Capturing threshold responses of marine benthos along gradients of natural and anthropogenic change

Elena Couce¹, Georg H. Engelhard¹,² and Michaela Schratzberger¹,²

¹Centre for Environment, Fisheries & Aquaculture Science (Cefas), Lowestoft NR33 0HT, United Kingdom
²Collaborative Centre for Sustainable Use of the Seas (CCSUS), University of East Anglia, Norwich NR4 7TJ, UK

Corresponding author: Elena Couce, CEFAS Laboratory, Pakefield Road, Lowestoft NR33 0HT, United Kingdom. E-mail: elena.couce@cefas.co.uk
Abstract

1. Ecologists and managers need to understand what types of communities emerge with continued human alterations to ecosystems against a background of natural change. Both natural and anthropogenic drivers are well known to affect organisms’ distributions; however it often remains unclear where along a range of environmental and anthropogenic gradients important compositional community changes occur.

2. We used a big-data approach, including over 175,000 presence records of benthic genera for the North Sea, to identify environmental (bed shear stress, sediment grain size, temperature) and anthropogenic parameters (trawling effort) driving benthic community composition over a 21-year period. We applied a Gradient Forest analysis, based on Random Forests, to estimate the locations and importance of thresholds where small cumulative increases in the predictors drive a much greater change in genus composition than would be expected from linear effects.

3. Shear stress was the most important predictor of benthic community composition. Trawling effort, temperature gradients and sediment grain size were of intermediate importance. This corroborates that current and wave effects (typically associated with seabed substrate types) are primary determinants of benthic communities.

4. Our results suggest a genus composition threshold for both infauna and epifaunal benthic communities is crossed when the seafloor is trawled as little as once every four years. Higher trawling levels corresponded with gradual compositional change without obvious thresholds, which would be consistent with chronic fishing in the North Sea over the last two centuries having caused persistent, long-term changes in ecosystem structure and functioning. This was corroborated by the large-scale spatial patterns of benthic community composition undergoing limited temporal changes during the 21-year study period.

5. Synthesis and applications. Although well established in theory, threshold effects are poorly validated in the field. We generated new information on multi-organism responses to environmental change at the scale of a continental shelf ecosystem and over a multi-decadal time period. This will help pure and applied scientists better understand the conditions under which community thresholds are
crossed and provide environmental managers with empirical evidence that is expected to reduce
uncertainty regarding decisions on the protection and sustainable use of the marine environment.

Resumen (Spanish abstract)

1. Ecologistas y administradores medioambientales necesitan entender el tipo de comunidades que
emergen en respuesta a los efectos de las acciones humanas en los ecosistemas, en un contexto de
cambio natural. Es bien sabido que tanto los efectos naturales como los antropogénicos impactan las
distribuciones de organismos; sin embargo, a menudo no está claro dónde se producen cambios
importantes en la composición de las comunidades en un gradiente de variaciones de factores
ambientales y antropogénicos.

2. Realizamos un análisis de “big data”, con más de 175000 registros de presencia de géneros
bentónicos en el Mar del Norte, para identificar parámetros naturales (tensión de cizalladura del
fondo marino, tamaño de partículas del sedimento, temperatura) y antropogénicos (intensidad de
pesca de arrastre) que influencian la composición de la comunidad bentónica en un período de 21
años. Aplicamos un análisis de Gradiente de Bosques Aleatorios (“Gradient Forests”), una técnica
basada en Bosques Aleatorios, para estimar la localización e importancia de los umbrales de
transición donde pequeñas variaciones de los predictores provocan cambios mucho mayores en la
composición de géneros de las comunidades bentónicas de lo que podría esperarse por efectos
lineales.

3. La tensión de cizalladura es el predictor más importante de la composición de la comunidad
bentónica. La intensidad de pesca de arrastre, la temperatura y el tamaño de partículas del sedimento
tienen importancia intermedia. Esto corrobora que los efectos de la corriente y las olas (muy
relacionados con los tipos de sustrato del fondo marino) son determinantes primarios de las
comunidades bentónicas.

4. Nuestros resultados sugieren que se producen cambios importantes de composición de géneros
bentónicos tanto para las comunidades epifaunales como infaunales cuando el fondo marino es
sometido a pesca de arrastre tan poco como una vez cada cuatro años. Niveles más altos de pesca de
arrastre provocan un cambio gradual de composición sin umbrales obvios, consistente con la
hipótesis de que la estructura y funcionamiento de los ecosistemas bentónicos en el Mar del Norte han sufrido cambios persistentes debido a los niveles crónicos de pesca de arrastre que han experimentado durante los dos últimos siglos. Esto es corroborado por las limitadas variaciones en la distribución espacial a gran escala de la composición de comunidades bentónica durante los 21 años del período estudiado.

