

***Final Draft***  
**of the original manuscript:**

Puls, W.; Bernem, K.-H. van; Eppel, D.; Kapitza, H.; Pleskachevsky, A.;  
Riethmueller, R.; Vaessen, B.:

**Prediction of benthic community structure from environmental  
variables in a soft-sediment tidal basin (North Sea)**

In: Helgoland Marine Research (2011) Springer

DOI: 10.1007/s10152-011-0275-y

## **Prediction of benthic community structure from environmental variables in a soft-sediment tidal basin (North Sea)**

W. Puls<sup>\*</sup>, K.-H. van Bernem, D. Eppel, H. Kapitza, A. Pleskachevsky<sup>a</sup>, R. Riethmüller, B. Vaessen<sup>b</sup>

Helmholtz-Zentrum Geesthacht

Max-Planck-Straße 1, D-21502 Geesthacht, Germany

<sup>\*</sup> Corresponding author. E-mail: [walter.puls@hzg.de](mailto:walter.puls@hzg.de)

<sup>a</sup> present address: German Aerospace Center, Oberpfaffenhofen, 82234 Wessling, Germany

<sup>b</sup> present address: Waterways and Shipping Office Cuxhaven, Am Alten Hafen 2, 27472 Cuxhaven, Germany

**Abstract** The relationship between benthos data and environmental data in 308 samples collected from the intertidal zone of the Hörnum tidal basin (German Wadden Sea) was analyzed. The environmental variables were current velocity, wave action, emersion time (all of which were obtained from a 2-year simulation with a numerical model) and four sediment grain-size parameters.

A grouping of sample stations into five benthos clusters showed a large-scale (> 1 km) zoning of benthic assemblages on the tidal flats. The zoning varied with the distance from the shore.

Three sample applications were examined to test the predictability of the benthic community structure based on environmental variables. In each application the dataset was spatially partitioned into a training set and a test set. Predictions of benthic community structure in the test sets were attempted using a multinomial logistic regression model. Applying hydrodynamic predictors, the model performed significantly better than it did when sediment predictors were applied. The accuracy of model predictions, given by Cohen's kappa, varied between 0.14 and 0.49. The model results were consistent with independently attained evidence of the important role of physical factors in Wadden Sea tidal flat ecology.

**Keywords** benthos · habitat suitability modeling · environmental predictor variables · bottom shear stress · multinomial logistic regression · German Wadden Sea

## **Introduction**

In his basic paper Beukema (1976) described the Wadden Sea as an area where zoobenthos is highly influenced by stressing environmental conditions. The observed species-environment relationships can be utilized to develop so-called habitat suitability models which relate the

spatial distribution of benthic species or benthic communities to environmental data.

Correlations between species distributions and the physical environment were first developed for terrestrial applications, for example distribution models for breeding-bird species or plant species. Species distribution models for marine or freshwater systems became more numerous after the year 2000. A comprehensive review on modeling distributions of individual species is given by Elith and Leathwick (2009), while Ferrier and Guisan (2006) review the spatial modeling of biodiversity at the community level. A general survey of predictive modeling in the biosciences is given by Fielding (2007), and Guisan and Zimmermann (2000) discuss the main statistical approaches of predictive habitat distribution modeling.

According to Elith and Leathwick (2009) early studies on species distribution models focused on ecological understanding, “seeking insight, even if indirectly, into the causal drivers of species distributions”. From 1985 onwards (Ferrier and Guisan 2006) studies focused more on the prediction of species or biotic communities. Three examples for the application of predictive habitat modeling are: (a) To generate full coverage spatial distributions of species or communities from point data (e.g. Degraer et al. 2008). The project “Mapping European Seabed Habitats (MESH)” provided a methodological framework for marine habitat mapping (<http://www.searchmesh.net>). (b) To create a basis for the selection of optimal locations for restoration or protection of endangered species (conservation planning, e.g. van Katwijk et al. 2000). (c) To predict future species distributions under climate-change scenarios, e.g. Pearson and Dawson (2003) or Araujo et al. (2005).

In the coastal zone of the southern North Sea, habitat suitability models for benthic species were developed for the Belgian continental shelf (Degraer et al., Willems et al. 2008), for the Schelde estuary (Ysebaert et al. 2002) and for the Dutch Wadden Sea (Brinkman et al. 2002, Bos et al. 2005). In the German Wadden Sea the association between benthic species and environmental data was investigated (a) based on time series on local scales (Damm-Böcker et al. 1993, Niemeyer and Michaelis 1997) and (b) within a study comparing 6 European tidal flats (Compton et al. 2009). The sediment grain-size was the standard environmental variable for the building of habitat suitability maps, but most studies also used hydrodynamic data (currents and/or waves, as recommended by Snelgrove and Butman 1994) calculated by numerical models.

In this study, two datasets were employed to analyze the relations between the benthic community structure and environmental variables in the Hörnum tidal basin, an intertidal soft-sediment environment in the German Wadden Sea. The first one was a benthos dataset of 308 sample stations. The second one was a dataset including (a) sediment grain-size parameters as point data and (b) full coverage distributions of model simulated current velocity, wave parameters and emersion time.

This study had two main objectives. The first one was to assess the general applicability of a habitat suitability model to predict the spatial pattern of benthic communities in a tidal basin of the German Wadden Sea. The second objective was to evaluate the prediction efficiency of a model using hydrodynamic variables as the only environmental predictors, and to compare this prediction efficiency with the prediction success of models using (a) a set of predictors which included sediment variables only and (b) a set of predictors which included both hydrodynamic and sediment variables.

## **Materials and methods**

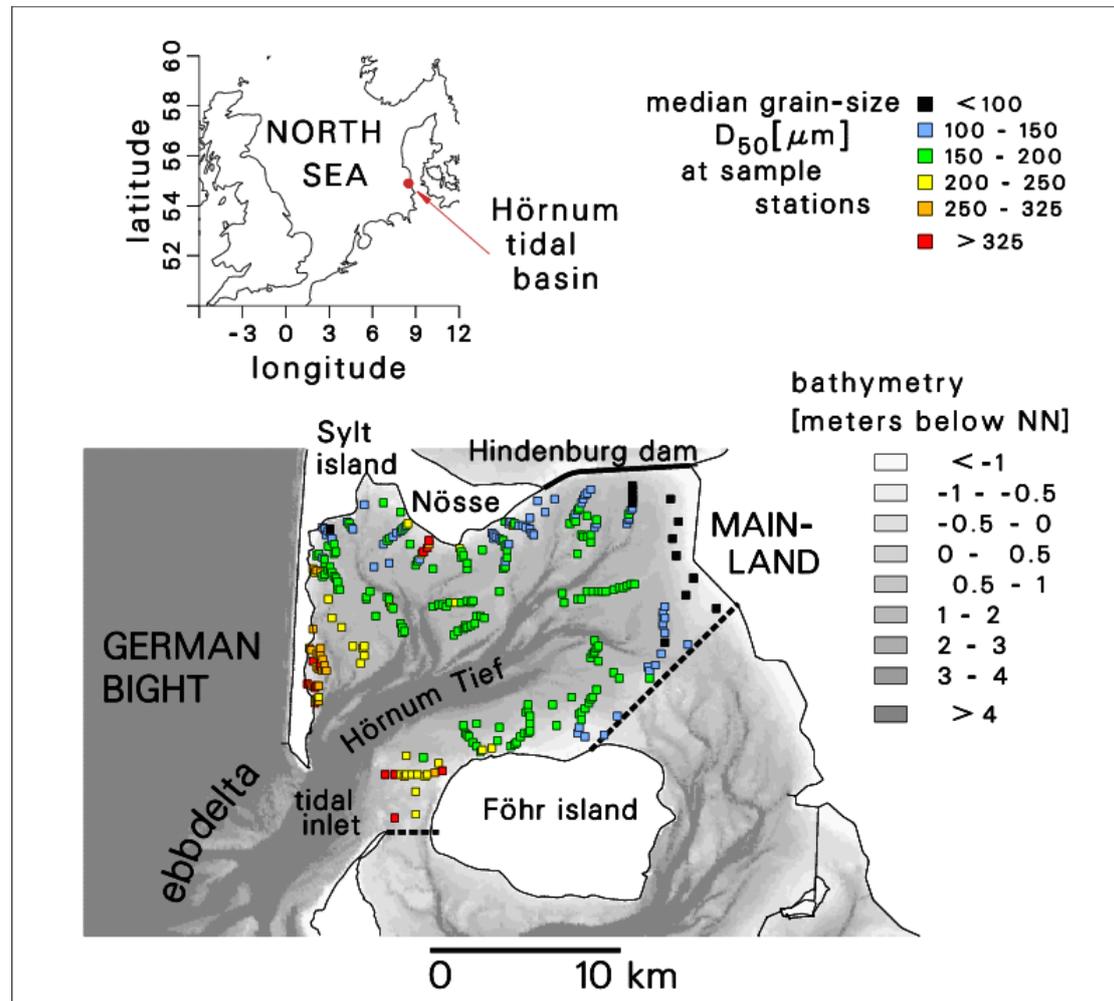
### **Study area**

The Hörnum tidal basin is part of the Wadden Sea which extends along the coast of the German Bight in the SE part of the North Sea (Fig. 1). The tidal basin is connected to the German Bight via the Hörnum Tief, a tidal inlet with a maximum water depth of 30 m. The basin is linked to the adjacent Wadden Sea in the south, while the connection to the Wadden Sea in the north is restricted by the Hindenburg dam. Figure 1 shows the bathymetry with special attention to tidal flats.

The tide in the Hörnum tidal basin is semi-diurnal with a mean tidal range of 2 m at the tidal inlet. The difference between the tidal ranges of spring and neap tide is only 10 % of the mean tidal range. Maximum tidal current velocities are between 1 and 1.4 m/s in the Hörnum Tief and a few decimeters/s on the tidal flats. Surface water waves in the basin are generated by local winds. Waves entering from the open sea are dissipated within the basin's extensive tidal ebb delta (Ross et al. 1998).

Climatically, the Wadden Sea belongs to the cold temperate region with rainy summers and mild winters. On tidal flats, the water temperatures in summer are between  $> 30$  °C (residual waters at low tide) and around 15 °C at high tide. In the Hörnum basin, rainfall ( $\approx 900$  mm  $y^{-1}$ ) and freshwater drainage through a floodgate ( $\approx 1$  m<sup>3</sup> s<sup>-1</sup>) are negligible

compared to the tidal prism of  $0.5 \text{ km}^3$ . The typical salinity is 28-30. The concentration of suspended particulate matter in the Wadden Sea is 5-40 mg/l during calm weather.



**Fig. 1** Map of the Hörnum tidal basin, including the bathymetry (gray shading) and the positions of 308 benthos sample stations. The bed elevation NN is the German ordnance level, approximately identical with the mean sea level. The colors of single stations indicate the distribution of median grain-sizes of bottom sediment. The two dashed lines indicate the southern boundary of the investigation area

About 50 % ( $207 \text{ km}^2$ ) of the Hörnum tidal basin consists of tidal flats. At  $20 \text{ km}^2$ , its fraction of salt marshes is comparatively small. The sediment on the tidal flats is mostly fine to medium sand, see Fig. 1. A special case is the Wadden Sea area along the southern spit of Sylt, where the predominantly westerly winds blow sand from the land-dunes onto the adjacent tidal flats. Another special case with respect to grain-size is the region south of Nösse peninsula. Here Pleistocene till was found in a small area. Altogether, the sediment in the Hörnum tidal basin is rather

coarse compared to other tidal flat systems in north-western Europe (Compton et al. 2009).

**Table 1** Characteristics of five sample station groups as classified by hierarchical clustering with group-average linking. For each cluster group, the occurrences (% present) of the single species are given. Bold presence numbers indicate the characteristic species of each cluster group. The lower part of the table shows the sample means and the standard deviations of the environmental variables within the cluster groups. The means and the standard deviations of the sediment grain-size data were calculated in logarithmic terms, thereafter the anti-logarithms were taken. For each environmental variable, both the maximum sample mean and the minimum sample mean are indicated using bold text.

