



# A 'fuzzy clustering' approach to conceptual confusion: how to classify natural ecological associations

Dario Fiorentino<sup>1,2,\*</sup>, Roland Pesch<sup>3</sup>, Carmen-Pia Guenther<sup>3</sup>, Lars Gutow<sup>4</sup>, Jan Holstein<sup>4</sup>, Jennifer Dannheim<sup>4</sup>, Brigitte Ebbe<sup>4</sup>, Tim Bildstein<sup>3</sup>, Winfried Schroeder<sup>5</sup>, Bastian Schuchardt<sup>3</sup>, Thomas Brey<sup>2,4</sup>, Karen Helen Wiltshire<sup>1</sup>

<sup>1</sup>Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Wadden Sea Station Sylt, Hafenstrasse 43, 25992 List, Germany

<sup>2</sup>Helmholtz Institute for Functional Marine Biodiversity at the University Oldenburg, 23129 Oldenburg, Germany

<sup>3</sup>Bioconsult Schuchardt & Scholle GbR, Reeder-Bischoff-Straße 54, 28757 Bremen, Germany

<sup>4</sup>Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany

<sup>5</sup>University of Vechta, Chair of Landscape Ecology, PO 1553, 49364 Vechta, Germany

**ABSTRACT:** The concept of the marine ecological community has recently experienced renewed attention, mainly owing to a shift in conservation policies from targeting single and specific objectives (e.g. species) towards more integrated approaches. Despite the value of communities as distinct entities, e.g. for conservation purposes, there is still an ongoing debate on the nature of species associations. They are seen either as communities, cohesive units of non-randomly associated and interacting members, or as assemblages, groups of species that are randomly associated. We investigated such dualism using fuzzy logic applied to a large dataset in the German Bight (southeastern North Sea). Fuzzy logic provides the flexibility needed to describe complex patterns of natural systems. Assigning objects to more than one class, it enables the depiction of transitions, avoiding the rigid division into communities or assemblages. Therefore we identified areas with either structured or random species associations and mapped boundaries between communities or assemblages in this more natural way. We then described the impact of the chosen sampling design on the community identification. Four communities, their core areas and probability of occurrence were identified in the German Bight: AMPHIURA-FILIFORMIS, BATHYPOREIA-TELLINA, GONIADELLA-SPISULA, and PHORONIS. They were assessed by estimating overlap and compactness and supported by analysis of beta-diversity. Overall, 62% of the study area was characterized by high species turnover and instability. These areas are very relevant for conservation issues, but become undetectable when studies choose sampling designs with little information or at small spatial scales.

**KEY WORDS:** Ecological communities · Benthic macrofauna · Fuzzy classification · Spatial scales

## INTRODUCTION

The classic approach towards spatial ecological patterns is the community concept as introduced by Möbius (1877), who was the first to designate benthic species associations as biocoenoses, which are characterized by 'typical' species. However, in ecological

theory there is an ongoing debate about the nature of species associations. One school of thought holds the view that species associations represent communities, i.e. entities highly structured by interactions and dependencies between the participating species (e.g. Paine 1980, Richardson 1980, Gotelli & McCabe 2002, Chase & Leibold 2003). The opposite position

\*Corresponding author: dario.fiorentino@awi.de

sees assemblages, i.e. species associated by random processes, with interactions playing a minor role (Whittaker 1952, Mills 1969, McCann et al. 1998, Berlow 1999). There is empirical evidence in support of both views. However, much of the corresponding research may suffer from methodological shortcomings. Many studies have been carried out on comparatively small, local scales and thus may have captured just a part of the association's regional species pool (Sommerfeld et al. 2009). Studies at scales below the spatial extent of the association may yield a biased or compromised view of the nature of the association (Ricklefs 2008). Accordingly, there is an increasing need for regional, large-scale approaches.

Consequently, most statistical methods used for community classifications are based on the simplified community concept, i.e. highly structured units based on functionality, phylogenetic similarity, species characteristics or species–environment interactions (e.g. Picard et al. 2012, Jackson et al. 2012, Dunstan et al. 2013, Leaper et al. 2014) which all aim at classifying samples according to their similarity in a multidimensional space (Dunn 1974, Pakhira et al. 2005, Everitt et al. 2011). Thus, traditional clustering methods (e.g. hierarchical clustering, *k*-means, partition around medoids) only provide hard classification—that is, classes with discrete and sharp boundaries. In marine science, defining boundaries is a widely recognized problem that stems from supply side ecology theory (Lewin 1986), e.g. in fields such as habitat, sediment and biotope mapping. Quite often, sharp boundaries are preferred, as maps are easier to read and thus more useful for stakeholders. But in nature, sharp ecological boundaries only occur in very specific (and rare) cases. Often we deal either with gradual transitions (i.e. ecotones) or with areas that are difficult to assign to one class because a little bit of everything is occurring (e.g. mixed sediments). The identification of boundaries may depend on how classes are defined, on the method used for achieving the classification and on the act of classifying (which is organizing objects into classes that are limited by sharp boundaries per definition). This prevents proper description of systems shaped by gradual variation from cluster to cluster and leads to a loss of important ecological information by the setting of inappropriate and artificial boundaries (Burrough et al. 2000). In contrast, fuzzy clustering (Zadeh 1965, Bezdek 1974, Kaufman & Rousseeuw 2008, Everitt et al. 2011) allows for partial truth rather than the concept of true/false, which makes this concept more suitable to test for uncertainty in the community concept which is typical of natural systems.

Instead of following the traditional community concept and its spatial distribution, we used the approach of fuzzy clustering. We believe that a fuzzy modelling approach is the most flexible and most suitable approach to classify communities in the most realistic way without assuming any particular association rule or constraint (e.g. environmental). As a result, we could define community core areas as those areas in which a community has high probability of occurring (e.g.  $p > 0.8$ ). Thus, our approach allows for the analysis of core communities and areas of mixed communities (transition zones and undefined zones) at the same time. We think that this approach reflects the natural occurrence of associations in the most realistic way.

As local sampling might impair our analysis of the distribution of associations (see above, e.g. Ricklefs 2008, Sommerfeld et al. 2009), we further explored the effect of different spatial scales and sampling effort on the corresponding loss of information on species occurrence. For future analysis, this might enable an estimate of how much data are generally needed to fulfil the requirements of a holistic view on the association concept.

In this study, we applied this flexible approach for the first time to the benthic associations in the German Bight, in the North Sea—one of the most intensively studied marine areas of the planet. Here, first studies on benthos were already carried out at the beginning of the last century (Petersen 1918, Blegvad 1922, Hagmeier 1925). Small-scale studies are very numerous (e.g. Armonies 2000, Wieking & Kröncke 2003), with only a few surveys covering the whole German Bight (e.g. Salzwedel et al. 1985, Rachor & Nehmer 2003) or even larger areas (e.g. Duineveld et al. 1991, Heip et al. 1992, Künitzer et al. 1992, Heip & Craeymeersch 1995, Rees et al. 2007). Furthermore, temporal dynamics (e.g. Kröncke et al. 2011), the influence of environmental variables (e.g. van Hoey et al. 2004, Kröncke 2006, Reiss et al. 2010) and anthropogenic effects (e.g. Rachor 1990, Jennings et al. 1999, van Dalfsen et al. 2000, Witt et al. 2004) on the benthos have been studied intensely over recent decades. Therefore, sound knowledge of benthic spatial (and temporal) variability is still missing, although essential for an understanding of ecological processes as well as for sustainable management and conservation of benthic systems. Such detailed information is required, e.g. in the context of nature protection policies in Europe (e.g. European Parliament and Council 2008), the convention on habitat classification (e.g. Davies & Moss 2004, HELCOM 2013), the Habitats Directive (Council of the European Union

1992), the Marine Strategy Framework Directive (European Parliament and Council 2008) and the German Federal Nature Conservation Act (1998). In fact, agencies and directives use the concept of community under the assumptions that it can act as a proxy for biodiversity, and underpin specific functions and ecosystem services, although some of these assumptions still require scientific support. However, besides conservation and protection (Gonzalez-Mirelis et al. 2011), the concepts of community and assemblages have mainly been devoted to studying the association's properties (e.g. Guillemot et al. 2011), environmental relationships (e.g. Leaper et al. 2014; Gogina et al. 2016) and assembly rules (e.g. Spasojevic & Suding 2012, Ricklefs 2015).

In the present study, we firstly aimed at describing the spatial distribution of benthic species associations. Regardless of community and assemblage concepts, we drew maps of both types of associations. Simulating different experimental designs, we secondly aimed at evaluating which sampling design was more likely to detect one association type or the other.