5. Síntesis y aplicaciones. Aunque bien establecidos en la teoría, hay pocos datos de campo que validen la existencia de umbrales de transición de ecosistemas. En este trabajo analizamos las respuestas de múltiples organismos al cambio ambiental a la escala de una plataforma continental y durante un período de varias décadas. Nuestros resultados ayudarán a científicos teóricos y aplicados a comprender mejor las condiciones bajo las cuales se cruzan los umbrales de transición de las comunidades bentónicas, y proporcionarán a los administradores ambientales datos empíricos para apoyar decisiones sobre la protección y el uso sostenible del medio marino.

**Key-words:** benthic invertebrates, genus distribution modelling, anthropogenic and environmental pressures, gradient forest, thresholds-based reference points, North Sea

**Introduction**

Natural environmental variation and multiple human activities and uses influence species distributions in aquatic and terrestrial ecosystems world-wide (Carpentier et al., 2009; Kenny et al., 2018; Kenny, Rune, Engelhard, Kershaw, & Reid, 2009). Many ecosystems are only studied intensively whilst already being subject to anthropogenic use and consequently there are no historical data or ‘baselines’ available against which to measure and predict change.

Past perspectives have often assumed that ecosystem changes brought about by manageable human activities are readily reversible and that nature would self-repair once the activity is managed (Folke, 2006). Empirical evidence across ecosystems, however, shows that natural and human variables often interact in complex and non-linear ways and are difficult to separate (Klein, 2013; Large, Fay, Friedland, & Link, 2015). Also, ecosystems are susceptible to abrupt change when small alterations in environmental
conditions produce large responses, and the resulting change might not be easily reversible (Groffman et al., 2006; Suding & Hobbs, 2009).

Similarly to many other temperate continental shelf ecosystems, the North Sea is highly productive, but also subject to a particularly broad range of human uses, including large-scale commercial and recreational fishing, shipping, oil and gas exploration, aggregate extraction, and marine renewables such as offshore wind farms (Emeis et al., 2015; Kenny et al., 2018). Surrounded by densely populated countries with extensive agriculture, the North Sea is further subject to pollution, eutrophication and various other pressures that may lead to biodiversity loss and/or habitat degradation (Capuzzo et al., 2018). Among these diverse uses, fishing is often seen as a major driver of ecosystem change in the North Sea, although varying considerably at local scales with type of seabed substrate, and according to habitats (e.g. Hiddink et al., 2006; Kenny et al., 2018). Trawl fishing in particular has been shown to impact benthic organisms living on (epifauna) or within (infauna) the seabed (Frid, Harwood, Hall, & Hall, 2000; Jennings & Kaiser, 1998).

We use a big-data approach to identify key natural and anthropogenic parameters that drive the occurrence of benthic organisms in the North Sea. In doing so, we establish characteristics of benthic communities in response to the natural hydrodynamic environment (shear stress, sediment size), climate variables (temperature) and anthropogenic activities (fishing effort). The past two decades have seen a plethora of studies in the North Sea, designed to establish the importance of environmental factors and human activities in structuring benthic communities and species distribution. Food availability, sediment structure, hydrodynamic regime, and a variety of human activities have been invoked to influence faunal distribution on various spatial scales (Callaway et al., 2002; Clare, Robinson, & Frid, 2015). There is often high spatial heterogeneity in these variables, with either gradual or abrupt changes in space – so-called environmental gradients. It is currently unclear where along the range of these environmental gradients important compositional changes in benthic communities occur. We investigate the nature of the faunal responses to environmental gradients and determine the extent to which these gradients predict their distribution. Quantifying the shape and magnitude of organisms’ responses allows us to identify community thresholds when a small change in a gradient causes sharp increases or decreases in the occurrence of a range of different types of organisms (Connell et al., 2017; Groffman et al., 2006; Tam et al., 2017). These threshold responses are often exacerbated by species interactions, where the loss or gain of one or more key species may precipitate the loss or appearance of others (e.g., Økland, Skarphaas, & Kausrud, 2009).
Although threshold effects are well established in theory (Groffman et al., 2006), they are poorly validated in the field. The purpose of this study is to collect empirical evidence to assess the existence of community thresholds in the benthic ecosystem of the North Sea. Specifically, we

(i) assess the relative importance of natural (i.e. shear stress, sediment grain size and sea bottom temperature) and anthropogenic variables (i.e. otter and beam trawling pressure) in predicting patterns of benthic community composition,

(ii) establish where along a range of environmental and trawling pressure gradients important compositional changes in benthic communities occur to identify critical thresholds along those gradients,

(iii) examine geographical patterns in community composition of marine benthos and the main factors (natural and anthropogenic) driving them.

Understanding the conditions under which community thresholds are likely to be crossed is critical for adaptive environmental management, as exceeding them may be undesirable, irreversible and, in some cases, may limit future options for management actions.

