic ecology (e.g. Glockzin et al. 2009, Gogina et al. 2010b, Reiss et al.

2011). First attempts focussed on the prediction of the probability of occurrence as the distribution of benthic invertebrates heavily varies in spaces and time. Studies focussing on the prediction of abundance or biomass of marine benthic invertebrates are still rare and are often linked with the demands of fishery (Wei et al. 2010, Siaulys et al. 2012). But the intended linkage with ecosystem function as well as the needs for management, e.g. the application of the recently developed HELCOM Biotope Classification System HUB for the Baltic Sea (HELCOM 2013b) requires a quantitative approach.

The actual case study targets the development of quantitative distribution maps of the most common bivalve species within the study area as a basis for the intended quantitative mapping of individual functions.

5.2. Material and Methods

5.2.1. Study area

Due to the highly variable environment including the salinity gradient, the south-western Baltic represents an ambitious area for this kind of studies. Nevertheless, the distribution of benthic invertebrates and their relation to abiotic parameters has already been subject to several studies (Forster & Zettler 2004, Glockzin & Zettler 2008, Gogina et al 2010a, b). The composition of surface sediments mainly results from postglacial processes. Shallow areas along the shore and on top of offshore glacial elevations are characterized by a mosaic of rocks, till, gravel and coarser sand. Substrate in general gets finer with increasing water depth. Muddy sediments dominate the basins and deeper part of trenches. These substrates are widely enriched with organic load. Additional parameters influencing the distribution and condition of benthic bivalves are water temperature and food availability. An important food source is the inflow of freshwater from the larger rivers such as Trave, Warnow and Oder. Negative effects on the population of soft-bottom bivalves are produced by the seasonal oxygen depletion events which occur especially in the deeper areas of the Bay of Kiel, and Mecklenburg Bay and in the Arkona Basin (Friedland et al. 2012).

5.2.2. Sampling and generation of data

Overall, 917 sampling events were included in the analysis. Samples were taken on behalf of different projects between 2004 and 2012. Sampling and sample treatment in the laboratory were performed using standard procedures as described in chapter 2.2.

Biomass is presented and used in models as ash-free dry mass (afdm) in g*m^{-2} for most of the larger bivalve species, but in mg*m^{-2} for smaller species. The blue mussels were not identified on species level as the population in the study area consists of both *Mytilus edulis*, *M. trossulus* and potentially of hybrids (Väinölä & Hvilsom 1991, Väinölä & Strelkov 2011, Riisgård et al. 2013)

Sorting of grain size fractions and presence of hard substrates were available as supplemental information on substrate characteristics provided by the maps of Tauber (2012). Additional environmental parameters, describing the compartment of the water column, were gained from oceanographic models. A regional adaptation of the ERGOM-model was used as source for the predictors light conditions, amount of detritus (sediment ratio) and oxic conditions (Table 9) as described in Neumann (2000) and Friedland et al. (2012). Information on salinity (mean, standard deviation), near-bottom water temperature (summer mean and winter mean) and the strength of near-bottom currents (mean and max shear stress and current velocity) were provided by a regionally adopted GETM-model (Klingbeil et al. 2013).

Table 9: Initially 15 proxies for different environmental variables from different sources were available.

Parameter	**Proxy**	**Unit**	**Source for model building and mapping**
Sediment condition	[1] median grain size [2] sorting	Phi-scaled -	measured per station and Tauber (2012) Tauber (2012)
Substrate type	[3] hard substrate	categories Yes/no	Tauber (2012)
[4] Light condition	Zonation	Categories (photic/aphotic)	Friedland et al. (2012)
[5] water depth	Depth	m	measured per station and IOW map
Salinity condition	[6] mean	psu	Klingbeil et al. (2013)

Parameter	Proxy	Unit	Source for model building and mapping
	[7] standard deviation	psu	Klingbeil et al. (2013)
bottom water temperature	[8] mean summer temperature	°C	Klingbeil et al. (2013)
	[9] mean winter temperature	°C	Friedland (2012)
Exposure to currents	[10] Mean current velocity	$m*s^{-1}$	Klingbeil et al. (2013)
	[11] Maximum current velocity	$m*s^{-1}$	Klingbeil et al. (2013)
	[12] Mean shear stress	Pascal	Klingbeil et al. (2013)
	[13] Maximum shear stress	Pascal	Klingbeil et al. (2013)
Oxygen condition	[14] Frequency of hypoxia	days per year with $O_2 < 2\ ml*l^{-1}$	Friedland et al. (2012)
food availability	[15] sink rate of detritus	$mm\ *year^{-1}$	Friedland et al. (2012)

5.2.3. Modelling process using random forests

All analyses were performed under the frame of the R environment (Version 2.15.2, R development Core Team 2012) using the package randomForest (Version 4.6-7, Liaw & Wiener 2002). Ash-free dry mass (afdm) was chosen as response variable instead of the wet weight in order to ignore anorganic weights like shells. The conversion factor from wet weight to afdm is about 1 : 0.05-0.1, i.e. wet weight is, depending on the species, approximately 10-20 times larger than values given in this study. Biomass values were $\log_{10}$ (x+1) transformed to downscale large values.

The first comprehensive introduction to random forests was given by Breiman (2001), who established the definition “A random forest is a classifier consisting of a collection of tree-structured classifiers with independent identically distributed random vectors and each tree casts a unit vote for the most popular class”. Thus, random forests technique (RF) is an extension of the statistical class of recursive portioning and is based on a set of regression trees.

Classification and Regression Trees (CART) have initially been published by Breiman et al. (1984) and can be applied both to categorical response variables (e.g. communities, classification tree) or numerical response variables (e.g. density of a species, regression tree). The basic principle of the CART-algorithms is to divide the dataset dichotomously into sub-nodes following a specified "if...then...else..." rule (Figure 29). The critical point is to find the "best" rule which is defined as the split point that minimises heterogeneity (Breiman et al. 1984). For numerical variables, least square deviation is generally used as a measure for heterogeneity. After the initial split, this procedure is subsequently repeated for all sub-nodes until a stop criterion is fulfilled. Stop criteria are mandatory, as the tree would otherwise keep on growing until every single value ends up in an own sub-group. They can be defined by setting a maximum number of splits or a minimum amount of increased homogeneity achieved by the split.

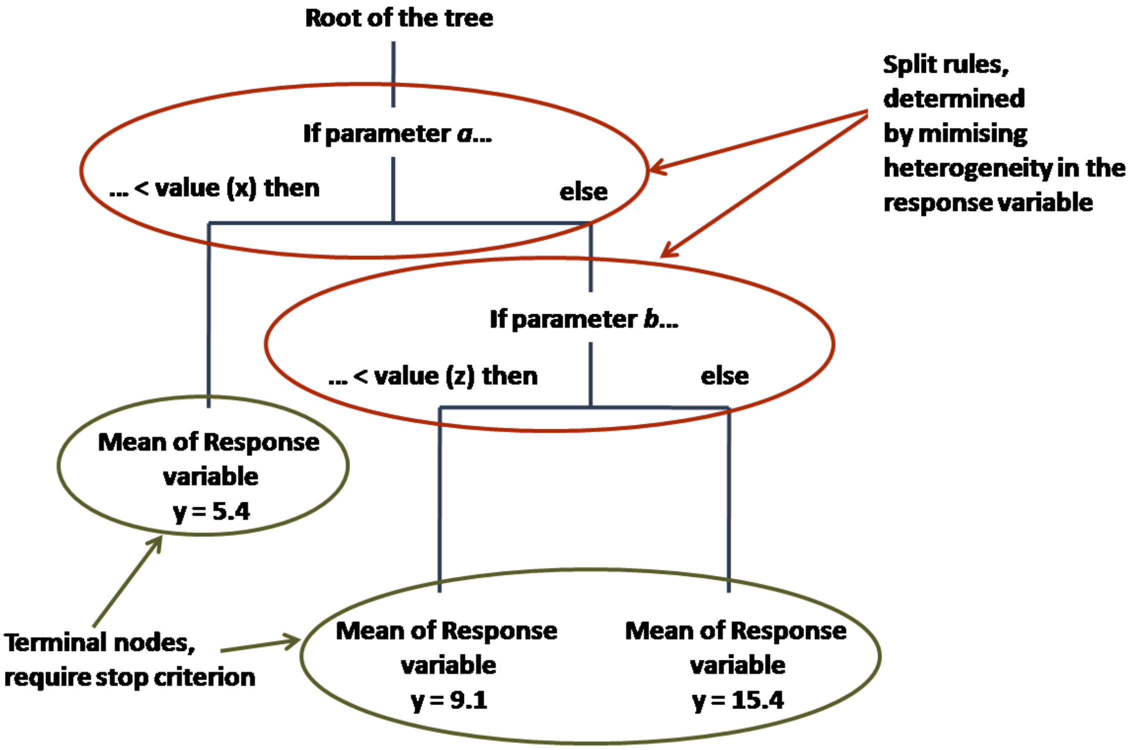


Figure 29: Sketch of a simple regression tree to illustrate the mentioned technical terms. All given numbers are exemplary without true background data.

.

In RF-algorithm a set of *n* randomly built and uncorrelated trees is drawn by randomly selecting two third of the training dataset (with replacement in *n*+1). Unlike in original CART-algorithm, only a subset *m* of the available variables is randomly pre-selected separately for each split. Internal validation is achieved by applying the tree to the remaining third of the training dataset and calculating the deviance of the prediction from the measured values. This error estimation is called “out-of-bag error” as it bases on the samples taken out of the “bag” which includes the entire training dataset before drawing the tree. It is calculated as Mean Squared (prediction) Error ($\mathrm{MSE}_{\mathrm{OOB}}$):

$$MSE_{OOB} = \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2 \tag{5}$$

where $\hat{Y}_i$ is the predicted and Y_i is the measured response value of sample *i*. A second index for internal assessment of the performance of the RF-algorithms is the explained variation:

$$\%var = 1 - \frac{MSE_{OOB}}{\sigma^2} \tag{6}$$

where σ is the unbiased estimate.

It has been proven that this error estimate in unbiased (Breiman 2001) and therefore it is assumed that an external validation is unnecessary for random forests. Nevertheless, some studies recommend an external validation to check for general applicability of the achieved RF-model (e.g. Vincenzi et al. 2011). This thesis follows these recommendations and uses two estimators for assessing the model performance with the external (“validation”) dataset:

(1) the root-mean squared error RMSE which is calculated as the root of the MSE for the validation set:

$$RMSE = \sqrt{MSE_{val}} \tag{7}$$

and (2) the Pearson-correlation coefficient between predicted and measured values.

The station set was randomly sub-divided into the training set (70%) and the validation set (30%, Figure 30) following the proposal of Franklin & Miller (2010). Training and validation set were checked for comparable frequency and mean biomass of the species.

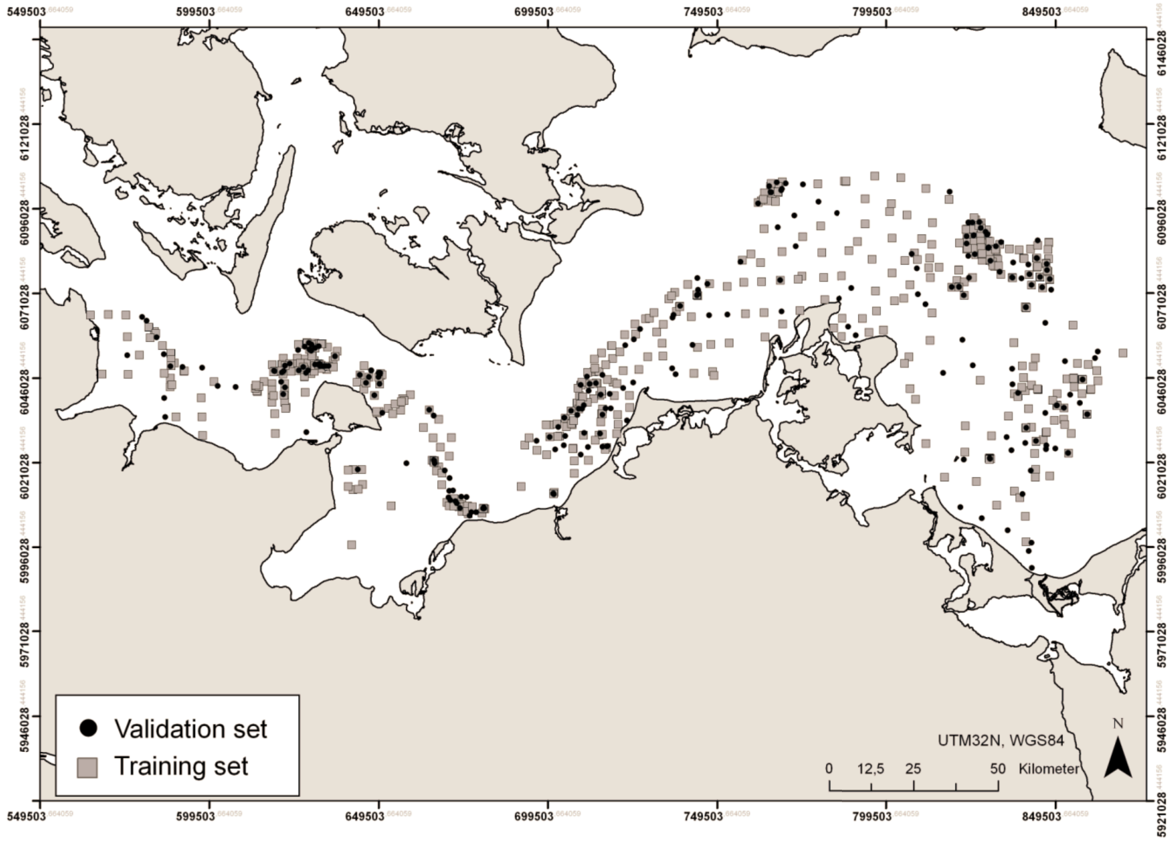


Figure 30: Map of the south-western Baltic Sea illustrating the position of the available stations and their attribution to training or validation set respectively.

As the importance of predictors and the relationship between predictors and response variable are not directly visible in RF in contrast to e.g. the equations of linear models or the split rules in CART-analyses, RF-specific indices were developed (Breiman 2001). One of the most often used measure for the importance of the variables is the increase of mean squared error (INC MSE_{OOB}). This measure is achieved by randomly permuting the values of a predictor within the OOB-dataset and recalculating the MSE_{OOB}. The larger the difference is between the original MSE_{OOB} and the MSE obtained by random permutation of the predictor, the larger is the loss of information due to randomisation. It is concluded that this directly correlates with the importance of the predictor for the set-up of the RF-model. As the INC MSE_{OOB}-values are dependent on the MSE-values of the particular model, the values are not comparable between the models attained for different species. The variable importance only provides information within the individual model. If all predictors share almost equal INC MSE_{OOB}-values, their importance in the RF-algorithm is almost the same. By contrast, single dominant predictors feature considerably larger INC MSE_{OOB}-values than all other predictors.

The relation between the response and a predictor in the obtained RF-algorithm can be shown by partial-dependence plots (e.g. Figure 38). The partial dependence on the y-axis of these plots at value *x* of the selected predictor depicts the average of all predicted responses of the training-set holding this particular value *x*, but vary in the values of all other predictors. Therefore the achieved values on the y-axis are in general much lower than the highest values predicted for the validation or the OOB-data. Nevertheless, these plots sufficiently display the relationship between the predictor and the response variable. Additionally, if the visible pattern of the most important variables in partial-dependence plots is serrated, it might indicate that the relation between predictor and response is not causal in an ecological sense, but that the predictor is used as a proxy for a missing variable.

Random forests are not affected by spatial autocorrelation, as they do not assume independence of the data (Evans et al. 2011). Nevertheless, a spatial bias in the model residuals might be another hint for missing important predictors. Therefore, an *a posteriori* test for spatial bias in the residuals was performed using Moran's Global I (Dormann et al. 2007). For theoretical background and derivation of Moran's Global I the explanations given by Gogina (2010) are recommended. Moran's Global I was calculated using a basic tool of ESRI ArcGIS10.

To avoid overfitting in model building and bias in the measurement of variable importance (Strobl et al. 2007), the number of predictors was *a priori* reduced using variance inflation factors (Zuur et al. 2010) and correlation analysis. From the 15 initially available predictors, only phi-scaled median grain size [1], sorting of the sediment fractions [2], occurrence of hard substrate [3], mean summer temperature [8], mean number of days with hypoxia [14], water depths [5], mean salinity [6], mean current velocity [10] and mean detritus rate [15] were chosen as predictors (numbers in brackets refer to those in Table 9). As the occurrence of stones is only relevant for epibenthic hard substrate species, this predictor was disregarded for soft-bottom species.

Separate RF-models were built for each species, varying in the number of available variables at each split between 2, 3, 4 and 5. Number of maximum trees was set to 500. As described above, the RF-algorithm produces a set of trees by randomly selecting a subset of available variables at each split. Consequently, (1) the performance of the resulting model can hardly be reproduced as the random choice is very likely to differ between the two runs and (2) the resulting model cannot be assumed to be the "best" model. Therefore, one should avoid choosing the best model of a randomly produced series of RF-models, as still one of the "unchecked" combinations might perform better. It is recommended to produce a series of RF-models, to calculate the mean and use standard deviation as a measure of variability. Thus, the procedure described above was repeated 5 times resulting in a set of 20 models for each species.

Mapping on a grid base with a cell size of 1000*1000 m was performed using ArcGIS10.

5.3. Results

5.3.1. Identified bivalve species

Overall, 29 bivalve species were identified in the samples (Table 10). Eight species occurred occasionally in very low densities in the westernmost part (Frequency < 1%: *Angulus tenuis, Barnea candida, Cerastoderma edule, Ensis directus, Musculus subpictus, Modiolus modiolus, Tellimya ferruginosa, Thracia pubescens*). These species were a priori excluded from modelling as the study area is situated at the border of their natural distribution.

Macoma balthica reached highest frequency in the data set and occurs in almost two third of the stations. A frequency of more than 30% was reached by *Mya arenaria*, *Mytilus* spp. (*M. edulis* and *M. trossulus*) and *Arctica islandica* with the last two revealing highest local biomass (afdm: > 200 g*m^{-2}). However, mean biomass of *Mytilus* spp. at places where it occured was much lower (10.8 g*m^{-2}) in comparison to *A. islandica* (30.0 g*m^{-2}). *Astarte borealis*, *A. elliptica*) and *Mya arenaria* showed a mean biomass (afdm) of more or about 5 g*m^{-2} and local maxima of more than 50 g*m^{-2}.

Another eleven species also only occurred in the Bay of Kiel and the Fehmarnbelt and were found in less than 10% of the samples (Table 10).