## MATERIALS AND METHODS

To aid the explanation of the logic we pursued, we provide the workflow of the steps we followed in Fig. 1. To achieve the first aim (Fig. 1A), we classified the samples into groups according to their species composition. Those groups were the species associations. Thus, samples had a probability of belonging to each association. This allowed the determination of whether samples belonged exclusively to one association or to more than one. The probability of belonging to more than one association was used to calculate an index of equitability. This was used to map transitional zones and zones with random species associations. Samples exclusively belonging to one association made the spatial distribution of communities. Communities were named according to their characteristic species and mapped in terms of their probability of occurring.

To achieve the second aim (Fig. 1B), we bootstrapped the dataset in 2 ways: (1) keeping the amount of infor-

mation constant, increasing the distance among samples and (2) keeping the distance between samples constant and reducing the amount of information. Each bootstrap resulted in a subsample that was submitted to fuzzy clustering, and results were evaluated in terms of the distribution of the probability of samples belonging to each association.

Methods used to achieve those aims are further detailed in the workflow in Fig. 2. Fig. 2A shows the steps followed to prepare the dataset, and Fig. 2B,C outlines the steps followed to achieve the first and second aims.

## Data preparation

The primary database for this study comprised infaunal grab samples collected at 1146 stations, mostly within the German Exclusive Economic Zone (EEZ) between 2000 and 2013 (Fig. 3). These samples were taken during various surveys carried out by the following organisations: Bundesamt für Naturschutz, Landesamt für Landwirtschaft Umwelt und ländliche Räume, Bundesministerium für Umwelt Naturschutz

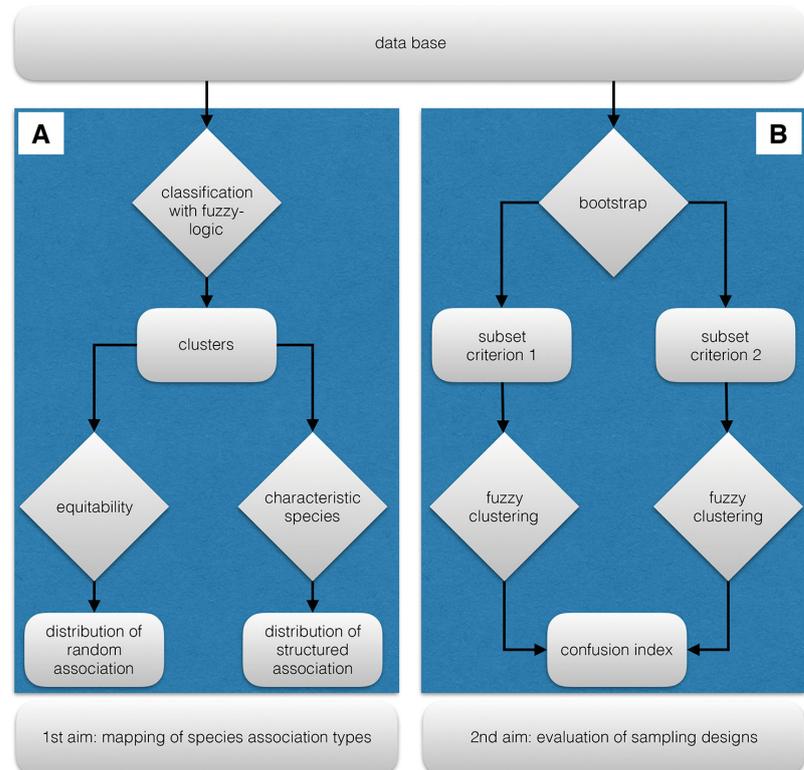


Fig. 1. Workflow outlining the logic followed to address the aims of the present study. Blue rectangles (A and B) represent the 2 different aims. Grey rhombi: conducted analyses; grey rectangles: results

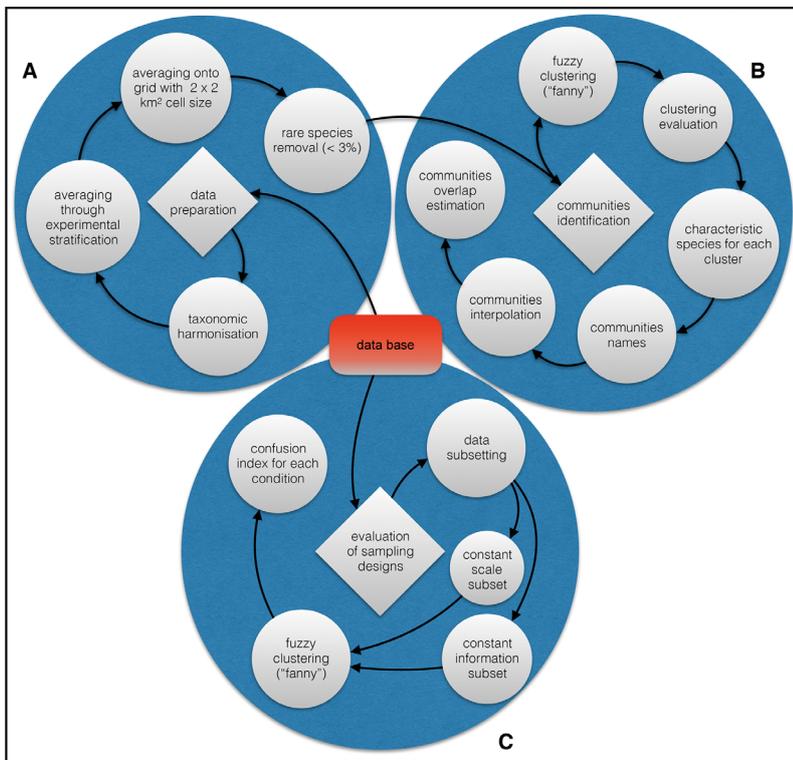
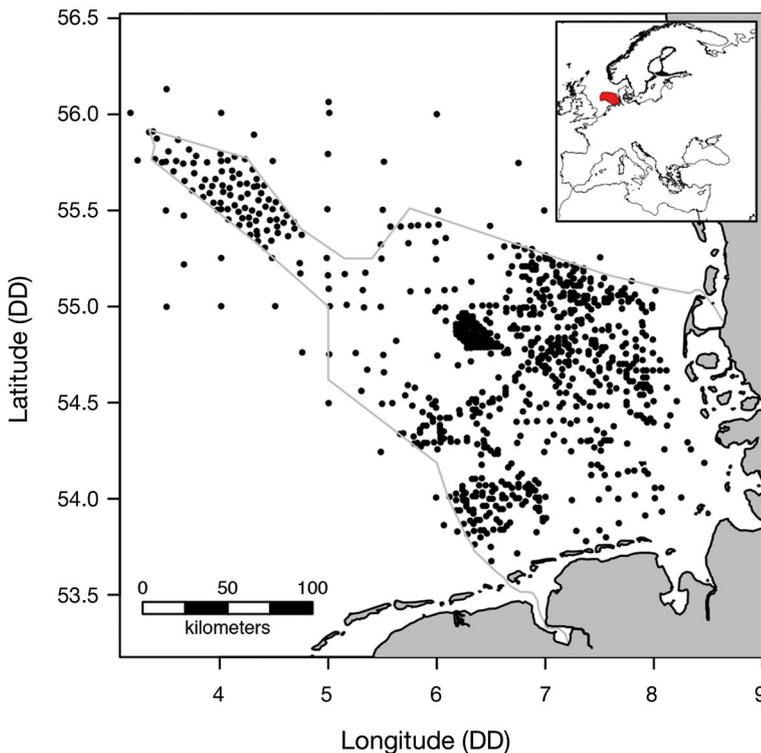


Fig. 2. Workflow outlining the methods followed by the present study showing 3 main groups of analyses that were achieved. (A) Steps followed to prepare the data; (B) steps followed to identify the associations; and (C) steps followed to evaluate the sampling design



Bau und Reaktorsicherheit, Arbeitsgemeinschaft Bund/Länder-Messprogramm für die Meeresumwelt, Mainstream Renewable Energy, TenneT Offshore GmbH, Energinet dk, and Institut für Angewandte Ökosystemforschung GmbH. Samples were collected and processed according to a standard protocol, i.e. 3 van Veen grab samples of 0.1 m<sup>2</sup> per station, sieved through a 1 mm mesh. The sampling method allowed retrieving data on soft bottom macrofauna. No information on hard bottom fauna is currently available, as hard substrates constitute only 1% of the German EEZ area (Laurer et al. 2012).

As shown in Fig. 2A, animals were identified to the lowest taxonomic level possible, and abundance was standardized to number of individuals m<sup>-2</sup> station<sup>-1</sup>. Taxonomic nomenclature was harmonized among surveys using the World Register of Marine Species (WoRMS; Boxshall et al. 2014).

For the purpose of the present study, to reduce the potential bias induced by such non-homogeneous distributions, we averaged species abundance according to the stratification of sampling designs adopted by different data sources onto a grid with 2 × 2 km<sup>2</sup> cells. Rare species (<3% relative abundance at each station) were excluded from the analysis following Mirza & Gray (1981). The 3% threshold was chosen to remove no more than the 70% of rare species from the dataset (Gray & Elliott 2009).