Materials and methods

STUDY AREA

We collated biological and environmental data sets from the North Sea, a temperate coastal shelf sea with a deep channel in the northeast, a permanently thermally mixed water column in the south and east, and seasonal stratification in the north. Seasonal fluctuations of environmental variables such as temperature, salinity and stratification are generally greater in the shallower southern parts (depths 0–50 m), which are characterised by large river inputs, than in deeper water towards the north (depths up to 500 m), which is strongly influenced by oceanic inflow (Reiss et al., 2010). The North Sea is one of the most intensively exploited and studied marine ecosystems in the world (Emeis et al., 2015). The dominant human activities in the northern parts are fishing, and oil and gas production whereas anthropogenic pressures in the south result from fishing, shipping, ports, gas production, wind farms and sand extraction (ICES, 2018). The North Sea has warmed by about 1.5°C over the past 100 years, and especially rapidly during the most recent four decades – with further increases predicted for decades to come (Dye et al., 2013; Tinker, Lowe, Pardaens, Holt, & Barciela, 2016).
BIOLOGICAL DATA SETS AND ENVIRONMENTAL VARIABLES

Data of the distribution of North Sea benthic organisms between 1990 and 2011 was obtained primarily from the Ocean Biogeographic Information System (OBIS, Grassle, 2000), the largest free-access online database on the occurrences of marine species worldwide. OBIS currently hosts over 45 million observations of nearly 120,000 marine species collated from scientific surveys, national monitoring programmes, museum collections and other sources, having been validated by expert reviewers. OBIS data was supplemented with benthic survey data collected by the Centre for Environment, Fisheries and Aquaculture Science (Cefas) in the same period (accessible from www.cefas.co.uk/cefas-data-hub/; see Fig 1). Using such a large and heterogenous dataset meant that abundance values could not be directly compared, and only presence information was considered. The benthic presence data was subject to an initial clean-up process in which duplicated entries were eliminated and OBIS data on pelagic larvae of benthic species (obtained from pelagic sampling) was discarded. In order to limit this study to open waters, records at less than 2 nautical miles from the shoreline were discarded. Taxonomic information for all data was checked on the World Register of Marine Species (WoRMS, 2018) and updated when necessary.

Macrobenthic infaunal and epifaunal genera were considered separately (Table 1. See also Table S1 in Supporting Information for the complete list of all the infauna and epifauna genera). The benthic invertebrate fauna is an important component of shallow shelf seas with tight bentho-pelagic coupling, such as the North Sea, where benthic organisms play a vital role in nutrient cycling, detrital decomposition, and as a food source for higher trophic levels (Reiss et al., 2010). Resulting from their ubiquitous distribution, benthic communities are subject to a wide range of environmental conditions across different habitats. They are thus well-suited for an analysis of their responses to environmental gradients on large spatial scales. We considered all benthic genera with at least 100 presence records, in order to limit the number of hard-to-detect taxa in the study for which absences may be due to lack of suitability of the sampling method (rather than being true absences). The study region was partitioned into a 0.25° x 0.25° regular grid (Fig. 1) with all presence records aggregated into “samples” by their grid cell and the year they were recorded (Table 2; see also Table S1 for the number of samples for which each genus was present). Samples were treated as independent and linked to the predictor data of the year they were taken, so that the analysis made no distinction of compositional changes over space or time. Only samples comprising at least 20 different genera were included, in order to exclude data from surveys targeting specific taxa and from surveys where insufficient or untargeted sampling effort resulted in an incomplete inventory of genera. It was considered an infauna sample when at least 33% of all its genera...
were infaunal, and similarly for epifauna. A sample could be both infaunal and epifaunal if it contained enough genera of both groups. Infauna (epifauna) genus absences were assumed where an infauna (epifauna) sample did not include that genus.

Fig. 1. Locations of benthic samples collected in the North Sea between 1990 and 2010, overlaid on the 0.25° x 0.25° grid. Blue scale background shows water depth. Benthic data sources: Ocean Biogeographic Information System (OBIS) and Centre for Environment, Fisheries and Aquaculture Science (Cefas).

Table 1. Overview and generalisation of key characteristics for macrobenthic infaunal and epifaunal organisms included here (MarLIN, 2006).

<table>
<thead>
<tr>
<th></th>
<th>Macrobenthic infauna</th>
<th>Macrobenthic epifauna</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>Sediment-dwelling invertebrates of &gt;1 mm in size living mainly burrowed into the sea floor</td>
<td>Invertebrates of &gt;1 mm in size living mainly on the surface of the sea floor or attached to hard substrates</td>
</tr>
</tbody>
</table>
Life history characteristics

| Generally of low mobility: must withstand the extremes of their local environment or perish; generally short-lived: changes in community composition can be observed in the short-term with population changes emerging in the longer-term |
| Relatively more mobile: potential to avoid unfavourable environmental conditions at least on smaller scales; relatively longer-lived: individuals integrate environmental changes in physical, chemical and ecological characteristics of their habitat over time |

Rationale for inclusion

| Play an important role in many ecological processes; interacting organisms cover many taxonomic and functional groups that have a range of sensitivities to any given environmental regime |

**Table 2.** Description of biological data sets used in the study. Number of records refers to the total number of genera found at different locations and/or times, while a sample is defined as a unique combination of grid cell and sample year.

