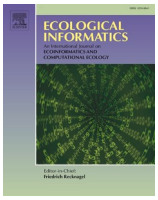
Ecological Informatics 92 (2025) 103510

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Propagating uncertainty from physical and biogeochemical drivers through to top predators in dynamic Bayesian ecosystem models improves predictions

Neda Trifonova [a,*], Juliane Wihsgott [b], Beth Scott [a]

[a] School of Biological Sciences, University of Aberdeen, Tillydrone Avenue, Aberdeen AB24 2TZ, UK
[b] Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK

ARTICLE INFO

Keywords:
Climate change
Hidden variable
Functional ecosystem change
Bio-physical
Indicators
Marine renewable energy

ABSTRACT

With the global rapid expansion of offshore renewable energies, there is an urgent need to assess and predict effects on marine species, habitats, and ecosystem functioning. Doing so will require dynamic, multitrophic, ecosystem-centric approaches coupled with oceanographic models that can allow for physical and/or biogeochemical indicators of marine ecosystem change to be included. However, in such coupled approaches, indicators carry uncertainties that can propagate and affect species higher up the trophic chain. Dynamic Bayesian networks (DBNs) are pragmatic approaches that probabilistically represent ecosystem-level interactions. They allow for uncertainties to be better estimated than mechanistic models that only account for expected values. In this study, we calculated variance as a measure of uncertainty from selected indicators and used them to build DBN models. A hidden variable was incorporated to model functional ecosystem change, where the underlying interactions dramatically change, following a disturbance. We wanted to assess whether propagating uncertainty into the modelling process affects the predictive accuracy of the models in the context of reconstructing the time series of the ecosystem dynamics. Model accuracy was improved for 60 % of the species once variance was added. The models were better in capturing the temporal inter-annual variability, once variance was calculated with a rolling window approach. The hidden variable successfully modelled previously identified ecosystem changes, however, now with the added uncertainty, the changes that implicated the ecosystem state were identified earlier in the time series. The results indicate that using DBNs is highly valuable as it gains accuracy with the addition of uncertainty.

1. Introduction

Marine ecosystems consist of complex dynamic interactions among species and the environment, the understanding of which has significant ecological and societal implications for predicting nature's response to changes in climate and biodiversity (Barange et al., 2014; García Molinos et al., 2016). Such interactions are further exacerbated by spatial and temporal variation of the ecosystem and its components (Doney et al., 2012; Hunsicker et al., 2011; Polis et al., 1996). Stressors such as, climate change, fishing, and resource exploitation have also been shown to modify the driving forces in ecosystems (Blanchard et al., 2012; Cheung et al., 2019; Lotze et al., 2019). Understanding and disentangling the drivers of ecosystem change can be challenging because of the variability in observations, for example due to imperfect methods of observation and uncertainty in potential associations due to external forces like climate change (Link et al., 2012). Such complexities are a major challenge for modellers, particularly as data are often rather scarce due to the relatively high costs of collecting field data, the practical difficulties of collecting samples from all parts of the ecosystem, and the lack of scientific understanding about the entire range of factors that may be relevant to the ecosystem functioning.

Significant progress has been made in developing ecosystem models that use traditional statistical approaches to understand the relationships between several variables (Lynam et al., 2017), including "end-to-end" ecosystem models to predict impacts of environmental change on the structure and function of marine food webs and the services they provide (Heath et al., 2021). However, all these models assume that the underlying functional relationships do not change their form over time.

* Corresponding author.
E-mail address: neda.trifonova@abdn.ac.uk (N. Trifonova).

https://doi.org/10.1016/j.ecoinf.2025.103510
Received 8 July 2025; Received in revised form 1 October 2025; Accepted 5 November 2025
Available online 6 November 2025

This assumption might not be true, as ecosystems are known to sometimes undergo relatively fast structural changes that have a major effect on the ecosystem dynamics (Möllmann et al., 2008). Further, it is possible that the changes are driven by unobserved components, i.e. ecosystem variables that we do not have data for. Allowing for uncertainties in the modelling process rather than just accounting for expected values is also an improvement to deterministic models which might remain fraught with errors when applied to problems with real data (Wikle, 2003). Thus, it is recommended that ecosystem models develop fuller non-mechanistic appreciation of ecological interactions across space and over time due to changing pressures at different levels of the trophic chain (Uusitalo et al., 2018).

Approaches using time series of multispecies population characteristics, as well as both physical and biological ecosystem components are useful methods to understand the drivers of ecosystem change, such that patterns of species population change can be quantified across space and over time, under different climate and/or anthropogenic scenarios (Lynam et al., 2017). In particular, coupling physical dynamics from high resolution oceanographic models (e.g., Finite Volume Community Ocean Model (FVCOM)) into ecosystem models can allow for critical physical (e.g., temperature) and/or biogeochemical (e.g., oxygen) indicators of marine ecosystem change to be included. However, all modelled and to some extent observed physical variables carry inherent biases and uncertainties. Additionally, model resolution and boundary forcing may contribute further to these errors (Powley et al., 2020). In coupled physics-ecosystem models, these errors can propagate through dependent physical and/or biogeochemical parameters and may affect variables higher up the trophic chain.

Other sources of error could come from data being assembled from different spatial and temporal scales, for example, fish stock data in one-year resolution, whilst climate data, such as sea surface temperature available on a daily resolution. The spatial and temporal scale of physical and biological indicators is a challenging issue with respect to understanding the multiplicity of mechanisms underlying observed patterns and variability changes (Levin, 1992; Wiens, 1989, 1990) and especially the trophic interactions of highly mobile marine animals (fish, seabirds, and marine mammals) within dynamic marine environments. However, the inclusion of physical and biological indicators is essential as they can be significant drivers of variation, or direction of change, in either behaviours, distributions and/or population dynamics of the highly mobile top predator marine species, thus delivering an understanding of the indicators of habitat and ecosystem change (Trifonova et al., 2022). For example, one of several possible physical indicators to measure stratification is the potential energy anomaly (PEA, J/m^3). PEA represents the amount of work required to bring about complete vertical mixing per unit of volume (Simpson and Bowers, 1981) and indicates the strength of stratification and mixing rate (De Boer et al., 2008; De Dominicis et al., 2018). The seasonal cycle of stratification underpins primary production cycles. Recent modelling outcomes reveal that PEA plays a significant role in predicting the abundance changes of both lower (e.g., sandeel larvae) and higher trophic level (e.g., harbour porpoise, black-legged kittiwake) marine species on a regional spatial scale (Carroll et al., 2015; Trifonova et al., 2021) and to a lesser extent, in determining habitat preferences on a North Sea scale (Sadykova et al., 2017; Wakefield et al., 2017). PEA can reflect more subtle spatial and temporal changes within a habitat type (Van Leeuwen et al., 2016) and season (Simpson and Bowers, 1981), thus further highlighting the importance of spatial and seasonal distribution of physical processes as good indicators up through the entire trophic chain and any changes that are affecting ecosystem functioning.

Being able to propagate uncertainty into ecosystem models can be very useful when applied to environmental challenges with real data. Understanding how levels of uncertainty affect the predictive accuracy of the ecosystem models could provide more insight into which variables are the causes and even if certain variables are relevant at all in efforts to reconstruct ecosystem dynamics. Most importantly, explicit accounting for uncertainty can add substantial practical insight to many real-life problems that can aid communicating theories and results to industry and policy (Uusitalo, 2007).

One way of dealing with uncertainty in environmental domains is using statistical indicators related to variability, autocorrelation and recovery time (Carpenter et al., 2011; Scheffer et al., 2009). Such approaches have been demonstrated as new tools for understanding nonlinear dynamics in ecosystems, thus revealing new indicators of vulnerability and improving ecosystem management in a rapidly changing environment. Previous studies have demonstrated the use of variance and autocorrelation in the early detection of data patterns that govern the temporal ecological dynamics (e.g., ecosystem shift: Scheffer et al., 2009). Specifically, studies have used an increase in the mean and increase in the variance in the Quickest detection method to account for the expected ecosystem shift and uncertainty, respectively (Carpenter et al., 2014). Also, it has been demonstrated that such metrics can improve the predictive accuracy of ecosystem models when trying to predict functional changes, i.e., regime shifts (Trifonova et al., 2014). To understand and predict ecosystem response to perturbation, it is necessary to unravel the ecological networks underlying ecosystem's stability and fragility (Dunne et al., 2002). However, identifying all the interactions and quantifying all the unexpected effects and interactions due to external pressures within complex real ecosystems can be rather challenging and beyond the scope of traditional fieldwork (Aderhold et al., 2012).