## Community identification

### Fuzzy clustering

We used fuzzy clustering for community classification. In particular, we applied the 'fanny' method (Kaufman

Fig. 3. Study area in the German Bight, southeastern North Sea; points indicate stations. Data courtesy of organisations listed in the text. Grey line: German EEZ

& Rousseeuw 2008), in which the algorithm first allocates the samples randomly into a set of *a priori* given clusters and uses their attribute values to calculate an objective function. This is repeated iteratively until the function is minimized and a stable solution is found (details on how to minimize the objective are described by Kaufman & Rousseeuw 2008, chapter 4, sections 4 and 5). The method only depends on the similarity between samples and is robust to outliers and non-spherical clusters (Kaufman & Rousseeuw 2008, Everitt et al. 2011). Furthermore, it reduces the bias introduced by zero-inflation because it uses an association measure (Clarke et al. 2006, Legendre & Legendre 2012). The minimized objective function allows for calculating the membership distribution of the samples. Sample membership may be interpreted as the distance of each sample to each cluster centroid. It describes the affinity of objects to each cluster and enables estimation of the uncertainty of an object's assignment (Pouw & Kwiatkowska 2013). In ecological terms, this procedure reproduces a species' feature of being related to more than one group of species (Jackson et al. 2010). The sample's membership distribution depends on the membership exponent, which sets the clustering degree of fuzziness. A membership exponent value close to 1 indicates hard clustering, whereas values tending to  $+\infty$  indicate complete fuzziness (Kaufman & Rousseeuw 2008). Fuzzy clustering was performed on Bray-Curtis distance (Clarke & Warwick 2001) of fourth-root transformed data to prevent very abundant species from masking the responses of those with low abundances (Clarke & Warwick 2001). Clusters were assigned to communities named according to characteristic species (see below for details) sensu Rachor et al. (2007). For visualization purposes, we interpolated the membership values to each cluster at each sampling station by means of kriging (Matheron 1963). We selected the best method among ordinary kriging, universal kriging or co-kriging according to the lowest root mean square error.

The following formulas detail the dependency ( $\sim$ ) of the target membership from the independent variables ( $x$ ,  $y$  and  $d$ ) used by kriging. Membership in clusters 1 to 4 is indicated by  $M_1$ ,  $M_2$ ,  $M_3$  and  $M_4$ ; water depth is abbreviated  $d$  and coordinates as  $x$  and  $y$ :  $M_1 \sim x + y$ ;  $M_2 \sim 1$ ;  $M_3 \sim x^2 + y^2 + x \cdot y + x + y$ ;  $M_4 \sim d$

Bathymetry was retrieved from the General Bathymetric Chart of the Oceans (www.gebco.net; the GEBCO\_2014 SID Grid, version 20141103). Overlaying the 4 maps of cluster memberships we assigned each  $1 \times 1 \text{ km}^2$  cell to the cluster with highest mem-

bership; that is, turning the fuzzy classification into a hard one.

### Clustering evaluation

Clustering was evaluated by estimating cluster compactness and separation (Wu & Yang 2005). These features were assessed using the normalized Dunn's partition coefficient (Dunn 1974), the silhouette averaged width (Kaufman & Rousseeuw 2008) and by visual examination of the principal coordinate ordination plot run on Bray-Curtis distance of species abundance (Borcard et al. 2011, Everitt et al. 2011). The normalized Dunn's partition coefficient (Dunn 1974) estimates the degree of fuzziness: it ranges from 0 (= total fuzziness), to 1 (= completely distinct clusters). The silhouette averaged width (Kaufman & Rousseeuw 2008) is the averaged ratio of the separation of each object from its cluster to the heterogeneity of the cluster. The closer the silhouette averaged width is to 1, the better the object was classified (Kaufman & Rousseeuw 2008, Everitt et al. 2011). For a more robust validation, we permuted the samples (i.e. resampled data without replacement) 1000 times and let the 'fanny' algorithm re-calculate the membership distribution and associated Dunn's partition coefficient and silhouette averaged width. We compared results among different partitions, letting the number of clusters vary between 2 and 10. Finally, we investigated the optimal clustering solution by statistically testing for differences of beta-diversity distribution within and between clusters (Anderson et al. 2006). We tested the null hypothesis that beta diversity did not change across the different clusters by using a test of homogeneity of variance (Anderson et al. 2006) on the 'Jaccard' association measure (Jaccard 1900) of species presence-absence data.

### Characteristic species

For the optimal clustering solution, we identified characteristic species of clusters. We used a multi-criteria approach in which for the  $i^{\text{th}}$  species we calculated fidelity in abundance  $FA_i$ , numerical dominance  $ND_i$ , presence  $P_i$ , and fidelity in presence  $FP_i$  as:

$$FA_i = \frac{A_{ic}}{A_i}, ND_i = \frac{A_{ic}}{A_c}, P_i = \frac{P_{ic}}{Nst_c}, FP_i = \frac{P_{ic}}{Nst_i} \quad (1)$$

where  $A_{ic}$  is the abundance of the  $i^{\text{th}}$  species within the  $c^{\text{th}}$  cluster,  $A_i$  is the abundance of the  $i^{\text{th}}$  species in the dataset,  $A_c$  is the abundance of all species

within the  $c^{\text{th}}$  cluster,  $P_{ic}$  is the presence of the  $i^{\text{th}}$  species in the  $c^{\text{th}}$  cluster,  $Nst_c$  is the number of stations in the  $c^{\text{th}}$  cluster and  $Nst_i$  is the number of stations in the dataset where the  $i^{\text{th}}$  species was present.

A species was accepted as characteristic if at least 3 of these indices exceeded a given threshold value ( $FA > 0.5$ ,  $ND > 0.01$ ,  $p > 0.6$ ,  $FP > 0.5$ ) as described by Salzwedel et al. (1985) and Rachor et al. (2007). We named communities according to the nomenclature provided by Salzwedel et al. (1985) and Rachor et al. (2007). We used the same names where characteristic species in our list matched those provided by Salzwedel et al. (1985) and Rachor et al. (2007). Community names are written in uppercase and not italic to differentiate their names from species names.

### Cluster overlap

Interpolated memberships to clusters were used to calculate the Pielou (1975) equitability index at each  $1 \times 1 \text{ km}^2$  cell. This provided estimates of the degree of overlap among clusters in a scale from 0 (no overlap) to 1 (all communities share same probability of occurring). For each cluster separately, we investigated the variation of beta diversity according to the degree of equitability. We tested the null hypothesis that beta diversity did not change across different degrees of equitability. A test of homogeneity of variance (Anderson et al. 2006) on the Jaccard association measure (Jaccard 1900) of species presence-absence data was used. To extract the data required to test each cluster, we selected those stations with membership to the selected cluster higher than 0.1. Furthermore, we explored whether the degree of equitability was related to a significant increase or decrease of the number of taxa within 4 taxonomic ranks: species, genus, family and order.

### Evaluation of sampling designs

In order to explore the effects of (1) distance between samples and (2) amount of information (i.e. number of stations) on the clustering results, we selected subsets of the main dataset in 2 ways. We either let the minimum distance among sampling stations change (minimum distance = 2, 3, 5, 6, 10 and 12 km), keeping the amount of information constant (number of stations = 110), or we let the amount of stations change (number of stations = 832, 600, 330, 250, 150, 110), regardless of the distance among sampling stations. For this subset, the stations were

selected to be equal in number in case of confused design, that is, when stations would be selected by letting the distance change without keeping the number of stations constant (i.e.  $d > 2, 3, 5, 6, 10$  and  $12 \text{ km}$  would generate sets with 832, 600, 330, 250, 150, 110 stations, respectively).

There were many combinations of stations that would have satisfied each condition. Therefore, we performed the fuzzy clustering on 500 possible combinations of stations for each condition and let the algorithm run with the previously selected optimal number of clusters. For each combination, the membership values were used to calculate the degree of confusion (CI) as defined by Burrough et al. (1997):

$$CI = \frac{\mu_{\max-1}}{\mu_{\max}} \quad (2)$$

where  $\mu_{\max}$  and  $\mu_{\max-1}$  are the maximum and second largest membership values given for the same pixel. The index ranges from 0 (lowest confusion) to 1 (highest confusion). This allowed calculating a confusion index distribution for each condition. We tested the null hypothesis that the confusion index did not change across the different conditions.