Table 10: Frequency within the dataset, mean and maximal biomass (afdm) of 21 bivalve species pre-selected for modelling

Species	Frequency	Biomass ($g*m^{-2}$) Mean[1]	 Max
Abra alba	15.0%	0.62	7.00
Arctica islandica	35.4%	30.05	267.30
Astarte borealis	24.9%	8.37	63.00
Astarte elliptica	12.6%	6.79	120.30
Astarte montagui	6.0%	0.99	10.40
Cerastoderma glaucum	20.4%	1.29	13.10
Corbula gibba	5.8%	0.22	2.10
Hiatella arctica	1.9%	0.38	1.00
Kurtiella bidentata	14.7%	0.15	0.80
Macoma balthica	61.6%	2.86	20.20
Macoma calcarea	6.7%	0.94	7.10
Musculus discors	1.7%	0.19	0.70
Musculus niger	3.4%	0.20	0.70
Mya arenaria	37.3%	4.88	76.70
Mya truncata	4.4%	3.21	23.40
Mytilus spp.	31.5%	10.81	219.50
Parvicardium pinnulatum	11.9%	0.42	8.40
Parvicardium scabrum	1.3%	0.13	0.30
Phaxas pellucidus	2.4%	0.12	0.30
Spisula subtruncata	3.2%	0.03	0.18
Thracia phaseolina	3.7%	0.03	0.13

[1]: mean biomass where it occurs, i.e. neglecting absence data

5.3.2. Single species Model

For half of the 21 evaluated species, none of the models was able to detect any relation between the predictors and the response variables. This was true for mostly smaller epibenthic species (*Musculus* spp., *Hiatella arctica*) and species with low frequency and an infrequent appearance in the western part of the study area (*Phaxas pellucidus, Spisula subtruncata, Thracia phaseolina*). Additionally, the performance of the models of some more frequent species like *Abra alba* and *Kurtiella bidentata* were unsatisfying with an explained variation within the training set < 10% or a weak correlation with the validation set (Table 11).

Best fitting with the training set was achieved by the models for *Macoma balthica* (64.7 ± 0.4% variation explained, MSE: 0.04 ± 0.0004). Highest correlations with the validation sets (>0.70, including a reasonable low RMSE) were reached for *Astarte borealis, Cerastoderma glaucum, Corbula gibba* and *Macoma calcarea*. Acceptable performance was also reached for *A. islandica, Astarte elliptica* and *Mya arenaria*. The models for *Mytilus* spp. showed acceptable performance within the training set (MSE: 0.11 ± 0.0002, variation explained: 41.9 ± 1.0%), but revealed a weak correlation with the validation set. Opposite to that, the correlation with the validation set was acceptable for the models of *Astarte montagui* whereas the performance with the training set was weak. However, the models for the 10 specified species were regarded to be of acceptable quality and were included in further analysis.

Salinity was identified as being the most important variable in the models of almost all species except for *M. arenaria* (Table 12). For this species, water depth and median grain size were the most influencing variables. One of these two parameters was also the second most important parameter for most of the other species. *Corbula gibba* and *C. glaucum* did not show any clear dominance of a single parameter. Several factors seemed to be of similar importance. The amount of available detritus was an important parameter only in the models for *M. arenaria*. The parameters with lowest impact on the models of all species were sorting of the sediment and oxygen conditions (mean days of hypoxia per year).

Highest biomass of the ocean quahog *A. islandica* was predicted to occur in the deeper parts of Kiel and Mecklenburg Bay with a biomass (afdm) of up to 72.2 g*m^{-2}, whereas it was absent in the most shallowest parts throughout the study area (Figure 31A). Confidence of the prediction was lowest in the areas with highest predicted biomass, near the Darss Sill and west of the Island of Fehmarn, whereas differences between the predictions where comparably low in the Mecklenburg Bay (Figure 31B).

The hotspots for *Astarte borealis* were identified in the west of the Island of Fehmarn and parts of the Bay of Kiel (max. 12.0 g*m^{-2}, Figure 32 [The confidence plots for the remaining species are available in the appendix.]). An almost opposite distribution with a focus on more eastern and shallower areas was predicted for *C. glaucum* and *M. arenaria*. While for the latter highest biomass (max. 15.6 g*m^{-2}) occurred in the shallowest areas close to the

River Oder mouth and some spots to the west of the Island of Rügen (Figure 35), the hotspot of *C. glaucum* was predicted to be around the Oderbank (Figure 33). Highest biomass of *Macoma balthica* (max. 8.3 g*m^{-2}) was predicted to occur along the slopes towards the Arkona Basin. In western parts of the study area, the species only appeared in shallower parts (Figure 34). The distribution of blue-mussels (*Mytilus* spp.) mainly depended on the availability of hard-substrates. Highest biomass values were displayed to be situated on the reef structures around the Island of Rügen and on the Adlergrund (Figure 36).

Table 11: Measure of performance of the models applied to the training set and the validation set presented as mean ± standard deviation

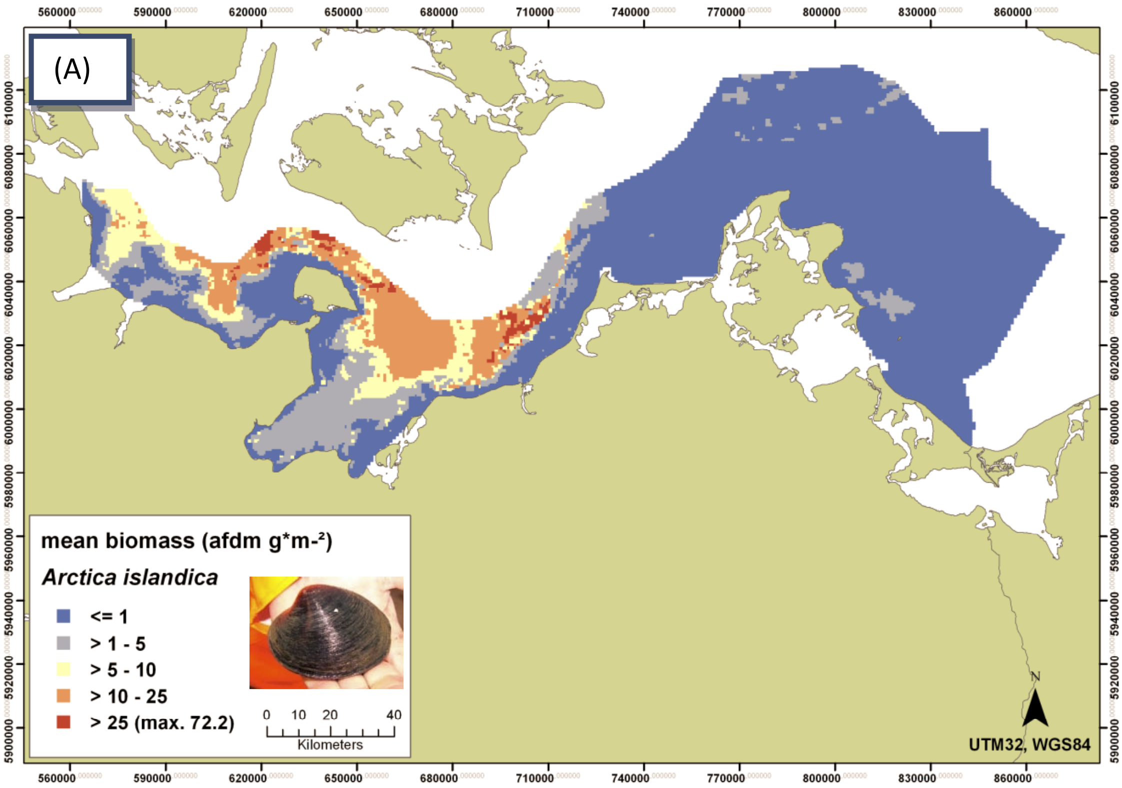
	training set		validation set	
species	**MSE**	**%Var**	**RMSE**	**Correlation**
Arctica islandica	0.20 ± 0.002	55.16 ± 0.51	19.09 ± 0.38	0.58 ± 0.02
Astarte borealis	0.08 ± 0.002	50.26 ± 1.08	2.95 ± 0.03	0.71 ± 0.01
Astarte elliptica	0.03 ± 0.001	52.09 ± 1.18	8.02 ± 0.03	0.35 ± 0.01
Astarte montagui	0.35 ± 0.01	19.70 ± 1.70	731.25 ± 5.07	0.48 ± 0.01
Cerastoderma glaucum	0.01 ± 0.0002	51.92 ± 0.98	0.77 ± 0.003	0.72 ± 0.01
Corbula gibba	0.17 ± 0.002	39.40 ± 0.60	44.49 ± 1.09	0.80 ± 0.01
Macoma balthica	0.04 ± 0.0004	64.69 ± 0.42	2.04 ± 0.02	0.70 ± 0.004
Macoma calcarea	0.25 ± 0.003	43.97 ± 0.62	358.59 ± 3.5	0.80 ± 0.02
Mya arenaria	0.07 ± 0.0005	48.40 ± 0.38	4.52 ± 0.05	0.60 ± 0.01
Mytilus spp.	0.11 ± 0.002	41.91 ± 1.03	15.27 ± 0.04	0.22 ± 0.01

MSE: mean of squared error, %Var: Variation explained by the model, RMSE: root-mean squared error

Table 12: Importance of the predictors expressed as increase of mean of squared error (mean ± standard deviation). The most important parameters for the individual species model are presented in bold numbers.

Species	phi	sort.	hard	T°C	hypoxia	depth	psu	veloc.	detritus
Arctica islandica	11,75 ± 1.09	13.57 ± 1.55	-	22.10 ± 1.54	14.79 ± 1.08	**35.77 ± 2.01**	**50.03 ± 6.44**	19.60 ± 0.99	16.45 ± 2.75
Astarte borealis	**33.82 ± 4.17**	12.30 ± 2.24	-	20.62 ± 2.35	9.89 ± 0.72	21.6 ± 1.19	**41.59 ± 4.82**	22.13 ± 1.14	20.21 ± 0.69
Astarte elliptica	19.54 ± 1.60	11.47 ± 1.38	-	17.06 ± 5.90	5.13 ± 0.89	15.3 ± 0.97	**34.15 ± 4.74**	19.49 ± 1.66	12.42 ± 0.93
Astarte montagui	14.62 ± 1.07	9.56 ± 0.63	-	16.50 ± 0.55	7.49 ± 0.78	16.25 ± 0.59	**27.74 ± 2.93**	14.17 ± 1.25	16.50 ± 0.87
Cerastoderma glaucum	**18.29 ± 1.33**	6.17 ± 1.94	-	**17.44 ± 1.64**	5.70 ± 0.82	**17.73 ± 1.64**	**16.31 ± 0.95**	12.27 ± 1.08	11.25 ± 1.14
Corbula gibba	**24.51 ± 1.44**	16.35 ± 1.24	-	24.02 ± 1.52	9.85 ± 0.61	20.24 ± 0.95	19.94 ± 0.84	**28.01 ± 2.98**	21.47 ± 1.32
Macoma balthica	27.55 ± 1.09	28.80 ± 2.15	-	29.47 ± 0.88	19.82 ± 1.04	**53.84 ± 8.84**	**51.02 ± 6.72**	22.40 ± 0.86	20.41 ± 1.79
Macoma calcarea	16.37 ± 1.09	8.49 ± 1.62	-	15.43 ± 1.05	1.83 ± 1.02	17.14 ± 1.15	34.88 ± 4.78	12.00 ± 0.80	12.34 ± 1.01
Mya arenaria	**29.14 ± 1.68**	22.36 ± 1.55	-	22.41 ± 1.10	7.25 ± 2.05	**39.16 ± 6.18**	25.71 ± 0.98	19.54 ± 1.13	**27.90 ± 2.60**
Mytilus spp.	**40.04 ± 3.69**	-	22.65 ± 1.18		-	22.37 ± 0.89	**52.88 ± 1.81**	-	-

Phi: median grain size (phi-scaled), sort.: sorting of sediment, hard: presence of hard substrates, T°C: mean summer temperature, hypoxia: days with hypoxia (< 2ml/l), psu: mean salinity, veloc: mean current velocity, detritus: mean sedimentation rate of detritus



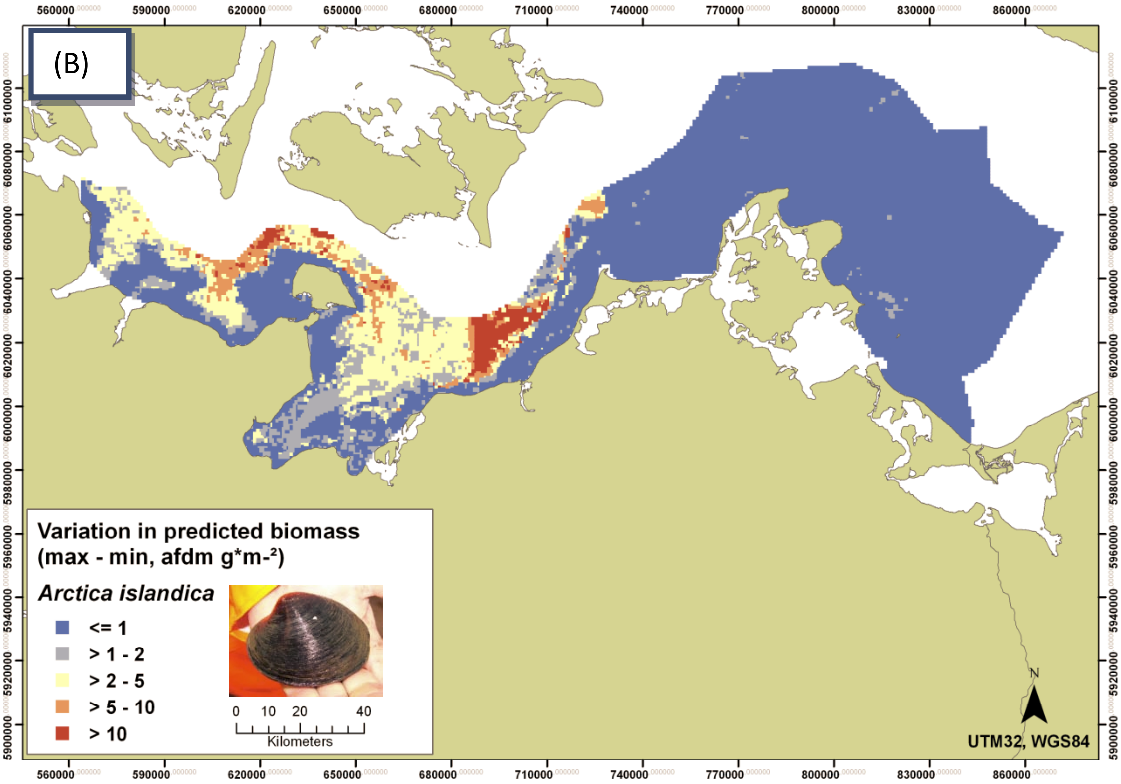


Figure 31: Mean predicted biomass for *Arctica islandica* (A) and the variation of the prediction as a measure of confidence (B) (species photo: IOW)

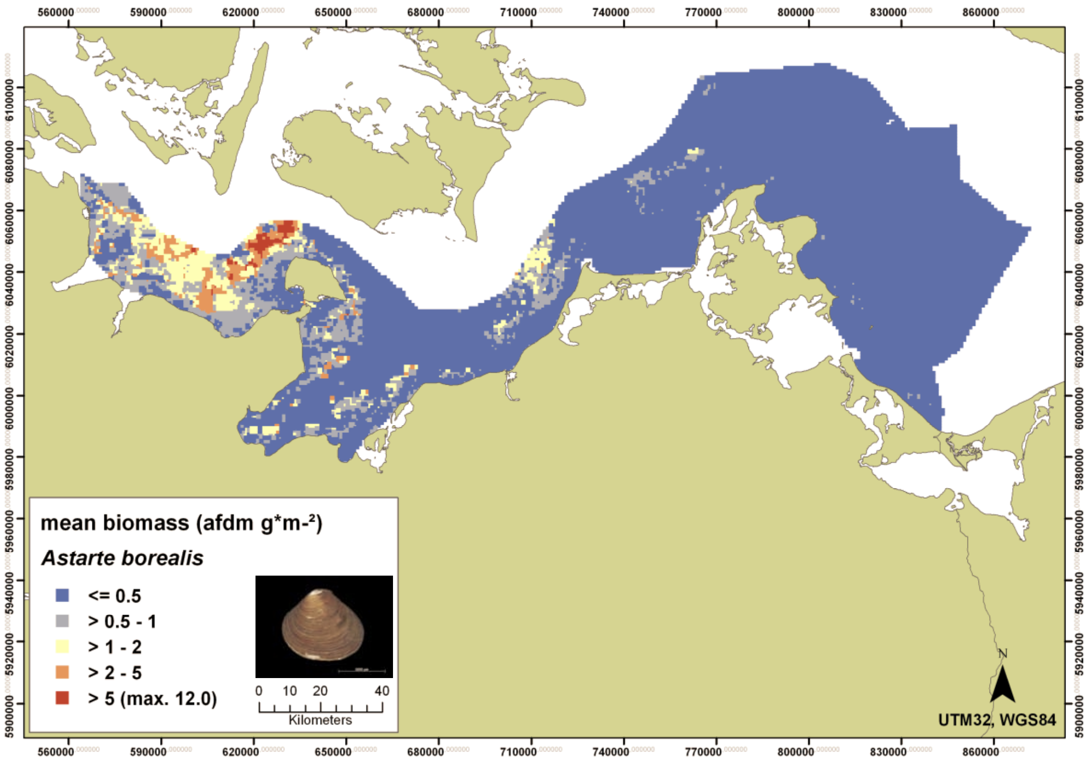


Figure 32: Mean predicted biomass of *Astarte borealis* (species photo: IOW)

