



Causal links between North Sea fish biomass trends and seabed structure

Mark Rademaker^{1,2,*}, Isabel M. Smallegange², Anieke van Leeuwen¹

¹NIOZ Royal Netherlands Institute for Sea Research, Department of Coastal Systems, and Utrecht University, 1797 SZ Texel, The Netherlands

²Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1012 WX Amsterdam, The Netherlands

ABSTRACT: Distinct areas of seabed, called seascapes, are known to shape benthic habitats and communities, yet little is known about the extent to which they affect the dynamics of marine fish populations. We explored the relationship between seascapes and trends in the biomass density of several North Sea fish species. We divided the North Sea into 10 seascapes using standardized methods. Time series of fish biomass density were derived from the North Sea International Bottom-Trawl Survey (NS-IBTS) and aggregated to the seascape level. We analysed the interdependencies between these time series using a causal association network and found independent biomass density trends between adjacent seascapes at a time interval of 0 yr in all species assessed. Long-term causal dependencies in biomass density occurred at time lags of 1–2 yr across different gradients of exchange: (1) both directions from North to South; (2) unidirectional, North–South; (3) unidirectional, South–North; (4) unidirectional, East–West; and (5) no clear direction. Our findings indicate that the separation in (a)biotic conditions between North Sea seascapes can represent relevant barriers to the processes determining the observed fish biomass density. We found that non-fusiform morphology and demersal habitat preferences best explained short-term causal dependencies. This combination is particular to the flatfish and ray species included in the present study. Contrarily, the movement of large, long-lived, benthopelagic species best explained long-term causal dependencies. Our work highlights how causal association networks can be used to study the temporal dependencies between spatial time series in ecology.

KEY WORDS: IBTS · International Bottom-Trawl Survey · PCMCI · Seascapes · Spatiotemporal analysis · Time series · Bottom-trawl

1. INTRODUCTION

The dynamics of marine fish populations are typically categorized as regulated primarily by top-down processes (i.e. predation) or by bottom-up processes (i.e. resource production) (Wollrab et al. 2012, Heath et al. 2014, Vinuela et al. 2014). For example, fish population dynamics in the North Sea were long thought to be driven by bottom-up processes (Lynam et al. 2017). In bottom-up regulated systems, the lowest-level driver is formed by abiotic factors (e.g. temperature, salinity, etc.) and their influence on

plankton biomass. The available plankton biomass in turn leads to cascades upwards through the food web (Gregory et al. 2009, Kirby & Beaugrand 2009, Olsen et al. 2011). Top-down interactions, represented by predation pressure from predators and fisheries exploitation, have also been shown to have a large impact on the dynamics of several North Sea fish populations (Fauchald et al. 2011, Engelhard et al. 2014). Furthermore, the interaction between top-down and bottom-up processes in the North Sea ecosystem can give rise to highly complex and non-linear dynamics in fish populations (Lynam et al.

*Corresponding author: mark.rademaker@nioz.nl

2017). To attain a better understanding of the complexity emerging from such interactions, it is important to have a complete overview of the bottom-up and top-down processes acting upon fish populations in the North Sea.

One potentially important bottom-up process that is often overlooked in regulating fish populations is seabed morphology. Physically distinct areas of seabed, also called seascapes, can act as a bottom-up controlling mechanism by shaping abiotic conditions, leading to the formation of different types of benthic habitat and associated species assemblages (Harris & Baker 2012). Due to the strong link between seabed morphology and community structure of benthic invertebrates (Beaman et al. 2005, Kaskela et al. 2017), seascapes are often used as a proxy for mapping marine benthic habitats (Kostylev et al. 2001, Brown & Collier 2008). Yet, we know little about the extent to which seascapes affect the dynamics of fish populations, despite the fact that the few case studies available show it could help explain the spatial distribution of fish resources (Pittman & Brown 2011, Getsiv-Clemons et al. 2012). Recent research revealed that the central and lower part of the North Sea comprises 8 distinct seascapes excluding the coastal zone (van der Reijden et al. 2018), and that these were selectively targeted by commercial fishing vessels. Although this preferential selection provides some indirect link to the availability and distribution of fish resources, a more complete study exploring the relationship between seascapes and the population trends of North Sea fish is missing. If seascapes are important, this changes the commonly held perspective on the types of bottom-up processes affecting North Sea fish populations.

To understand the relationship between seascapes and the population trends of North Sea fish, we investigated the temporal interdependencies in fish biomass density between seascapes. We hypothesized that if seascapes are not related to trends in fish populations in the North Sea, biomass densities in separate seascapes will be directly dependent on each other. Ecologically, this result could be interpreted as the 'borders' between separate seascapes being diffuse, i.e. not relevant, to the processes determining the amount of fish biomass observed. In contrast, if seascapes define distinct abiotic environments that are important for the distribution of fish populations, we expect biomass densities in different seascapes to be unrelated to each other, although they can still be related across longer time lags. Ecologically, this result could be interpreted as the separation in conditions between seascapes representing relevant

barriers to the processes determining the amount of fish biomass observed.

To examine the temporal interdependencies in fish biomass density between seascapes, we used a recently developed causal association network approach (Runge et al. 2019a). The advantage of this data-driven method is that it learns causal associations at various time lags directly from aggregated sets of time series, rather than establishing causality through computer simulation experiments on correlative relationships found in observational data (Runge et al. 2019b). As such, our study also tests a new approach to examine intraspecific dependencies in the biomass trends of North Sea fish.