All analyses were conducted using R v.3.2.3 (R Development Core Team 2015); the R script as well as all georeferenced maps and graphs are available in the Supplements at [www.int-res.com/articles/suppl/m584p017\\_supp/](http://www.int-res.com/articles/suppl/m584p017_supp/). Fuzzy clustering was performed using the function 'Fanny' as described by Kaufman & Rousseeuw (2008) and implemented in R package 'cluster' (Maechler et al. 2015). Package 'gstat' (Pebesma 2004) was used for interpolation of the communities' distributions and 'vegan' for analysis of beta-diversity (Oksanen et al. 2015). Maps used the WGS84 (EPGS 4326) projection. On the maps, the coastline was provided by Global Self-consistent Hierarchical High-resolution Geography (version 2.3.4 Jan 1, 2015; <https://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html>; Wessel & Smith 1996), and the grey line indicates the border of the German EEZ.

## RESULTS

### Communities in the German Bight

Our analysis identified 4 core benthic communities in the southern North Sea: The AMPHIURA-FILIFORMIS community, which occupies areas towards the central North Sea and along the borders of the paleo Elbe river valley (Fig. 4A); the BATHYPOREIA-TELLINA community, which is situated mostly on sand banks

(Fig. 4B); the GONIADELLA-SPISULA community (Fig. 4C) in shallower areas and the PHORONIS community on the fringe of the Sylt Outer Reef (Fig. 4D). Georeferenced raster data for each community are available in Supplement 3.

### Clustering evaluation

The normalized Dunn's partition coefficient (Fig. 5A) and the silhouette averaged width (Fig. 5B) similarly revealed that the classification with 4 clusters had the lowest degree of fuzziness ( $0.764 \pm 0.024$ ) and the highest compactness ( $0.17 \pm 0.01$ ).

Beta diversity analysis showed significant differences both between and within groups (see Table S1 in Supplement 1). Since beta-diversity provides an estimate of species turnover, the detection of significant differences between groups supports the claim that communities differ in their species composition (Fig. 6).

Differences within groups (Fig. 6) were mostly driven by the PHORONIS and GONIADELLA-SPISULA communities, which displayed the shortest and highest distances to the centroid respectively (Fig. 6).

Principal coordinate ordination plots confirmed these findings (see Fig. S1 in Supplement 1). In fact, visual examination of sample distribution and the

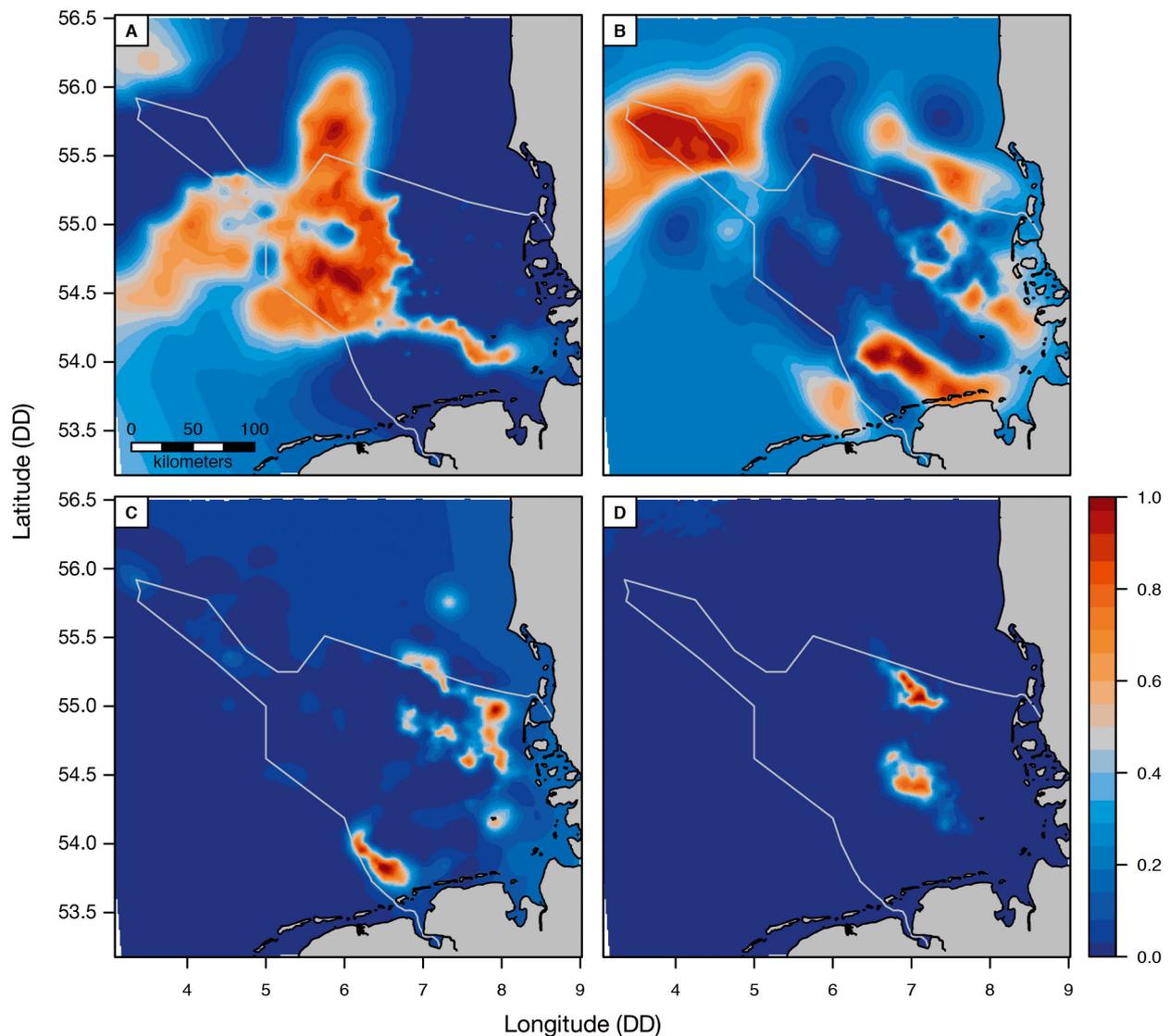


Fig. 4. Probability of occurrence of the 4 groups: (A) AMPHIURA-FILIFORMIS, (B) BATHYPOREIA-TELLINA, (C) GONIADELLA-SPISULA, and (D) PHORONIS. Maps were created by interpolation of probability of occurrence estimated with fuzzy clustering

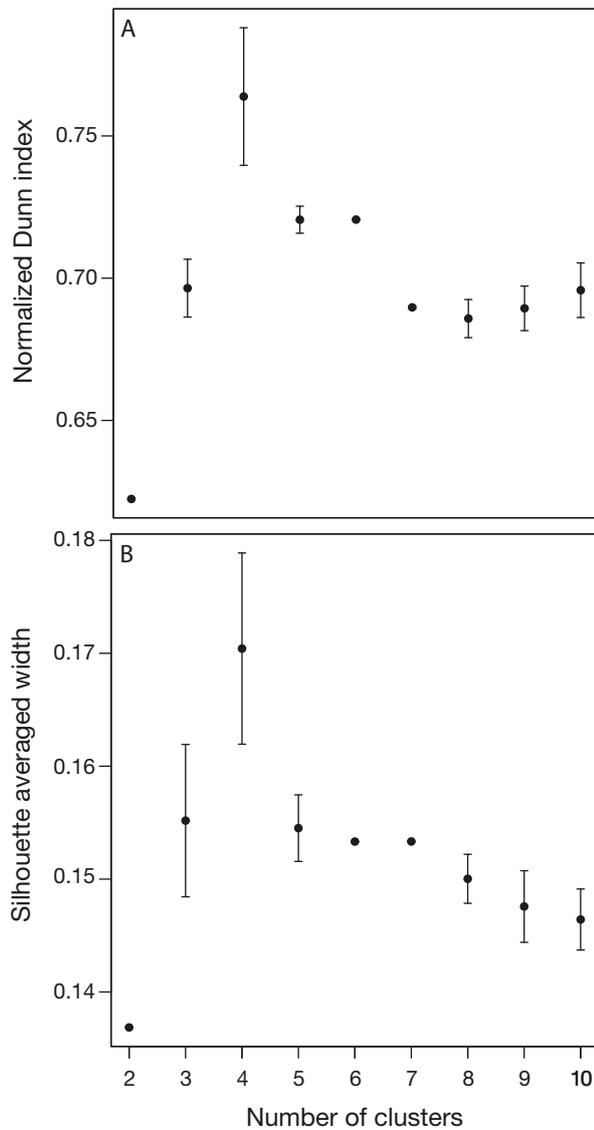


Fig. 5. Normalized (A) Dunn's index and (B) silhouette averaged width for optimal clustering selection. Error bars: index estimation over 1000 runs of fuzzy clustering

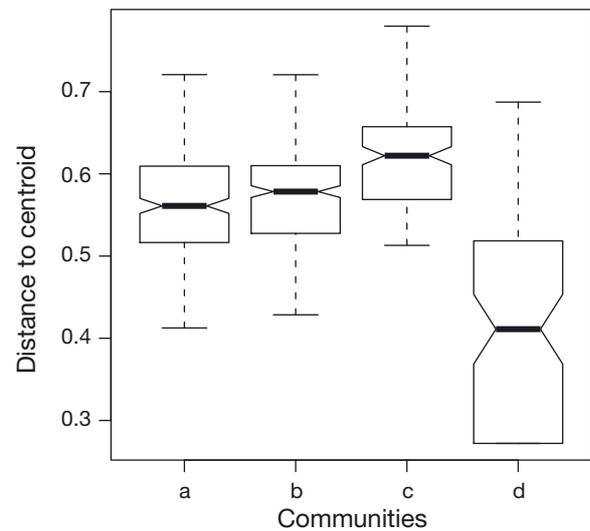


Fig. 6. Boxplot of beta diversity distance from stations to group centroids. AMPHIURA-FILIFORMIS, BATHYPOREIA-TELLINA, GONIADELLA-SPISULA, and PHORONIS is on the x-axis indicated by letters from 'a' to 'd' respectively

related cluster hulls showed that a 4-cluster arrangement provided the least overlap and the highest compactness (Fig. S1). Hard classification of the communities' distributions is provided in Fig. S2A. A comparison with previous hard classification of community distributions (Fig. S2B, modified from Rachor & Nehmer 2003) provided a further evaluation.