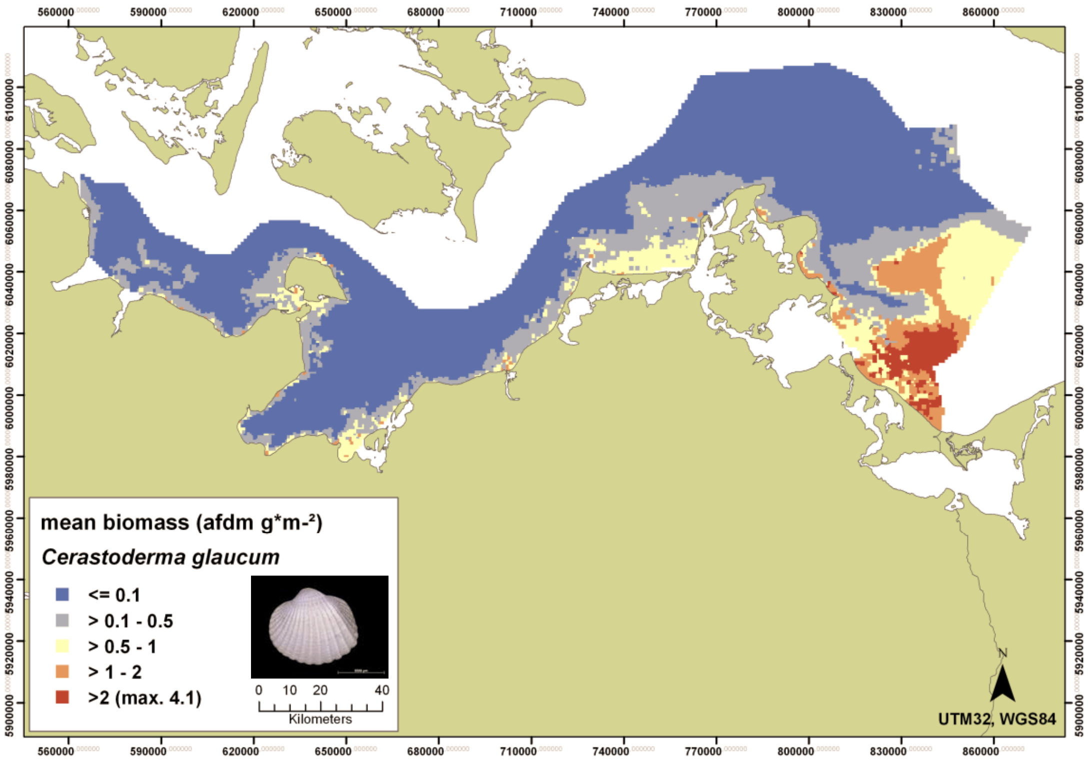


Figure 33: Mean predicted biomass of *Cerastoderma glaucum* (species photo: IOW)

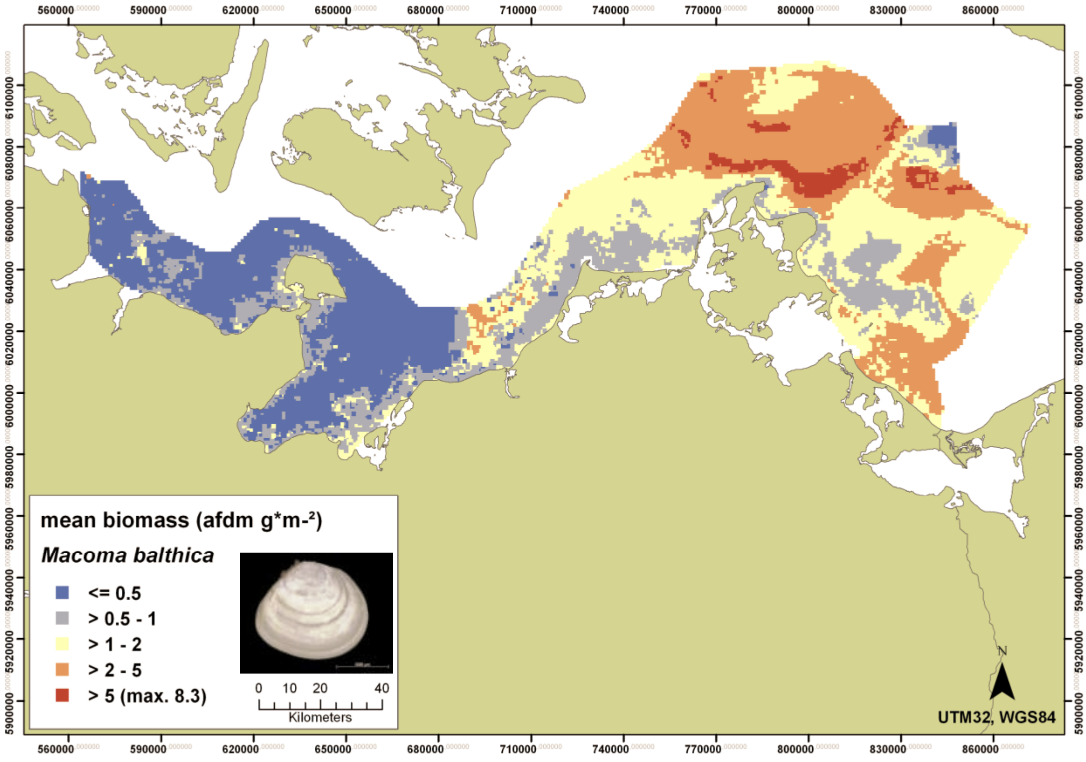


Figure 34: Mean predicted biomass of *Macoma balthica* (species photo: IOW)

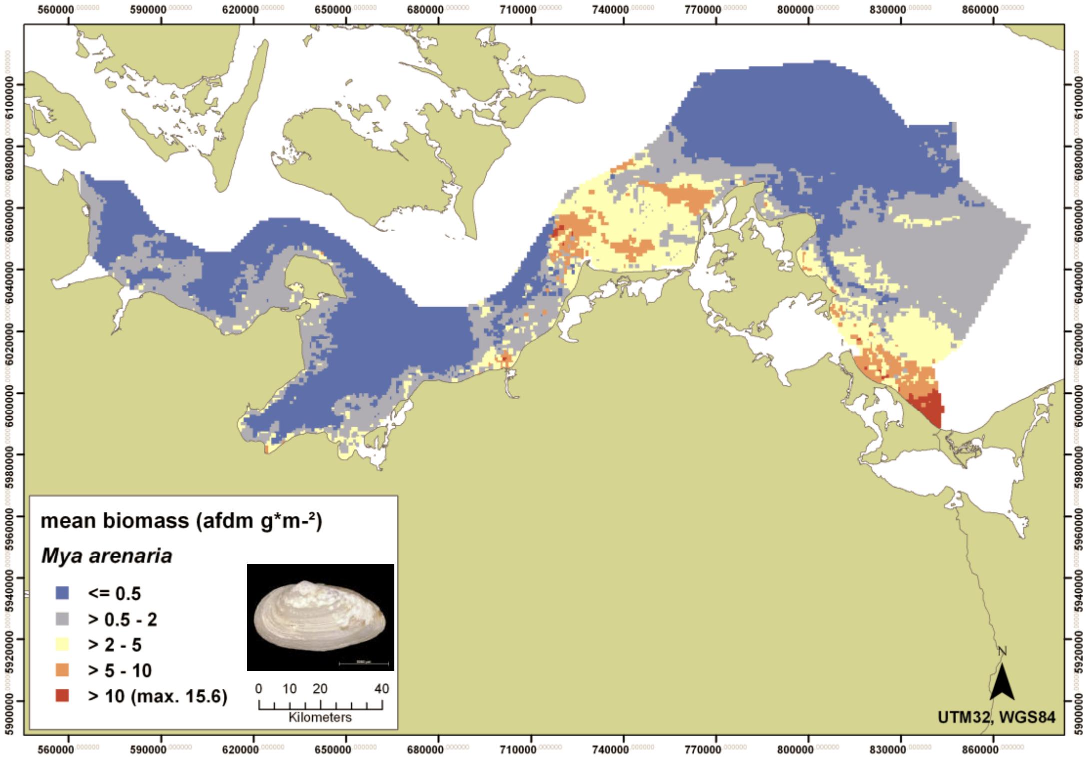


Figure 35: Mean predicted biomass of *Mya arenaria* (species photo: IOW)

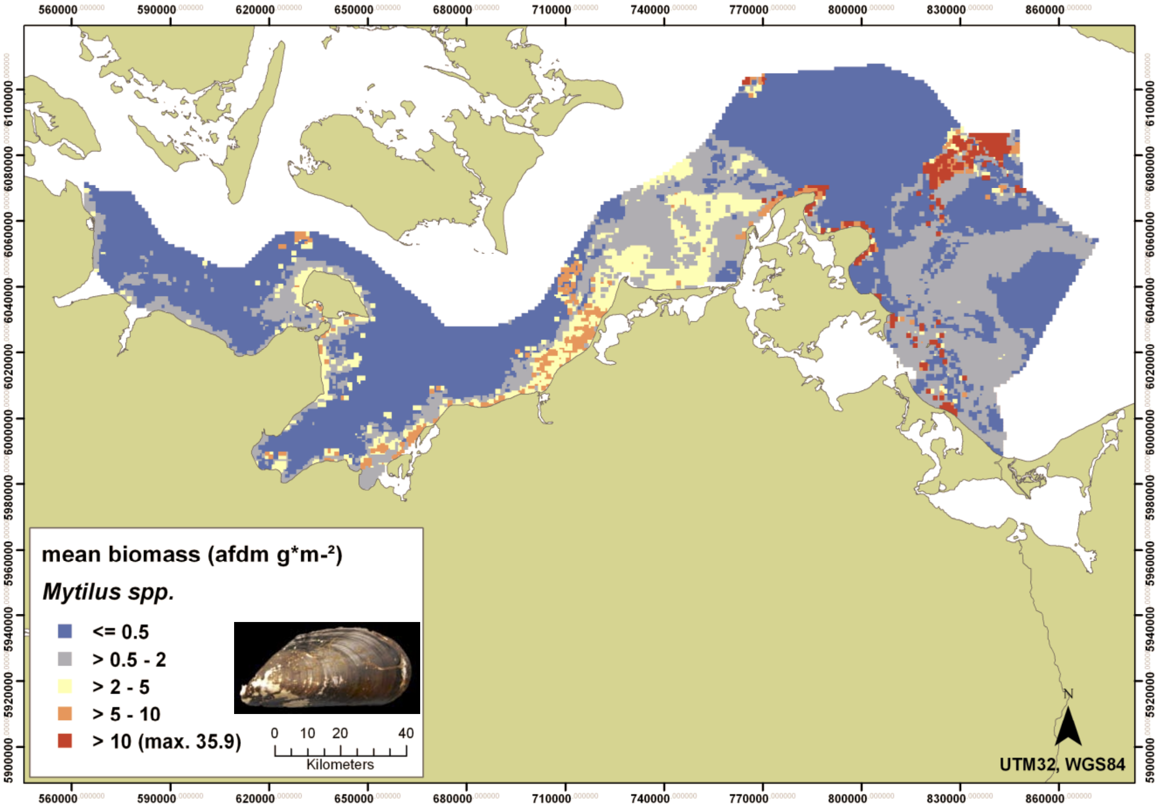


Figure 36: Mean predicted biomass of *Mytilus* spp. (species photo: IOW)

5.3.3. Detection of hotspots

For detection of the hotspots of bivalve biomass in the study area, the predicted biomass of ten species (listed in Table 12) was summed up (Figure 37). Highest biomass (afdm) of more than 25 g*m^{-2} up to 84.5 g*m^{-2} was predicted to occur in deeper parts of the Bay of Kiel, the Fehmarnbelt and the western part of the Kadet Trench as well as on the reefs around the Island of Rügen and on the Adlergrund. Bivalve biomass in the Pomeranian Bay was estimated to be highest close to the mouth of the River Oder. Lowest Bivalve biomass values were calculated for the deepest parts of the inner Bay of Lübeck, deepest parts of the Arkona Basin and the shallow areas of the Bay of Kiel.

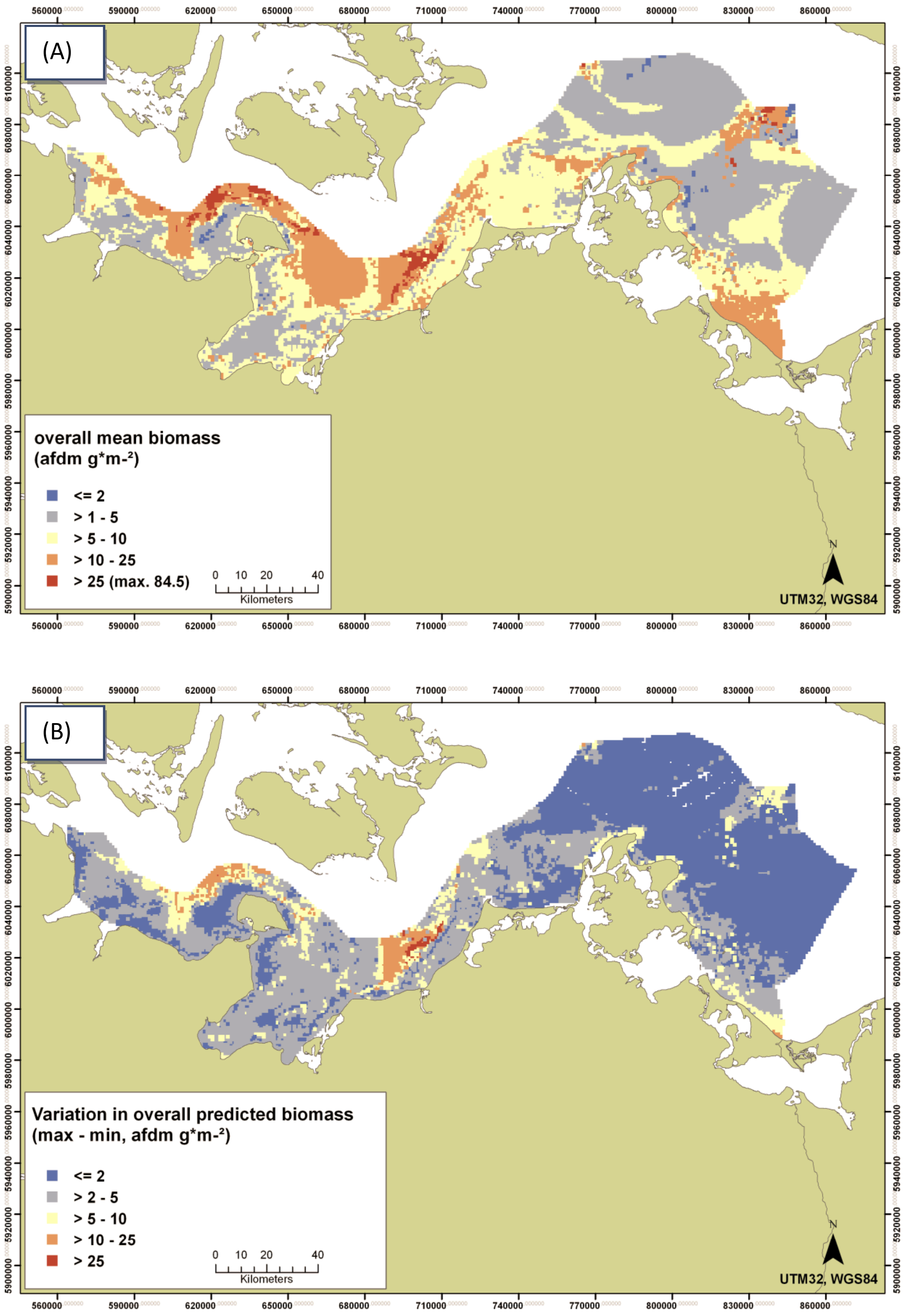


Figure 37: Predicted mean overall bivalve biomass (A) und variation in the predictions as a measure of confidence of the prediction (B)

5.4. Discussion and conclusions

5.4.1. Model performance

Several studies have demonstrated that random forests are not only a reliable tool to predict and map general pattern of the distribution of different species groups (e.g. Cutler et al. 2007, Swatantran et al. 2012, Musters et al. 2013, Olaya-Marin et al. 2013), but moreover, it was shown that the predictions of random forests were more accurate than those of most other available techniques also for predicting the distribution of marine benthic invertebrates (e.g. Reiss et al. 2011).

The mean variation within the training set explained by the selected RF-models ranged from 39.4% to 64.7% except for *A. montagui* which was explicitly lower. The correlation coefficient between predicted values and the measured values in the validation set lied within the range of 0.35 and > 0.80 for the most important species of the study area except for *Mytilus* spp. Considering the natural variability of invertebrates' density within a habitat (Thrush et al. 1994) and the variability added by including data from almost one decade and different seasons, the attained models performed remarkably well. This might indicate both rather stable distribution patterns of bivalve biomass over the last ten years and the suitability of the used modelling techniques.

No significant bias was detected in the residuals ($p > 0.05$ in Moran's I test for spatial autocorrelation) except for *Astarte borealis*. As also for this species spatial autocorrelation of the residuals was small ($Z = 2.83$, $p < 0.001$), it can be assumed that the dependency of the selected species on the environment was well reflected by the chosen proxies. Nevertheless, the correlation might be further improved by adding or substituting some of the predictors. The variable importance measure detected salinity, water depth and median grain size (d50) as the most important predictors for most species. This finding is consistent with the results of earlier studies, analysing the probability of occurrence of benthic invertebrate species within the study area (Glockzin et al. 2009, Gogina et al. 2010b). However, the measures of variable importance in RF- models only show tendencies of the true correlation between the response variable and the individual predictors illustrated e.g. in partial dependence plots (Figure 38). Mean salinity was used as proxy for salinity conditions although the variability of salinity is known to have also an important impact on the physiology of benthic invertebrates (Atrill 2002). But as the two parameters were derived from the same model and were highly auto-correlated, only one of them could be included in the model. Median grain size and degree of sorting provided by Tauber (2012) were the only proxies for sediment characteristics that were available for mapping. In general, it may be expected that the composition of different sediment fractions (e.g. silt-content, gravel-content) is more important than the median grain size. Also the importance of the organic load of the sediment was highlighted in several studies (e.g. Hyland et al.2005, Magni et al. 2009, Rakocinski 2012).

Food availability is often neglected in habitat suitability models, as sufficient data is also rarely available in the required resolution. Nevertheless, this parameter is of major importance regarding the biomass and might outreach those of salinity or substrate characteristics on a regional scale (Rosenberg 1995, Kube et al. 1996). In this study, a modelled detritus accumulation rate was involved as a proxy for food availability. The underlying assumptions by Friedland et al. (2012) are first estimations for this parameter and new approaches might improve the suitability of this parameter in the models. Also interspecific interactions were not reflected by the chosen proxies, although they heavily influence the occurrence and density of all species (Soberón 2007). Especially predation and competition for space may have major impacts on the biomass of bivalve species.