Cluster No. →	1	2	3	4	5
number of samples in cluster →	74	106	103	11	10
	occurrence of benthic species (% present)				
<i>Arenicola marina</i> (6 classes)	93	97	<b>100</b>	73	30
<i>Cerastoderma edule</i> (ind./m <sup>2</sup> )	91	78	40	55	100
<i>Lanice conchilega</i> (3 classes)	4	28	7	<b>100</b>	20
Diatoms (% bottom cover)	36	14	5	9	<b>100</b>
<i>Zostera marina</i> (% bott. cover)	<b>85</b>	8	2	-	-
<i>Zostera noltii</i> (% bottom cover)	<b>91</b>	5	2	-	-
<i>Enteromorpha</i> spp. (% bott. co.)	31	<b>94</b>	1	45	20
<i>Ulva lactuca</i> (3 classes)	35	<b>82</b>	1	18	80
<i>Fucus</i> spp. (3 classes)	8	8	5	10	-
<i>Porphyra</i> spp. (3 classes)	1	10	-	-	-
<i>Salicornia europaea</i> (3 classes)	3	-	-	-	-
<i>Spartina anglica</i> (3 classes)	7	1	-	-	-
<i>Mytilus edulis</i> (6 classes)	7	25	2	27	30
	Sample-means ± standard deviations of environmental data				
emersion time “dry” (%)	<b>41</b> ± 17	21 ± 17	20 ± 19	<b>8</b> ± 5	13 ± 14
$\tau_c^{\text{mean}}$ (N/m <sup>2</sup> )	0.037	0.069	0.091	<b>0.099</b>	<b>0.031</b>
	± 0.026	± 0.027	± 0.041	± 0.028	± 0.005
$\tau_c^{\text{max}}$ (N/m <sup>2</sup> )	0.28	0.40	<b>0.55</b>	0.44	<b>0.19</b>
	± 0.14	± 0.16	± 0.25	± 0.12	± 0.03
$\tau_w^{\text{mean}}$ (N/m <sup>2</sup> )	<b>0.25</b>	0.29	0.30	<b>0.31</b>	0.26
	± 0.06	± 0.05	± 0.10	± 0.07	± 0.02
$\tau_w^{\text{max}}$ (N/m <sup>2</sup> )	<b>1.43</b>	1.51	1.60	<b>1.61</b>	1.50
	± 0.21	± 0.15	± 0.32	± 0.19	± 0.04
10th percentile grain-size D <sub>10</sub> (µm)	85	116	<b>138</b>	100	<b>41</b>
	+100/-46	+37/-28	+61/-42	+85/-46	+42/-21
median grain-size D <sub>50</sub> (µm)	156	166	<b>201</b>	173	<b>72</b>
	+86/-55	+37/-30	+59/-46	+21/-19	+70/-36
90th percentile grain-size D <sub>90</sub> (µm)	245	235	<b>308</b>	274	<b>175</b>
	+172/-	+76/-57	+142/-97	+83/-64	+89/-59
	101				
sediment grain-size sorting D <sub>90</sub> /D <sub>10</sub>	2.89	<b>2.04</b>	2.23	2.74	<b>4.29</b>
	+2.5/-1.3	+0.7/-0.5	+1.0/-0.7	+2.8/-1.4	+1.7/-1.2

The Hörnum tidal basin is a heterotrophic, nutrient-rich system (e.g. Asmus and Asmus 1998). The water column is always well mixed,

therefore problems linked with oxygen deficiency do not occur. Due to its dominance on sandy tidal flats, the lugworm *Arenicola marina* is the key species in the tidal basin (Reise 1985). The fishing of the cockle *Cerastoderma edule* has been prohibited since 1990. In 1999, an area of complete “no-use” was established in the eastern half of the Hörnum tidal basin. In the tidal channels of the western basin, however, there is fishing for shrimp *Crangon crangon* and farming of the blue mussel *Mytilus edulis*.

## Benthos data

Benthos data were sampled by the Helmholtz-Zentrum Geesthacht and by the Institute of Applied Biology, Freiburg/Elbe. The investigated area was the entire German Wadden Sea, the aim being to provide data for assessing the vulnerability of the German North Sea coast to oil spills (van Bernem et al. 2007). In order to capture the diversity of habitats on the wide tidal flats within a reasonable time, the collection of data at each sample station was restricted to (a) visually classifying the abundances of benthic species at the bed surface (plot area about 50·50 m<sup>2</sup>) and (b) counting *C. edule* by raking the sediment up to a depth of 10 cm in four areas of 0.25 m<sup>2</sup> each. In the Hörnum tidal basin, 14 benthos species were recorded at 391 sample stations during the summer months of 2001, 2002 and 2003. As the soft-shell clam *Mya arenaria* was only recorded at some of the sample stations, *M. arenaria* was not considered here. Among the 13 remaining species (listed in the upper part of Table 1), eight species were categorized into classes:

- three classes (c = 0, 1, 2) as “absent”, “sporadic” and “dense” for the sand mason worm *Lanice conchilega*, the green macroalga *Ulva lactuca*, the brown macroalgae *Fucus* spp., the red macroalgae *Porphyra* spp., the glasswort *Salicornia europaea*, the common cordgrass *Spartina anglica*
- six classes (from c = 0 for “absent” to c = 5 for “compact mussel bank”) for *M. edulis*.
- six classes for *A. marina*. The classes c = 0 to 5 mean 0, <1, 1-10, 11-50, >50 and >>50 individuals/m<sup>2</sup>. The categories 4 and 5 were discarded (and replaced by category 1) because they are typical for juvenile lugworms. The main influence on the population density of juvenile lugworms is the presence of adult lugworms, as juveniles settle where adult densities are low and vice versa (Flach and Beukema 1994, Reise et al. 2001). Highest abundances of (juvenile) *A. marina* thus exist in areas which are less attractive for adult *A. marina*. The inclusion of juvenile lugworms into the data analysis would thus have lead to confusing results. The

replacement of categories 4 and 5 by category 1 was undertaken because adult lugworms were sporadically observed (i.e.  $c = 1$ ) where high abundances of juvenile lugworms were present.

The other five species were quantified as

- individuals/m<sup>2</sup> (*C. edule*)
- % bottom cover (diatoms, the green macroalgae *Enteromorpha* spp., the seagrasses *Zostera marina* and *Zostera noltii*). “Diatoms” is a group of a large number of microalgae, but for the sake of simplicity this is not mentioned any further.

Similarities between samples were calculated as Bray-Curtis coefficients (e.g. Clarke and Warwick 1994). Bray-Curtis coefficients are only computable if there are no missing values in the species data. There were 308 sample stations in the Hörnum tidal basin which included data for each of the 13 species. These sample stations are shown in Fig. 1. The benthos data sampled in the years 2001, 2002 and 2003 were combined into a single dataset. The justification for this was shown by a three-fold cross-validation test as presented in Degraer et al. (2008).

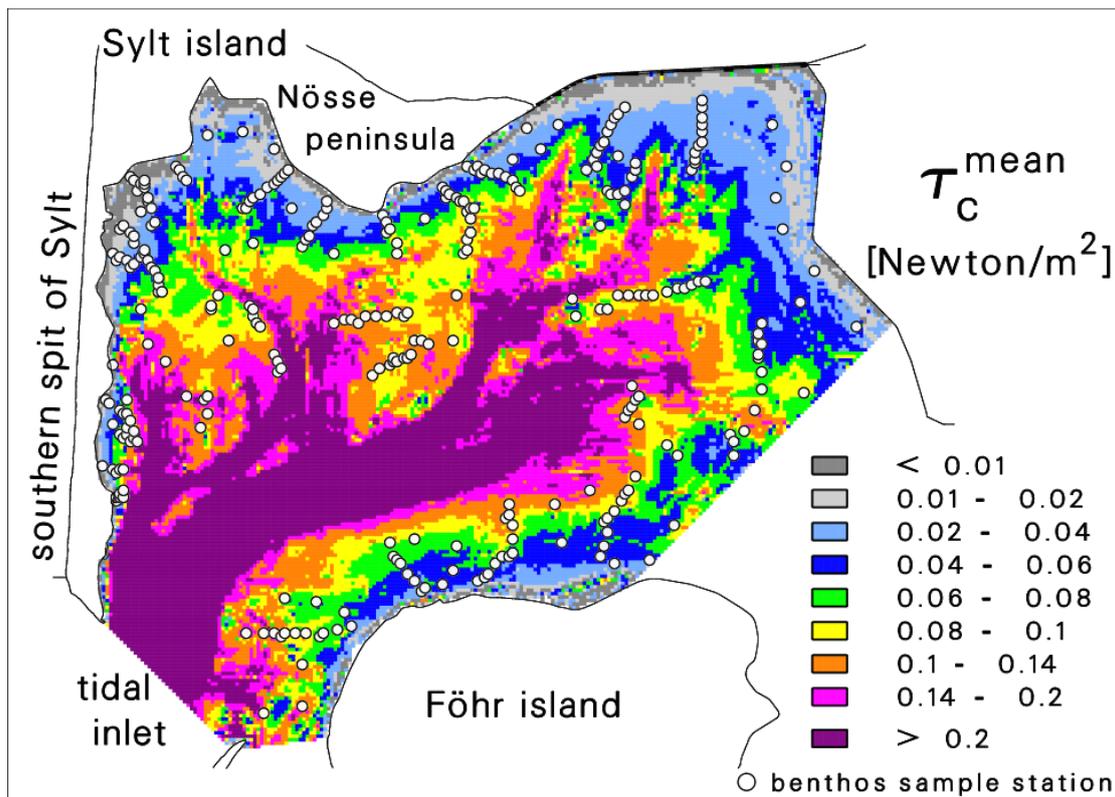
Distances between sample stations were roughly 100 to 1000 m. Ysebaert and Herman (2002) identified the dominating spatial patterns observed in the macrobenthic assemblages of the Schelde estuary at a scale of 100 to 500 m. According to Ysebaert et al. (2002), the benthic community structure at this scale was “strongly and directly coupled to physicochemical processes” (in the Schelde estuary).

#### Transformation of benthos data

For a calculation of Bray-Curtis similarities between sample stations the attributes (values) of the 13 benthic species should be of comparable magnitude; otherwise the continuous abundances of certain species (high coverage percentages of diatoms, *Z. marina*, *Z. noltii*, *Enteromorpha* spp. and high numbers of individuals per m<sup>2</sup> of *C. edule*) would dominate the low category-numbers of the other species. The data transformation from the original data “c” to the transformed data “c\*” was carried out as follows:

- a. The data of the species with 3 classes were not changed; their original category numbers ( $c = c^* = 0, 1, 2$ ) were used as a reference for the other species. In particular, the arithmetic mean (not taking into account the stations with absence “0”) of *L. conchilega* was used as a yardstick for transforming the five species with continuous abundances. This arithmetic mean was 1.33.

- b. The original *A. marina*-attributes “c = 1” and “c = 2” were united to “c\* = 1”, the original attribute “c = 3” was converted to “c\* = 2”.
- c. The original *M. edulis*-attributes “c = 1 to 5” were converted to “c\* = 1, 1.2, 1.5, 1.7, 2”.
- d. The data of the five species with continuous abundances were transformed using power transformations  $c_* = c^\lambda$  with a specific  $\lambda$  for each of the five species. The aim of the transformations was to obtain the same arithmetic  $c_*$ -mean (= 1.33, not taking into account the stations with absence “0”) for each of the five species.