Computational inference of ecological interactions presents an alternative route to unravel ecosystem dynamics. Specifically, one way forward of dealing with these issues is to use probabilistic methods such as Bayesian networks (BNs) that can be used to capture ecological patterns between variables (Hui et al., 2022) and reveal spatiotemporal trends (Tucker and Duplisea, 2012), without requiring specific information on mechanisms and vast amounts of observational data used in traditional ecosystem models (Uusitalo, 2007). Modelling time series is achieved by using an extension of the BN known as the Dynamic Bayesian Network (DBN) which allow predictions to be made across different spatial and temporal scales and with a range of indicator species or functional groups representing all trophic levels (Trifonova et al., 2015). A hidden variable can be used to enable the modelling of non-stationary dynamics (Tucker and Liu, 2004), which is potentially highly useful in ecological analyses where complex ecological interactions change in time due to changing pressures at different levels of the trophic chain. Its value depends on all the observed variables it is linked to, and a change in the pattern of the hidden variable indicates a change in the system interactions. BNs use probability as a measure of uncertainty: beliefs about values of variables are expressed as probability distributions, and the higher the uncertainty, the wider is the probability distribution. As information accumulates, knowledge of the true value of the variable usually increases, i.e., the uncertainty of the value diminishes and the probability distribution grows narrower (Gelman et al., 1995; Sivia and Skilling, 2006).

In this study, we focused on providing a method to propagate uncertainty (i.e., calculated as variance) from a set of physical and biological indicators that included critically important factors of ecosystem change (e.g. stratification, primary production, temperature; Trifonova et al., 2021). Understanding how levels of uncertainty affect the predictive accuracy of the ecosystem models could provide more insight into which indicators are more relevant when evaluating ecosystem structure and function in efforts to determine the ecosystem state. In this way, we wanted to provide a pragmatic yet powerful methodology that can be used within marine spatial planning considerations of the relevant implications of future climate change versus anthropogenic impacts (e.g., offshore large-scale wind developments). Firstly, the indicators were used to build dynamic hidden BN models, and we wanted to assess whether bringing in uncertainty into the modelling process would affect the predictive accuracy of the models in the context of reconstructing the ecosystem dynamics. To be able to do so, we used a

machine learning optimization technique to find the data-driven estimates of interactions among the physical and biological indicators. We used the learned data-driven interactions to construct a dynamic BN, i.e. one that explicitly represents the behaviour of the system over time, that incorporates a hidden variable to enable the modelling of non-stationary dynamics. A hidden variable was incorporated in the model to see whether we can detect a change in the interactions of the observed variables over time. Therefore, we wanted to see whether the hidden variable can be used to model changes in the ecosystem state. We then calculated the variance as a measure of uncertainty for selected physical and biological indicators and used it to define their conditional probability distributions when learning the model variants of the hidden dynamic BN model. We used a rolling window approach to calculate variance and autocorrelation for all the ecosystem components to build a separate model variant. We examine the models' accuracy in terms of their ability to reproduce observations of the trends (increases versus decreases) in all the ecosystem components. We evaluate the potential usefulness of Bayesian inference for ecosystem-level interactions by examining whether using statistical metrics, such as variance and autocorrelation, improves the predictive accuracy and modelling of the ecosystem state.

2. Materials and methods

2.1. Study region and ecosystem components

Focus of the study are UK coastal waters and specifically the region of Firth of Forth, as defined by the spatial boundaries in Fig. 1, which currently contains Scotland's largest operational offshore wind farm. The marine environment in this area is very complex due to a composite bathymetry exhibiting localized shelf banks, suggesting that the region might be subjected to small-scale processes defining a "mosaic" of heterogenous hydrodynamic conditions (Zampollo et al., 2025). The Forth is known to support overwintering populations and juveniles of estuarine fish communities and represents a significant percentage of the commercial activity in the North Sea (Elliott et al., 1990). The Isle of May, located about 40 km east of the Firth of Forth, is known to be a National Nature Reserve hosting >250,000 breeding seabirds and a population of breeding grey seals that feed upon eastern productive offshore waters (Harris and Wanless, 1998).

The time series input data consisted of annual values (1990–2022) as either seasonal mean values of physical variables (e.g. temperature) or cumulative values of biological variables (e.g. net primary production), or maximum values of physical and biological variables: current speeds and maximum Chl-*a*, respectively. Biological variables for population dynamics included total annual abundance, recruitment or mean breeding/pupping success (Table 1). Individual zooplankton species were grouped by summing up the abundance into assemblages, based on the geographical distribution and ecological characteristics of the underlying species, based on Beaugrand, 2004. All other trophic levels (fish, seabirds and mammals) were not grouped but used as individual species in the ecosystem models (Table 1). We refer to all the variables in the study as "ecosystem components" but distinguish components based on them being either physical (e.g. horizontal currents speed) or biological (e.g. sandeel recruitment) indicators. The ecosystem components in the study were chosen as they cover the main physical and biological variables that have been shown to be important to marine mammals and seabirds and their prey (Carroll et al., 2015; Chavez-Rosales et al., 2019; Wakefield et al., 2017). These will alter with climate change (Holt et al., 2016; Sadykova et al., 2017; Wakelin et al., 2015), and with the next biggest change to our shallow seas: very large number of new structures and substantial (100's GWs) extraction of energy from ORE (Boon et al., 2018; Daewel et al., 2022; De Dominicis et al., 2018; Dorrell et al., 2022; van der Molen et al., 2014). Previous studies were also conducted on a larger spatial scale (i.e., > 1000 km^2) identifying the indicators used here as key regarding ecosystem change (Trifonova et al., 2021) and assessing both ecosystem status and resilience to natural and anthropogenic changes (Trifonova and Scott, 2024).

2.2. Bayesian networks

Formally, a Bayesian network (BN) describes the joint distribution (a way of assigning probabilities to every possible outcome over a set of variables, $X_1 \ldots X_N$) by exploiting conditional independence relationships, represented by a directed acyclic graph (DAG) (Friedman et al., 1999). The conditional probability distribution (CPD) associated with each variable *X* encodes the probability of observing its values given the values of its parents and can be described by a continuous or a discrete distribution. In this case, the CPD is called a Conditional Probability Table (CPT) and all the CPTs in a BN together provide an efficient

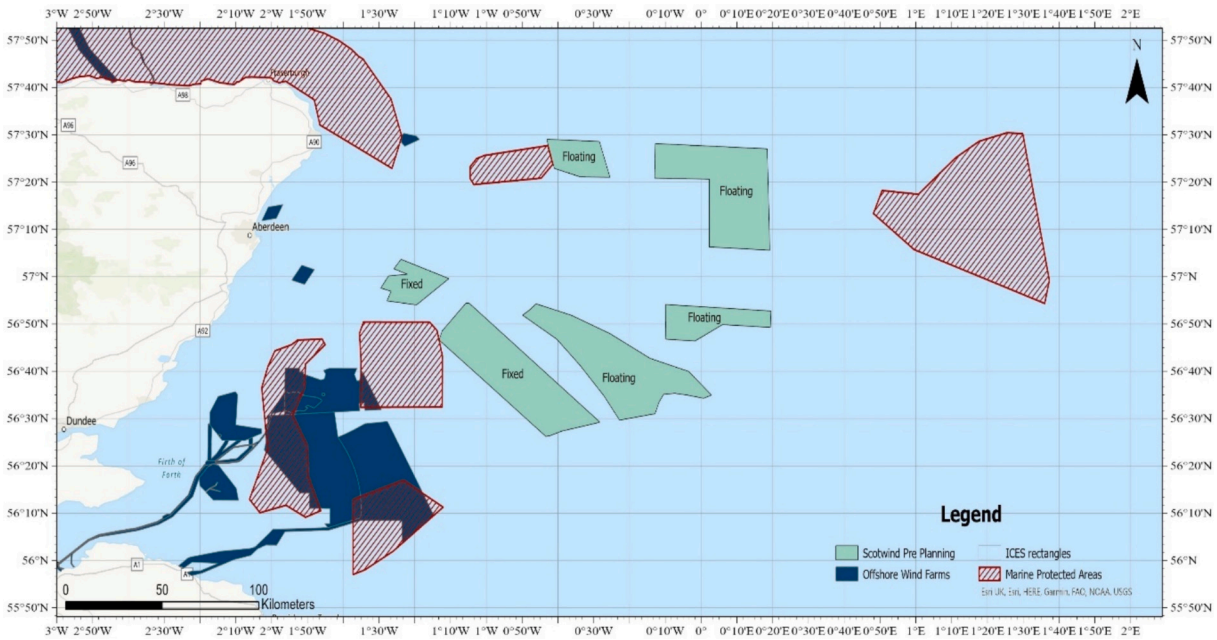


Fig. 1. The spatial boundaries of the study region: Firth of Forth.