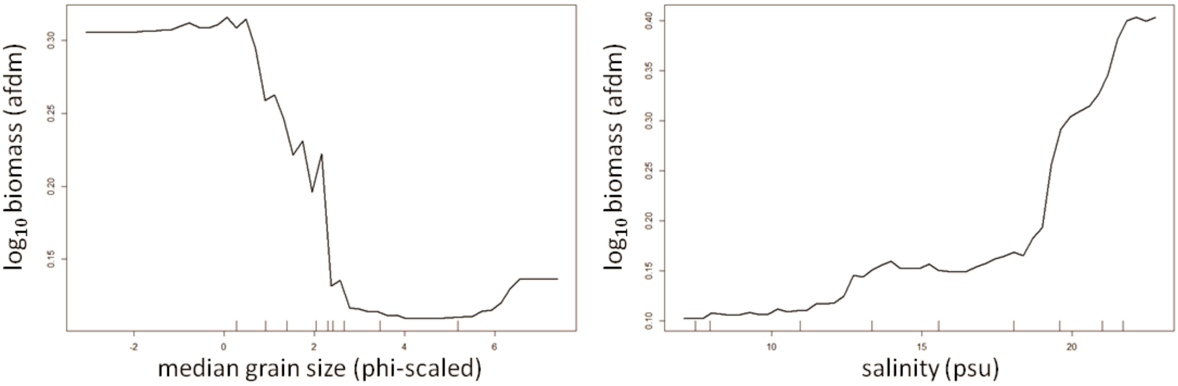


Figure 38: Partial dependence plot of the biomass of *Astarte borealis* (log$_{10}$ afdm, in g*m^{-2}) on median grain size (phi-scaled, left) and salinity (psu, right)

5.4.2. Single species models

For almost two third of the identified bivalve species, no satisfying model could be achieved. Some of these species only occur occasionally in the western part of the study area (e.g. *Phaxas pellucidus, Thracia phaseolina, Spisula subtruncata*) as their appearance strongly depends on larval supply with the inflow of euhaline waters. These species are not able to establish autochthon populations within the study area and frequently disappear in years with lower salinities (own observations). A second group of species were small, short-lived species like *Kurtiella bidentata*. Patterns in biomass distribution of these species are hardly detectable as they are highly variable due to the strong dependency on the variability of salinity. This variability was not reflected in the available parameter for salinity conditions, which aggregates over a time scale of 7 years. The third group of species without satisfying model performance comprised epibenthic bivalves settling on hard-substrates, macrophytes or commensalism on ascidians (e.g. *Hiatella arctica, Musculus* spp.). Information on these parameters was not available for all stations. The map by Tauber (2012) provides only coarse information on the presence of stones and other hard substrates without detailed information on their size or density which strongly limits its correlation with the density of the associated fauna. Additionally, as hard-substrates are randomly and not quantitatively sampled by Van-Veen grabs, the methodical error in sampling procedure is tremendous. However, specific sampling of hard-substrates by divers in offshore areas down to 35 m (Adlergrund, Kriegers Flak) is cost and time-consuming and impossible in areas with high ship-traffic (Kadet Trench).

This weakness also severely affects the model performance of the most common epibenthic bivalve of the study area, the blue mussels *Mytilus* spp. The detected importance of the variable “hard substrate” in the model is rather low (Table 12). Complexity in the distribution of blue mussels in the study area is added by their ability to survive after their detachment from the hard-substrates. The loose, floating conglomerates often aggregate on soft-bottoms close to the originating hard-substrates or in areas with low currents, disabling the conglomerates to keep on flowing. These conglomerates are randomly sampled by the used method and their distribution is hardly linked with any of the available parameters. Nevertheless, the model achieved a good fitting in the general distributional pattern e.g. by detecting the reef structures in the eastern part of the study area as hotspots (Figure 39).

The largest underestimations of the model (negative numbers in Figure 39) occur at the edge of the hotspot area, which might indicate an underestimation of the extent of the area with high blue mussel biomass. On the other hand, large positive numbers (overestimation) mainly occur in heterogeneous areas and within the predicted hotspot areas. This might be both an effect of local spatial variability in the density of blue mussels and of the sampling error produced by the focus on soft-substrates also within reef areas.

An alternative approach to avoid these problems is to exclude substrate characteristics and to relate the distribution of blue mussels on food availability and hydrodynamic models as proposed by Møhlenberg & Rasmussen in Skov et al. (2012). They predict hotspots of *Mytilus*-biomass in the actual study area along the Darss Sill and on the Oderbank whereas the predicted *Mytilus*-index is comparably low on the Adlergrund and on the reefs around the Island of Rügen. This result supports the need for reliable data on spatial distribution and density of hard-substrates.

Pomeranian Bay and adjacent areas were identified as main distributional areas of the three soft-bottom species *C. glaucum*, *M. balthica* and *M. arenaria*. While highest biomass of *Mya arenaria* and *C. glaucum* were predicted for the shallow sandy area in the Pomeranian Bay and the shallowest areas along the shore-line in large parts of the study area, *Macoma balthica* seems to reach high biomass both in the muddy substrate of the Arkona Basin and in the shallow sandy area close to the mouth of the river Oder. These distributional patterns coincide with the findings of previous studies from Kube et al (1997), Forster & Zettler (2004), Glockzin & Zettler (2008) and Gogina et al. (2010b). Kube et al. (1997) described the increase of the biomass of *M. arenaria* between the 1960s and 1990s as a consequence of the higher nutrient load of the Oder plume. Consequently, they detect highest biomass of *M. arenaria* close to the mouth of the Oder. Kube et al. (1997) and Forster & Zettler (2004) described much higher local biomass of *M. arenaria* than predicted by the model presented in this study (afdm > 75 $g*m^{-2}$ in Oder mouth and > 30 $g*m^{-2}$ close to Rostock and in the Kadet Trench were reported). This discrepancy might be caused by a different statistical approach as both studies used interpolation methods to display spatial distribution of invertebrate biomass. However, Powilleit et al. (1996) calculated a mean biomass (afdm) of 20 $g\ m^{-2}$ for parts of the Pomeranian Bay, which approximates the results of the present study. The biomass distribution of *M. balthica* and *C. glaucum* had to a lesser extent been effected by eutrophication (Kube et al. 1996). Although *C. glaucum* is known to tolerate higher organic load, it is solely found on clear sand in the offshore waters of the Baltic Sea (Zettler et al. 2013c). By contrast, *M. balthica* was found in high densities both on clear sands and in muddy areas. This species is known to be able to switch between suspension feeding and deposit feeding depending on food availability and current flow (e.g. in Petterson & Skilleter 1994). This behavioural or even genetic adaptation (Nikula et al. 2008) enables *M. balthica* to compete against a variety of other invertebrate species in different habitats in brackish waters.

The arctic-boreal origin of *Arctica islandica* and *Astarte borealis* was reflected in the occurrence of both species in deeper areas with lower summer temperatures and the polyhaline or β-mesohaline salinity. Zettler (2002) described a scattered distribution of *A. borealis* in the Mecklenburg Bay and adjacent areas without a clear substrate preference and local biomasses (afdm) of 5-16 $g*m^{-2}$. Although the species is frequently found on muddy and sandy substrates throughout the southern Baltic, medium sand seems to be preferred in our study area (Gogina et al. 2010a). Mud in deep basins with a high risk of oxygen depletion is avoided. In contrast, *A. islandica* is known to be resistant against oxygen depletion. But if oxygen depletion events last too long or are too frequent, successful larvae recruitment is prohibited, the population over-ages and finally disappears. This phenomenon has already been described for part of the inner Bay of Lübeck by Zettler et al. (2001) and was reflected by the model with considerable lower predicted biomass than in other areas with comparable salinity and substrate characteristics.

In conclusion, the predictions of all selected species showed a good fit to the general pattern of biomass distribution as described in earlier studies within the study area and provide a profound basis for the cumulative biomass map.

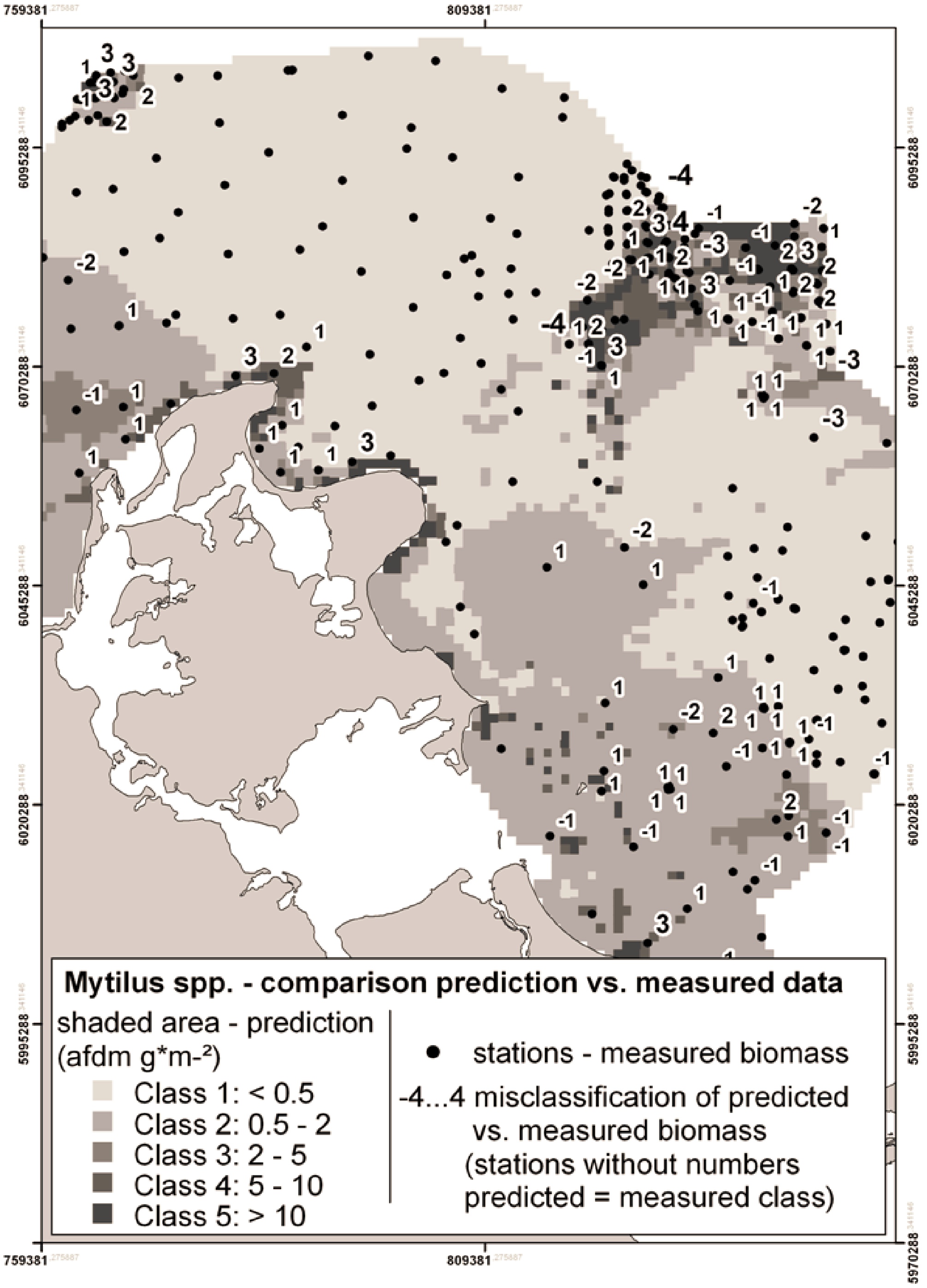


Figure 39: Comparison of the mean prediction for blue-mussel biomass (shaded area) and measured biomass at the station in the eastern part of the study area. Negative values indicate an underestimation of the blue mussel biomass by the model at the stations, positive values an over estimation. The larger the numbers are the larger is the difference between predicted and measured biomass class.

5.4.3. Cumulative biomass map

The cumulative biomass map includes the predicted biomass of ten of the most important bivalve species in the south-western Baltic. Two different hotspots of bivalve biomass are visible: the muddy substrate of the Bay of Kiel and parts of the Mecklenburg Bay, including the southern Kadet trench and the reef structures of the Adlergrund and the Island of Rügen. As the genesis of the bivalve biomass totally differs between these hotspots, one should not confuse benthic biomass with benthic productivity. The ocean quahog *A. islandica* is dominant in the muddy substrates in the deep western part of the study area. It is a slowly growing species, reaching the vertex of its growing curve in the study area after 40-50 years (Zettler et al. 2001). On the other hand, the blue highest biomass on the shallow reefs in the western part of the investigated area was mainly reached by the blue mussels *Mytilus* spp. Blue mussels are rather fast growing bivalves (Bayne & Worrall 1980).

The influence of food availability on the distribution of bivalve species is not well pronounced in the model results. This is most probably due to the poor fitting of the available proxy. The southern Pomeranian Bay with the plume from the River Oder is the only area where the positive effect of the increased food availability is visible.

5.5. Conclusions and outlook

It is concluded that the presented map provides a sound basis for further analysis e.g. on functions of benthic compartment of the Baltic Sea ecosystem. An approximation for filter feeding capacity of the bivalves in the study area might be calculated using biomass-related equations (e.g. Riisgård & Seerup 2004). A crucial next step is to model the biomass distribution of important bioturbating or bio-irrigating species, respectively, as a base for spatial analysis of a second important ecological function of macrobenthic soft-bottom communities.

6. Fourth case study: Calculating filtering capacity – the *Mya arenaria* example

6.1. Introduction

The last case study directly picks up the proposal given in the preceding chapter. The prepared distribution maps are a first step towards the development of quantified maps for single functions, for example of the potential filtering capacity of the benthic community.

It was shown in laboratory experiments that the ratio of filtering capacity and individual biomass changes over the life-time, therefore it is more appropriate to use individual biomass instead of overall biomass for calculations (Forster & Zettler 2004, Riisgård & Steerup 2004). As individual biomass is rarely measured in standard procedures, calculating a mean individual weight by dividing the overall biomass has been proposed as a proxy by Forster & Zettler (2004). Additionally, the correlation between individual body mass and filtering capacity varies between the species, thus individual equations have to be used to link species' biomass with the filtering function (Møhlenberg & Riisgård 1979).

Unfortunately, this kind of equations is not available for all of the most dominant species in the study area. Therefore, it was impossible to calculate the filter feeding capacity of the full benthic community. Relevant equations have been published for *Mya arenaria* by Riisgård & Steerup (2004) and by Riisgård et al. (2013) for *Mytilus* spp. First estimations for *Arctica islandica* were provided by Møhlenberg & Riisgård (1979), but have so far not been revised. As this thesis focuses on soft-bottom communities, exemplary calculations were conducted for *Mya arenaria.*

The softshell clam *M. arenaria* is the largest bivalve and one of the most important bivalve species in the Pomeranian Bay and in shallow coastal waters throughout the coast of Mecklenburg-Vorpommern (Kube et al. 1997, Zettler et al. 2000). The species was identified as being the bivalve species benefiting most from additional food supply by eutrophication (Kube et al. 1997). Consequently, largest biomass was predicted to occur in the southern Pomeranian Bay near the mouth of the River Oder were food availability was highest (chapter 5.3.2).

6.2. Material and Methods

The analysis was based on the result of the 3rd case study. Therefore the same study area, database and technical approaches as described in chapter 5.2 were used. As the filtering capacity of *M. arenaria* has to be linked with the individual biomass two approaches were followed:

(1) calculate the mean biomass for each station by dividing the measured biomass by the measured density and model/ map the distribution of the mean individual biomass M_i *(station$_x$)* (in g afdm) or

(2) develop separate models and maps for biomass and density and calculate the mean individual biomass M_i for each grid cell using the predicted values:

$$M_i\ (cell_x) = \frac{predicted\ biomass\ \ (cell_x)}{predicted\ abundance\ (cell_x)} \qquad (8)$$

Following the first approach, the achieved model showed an unsatisfying performance due to high variability of the calculated mean individual biomass at the sampled stations. Thus, only the second approach was pursued. As the predicted biomass of *M. arenaria* was taken from the results of the 3rd case study, only RF-models for the abundance of *M. arenaria,* using the same methods as described in chapter 5.2.3 were designed and applied.

In a second step, the calculated mean individual biomass was linked with the filtering capacity using the equation provided by Riisgård & Seerup (2004). Individual filtration rate F_i (in l * ind^{-1} * h^{-1}) was calculated by

$$F_i = 4.76\ x\ M_i^{0.71} \qquad (9)$$

.

To obtain the population filtration rate F_{pop} (in m^3 * m^{-2} * d^{-1}), individual filtration rate was multiplied with the predicted density *N*. Two factors were used to align for required changes in temporal (from hour to day) and spatial reference (from liter to m³):

$$F_{pop}\ =\ F_i\ x\ N\ x\ 24\ x\ 10^{-3} \qquad (10)$$

6.3. Results

The abundance model for *M. arenaria* performed well, both within the training set (MSE: 0.32 ± 0.003, variation explained: 74.3 ± 0.24%) and in application on the validation set (correlation coefficient: 0.51 ± 0.01, RMSE: 585 ± 3.48).

The predicted abundance of *Mya arenaria* resembled the predicted biomass distribution of the same species (Figure 40). Analogous to the prediction for biomass, highest densities for *M. arenaria* (500 - 1 000 ind.*m^{-2}) were predicted for the southern Pomeranian Bay and along the shore of Mecklenburg-Vorpommern. In contrast, the species was predicted to be absent in the Arkona Basin and in the deeper parts of the Bay of Kiel and the Mecklenburg Bay. Water depth and mean salinity were the most important predictors (Table 13). The relative importance of sediment characteristics and detritus was lower than for the biomass model.

Table 13: Importance of the predictors expressed as increase of mean squared error (% Inc MSE) for in the RF-models predicting the abundance of *Mya arenaria*

predictor	%IncMSE
Median grain size	23.38 ± 1.71
Sorting	14.19 ± 1.83
Mean summer temperature	22.58 ± 0.87
Oxygen conditions	14.94 ± 0.88
Water depth	38.21 ± 3.93
Mean salinity	37.98 ± 3.66
Velocity	14.13 ± 1.38
Detritus	18.25 ± 1.68

Based on the predicted biomass and density, mean individual biomass was calculated by applying equation (8). Filtering capacity was calculated by applying equations (9) and (10). In cases where the predicted abundance was 0, M_i was set to 0 “by hand” to avoid undefined cell entries. Hotspots of filtering capacity were detected close to the mouth of the River Oder, in the shallow area north of Zingst Peninsula and close to the City of Rostock with maxima of 3 - 4.25 m^3 * m^{-2} * d^{-1} (Figure 41).