<table>
<thead>
<tr>
<th>OBIS</th>
<th>Cefas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Infauna</td>
</tr>
<tr>
<td>Depth range [m]</td>
<td>1 – 378</td>
</tr>
<tr>
<td>No. of records</td>
<td>97,135</td>
</tr>
<tr>
<td>No. of samples</td>
<td>9,610</td>
</tr>
<tr>
<td>No. of genera</td>
<td>179</td>
</tr>
</tbody>
</table>

We quantified the shape and magnitude of multi-genus responses along environmental gradients, including variables related to current patterns, sediment grain size, sea bottom temperature and sea bottom trawling effort. All of these variables have been shown to affect benthic community structure at various spatio-temporal scales (Hall, 1994; Jennings & Kaiser, 1998; Snelgrove & Butman, 1994). In order to provide spatially and temporally meaningful predictions of genus distributions, modelled environmental variables were used because these are available as full coverage layers at the spatio-temporal scales of the biological data.
Shear stress data were obtained from the Cefas North Sea Hindcast (van der Molen, Ruardij & Greenwood, 2015), a 20-year simulation with the hydrodynamic General Estuarine Transport Model (GETM; Burchard & Bolding, 2002) spanning 1990-2011. GETM is a fully baroclinic 3D open-source model (www.getm.eu) which includes sea surface elevations, currents, temperature and salinity. The hindcast was run using timesteps of 10 seconds and a spatial resolution of 0.08° longitude by 0.05° latitude (approximately 5.5 by 5.5 km) over the greater North Sea area. For full details of model implementation and validation of the physical component of the model (e.g., via comparison with observational data from tide gauges and current meters) see van der Molen et al. (2015). Wave-induced and current-induced components of the shear stress were considered separately, and the data can be accessed at: https://cefasbfmdata.cefas.co.uk/thredds/catalog/catalog.html?dataset=SH_BD_ST_3D.

Sea Bed Temperature (SBT) data was obtained from the “North West shelf reanalysis”, modelled using version 3.4 of the Nucleus for European Modelling of the Ocean (NEMO) ocean model code (Madec, 2008) coupled to the European Regional Seas Ecosystem Model (ERSEM; Butenschön et al., 2016). The data was generated with a spatial resolution of \(\frac{1}{9}\)° longitude by \(\frac{1}{15}\)° latitude covering the North West European shelf. It is accessible online from the online catalogue of the Copernicus Marine Environment Monitoring Service (http://marine.copernicus.eu/; dataset “NORTHWESTSHELF_REANALYSIS_PHYS_004_009”). For a detailed description of model implementation and the validation of modelled temperature via comparison with independent climatologies and satellite data products see Wakelin et al. (2016). Mean annual SBT was considered, together with the monthly average SBT of both the hottest and coldest months of the year.

Beam and otter bottom trawling effort data was derived from Couce, Schratzberger and Engelhard (2019, under review), a 31-year reconstruction of total international beam and otter trawling effort in the North Sea region at a resolution of 0.5° longitude by 1° latitude. This dataset is a compilation of various other datasets for different countries and periods largely based on logbook data, with estimates for missing data to provide a complete picture of total international trawling effort. Effort is measured as the total number of hours trawled by individual vessels per grid cell. The full dataset is available on the Cefas Data Hub at: https://doi.org/10.14466/CefasDataHub.61.

Sediment median grain size was collated from Wilson et al. (2018), a synthetic map derived from a blend of survey data, statistically modelled values derived from bed shear stress, and bathymetry, with a spatial
resolution of 0.125° x 0.125°. The full dataset can be accessed at: https://doi.org/10.15129/1e27b806-1eae-494d-83b5-a5f4792c46fc.

In order to assess multicollinearity between the different environmental variables, pairwise Pearson correlation coefficients were computed for all pairs of variables. We removed variables with multiple correlations greater than 0.70, leading to the exclusion of the mean SBT of the hottest month in the year (see Fig. S5 for the correlation coefficients between the remaining variables).

DATA ANALYSIS

We applied a Gradient Forest approach, developed by Ellis et al. (2012). This is an extension of the Random Forest analysis (Breiman, 2001), a machine-learning modelling technique based on fitting an ensemble of regression or classification trees for the prediction of the distribution of a taxon as a function of relevant environmental variables. For Random Forests each tree is trained on bootstrap samples of the training set, and each split of a tree is determined using only a randomly-chosen subset of the environmental variables. Tree-based techniques are capable of modelling nonlinear responses and interactions effects, and have been shown to be among the top performing techniques for species distribution analysis (e.g., Lawler, White, Neilson, & Blaustein, 2006).