### Characteristic species

The characteristic species of the community areas are listed in Table 1. As these are almost identical to those reported by Salzwedel et al. (1985) and Rachor et al. (2007), we chose similar community

Table 1. Characteristic species for the 4 groups AMPHIURA-FILIFORMIS, BATHYPOREIA-TELLINA, GONIADELLA-SPISULA, and PHORONIS according to the criteria of Rachor et al. (2007). For group 4, only *Phoronis* spp. satisfied the given criteria. Groups were named following Rachor et al. (2007)

Group 1 (AMPHIURA-FILIFORMIS)	Group 2 (BATHYPOREIA-TELLINA)	Group 3 (GONIADELLA-SPISULA)	Group 4 (PHORONIS)
<i>Abra nitida</i>	<i>Bathyporeia elegans</i>	<i>Aonides paucibranchiata</i>	<i>Phoronis</i> spp.
<i>Amphictene auricoma</i>	<i>Bathyporeia guilliamsoniana</i>	<i>Branchiostoma lanceolatum</i>	
<i>Amphiura filiformis</i>	<i>Echiurus echiurus</i>	<i>Echinocyamus pusillus</i>	
<i>Corbula gibba</i>	<i>Lanice conchilega</i>	<i>Ophelia borealis</i>	
<i>Kurtiella bidentata</i>	<i>Magelona johnstoni</i>	<i>Ophelia limacina</i>	
<i>Lagis koreni</i>	<i>Scoloplos Scoloplos armiger</i>	<i>Pisione remota</i>	
<i>Nucula nitidosa</i>	<i>Spiophanes bombyx</i>	<i>Protodorvillea kefersteini</i>	
<i>Phaxas pellucidus</i>	<i>Tellina fabula</i>	<i>Spio filicornis</i>	
<i>Pholoe baltica</i>	<i>Urothoe poseidonis</i>	<i>Spio gonocephala</i>	

names accordingly: AMPHIURA-FILIFORMIS, BATHYPOREIA-TELLINA, GONIADELLA-SPISULA and PHORONIS.

Characteristic species of each community are exclusive to each community. The AMPHIURA-FILIFORMIS and BATHYPOREIA-TELLINA communities are characterized by 9 species, GONIADELLA-SPISULA by 8 species, and the PHORONIS community is characterized by the species *Phoronis* spp. only.

### Overlap in communities

We used interpolated memberships to each of the 4 clusters to set up a map of community equitability (Fig. 7). Equitability ranged between 0 (a location can be assigned to one community with high probability) and 1 (a location could be assigned to each of the 4 communities with equal probability). Georeferenced rasters of community equitability are available in Supplement 3. An equitability threshold value of 0.4 was used to separate areas in which one community clearly dominated (blue areas in Fig. 7). Those areas covered about the 38% of the study area (i.e. the region actually covered by data), while the remaining 62% could be assigned to more than one community (red areas in Fig. 7). Each community's species composition changed significantly according to the degree of equitability (Fig. 8, Table S2). Maps in Fig. S3 show the subsets used to conduct the tests. Beta diversity showed a significant increase from areas characterized by low equitability to areas with high equitability (Table S2). Patterns of change of taxa with increasing equitability were not consistent across communities (Table S3). Significant taxonomic simplification was observed for AMPHIURA-FILIFORMIS and GONIADELLA-SPISULA, whereas an increase of the number of taxa was observed for BATHYPOREIA-TELLINA and PHORONIS (Table S3).

### Experimental design assessment

The experimental design assessment exercise, with the 12 settings shown in the examples provided in Figs. S4 & S5, indicates that the confusion index significantly increased with increasing number of stations (Fig. 9a,

Table S4) and with increasing distance between stations (Fig. 9b, Table S4). Accordingly, both low amounts of information and small spatial scale gener-

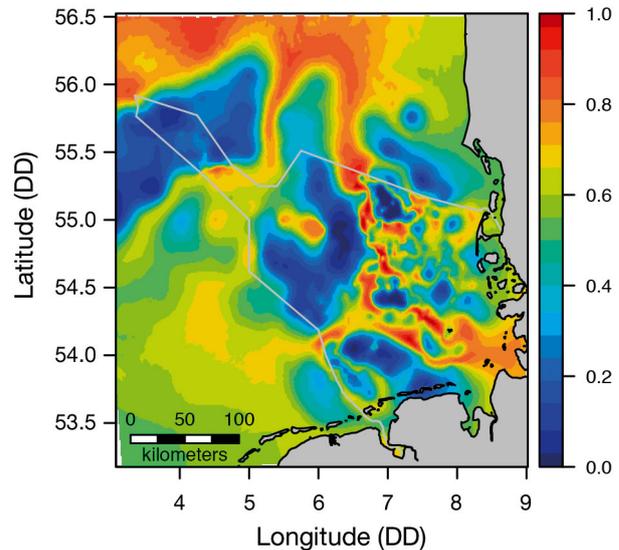


Fig. 7. Map of community equitability. Blue shading indicates areas where only one group occurs or dominates over the other 3 (low equitability); red shading indicates areas where all defined communities have similar probability of occurrence (high equitability)

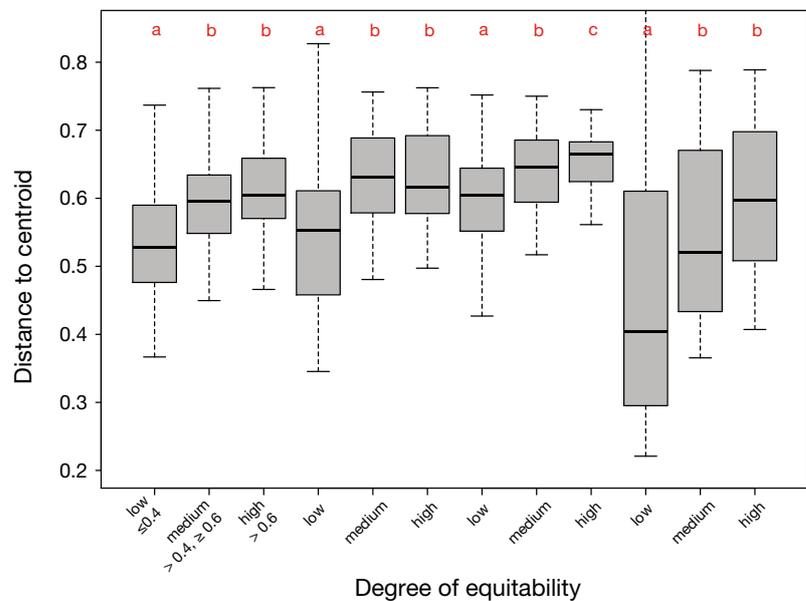


Fig. 8. Boxplot of beta diversity distance from stations to centroids low medium and high equitability. Degrees of equitability are provided for each of the 4 groups: AMPHIURA-FILIFORMIS, BATHYPOREIA-TELLINA, GONIADELLA-SPISULA, and PHORONIS. Low: equitability  $\leq 0.4$ ; medium: equitability  $> 0.4$  and  $\leq 0.6$ ; high: equitability  $> 0.6$ . Lowercase letters above boxes result from pairwise comparison within each group; same letters indicate no significant differences between the 2 degrees of equitability compared; different letters indicate significant difference at  $\leq 0.001$

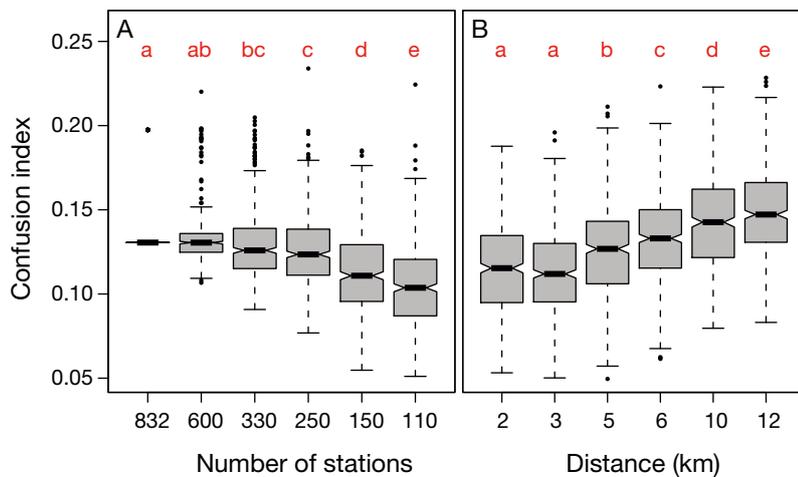


Fig. 9. Estimated confusion index showing experiments with (A) simulated change of amount of information and (B) simulated change of spatial scale and constant information

ate low overlap (= low confusion), while a high amount of information and samples distributed evenly over large scales generate high overlap (= high confusion index).