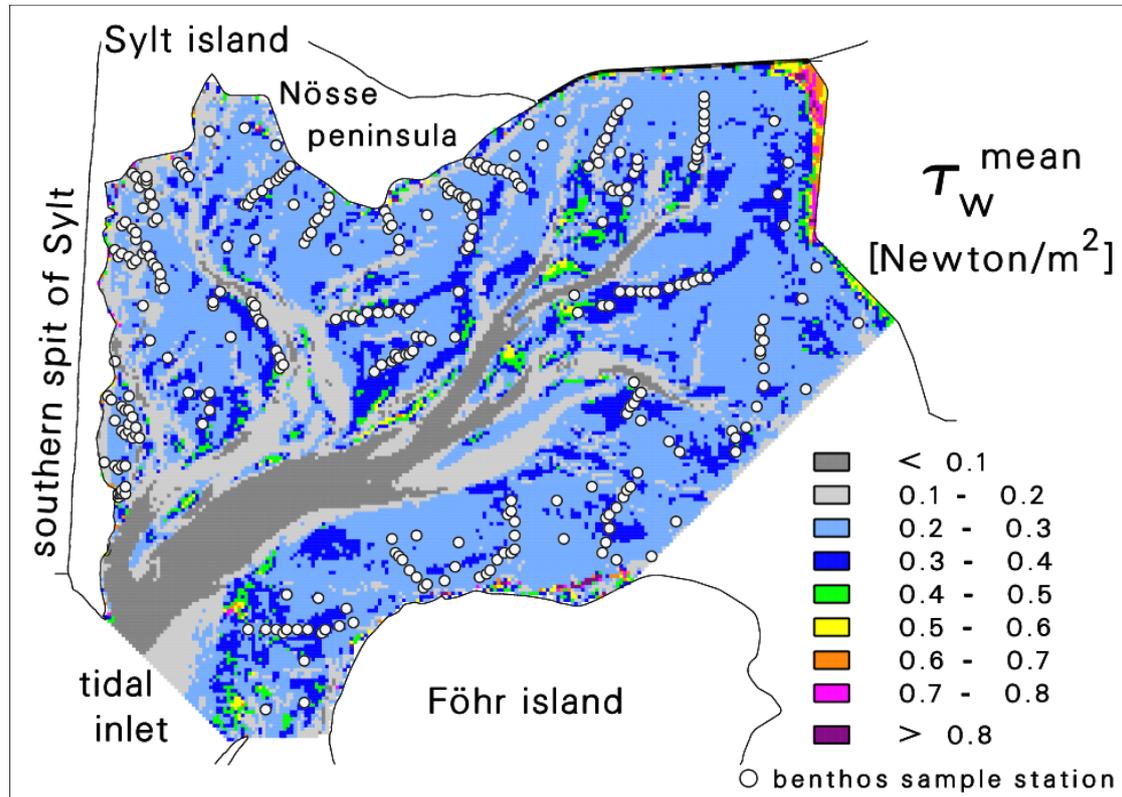


**Fig. 2** Distribution of time-averaged bed shear stress  $\tau_c^{\text{mean}}$  (generated by near-bed currents) in the Hörnum tidal basin. Highest  $\tau_c^{\text{mean}}$ -values appear in the tidal channels. On the tidal flats,  $\tau_c^{\text{mean}}$  decreases continuously from the tidal channels towards the coast

### Abiotic variables

A hydrodynamic model was used to calculate a 2-year time-series (November 1999 – October 2001) of water levels, current velocities and wave parameters (e.g. significant wave height, mean wave period) in the Hörnum tidal basin. The hydrodynamic model consisted of the 3-dimensional current model TRIM3D (Casulli and Stelling 1998) with 100 m horizontal grid size and the Wadden Sea version of the “k-model” (400 m grid size) for computing the wave spectrum (e.g. Moghimi et al. 2005).

The current model and the wave model were synchronously coupled meaning they were run in parallel, permanently exchanging information. For increasing the model speed, waves were not calculated in the Wadden Sea areas south of the dotted lines in Fig. 1. In order to have a complete and consistent hydrodynamic data set, the wave data were interpolated to the 100 m grid cells of the current model.



**Fig. 3** Distribution of time-averaged bed shear stress  $\tau_w^{\text{mean}}$  (generated by surface wave action) in the Hörnum tidal basin. The highest  $\tau_w^{\text{mean}}$ -values, e.g. along the mainland coast, were generated on salt marshes. The reason for those high  $\tau_w^{\text{mean}}$ -values is that only times of inundation were taken into account for time averaging. On the salt marshes, flooding only happened when there were strong westerly winds with correspondingly strong wave action. High  $\tau_w$  were also generated along the margins of the tidal channels where deeper-water waves “abruptly” came into contact with the shallow bottom of the tidal flat. The  $\tau$ -scales in this figure and in Fig. 2 are not identical

The model results were validated against observed data (Eppel et al. 2006). The observed data included water levels measured at tidal gauges within the basin, current velocities measured within cross-sections in the Hörnum Tief and waves measured by a waverider buoy, also positioned in the Hörnum Tief. There were no wave or current measurements on tidal flats for the two simulated years.

Computed current velocities and computed wave parameters were transformed into bed shear stresses  $\tau_c$  and  $\tau_w$ , respectively, with  $\tau_c$  being the (absolute) bed shear stress due to near-bed currents and  $\tau_w$  being the amplitude of the oscillatory bed shear stress due to waves. The bed shear stress  $\tau$  describes the impact of hydrodynamics on the sea bed; it is traditionally used for calculating sediment transport. The formulas used for calculating  $\tau_c$  and  $\tau_w$  are given in Eppel et al. (2006); they are based on Soulsby (1997). By using the bed shear stresses  $\tau_c$  and  $\tau_w$  instead of current velocities and wave data, the impacts of currents and waves on the sea bed are directly comparable. It must be kept in mind that an oscillating near-bed velocity generated by a wave produces a much higher bed shear stress than the same near-bed current velocity of a unidirectional current.

The bed shear stress depends on a friction factor, which, in turn, depends on a bed roughness. In this study, a skin roughness length of  $z_0 = 17 \mu\text{m}$  was used in the entire model area, which corresponds to a median sediment grain-size of  $200 \mu\text{m}$  (typical for Hörnum tidal basin). By using one single bed roughness for the entire basin, the spatial distributions of  $\tau_c$  and  $\tau_w$  only depend on the currents and wave parameters calculated by the model, not on the spatial sediment distribution.

For each of the 308 sample stations, nine environmental variables were added to the benthos data. The means of  $\tau_c$  and  $\tau_w$  were average values of the 12 summer months April-September for the years 2000 and 2001. The same applies to the emersion time “dry”. These mean values represent regular (tidal) conditions. For the same time periods, maximum values of  $\tau_c$  and  $\tau_w$  represent stormy conditions. The nine environmental variables were:

**(1)** The time fraction “dry” (in %) of a station’s emersion. The variable “dry” is comparable to “depth”, “elevation” or “bed level height” used in other studies. The spatial distribution of “dry” is similar to the bathymetry shown in Fig. 1.

**(2)** The mean bed shear stress  $\tau_c^{\text{mean}}$  generated by near-bed currents. The distribution of  $\tau_c^{\text{mean}}$  is shown in Fig. 2. Over the tidal flats, currents are mainly in cross-slope direction, they are controlled by the continuity equation. The shown distribution of  $\tau_c^{\text{mean}}$  is typical for a concave tidal flat bottom profile (Friedrichs and Aubrey 1996) as it prevails in the Hörnum basin and also in the List tidal basin located to the north (Reise 1998).

**(3)** The maximum bed shear stress  $\tau_c^{\text{max}}$  generated by maximum near-bed currents. The spatial pattern of  $\tau_c^{\text{max}}$  (not shown) is roughly similar to that of  $\tau_c^{\text{mean}}$ .

(4) The mean bed shear stress  $\tau_w^{\text{mean}}$  generated by surface waves, shown in Fig. 3. The accuracy of modeled wave parameters on the tidal flats could not be tested because wave measurements were not available. But there are two considerations which suggest that the spatial distributions (not necessarily the absolute values) of calculated wave parameters (and thus also of  $\tau_w$ ) are reasonable: (a) The tidal flat morphology in Hörnum tidal basin is in a dynamic equilibrium (Hirschhäuser and Zanke 2001) which means, according to Friedrichs and Aubrey (1996), that the bottom shear stress is expected to be rather uniform over the tidal flats. As  $\tau_w^{\text{mean}}$  was much higher than  $\tau_c^{\text{mean}}$  on the tidal flats, tidal flat sediment erosion was governed by wave activity and not by currents. This is also indicated by the concave tidal flat bottom profiles, Friedrichs and Aubrey (1996). It is thus  $\tau_w^{\text{mean}}$  which is expected to be nearly uniform over the tidal flats. As shown in Fig. 3, this is indeed the case. (b) On tidal flats the modeled significant wave height was linearly correlated with the water depth (correlation coefficient = 0.81). This model result agrees with the findings obtained from field measurements in the German Wadden Sea (Niemeyer 1979, Grüne 2009).

(5) The maximum bed shear stress  $\tau_w^{\text{max}}$  generated by waves. The spatial pattern of  $\tau_w^{\text{max}}$  (not shown) is approximately similar to that of  $\tau_w^{\text{mean}}$ .

(6 → 9) The bottom sediment percentile grain-sizes  $D_{10}$ ,  $D_{50}$  and  $D_{90}$  plus the degree of sediment sorting  $D_{90}/D_{10}$ . The sediment was collected with tubes (diameter 6 cm) from 0 to 7 cm sub-bottom depth at the benthos sample stations. The dry sediment sample was weighed (mass 1) before the sample was analyzed. After removal of organic sediment with  $\text{H}_2\text{O}_2$ , fine sediment was washed away through a 53  $\mu\text{m}$  sieve. The remaining mineral sediment  $> 53 \mu\text{m}$  was again dried and weighed (mass 2); thereafter the grain-size distribution  $> 53 \mu\text{m}$  was determined by dry sieving. There were sediment percentile grain-sizes for 232 of the 308 benthos stations. For the remaining 76 sample stations, sediment parameters were obtained from data interpolation, from a more comprehensive dataset covering the years 1972-73 (Figge 1981), 1987 (van Bernem et al. 1994), and 2001-2003. The median grain-sizes  $D_{50}$  at the 308 sample stations are given in Fig. 1. By and large, the sediment grain-size in the Hörnum tidal basin decreases with increasing distance from the tidal inlet, see also Hirschhäuser and Zanke (2004). For using the four sediment parameters in data processing, they were logarithmized to base 10. A common sediment parameter, the combined percentage of silt and clay (“mud content”, fraction  $< 63 \mu\text{m}$ ), was not included because the difference of mass 1 and mass 2 turned out to be afflicted with a large relative error, in particular for the coarsest samples. However,  $D_{10}$  can be taken as a surrogate for mud content.

Hydrodynamic data were available in each grid cell (grid size 100 m) of the numerical current model, while sediment data were present at scattered sample points. To create full coverage sediment maps which were consistent with the gridded hydrodynamic data, the point data were interpolated to the model grid cells. The interpolation was done using an inverse distance algorithm. In some parts of the tidal flats, the observed sediment data were not dense enough for an interpolation. So the set of environmental variables was incomplete in those parts and a prediction of the benthic community structure was not possible there.

## Data analysis

It was decided not to model distributions of individual species one at a time, but to model biodiversity at the community-level (Ferrier and Guisan 2006). The benefit of community-level modeling was that the complete set of 13 species was included in the model. When distributions of individual species are modeled, species with too low occurrences (e.g. *S. europaea*) “are usually excluded from further analysis (for statistical reasons)”. In this study, the modeling strategy “assemble first, predict later” (Ferrier and Guisan 2006) was applied. This means that in a first step the benthos survey data were “subjected to some form of ... aggregation, without any reference to the environmental data”. In a second step the aggregates were “modeled as a function of environmental predictors”. For alternative strategies see Ferrier and Guisan (2006).