Gradient Forest extends the Random Forest concept to model a community instead of a single taxon. Multiple Random Forest models are trained simultaneously, each fit to a different taxon. The combined results are analysed by performing standardized measurements of change along environmental gradients for all taxa simultaneously. This information is then used to build ‘response curves’: empirical functions of compositional change for each environmental variable. For each of the predictive variables, thresholds for which significant changes of community composition take place can be identified from frequency histograms of the values at which splits happen in the individual decision trees, after accounting for the distribution of the data along the gradient of the predictor. Gradient Forests also allow for the quantification of the relative importance of the predictor variables by averaging across all taxa.

The concept of community threshold used here is defined as a zone along an environmental gradient where the change in community genus composition is enhanced as a result of sharp increases or decreases in the occurrence of several genera. Therefore, Gradient Forest enabled us to identify critical values along environmental gradients that correspond to threshold changes in genus composition. The
Gradient Forest analysis was carried out in R (R Development Core Team, 2009), with package `gradientForest` (Ellis et al., 2012).

Results

We applied the Gradient Forest approach to the presence/absence benthic data and the environmental and anthropogenic predictive data to estimate the location and importance of benthic community composition thresholds along environmental gradients in the North Sea.

RELATIVE IMPORTANCE OF NATURAL AND ANTHROPOGENIC VARIABLES IN PREDICTING BENTHIC COMMUNITY COMPOSITION

All environmental and trawling variables combined predicted 52% and 51% of the variation in occurrence of infaunal and epifaunal genera, respectively (obtained by adding the $R^2$-square values for all predictors; Fig. 2). Due to correlation between explanatory variables (Fig. S5) care should be taken to interpret the $R^2$ values in Fig. 2 to assess relative contributions. Variables related to current patterns (i.e. wave- and current-induced shear stress) ranked high in their influence on both infaunal and epifaunal benthic genera composition, a result consistent across multiple variations of the analysis using different subsets of explanatory variables. In the case of infauna, beam trawling pressure ranked second amongst the variables examined, and third for epifauna. Other variables were of lesser importance (Fig. 2).
SHAPE AND MAGNITUDE OF COMPOSITIONAL CHANGES ALONG ENVIRONMENTAL GRADIENTS AND LOCATION OF COMMUNITY COMPOSITION THRESHOLDS

The compositional change of infauna along environmental gradients is shown in Fig. 3 for each genus and averaged over all genera. Ldots are standardised by density of observations. Each genus curve (e.g., blue line) in the figure is scaled by the individual goodness-of-fit $R^2$ values. Simultaneous steep slopes for many genera in Fig. 3 indicate thresholds where the community is undergoing significant change. Although responses of individual genera clearly vary in shape and magnitude, suggesting a mosaic of genus-specific responses to environmental gradients, they are consistently non-linear.
Fig. 3. Compositional change along each environmental gradient for all infauna genera. Each blue line denotes a genus while the thicker black line denotes the average (e.g., the overall pattern of genus compositional change along the gradient). The y-axes have been normalised so that the maximum corresponds to the relative variable importance. Individual plots are arranged (left to right) from the most to the least important predictor.
Frequency histograms of the values used by the classification trees for splits (i.e., the splits density plots in Fig. 4) help to quantify thresholds. These plots indicate where along the environmental gradients community changes are happening. However, the histograms of splits can be biased if environmental data is not uniformly distributed along the gradient, as is typically the case. To visualise this, the red lines in Fig. 4 show the data distribution (normalised to 1), and the blue line shows the ratio between the density of splits (black line) and the density of the data. The gradient values at which the blue line reaches values above 1 are suggestive of thresholds, particularly when this is not driven by mere lack of data (i.e., when the red line is close to 0).

Along the gradient for otter trawling no obvious thresholds for infauna are apparent, as the density of splits largely matches the distribution of the data. Conversely, a clear threshold exists for beam trawling of 0-4,000 hours, suggesting that the most significant infaunal compositional changes occur between samples with no recorded beam trawling activity that year and those that have experienced any beam trawling effort, even at very low levels. Community changes brought about by additional increases in trawling effort are more gradual. An important SBT threshold lies at around a mean annual SBT of 8.5°C, with notable community changes occurring at a degree above and below that temperature, and at 7°C for the SBT of the coldest month. Along a gradient of shear stress, conditions of high relative community change were observed around 1e-04 to 0.01 J/Kg for wave-induced shear stress and 0.014, 0.07 and 0.15 to ~0.25 J/Kg for current-induced shear stress, respectively. A less significant compositional threshold occurred at median grain sizes around 0.4 mm.
Fig. 4. Frequency histograms of gradient values at which splits occur in the regression trees for all infauna genera, showing where along environmental gradients important compositional changes are taking place. Black lines are the kernel density of the histograms, red lines show the (normalised) distribution of the data along the environmental gradients, and blue lines indicate the ratio between splits and data (ratio between black and red lines). Thus ratios >1 (above the dotted line) indicate conditions of relatively
greater change in genus composition (i.e. community thresholds). Individual plots are arranged (left to right) from the most to the least important predictor.