## DISCUSSION

The classification of marine communities constitutes one of those cases where multiple causes hamper the development of a rather realistic picture of spatial patterns. Firstly, the community concept is sometimes used to indicate structured units, and other times to indicate random species association (i.e. assemblages). Secondly, standard classification methods are designed to interpret communities as structured units and therefore emphasize sharp boundaries. Inappropriate data design may further emphasize this trend. Thirdly, due to relatively recent regulations (such as European Habitats Directive and European Marine Strategy Framework Directive), decision makers are in need of 'clear and distinct' data and information they can work with, and sharp community boundaries greatly simplify spatial management.

Fuzzy logic is one attempt to overcome the dilemma of sharp borders without creating too much complexity. It does not rely on any specific community concept and is flexible enough to provide both classical classification (true/false logic) and fuzzy classification (partially true) typical of natural systems. Furthermore, fuzzy clustering is a site-based approach, which we prefer over others because it implies looking at species occurrence at a specific

location without assuming any particular species–environment relationship. In fact, weak relationships do not necessarily imply that a community structure does not exist. Many studies have used environmental constraints to define community distributions (e.g. Leaper et al. 2014), but we presume that a community definition, and hence its related concept, might be independent of any environmental constraints. Indeed, other factors may play a more important role (Ricklefs 2015) such as evolutionary processes (Wilson 1961), historical processes (Gotelli et al. 1997, Bloch et al. 2007), species interactions (Peres-Neto et al. 2001, Comita et al. 2010, Mangan et al. 2010) and species dispersal capabilities (Cook & Quinn 1998, Loo et al. 2002, Ulrich 2004, McAbendroth et al. 2005).

The size of our dataset eliminated the potential bias of localness (as described by Ricklefs 2008) and made us confident that all species relevant to compose the description of the communities were included. The saturation reached by the confusion index both according to the changing amount of information (Fig. 9A) and distance (Fig. 9B) suggests that no further information (no larger dataset) and no higher resolution (no smaller distance among samples) are required to improve the classification and related boundary definitions.

The actual amount of information available may be relevant for the bias of localness as well (Ricklefs 2008). Although the concept of spatial scale is closely related to the amount of information available (i.e. sampling effort), we disentangled the potential effects of spatial scale and amount of information on the community identification process by treating them separately.

Subsets of our main dataset, created by keeping constant either the amount of information or the scale (see Figs. S4 & S5 in Supplement 1), let us simulate different sampling designs, which were further evaluated in terms of the returned degree of confusion (Fig. 9). A higher index of confusion may be expected when there is more information, simply because the variance contained in the data would be higher, and vice versa. Regarding spatial designs (Fig. S4), clumped designs favor the detection of the core of the community, whereas more evenly distributed station patterns may equally entail portions of the core of the community and transitional areas, therefore return-

ing higher confusion indices. Sampling design, however, is a consequence of the actual research question, and thus the community concept might change according to the field of its application (Mills 1969). For instance, many monitoring programs sample few stations in relatively small areas and thus tend to 'see' a community as a unit (Fig. 9; low confusion). On the contrary, large-scale projects that aim at obtaining representative data of large areas are prone to find areas where species tend to associate at random (Fig. 9; high confusion). Therefore, not only localness but also sampling effort (i.e. amount of information) may impair our picture of a species association.

Our maps of community probability of occurrence (Fig. 4) result from interpolation. We interpreted the gradual change in community occurrence as a gradient of species association strength. The association becomes weaker with increasing distance from the core area. This view coincides with the concept of species nestedness (Patterson & Atmar 1986), i.e. the strength of the mutualistic network structure (Bastolla et al. 2009) decreases along this gradient. However, species segregation, i.e. the tendency of species to exclude each other (see Ulrich & Gotelli 2007 for further details), might act as an additional but opposite effect. In the present study, fuzzy clustering allowed us to handle both phenomena equally well by emphasizing neither nestedness nor segregation. Nested sites (those where most species were shared), obtained a higher probability of belonging to one cluster. Those sites showing high species segregation (few or no species shared) obtained a lower probability of belonging to all clusters. This is the reason why fuzzy logic can mirror natural patterns so well by showing the gradients of community distribution, from structured to unstructured units.

The evaluation of fuzzy classification aimed at ensuring optimal cluster configuration. The correct choice of clusters minimizes the portion of uncertainty related to algorithm error. For example, in Fig. S1, we show that configurations different from 4 clusters have a much higher degree of between-cluster similarity. However, we are not aware of any method that would enable us to separate different sources of uncertainty. Reducing the misclassification-due-to-algorithm error provides more accurate classification and related cluster membership. This lets us interpret membership values at a given location as being mostly representative of a natural pattern, not as a result of an algorithm mistake. In Fig. 7, we show that the fuzzy approach allows us to bridge the gap between the concepts

of community (structured unit) and assemblages (random association of species). Fig. 7 is a map of equitability. Since low equitability is given when one community occurs, we considered such values as indications of structured species associations. On the contrary, high equitability is given by an equal probability of communities occurring. We considered such values as indications of random species associations. Blue shaded areas (equitability < 0.4, dominance > 80%) support the view of communities as highly structured units (sensu Paine 1980, Richardson 1980, Gotelli & McCabe 2002, Chase & Leibold 2003). Those areas may be seen as the core area of a community. Green to red shades (equitability > 0.4) may indicate areas of a random, unstructured species association (sensu Whittaker 1952, Mills 1969, McCann et al. 1998, Berlow 1999), e.g. a transition state spatially situated between 2 or more core areas, or an unstable state with a high degree of species turnover (caused, for example, by non-chronic human disturbance). Therefore, Fig. 7 might be interpreted as a map of both stable and dynamic areas. Beta diversity may be a reasonable measure of species turnover in time (Anderson et al. 2011). A significant increase in beta diversity with increasing equitability (Fig. 8, Table S2) further supports the idea that areas with random species associations are actually characterized by high species turnover. We speculate that causes of high equitability might be driven by more local events (e.g. fisheries), because patterns of change in taxa with increasing equitability were not consistent across communities (Table S3). Despite our supposition that the pattern shown might be caused by fishery activities, we are not aware of any studies that could provide further comparison. To our knowledge, the present study is unique in showing patterns of types of associations and areas with different regimes of species turnover (Fig. 7).

The observed taxonomic simplification, presented to order level for AMPHIURA-FILIFORMIS and GONIADELLA-SPISULA (Table S3), suggests that high equitability for those groups may also lead to an ecosystem function simplification.

High values of beta diversity within the GONIADELLA-SPISULA community (Fig. 6c) implies higher within-community dynamics and a facilitation of species colonization. On the contrary, low beta-diversity within the PHORONIS community (Fig. 6D) indicates low species turnover, apparently because in this community the habitat is thoroughly modified by *Phoronis* spp. (Niermann 1996), which hampers colonization by other species.

Our 'hard' classification confirms the overall patterns identified in previous studies (e.g. Rachor & Nehmer 2003) (Fig. S2). However, our analysis reduced the number of communities from 6 (Rachor & Nehmer 2003, Pesch et al. 2016) to 4. Our analysis found only minor differences between BATHYPOREIA-TELLINA and the TELLINA-FABULA communities in terms of species composition, and thus merged them into the BATHYPOREIA-TELLINA community. The same holds true for the NUCULA-NITIDOSA and AMPHIURA-FILIFORMIS communities (Rachor et al. 2007), which were merged into the AMPHIURA-FILIFORMIS community. Our fuzzy classification found one case of a community characterized by *Phoronis* spp. (Table 1). This species occurs in very high abundances (Niermann 1996) and builds chitinous tubes that distinctly change the characteristics of the soft bottom habitat. *Phoronis* spp. displays strong annual density fluctuations owing to its high reproductive rate, that compensates for decimation events (Niermann 1996).