Table 1
Summary of physical and biological data.

Category	Ecosystem component	Explanation	Source
Physical	Bottom temperature (BT)	Annual summer (May–August) mean bottom temperature (°C)	Atlantic-European North West Shelf-Ocean Physics Reanalysis provided by E.U. Copernicus Marine Service Information using the Nucleus for European Modelling of the Ocean (NEMO) model (CMEMS-NWS physics)
Physical	Sea surface temperature (SST)	Annual summer (May–August) mean sea surface temperature (°C)	CMEMS-NWS physics
Physical	Potential Energy Anomaly (PEA)	Annual summer (May–August) mean potential energy anomaly (J/m^3). The energy required to mix the water column completely and commonly used as an indicator for the strength of stratification (De Dominicis et al., 2018)	CMEMS-NWS physics
Physical	Horizontal current speed (HSpeed)	Annual summer (May–August) mean maximum depth averaged horizontal currents speed (m/s)	CMEMS-NWS physics
Physical	Vertical current speed (VSpeed)	Annual summer (May–August) mean maximum depth averaged horizontal currents speed (m/s)	CMEMS-NWS physics
Primary production	Chlorophyll-*a* (max Chl-*a*)	Annual summer (May–August) mean maximum chlorophyll-at any depth (mgC/m^3)	Atlantic-European North West Shelf-Ocean Biogeochemistry Reanalysis provided by E.U. Copernicus Marine Service Information using the European Regional Seas Ecosystem Model (ERSEM) (CMEMS-NWS biogeochemistry)
Primary production	Net primary production (NetPP)	Annual summer (May–August) mean depth averaged net primary production (gC m^{-2} $year^{-1}$)	CMEMS-NWS biogeochemistry
Physical	Mixed Layer Depth (MLD)	Annual summer (May–August) mean mixed layer depth (m). The deepest layer affected by surface turbulent mixing; indicator for the variations of primary production	CMEMS-NWS physics
Biogeochemical	Oxygen within the Bottom Mixed Layer (Oxy)	Annual summer (July–October) mean minimum oxygen (μmol L^{-1})	CMEMS-NWS biogeochemistry
Abundance	A2 zooplankton assemblage	Annual summer (May–August) total sum count of zooplankton species (e.g. *Calanus helgolandicus*)	Continuous plankton recorder (CPR) survey
Abundance	A4 zooplankton assemblage	Annual summer (May–August) total sum count of zooplankton species (e.g. *Para-Pseudocalanus* spp.)	CPR Survey
Abundance	A5 zooplankton assemblage	Annual summer (May–August) total sum count of zooplankton species (e.g. *Acartia* spp.)	CPR Survey
Abundance	A6 zooplankton assemblage	Annual summer (May–August) total sum count of zooplankton species (e.g. *Calanus finmarchicus*)	CPR Survey
Recruitment	Sandeel	Annual number of individuals to enter the fished component of the stock	ICES Stock Assessment
Recruitment	Herring	Annual number of individuals to enter the fished component of the stock	ICES Stock Assessment
Recruitment	Sprat	Annual number of individuals to enter the fished component of the stock	ICES Stock Assessment
Recruitment	Mackerel	Annual number of individuals to enter the fished component of the stock	ICES Stock Assessment
Recruitment	Haddock	Annual number of individuals to enter the fished component of the stock	ICES Stock Assessment
Recruitment	Cod	Annual number of individuals to enter the fished component of the stock	ICES Stock Assessment
Human pressure	Catch of pelagic fish species (herring, sandeel, sprat, mackerel; Catch PEL)	Annual total sum of nominal catches (tonnes live weight)	ICES Historical Nominal Catches (1950–2010) and Official Nominal Catches (2006–2022)
Human pressure	Catch of demersal species (cod, haddock; Catch DEM)	Annual total sum of nominal catches (tonnes live weight)	ICES Historical Nominal Catches (1950–2010) and Official Nominal Catches (2006–2022)
Human pressure	Catch of shellfish species (scallops,	Annual total sum of nominal catches	ICES Historical Nominal Catches (1950–2010) and

(*continued on next page*)

Table 1 (*continued*)

Category	Ecosystem component	Explanation	Source
	Nephrops; Catch Shell)	(tonnes live weight)	Official Nominal Catches (2006–2022)
Human pressure	Landings of pelagic fish species (herring, mackerel; Landings PEL)	Annual summer (May–August) total sum of landed fish (tonnes live weight)	Marine Management Organization (MMO)'s annual UK Sea Fisheries Statistics
Human pressure	Landings of demersal fish species (cod, haddock; Landings DEM)	Annual summer (May–August) total sum of landed fish (tonnes live weight)	MMO Fisheries Statistics
Human pressure	Landings of shellfish species (scallops, *Nephrops*; Landings Shell)	Annual summer (May–August) total sum of landed shellfish (tonnes live weight)	MMO Fisheries Statistics
Breeding success	Northern gannet (*Morus bassanus*)	Annual summer mean number of chicks fledged per pair	Seabird monitoring programme
Breeding success	Black-legged kittiwake (*Rissa tridactyla*)	Annual summer mean number of chicks fledged per pair	Seabird monitoring programme
Breeding success	Common guillemot (*Uria aalge*)	Annual summer mean number of chicks fledged per pair	Seabird monitoring programme
Breeding success	Razorbill (*Alca torda*)	Annual summer mean number of chicks fledged per pair	Seabird monitoring programme
Abundance	Harbour porpoise (*Phocoena phocoena*)	Annual summer (May–August) mean of encounter rate	Waggitt et al. (2020).
Productivity	Grey seal (*Halichoerus grypus*)	Annual summer mean estimates of pup production	Special Committee on Seals (SCOS, 2022)
Abundance	Harbour seal (*Phoca vitulina*)	Annual summer (August) total sum count of harbour seals	SCOS (2022)

factorization of the joint probability:

$$p(x) = \prod_{i=1}^{n} p(x_i|pa_i)$$

where pa_i are the parents of the node x_i (which denotes both node and variable).

The DAG consists of nodes (or variables) and edges (or links) between the nodes. "Parent" nodes are those from which arrows originate, and "child" nodes are those to which arrows are pointing. Edges between nodes represent dependency relationships. Each node in the DAG is characterized by a state which can change depending on the state of other nodes and information about those states propagated through the DAG. By using this kind of inference, one can change the state or introduce new data or evidence (change a state or confront the DAG with new data) into the network, apply inference and inspect the posterior distribution (which represents the distributions of the variables given in the observed evidence). The graphical structure of BNs is particularly convenient when we aim to describe an ecological network to model all the interactions between species and their environment that also provides a user-friendly framework to communicate the results (Chen and Pollino, 2012). It is relevant to think of the BN as a "graph", describing species as the "nodes" within the graph, and interactions as the links or "edges" that join the nodes (Faisal et al., 2010).

2.3. Dynamic Bayesian networks

Modelling time series is achieved by using an extension of the BN known as the Dynamic Bayesian Network (DBN), where nodes represent variables at time slices. DBNs are directed graphical models of stochastic processes that characterize the unobserved and observed state in terms of state variables, which can have complex interdependencies (Murphy, 2001). DBNs can model the dynamics of a dataset using a hidden variable.