2. MATERIALS AND METHODS

2.1. Study area

The North Sea is a semi-enclosed shelf extending from 4° W and 62° N (Fig. 1a) (Daan et al. 2005). In this study, we focussed on the region that includes International Council for the Exploration of the Sea (ICES) statistical rectangles 4a, b and c. We excluded the Norwegian trench, the Skagerrak and the waters south of the English Channel, as their distinctiveness is well recognized (Bergstad 1990, ICES 2018). Extending the method of van der Reijden et al. (2018), we derived 10 separate seascapes (Fig. 1b). These seascapes were identified based on the weighted sum of 5 bathymetry positioning indices (BPIs), which express the depth of a pixel relative to the depth of its surroundings. The specific weighted-sum BPI used (Eq. 1) highlights large-scale patterns in seabed morphology and is strongly linked to benthic communities (Reiss et al. 2010). A full description of the weighted BPI can be found in van der Reijden et al. (2018).

$$\text{BPI 75 km} + (0.9 \times \text{BPI 50 km}) + (0.8 \times \text{BPI 30 km}) + (0.7 \times \text{BPI 10 km}) + (0.6 \times \text{BPI 5 km}) \quad (1)$$

2.2. Data preparation

We extracted data on the catch per unit effort (CPUE) of 9 species of fish collected in the North Sea International Bottom Trawl Survey (NS-IBTS) from 1978 to 2019 (ICES 2019) (Table 1). We chose 1978 as the starting year of the time series as all seascapes were systematically sampled within the NS-IBTS from this year onward. The 9 selected species covered a variety of different life history strategies and

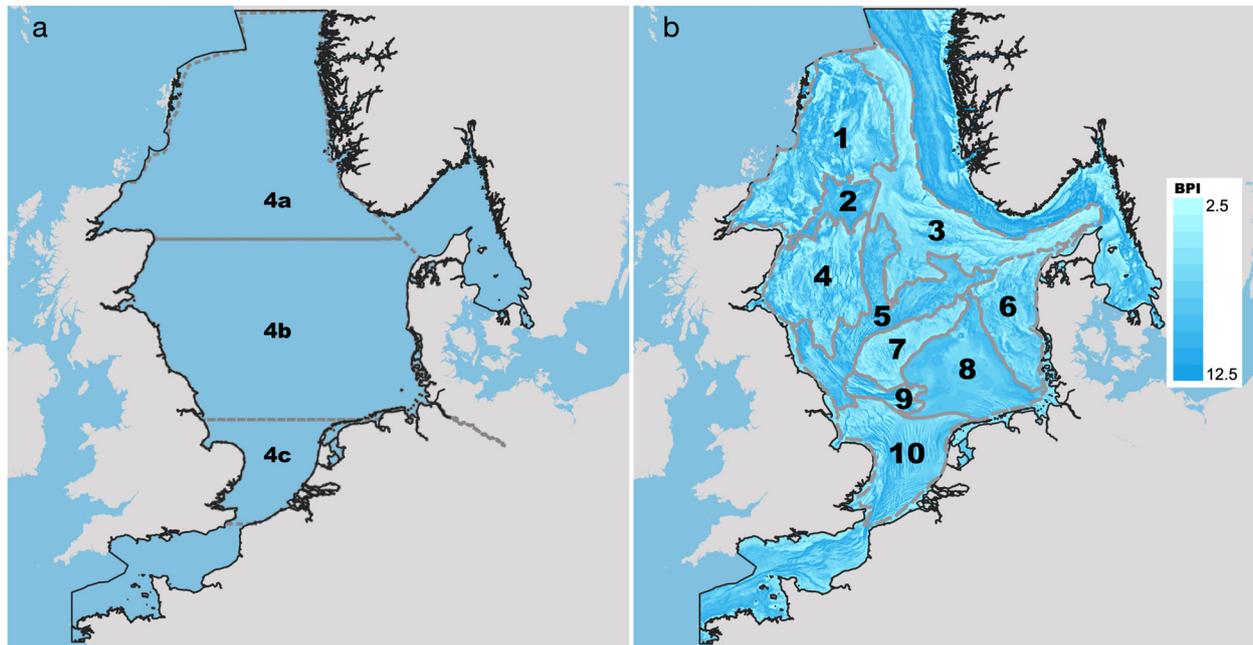


Fig. 1. Greater North Sea Ecoregion (black coastline) (a) divided into ICES statistical rectangles 4a, 4b and 4c (boundaries are indicated by the grey dashed lines), and (b) its division into 10 seascapes based on seabed morphology, following the method of van der Reijden et al. (2018). BPI: bathymetry positioning index, which expresses the depth of a pixel relative to the depth of its surroundings

had sufficient positive catch data available in all seascapes across selected years. To minimize bias from potential variability in sampling gear and timing, we only included observations that had been sampled in the same quarter of the year and using the GOV trawl. Next, we determined in which seascape each haul was performed based on the shoot longitude and latitude. We estimated the biomass in kg for a given species for each haul by multiplying the number of individuals of a given length class by their length–weight relationship, as derived from FishBase (Froese & Pauly 2019), and summing over all length

classes present in the haul. We then aggregated the data to a single time series per seascape and species, representing the average biomass in $\text{kg haul}^{-1} \text{yr}^{-1}$. We refer to this metric as biomass density.

2.3. Data analysis

2.3.1. Causal association network

We define temporal interdependencies in population trends as the existence of time-lagged causal

Table 1. Fish species, records and length to weight conversion used in the study. All length–weight (L–W) relationships were taken directly from FishBase (Froese & Pauly 2019). Slight differences in the number of filtered (unique) hauls between species originate from small differences in the raw North Sea International Bottom Trawl Survey CPUE-length-per-hour datasets pulled from the ICES/DATRAS online database (ICES 2019) for each species (mean \pm SD: $13\,667.2 \pm 104.9$ filtered hauls, range 13 430–13 760 filtered hauls)

Species	Scientific name	Filtered hauls	Non-zero hauls	L–W relationship	<i>n</i>
Common dab	<i>Limanda limanda</i>	13 430	11 847	$0:0068(L)^{3.14}$	19
Whiting	<i>Merlangius merlangus</i>	13 760	13 364	$0:0063(L)^{3.06}$	33
Atlantic herring	<i>Clupea harengus</i>	13 746	12 085	$0:0060(L)^{3.08}$	114
Atlantic cod	<i>Gadus morhua</i>	13 743	11 197	$0:0071(L)^{3.08}$	42
European sprat	<i>Sprattus sprattus</i>	13 747	8135	$0:0056(L)^{3.09}$	36
Dragonet	<i>Callionymus lyra</i>	13 682	3825	$0:0214(L)^{2.59}$	12
European plaice	<i>Pleuronectes platessa</i>	13 615	9216	$0:0089(L)^{3.04}$	14
Grey gurnard	<i>Eutrigla gurnardus</i>	13 624	8607	$0:0079(L)^{3.02}$	15
Starry ray	<i>Amblyraja radiata</i>	13 658	4308	$0:0105(L)^{2.94}$	6

dependencies in fish biomass density between seascapes. If seascapes are not related to fish populations in the North Sea, we expect that the time series in fish biomass densities between separate seascapes are causally related to each other at the time interval of zero. Contrarily, if seascapes are related to fish populations in the North Sea, we expect that the time series in fish biomass densities in separate seascapes are independent of each other at the time interval of zero, although they can be causally dependent across longer time scales.