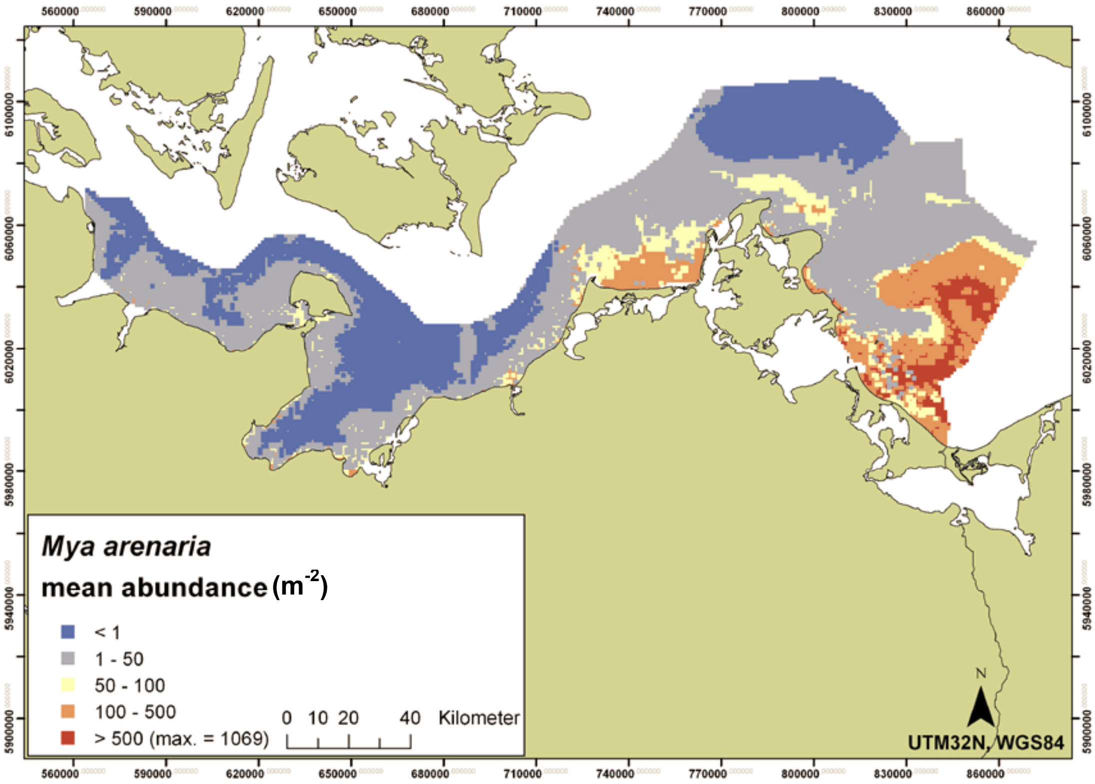


Figure 40: Mean predicted abundance of *Mya arenaria* in the south-western Baltic Sea

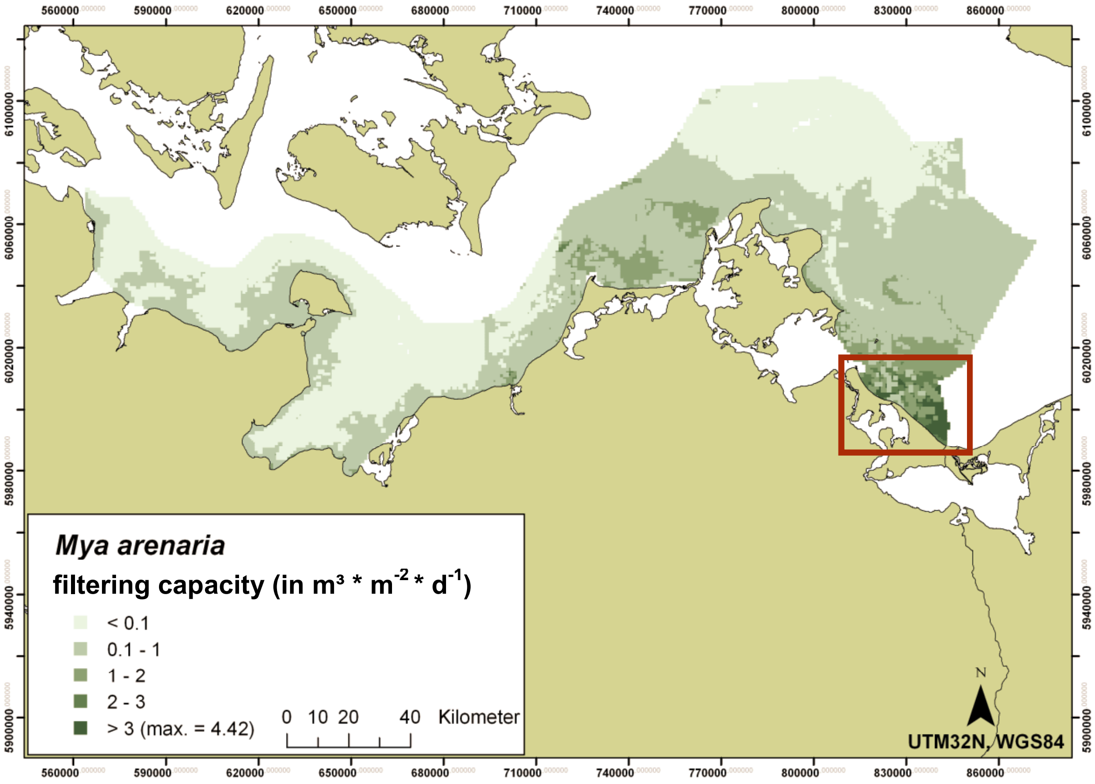


Figure 41: Modelled filtering capacity ($m^3*m^{-2}*d^{-1}$) of *Mya arenaria* in the south-western Baltic Sea. The red rectangle indicates the area selected for further calculations.

In order to put this figure into context, the filtering capacity was calculated exemplarily for the area where the filtering capacity of *M. arenaria* is highest, the southern Pomeranian Bay between the mouth of River Oder and Peenemünder Haken (northern tip of the Isle of Usedom, red rectangle in Figure 41). This part of the study area covers approximately 500 km². Within this area, the mean predicted filtering capacity is F_{pop} = 2.11 $m^3 * m^{-2} * d^{-1}$. Thus, the potential filtering capacity of the whole area is

$$F_{area} = 2.11 * 500 * 10^6 = \underline{1.06 * 10^9 \text{ m}^3 * \text{d}^{-1}}$$

This is about one billion cubic-meters per day. In a second step, this figure was related to (1) the outflow of the River Oder and (2) to the water volume of this area.

(1) Assuming that the mean outflow of the River Oder (MQ_{Oder}) is about 574 $m^3 * s^{-1}$ (HELCOM 2010), the water supply by the River Oder per day can be calculated as

$$MQ_{\ Oder} = 574 * 60 * 60 * 24 = \underline{4.96 * 10^7 \text{ m}^3 * \text{d}^{-1}}$$

This reveals that the filtering capacity of *Mya arenaria* in this area is about 20 times higher than the water supply of the River Oder.

(2) As the mean water depth of the selected area is about 11.5 m, an approximation of the volume can be calculated as

$$V_{area} = 11.5 * 500 * 10^6 = \underline{5.75 * 10^9 \text{ m}^3}$$

The theoretical clearance time of the water T_c by *M. arenaria* can be calculated by

$$T_c = \frac{V_{area}}{F_{area}} = \frac{5.75 * 10^9 \ m^3}{1.06 * 10^9 m^3 * d^{-1}} \approx 5.4 \ d \ .$$

6.4. Discussion and conclusions

The calculation of the potential filtering capacity of *Mya arenaria* in the German sector of the Baltic Sea was a first attempt to spatially quantify basic benthic functions for the whole area. Although the combination of two different models (biomass and abundance) might have increased the uncertainty of the prediction, scale and distributional pattern of the predicted filtering capacity sufficiently fit with the results of Forster & Zettler (2004) in the Mecklenburg Bay. Forster & Zettler (2004) used the same equation provided by Riisgård & Seerup (2004), but achieved spatial predictions by interpolation methods. The substantial progress of the method used in this study is the independency of the prediction from the spatial distribution of the stations. While in spatial interpolation methods the prediction for the space between the stations directly depends on both the measured response value at the surrounding stations and the distance between the stations, the prediction in the actual approach was calculated using the predictors in the RF-algorithm. Therefore, the value of the response variable solely depended on the values of the relevant environmental parameters for the individual grid cell. Consequently, a much higher spatial accuracy of the prediction was reached.

Taking into account that the Polish part of the southern Pomeranian Bay, which is about twice as large as the German part, has been neglected in the comparison of filtering capacity and the Oder outflow, the presented figure becomes even more remarkable. As the selected area was artificially determined due to the availability of the data and did not represent a geographical unit, calculations of the residence time of the water body might be improper. Therefore no conclusions can be drawn whether the selected area is regulated by benthic suspension feeders *sensu* Grall & Chauvaud (2002, Figure 3). Additionally, as phytoplankton composition and density heavily varies throughout the year (e.g. Wasmund et al. 2011), the interrelation between pelagic primary production and the effect of filter feeding bivalves might be more complex than expressed in the simple figure provided by Grall & Chauvaud (2002).

Although the study considered only a single species and an artificially delimited area, the results illustrate the possibilities of the chosen approach to link species distribution models with single benthic functions. A first step for a full estimation of the filtering capacity of the soft-bottom biotopes within a selected area requires reliable estimations for the filtering capacity of all relevant species which are not yet available in literature and still have to be measured *in vitro* or *in situ*. Møhlenberg & Riisgård (1979) provided equations for about a dozen bivalve species, indicating that the relation between body weight and filtering capacity might be quite similar between different species. In contrast, their results for *Cerastoderma edule* (used synonym: *Cardium edule*) and *Acanthocardia echinata* (used synonym: *Cardium echinata*) evidenced that the ratio might also be rather different between closely related species. Consequently, these equations should only be used for the species where they were derived from. Additionally, recent investigation e.g. by Riisgård et al. (2013) highlighted the intraspecific dependency of the filtering capacity on different abiotic parameter such as temperature, salinity and - the most important - on the density of particles. The application of this improvement might be implemented in the actual approach by including logical algorithms as the required information is available for most of the parameters for each grid cell.

7. Synthesis: Conclusions and future challenge

The target of this thesis was to investigate the impact of the salinity gradient on the functioning of macrobenthic soft-bottom community in the south-western Baltic Sea. The problem was tackled using two approaches: the appraisal of functional diversity and composition under consideration of the whole benthic community and the quantification of species distribution as a tool for mapping single key functions. The achieved results and conclusions were foreseen to provide a basis for the later development of function-based indicators in state assessment.

7.1. Changes of benthic function along the salinity gradient

The results of the 2nd case study clearly showed that functional composition and diversity were affected by the salinity gradient, but apparently to a much lesser extent than species diversity. Unfortunately, the results of species-based and function-based analyses are hardly comparable as the approaches differ statistically. In species-based community-analysis of this area, usually 200 - 400 species are considered. This number of species is squeezed into less than 50 biological traits in the present approach, leading to a severe decrease in available dimension in multivariate space - in other words: it is more difficult to be different or to be more/less diverse in the traits-based approach. Nevertheless, the loss of species-diversity from the Bay of Kiel towards the Pomeranian Bay and the dramatic changes in community composition are followed by almost no functional loss and only a weak change in functional composition. The main alteration in functional composition along the salinity gradient does to a much lesser extent effect traits describing the ecosystem relevant activities such as feeding type or sediment transport, but is more pronounced for life-history traits (e.g. life-span) and traits describing life-strategies (e.g. larval development). This might indicate that the stability of the community is affected, but not the quantity of its key roles in ecosystem functioning. On the other hand, the reduced maximal life-span is, at least within the taxonomic group of bivalves, in general followed by a reduced maximum size and individual biomass (e.g. *Arctica islandica* vs. *Macoma balthica* as the dominant species in muddy substrates of the Bay of Kiel and the Arkona Basin, respectively). A distinct trend of decreasing bivalve biomass at about 17 – 20 psu is visible in the predicted overall biomass of the 3rd case study (Figure 42). But also other parameters, like the visible positive effect of additional food in the 6 - 7-psu class and the negative effect of hypoxia events at higher salinities, apparently influence bivalve biomass. However, in order to answer the question whether a decline in filtering capacity or other basic functions along the salinity gradient does exist, further investigations *in vitro* (e.g. to set up species-specific filtering capacity biomass ratios) and *in situ* to validate the laboratory experiments are still needed.

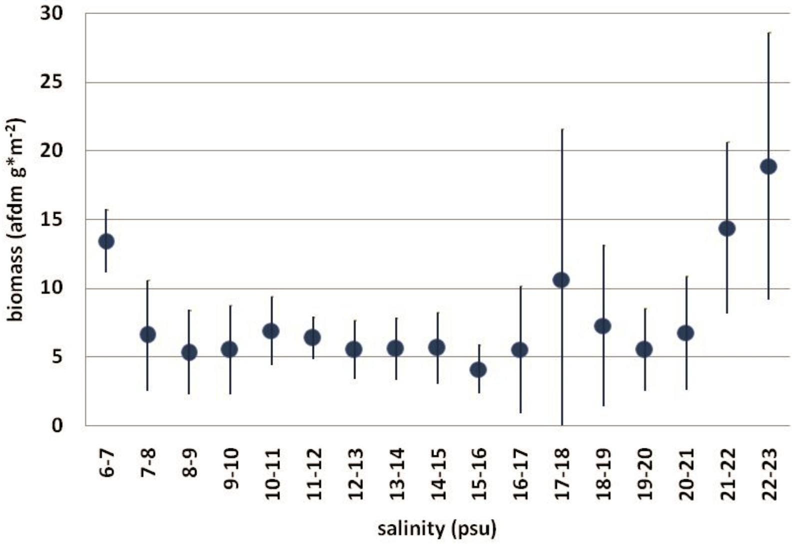


Figure 42: Mean (dots) and range (error bars) of predicted overall biomass of the 3rd case study aggregated into salinity classes

Despite the detected changes in functional composition, the shift from highly specialised to ubiquitous species in the community remained invisible in the chosen BTA-approach. As clearly shown in the 1st case study, the apparent habitat preferences of many species change along the salinity gradient. As discussed in chapter 3.4.2, these changes might be linked with behavioural or genetic differences. Besides the species selected in the 1st case study, similar opportunities are well known for many common species of the study area such as *Macoma balthica*, *Peringia ulvae* and *Scoloplos armiger*. In general, intra-specific genetic and behavioural variability have to be more considered in functional approaches (Bolnick et al. 2012, Dowdall 2012). Future revised versions of the BTA-approach have to incorporate these behavioural changes as the current version only reflects the “behavioural potential” of the species and not the actual function provided by the species under the given environmental conditions. Additionally, the quantification of the functions has to be stronger reflected than in the current approach. As discussed in the 4th case study, the individual filtering capacity per individual biomass differs between different bivalve species. The differences might be much more pronounced between different taxonomic groups and especially between active and passive suspension feeders and are only partly reflected by the usage of overall biomass of the species. The introduction of species-specific “functions intensity” factors into the BTA-table might contradict the rule of un-biased importance between the traits, but has a great potential to improve the explanatory power.

7.2. Mapping of benthic function

The limitation not to reflect behavioural changes also applies for the mapping of single benthic functions, although the solution might be less complicated once the dependency of the filtering capacity on the relevant environmental parameters have been described (compare chapter 6.4). The basic requirement to improve the prediction of benthic functions is to enhance the accuracy of the predictions for the underlying response variables biomass and abundance. In addition to the improvement of the available proxies describing the environmental conditions, the incorporation of interspecific interactions is a crucial demand for future development (e.g. Robinson et al. 2011, Reiss et al. *in prep*). Besides “negative” impacts of interspecific interactions by competition or predation, also “positive” effects have to be reflected. As stated in the introduction, the presence or activity of particular species might have broad effects on the composition of the whole macrobenthic community. For example, small annelid species (e.g. *Pygospio elegans*, Tubificidae) appear to be attracted by high densities of blue-mussels (own observations). The small deposit feeding annelids might benefit both from the additional available space/shelter and the additional food source (pseudo faeces).

An additional improvement would be to develop separate models for different age cohorts of a particular species. Up to now, the mean individual biomass had to be calculated by dividing the overall biomass through overall density as comprehensive length measurements were not available (see chapter 6.2). But as the length-distribution of long-lived species is naturally multimodal and not normally distributed the usage of the mean individual biomass might incorporate a bias into the model. The subdivision of the population into simple size classes (small, medium, large) and the development of separate models for each individual size class might be a useful step towards higher accuracy in the prediction of benthic functions.

7.3. Towards the development of function-based indicators and further applications

As mentioned above, the critical point in both approaches (BTA and mapping single functions) is the necessity to incorporate potential behavioural adaptations of the species into the assessment system as it might affect the capability to detect occurring functional changes along other natural or anthropogenic induced gradients. This disguise of functional changes by “multifunctional” species has to be considered in the development of indicators basing on functional diversity or composition of benthic communities in brackish waters. Nevertheless, BTA-approaches have successfully been used to describe the impact of different anthropogenic activities or severe natural events (e.g. cold winters) onto the functional composition of benthic communities (e.g. Tillin et al. 2006, Bremner et al. 2008, Frid et al. 2008, Dimitriadis & Koutsoubas 2011).

Severe hypoxia was shown to result in significant changes in functional composition and diversity at two monitoring stations within the study area of this thesis (Gogina et al. 2014). While the response of the functional diversity index was not fully specific and partly contained a time lag (Figure 43), the changes in functional composition after hypoxia were much more distinct (Gogina et al. 2014). This result might be a first indication for the suitability of functional composition as a basis for state assessment indicators within the considered area. A next step might be to relate an updated BTA-approach (incorporating the modifications mentioned above) to other specific impacts, such as physical disturbance by bottom trawling, sediment extraction and dumping. Schröder et al. (2008) showed that the permanent impact of bottom trawling onto the communities in the German Bight had differing effects for different functional groups and therefore led to a distinct shift in the functional composition.

However, the incorporation of the functional approach into recent indicator systems requires a “translation” of the discovered pattern into resilient figures and finally into state categories. Analogous to taxonomy-based approaches, this might be rather simple for univariate diversity indices, but considerably more complicated for the multivariate BTA-approach. A potential procedure is to use the BTA-approach to detect the relevant affected traits and to build up the indicator by categorizing the changes within the individual traits. According to the target of the study/indicator, this procedure might be coupled with the approach of mapping single functions.

Actually, the potential applications of maps providing quantitative information on the distribution of single functions are manifold. Just to mention a few examples, they might be incorporate into “ecosystem service” assessment or might be used for designation of marine protected areas as hotspots of benthic functions should be considered as valuable as hotspots of species diversity. The filtering capacity of bivalve populations can also be used in restoration plans for highly eutrophic lagoons by calculating the potential impact of the bivalves on the pelagic primary production and the potentially required additional filtering capacity e.g. added by mussel farming as proposed by Schernewski et al. (2012).