This hidden variable is used to model unobserved variables and missing data and can infer some underlying state of the series when applied through an autoregressive link that can capture relationships of a higher order (Murphy, 2001). The hidden variable allows us to examine unmeasured effects that would bring further insight on the importance of ecosystem dynamics to better understand community structure and resilience in an exploited ecosystem (Trifonova et al., 2015; Uusitalo et al., 2018). In most domains, the observed variables represent only some characteristics of a system, which can have a negative effect on the learning procedure. For example, the apparent complexity of a predicted variable can be explained imagining it is a result of two simple processes, the "true" underlying state, which may evolve deterministically, and our measurement of the state, which is often noisy (Murphy, 2002). We can then "explain away" unexpected outliers in the observations, as opposed to strange fluctuations in "reality".

A hidden variable can be linked to one, multiple, or all, of the observed ecosystem components in the model. Then, the hidden variable value depends on all the observed ecosystem components it is linked to, and a change in the pattern of the hidden variable indicates a change in the system interactions. This is highly useful in ecological analyses where nonstationary dynamics are common and complex ecological interactions change with time due to changing pressures e.g., climate change (Chen and Pollino, 2012). In this work, the hidden variable was included in the models, to capture complex interdependencies between and among ecosystem components that might represent something external to the community, which is not purely constrained within the model structure. We use the hidden variable in this study, to represent a change in the underlying ecosystem dynamics (i.e. ecosystem state), following a natural or anthropogenic disturbance to the system interactions in the study region.

2.4. Uncertainty propagation

We used the variance as a measure of uncertainty (i.e., a high variance meaning greater uncertainty about the outcome of X given its parents or a low variance meaning that X is more tightly constrained by its parents). For example, in the case of two variables X and Y, with Y influencing X. The conditional distribution $P(X|Y)$ can be modelled as a Gaussian distribution with mean ($\mu_X(Y)$) and variance ($\sigma_X^2(Y)$):

$$P(X|Y) = N\left(\mu_X(Y), \sigma_X^2(Y)\right)$$

In a Bayesian network, uncertainty about a variable propagates through the network from the parent nodes to the child nodes. We defined variance in the CPD which reflects the uncertainty associated with predicting the value of a variable given its parents. When set up in this way, the uncertainty, in turn, influences the accuracy and reliability of the model predictions (i.e., if the conditional variance of a node is high, the distribution of possible outcomes for that node will be wide). The accuracy of a Bayesian network model can be affected by how well it handles uncertainty and variance, especially in terms of its predictive performance.

However, a more complex model which accounts for varying variance in different parts of the network (e.g., different variances for

different conditional distributions), can sometimes improve predictions by allowing the model to better reflect the underlying uncertainties in the system (Nabney and Cheng, 1997; Simoen et al., 2015; Montesinos López et al., 2022). If variance is assumed to be constant or ignored in certain parts of the model, it may lead to overly simplistic models that fail to capture important nuances in the data.

There is copious literature that addresses statistical metrics, such as variance and autocorrelation and their use as indicators of an approaching regime shift (Carpenter et al., 2014; Jiao, 2009). In previous work, they have been used to identify a functional collapse (i.e., regime shift) by modelling early-warning signals in the time-series (Trifonova et al., 2014). However, here, we use them as an alternative approach to account for uncertainty and examine whether their inclusion influences the model accuracy. We are also interested in identifying to what extent including them in our model impacts the expected values of the hidden variable. Previous work has shown that after the addition of the metrics in the model, the hidden variable was more stable and more likely to reflect the underlying ecosystem dynamics (i.e., capture a regime shift) (Trifonova et al., 2014).

3. Experiments

3.1. Learning Bayesian networks

We learn the Bayesian network structure for each of the spatial regions by applying a hill-climb optimization technique. The hill-climb search begins with an empty network. In each stage of the search, networks in the current neighbourhood are found by applying a single change to a link in the current network such as "add arc" or "delete arc" and choose the one change that improves the score the most. We used the Bayesian Information Criterion (BIC) for scoring candidate networks:

$$BIC = logP(\theta) + logP(\theta|D) - 0.5klog(n)$$

where Θ represents the model, D is the data, n is the number of observations (sample size) and k is the number of parameters. Log $P(\Theta)$ is the prior probability of the network model Θ, log $P(\Theta|D)$ is the log likelihood whilst the term k log(n) is a penalty term, which helps to prevent overfitting by biasing towards simpler, less complex models. The learned Bayesian network links represent dependence, these are spatial relationships that are predictive in an informative, not causal aspect (Milns et al., 2010; Trifonova et al., 2015). The method identifies similarity in the temporal trend of the paired variables (i.e. both variables increase, or as one increases, the other decreases over time). We performed the hill-climb with random restart ($n = 10$), which conducts several hill-climbing runs, perturbing the result of each one as the initial network for the next. Then, we apply the learning for 1300 iterations. The maximum number of "parent" nodes (learned from the hill-climb) was limited to three to avoid over-fitting (Trifonova et al., 2015). This hill-climb approach produces a probabilistic dependency output per year (i.e., value from 0 to 1) and for all the possible combinations of paired relationships between the observed variables. We define a confidence threshold - the minimum confidence (estimate of the probability of finding a relationship) for a relationship to be accepted in the learned network structure. We defined relationships of high confidence in time as those in which we have the greatest mean confidence (calculated from all the years per identified relationship) of being in the generated network (threshold ≥ 0.25). We use the confidence value to represent the strength of each dependency relationship between a pair of two variables. The confidence of the identified relationship represents the level of similarity in the temporal trend of the paired variables. In addition, to learn the network structure for each year in the time window, the hill-climbing was conducted on a window of data (size of window = 10). In this way, we would be able to capture any significant interactions over the previous 10 years. Based on the level of confidence and the number of the identified relationships between the observed indicators and between the observed indicators with the hidden variable, we define "best" indicators, which represent the most confident data-driven estimates of indicators of ecosystem dynamics and their changes across space and time (Fig. 2).

3.2. Ecosystem models comparison

3.2.1. HDBN ecosystem model

The modelling approach is a dynamic Bayesian network model with a hidden variable (HDBN, Fig. 2) that is a modified version of the model developed in Trifonova et al. (2015, 2017). The model was developed from the identified consistent physical and biological indicators from Section 3.1. From the strongest relationships, up to three indicators (i.e. "parent" nodes) were selected that drive each target ecosystem component (i.e. "child" node) and were used to build the modelling structure. Therefore, in this way the HDBN ecosystem model captures the spatial and temporal variability of multiple biophysical interactions throughout the trophic chain, ensuring that the strongest relationships (i.e. relationships of high dependency that are predictive in an informative, not causal aspect), and so the most consistent indicators of ecosystem change, are the ones identified in this process. The model included a single hidden variable that was modelled as a discrete node with two states.

When the model parameters are fitted with data, the value of the hidden variable is set so that it maximizes the fit of the model to the data (e.g. the log-likelihood). If the patterns of the observed variables change in the time series, e.g. the slope of a dependency between two variables changes, the value of the hidden variable linked to these variables' changes. Thus, we use the hidden variable in this study, to represent a change in the underlying ecosystem dynamics (i.e. ecosystem state). To do this, the hidden variable was linked to all the ecosystem components in the model.

We want to compute $P\ (H^t|X^t, X^{t-1})$, where H^t represents the hidden variable and X^t represents all observed variables at times t. We use the predicted variable states from time t to infer the hidden state at time t. The hidden variable was parameterized using the Expectation Maximization (EM) algorithm (Bilmes, 1998). In this case, the log-likelihood is:

$$L(\Theta) = \log P(X|\Theta) = \log \sum_j P(X, H|\Theta)$$

where $\sum_H$ is the sum over the set of hidden variables H, required to obtain the marginal probability of the data. In the first step of the EM, the hidden variable is inferred using the predicted states, whilst in the second step the estimated likelihood function is maximized. When the algorithm converges to a local maximum, the parameters are estimated. We used an exact inference method: the junction tree algorithm (Murphy, 2001).