We applied a causal association network approach recently developed by Runge et al. (2019a), to evaluate the underlying causal dependency structure between the fish biomass density time series from different seascapes (Fig. 2). The linear and non-linear time-lagged dependencies between sets of time series can cause inflated or spurious associations to be included when analysing their interdependence (Pearl 2009). These need to be accounted for in order to arrive at an approximation of the ‘true’ underlying dependency structure. The causal association network utilizes the PCMCI algorithm, as implemented in the package ‘TIGRAMITE’ in Python (Runge 2019), that arrives at a solution of the underlying dependency structure through iterative conditional independence testing. Internally, the PCMCI algorithm takes 2 processing steps. First, given a set of time series, the PC1 algorithm constructs a graphical model in which all time series and their time-lagged components form separate nodes. For each of these nodes or time series variables, the PC1 algorithm then identifies all relevant conditions or ‘preliminary parents’ $\hat{P}(X_i^j)$ in the set $X_i^j \in \{X_1^1, \dots, X_N^N\}$, through iterative conditional independence testing. This conditional dependence can be interpreted as the relation between time series variables 1 and 2, after accounting for the

relations existing between all time series variables within the set. The time series variables with significant relations to the focal time series after accounting for all relationships in the set form the preliminary parents of that focal time series. This first step reduces dimensionality, i.e. it identifies all relevant connections in the graph for a given time series variable and a first estimation of their strength, and will thereby increase the detection power for underlying causal relations. In the second step, the MCI algorithm addresses the false positives, i.e. spurious or inflated links found between the time series variables. It does so by repeating the iterative conditional independence tests, but now conditioning on both the preliminary parents of X_i^j and the time-shifted parents of $X_{i-\tau}^j$ (Runge et al. 2019b). Thus the graphical model contains the focal time series variable, its preliminary parents and their preliminary parents. This second step is therefore well suited to identifying spurious or inflated relationships originating from common drivers. Those time series variables with significant relations to the focal time series variable after this second iterative testing phase are included in the final causal association network. The final causal association network thus represents the algorithm’s approximation of the underlying causal dependency structure existing between the set of time series.

Our time series covered 42 yr, which is extensive in ecology, but relatively limited in data science. We therefore used the linear partial correlation test, a maximum time lag of 2 yr, and let the PC1 algorithm choose the optimal α level for condition selection. As the linear partial correlation test is sensitive to outliers, we applied a 95% winsorization to the data prior to analysis (Reifman & Keyton 2010). Confidence intervals were generated using bootstrapping, and inferences on the conditional dependence be-

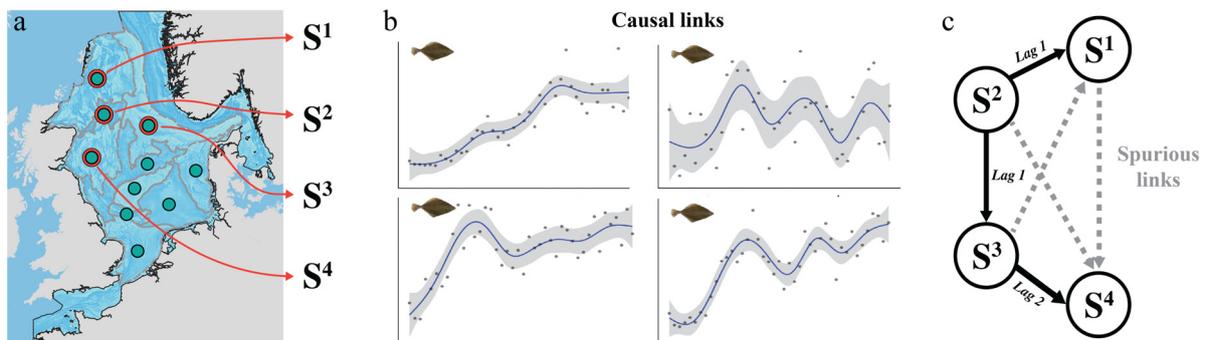


Fig. 2. (a) Each of the 10 seascapes (green circles) has (b) an associated time series of fish biomass density, e.g. S1–S4, of which we estimate the underlying causal dependency structure in (c) a causal association network. Common drivers and indirect effects can yield spurious associations that are accounted for in our approach. The dots in panel b represent observed biomass density values, the blue line is a fitted trend with 95% confidence bounds indicated in gray shading. Figure adapted from Runge et al. (2019a)

tween time series were made using a shuffle test with 3000 trials. We further corrected the results for the total number of pairwise tests performed by applying the false discovery rate procedure of Benjamini & Hochberg (1995). The final causal association network included those links between time series significant at an α level of 0.05. If links occurred between 2 time series at multiple lags, the strongest link was included in the network. Rather than printing the full causal association network found for each species, which would result in a cluttered and unclear overview, we report the results as follows: (1) For each species, we report the number of seascapes with biomass density trends independent of their adjacent seascapes at time lag zero. This provides a link back to our initial hypothesis that the contrasting conditions in seascapes are reflected in contrasting biomass density trends. (2) We produce a species-specific map where relationships at time lag zero are indicated by the borders between adjacent seascapes, and relationships over longer time lags between all seascapes are indicated by arrows.

2.3.2. Post-hoc test to link life histories to causal relationships

We conducted a post-hoc test to better understand the variables driving the observed dependencies in biomass density between seascapes at different time lags. More specifically, we used a random forest classifier (RFC) model to predict the presence of causal relationships between seascapes based on a set of 17 variables pertaining to life history, the biomass density of different life stages and environmental conditions (see Table S1 in the Supplement at www.int-res.com/articles/suppl/m677p129_supp.pdf).

2.3.3. Variable collection

We extracted 9 features related to life-history strategies from the marine fish trait database of Beukhof et al. (2019). Next, for every year and seascape combination, we calculated the absolute difference in biomass density of small and large size classes in the juvenile and adult life stages. To do this, we first determined the length at maturation to split the IBTS data between adult and juvenile biomass density. Next, for each of these groups, we made an equal split between the 50% largest and 50% smallest individuals in terms of length. We then aggregated the data into a single data frame listing the biomass

density of the 4 groups (small juveniles, large juveniles, small adults, large adults) in each haul. Next, we computed the mean biomass density of each group per haul per year. Finally, we calculated the difference in mean biomass density per haul per year between different seascapes for each group. This dataframe was combined with the life-history trait dataframe to produce the final feature dataset for training. As there was a class imbalance in the label data between the number of samples with and without causal lag, we used a cost-sensitive learning approach and weighted the classes according to their representation in the dataset (Weiss et al. 2007).