GEOGRAPHICAL PATTERNS IN COMMUNITY COMPOSITION OF MARINE BENTHOS
The modelled relation between diversity patterns and explanatory variables (i.e., the response curves in Fig. 3) can be used to produce predictive maps in geographical space of changes in benthic communities and identify the most important variables driving those changes (e.g., see Pitcher et al., 2011 for details of the procedure and a worked-out example). This process is similar to a community composition analysis, but captures the non-linear responses of Random Forest models. Fig. 5 illustrates the model predictions for the year 2000, obtained after re-interpolating the environmental variables for that year to match the study grid. Several distinctive regions and the associated drivers are apparent: the deeper region near the Norwegian coast with community composition influenced mainly by SBT of the coldest month, the southern area governed by both mean SBT and beam trawling effort, the north west sector with a strong influence by SBT of the coldest month that gradually gives way to other explanatory variables acting at more local scales as we advance westwards towards the UK coast. To illustrate temporal variation, Fig. S3 shows the predicted spatial patterns for the beginning, middle and end of the 21-year study period (years 1990, 2000 and 2010).
Fig. 5. Changes of infauna community composition predicted by the Gradient Forest analysis for the environmental and trawling pressure data for the year 2000. These changes have been mapped over the first two dimensions of a biologically transformed environmental space that accounts for their respective influence in dictating compositional patterns (together capturing 80% of the explained variance). The colour key shows the environmental variables driving those compositional changes (longer arrows denote stronger influences; arrows for median grain size and otter trawling are not shown because their contribution is less significant).

Figs. 3-5 refer only to the infauna analysis, the equivalent plots for epifauna show similar patterns (Figs. S1, S2 and S4).

Discussion

Marine and coastal data are collected by many organisations, across a range of academic, industry and public sectors, covering a wide variety of natural science disciplines. Some of these data are difficult and expensive.
to obtain. Despite the ever-present need for more data, there is an equally strong need to make better use of existing data to generate new information on multi-organism responses to environmental change. The fundamental importance of documenting and understanding ecological patterns at large spatio-temporal scales has come increasingly to the fore in recent decades because it has become clear that many of the pressing issues facing ecosystems are operating on large spatial and temporal scales. Using Gradient Forest analyses, we integrated quantitative results of disparate surveys across the North Sea spanning 21 years to delineate compositional changes in benthic communities along environmental gradients.

METHODOLOGICAL CONSIDERATIONS

This study depended on spatially and temporally extensive data on benthos and environmental variables compiled from many, sometimes disparate, data sources, which necessarily raises concerns over data quality and consistency. The OBIS dataset is a large open-access dataset which compiles data from many different sources, including scientific, commercial and citizen science (Grassle, 2000). These datasets might suffer from various biases, such as sampling bias favouring more populated areas near coastal zones, reporting bias favouring well-known, charismatic or otherwise significant species, or identification bias. Moreover the data would have been collected with a variety of sampling and observation methods, and its quality is not consistent, despite standard quality control measures that may be applied (e.g., see Vandepitte et al., 2015). The long-term trawling data, compiled from a variety of sources, might also have data quality and consistency issues (e.g., inconsistencies between countries and time periods in the minimum sizes of fishing vessels and in types of trawling gear included in the compilation; for discussion see Couce et al. [in review]). While acknowledging these challenges, the use of large-scale, heterogeneous datasets also makes it possible to carry out studies over much larger areas and longer time periods than would have been possible if the study would have been limited to data collected under rigorous sampling protocols.

Benthic faunal composition can be expected to change significantly over the course of a year due to processes such as seasonal recruitment, growth and/or migrations. These changes will have increased the variability of our dataset and influenced the detectability of the responses of benthic communities along environmental gradients. However, this also was an inevitable consequence of using a spatially and temporally extensive data set combining observations from disparate sources rather than a significant limitation of our study itself. Further, if the effect of environmental and anthropogenic pressure gradients is obscured by temporal variability, this would imply that any detected ecosystem threshold refers to
community composition changes which are more significant than those caused by seasonal variability alone.

Assessing the relative importance of predictive variables is problematic when some of these are correlated. For example, mean SBT was moderately correlated with current and wave-induced shear stress and with beam trawling effort \((r > 0.6; \text{Fig. S5})\). While it ranked last in importance among predictive variables (Fig. 2), some of its real importance may have been erroneously assigned to the others. Shear stress variables invariably ranked at the top of the \(R^2\)-weighted average of predictors for versions of the model trained with different subsets of explanatory variables, and had the most significant impact over the two dimensional plane of the first two main components explaining modelled changes in community composition (black arrows in Fig. 5; notice also wave and current-induced shear stress components have nearly orthogonal effects). Beam trawling effort also tended to rank high in importance across all model versions, but the relative importance of the remaining variables was less consistent and remains difficult to establish.