3.2.2. HDBN + physics and HDBN + ecology ecosystem models

The variance (i.e., the square root of the standard deviation) was calculated for all the physical and biogeochemical drivers and used to populate their CPDs (i.e., gaussian distributions) when building the first model variant, the HDBN + physics ecosystem model. In the second ecosystem model variant (HDBN + ecology), the variance was calculated for the biological indicators (i.e., zooplankton abundance, fish recruitment, birds breeding success and mammals' abundance and/or harbour porpoise encounter rate) and was used to populate their CPDs when building the HDBN + ecology ecosystem model. In this way, we can account for varying variance in different parts of the network (e.g., different variances for different conditional distributions), which will potentially improve predictions by allowing the models to better reflect the underlying uncertainties in the system. In contrast, parameters for the HDBN model were not specifically assigned. In the case of gaussian nodes, the following was assumed: if node is called Y, its continuous parents (if any) are called X, and its discrete parents (if any) are called Q.

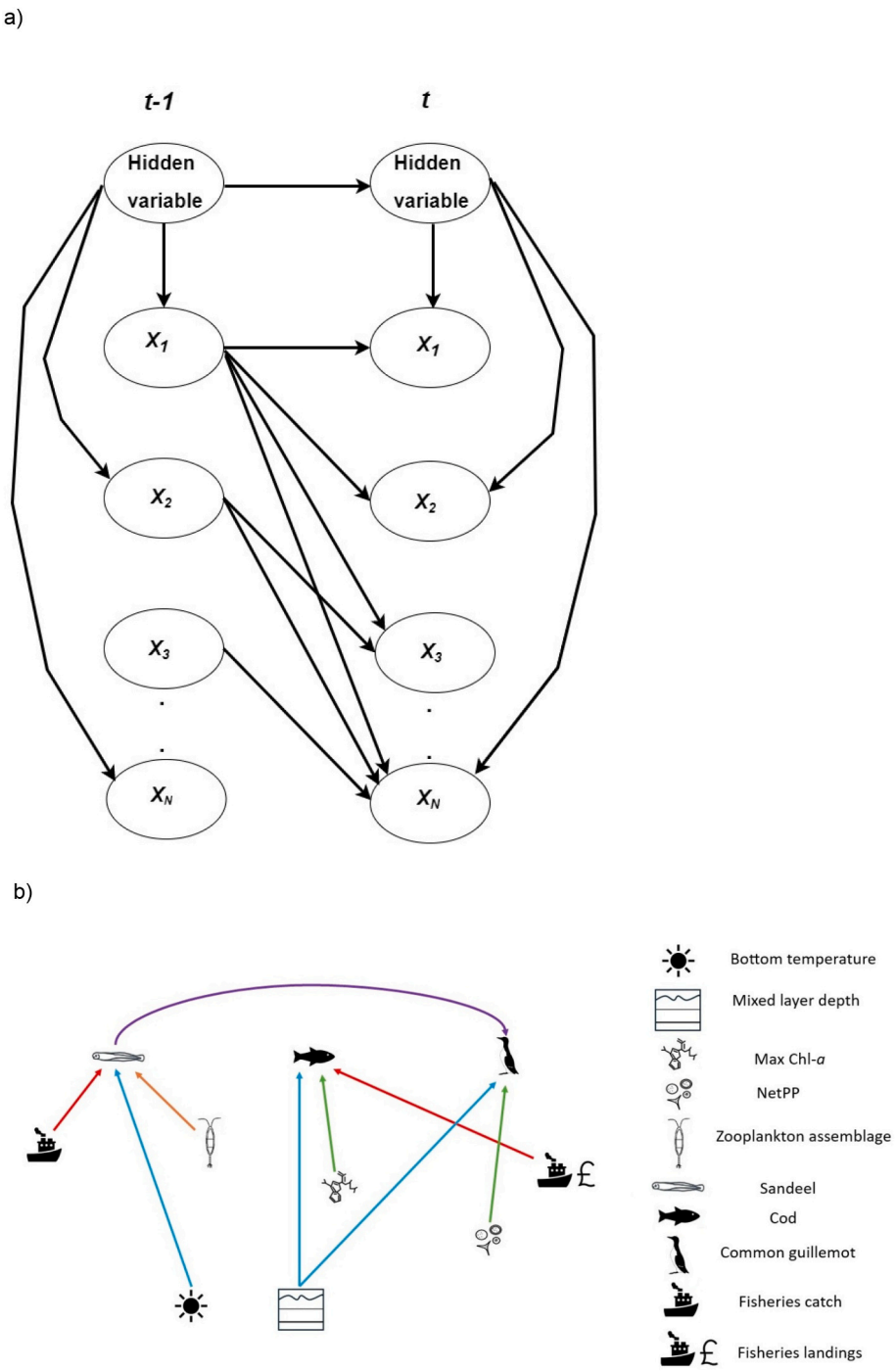


Fig. 2. (a) General structural form of the dynamic Bayesian network model with a hidden variable (HDBN) where $X_1 \ldots X_N$ represents the set of variables and arrows denote conditional independence relationships. (b) The strongest data-driven relationships, that were used to build the ecosystem models, however, only some illustrative ecosystem components are shown, to maximise visual clarity. The same illustrative ecosystem components are presented in sub-section 4.12. Blue-coloured links indicate relationships with the physical indicators, red with fisheries catch and landings, green with primary production components, orange with zooplankton assemblages and purple with higher trophic levels. Symbols used to denote the ecosystem components are next to the relationships. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The distribution on *Y* is defined as follows:

$$Y|X = x, Q = i \sim N(mu(:, i) + W(:, :, i)^{*} x, Sigma(:, :, i))$$

where *N (mu, Sigma)* denotes a normal distribution with mean *mu* and covariance *Sigma* (Murphy, 2001).

3.2.3. *HDBN + VAR + AC ecosystem model*

Finally, variance and autocorrelation were calculated on a window of data, set to size 10, so that each metric captures the value of interest over the previous 10 years. The size of the window was found to be optimum due to the length of the time series Note, the predictions from this model variant will be available for a shorter time series due to the windowing approach. The metrics were included in another ecosystem model variant (HDBN + VAR + AC), and they were both linked to all the remaining ecosystem components. We explore to what extent including the two statistical metrics in our model impacts the expected values of

the hidden variable but also whether including the two metrics influences the model accuracy. This model will not be addressed as a competing model in the results and discussion, we simply wanted to state the overall predictive accuracy during the learning process using a windowing approach, however, the learned value of the hidden variable will be discussed in the following sub-section.

We examine the models' (i.e., HDBN, HDBN + physics, HDBN + ecology and HDBN + VAR + AC, Table 2) accuracy in terms of their ability to reproduce observations of the trends (increases versus decreases each year of the time series) in all the ecosystem components (oceanographic processes as well as species/functional groups at all trophic levels). The same modelling structure (learned from the identified relationships from Section 3.1) was used for all the model variants, except for the HDBN + VAR + AC model, where variance and autocorrelation were both linked to all the remaining ecosystem components. In the HDBN + physics and HDBN + ecology models, the difference was the added variance to the CPDs, but the graphical structure was identical. Model performance, in terms of sum of squared error (SSE), was assessed for each model and predictions were compared on a year-to-year basis versus the original input data. Non-parametric bootstrap (re-sampling with replacement from the training set, (Friedman et al., 1999) was applied 250 times for the model and its variants to obtain statistical validation in the predictions. The data were standardised prior to conducting the experiments to a mean of 0 and standard deviation of 1. We conduct all experiments using the Bayes Net Toolbox in MATLAB (Murphy, 2001).