In order to specify the structure of the benthic community, the sample stations were classified into 5 groups, using hierarchical clustering with group-average linking of Bray-Curtis similarities calculated on transformed benthos data. After hierarchical clustering, the membership of some sample stations was “corrected” using a partitioning method (e.g. Fielding 2007). The reason for the correction was as follows: After classifying 308 sample stations into five cluster groups (using group-average linking), 40 sample stations showed a higher Bray-Curtis similarity to another group than to the group they were assigned to. The cluster memberships were changed for those 11 stations which showed the greatest mismatch of similarities. In the logistic fits to environmental covariates (see below) the ambiguity of cluster memberships was taken into account by allocating weights to sample stations.

As an alternative to “hierarchical clustering with group-average linking”, the following cluster analysis methods were tested: (a) the agglomerative hierarchical methods “nearest neighbor”, “furthest neighbor” and “minimum variance”, (b) the partitioning method “K-means” and (c) the

divisive hierarchical method “TWINSPAN” (e.g. Fielding 2007). Hierarchical clustering with group-average linking was selected because cluster memberships obtained with this method could best be fitted by a multinomial logistic regression using a *small* number (up to three) of benthic species as explanatory variables. A best fit meant that the cluster grouping was best interpretable in terms of a few major structuring species.

Multinomial logistic regression (Hosmer and Lemeshow 2000) was used for the classification of cluster-memberships as a response to environmental variables. The environmental predictor variables were 18 so-called covariates combining (1) the z-score transformed environmental variables plus (2) the squares of the z-score transformed variables. By using the environmental variables both in linear and squared terms, there was a chance to describe both linear and unimodal relations between benthic community composition and environmental variables (Ysebaert et al. 2002). The response variables were the (weighted) cluster memberships of sample stations. The weight of a particular station belonging to cluster “i” was its (Bray-Curtis) similarity to cluster “i”, divided by the sum of the station’s similarities to all five clusters. The results of a multinomial logistic regression (at a particular site) were classification probabilities  $q_i$  for each cluster  $i$  ( $i = 1 \dots 5$ ), based on the environmental data for that particular site. The five probabilities always sum up to unity.

The analytical design consisted of three parts (Randin et al. 2006): training, spatial prediction and testing. The complete set of sample stations (in the entire basin) was subdivided into a training set and a test set. Three different subdivisions were applied, see section “Sample applications and groups of predictor variables”. The task was to generate a relation between the environmental covariates and the cluster memberships, based on the sample station data of the training set. This relation was then used (a) for predicting a cluster membership for each grid cell of the basin’s tidal flats and (b) for classifying the (a priori unknown) cluster memberships of the test set stations. The validity of a modeled classification was quantified by CCR, the “correct classification rate”, and by Cohen’s kappa (e.g. Fielding and Bell 1997).

*Model training:* The training set was subdivided (by random (unstratified) selection of stations) into a fit sample and a hold-out sample of equal size. A logistic regression was conducted with the data of the fit sample, testing each possible combination of a specified number of covariates (exhaustive search). The final result was a “best predicting set of

covariates”. This “best predicting set of covariates” was found using two criteria. First criterion was to achieve the highest CCR not for the fit sample, but for the hold-out sample. In this way *that* set of covariates was determined which was best able to generalize beyond the fitting data (see also Elith and Leathwick 2009). Second criterion was to avoid overfitting. This was done by inspecting the result of that spatial prediction (see below) which was generated by the applied set of covariates. If two sets of covariates performed equally, the one with less covariates was selected for the sake of model robustness. After having found the “best predicting set of covariates”, the next step was to use the data of the complete training set for a multinomial logistic regression. This logistic regression model (with the “best predicting set of covariates”) was the “habitat suitability model”.

*Spatial prediction:* Environmental data were present in (almost) each grid cell (grid size 100 m) on tidal flats. Depending on the covariate values in a particular grid cell, the habitat suitability model predicted cluster-probabilities  $q_i$  ( $i=1\dots5$ ) for this grid cell. The cluster group with the highest  $q_i$  was allocated to this grid cell (“majority decision”). A decision threshold was not applied. The result of spatial prediction was an (almost) full coverage distribution of cluster-memberships on the tidal flats.

*Model testing:* At each test set sample station, the result of spatial prediction (cluster group with the highest predicted probability  $q_i$ ) was compared with the station’s “observed” (Fig. 5) cluster group membership. The goodness of the comparison was interpreted by Cohen’s kappa which takes values between -1 and +1. A value of “0” means “no correlation”, while “+1” means perfect agreement. According to Landis and Koch (1977),  $0.4 < \kappa < 0.6$  points towards a “moderate agreement” between observed and predicted cluster memberships, while  $0.2 < \kappa < 0.4$  means “fair agreement”. In an abridged interpretation of Landis and Koch (1977), Fielding (2007) suggested that “ $\kappa < 0.4$  indicates poor agreement, whilst a value above 0.4 is indicative of good agreement”. Araujo et al. (2005) used the same classification limit of 0.4 to distinguish between “good” and “poor” agreement.

The classification of sample stations into benthos clusters was conducted employing the NAG-routines G03ECF and G03EFF (Numerical Algorithms Group Ltd., Oxford, UK). TWINSPAN was downloaded from <http://cc.oulu.fi/~jarioksa/softhelp/ceprog.html>. Multinomial logistic regression was carried out with the R-routine “multinom” (R-library “nnet”).

## Sample applications and groups of predictor variables

The model building procedure as described above was employed in the following sample applications:

- A1: The dataset was subdivided by (unstratified) random selection into a training set and a test set of equal size. Such a random split into two groups is a “classic” method (“2-fold cross-validation”) to test the predictive performance of a model. Elith and Leathwick (2009) characterized such a prediction as “model-based interpolation to unsampled sites”.
- A2: A spatial “gap” was assumed in the sampled benthos data. Logistic regression was used to fill this gap. The chosen test set (the gap) was a rectangle including 60 sample stations, covering the Steenack tidal flat as shown in Fig 6.
- A3: Application A3 was, in principle, the same as application A2. With 149 sample stations the test set had now a similar size as the training set. A separation line divided the tidal basin into a western test area and an eastern training area. The separation line ran across the Steenack tidal flat, thus separating the two major drainage areas of Hörnum tidal basin (Fig. 8).

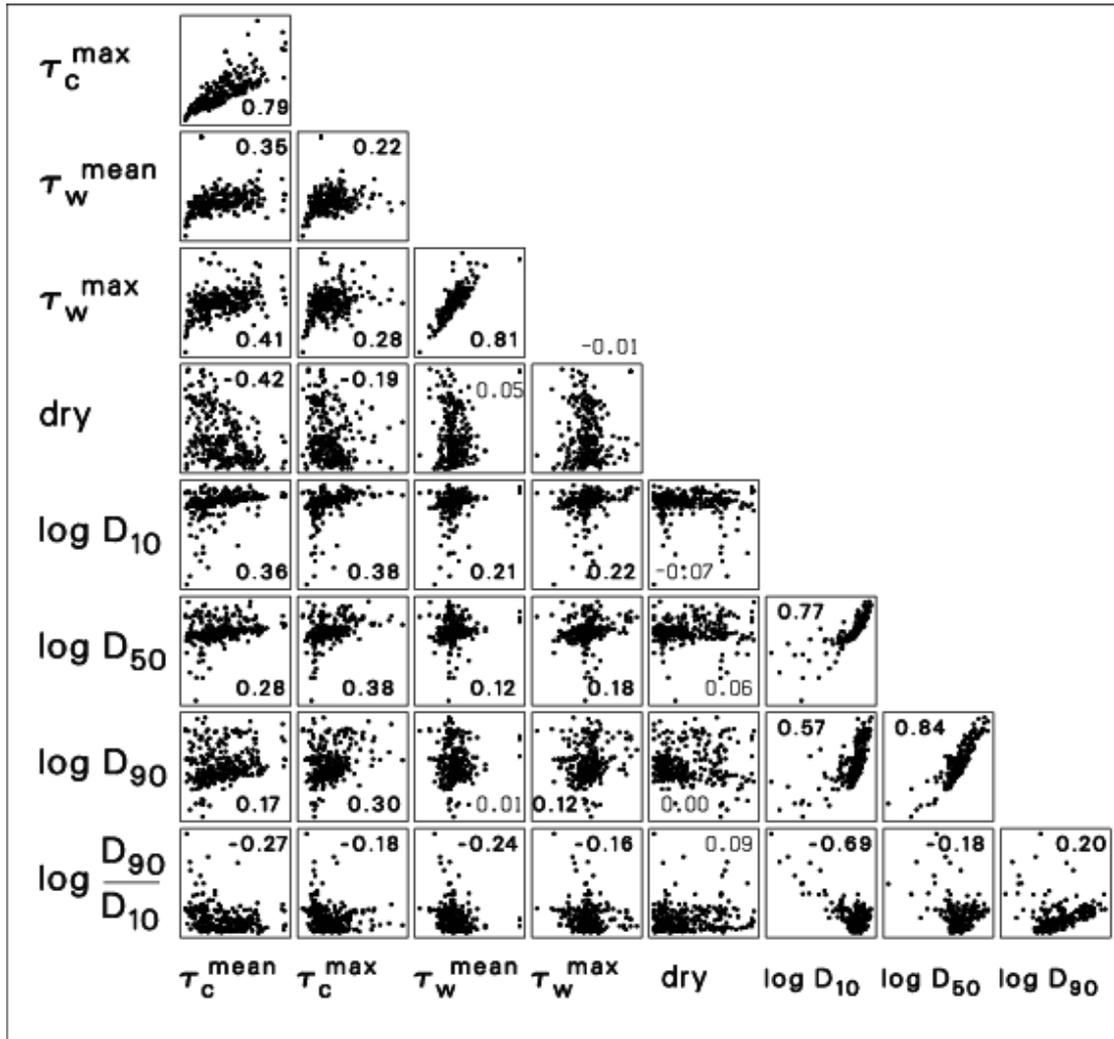
Three groups of predictor variables were applied to each of the three sample applications:

- H: Hydrodynamic variables: dry,  $\tau_c^{\text{mean}}$ ,  $\tau_c^{\text{max}}$ ,  $\tau_w^{\text{mean}}$ ,  $\tau_w^{\text{max}}$
- S: Logarithms of the sediment variables:  $D_{10}$ ,  $D_{50}$ ,  $D_{90}$ ,  $D_{90}/D_{10}$
- H+S: Combination of the five hydrodynamic variables H and the four sediment variables S.