A major improvement of our approach is the capability of identifying stable and less stable/dynamic areas. There are core areas of the distribution of all 4 communities (Fig. 4) in which they were stable, at least for the time interval covered by our dataset (13 yr). These pictures add up to an overall spatial pattern of stability (Fig. 7). For example, the Dogger Bank and sand banks were demonstrated to be stable areas, whereas the paleo-Elbe river valley appeared unstable. Potential causes of such stability and instability still need to be detected and assessed. If nature protection policies, such as the Habitats Directive (Council of the European Union 1992) and the Marine Strategy Framework Directive (European Parliament and Council 2008), move toward diversity protection through management of hard classified communities, only stable areas (38% of studied area) would be considered and protected, while those being more dynamic, with higher species turnover (62% of the studied area), would tend to be neglected. Furthermore, we do not know whether these dynamic areas result from natural or anthropogenic processes and whether they require specific management, e.g. because of triggering species invasion. These are open questions that need to be addressed properly before proceeding further with the development and implementation of conservation measures.

**Acknowledgements.** D.F. was partially funded by the Bundesamt für Naturschutz (German Federal Agency for Nature Conservation, [www.bfn.de](http://www.bfn.de)) with the project 'Cluster 6 BfN', 'Kartierung und Registrierung der marinen Lebensraumtypen (LRT) bzw. Biotope in der Ausschließlichen Wirtschaftszone (AWZ)', nr. Z 1.2-53302/2010/15. Thanks are due to

data providers (courtesy of Bundesamt für Naturschutz, Landesamt für Landwirtschaft Umwelt und ländliche Räume, Bundesministerium für Umwelt Naturschutz Bau und Reaktorsicherheit, Arbeitsgemeinschaft Bund/Länder-Messprogramm für die Meeresumwelt, Mainstream Renewable Energy, TenneT Offshore GmbH, Energienet dk, Institut für Angewandte Ökosystemforschung GmbH).

#### LITERATURE CITED

- ✦ Anderson MJ, Ellingsen KE, McArdle BH (2006) Multivariate dispersion as a measure of beta diversity. *Ecol Lett* 9: 683–693
- ✦ Anderson MJ, Crist TO, Chase JM, Vellend M and others (2011) Navigating the multiple meanings of diversity: a roadmap for the practicing ecologist. *Ecol Lett* 14:19–28
- ✦ Armonies W (2000) On the spatial scale needed for benthos community monitoring in the coastal North Sea. *J Sea Res* 43:121–133
- ✦ Bastolla U, Fortuna MA, Pascual-García A, Ferrera A, Luque B, Bascompte J (2009) The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* 458:1018–1020
- ✦ Berlow EL (1999) Strong effects of weak interactions in ecological communities. *Nature* 398:330–334
- ✦ Bezdek DJC (1974) Numerical taxonomy with fuzzy sets. *J Math Biol* 1:57–71
- ✦ Blegvad H (1922) Animal communities in the southern North Sea. *Proc Zool Soc Lond* 92:27–32
- ✦ Bloch CP, Higgins CL, Willig MR (2007) Effects of large-scale disturbance on metacommunity structure of terrestrial gastropods: temporal trends in nestedness. *Oikos* 116:395–406
- Borcard D, Gillet F, Legendre P (2011) *Numerical ecology with R*, 1<sup>st</sup> edn. Springer, New York, NY
- ✦ Boxshall GA, Mees J, Costello MJ, Hernandez F and others (2014) World register of marine species (WoRMS) [www.marinespecies.org](http://www.marinespecies.org)
- ✦ Burrough PA, van Gaans PFM, Hootsmans R (1997) Continuous classification in soil survey: spatial correlation, confusion and boundaries. *Geoderma* 77:115–135
- ✦ Burrough PA, van Gaans PFM, MacMillan RA (2000) High-resolution landform classification using fuzzy *k*-means. *Fuzzy Sets Syst* 113:37–52
- Chase JM, Leibold MA (2003) *Ecological niches: linking classical and contemporary approaches*, 1st edn. University of Chicago Press, Chicago, IL
- Clarke KR, Warwick RM (2001) *Change in marine communities: an approach to statistical analysis and interpretation*, 2nd edn. PRIMER-E, Plymouth
- ✦ Clarke KR, Somerfield PJ, Chapman MG (2006) On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. *J Exp Mar Biol Ecol* 330:55–80
- ✦ Comita LS, Muller-Landau HC, Aguilar S, Hubbell SP (2010) Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* 329: 330–332
- ✦ Cook RR, Quinn JF (1998) An evaluation of randomization models for nested species subsets analysis. *Oecologia* 113:584–592
- Council of the European Union (1992) Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural

- habitats and of wild fauna and flora. *Off J Eur Comm L* 206:7–50
- Davies CE, Moss D (2004) EUNIS habitat classification marine habitat types: revised classification and criteria, September 2004. Centre for Ecology and Hydrology, Dorset
- ✦ Duineveld GCA, Künitzer A, Niermann U, De Wilde PAWJ, Gray JS (1991) The macrobenthos of the North Sea. *Neth J Sea Res* 28:53–65
- ✦ Dunn JC (1974) Some recent investigations of a new fuzzy partitioning algorithm and its application to pattern classification problems. *J Cybern* 4:1–15
- ✦ Dunstan PK, Foster SD, Hui FKC, Warton DI (2013) Finite mixture of regression modeling for high-dimensional count and biomass data in ecology. *J Agric Biol Environ Stat* 18:357–375
- European Parliament and Council (2008) Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). *Off J Eur Union L* 164:19–40
- Everitt BS, Landau S, Leese M, Stahl D (2011) Cluster analysis, 5<sup>th</sup> edn. John Wiley & Sons, London
- Federal Nature Conservation Act (Bundesnaturschutzgesetz, BNatSchG) (1998) Conservation of Nature and of Landscapes Act (Gesetz über Naturschutz und Landschaftspflege). *Federal Law Gazette I*, p 2994
- Gogina M, Nygård H, Blomqvist M, Daunys D and others (2016) The Baltic Sea scale inventory of benthic faunal communities. *ICES J Mar Sci* 73:1196–1213
- ✦ Gonzalez-Mirelis G, Bergström P, Lindegarth M (2011) Interaction between classification detail and prediction of community types: implications for predictive modeling of benthic biotopes. *Mar Ecol Prog Ser* 432:31–44
- ✦ Gotelli NJ, McCabe DJ (2002) Species co-occurrence: a meta-analysis of JM Diamond's assembly rules model. *Ecology* 83:2091–2096
- ✦ Gotelli NJ, Buckley NJ, Wiens JA (1997) Co-occurrence of Australian land birds: Diamond's assembly rules revisited. *Oikos* 80:311–324
- Gray JS, Elliott M (2009) Ecology of marine sediments from science to management. Oxford University Press, Oxford
- ✦ Guillemot N, Kulbicki M, Chabanet P, Vigliola L (2011) Functional redundancy patterns reveal non-random assembly rules in a species-rich marine assemblage. *PLOS ONE* 6:e26735
- Hagmeier A (1925) Vorläufiger Bericht über die vorbereitenden Untersuchungen der Bodenfauna der Deutschen Bucht mit dem Petersen-Bodengreifer. *Ber Dtsch Wiss Komm Meeresforsch* 1:247–272
- ✦ Heip C, Craeymeersch JA (1995) Benthic community structures in the North Sea. *Helgol Meeresunters* 49:313–328
- ✦ Heip C, Basford D, Craeymeersch JA, Dewarumez JM and others (1992) Trends in biomass, density and diversity of North Sea macrofauna. *ICES J Mar Sci* 49:13–22
- HELCOM (2013) Technical report on the HELCOM underwater biotope and habitat classification. Helsinki Commission, Helsinki
- Jaccard P (1900) Contribution au problème de l'immigration post-glaciaire de la flore alpine: étude comparative de la flore alpine du massif de Wildhorn, du haut bassin du Trient et de la haute vallée de Bagnes. *Bull Société Vaudoise Sci Nat* 36:87–130
- Jackson DA, Walker SC, Poos MS (2010) Cluster analysis of fish community data: new tools for determining meaningful groupings of sites and species assemblages. *Am Fish Soc Symp* 73:503–527
- ✦ Jackson MM, Turner MG, Pearson SM, Ives AR (2012) Seeing the forest and the trees: multilevel models reveal both species and community patterns. *Ecosphere* 3:1–16
- ✦ Jennings S, Alvsvåg J, Cotter AJR, Ehrich S and others (1999) Fishing effects in northeast Atlantic shelf seas: patterns in fishing effort, diversity and community structure. III. International trawling effort in the North Sea: an analysis of spatial and temporal trends. *Fish Res* 40:125–134
- Kaufman L, Rousseeuw PJ (2008) Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, Hoboken, NJ
- ✦ Kröncke I (2006) Structure and function of macrofaunal communities influenced by hydrodynamically controlled food availability in the Wadden Sea, the open North Sea, and the deep-sea. A synopsis. *Senckenb Marit* 36:123–164
- ✦ Kröncke I, Reiss H, Eggleton JD, Aldridge J and others (2011) Changes in North Sea macrofauna communities and species distribution between 1986 and 2000. *Estuar Coast Shelf Sci* 94:1–15
- ✦ Künitzer A, Basford D, Craeymeersch JA, Dewarumez JM and others (1992) The benthic infauna of the North Sea: species distribution and assemblages. *ICES J Mar Sci* 49:127–143
- Laurer WU, Naumann M, Zeiler M (2012) Erstellung der Karte zur Sedimentverteilung auf dem Meeresboden in der deutschen Nordsee nach der Klassifikation von Figge (1981). LBEG–BSH–BGR, Hannover/Hamburg
- ✦ Leaper R, Dunstan PK, Foster SD, Barrett NS, Edgar GJ (2014) Do communities exist? Complex patterns of overlapping marine species distributions. *Ecology* 95:2016–2025
- Legendre P, Legendre L (2012) Numerical ecology, 3rd edn (revised). Elsevier, Oxford
- ✦ Lewin R (1986) Supply-side ecology. *Science* 234:25–27
- ✦ Loo SE, Nally RM, Quinn G (2002) An experimental examination of colonization as a generator of biotic nestedness. *Oecologia* 132:118–124
- ✦ Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2015) Package 'cluster': cluster analysis basics and extensions. R package version 2.0.6. <https://cran.r-project.org/web/packages/cluster/>
- ✦ Mangan SA, Schnitzer SA, Herre EA, Mack KML, Valencia MC, Sanchez EI, Bever JD (2010) Negative plant-soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* 466:752–755
- ✦ Matheron G (1963) Principles of geostatistics. *Econ Geol* 58:1246–1266
- ✦ McAbendroth L, Foggo A, Rundle SD, Bilton DT (2005) Unravelling nestedness and spatial pattern in pond assemblages. *J Anim Ecol* 74:41–49
- ✦ McCann K, Hastings A, Huxel GR (1998) Weak trophic interactions and the balance of nature. *Nature* 395:794–798
- ✦ Mills EL (1969) The community concept in marine zoology, with comments on continua and instability in some marine communities: a review. *J Fish Res Board Can* 26:1415–1428
- ✦ Mirza FB, Gray JS (1981) The fauna of benthic sediments from the organically enriched Oslofjord, Norway. *J Exp Mar Biol Ecol* 54:181–207
- Möbius K (1877) Die Auster und Austernwirtschaft. Wie-gandt, Hempel und Parey, Berlin