Finally, it can be concluded that despite the mentioned limitations of the actual procedures, both chosen approaches provide a sound basis for a wide variety of application in state assessment, nature conservation and ecosystem analysis.

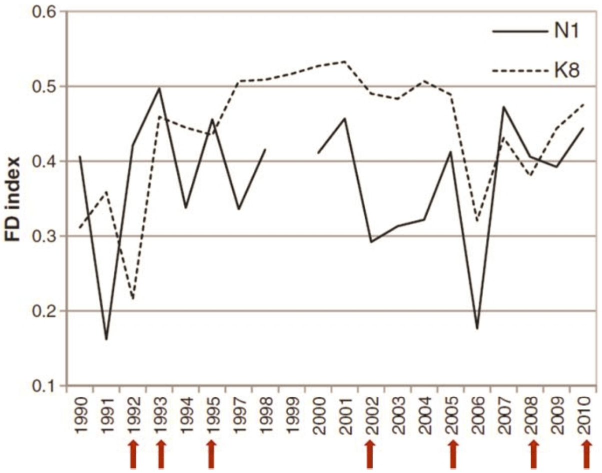


Figure 43: Interannual variability of functional diversity at monitoring stations N1 (Fehmarnbelt) and K8 (near Darss Sill). Red arrows indicate hypoxia measured during sampling event at station N1. No hypoxia was detected at station K8 (adapted from Gogina et al. 2014).

References

Aarnio K, Mattila J, Törnroos A, Bonsdorff E (2011) Zoobenthos as an environmental quality element: the ecological significance of sampling design and functional traits. Mar Eco 32 (Suppl. 1):58-71

Anderson MJ (2008) Animal-sediment relationships re-visited: Characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. J Exp Mar Biol Ecol 366:16-27

Atrill MJ (2002) A testable linear model for diversity trends in estuaries. J Animal Ecol 71:262-269

Barbone E, Rosati I, Reizopoulou S, Basset A (2012) Linking classification boundaries to sources of natural variability in transitional waters: A case study of benthic macroinvertebrates. Ecol Indic 12:105-122

Barrio Froján CRS, Cooper KM, Bremner J, Defew EC, Hussin WMRW, Paterson DM (2011) Assessing the recovery of functional diversity after sustained sediment screening at an aggregate dredging site in the North Sea. Est Coast Shelf Sci 92(3):358-366

Bayne BL, Worrall CM (1980) Growth and production of mussels *Mytilus edulis* from two populations. Mar Ecol Prog Ser 3:317-328

Bleich S, Powilleit M, Seifert T, Graf G (2011) β-diversity as a measure of species turnover along the salinity gradient in the Baltic Sea, and its consistency with the Venice System. Mar Ecol Prog Ser 436:101-118

Bolnick DI, Amarasekare P, Araújo MS, Bürger R, Levine JM, Novak M, Rudolf VHW, Schreiber SJ, Urban MC, Vasseur DA (2012) Why intraspecific trait variation matters in community ecology. Trends Ecol Evol 26:183-192

Bonaglia S, Bartoli M, Gunnarsson JS, Rahm L and others (2013) Effect of reoxygenation and *Marenzelleria* spp. bioturbation on Baltic Sea sediment metabolism. Mar Ecol Prog Ser 482:43-55

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

Boström C, Törnroos A, Bonsdorff, E (2010) Invertebrate dispersal and habitat heterogeneity: Expression of biological traits in a seagrass landscape. J Exp Mar Biol Ecol 390:106-117

Botta-Dukat Z (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits J Veg Sci 16:533-540

Breiman L (2001) Random Forests. Machine Learning 45:5-32

Breiman L, Freidman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth.

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

Bremner J, Rogers SI, Frid CLJ (2003) Assessing functional diversity in marine benthic systems: a comparison of approaches. Mar Ecol Prog Ser 254:11–25

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

Brey T, Arntz WE, Pauly D, Rumohr H (1990) *Arctica (Cyprina) islandica* in Kiel Bay (Western Baltic): growth, production and ecological significance. J Exp Mar Bio Ecol 136:217-235

Brown CJ, Smith SJ, Lawton P, Anderson JT (2011) Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Est Coast Shelf Sci 92:502-520

Bučas M, Bergström U, Downie A-L, Sundblad G, Gullström M, von Numers M, Šiaulys A, Lindegarth M (2013) Empirical modelling of benthic species distribution, abundance, and diversity in the Baltic Sea: evaluating the scope for predictive mapping using different modelling approaches. – ICES J Mar Sci, doi:10.1093/icesjms/fst036.

Buckley LB, Urban MC, Angilletta MJ, Crozier LG, Rissler LJ, Sears MW (2010) Can mechanism inform species' distribution models? Ecol Lett 13:1041-1054

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. Sociol Methods Res 33(2):261-304

Cade BS (2003) Quantile regression models of animal habitat relationship. Dissertation at the Colorado State University, Fort Collins, Colorado 177 pp

Cade BS, Terrell JW, Schroeder, RL (1999). Estimating effects of limiting factors with regression quantiles. Ecology 80:311-323

Champeley S, Chessel D (2002) Measuring biological diversity using Euclidean metrics. Environ Ecol Stat 9:167-177

Chessel D, Dufour AB, Thioulouse J (2004) The ade4 package - One-table methods. R News 4:5-10

Chevenet F, Doledec S, Chessel D (1994) A fuzzy coding approach for the analysis of long-term ecological data. Freshwater Biol 31:295-309

Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random Forests for classification in ecology. Ecology 88(11):2783-2792

d'Udekem d'Acoz C (2004) The genus *Bathyporeia* Lindström, 1855, in western Europe (Crustacea: Amphipoda: Pontoporeiidae). Zool Verhand 348:1-164

Dauvin JC, Alizier S, Rolet C, Bakalem A, Bellan G, Gomez Gesteira JL, Grimes S, de-la-Ossa-Carretero JA, Del-Pilar-Ruso Y (2012) Response of different benthic indices to diverse human pressures. Ecol Indic 12:143-153

Dauvin JC, Bellan G, Bellan-Santini D (2010) Benthic indicators: From subjectivity to objectivity – Where is the line? Mar Poll Bull 60:947-953

Díaz S, Cabido M (2001) Vive la différence: plant functional diversity matters to ecosystem processes. Trends Ecol Evol 16:646-655

Dimitriadis C, Koutsoubas D (2011) Functional diversity and species turnover of benthic invertebrates along a local environmental gradient induced by an aquaculture unit: the contribution of species dispersal ability and rarity. Hydrobiologia 670:307-315

Dormann CF, McPherson JM, Araujo MB, Bivand R, Bolliger J, Carl G, Davies RG, Hirzel A, JetzW, Kissling WD, Kühn I, Ohlemüller R, Peres-Neto PR, Reineking B, Schröder B, Schurr FM, Wilson R (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609-628

Dowdall JRT (2012) Intra- and interspecific variation along environmental gradients: Adaption, plasticity, and range limits. Dissertation Colorado State University: 162pp

Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32:66-77

Elliott M, Quintino V (2007) The Estuarine Quality Paradox, Environmental Homeostasis and the difficulty of detecting anthropogenic stress in naturally stressed areas. Mar Poll Bull 54:640-645

Evans JS, Murphy MA, Holden ZA, Cushman SA (2011): Modelling species distribution and change using randon forests. In: Drew CA, Wiersma YF, Huettmann F (ed.): Predictive Species and habitat modelling in landscape ecology – concepts and applications. Springer New York Dodrecht Heidelberg London. 328 pp. ISBN 1441973907

Feistel R, Nausch G, Wasmund N (2008) State and Evolution of the Baltic Sea 1952-2005: A detailed 50-year survey of meteorology and climate, physics, chemistry, biology and marine environment. Wiley Hoboken, New Jersey, 703 pp

Fleischer D, Zettler ML (2009) An adjustment of benthic ecological quality assessment to effects of salinity. Mar Poll Bull 58:351-357

Forster S, Zettler ML (2004) The capacity of the filter-feeding bivalve *Mya arenaria* L. to affect water transport in sandy beds. Mar Biol 144:1183-1189

Franklin J & Miller JA (2010): in Franklin (ed.): Mapping Species Distributions - Spatial Inference and Prediction. Cambridge Academic Press. ISBN: 9780521700023

Frid CLJ, Paramor OAL, Brockington S, Bremner J (2008) Incorporating ecological functioning into the designation and management of marine protected areas. Hydrobiologia 606:69-79

Friedland R, Neumann T, Schernewski G (2012) Climate change and the Baltic Sea action plan: Model simulations on the future of the western Baltic Sea. J Mar Sys 105:175-186

Gamenick I, Vismann B, Grieshaber MK, Giere O (1998) Ecophysiological differentiation of *Capitella capitata* (Polychaeta). Sibling species from different sulfidic habitats. Mar Ecol Progr Ser 175:155-166

Glockzin M & Zettler ML (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

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

Gogina MA (2010) Investigation of interrelations between sediment and near-bottom environmental parameters and macrozoobenthic distribution patterns for the Baltic Sea. Dissertation at EMAU Greifswald: 58 pp + appendices

Gogina MA, Darr A, Zettler ML (2014): Approach to assess consequences of hypoxia disturbance events for benthic ecosystem functioning. J Mar Sys 129:203-213

Gogina MA, Glockzin M, Zettler ML (2010a) Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. Part 1. Causal analysis. J Mar Sys 79:112-123

Gogina MA, Glockzin M, Zettler ML (2010b) Distribution of benthic macrofaunal communities in the western Baltic Sea with regard to near-bottom environmental parameters. 2. Modelling and prediction. J Mar Sys 80:57-70

Gogina MA, Zettler ML (2010) Diversity and distribution of benthic macrofauna in the Baltic Sea. Data inventory and its use for species distribution modelling and prediction. J Sea Res 64:313-321

Gosling, E (2003) Bivalve Molluscs: Biology, Ecology and Culture. Blackwell Publishing Oxford, Malden, Carlton. ISBN 0-85238-234-0, 443 pp

Graf G (1992) Bentho-pelagic coupling – a benthic view. Oceanogr Mar Biol Ann Rev 30:149-190

Grall J, Chauvaud L (2002) Marine Eutrophication and benthos: the need for new approaches and concepts. Global Change Biology 8:813-830

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

Grémare A, Labrune C, Vanden Berghe E, Amouroux JM, Bachelet G, Zettler ML, Vanaverbeke J, Fleischer D, Bigot L, Maire O, Deflandre B, Craeymeersch J, Degraer S, Dounas C, Duineveld G, Heip C, Herrmann M, Hummel H, Karakassis I, Kędra M, Kendall M, Kingston P, Laudien J, Occhipinti-Ambrogi A, Rachor E, Sardá R, Speybroeck J, Van Hoey G, Vincx M, Whomersley P, Willems W, Wlodarska-Kowalczuk M, Zenetos A (2009) Comparison of the performances of two biotic indices based on the MacroBen database. Mar Ecol Prog Ser 382:297-311

Groth H, Theede H (1989) Does brackish water exert long-term stress on marine immigrants in the Baltic Sea? Scient Mar 53 (2-3):677-684

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147-186

Hartmann-Schröder G (1996) Annelida, Borstenwürmer, Polychaeta. (2. Aufl.) In: Dahl F, editor, Die Tierwelt Deutschlands und der angrenzenden Meeresteile nach ihren Merkmalen und ihrer Lebensweise. Teil 58, Gustav Fischer Verlag, Jena: 648 pp

Hastie TJ (1992) Generalized additive models. Chapter 7 of Statistical Models in S. In: Chambers JM, Hastie TJ, editors: Wadsworth & Brooks/Cole

Heink U, Kowarik I (2010) What are indicators? On the definition of indicators in ecology and environmental planning. Ecol Indic 10:584-593

HELCOM (2010) HELCOM Atlas of the Baltic Sea. Helsinki Commission 189 pp

HELCOM (2013a) Manual for Marine Monitoring in the COMBINE Programme of HELCOM. Version 26.09.2013. download from www.helcom.fi/Documents/Action%20areas/Monitoring%20and%20assessment/Manuals%20and%20Guidelines/Manual%20for

%20Marine%20Monitoring%20in%20the%20COMBINE%20Programme%20of%20HELCOM.pdf

HELCOM (2013b) Red List of Baltic Sea underwater biotopes, habitats and biotope complexes Baltic Sea Environmental Proceedings 138:69 pp

Hewitt JE, Thrush SF, Dayton PD (2008) Habitat variation, species diversity and ecological functioning in a marine system. J Exp Mar Biol Ecol 366:116-122

Hrs-Brenko M (2006) The basket shell, *Corbula gibba* Olivi, 1792 (Bivalve Mollusks) as a species resistant to environmental disturbances: A review. Acta Adriat 47:49-64

Hurlbert S (1971) The Nonconcept of Species Diversity: A Critique and Alternate Paradigms. Ecology 52(4):577-586

Hyland J, Balthis LW, Karakassis I, Magni P, Petrov A, Shine JR, Vestergaard O, Warwick R (2005) Organic carbon content of sediments as an indicator of stress in the marine benthos. Mar Ecol Prog Ser 295:91-103

Ieno EN, Solan M, Batty P, Pierce GJ (2006) How biodiversity affects ecosystem functioning: roles of infaunal species richness identity and density in marine benthos. Mar Ecol Prog Ser 311:263-271

Keely, Macleod CK, Forrest BM (2012) Combining best professional judgement and quantile regression splines to improve characterisation of macrofaunal responses to enrichment. Ecol Indic 12:154-166

Kesaniemi JE, Rawson PD, Lindsay SM, Knot KE (2012) Phylogenetic analysis of cryptic speciation in the polychaete *Pygospio elegans*. Ecol Evol 2:994-1007

Klingbeil K, Mohammadi-aragh M, Gräwe U, Burchard H (2013) Quantification of spurious dissipation and mixing – discrete variance decay in a finite-volume framework. Ocean Mod (submitted for publication)

Koenker R (2012) quantreg: Quantile Regression. R package version 4.79. http://CRAN.R-project.org/package=quantreg

Koenker R, Bassett, G (1978) Regression quantiles. Econometrica 46:33-50

Koenker R, Ng P, Portnoy S (1994). Quantile smoothing splines. Biometrika 81:673-680

Krause JC (2002) The effects of marine sediment extraction on sensitive macrozoobenthic populations in the southern Baltic Sea. PhD Thesis University of Rostock

Kruse, I, Reusch TBH, Schneider MV (2003) Sibling species or poecilogony in the polychaete *Scoloplos armiger*? Mar Biol 142:937-947

Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. J American Statistical Association 47 (260):583-621

Kube J, Gosselck F, Powilleit M, Warzocha J (1997) Long-term changes in the benthic communities of the Pomeranian Bay (Southern Baltic Sea). Helgol Meeresunters 51:399-416

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

Lewis TL, Esler D, Boyd WS (2007) Effects of predation by sea ducks on clam abundance in soft-bottom intertidal habitats. Mar Ecol Prog Ser 329:131-144

Liaw A, Wiener M (2002) Classification and Regression by randomForest. R News 2(3):18-22

Luttikhuizen PC, Bol A, Cardoso JFMF, Dekker R (2011) Overlapping distributions of cryptic *Scoloplos* cf. *armiger* species in the western Wadden Sea. J Sea Res 66:231-237

Magni P, Tagliapietra D, Lardicci C, Balthis L, Castelli A, Como S, Frangipane G, Giordani G, Hyland J, Maltagliati F, Pessa G, Rismondo A, Tataranni M, Tomassetti P, Viaroli P (2009) Animal-sediment relationships-Evaluating the Pearson–Rosenberg paradigm-organic-Mediterranean). Mar Poll Bull 58:478-486

Mann HB, Whitney DR (1947) On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann of Math Statistics 18 (1):50-60

MarLIN (The Marine Life Information Network): http://www.marlin.ac.uk/

Mason NWH, Mouillot D, Lee WG and Wilson JB (2005) Functional richness, functional evenness and functional divergence: the primary components of functional diversity. Oikos 111:112-118

Meyer T, Berg T, Fürhaupter K (2008) Ostsee Makrozoobenthos Klassifizierungssystem für die Wasserrahmenrichtlinie. Referenz-Artenlisten, Bewertungsmodell und Monitoring. MARILIM Gewässeruntersuchung: 131 pp. http://www.marilim.de/marbit/download/files/MarBIT.pdf

Miloslavich P, Klein E (2009) The World Conference on Marine Biodiversity: current global trends in marine biodiversity research. Mar Biodiv 39:147-152

Møhlenberg F, Riisgård HU (1979) Filtration Rate, Using a New Indirect Technique, in Thirteen Species of Suspension-Feeding Bivalves. Mar Biol 54:143-147

Mosteller F, Tukey JW (1977). Data analysis and regression. Addison-Wesley, New York

Mouillot D, Spatharis S, Reizopoulou S, Laugier T, Sabetta L, Basset A, Do Chi T (2006) Alternatives to taxonomic-based approaches to assess changes in transitional water communities. Aquat Conserv 16:469-482

Musters CJM, Kalkman V, van Strien A (2013) Predicting rarity and decline in animals, plants, and mushrooms based on species attributes and indicator groups. Ecol Evol 3(10):3401-3414

Naeem S (1998) Species redundancy and ecosystem reliability. Conserv Biol 12(1):39-45

Neumann T (2000) Towards a 3D-ecosystem model of the Baltic Sea. J Mar Syst 25:405-419

Nikula R, Strelkov P, Väinölä R (2008) A broad transition zone between an inner Baltic hybrid swarm and a pure North Sea subspecies of *Macoma balthica* (Mollusca, Bivalvia). Mol Ecol 17:1505-1522

Nikula R, Väinölä R (2003) Phylogeography of *Cerastoderma glaucum* (Bivalvia: Cardiidae) across Europe: a major break in the Eastern Mediterranean. Mar Biol 143:339-350

Norling K, Rosenberg R, Hulth S, Gremare A, Bonsdorff E (2007) Importance of functional biodiversity and species-specific traits of benthic fauna for ecosystem functions in marine sediments. Mar Ecol Prog Ser 332:11-23

Norkko A, Hewitt JE, Thrush SF, Funnell GA (2001) Benthic-pelagic coupling and suspension-feeding bivalves: Linking site-specific sediment flux and biodeposition to benthic community structure. Limnol Oceanogr 46(8):2067-2072

Odum, EP (1985) Trends expected in stressed ecosystems. Bioscience 35:419-422

Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2011) vegan: Community Ecology Package. R package version 1.17-11

Olaya-Marín EJ, Martínez-Capel F, Vezza P (2013) A comparison of artificial neural networks and random forests to predict native fish species richness in Mediterranean rivers. Knowl Manag Aquat Ecosys 409(07)