2.3.4. Model training and performance

We further optimised hyperparameters of the RFC through a grid search on the maximum depth, minimum samples per split, minimum samples per leaf and the number of trees. The final RFC model with causal lag zero as the label feature included 300 trees, with a maximum depth of 40 splits, a minimum 10 samples per leaf and split. The final RFC model with causal lag 1–2 as the label feature consisted of 500 trees, with a maximum depth of 40 splits and a minimum of 10 samples per leaf and 15 per split. Before training, we split 50% of the data into a validation set, and, during training, we used stratified K-fold cross validation with 5 splits (Refaeilzadeh et al. 2009). We used the area under the receiver operating characteristic curve (AUC) values and a confusion matrix, from which we could calculate recall and precision, to evaluate model performance. The model was run twice, once with causal lag zero as the label feature, and once with causal lag 1–2 as the label feature. The model performed well on the validation set in both runs, with an AUC of 98.0 and 98.2%, a recall of 95.2 and 97.1% and precision of 93.0 and 91.6%, respectively. This is further illustrated in the confusion matrixes in Fig. S1.

2.3.5. Variable importance

We calculated the importance of each of the 17 variables in predicting the presence of a causal dependency between seascapes at various time lags using the 'TreeExplainer' function in the 'SHAP' package (Lundberg et al. 2020). 'TreeExplainer' is based on game theory and approximates Shapley values to compute the contribution of a target variable to the predicted output of tree-based machine learning

the absence of a causal dependency between fish biomass density trends in adjacent seascapes at a time lag of zero. This finding provides the insight

that seascapes can function as a bottom-up driver of North Sea fish population dynamics. However, a considerable level of caution is warranted in interpreting

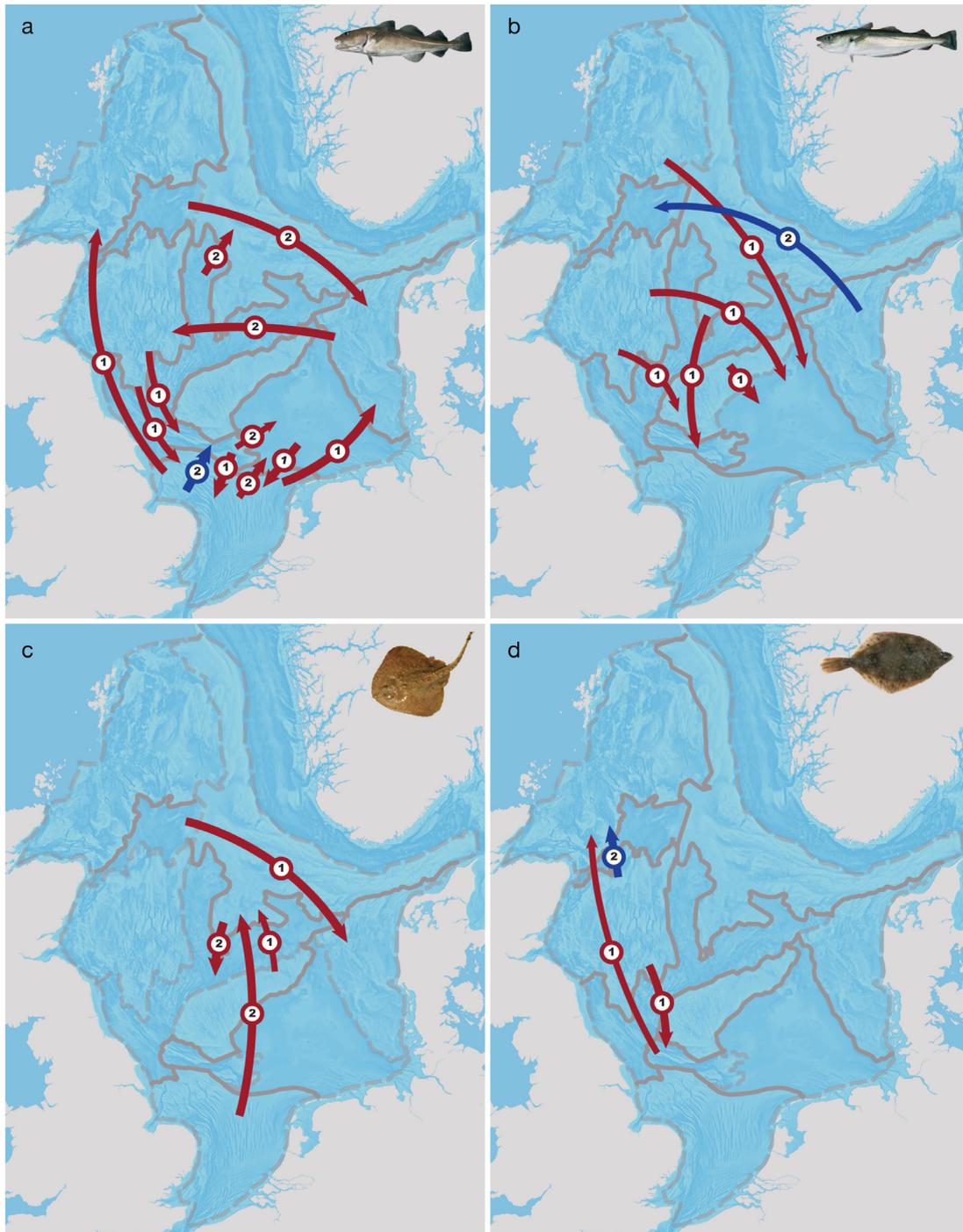


Fig. 3. Causal relationships in mean fish biomass density ($\text{kg haul}^{-1} \text{yr}^{-1}$) between different seascapes for (a) Atlantic cod, (b) whiting, (c) starry ray and (d) European plaice. In all 4 species, time-lagged exchanges occur predominantly in both directions between more northern and southern seascapes. Arrow colour and width represent the type and strength of the relationship (red = positive, blue = negative, thin = weak, thick = strong) between different areas. The associated time-lag is listed on the arrow, and if a contemporaneous relationship was present, the border between the areas was removed

the results, as the relative standard error in estimated biomass density over time frequently exceeded 30% for species and seascape combinations (Table S2).