4. Results

4.1. Model comparison

4.1.1. Lowest SSE values

SSE per species, per model were compared to assess how well each model performed against the annual input data values. The HDBN model reported the highest number of ecosystem components (n = 9, 47 %) predicted most accurately (least SSE per species), closely followed by the HDBN + physics (n = 6, 32 %). These results are reassuring that the inference scheme can handle the increased model complexity. The model with the least accurately predicted ecosystem components was the HDBN + ecology (n = 4, 21 %) with some SSEs higher than 30.00, which highlights the importance of the level of variance when running these types of models. We found the threshold of 30.00 to be most appropriate based on examining the range of SSE values across models as well as across ecosystem components. Although the general improvement in predictive accuracy of the HDBN model over the competing models, there is a similar level of accuracy (i.e., least SSE difference: less than 5.00 between the generated overall predictions of two models) for most of the ecosystem components. Two exceptions to that were haddock and razorbill. For these species, the SSE difference across the competing models was always higher than 5.00. Some ecosystem components (e.g., the zooplankton functional group A6), were generally predicted with higher SSEs values from two of the competing models (i.e., SSE higher than 30.00) in comparison to the remaining ecosystem components. Overall, the HDBN + VAR + AC also performed well, specifically reporting some low SSE values for the lower trophic levels (e.g., A2 zooplankton group) and higher trophic level species like haddock and harbour seal (i.e., SSE less than 10.00). At the same time, a higher SSE (i.e., SSE higher than 30.00) was reported for the seabirds (e.g., razorbill).

Overall, for higher trophic levels of fish and above (i.e., seabirds and mammals) the addition of either physical or biological variance saw a better fit (decrease in SSE) for 77 % of those 13 species (10/13). When comparing predictions across the ecosystem components, mammal species were most accurately predicted by either the HDBN + physics or HDBN model. The seabird species were most accurately predicted by either the HDBN + physics or HDBN + ecology, highlighting the importance of including variance in the case of these higher trophic level ecosystem components. Across the fish species, it was harder to find any specific patterns in terms of which model reported most accurate performance per species, however, some patterns were identified based on what specific indicators (physical vs biological) were driving the fish recruitment in the models. For example, sandeel and herring were both most accurately predicted by the HDBN + physics model and their levels of recruitment are both driven by a combination of physical and biogeochemical indicators, whilst the sprat, which was most accurately predicted by the HDBN was driven by biological indicators and catch.

Table 2
Summary of HDBN models.

Models	Name	Comments
HDBN	Hidden Dynamic Bayesian network model	A hidden dynamic Bayesian network model with no variance specified in the CPDs
HDBN + physics	Hidden Dynamic Bayesian network model with conditional variance on physical and biogeochemical indicators	Variance was calculated for BT, SST, PEA, max Chl-a, NetPP, MLD, Hspeed, Vspeed, Oxygen and was used to populate their CPDs
HDBN + ecology	Hidden Dynamic Bayesian network model with conditional variance on biological indicators	Variance was calculated for zooplankton abundance, fish recruitment, birds breeding success and mammals' abundance and/or productivity and was used to populate their CPDs
HDBN + VAR + AC	Hidden Dynamic Bayesian network model with statistical metrics: variance and autocorrelation	Variance and autocorrelation were calculated on a rolling window of data

Table 2. Sum of squared error (SSE) of the ecosystem components predictions generated by the HDBN and its variants (a). The component-specific interactions that are used to build the HDBN models are shown inside the brackets. The * symbol indicates most accurate predictions for ecosystem components across the three models (values of SSEs that are less than 30.00). (b) shows SSE of the ecosystem components generated by the HDBN + VAR + AC. In a separate table as the values are not directly comparable to the other three models.

a) Ecosystem components	HDBN	HDBN + physics	HDBN + ecology
Max Chl-*a* (PEA, Hspeed, MLD)	**22.48***	29.85	25.39
NetPP (Oxy, Hspeed)	**6.34***	6.92	6.43
A2 (Hspeed, Vspeed)	**29.96***	31.14	30.99
A4 (PEA, NetPP, Vspeed)	**10.08***	10.13	11.33
A5 (PEA, Vspeed)	**9.63***	10.47	9.85
A6 (BT, NetPP, Hspeed)	**29.35***	32.50	41.32
Sandeel (BT, A4, Catch DEM)	20.15	**17.59***	19.52
Herring (MLD, NetPP, Land PEL)	26.85	**25.80***	29.50
Sprat (A2, A6, Catch PEL)	**26.99***	27.96	29.46
Mackerel (BT, A2)	30.99	30.50	**27.94***
Haddock (NetPP, Catch DEM)	26.01	20.03	**18.03***
Cod (MLD, Max Chl-*a*, Land DEM)	**14.64***	24.68	14.92
Kittiwake (Vspeed, Sandeel, Sprat)	29.1	**27.04***	43.13
Guillemot (MLD, NetPP, Sandeel)	29.79	**24.36***	30.57
Gannet (A6, Herring, Sprat)	29.44	36.13	**29.84***
Razorbill (Mackerel, Haddock, Cod)	29.71	31.90	**21.48***
Grey seal (BT, NetPP, A6)	14.80	**13.12***	20.93
Harbour seal (Max Chl-*a*, A5, Cod)	**29.41***	31.55	42.90
Harbour porpoise (Max Chl-*a*, Oxy, Sprat)	15.51	**11.55***	13.44

b)	HDBN + VAR + AC
Max Chl-*a*	22.34
NetPP	16.72
A2	9.31
A4	16.17
A5	26.88
A6	22.52

(*continued on next page*)

(*continued*)

b)	HDBN + VAR + AC
Sandeel	17.07
Herring	26.69
Sprat	38.38
Mackerel	14.38
Haddock	9.11
Cod	23.87
Kittiwake	29.65
Guillemot	32.66
Gannet	10.39
Razorbill	37.64
Grey seal	10.92
Harbour seal	9.68
Harbour porpoise	11.87

4.1.2. Temporal trends

We compared example ecosystem components and their population trend predictions in time across the three model variants. We want to visually demonstrate how well the model variants performed in reproducing the inter-annual variability and long-term patterns (always shown as blue lines) versus the original input data (red lines). Note, we only show some illustrative examples, with their 95 % confidence intervals calculated from the bootstrap predictions' mean and standard deviation, shown in the Supporting information (SI). The models were able to capture many of the changes (increases versus decreases) of the ecosystem components across over time, predicting the general trends in population dynamics for all lower trophic level functional groups and higher trophic level species using three or fewer indicators.

For sandeel, the HDBN and HDBN + physics were better able to capture the declining trend in the early 2000s, in comparison to the HDBN + ecology (Fig. 3). The HDBN + physics (least SSE value) in comparison to the HDBN was better in modelling some of the individual yearly variations (e.g., years 2014, 2019), however, the declining trend (from the 1990s to early 2000s) over time was better captured by the HDBN model. The HDBN + VAR + AC model variant performed well in capturing some of the individual yearly variations (e.g., 2005).

Similarly, the two best performing models (least SSE values) were able to capture the long-term trend in the case of cod recruitment, specifically, the declining trend from the early 2000s, whilst the HDBN + physics model was better in capturing some of the specific yearly variations (mid to late 2000s, Fig. 4). Similarly, to the sandeel, the HDBN + VAR + AC model for cod, was able to capture well a lot of the individual yearly variations.

In the case of guillemot, all three model variants were able to capture the declining trend from the mid 1990s- early 2000s (Fig. 5). However, none of the models were able to capture the steep decline in 2006, with the exception of the HDBN + VAR + AC model. After 2006, it was the HDBN + physics, followed by the second accurate model (HDBN) that were able to capture the inter-annual variabilities in the time series.

4.2. Hidden variable

To assist in characterizing the ecosystem state, we examine the learned hidden variables from the HDBN and HDBN + VAR+ AC models (Fig. 6). The hidden variable from the HDBN model was relatively stable for majority of the time series: it modelled one state until 2010, followed by some fluctuations in 2011 and 2018. The hidden variable from the HDBN + VAR+ AC model was a bit more varied than the hidden variable from the HDBN model.It still identified a change in the ecosystem state around the same time, even, a year earlier in 2009, followed by two changes in the ecosystem state in 2013–2015 and then in 2018 that remained until the end of the time series.