Osowiecki A, Lysiak-Pastuszak E, Piatkowska Z (2008) Testing biotic indices for marine zoobenthos quality assessment in the Polish sector of the Baltic Sea. J Mar Sys 74:124-132

Papageorgiou N, Sigala K, Karakassis I (2009) Changes of macrofaunal functional composition at sedimentary habitats in the vicinity of fish farms. Est Coast Shelf Sci 83:561-568

Parapar J, Helgason GV, Jirkov I, Moreira J (2012) Polychaetes of the genus *Ampharete* (Polychaeta: Ampharetidae) collected in Icelandic waters during the BIOICE project. Helgol Mar Res 66:331-344

Pearson TH, Rosenberg, R (1978) Macrobenthic succession in relation to organic enrichment and pollution on the marine environment. Oceanogr Mar Biol Ann Rev 16:229-311

Petchey OL and Gaston KJ (2006) Functional diversity: back to basics and looking forward Ecology Letters 9:741-758

Pilgrim EM, Darling JA (2010) Genetic diversity in two introduced biofouling amphipods (*Ampithoe valida* & *Jassa marmorata*) along the Pacific North American coast: investigation into molecular identification and cryptic diversity. Diversity Distrib 16:827-839

Pollack DSG (1993) Smoothing with cubic splines. Economics working paper series, London University, Queen Mary and Westfield College, Department of Economics - no.291

Powilleit M, Kube J, Maslowski J, Warzocha J (1996) Distribution of macrobenthic invertebrates in the Pomeranian Bay (Southern Baltic sea) in 1993/94. Bull Sea Fish Inst Gdynia 3:75-87

Pulliam HR (2000) On the relationship between niche and distribution. Ecol Lett 3:349-361

Queirós AM, Birchenough SNR, Bremner J, Godbold JA, Parker RE, Romero-Ramirez A, Reiss H, Solan M, Somerfield PJ, Van Colen C, Van Hoey G, Widdicombe S (2013) A bioturbation classification of European marine infaunal invertebrates. Ecol Evol doi: 10.1002/ece3.769

R Development Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Rakocinski-CF (2012)-Evaluating macrobenthic process indicators in relation to organic enrichment and hypoxia. Ecol Ind 13:1-12

Rao CR (1982) Diversity and dissimilarity coefficients: a unified approach. Theor Pop Biol 21:24-43

Reiss H, Birchenough S, Borja A, Buhl Mortensen L, Craeymeersch J, Dannheim J, Darr A, Degraer S, Galparsoro I, Goggina M, Neumann H, Populus J, Rengstorf A, Valle Tobar M, van Hoey G, Zettler ML (*in prep.*) Benthos Distribution Modelling and its relevance for marine ecosystem management

Reiss H, Cunze S, König K, Neumann H, Kröncke I (2011). Species distribution modelling of marine benthos: a North Sea case study. Mar Ecol Prog Ser 442:71-86

Remane A (1934) Die Brackwasserfauna. Zool Anz 7(Suppl):34-74

Reusser DA, Lee H (2011) Evolution of natural history information in the 21st century – developing an integrated framework for biological and geographical data. J Biogeogr 38:1225-1239

Ridgway ID, Richardson CA (2010) *Arctica islandica*: the longest lived non colonial animal known to science. Rev Fish Biol Fisheries (2011) 21:297-310

Riisgård HU, Lüsskow F, Pleissner D, Lundgreen K, López MAP (2013) Effect of salinity on filtration rates of mussels *Mytilus edulis* with special emphasis on dwarfed mussels from the low-saline Central Baltic Sea. Helgol Mar Res 67:591-598

Riisgård HU, Seerup DF (2004) Filtration rates of the soft clam *Mya arenaria*: effects of temperature and body size. Sarsia 88:416-428

Robinson LM, Elith J, Hobday AJ, Pearson RG, Kendall BE, Possingham HP, Richardson AJ (2011) Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. Global Ecol Biogeo 20:789-802

Röhner M, Bastrop R, Jürss K (1997) Genetic differentiation in *Hediste diversicolor* (Polychaeta: Nereididae) for the North Sea and the Baltic Sea. Mar Biol 130:171-180

Rosenberg R (1995) Benthic marine fauna structured by hydrodynamic processes and food availability. Neth J Sea Res 34 (4):303-317

Rosenberg R, Blomqvist M, Nilsson HC, Cederwall H, Dimming A (2004) Marine quality assessment by use of benthic species-abundance distributions: a proposed new protocol within European Union Water Framework Directive. Mar Poll Bull 49:728-739

Schernewski G, Stybel N, Neumann T (2012) Zebra mussel farming in the Szczecin (Oder) Lagoon: water-quality objectives and cost-effectiveness. Ecol Soc 17(2):4

Schleuter D, Daufresne M, Massol F, Argillier C (2010) A User's guide to functional diversity indices. Ecol Mono 80(3):469-484

Schröder A, Gutow L, Gusky M (2008) Auswirkungen von Grundschleppnetzfischereien sowie von Sand- und Kiesabbauvorhaben auf die Meeresbodenstruktur und das Benthos in den Schutzgebieten der deutschen AWZ der Nordsee (MAR 36032/15). Abschlussbericht für das Bundesamt für Naturschutz: 124 pp

Shannon CE (1948) A mathematical theory of communication. The Bell System Technical Journal 27:379-423

Siaulys A, Daunys D, Bučas M, Bacevicius E (2012) Mapping an ecosystem service: A quantitative approach to derive fish feeding ground maps. Oceanologia 54(3):491-505

Skov H, Dahl K, Dromph K, Daunys D, Engdahl A, Eriksson A, Floren K, Gullström M, Isaeus M, Oja J (2012) MOPODECO - Modelling of the Potential coverage of habitatforming species and Development of tools to evaluate the Conservation status of the marine Annex I habitats. Report on behalf of the Nordic Council of Ministers, 205pp. http://dx.doi.org/10.6027/TN2012-532

Soberón J (2007) Grinnellian and Eltonian niches and geographic distributions of species. Ecol Lett 10:1115-1123

Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics 8:25

Swatantran A, Dubayah R, Goetz S, Hofton M, Betts MG (2012) Mapping Migratory Bird Prevalence Using Remote Sensing Data Fusion. PLoS ONE 7(1): e28922. doi:10.1371/journal.pone.0028922

Tauber F (2012) Seabed sediments in the German Baltic Sea. Ed. by M. Zeiler, Bundesamt für Seeschifffahrt und Hydrographie, Hamburg, Rostock. http://gdi.bsh.de/mapClient/initParams.do

Tillin HM, Hiddink JG, Kaiser MJ, Jennings S (2006) Chronic bottom trawling alters the functional composition of benthic invertebrate communities on a sea basin scale. Mar Ecol Prog Ser 318:31-45

Thrush SF, Hewitt JE, Herman PMJ, Ysebaert T (2005) Multi-scale analysis of species-environment relationships. Mar Ecol Prog Ser 302:13-26

Thrush SF, Pridmore RD, Hewitt JE, (1994) Impacts on soft-sediment macrofauna: the effects of spatial variation on temporal trends. Ecol Appl 4(1):31-41

Törnroos A, Bonsdorff E (2012) Developing the multitrait concept for functional diveristy: lessons from a system rich in functions but poor in species. Ecol Appl 22(8):2221-2236

Trutschler K, Samtleben C (1988) Shell growth of *Astarte elliptica* (Bivalvia) from Kiel Bay (Western Baltic Sea). Mar Ecol Prog Ser 42:155-162

van der Linden P, Patricio J, Marchini A, Cid N, Neto JM, Marques JC (2012) A biological trait approach to assess the functional composition of subtidal benthic communities in an estuarine ecosystem Ecol Ind 20:121-133

Väniölä R, Hvilsom MM (1991) Genetic divergence and a hybrid zone between Baltic and North Sea *Mytilus* populations (Mytilidae; Mollusca). Biol J Linn Soc 43:127-140

Väniölä R, Strelkov P (2011) *Mytilus trossulus* in Northern Europe. Mar Biol 158:817-833

Vincenzi S, Zucchetta M, Franzoi P, Pellizzato M, Pranovi F, De Leo GA, Torricelli P (2011) Application of a Random Forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy. Ecol Mod 222:1471-1478

Villnäs A, Perus J, Bonsdorff E (2011) Structural and functional shifts in zoobenthos induced by organic enrichment – Implications for community recovery potential. J Sea Res 65:8-18

Virgilio M, Fauvelot C, Costantini F, Abbiati M, Backeljau T (2009) Phylogeography of the common ragworm *Hediste diversicolor* (Polychaeta: Nereididae) reveals cryptic diversity and multiple colonization events across its distribution. Mol Ecol 18:1980-1994

Wang Y (2011) Smoothing Splines – Methods and Apllications. Taylor and Francis 384 pp

Wasmund N, Tuimala J, Suikkanen S, Vandepitte L, Kraberg A (2011) Long-term trends in phytoplankton composition in the western and central Baltic Sea. J Mar Sys 87:145-159

Wei CL, Rowe GT, Escobar-Briones E, Boetius A, Soltwedel T (2010) Global patterns and predictions of seafloor biomass using random forests, PLoS ONE, 5 (12), e15323

Word JQ (1979) The infaunal trophic index. In: Annual report, 1978: Coastal Water Research Project, El Segundo, Calif 19-39

Zettler ML (2002) Ecological and morphological features of the bivalve *Astarte borealis* (Schumacher, 1817) in the Baltic Sea near its geographical range. J Shellfish Res 21:33-40

Zettler ML, Bönsch R, Gosselck F (2000) Das Makrozoobenthos der Mecklenburger Bucht - rezent und im historischem Vergleich. Meereswissenschaftliche Berichte 42: 144pp

Zettler ML, Bönsch R, Gosselck F (2001) Distribution, abundance and some population characteristics of the ocean quahog, *Arctica islandica* (Linnaeus, 1767), in the Mecklenburg Bay (Baltic Sea). J Shellfish Res 20:161-169

Zettler ML, Darr A, Labrenz M, Sagert S, Selig U, Siebert U, Stybel N (2013a) The indicator value of organisms. In: Snoeijs P, Schubert H, Radziejewska T, editors, Biological Oceanography of the Baltic Sea: Springer (in press)

Zettler ML, Karlsson A, Kontula T, Gruszka P, Laine A, Herkül K, Schiele K, Maximov A, Haldin J (2013b) Biodiversity gradient in the Baltic Sea: A comprehensive inventory of macrozoobenthos data. Helgol Mar Res (in press)

http://link.springer.com/article/10.1007%2Fs10152-013-0368-x

Zettler ML, Proffitt CE, Darr A, Degraer S, Devriese L, Greathead C, Kotta J, Magni P, Martin G, Reiss H, Speybroeck J, Tagliapietra D, Van Hoey G, Ysebaert T (2013c) On the myths of indicator species: issues and further consideration in the use of static concepts for ecological applications. PlosONE 8(10):e78219

Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. Methods Ecol Evol 1:3-14

Zuur, AF, Ieno EN, Smith GM (2007) Analysing Ecological Data. Series: Statistics for Biology and Health. Springer Science and Business Media New York 672 pp

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed Effects Models and Extensions in Ecology with R. Springer Science and Business Media New York 574 pp

Zuur AF, Saveliev AA, Ieno EN (2012) Zero Inflated Models and Generalized Linear Mixed Models with R. Highland Statistics Newburgh 323 pp

Danksagung

Ich danke Herrn Prof. Gerhard Graf für die bereitwillige Betreuung meiner Dissertation und die aufmunternden Gespräche. Maren Graf danke ich für die wertvollen Anmerkungen zur englischen Grammatik und zum englischen Redefluss.

Ein ganz besonderer Dank gilt Dr. Michael Zettler für die wohlwollende Aufnahme in seine Arbeitsgruppe, das große mir entgegengebrachte Vertrauen, den permanenten Einsatz für seine Mitarbeiter und das große Verständnis für die Belange praktizierender Väter.

Ich danke allen KollegInnen der AG Benthos im IOW für das gute, familiäre Arbeitsklima und den unermüdlichen Einsatz auf See und im Labor, ohne den, ich sag mal: JAWOLL, weder die Projekte noch diese Arbeit hätten erfolgreich abgeschlossen werden können. Ein besonderer Dank gilt meinen Mitinsassinnen in Zelle 233, die mein gelegentliches Gebrummel stets klaglos über sich ergehen ließen und geduldig Telefonate für mich wechselweise annahmen oder abwimmelten.

Ein großes Dankeschön geht auch an meine alten Weggefährten und Lehrmeister aus Broderstorfer Zeiten, insbesondere an Fritz Gosselck, Jan Kube und Karin Meissner, die mir nicht nur die Grundlagen der anwendungsorientierten Benthosforschung und den Umgang mit konträren „Reviewer“-Meinungen, sondern auch das *eigen*strukturierte Arbeiten unter Zeitdruck gelehrt haben.

Desweiteren danke ich allen Projektpartnern vom BfN und vom AWI sowie den Benthoskollegen der ICES BEWG und der HELCOM-Expertengruppen für die vielen fruchtbaren Diskussionen und dem damit verbundenen unschätzbaren Wissensgewinn! Hervorzuheben sind insbesondere Brigitte Ebbe und ihr Frohsinn, der selbst beim wiederholten Ausfüllen des Begleitdokuments zum Behuf der Erläuterung der Ausfüllhilfe als erklärender Anlage der Basisdatenbank zur Erstbewertung der dem Monitoring unterliegenden bewertungsrelevanten Flächen der FFH-LRT auch *ohne* Monitoringdaten nur kurzfristig ins Wanken geriet.

Edward A. Murphy danke ich für die Formulierung seiner Lebensweisheit, unser aller Fels in der Brandung.

Ein ganz dickes Thanxx! geht an meinen Freund Stefan, der mir rund um die Uhr (*s.l.*) mit seinem unschätzbaren p*skype*ologischen Rat zur Seite stand und steht und der früher oder später auch noch BVB-Fan wird.

Meinen Eltern und meinem Opa gilt mein tiefster Dank, denn ohne ihre finanzielle und moralische Unterstützung wäre ich schon im Grundstudium gescheitert.

Der allerdickste Dank aber gilt meiner geliebten Familie, Johanna, Frederik und Sonja, die jederzeit in der Lage sind, meinen Blick auf die wirklich wichtigen Dinge des Lebens zu lenken und doch in den letzten Jahren auf so viel verzichten mussten, damit ich endlich FERTIG werde.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig angefertigt und ohne fremde Hilfe verfasst habe, keine außer den von mir angegebenen Hilfsmitteln und Quellen dazu verwendet habe und die den benutzten Werken inhaltlich und wörtlich entnommenen Stellen als solche kenntlich gemacht habe.

Rostock, den

Appendix: Additional tables and figures

Table 14: More than 250 additional stations were realized in this thesis sampled during seven cruises.

cruise	Vessel	date	Project	stations
07PE0923	Professor Albrecht Penck	03.-07.08.09	Natura 2000	20
-	Miljö	06.-11.09.09	FEMA	59
07PE1009	Professor Albrecht Penck	28.04.-03.05.10	Natura 2000	39
-	Laura	09.-10.06.12	Cluster 4	9
06EZ1225	Elisabeth Mann-Borgese	19.-22.10.12	Cluster 6	33
EMB047	Elisabeth Mann-Borgese	04.-10.06.13	Cluster 4	52
EMB049	Elisabeth Mann-Borgese	24.-30.06.13	Cluster 6	43

Table 15: Overview of the species mentioned in the text

Cnidaria
- *Edwardsia danica* Carlgren, 1921

Mollusca - Bivalvia
- *Abra alba* (W. Wood, 1802)
- *Acanthocardia echinata* (Linaaeus, 1758)
- *Angulus tenuis* (da Costa, 1778)
- *Arctica islandica* (Linnaeus, 1767)
- *Astarte borealis* (Schumacher, 1817)
- *Astarte elliptica* (Brown, 1827)
- *Astarte montagui* (Dillwyn, 1817)
- *Barnea candida* (Linnaeus, 1758)
- *Cerastoderma edule* (Linnaeus, 1758)
- *Cerastoderma glaucum* (Bruguière, 1789)
- *Corbula gibba* (Olivi, 1792)
- *Ensis directus* (Conrad, 1843)
- *Hiatella arctica* (Linnaeus, 1767)
- *Kurtiella bidentata* (Montagu, 1803)
- *Macoma balthica* (Linnaeus, 1758)
- *Macoma calcarea* (Gmelin, 1791)
- *Musculus subpictus* (Cantraine, 1835)
- *Modiolus modiolus* (Linnaeus, 1758)
- *Musculus discors* (Linnaeus, 1767)
- *Musculus niger* (J.E. Gray, 1824)
- *Mya arenaria* Linnaeus, 1758
- *Mya truncata* Linnaeus, 1758

Mollusca - Gastropoda
- *Amauropsis islandica* (Gmelin, 1791)
- *Odostomia scalaris* MacGillivray, 1853
- *Peringia ulvae* (Pennant, 1777)
- *Theodoxus fluviatilis* (Linnaeus, 1758)

Annelida - Polychaeta
- *Alitta virens* (M. Sars, 1835)
- *Ampharete baltica* Eliason, 1955
- *Euchone papillosa* (Sars, 1851)
- *Hediste diversicolor* (O.F. Müller, 1776)
- *Lagis koreni* Malmgren, 1866
- *Marenzelleria viridis* (Verrill, 1873)
- *Neoamphitrite figulus* (Dalyell, 1853)
- *Nephtys caeca* (Fabricius, 1780)
- *Nephtys ciliata* (Müller, 1788)
- *Pygospio elegans* Claparède, 1863
- *Scoloplos (Scoloplos) armiger* (Müller, 1776)
- *Terebellides stroemii* Sars, 1835
- *Travisia forbesii* Johnston, 1840

Annelida - Oligochaeta
- *Heterochaeta costata* Claparède, 1863

Crustacea
- *Amphibalanus improvisus* (Darwin, 1854)
- *Bathyporeia pilosa Lindström, 1855*

Mytilus edulis Linnaeus, 1758

Mytilus trossulus Gould, 1850

Parvicardium pinnulatum (Conrad, 1831)

Parvicardium scabrum (Philippi, 1844)

Phaxas pellucidus (Pennant, 1777)

Spisula subtruncata (da Costa, 1778)

Tellimya ferruginosa (Montagu, 1808)

Thracia phaseolina (Lamarck, 1818)

Thracia pubescens (Pulteney, 1799)

Corophium volutator (Pallas, 1766)

Diastylis rathkei (Krøyer, 1841)

Tunicata

Ciona intestinalis (Linnaeus, 1767)

Dendrodoa grossularia (Van Beneden, 1846)