We further found both positive and negative causal dependencies between seascapes at time lags of 1–2 yr in most species. Species such as Atlantic cod, European plaice, whiting and starry ray displayed clear directional gradients in these longer-

term dependencies along a North–South gradient. Positive long-term dependencies can potentially be explained by both movement of adult biomass to, and recruitment of juvenile biomass in, the dependent seascapes. Negative long-term dependencies cannot be explained by the process of recruitment. However, they can still be explained by movement of biomass away from a dependent seascape to sur-

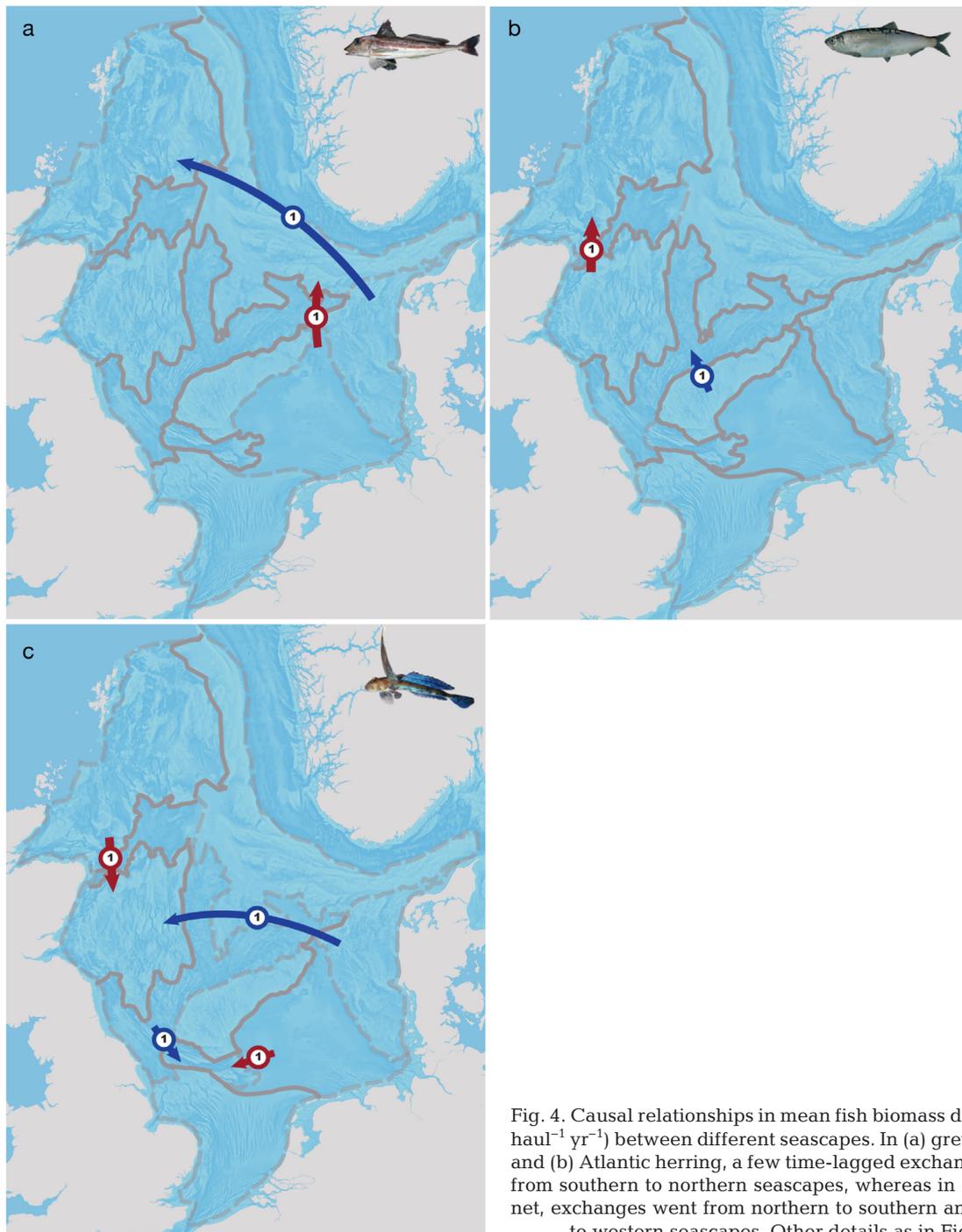


Fig. 4. Causal relationships in mean fish biomass density ($\text{kg haul}^{-1} \text{yr}^{-1}$) between different seascapes. In (a) grey gurnard and (b) Atlantic herring, a few time-lagged exchanges occur from southern to northern seascapes, whereas in (c) dragoonet, exchanges went from northern to southern and eastern to western seascapes. Other details as in Fig. 3