## Results

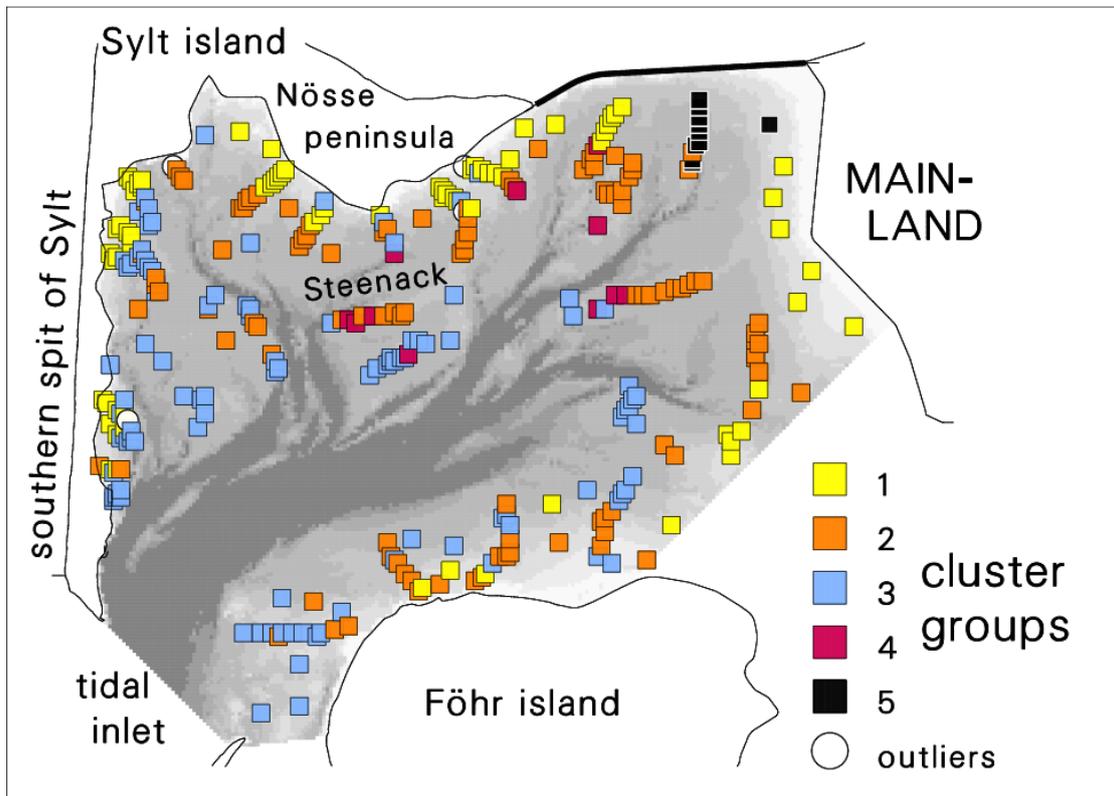
### Representation of the physical environment

Taking into account the data of the entire tidal flats and of the summer months, spatial average  $\pm$  standard deviation of  $\tau_c^{\text{mean}}$  was  $0.071 \pm 0.047$  Newton/m<sup>2</sup>, corresponding (for a water depth of 0.5 m) to a current velocity of  $\approx 0.19 \pm 0.07$  m/s. The high standard deviation reflected the high spatial variability of  $\tau_c^{\text{mean}}$  as shown in Fig. 2. Compared to the spatial average of  $\tau_c^{\text{mean}}$ , the spatial averages of  $\tau_c^{\text{max}}$ ,  $\tau_w^{\text{mean}}$  and  $\tau_w^{\text{max}}$  were higher by factors 6, 4 and 21, respectively. Sediment grain-size sorting  $D_{90}/D_{10}$  was mostly between 2 and 3 which means that the sediment was almost “well sorted” (Soulsby 1997) and which in turn is indicative of a dynamic hydraulic regime with mobile sediments.



**Fig. 4** Scatter plot matrix showing the relations among the nine used environmental variables at 308 sample stations, including the linear correlation coefficient in each scatter plot. The correlation coefficients presented in bold characters are significantly ( $p < 0.05$ ) different from zero

The mutual relations among the nine abiotic variables are shown by pairwise scatter plots in Fig. 4, including linear correlation coefficients  $r$ . The highest  $r$ -value was 0.84, calculated between  $\log D_{50}$  and  $\log D_{90}$ . This  $r$ -value was still below the rule-of-thumb limit of  $r \approx 0.95$  suggested by Clarke and Ainsworth (1993) for the reduction of datasets. A permutation method as, for instance, described in Clarke and Warwick (1994) showed that most correlation coefficients given in Fig. 4 were significantly different from zero (error probability  $p < 0.05$ ). Those correlation coefficients are presented in bold characters. A permutation method was used because, with regard to possible deviations from normality in the data, it is more robust than “classical” parametric statistical tests.



**Fig. 5** Distribution of “observed” cluster memberships in the Hörnum tidal basin. The gray shading indicates the bathymetry

### Classification of the benthic community

A hierarchical cluster analysis with group-average linking resulted in a division into five clusters which existed at a Bray-Curtis dissimilarity level of 54 %. At this level there were three dominant and two small clusters. The clustering process was stopped at this level to prevent a further fusing of two big clusters. The spatial distribution of the five cluster groups is shown in Fig. 5. There were four sample stations (outliers) which were not grouped into one of the five clusters. The set of sample stations used in the model thus consisted of  $308 - 4 = 304$  stations. An intertidal zoning of clusters 1, 2 and 3 can be recognized (see also Reise et al. 2008). The characteristics of the cluster groups are summarized in Table 1. Characteristic species of each cluster are indicated by bold presence numbers. A description of cluster groups is done in the next section.

### Community habitat preferences

The averages of the environmental variables in the five benthic community cluster groups are given in the lower part of Table 1. The intertidal zoning of clusters 1, 2, and 3 is reflected in increasing

hydrodynamic impact and increasing sediment grain-sizes across all three zones.

The cluster 1 zone featured nearshore *Zostera* meadows with high biodiversity. The low hydrodynamic impact was favorable to the development of seagrass beds (Schanz and Asmus 2003). Rather unusual, *Zostera* spp. were associated with the “sand-follower” *A. marina*. The reason for that was the comparatively high sand fraction in the silt-sandy sediment of the cluster 1 zone in the western part of the basin. The sand fraction was still increased by deposition of coarser airborne sand.

The transitional spatial position of cluster 2 (between clusters 1 and 3) was reflected in medium cluster averages for nearly each environmental variable in Table 1. The fine sand tidal flats of the cluster 2 zone were characterized by a high presence of *A. marina* and *C. edule* and high deposition of floating macroalgae. Moreover, the cluster 2 zone was a pioneer-area for *Zostera* spp. Highest occurrences of green macroalgae *Enteromorpha* spp. and *U. lactuca* were found in the Cluster 2 zone, often in front of *Zostera* meadows which can serve as barrier for the landward transport of macroalgae. The high occurrence of green macroalgae in the cluster 2 zone was favored by two “mechanisms”. First the deposition and adherence of macroalgae to the bottom was enabled by burrows of *A. marina*. Second the “medium” currents in the cluster 2 zone were strong enough to transport macroalgae into the zone, but not strong enough to erode (together with wave action) stuck macroalgae from the ground.

The cluster 3 zone was closest to the low water line; it was characterized by high hydrodynamic impact. The zone consisted of exposed sandy habitats with high occurrences of *A. marina* only.

Within the cluster 2 and cluster 3 zones, the cluster 4 community occurred as low-lying isles within sandy habitats. The cluster 4 isles were characterized by small emersion time and higher hydrodynamic exposure. Such conditions are preferred by suspension feeders as *L. conchilega*, the key species of cluster 4. The reason for the higher occurrence of *Enteromorpha* spp. in cluster 4 was that algae-fascicles were retained by *Lanice* tubes.

The sample stations of cluster 5 were mainly positioned near the low water line of a gully in the NE corner of the basin (Fig. 5). At this site, sediment grain-size, current velocity, wave exposure and emersion time were all low, which was favorable to diatoms, the characteristic

organisms of cluster 5, as well as favorable to (mostly juvenile) cockles. Both of these species prefer a fine-grained sediment regime because of its higher stability, nutrient richness (diatoms) and nutrient content (cockles). The low hydrodynamics in the cluster 5 area favored the deposition of *U. lactuca*.

The blue mussel, *M. edulis*, showed no significant peak of occurrence across the five clusters. This species was mainly found in form of individuals or clumps, broken free from banks and driven over the entire intertidal flats.

### Selected predictor variables

For three sample applications and three groups of environmental predictors, Table 2 shows the performance of  $3 \cdot 3 = 9$  habitat suitability models. Altogether, 18 covariates were tested in each model building process: nine environmental variables both in linear and squared terms. Regarding the group of the five hydrodynamic predictors H, the “best predicting sets of covariates” in Table 2 contain “dry”,  $\tau_c^{\text{mean}}$  and  $\tau_w^{\text{max}}$ , while  $\tau_c^{\text{max}}$  and  $\tau_w^{\text{mean}}$  were never selected. Concerning the group of the four sediment predictors S, sediment sorting was the only one which was never selected. From the combined group of hydrodynamic and sediment predictors, H+S,  $\tau_c^{\text{mean}}$  and the emersion time “dry” were selected in each of the three sample application of Table 2, as well as at least one sediment predictor. It is interesting to note that currents representing regular tidal conditions ( $\tau_c^{\text{mean}}$ ) were more important for benthic communities than currents during stormy conditions, represented by  $\tau_c^{\text{max}}$ . For waves the situation was vice versa: the wave impact during storms ( $\tau_w^{\text{max}}$ ) was more important than the average wave climate, represented by  $\tau_w^{\text{mean}}$ .

### Spatial prediction of the benthic community structure

A full coverage distribution (spatial prediction) of cluster memberships (i.e. of the benthic community structure) was calculated with multinomial logistic regression which was solely based on the training set. Fig. 6 shows the spatial prediction for application A2, using the group of combined predictors H+S. The spatial prediction shows a clear intertidal zoning of clusters 1, 2 and 3. Cluster 4 was predicted in a very small number of grid cells only. As the training set of application A2 included 244 of the complete set of 304 sample stations, the predicted benthic community structure in Fig. 6 is similar to a map which is based on the complete set of stations. With  $\kappa_{\text{TEST}} = 0.44$ , the agreement between

<b>Group of predictors</b> →	H hydrodynamics	S sediment	H+S hydro + sedim.
------------------------------	--------------------	---------------	-----------------------

**Application A1:** spatial interpolation

Best predicting set of covariates, based only on training set →	dry, $\tau_c^{\text{mean}}$ , $(\tau_w^{\text{max}})^2$	logD <sub>10</sub>	dry, $\tau_c^{\text{mean}}$ , $(\tau_w^{\text{max}})^2$ , logD <sub>10</sub>
CCR for training set	0.63 ± 0.030	0.50 ± 0.049	0.68 ± 0.031
CCR for test set	0.61 ± 0.031	0.49 ± 0.058	0.64 ± 0.031
Cohen's kappa for training set	0.47 ± 0.043	0.26 ± 0.079	0.54 ± 0.046
Cohen's kappa for test set, $\kappa_{\text{TEST}}$	0.44 ± 0.044	0.25 ± 0.077	0.49 ± 0.044

**Application A2:** extrapolation to a spatial gap in the benthos data

Best predicting set of covariates, based only on training set →	dry, $\tau_c^{\text{mean}}$ , $(\tau_w^{\text{max}})^2$	logD <sub>10</sub> , logD <sub>50</sub>	dry, $\tau_c^{\text{mean}}$ , $(\tau_w^{\text{max}})^2$ , logD <sub>50</sub> , $(\log D_{50})^2$
CCR for training set	0.65 ± 0.018	0.53 ± 0.028	0.71 ± 0.017
CCR for test set	0.58 ± 0.034	0.42 ± 0.051	0.62 ± 0.054
Cohen's kappa for training set	0.50 ± 0.026	0.29 ± 0.043	0.58 ± 0.025
Cohen's kappa for test set, $\kappa_{\text{TEST}}$	0.40 ± 0.051	0.14 ± 0.077	0.44 ± 0.077