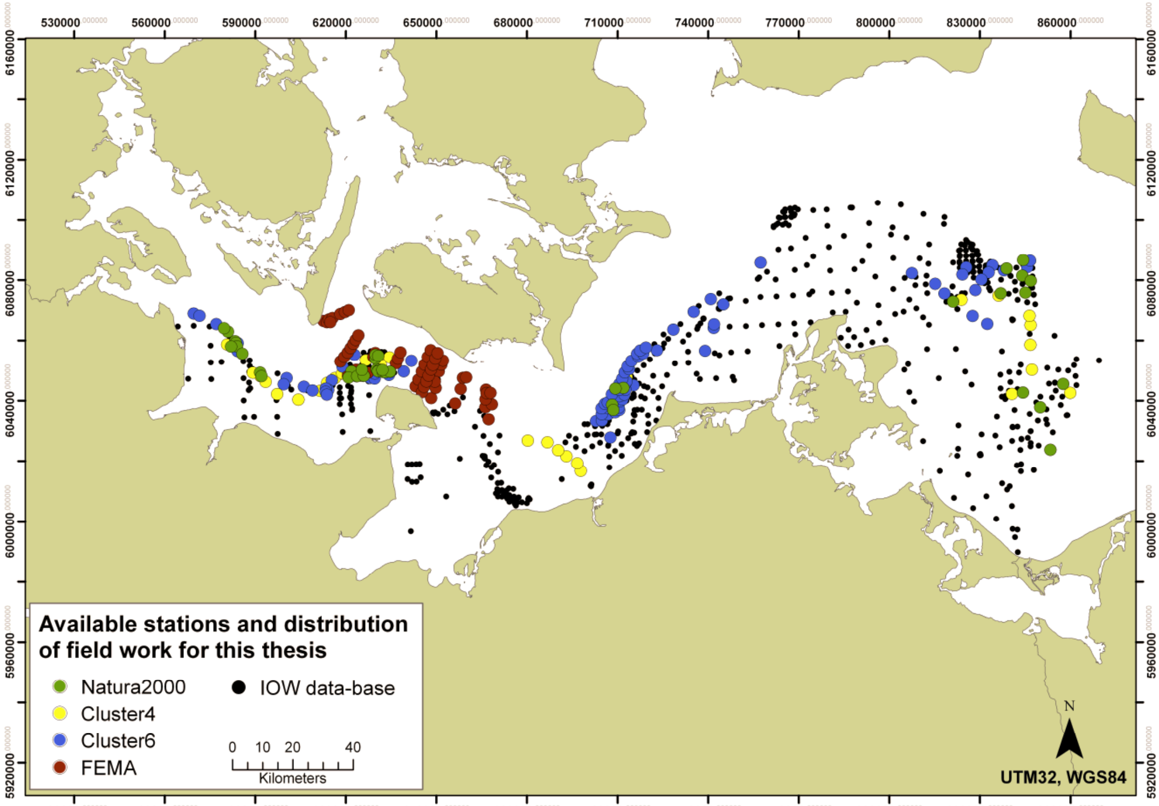


Figure 44: Map of the south-western Baltic Sea showing the available stations of the IOW data base and the stations sampled during this thesis

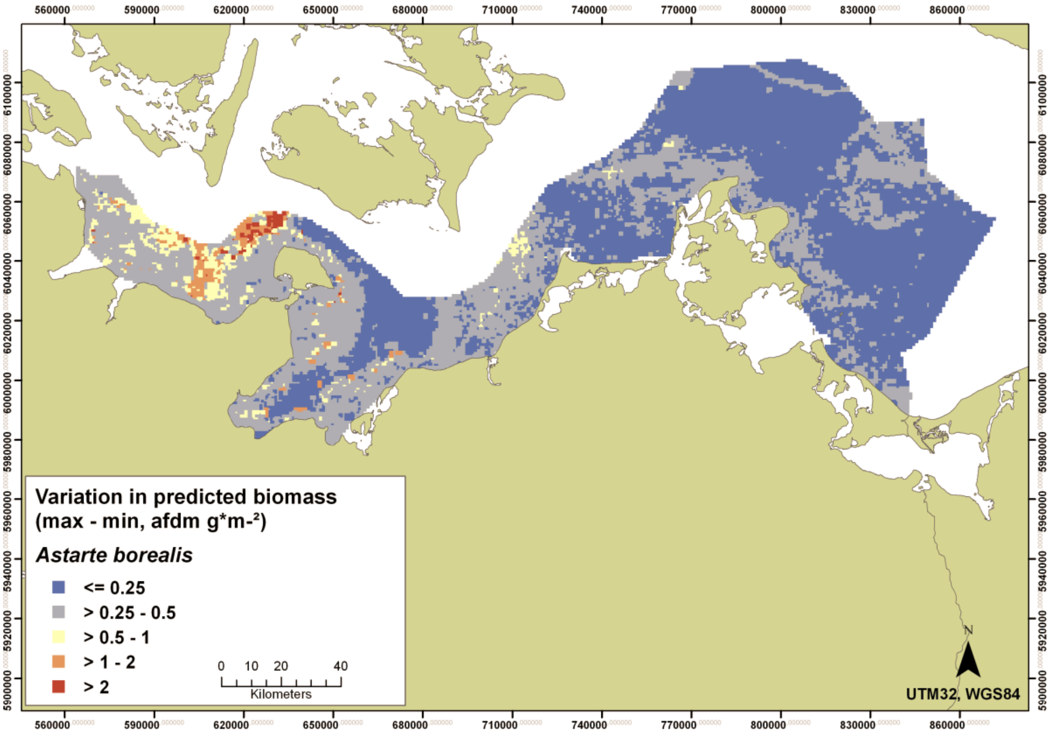


Figure 45: Variation in the predictions for *Astarte borealis* (biomass) as a measure of confidence

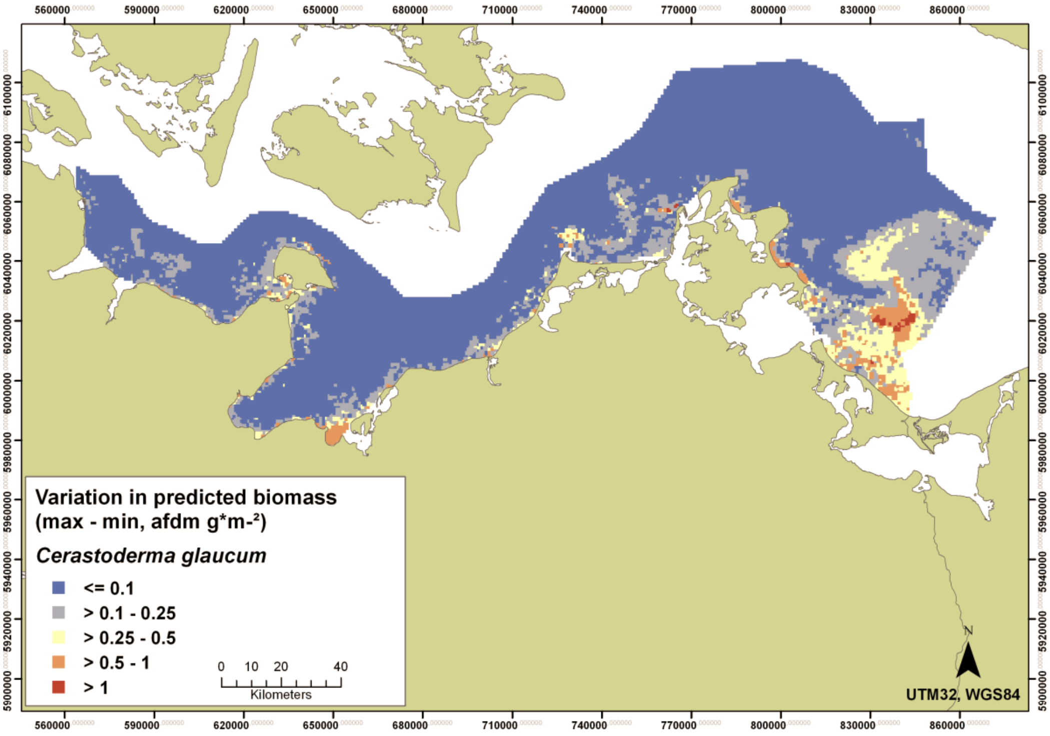


Figure 46: Variation in the predictions for *Cerastoderma glaucum* (biomass) as a measure of confidence

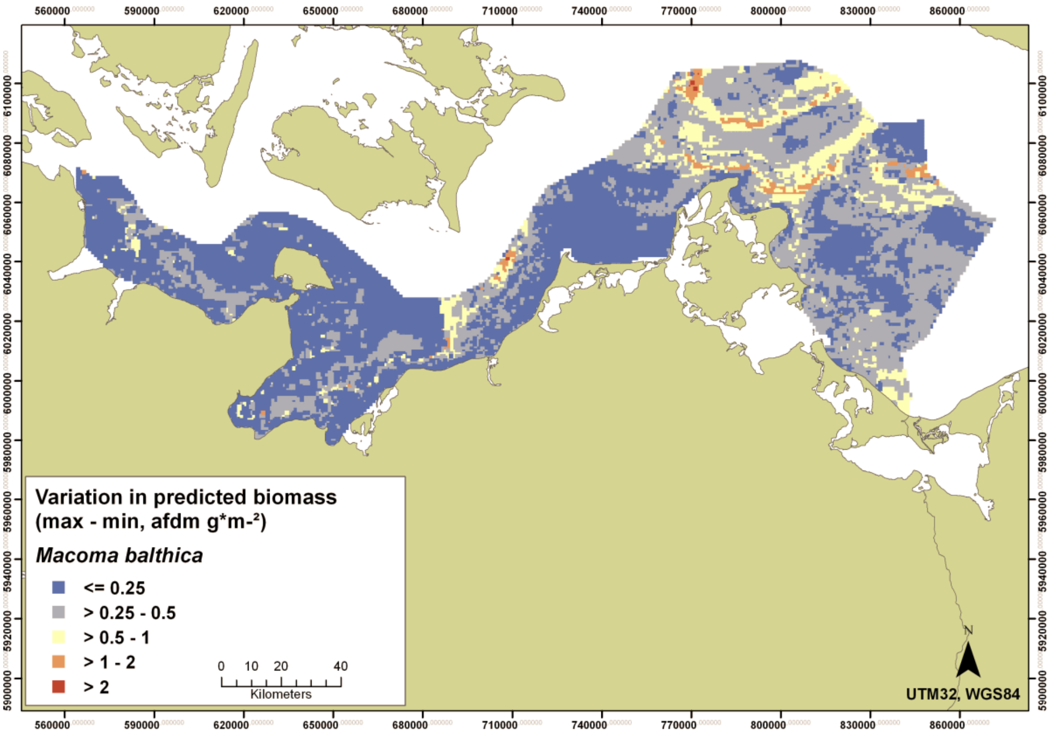


Figure 47: Variation in the predictions for *Macoma balthica* (biomass) as a measure of confidence

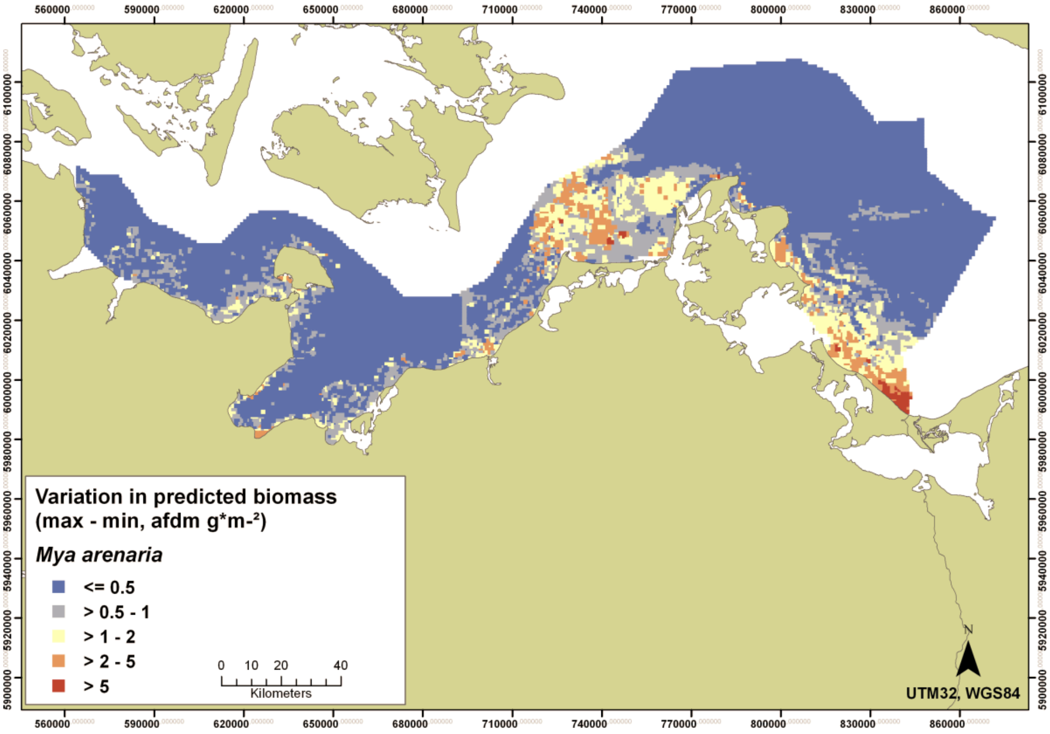


Figure 48: Variation in the predictions for *Mya arenaria* (biomass) as a measure of confidence

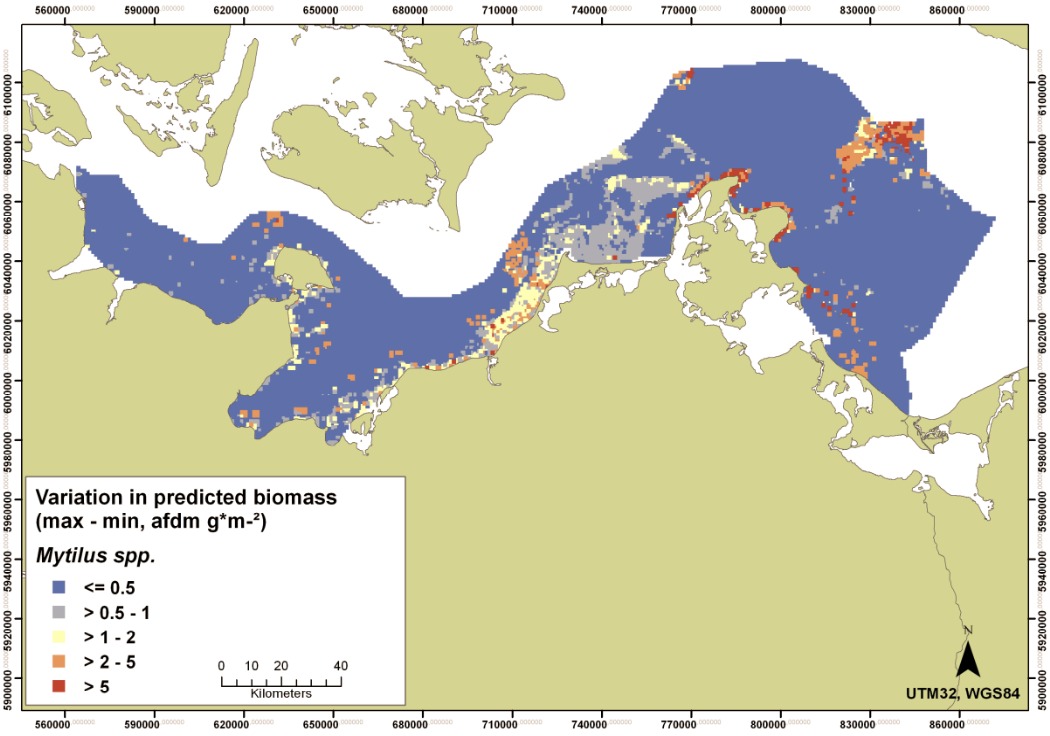


Figure 49: Variation in the predictions for *Mytilus* spp. (biomass) as a measure of confidence

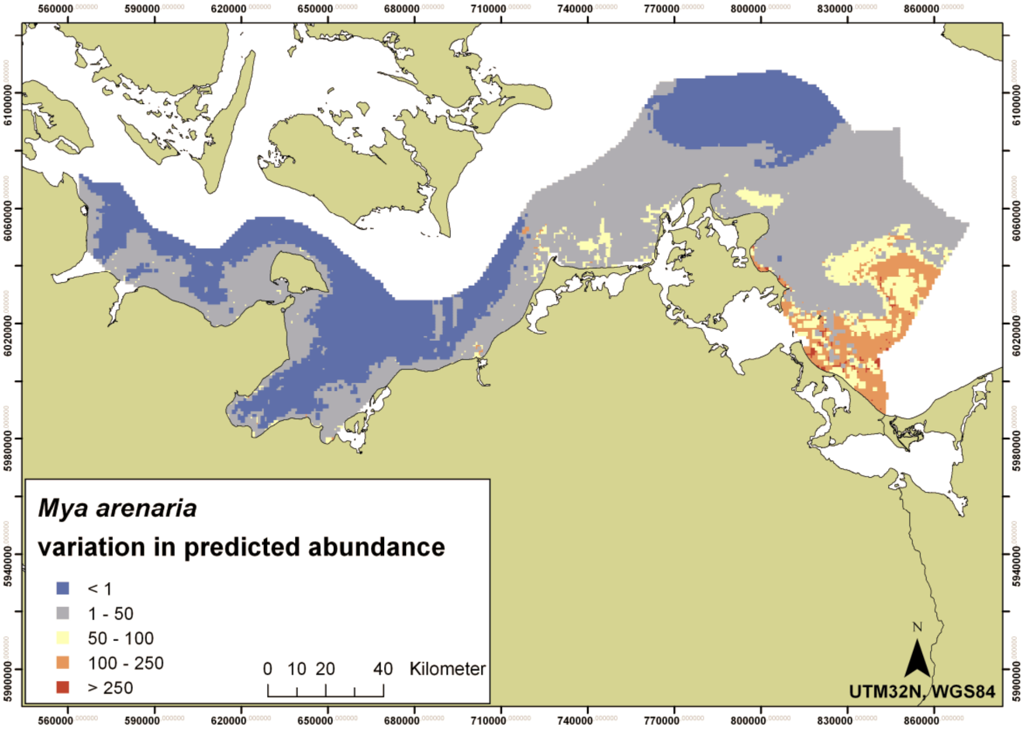


Figure 50: Variation in the predictions for *Mya arenaria* (abundance) as a measure of confidence