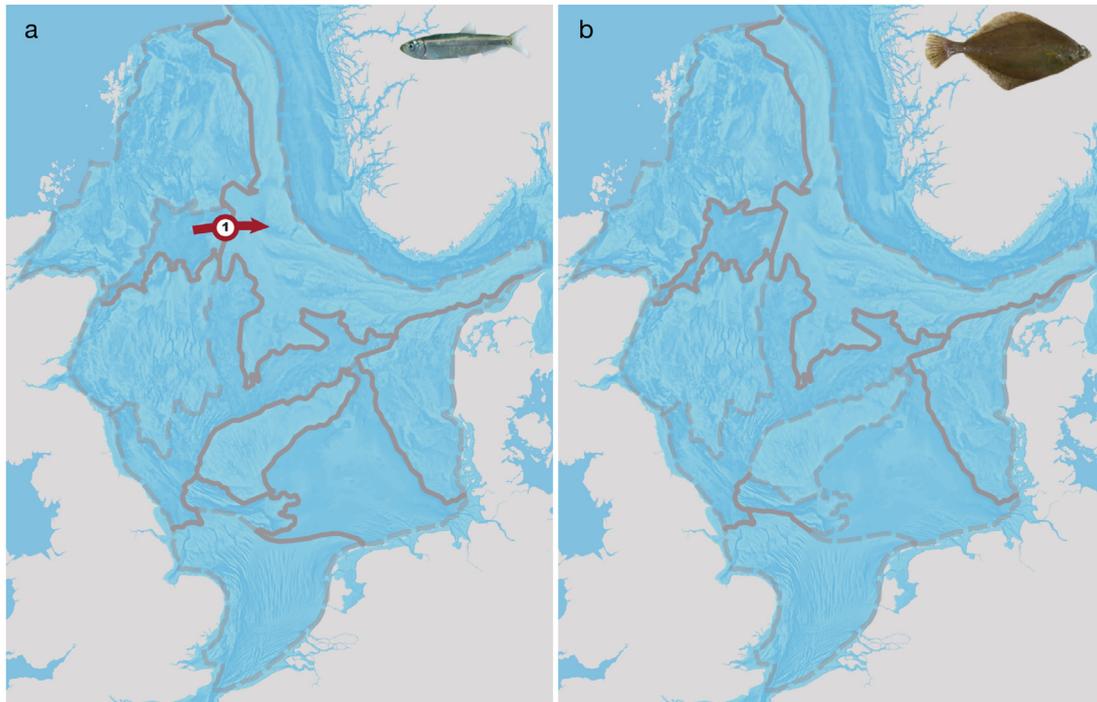


Fig. 5. Causal relationships in mean fish biomass density ($\text{kg haul}^{-1} \text{yr}^{-1}$) between different seascapes. (a) In European sprat, there is a single time-lagged exchange and (b) in common dab, there is no time-lagged exchange; both species therefore lack a clear gradient of exchange. Other details as in Fig. 3

rounding areas following accumulation of biomass in the driving seascape. The dependent seascape could in this case be considered suboptimal, and the driving seascape more optimal. This would be the inverse process of that described by the basin model theory (McCall 1990), in which there is adult movement away from optimal basins, or seascapes in this case, to suboptimal surrounding basins following the buildup of biomass in the former (Bertrand et al. 2008, Hintzen et al. 2014).

Our post-hoc test provided a deeper understanding of the ecological meaning of the observed dependencies at different time lags. We found that non-fusiform morphology and demersal habitat preferences best explained short-term causal dependencies. This combination is particular to the flatfish and ray included in the present study. This might be due to multiple factors. Flatfish such as European plaice exhibit more restricted movement patterns and stronger selection of small-scale variations in sediment type compared to benthopelagic species such as Atlantic cod (Gibson & Robb 2000, Hinz et al. 2006, Griffiths et al. 2018). Therefore, smaller-scale processes than those delineated by the seascapes might be more important in these species at the 0 yr time lag. Long-term causal dependencies in biomass density between seascapes were best explained by the adults

of large, long-lived, benthopelagic species. In these species, the recruitment of juveniles is unlikely to drive the observed dynamics because differences in juvenile biomass negatively affected the probability of long-term causal dependencies.

In this study, we also addressed a missing link between seascape-specific fisheries activities in the North Sea (van der Reijden et al. 2018) and the biomass of harvested stocks. van der Reijden et al. (2018, their Fig. 3) showed fishing intensity with a beam trawl to target plaice was highest in seascapes 10 and 7, and the northern part of seascape 6, while otter-mix fishing was highest in the centre of seascape 9 and along the northern and southern borders of seascape 8. If these fishing activities affect the distribution of fish biomass, we might have expected to see some effect of this in the post-hoc test. An example would be that short-term exchanges in (adult) biomass between intensively and less intensively fished adjacent seascapes are important in explaining the observed causal dependencies. However, no specific combination of seascape pairs was found to be important in explaining the observed dependencies. Although it might be tentative to place an explicit link between seascape-specific fisheries activities and exchanges in fish biomass, separate quantitative analysis, including North Sea-wide fishing intensity,