**Application A3:** extrapolation to a new drainage area

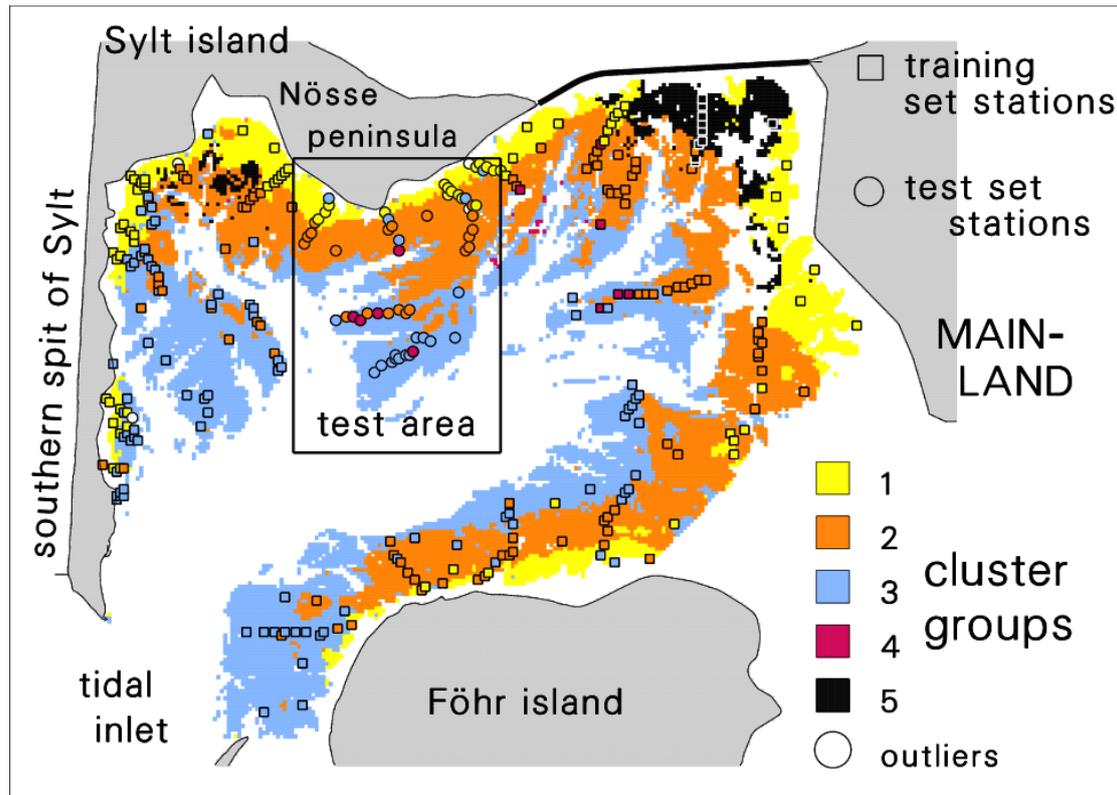
Best predicting set of covariates, based only on training set →	dry, $\tau_c^{\text{mean}}$	logD <sub>10</sub> , logD <sub>50</sub> , logD <sub>90</sub>	dry, $\tau_c^{\text{mean}}$ , $\tau_c^{\text{mean}}$ , logD <sub>50</sub>
CCR for training set	0.67 ± 0.032	0.62 ± 0.028	0.73 ± 0.021
CCR for test set	0.53 ± 0.041	0.51 ± 0.043	0.51 ± 0.038
Cohen's kappa for training set	0.52 ± 0.047	0.42 ± 0.049	0.61 ± 0.030
Cohen's kappa for test set, $\kappa_{\text{TEST}}$	0.34 ± 0.048	0.25 ± 0.053	0.26 ± 0.045

**Table 2** Results of applying multinomial logistic regression to three sample applications and three groups of environmental predictor variables. The results of application A1 are averages and standard deviations of 200 different (random) subdivisions of the full dataset. The standard deviations of applications A2 and A3 were calculated from 200 bootstrap replications of the training set. CCR is the “correct classification rate”.

“observed” and predicted cluster memberships within the rectangle of Fig. 6 was “good”.

As a counter-example, Fig. 7 shows the predicted pattern of benthic communities with a poor model performance: for the combination “application A2 / predictor group S”,  $\kappa_{\text{TEST}}$  was only 0.14. There were two reasons for the poor agreement between observed and predicted cluster memberships within the test area of Fig. 7. The first was the

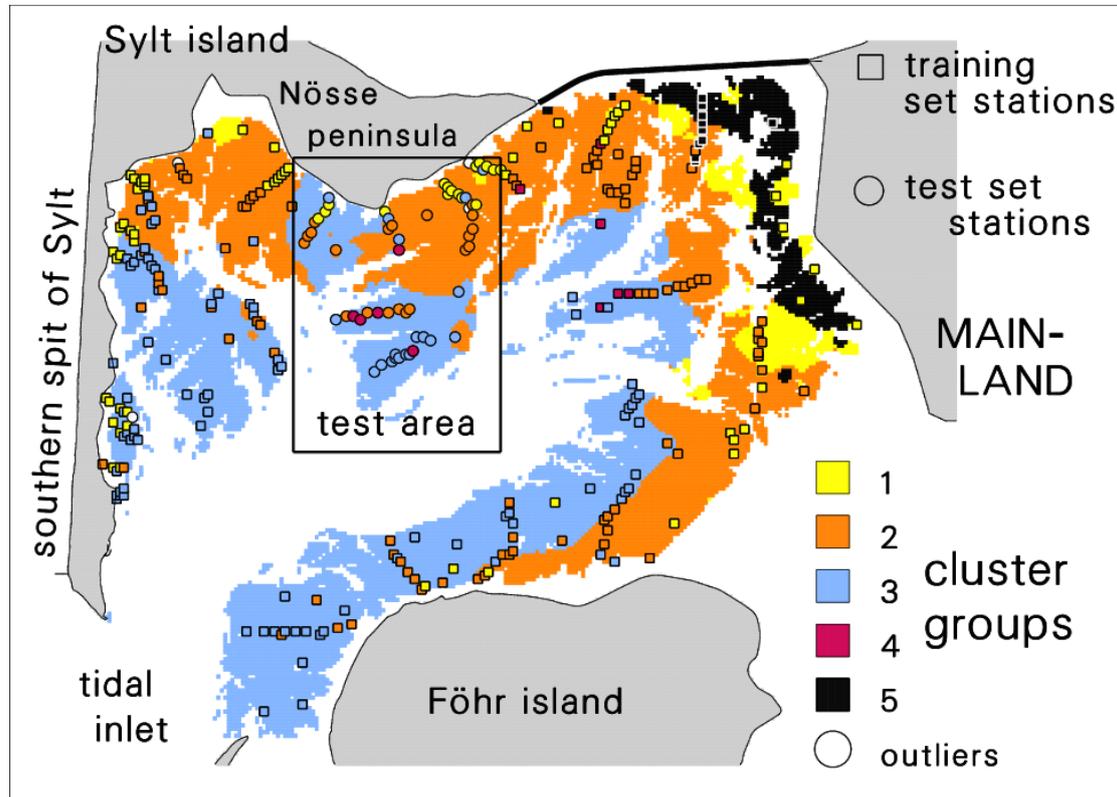
generally low correlation between the spatial distribution of “observed” cluster groups (Fig. 5) and the basin-wide sediment grain-size distribution (Fig. 1). This low correlation is indicated in Table 2 by the poor Cohen’s kappas obtained for the *training* sets of combinations “A1/S” or “A2/S”. The second reason was “bad luck”: The test set of application A2 (the stations within the rectangle) contained a large percentage of stations with “observed” cluster 1 memberships. Cluster 1, however, was generally predicted to be rare when using predictor group “sediments only”.



**Fig. 6** Cluster memberships on the tidal flats of the Hörnum tidal basin, sample application A2 using predictor group “H+S”, i.e. both hydrodynamic and sediment predictor variables. The area within the rectangle is the test area (station symbols as circles), the rest of the basin is the training area (station symbols as squares). The cluster memberships at the sample stations are “observed” data, they are identical to those given in Fig. 5. A cluster membership was not predicted (a) in those parts which are continuously covered with water and (b) in some (shallow) areas near the shore. In those near shore areas the set of environmental variables was incomplete because the observed sediment data was not dense enough for an interpolation

Fig. 8 shows the spatial prediction for application A3, using the group of combined predictors H+S. When comparing the “observed” cluster memberships at the sample stations with the full coverage map of predicted cluster memberships, the agreement was substantially higher in the training area (east of the separation line) than in the test area. In Table 2 this difference is reflected in a drastic  $\kappa$ -decrease from 0.61 for the

training set to 0.26 for the test set. As an example, along the southern spit of Sylt, cluster 1 was frequently observed, but not predicted. Instead, cluster 2 was predicted there, but rarely observed.



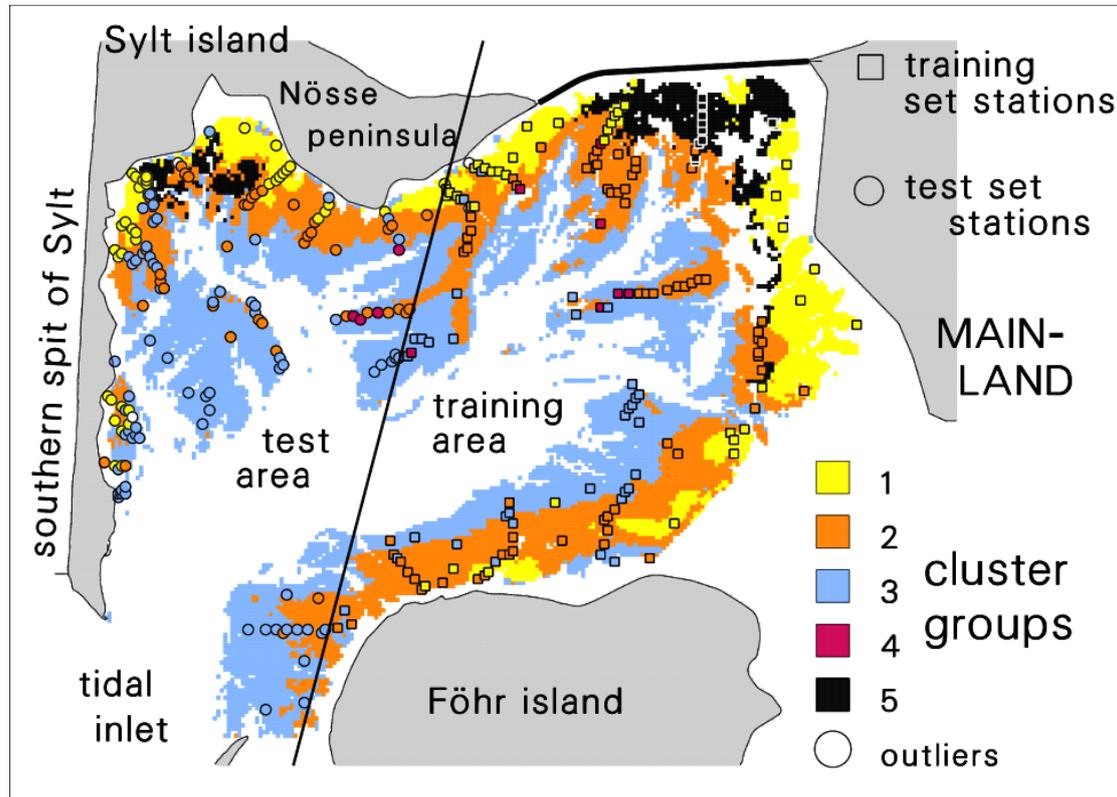
**Fig. 7** Cluster memberships on the tidal flats of the Hörnum tidal basin, sample application A2 using predictor group “S”, i.e. sediment predictor variables only. This is the spatial prediction with the lowest test set accuracy ( $\kappa_{\text{TEST}} = 0.14$ ) in Table 2. See the caption of Fig. 6 for more information

### Model evaluation

The most informative model performance indicator in Table 2 is Cohen’s kappa for the test set,  $\kappa_{\text{TEST}}$ . The prediction success was “good” ( $\kappa_{\text{TEST}} \geq 0.40$ ) for applications A1 and A2 when using the predictor groups H and H+S. However, the prediction success was consistently poor ( $\kappa_{\text{TEST}} < 0.4$ ), (a) for application A3 and (b) when only sediment variables were used as predictors (group S).