5. Discussion

5.1. Summary of model variants

In this study, we examined whether adding variance as a measure of uncertainty to the CPDs of specific ecosystem components would result

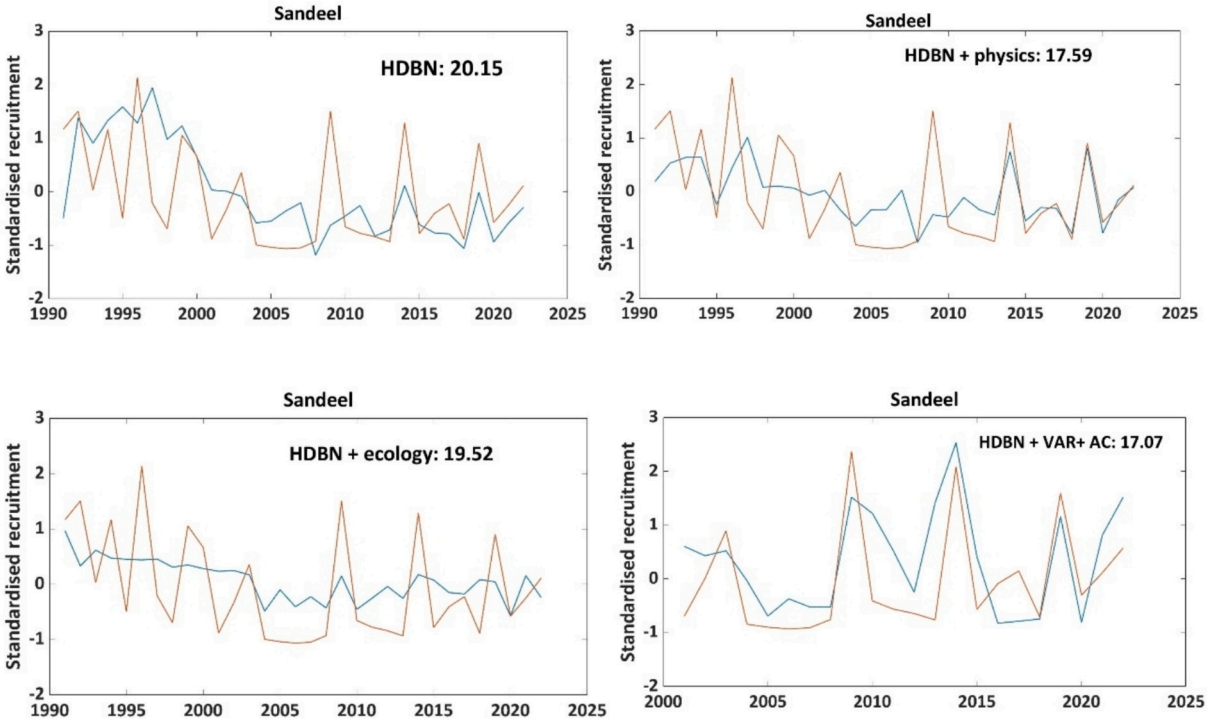


Fig. 3. Model predictions (blue line) from the HDBN and its model variants (HDBN + physics, HDBN + ecology and HDBN + VAR + AC) versus real data (red line) for sandeel recruitment. SSEs are shown under the panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

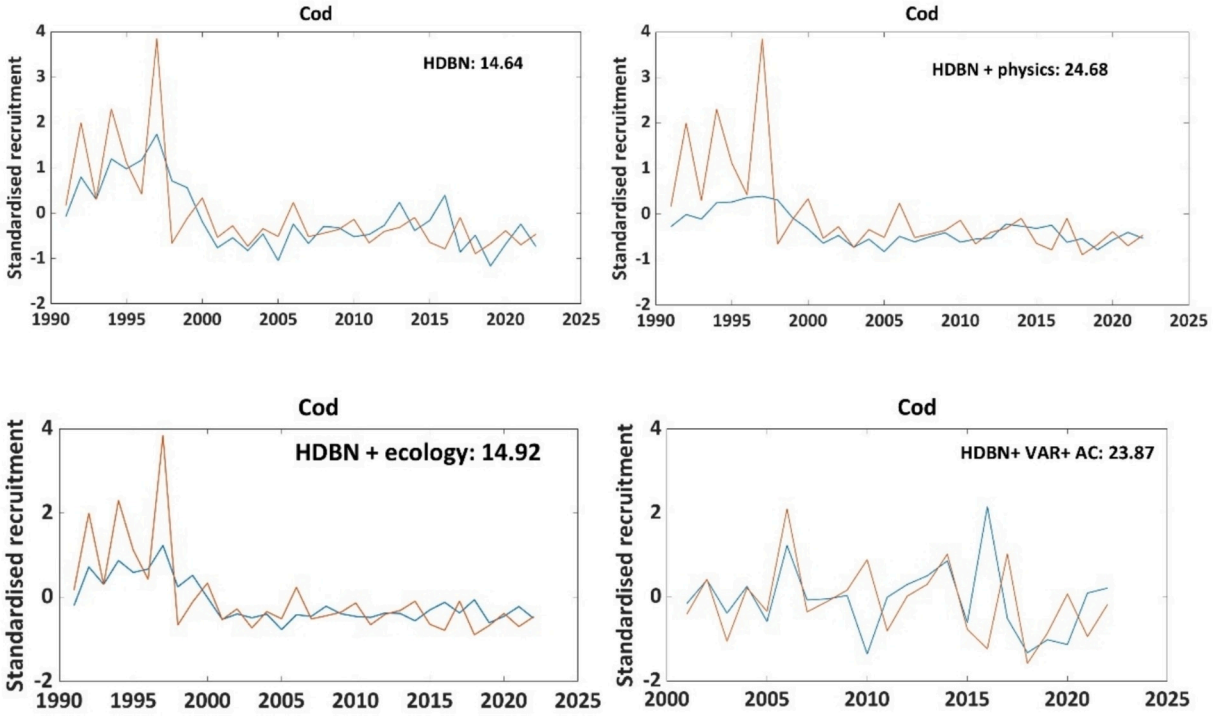


Fig. 4. Model predictions (blue line) from the HDBN and its model variants (HDBN + physics, HDBN + ecology and HDBN + VAR + AC) versus real data (red line) for cod recruitment. SSEs are shown under the panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

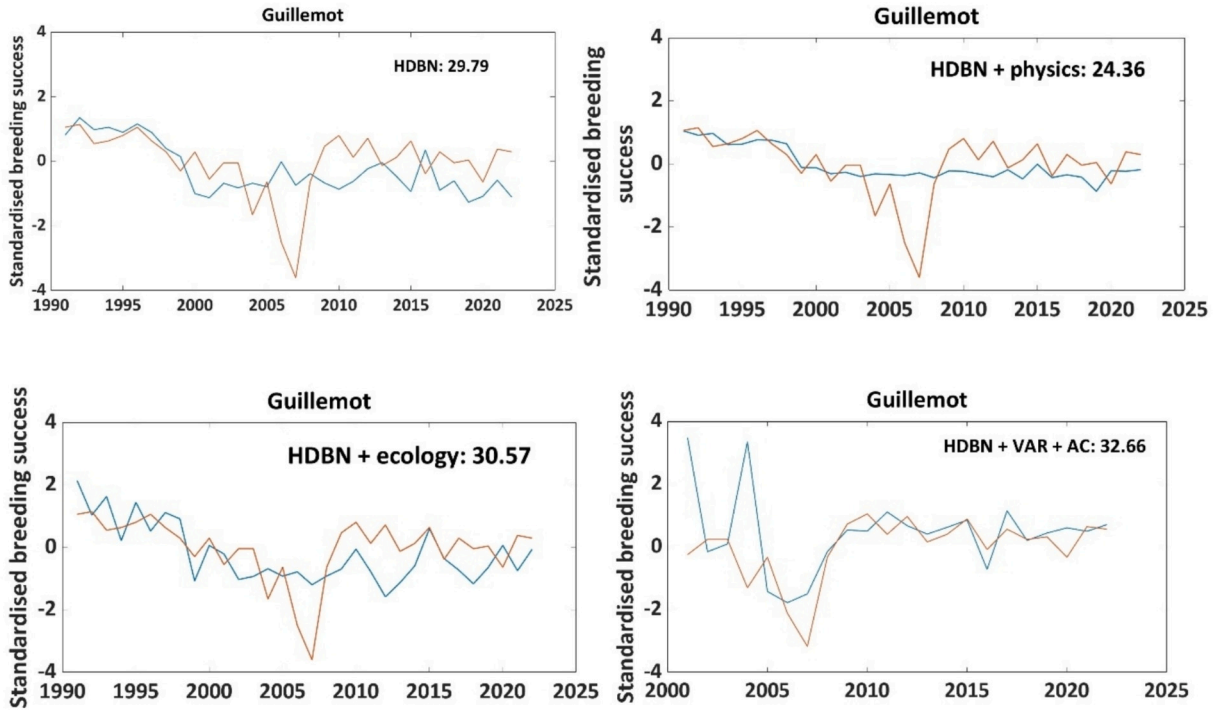


Fig. 5. Model predictions (blue line) from the HDBN and its model variants (HDBN + physics, HDBN + ecology and HDBN + VAR + AC) versus real data (red line) for guillemot breeding success. SSEs are shown under the panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in a better performance of a hidden dynamic Bayesian network model (HDBN). The HDBN model showed consistently accurate predictions of the ecosystem components. The HDBN + physics (variance added to the CPDs of selected physical and biogeochemical indicators) was the second-best performing model which was reassuring that the increased model complexity applied here has resulted in revealing some genuine patterns of the underlying lower and higher trophic level relationships identified by our approach and their heterogeneity. In addition, these