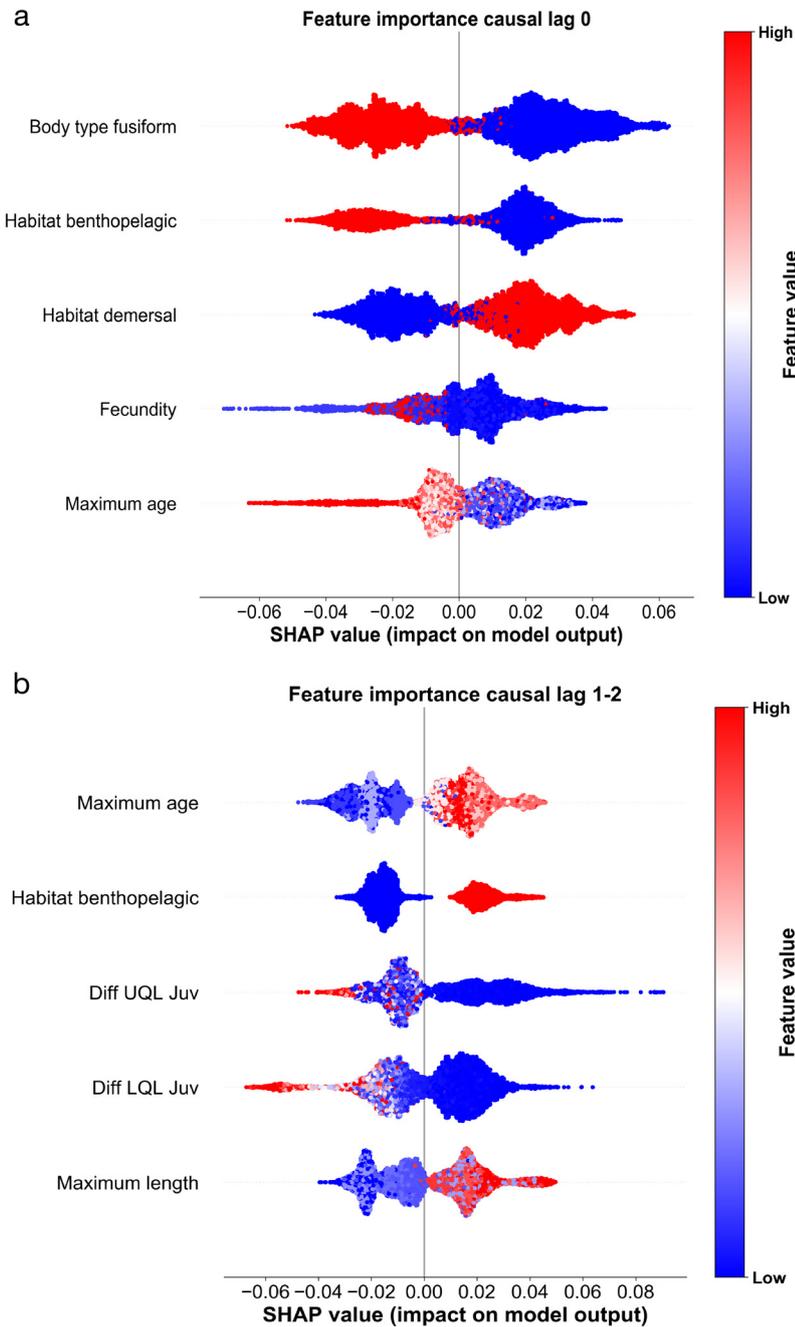


Fig. 6. Ordered variable importance plots for random forest classifier models predicting the presence of a causal relationship between seascapes at a time-lag of (a) 0 yr or (b) 1–2 yr. Each dot represents an individual sample from the validation dataset. Dot colour indicates the value of the variable. The position on the x-axis indicates the impact on the predicted probability of a causal lag existing between seascapes. SHAP: SHapley Additive exPlanations; Diff UQL Juv: difference in biomass density of large juveniles (individuals in the upper 50% body length quantile); Diff LQL Juv: difference in biomass density of small juveniles (individuals in the lower 50% body length quantile)

would be required to provide more conclusive evidence. The study of van der Reijden et al. (2018) was based on Dutch fishing fleet data. Consequently, sea-

scapes that are assumed to be free from high fishing pressure, based on Dutch data, might turn out to be intensively used by the fleets of other nations.

The methodological novelty of our study lies in the application of a causal association network for hypothesis testing using ecological time series. These types of data are typical of long-term monitoring programmes, and causal association networks therefore have a potentially wide applicability in ecology. The advantage of this specific causality-based approach is that it explicitly accounts for inflated or spurious correlations between time series (Pearl 2009). Using a method that accounts for this confounding effect differentiates our work from other, recent research examining dependencies between time series of fish populations and the environment. For example, Milligan et al. (2020) examined time-lagged dependencies between resource productivity and abundances in deep-sea fish using additive models in combination with cross-correlation analysis. Additive models are a popular tool in ecology because they can include the effects of categorical and continuous covariates in fitting statistical trends to population time series (Zuur et al. 2017, Pedersen et al. 2019). Cross-correlating such fitted statistical trends to extract time-lagged dependencies increases the risk of finding spurious relationships inherent to time series analysis (Yule 1926, Olden & Neff 2001, Runge et al. 2014). Although causal association networks explicitly account for this confounding effect, they are not able to include both categorical and continuous covariates *a priori*. Additional insights can be gained indirectly through post-hoc testing as was done in the present

study. However, this approach does not provide the explanatory power on the effects of environmental covariates that additive modelling provides. Therefore, a decision on which combination of methods to apply in future ecological

studies will depend on both (1) the types of environmental effects a researcher wants to include in examining trends and (2) the need to explicitly model time-lagged dependencies.

In conclusion, our study provides an indication that seascapes can function as a bottom-up driver of North Sea fish population dynamics. The presence of causal dependencies between seascapes was best explained by different types of life history traits at different time lags. In a qualitative comparison with previous research, we did not find a clear link between seascape-specific fishery activities and biomass trends. Finally, the suitability of using causal association networks to model spatial time series in ecology will depend both on the types of effects a researcher wants to include and the need to model time-lagged dependencies.

Acknowledgements. We thank Karin van der Reijden for explaining seascape delineation and Misha Zhemchuzhnikov for his advice on data visualization.

LITERATURE CITED

- Beaman RJ, Daniell JJ, Harris PT (2005) Geology–benthos relationships on a temperate rocky bank, eastern Bass Strait, Australia. *Mar Freshw Res* 56:943–958
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300
- Bergstad OA (1990) Ecology of the fishes of the Norwegian Deep: distribution and species assemblages. *Neth J Sea Res* 25:237–266
- Bertrand A, Gerlotto F, Bertrand S, Gutiérrez M and others (2008) Schooling behaviour and environmental forcing in relation to anchoveta distribution: an analysis across multiple spatial scales. *Prog Oceanogr* 79:264–277
- Beukhof E, Dencker TS, Palomares ML, Maureaud A (2019) A trait collection of marine fish species from North Atlantic and Northeast Pacific continental shelf seas. PANGAEA. <https://doi.pangaea.de/10.1594/PANGAEA.900866>
- Brown CJ, Collier JS (2008) Mapping benthic habitat in regions of gradational substrata: an automated approach utilising geophysical, geological, and biological relationships. *Estuar Coast Shelf Sci* 78:203–214
- Daan N, Gislason HG, Pope JC, Rice J (2005) Changes in the North Sea fish community: evidence of indirect effects of fishing? *ICES J Mar Sci* 62:177–188
- Engelhard GH, Peck MA, Rindorf A, Smout SC and others (2014) Forage fish, their fisheries, and their predators: Who drives whom? *ICES J Mar Sci* 71:90–104
- Fauchald P, Skov H, Skern-Mauritzen M, Johns D, Tveraa T (2011) Wasp-waist interactions in the North Sea ecosystem. *PLOS ONE* 6:e22729
- Froese F, Pauly D (2019) FishBase: the length–weight table. https://www.fishbase.de/manual/English/FishbaseThe_LENGTH_WEIGHT_table.htm (accessed 10 October 2019)
- Getsiv-Clemons JE, Wakefield WW, Whitmire CE, Stewart IJ (2012) Identifying potential habitats from multibeam echosounder imagery to estimate abundance of groundfish: a case study at Heceta Bank, OR, USA. In: Harris PT, Baker EK (eds) *Seafloor geomorphology as benthic habitat: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam, p 569–586
- Gibson R, Robb L (2000) Sediment selection in juvenile plaice and its behavioural basis. *J Fish Biol* 56:1258–1275
- Gregory B, Christophe L, Martin E (2009) Rapid biogeographical plankton shifts in the North Atlantic Ocean. *Glob Change Biol* 15:1790–1803
- Griffiths CA, Patterson TA, Blanchard JL, Righton DA, Wright SR, Pitchford JW, Blackwell PG (2018) Scaling marine fish movement behavior from individuals to populations. *Ecol Evol* 8:7031–7043
- Harris PT, Baker EK (2012) Why map benthic habitats? In: Harris PT, Baker EK (eds) *Seafloor geomorphology as benthic habitat: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam, p 3–22
- Heath MR, Speirs DC, Steele JH (2014) Understanding patterns and processes in models of trophic cascades. *Ecol Lett* 17:101–114
- Hintzen N, Corten AAHM, Gerlotto F, Habasque J and others (2014) Hydrography and jack mackerel stock in the South Pacific – final report. Tech Rep C176.14. IMARES, IJmuiden
- Hinz H, Bergmann M, Shucksmith R, Kaiser MJ, Rogers SI (2006) Habitat association of plaice, sole, and lemon sole in the English Channel. *ICES J Mar Sci* 63:912–927
- ICES (2018) Greater North Sea Ecoregion—ecosystem overview. https://www.ices.dk/sites/pub/Publication%20Reports/Advice/2018/2018/GreaterNorthSeaEcoregion_EcosystemOverview.pdf (accessed 28 April 2021)
- ICES (2019) ICES Data Portals: DATRAS. Public trawl survey data. https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx (accessed 10 October 2019)
- Kaskela A, Rousi H, Ronkainen M, Orlova M and others (2017) Linkages between benthic assemblages and physical environmental factors: the role of geodiversity in Eastern Gulf of Finland ecosystems. *Cont Shelf Res* 142:1–13
- Kirby RR, Beaugrand G (2009) Trophic amplification of climate warming. *Proc R Soc B* 276:4095–4103
- Kostylev VE, Todd BJ, Fader GBJ, Courtney RC, Cameron GDM, Pickrill RA (2001) Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Mar Ecol Prog Ser* 219:121–137
- Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. In: *Proceedings of the 31st international conference on neural information processing systems*. NIPS, Long Beach, CA, p 4768–4777
- Lundberg SM, Erion G, Chen H, DeGrave A and others (2020) From local explanations to global understanding with explainable AI for trees. *Nat Machine Intellig* 2:56–67
- Lynam CP, Llope M, Möllmann C, Helauouët P, Bayliss-Brown GA, Stenseth NC (2017) Interaction between top-down and bottom-up control in marine food webs. *Proc Natl Acad Sci USA* 114:1952–1957
- McCall A (1990) *Dynamic geography of marine fish populations*. Washington Sea Grant program. University of Washington Press, Seattle, WA
- Milligan RJ, Scott EM, Jones DOB, Bett BJ and others (2020) Evidence for seasonal cycles in deep-sea fish abundances: a great migration in the deep SE Atlantic? *J Anim Ecol* 89: 1593–1603
- Molnar C (2020) Interpretable machine learning—a guide