THRESHOLD RESPONSES OF BENTHOS ALONG ENVIRONMENTAL AND ANTHROPOGENIC GRADIENTS

Benthic communities in any sedimentary habitat have morphological, physiological and life history adaptations to a naturally occurring disturbance regime, including physical perturbations such as waves and currents. Physical perturbations of anthropogenic origin (e.g. fishing) are therefore likely to have a more significant and detectable impact if they exceed the background levels of intensity and frequency of natural physical disturbance (Kaiser et al., 2006). Using modelled shear stress and fishing effort in the Greater North Sea, Diesing et al. (2013) demonstrated that natural disturbance exceeded disturbance generated by fishing in about half the area characterised by soft sediments. This, in combination with results from our Gradient Forest analyses suggests that the response of benthic genera in our study area to fishing disturbance was greatly influenced by the extent to which these genera and their communities are preconditioned to disturbance by natural processes, and in particular waves and currents (Hall, 1994).

The combined forces waves and currents exert on the seafloor (i.e. shear stress) shape the physical habitat of benthic invertebrates by altering the surface sediment texture and micro-topography of sedimentary environments, thereby changing their suitability for settling organisms and/or by increasing the resuspension of organisms as a result of sediment deposition and/or sediment instability (Heath, Sabatino, Serpetti, McCaig, & O’Hara Murray, 2017; Ward, Neill, Van Landeghem, & Scourse, 2015). Along a gradient of shear stress, we observed high relative changes in genus assemblages at locations where
waves and currents were strong enough to scour away fine sediments exposing more consolidated substrata (as identified by Dalyander, Butman, Sherwood, Signell, & Wilkin, 2013). A less significant compositional threshold was also seen for values of sediment median grain size close to 0.4 mm.

Whilst these are not new findings, our results emphasise the limitations of previous attempts to explain spatial distributions of organisms from physical conditions measured in studies that are spatially and temporally more limited. There have been many scientific studies, both in fishing grounds and using experimental trawling, investigating the impacts of trawling on benthic communities at selected stations in shelf seas (see Lengkeek & Bouma, 2010 and references therein). Individual studies often fail to detect trawling-specific community changes in areas exposed to high levels of natural hydrodynamic disturbance. This has led to the conclusion that natural and trawling-induced physical disturbance affect benthic communities in comparable ways and that trawling disturbance in hydrodynamically active environments causes limited additional changes in benthic communities (Lambert et al., 2017; Lengkeek & Bouma, 2010; Van Denderen et al., 2015). Attempts to determine how consistent and applicable these study-specific conclusions are at larger spatio-temporal scales have been hampered by the lack of reliable estimates of the historic trawling effort and therefore relied on meta-analyses of existing studies (e.g. Collie, Hall, Kaiser, & Poiner, 2000; Kaiser et al., 2006; Sciberras et al., 2018).

Using a 31-year reconstruction of international beam and otter trawling effort (Couce et al., in review) we demonstrate for the first time that, at the scale of the North Sea, natural physical disturbance regimes dominate in structuring benthic communities, potentially confounding changes in the occurrence of many genera along gradients of beam trawling effort. Beam trawling alters seabed complexity, removes, damages or kills biota, and reduces benthic production, thereby causing changes in benthic ecosystems (Sciberras et al., 2018 and references therein). We show that the effects of such changes on benthic genera are particularly persistent and, even against a background of varying levels of natural physical disturbance, already measurable at low levels of beam trawling effort, i.e. between 0 and 4000 vessel fishing hours per year per spatial rectangle of 1° longitude by 0.5° latitude (approximately 3000 km²). Considering a typical swept area for beam trawlers of approximately 0.2 km² per hour (Eigaard et al., 2016), this equates to a probability of 25% per year, implying the compositional threshold is crossed when the seafloor is subject to the impact of a beam trawl even only once in four years. The lack of steep compositional changes in response to higher levels of trawling (Fig. 4) could be indicative of a benthic community of resilient genera. Such communities would respond primarily to natural environmental gradients, similar to communities present in areas subjected to chronic beam trawling for the last two
centuries, where ecosystem structure and functioning has been altered significantly (Callaway, Engelhard, Dann, Cotter, & Rumohr, 2007). Being able to pinpoint threshold responses of communities to trawling effort is particularly relevant given the drive towards ecosystem approaches to fisheries management that integrate the wider ecological effects of fishing (Pikitch et al., 2004).