Concerning the model performance across the three sample applications A1 to A3, the differences {“ $\kappa$  of training set” minus  $\kappa_{\text{TEST}}$ } uniformly increased across A1 to A3. This increase was expected, as it corresponded to an increase in spatial distances between training sets and test sets

across A1 to A3. The prediction efficiency  $\kappa_{\text{TEST}}$  itself showed a uniform behavior across A1 to A3 when hydrodynamic predictors were applied, while the behavior was irregular when sediment predictors were used. Most surprising, when predictors H+S were applied, there was an “unexpected” poor  $\kappa_{\text{TEST}}$  ( $= 0.26$ ) for application A3 (unexpected, because



**Fig. 8** Cluster memberships on the tidal flats of the Hörnum tidal basin, sample application A3 using predictor group “H+S” (both hydrodynamic and sediment predictors). The solid line subdivides the basin into a western test area and an eastern training area. See the caption of Fig. 6 for more information

with the combined group of predictors H+S, the model was expected to perform better than with the single predictor group H). This poor  $\kappa_{\text{TEST}} = 0.26$  was caused by inconsistent sediment grain-size ranges in training set and test set. A comparison of median grain-sizes  $D_{50}$  (see Fig. 1) showed that at 41 % of the A3 test set stations (west of the separation line, Fig. 8),  $D_{50}$  was above  $200 \mu\text{m}$ , while this was true for only 4 % of the training set stations. When predicting the spatial distribution of cluster groups in the test area, the logistic regression’s response curve of cluster memberships was extrapolated into the test area beyond the realized range of grain-sizes within the training set. The existence of  $\log D_{50}$  in the “best predicting set of covariates” spoiled the prediction success for the combination “A3/H+S”. Inconsistent ranges of environmental predictors

are a common hazard in spatial extrapolation, see Zimmermann and Kienast (1999), Randin et al. (2006) or Elith and Leathwick (2009).

Comparing the predictive efficiencies across the three different predictor groups H, S and H+S, Table 2 shows that (a) hydrodynamic predictors performed much better than sediment predictors, (b) compared to the performances of “hydrodynamic predictors only” or “sediment predictors only”, the combined group of predictors (H+S) performed best in applications A1 and A2, but not in A3. The reason for this “unexpected” result was explained in the paragraph above.

## Discussion

The most frequently applied method to “validate” the prediction efficiency of a statistical model is a cross-validation test, i.e. the random selection of a test set from the entire dataset as in application A1. For application A1 using predictors group H+S,  $\kappa_{\text{TEST}}$  was 0.49, see Table 2. The  $\kappa_{\text{TEST}}$  of a cross-validation test assesses “how well the model predicts over ... the average distance between samples” (Thrush et al. 2005). Strictly speaking, a cross-validation test is not a prediction test as there is no “extrapolation to new conditions” (Elith and Leathwick 2009). The result of a cross-validation test (in this study) is rather a yardstick for the degree of correlation between the applied environmental variables and the benthic community structure. The question is whether this degree of correlation was also found in other coastal regions of the southern North Sea. According to Zimmermann and Kienast (1999), this comparison should be made using a model which also used benthic communities as a response variable, as statistical model fitting is easier for communities than for individual species due to the communities’ more uniform response to environmental gradients. From 3-fold cross-validation tests, Degraer et al. (2008) obtained a  $\kappa_{\text{TEST}}$ -value of about 0.7 (4 macrobenthic communities, Belgian continental shelf, 364 samples). Compared to this model performance, the present study’s  $\kappa_{\text{TEST}}$  of 0.49 for the combination “A1/H+S” was substantially smaller. It may be taken into account, however, that Degraer et al. (2008) excluded 409 “inconsistent” samples from their starting dataset of 773 samples, which possibly improved their model performance. This study used (apart from four outliers) the original starting dataset.

Using hydrodynamic predictors (group H) only, the values for  $\kappa_{\text{TEST}}$  in Table 2 were  $\geq 0.4$  and thus assessed as “good” for combinations “A1/H” and “A2/H”. For the combination “A3/H”,  $\kappa_{\text{TEST}}$  dropped below  $\kappa_{\text{TEST}} = 0.4$ , the transition line between good and poor agreement. The model

performance ( $\kappa_{\text{TEST}}$ ) uniformly decreased across the three sample applications A1 to A3 corresponding to the increasing spatial distance between training set and test set across A1 to A3. The results suggest (a) that the relationships between benthic communities and hydrodynamics were strong enough to justify the application of a habitat suitability model in the Wadden Sea as long as training and prediction areas are closely associated in space and (b) that the transferability between training and prediction area has to be carefully assessed if the two areas are different tidal drainage areas.

### Hydrodynamic versus sediment predictors

The most interesting result shown in Table 2 is the low correlation between sediment predictors and the benthic community structure: it is significantly below the correlation between hydrodynamic predictors and benthic communities. The low performance of sediment predictors is surprising as in most studies the use of sediment parameters is considered the most reliable way to a prediction of benthic communities or benthic species in marine waters. In Thrush et al. (2005) the sediment mud content was used as the only environmental predictor for analyzing the effect of different spatial scales on environment-species relationships in New Zealand estuaries. In Compton et al. (2009) the median grain-size was the only predictor used to investigate the repeatability of bivalves across six European tidal flats. Degraer et al. (2008) selected two sediment predictors, median grain-size and mud content, for mapping the benthic community structure in the Belgian part of the North Sea.

Having reviewed numerous studies about the relationships between marine infaunal species and sediment parameters, Snelgrove and Butman (1994) argued that sediment grain-size was used as a substitute for the true causative (ecologically relevant) factors influencing infaunal species distributions, namely the near-bed flow conditions or rather the hydrodynamic regime. They further argued that “to act as a substitute” was feasible for sediment predictors because they were “to a large extent the direct result of near-bed flow conditions”, and thus correlated with the hydrodynamic regime.

In the Hörnum tidal basin, sediment parameters only moderately correlated with the hydrodynamic variables (Fig. 4). This affords the opportunity to separate between the influences of hydrodynamic and sediment predictors on benthic communities. The result is shown in Table 2: The influence of hydrodynamic variables on the benthic community structure (in the Hörnum tidal basin during the summer months of 2001-

2003) was significantly higher than the influence of sediment parameters. This finding supports the view of Snelgrove and Butman (1994). Low correlation between hydrodynamics (model calculated ebb and flood current velocities) and sediment parameters (mud content and median grain-size) were also observed in Ysebaert and Herman (2002). Hydrodynamic and sediment predictors were, however, not tested separately in a logistic model in Ysebaert et al. (2002).

#### Ecologically relevant predictor variables

The description of relationships between benthic communities and the physical environment in section “Community habitat preferences” suggests that hydrodynamic impact was an ecologically relevant driver of benthic ecosystems in the Hörnum tidal basin. The same, but to a less degree, was true for sediment, e.g. with regard to clusters 1 (*A. marina*) and 5 (diatoms). It must be considered that sediment conditions (apart from being influenced by hydrodynamics) may also be influenced by bioengineering activities of benthic species, i.e. by the “response variable” itself. For example, the lugworm *A. marina* maintains mobile permeable sand (Volkenborn et al. 2007) and seagrass accumulates fine sediment particles. This means that, in contrast to hydrodynamics, sediment is not always a *direct* driver of Wadden Sea ecosystems.

This study focuses on the prediction of the spatial distribution of benthic communities. In contrast to an explanatory model, in a predictive model the environmental variables must not necessarily be ecologically relevant (Fielding 2007). According to Fielding (2007), in a predictive model “only the model’s accuracy is important ... as long as it is robust”. Elith and Leathwick (2009), however, argue that if predictive models use ecologically irrelevant variables, “extrapolation in space or time will be particularly error-prone”. In any case the identification of hydrodynamics as being of primary ecological relevance is compatible with the finding (Table 2) that hydrodynamics were more efficient than sediment parameters in predicting benthic communities.

#### Temporal prediction of ecosystem changes

Concerning temporal prediction of ecosystems, there are increasing demands for reliable and quantitative models, see Clark et al. (2001). There are numerous studies, mostly in terrestrial environments, which predict the potential impact of climate change on ecosystems. A review on so-called bioclimate envelope models was given by Pearson and Dawson (2003). Harley et al (2006) gave a survey on the future

implications of climate change for marine ecosystems, but predictive mathematical models were mentioned only briefly. A “habitat requirement model” was developed by Fourquaran et al. (2003) for predicting changes in Florida’s seagrass beds as a result of future changes in water management practices. For the German Wadden Sea, trends for long-term changes in the ecosystem (over several decades) were outlined in Reise and van Beusekom (2008). The environmental drivers for those trends in the Wadden Sea were temperature rise, sea level rise (both related to climate change), a declining nutrient supply and the invasion of alien species.

This study deals with the *spatial* prediction of the benthic community structure. Ysebaert and Herman (2002) showed that the variation of benthic macrofauna in the Schelde estuary was mainly spatially structured, only a small part of the species variation was temporally (between-years) structured. The results of Ysebaert and Herman (2002) were based on a survey period of 7 years. Their results suggest that a temporal prediction of the benthic community structure (using environmental variables as predictor variables) should at least have the same potential as a spatial prediction. Consequently, with regard to the results of this study, a temporal prediction should have a good chance to achieve an accuracy of  $\kappa_{\text{TEST}} > 0.4$ . However, it has to be assumed that such an accuracy can only be “guaranteed” for short time periods (i.e. a few years), considering the ongoing invasion of alien species in the Wadden Sea (Reise 2005, Reise et al. 2006). The dispersal of species is one (of three) “fundamental limitations to the predictive capacity” of long-term predictions addressed in Pearson and Dawson (2003). The other two limitations are “biotic interactions” and “evolutionary change”.

A scenario for a short-term prediction would be that a coastal management scheme changes the hydrodynamic regime which in turn modifies benthic habitats. The German Wadden Sea, however, is on the list of German National Parks (since 1985-1990) and was added (in 2009-2011) to the list of UNESCO World Heritage sites. New engineering constructions (e.g. to build a dam to prevent the erosion of an island’s geological basement) in the Wadden Sea are thus extremely unlikely. However, when deepening the fairway of the Elbe or Weser estuaries, the hydraulic regime is changed which in turn may change benthic habitats on adjacent tidal flats (Esser et al. 2002). Other “coastal mitigation and adaption technologies” are conceivable for German estuaries which provide management strategies against coastal erosion and flooding and which, in parallel, aim at “healthy coastal habitats” (<http://www.theseusproject.eu>).

## Environmental variables for temporal prediction

The classical parameters used for the spatial prediction of the benthic community structure in a coastal environment are parameters measured in the field (Snelgrove and Butman 1994). In particular, sediment parameters proved to be successful environmental predictor variables in the relevant literature, see section “Hydrodynamic versus sediment predictors”. One further example is the study of Willems et al. (2008): their habitat models selected only sediment parameters (median grain-size, % mud, % coarse fraction) as environmental predictor variables. Yet when a *temporal* prediction is concerned, it is not possible to apply data collected in the field as predictor variables, simply because such data are not available for a future scenario. Only those environmental variables which are themselves forecast, e.g. by a numerical model, can be used for a temporal benthos prediction.

In a typical real application the target area of a benthos prediction is defined by a coastal authority who plans an engineering construction or a deepening of a navigation channel. In many cases the planning includes an environmental impact assessment. The planning will usually be supported by a prediction of hydrodynamics with an appropriate model (e.g. Carniello et al. 2005). In a subsequent step, based on the hydrodynamic model results, a bottom sediment distribution may be predicted (a) using a morphodynamic model (e.g. de Swart and Zimmerman 2009) or (b) using a statistical approach (e.g. Zwarts 2004, Escobar and Mayerle 2007). In any case the predicted hydrodynamic data are of first order accuracy while the predicted sediment data are of second order accuracy. Therefore it appears reasonable to use hydrodynamic data only for a temporal prediction of benthic community structure. As shown in Table 2 for sample application A1 and A2,  $\kappa_{\text{TEST}}$  was (a) 0.49 and 0.44 when hydrodynamic *and* sediment predictors were applied, and (b) 0.44 and 0.40 when only hydrodynamic predictors were applied. This shows that hydrodynamic predictors performed only by 10 % worse than the combined set of hydrodynamic and sediment predictors.

Another argument for the exclusive use of modeled hydrodynamic predictor variables (now for a spatial prediction) is the consideration of costs. If appropriate measured (sediment) data are available, they will of course be included in the list of potential environmental predictor variables. However, if measured field data are lacking, it is worth considering that the acquisition and the lab-analysis of new field samples may be more costly than the setting up of a numerical model. One could

thus make the case that completely relying on the results of numerical modeling might be preferable to the acquisition of costly field data, despite the hazard of a less accurate prediction. The consideration of costs is a “pragmatic issue” discussed in Fielding (2007).