- for making black box models explainable. <https://christophm.github.io/interpretable-ml-book/>
- ✦ Olden JD, Neff BD (2001) Cross-correlation bias in lag analysis of aquatic time series. *Mar Biol* 138:1063–1070
- ✦ Olsen EM, Ottersen G, Llope M, Chan KS, Beaugrand G, Stenseth NC (2011) Spawning stock and recruitment in North Sea cod shaped by food and climate. *Proc R Soc B* 278:504–510
- Pearl J (2009) *Causality: models, reasoning and inference*, 2nd edn. Cambridge University Press, Cambridge
- ✦ Pedersen EJ, Miller DL, Simpson GL, Ross N (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ* 7:e6876
- ✦ Pittman SJ, Brown KA (2011) Multi-scale approach for predicting fish species distributions across coral reef seascapes. *PLOS ONE* 6:e20583
- Refaeilzadeh P, Tang L, Liu H (2009) Cross-validation. In: Liu L, Özsu MT (eds) *Encyclopedia of database systems* 5. Springer, Boston, MA, p 532–538
- Reifman A, Keyton K (2010) Winsorize. In: Salkind NA (ed) *Encyclopedia of research design* 3. Sage Publications, Thousand Oaks, CA, p 1636–1638
- ✦ 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
- ✦ Runge J (2019) TIGRAMITE—causal discovery for time series datasets. <https://github.com/jakobrunge/tigramite/> (accessed 20 January 2019)
- ✦ Runge J, Petoukhov V, Kurths J (2014) Quantifying the strength and delay of climatic interactions: the ambiguities of cross correlation and a novel measure based on graphical models. *J Clim* 27:720–739
- ✦ Runge J, Nowack P, Kretschmer M, Flaxman S, Sejdinovic D (2019a) Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci Adv* 5: eaau4996
- ✦ Runge J, Bathiany S, Bollt E, Camps-Valls G and others (2019b) Inferring causation from time series in Earth system sciences. *Nat Commun* 10:2553
- Shapley LS (1953) A value for n-person games. In: Kuhn HW, Tucker AW (eds) *Contributions to the theory of games*, Vol 2. Princeton University Press, Princeton, NJ, p 307–318
- ✦ van der Reijden KJ, Hintzen NT, Govers LL, Rijnsdorp AD, Olf H (2018) North Sea demersal fisheries prefer specific benthic habitats. *PLOS ONE* 13:e0208338
- ✦ Vinuela LR, Menge BA, Ruiz D, Palacios DM (2014) Oceanographic and climatic variation drive top-down/bottom-up coupling in the Galápagos intertidal meta-ecosystem. *Ecol Monogr* 84:411–434
- Weiss GM, McCarthy K, Zabar B (2007) Cost-sensitive learning vs. sampling: Which is best for handling unbalanced classes with unequal error costs? In: Stahlbock R, Crone SF, Lessman S (eds) *Proceedings of the 2007 International Conference on Data Mining, DMIN 2007*, June 25–28, 2007, Las Vegas, Nevada, USA, p 35–41
- ✦ Wollrab S, Diehl S, De Roos AM (2012) Simple rules describe bottom-up and top-down control in food webs with alternative energy pathways. *Ecol Lett* 15:935–946
- ✦ Yule GU (1926) Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *J R Stat Soc* 89:1–63
- Zuur AF, Ieno EN, Saveliev AA (2017) *Beginner's guide to spatial, temporal and spatial-temporal ecological data analysis with R-INLA*. Vol II: GAM and zero-inflated models. Highland Statistics, Newburgh

Editorial responsibility: Konstantinos Stergiou, Thessaloniki, Greece
Reviewed by: E. Johnsen and 2 anonymous referees

Submitted: September 1, 2020
Accepted: July 27, 2021
Proofs received from author(s): September 27, 2021