The main factor determining the occurrence of benthic genera is the physical habitat. Where a habitat is very restricted in occurrence, the distribution of a genus will generally reflect occurrence of the habitat and may not be primarily influenced by other factors such as temperature (Hiscock, Southward, Tittley, & Hawkins, 2004). Human activities such as trawling can amplify or mediate the effects of temperature on benthic communities, and so can diverse biological requirements not considered in this study (including distribution of predators, food availability etc.). A previous study by Hiddink et al. (2015) tracking thermal niches of benthos in the North Sea, showed that extended distribution of southern species alongside retreat in the distribution of northern species are the most likely effects of increased seabed temperatures resulting from climate change. However, the rate at which, or whether, change occurred varied greatly from species to species. Results from our Gradient Forest analyses pinpointed threshold responses of benthic communities at approximately 1°C above and below a mean annual seabed temperature of 8.5°C for both infaunal and epifaunal genera (Figs. 4 and S2). An additional threshold at 12°C was particularly evident for epifaunal genera (Fig. S2). Alterations in the occurrence of many genera at these thresholds were most likely a result of changes in the development of eggs or other propagules, survival of larval stages, survival of post-settlement juveniles and survival of adults (Hiscock et al., 2004).

Geographical patterns in benthic community composition changed moderately over the 21-year study period (Figs. S3 and S4). The representation of the variation in genus composition and the main predictive variables driving it reflects the strong constraints the physical environment in the Greater North Sea exerts on the occurrence of benthic genera, and how changes in these constraints are driving changes in community composition. This is critical for applications such as spatial planning because biophysical limits of the marine environment set the ultimate boundaries within which sustainability goals must be achieved. Integrating quantitative biological response information into mapped environmental variables, as presented here, will help to target and facilitate future survey designs.

Conclusions

In order to avoid unexpected changes to the structure and function of communities, ecologists and managers need to obtain a more robust understanding of the types of communities that will emerge with
continued human alterations to ecosystems caused by perturbations such as fishing and climate change, and against a background of natural change (Beisner, Haydon, & Cuddington, 2003). Here we demonstrate that small cumulative changes in natural environmental parameters (e.g., shear stress), climate variables (e.g., sea bottom temperature) and anthropogenic activities (e.g., bottom trawling) can drive much larger responses than can be predicted from linear effects. The resulting communities thresholds can act as early-warning signals of loss of resilience, leading to potentially massive change in ecosystem state which may come at a cost to society. Knowledge of the specific values when thresholds are expected to be crossed can additionally help managers understand when benthic ecosystem changes are due to manageable or non-manageable pressures.

Authors' contributions
M.S. and E.C. conceived the research idea, E.C. designed the methodology and carried out the data collation and analysis, M.S., E.C. and G.E. participated in the interpretation of results and the writing of the manuscript and gave final approval for publication.

Acknowledgements
We thank Johan van der Molen for creating the GETM Cefas hindcast spanning 1958-2008 and the additional years of 2009-2011, and Sonja van Leeuwen and Luz Garcia for making the shear stress data available to us. The work was funded by the UK Department for Environment, Food and Rural Affairs via project SLA44: Marine Biodiversity Advice, and by the Natural Environment Research Council and Department for Environment, Food and Rural Affairs via grant NE/L003279/1 (Marine Ecosystems Research Programme). The manuscript also benefitted greatly from comments and suggestions from Associate Editor Rute Pinto, Ruth Callaway, Murray Thompson, and two additional anonymous reviewers.

Data availability statement
All datasets used to train the Gradient Forest models are available online: the processed benthic genera presence/absence dataset via the Cefas Data Hub at https://doi.org/10.14466/CefasDataHub.94 (Couce, Schratzberger, & Engelhard, 2020); shear stress data via the Cefas THREDDS Catalog (https://cefasbfmdata.cefas.co.uk/thredds; dataset “Shear Stress exerted by Waves and Currents on Sea Bed”, with dataset id “SH_BD_ST_3D”), Sea Bottom Temperature via de COPERNICUS repository (http://marine.copernicus.eu/; dataset “NORTHWESTSHELF_REANALYSIS_PHYS_004_009”); beam and otter trawling effort reconstruction data via the Cefas Data Hub at https://doi.org/10.14466/CefasDataHub.61 (Couce, Schratzberger, & Engelhard, 2019); sediment median grain size via the University of Strathclyde at https://doi.org/10.15129/1e27b806-1eae-494d-83b5-a5f4792c46fc (Wilson, Heath, Speirs, & Sabatino, 2017).
References


This article is protected by copyright. All rights reserved


Klein, E. S. (2013). *Change in nonlinear dynamics and spatial structure of coastal socio-ecological systems: Bay of Fundy as case study*. University of New Hampshire, Durham, NH, USA.


model differences and model reliability. Global Change Biology, 12(8), 1568–1584. doi:10.1111/j.1365-2486.2006.01191.x


Wilson, R., Heath, M., Speirs, D., & Sabatino, A. (2017). Data for: “A synthetic map of the northwest European Shelf sedimentary environment for applications in marine science.” University of Strathclyde. doi:10.15129/1e27b806-1eae-494d-83b5-a5f4792c46fc


This article is protected by copyright. All rights reserved