**Acknowledgements** The research was carried out within the framework of the BELAWATT project (BELAWATT = German acronym of “The hydrodynamic impact on Wadden Sea areas”). This KFKI-project was funded by the Bundesministerium für Bildung und Forschung (BMBF) under grant No. 03KIS038. The bathymetric data for the model topography were provided by the Bundesamt für Seeschifffahrt und Hydrographie (BSH) and by the Landesbetrieb für Küstenschutz, Nationalpark und Meeresschutz Schleswig-Holstein (LKN-SH), Husum.

## References

- Araujo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species-climate impact models under climate change. *Glob Change Biol* 11: 1504-1513
- Asmus H, Asmus R (1998) The role of macrobenthic communities for sediment-water material exchange in the Sylt-Rømø tidal basin. *Senckenbergiana maritima* 29 (1/6): 111-119
- Beukema JJ (1976) Biomass and species richness of the macro-benthic animals living on the tidal flats of the Dutch Wadden Sea. *Neth J Sea Res* 10(2): 236-261
- Bos AR, Dankers N, Groeneweg AH, Hermus DCR, Jager Z, de Jong DJ, Smit T, de Vlas J, van Wieringen M, van Katwijk MM (2005) Eelgrass (*Zostera marina* L.) in the western Wadden Sea: monitoring, habitat suitability model, transplantations and communication. In: Herrier J-L, Mees J, Salman A, Seys J, van Nieuwenhuysse H, Dobbelaere I (eds) Proceedings “Dunes and Estuaries 2005”, Int Conf on Nature Restoration Practices in European Coastal Habitats, Koksijde, Belgium, VLIZ Special Publication 19, pp 95-109
- Brinkman AG, Dankers N, van Stralen M (2002) An analysis of mussel bed habitats in the Dutch Wadden Sea. *Helgol Mar Res* 56: 59-75
- Carniello L, Defina A, Fagherazzi S, D’Alpaos L (2005) A combined wind wave – tidal model for the Venice lagoon, Italy. *J Geophys Res* 110, F04007, doi:10.1029/2004JF000232
- Casulli V, Stelling GS (1998) Numerical simulations of 3D quasi-hydrostatic, free-surface flows. *J Hydr Eng* 124 (4): 678-686
- Clark JS et al. (2001) Ecological forecasts: An emerging imperative. *Science* 293: 657-660

- Clarke KR, Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. *Mar Ecol Prog Ser* 92: 205-219
- Clarke KR, Warwick RM (1994) *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*. Plymouth Marine Laboratory, UK
- Compton TJ, Troost TA, Drent J, Kraan C, Bocher P, Leyrer J, Dekinga A, Piersma T (2009) Repeatable sediment associations of burrowing bivalves across six European tidal flat systems. *Mar Ecol Prog Ser* 382: 87-98
- Damm-Böcker S, Kaiser R, Niemeyer HD (1993) Determination of benthic Wadden Sea habitats by hydrodynamics. In: *Proc ICC-Kiel'92, Germany*, pp 430-441
- Degraer S, Verfaillie E, Willems W, Adriaens E, Vincx M, van Lancker V (2008) Habitat suitability modelling as a mapping tool for macrobenthic communities: An example from the Belgian part of the North Sea. *Cont Shelf Res* 28(3): 369-379
- De Swart HE, Zimmerman JTF (2009) Morphodynamics of tidal inlet systems. *Annu Rev Fluid Mech* 41: 203-229
- Eppel D, Kapitza H, Onken R, Pleskachevsky A, Puls W, Riethmüller R, Vaessen B (2006) *Watthydrodynamik: Die hydrodynamische Belastung von Wattgebieten*. Report 2006/8, GKSS-Forschungszentrum Geesthacht, Germany
- Elith J, Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40: 677-679
- Escobar CA, Mayerle R (2007) Procedures for improving the prediction of equilibrium grain sizes, bed forms and roughness in tidally-dominated areas. In: McKee Smith J (ed) *Proc 30<sup>th</sup> Int Conf Coastal Eng 2006, San Diego, USA*, pp 3092-3104
- Esser B et al. (2002) *Ökologische Aspekte bei Fahrrinnenanpassungen*. Jahrbuch der Hafentechnischen Gesellschaft. Schiffahrts-Verlag Hansa, Hamburg. Band 53: 98-122
- Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. *J Appl Ecol* 43: 393-404
- Fielding AH (2007) *Cluster and classification techniques for the biosciences*. Cambridge University Press, UK
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conservation* 24(1): 38-49
- Figge K (1981) *Begleitheft zur Karte der Sedimentverteilung in der Deutschen Bucht (Karte Nr. 2900)*. Deutsches Hydrographisches Institut, Hamburg, Germany

- Flach EC, Beukema JJ (1994) Density-governing mechanisms in populations of the lugworm *Arenicola marina* on tidal flats. *Mar Ecol Prog Ser* 115: 139-149
- Fourquaran JW, Boyer JN, Durako MJ, Hefty LN, Peterson BJ (2003) Forecasting responses of seagrass distributions to changing water quality using monitoring data. *Ecol Appl* 13(2): 474-489
- Friedrichs CT, Aubrey DG (1996) Uniform bottom shear stress and equilibrium hypsometry of intertidal flats. In: Pattiaratchi C (ed) *Mixing Processes in Estuaries and Coastal Seas*, *Coast Estuar Studies* 50, American Geophysical Union, pp 405-429
- Grüne J (2009) Safety analysis model and procedure for a coast protection master plan of North Sea coast of Schleswig-Holstein. In: McKee Smith J (ed) *Proc 31<sup>st</sup> Int Conf Coastal Eng 2008*, Hamburg, Germany, pp 2899-2911
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Mod* 135: 147-186
- Harley CDG, Hughes AR, Hultgren KM, Miner BG, Sorte CJB, Thornber CS, Rodriguez LF, Tomanek L, Williams SL (2006) The impacts of climate change in coastal marine systems. *Ecol Letters* 9: 228-241
- Hirschhäuser T, Zanke UCE (2001) Morphologische Langfristprognose für das System Tidebecken-Außensände am Beispiel Sylts und der Dithmarscher Bucht. *Die Küste* 64: 127-160
- Hirschhäuser T, Zanke U (2004) Langfristige Sedimentdynamik des Systems Tidebecken-Ebbdelta unter besonderer Berücksichtigung von verändertem Seegang und Wasserständen. *Die Küste* 68: 201-248
- Hosmer DW, Lemeshow S (2000) *Applied Logistic Regression*. John Wiley & Sons, New York, USA, 2<sup>nd</sup> edition
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174
- Moghimi S, Gayer G, Günther H, Shafieefar M (2005) Application of third generation shallow water wave models in a tidal environment. *Ocean Dyn* 55: 10-27
- Niemeyer HD, Michaelis H (1997) Fluktuationen von Makrozoobenthospopulationen in Wattgebieten infolge variierender hydrodynamischer Randbedingungen. ELAWAT Teilprojekt B8, Abschlußbericht, Forschungsstelle Küste, Norderney, Germany
- Niemeyer HD (1979) Untersuchungen zum Seegangsklima im Bereich der Ostfriesischen Inseln und Küste. *Die Küste* 34: 53-70
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob Ecol & Biogeogr* 12: 361-371

- Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A (2006) Are niche-based species distribution models transferable in space? *J Biogeogr* 33: 1689-1703
- Reise K (1985) Tidal flat ecology. An experimental approach to species interactions. Springer-Verlag, Berlin, Germany
- Reise K (1998) Coastal change in a tidal backbarrier basin of the northern Wadden Sea: are tidal flats fading away? *Senckenbergiana maritima* 29 (1/6): 121-127
- Reise K (2005) Coast of change: habitat loss and transformations in the Wadden Sea. *Helgol Mar Res* 59: 9-21
- Reise K, Simon M, Herre E (2001) Density-dependent recruitment after winter disturbance on tidal flats by the lugworm *Arenicola marina*. *Helgol Mar Res* 55: 161-165
- Reise K, Olenin S, Thieltges DW (2006) Are aliens threatening European aquatic coastal ecosystems? *Helgol Mar Res* 60: 77-83
- Reise K, van Beusekom JEE (2008) Interactive effects of global and regional change on a coastal ecosystem. *Helgol Mar Res* 62: 85-91
- Reise K, Herre E, Sturm M (2008) Mudflat biota since the 1930s: change beyond return? *Helgol Mar Res* 62: 13-22
- Ross J, Mittelstaedt E, Klein H, Berger R, Ricklefs K (1998) Der Wasseraustausch im Tidebecken Hörnum-Tief. Bericht Nr. 16 des Bundesamtes für Seeschifffahrt und Hydrographie, Hamburg, Germany
- Schanz A, Asmus H (2003) Impact of hydrodynamics on development and morphology of intertidal seagrasses in the Wadden Sea. *Mar Ecol Prog Ser* 261: 123-134
- Snelgrove PVR, Butman CA (1994) Animal-sediment relationships revisited: cause versus effect. In: Ansell AD, Gibson RN, Barnes M (eds) *Oceanography and Marine Biology: an Annual Review*. Volume 32, UCL Press, London, UK, pp 111-177
- Soulsby R (1997) *Dynamics of marine sands*. Thomas Telford Publications, London, UK
- Thrush SF, Hewitt JE, Herman PMJ, Ysebaert T (2005) Multi-scale analysis of species-environment relationships. *Mar Ecol Prog Ser* 302: 13-26
- van Bernem K-H, Grotjahn M, Knüpling J, Krasemann HL, Müller A, Neugebohrn L, Patzig S, Ramm G, Riethmüller R, Sach G, Suchrow S (1994) Thematische Kartierung und Sensitivitätsraster im deutschen Wattenmeer, Juni 1987 – Juni 1993. Report 94/E/10, GKSS-Forschungszentrum Geesthacht, Germany
- van Bernem K-H, Doerffer R, Grohnert A, Heymann K, Kleeberg, U, Krasemann H, Reichert J, Reichert M, Schiller H (2007) Sensitivitätsraster Deutsche Nordseeküste II – Aktualisierung und

- Erstellung eines operationellen Modells zur Vorsorgeplanung bei der Ölbekämpfung. Report 2007/2, GKSS-Forschungszentrum Geesthacht, Germany
- van Katwijk MM, Hermus DCR, de Jong DJ, Asmus RM, de Jonge VN (2000) Habitat suitability of the Wadden Sea for restoration of *Zostera marina* beds. Helgol Mar Res 54: 117-128
- Volkenborn N, Hedtkamp SIC, van Beusekom JEE, Reise K (2007) Effects of bioturbation and bioirrigation by lugworms (*Arenicola marina*) on physical and chemical sediment properties and implications for intertidal habitat succession. Estuar Coast Shelf Sci 74: 331-343
- Willems W, Goethals P, van den Eynde D, van Hoey G, van Lancker V, Verfaillie E, Vincx M, Degraer S (2008) Where is the worm? Predictive modelling of the habitat preferences of the tube-building polychaete *Lanice conchilega*. Ecol Mod 212: 74-79
- Ysebaert T, Herman PMJ (2002) Spatial and temporal variation in benthic macrofauna and relationships with environmental variables in an estuarine, intertidal soft sediment environment. Mar Ecol Prog Ser 244: 105-124
- Ysebaert T, Meire P, Herman PMJ, Verbeek H, 2002. Macrobenthic species response surfaces along estuarine gradients: prediction by logistic regression. Mar Ecol Prog Ser 225: 79-95
- Zimmermann NE, Kienast F (1999) Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. J Veg Sci 10: 469-482
- Zwarts L (2004) Bodemgesteldheid en mechanische kokkelvisserij in de Waddenzee. Rapport RIZA/2004.028, Ministerie van Verkeer en Waterstaat, The Netherlands