- Niermann U (1996) Fluctuation and mass occurrence of *Phoronis mülleri* (Phoronidea) in the south-eastern North Sea during 1983–1988. *Senckenb Marit* 28:65–79
- Oksanen J, Blanchet FG, Kindt R, Legendre P and others (2015) Package 'vegan': community ecology package. R package version 2.4-4. <https://cran.r-project.org/web/packages/vegan/>
- Paine RT (1980) Food webs: linkage, interaction strength and community infrastructure. *J Anim Ecol* 49:667–685
- Pakhira MK, Bandyopadhyay S, Maulik U (2005) A study of some fuzzy cluster validity indices, genetic clustering and application to pixel classification. *Fuzzy Sets Syst* 155:191–214
- Patterson BD, Atmar W (1986) Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biol J Linn Soc* 28:65–82
- Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. *Comput Geosci* 30:683–691
- Peres-Neto PR, Olden JD, Jackson DA (2001) Environmentally constrained null models: site suitability as occupancy criterion. *Oikos* 93:110–120
- Pesch R, Propp C, Darr A, Bartholomä A and others (2016) Progress in marine biotope mapping in Germany. In: Von Nordheim H, Wollny-Goerke K (eds) *Proceedings of the 4th international conference on progress in marine conservation in Europe 2015, 14–18 September 2015, Stralsund*. Federal Agency for Nature Conservation, Bonn, p 115–120
- Petersen CGJ (1918) The sea bottom and its production of fish-food, a survey of the work done in connection with valuation of the Danish waters from 1883–1917. C Ferslew, Copenhagen
- Picard N, Köhler P, Mortier F, Gourlet-Fleury S (2012) A comparison of five classifications of species into functional groups in tropical forests of French Guiana. *Ecol Complex* 11:75–83
- Pielou EC (1975) *Ecological diversity*. Wiley & Sons, New York, NY
- Pouw FA, Kwiatkowska M (2013) An overview of fuzzy-logic based approaches to ecology: addressing uncertainty. In: *Proceedings of the IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013 Joint, 24–28 June 2013, Edmonton*. North American Fuzzy Information Processing Society, Edmonton, p 540–545
- R Development Core Team (2015) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna
- Rachor E (1990) Changes in sublittoral zoobenthos in the German Bight with regard to eutrophication. *Neth J Sea Res* 25:209–214
- Rachor E, Nehmer P (2003) Erfassung und Bewertung ökologisch wertvoller Lebensräume in der Nordsee (Survey and assessment of ecologically valuable habitats in the North Sea). Alfred Wegener Institute for Polar and Marine Research, Bremerhaven
- Rachor E, Reiss H, Duineveld GCA and others (2007) Structure, distribution, and characterizing species of North Sea macro-zoobenthos communities in 2000. In: Rees HL, Eggleton JD, Rachor E, Vander Berghe E (eds) *Structure and dynamics of the North Sea benthos*. ICES Cooperative Research Report No. 288. ICES, Copenhagen, p 46–59
- Rees HL, Eggleton JD, Rachor E, Vanden Berghe E (2007) *Structure and dynamics of the North Sea benthos*. ICES Cooperative Research Report No. 288. ICES, Copenhagen
- Reiss H, Degraer S, Duineveld GCA, Kröncke I and others (2010) Spatial patterns of infauna, epifauna, and demersal fish communities in the North Sea. *ICES J Mar Sci* 67: 278–293
- Richardson AMM (1980) Notes on the occurrence of *Talitrus dorrieni* Hunt (Crustacea Amphipoda: Talitridae) in south-west England. *J Nat Hist* 14:751–757
- Ricklefs RE (2008) Disintegration of the ecological community. *Am Nat* 172:741–750
- Ricklefs RE (2015) Intrinsic dynamics of the regional community. *Ecol Lett* 18:497–503
- Salzwedel H, Rachor E, Gerdes D (1985) Benthic macrofauna communities in the German Bight. *Veröff Inst Meeresforsch Bremerh* 20:199–267
- Somerfield PJ, Arvanitidis C, Faulwetter S, Chatzigeorgiou G and others (2009) Assessing evidence for random assembly of marine benthic communities from regional species pools. *Mar Ecol Prog Ser* 382:279–286
- Spasojevic MJ, Suding KN (2012) Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *J Ecol* 100:652–661
- Ulrich W (2004) Species co-occurrences and neutral models: reassessing JM Diamond's assembly rules. *Oikos* 107: 603–609
- Ulrich W, Gotelli NJ (2007) Disentangling community patterns of nestedness and species co-occurrence. *Oikos* 116:2053–2061
- van Dalen JA, Essink K, Madsen HT, Birklund J, Romero J, Manzanera M (2000) Differential response of macrozoobenthos to marine sand extraction in the North Sea and the Western Mediterranean. *ICES J Mar Sci* 57: 1439–1445
- van Hoey G, Degraer S, Vincx M (2004) Macrobenthic community structure of soft-bottom sediments at the Belgian Continental Shelf. *Estuar Coast Shelf Sci* 59:599–613
- Wessel P, Smith WHF (1996) A global, self-consistent, hierarchical, high-resolution shoreline database. *J Geophys Res Solid Earth* 101:8741–8743
- Whittaker RH (1952) A study of summer foliage insect communities in the Great Smoky Mountains. *Ecol Monogr* 22:1–44
- Wieking G, Kröncke I (2003) Macrofauna communities of the Dogger Bank (central North Sea) in the late 1990s: spatial distribution, species composition and trophic structure. *Helgol Mar Res* 57:34–46
- Wilson EO (1961) The nature of the taxon cycle in the Melanesian ant fauna. *Am Nat* 95:169–193
- Witt J, Schroeder A, Knust R, Arntz WE (2004) The impact of harbour sludge disposal on benthic macrofauna communities in the Weser estuary. *Helgol Mar Res* 58:117–128
- Wu KL, Yang MS (2005) A cluster validity index for fuzzy clustering. *Pattern Recognit Lett* 26:1275–1291
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8:338–353