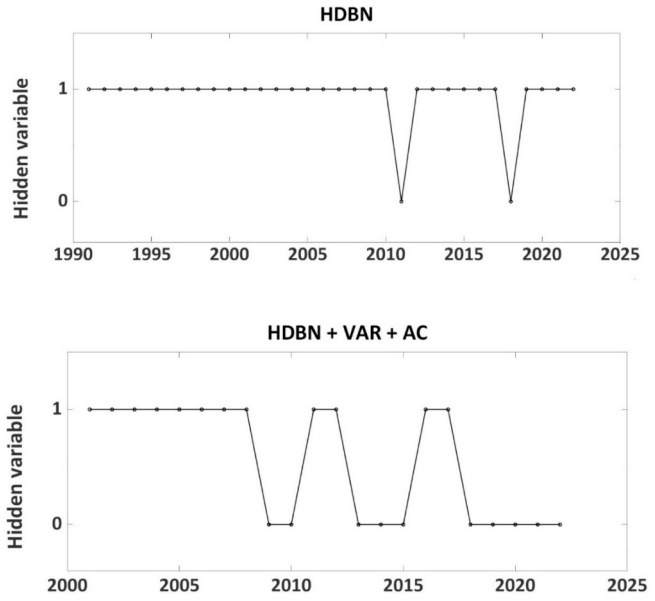


Fig. 6. The learned hidden variables from the HDBN (top) and HDBN + VAR+ AC (bottom) models. Note, the shorter time series in the bottom plot due to the windowing approach.

results also suggest more insight into the relevance of physical and biogeochemical indicators, such as stratification, temperature and NetPP when evaluating ecosystem structure and function in efforts to determine the ecosystem state. It is evident that changes in the population dynamics of higher trophic levels are likely to reflect those of their preferred prey, which may, in turn, be bottom–up driven by dynamic bio-physical oceanographic processes like across spatial and temporal scales (Bertrand et al., 2014; Boyd et al., 2015; Cox et al., 2018). Temperature is another major driver of marine ecosystems and one of the key factors affecting the physiology and ecology of all marine organisms (Edwards et al., 2020; Evans and Waggitt, 2020; Simpson et al., 2011). BT was found to be a better indicator, in comparison to SST (Table 2a), with the mechanism for that potentially being that BT reflects steady changes over longer periods of time, including integrated trends in warming/cooling, whilst changes in SST reflect a wider range of the daily/seasonal extremes (Trifonova et al., 2022). NetPP was another indicator that was found to be a key driver for majority of the higher trophic level species dynamics. It is through energy transfer along the food web, that NPP provides bottom–up control on fisheries production, identified within the North Sea and across other large marine ecosystems around the globe (Barange et al., 2014; Blanchard et al., 2012; Chassot et al., 2007). NPP plays a significant role in determining habitat preferences and facilitating foraging for both fish prey and top predator species, highlighting that prey and predators are selecting aspects of the habitat type very differently and that might be a reflection of prey species avoiding areas with predators (Cox et al., 2018; Sadykova et al., 2017). With the improved understanding of the exact bottom-up (e.g. levels of mixing and stratification) versus top-down (e.g. predators and fishing) mechanisms that influence habitat use by marine animals across spatial (< 1 km through to 1000 km) and temporal (days through to years) scales, the effects of biophysical interactions on populations and ecosystems and how these vary with climate change can be better understood.

Predictions from all models were sensitive to the observed variables incorporated due to the complex natural processes involved in generating the ecological input data. We reported some higher SSEs (e.g., zooplankton functional group A6, Table 2) which could be due to structural uncertainties (i.e. species-specific relationships and/or colony-specific drivers used to build the models) but also due to empirical data uncertainties perhaps due to some sampling variation in the survey data. In addition, there was some similarity in accuracy of generated predictions from different models that might be attributed to the similar effects of changing climate on many species (Fernandes et al., 2013), as previous work has demonstrated the potential effects of continuous warming with cold-water anomalies and salinity changes (Trifonova et al., 2021).

The HDBN model performed consistently well across the ecosystem components, because the model evaluates the relative influence of different driving factors when modelling ecosystem-level dynamics. Our results highlight the need to include region-specific ecosystem level changes and dynamics of the multiplicity of interactions when building predictive models of complex and heavily exploited ecosystems within shallow seas, such as the North Sea. The recognition of a hidden variable is important which was adopted here to capture unmeasured effects and changes in ecosystem components variance that are not purely constrained within the model structure. This is very different from mass balance model approaches (Christensen and Walters, 2004) whose fitting is conditioned completely upon the model structure.

However, for some of the species, prediction accuracy was improved once variance was included by the HDBN + physics model, suggesting that accounting for additional sources of variation removed spurious interactions and let to a more plausible network structure (Aderhold et al., 2012; Faisal et al., 2010). The successful performance of the model variant highlights the heterogeneous nature of the ecosystem component-specific lower and higher trophic level interactions (e.g., driven by physical, biological and/or combination of both) and gives us more accurate insights on the structure of the underlying ecological system. Through the applied BN approach, we were able to make tractable predictions of the true dynamic nature of physical, biogeochemical and biological relationships and their patterns across trophic levels, and their changes over time. This increases knowledge necessary to add to the traditional use of top predator population dynamics as separate aspects of marine systems and will reduce uncertainties of the level of direct and indirect effects on populations across a range of trophic levels. By accounting for varying variances in the CPDs of the drivers, the model becomes better equipped to handle uncertainty in a way that reflects the true underlying system (Yıldırım and Liaw, 2024).

Some dependencies between variables might be complex and nonlinear. It is often assumed by statistical and mechanistic modelling approaches that the underlying functional relationships (i.e., the governing ecological processes) are assumed to be static. This assumption, of course, is overly simplistic, and maybe inaccurate, as ecosystems are subject to increasing human pressures that can lead to drastic changes, including regime shifts (Scheffer et al., 2009). However, the assumption of static relationships is a necessary feature of mechanistic models, as these models are built on the best available current understanding of how systems function. While this static representation reflects our best knowledge at the time, we are unable to account for 'unknown unknowns'.

From the modelling and data analysis perspective, these pose a challenge, since the same functional forms may not describe the relationships between the variables before and after the change (Blenckner et al., 2015). By allowing variances in the physical and/or biogeochemical drivers to vary, our HDBN model + physics can better capture complex and nonlinear lower and higher trophic level relationships where the uncertainty might vary depending on certain conditions. For example, in a previous study using a larger ecosystem region (i.e., deep central North Sea), that enclosed the study region in this work, the region was shown to be controlled by both types of forcing: bottom-up (e.g., primary production) and top-down (e.g., fisheries exploitation) leading to complex patterns of control on the ecosystem (Trifonova and Scott, 2024). This further explains the better performance of the models with added variance for some of the ecosystem components, i.e., the Firth of Forth region is characterized by a changeable ecosystem state prone to variability, therefore, a more

complex model is needed to capture the underlying ecosystem dynamics. For example, in the case of breeding success, every bird species was better predicted by either the HDBN + physics or HDBN + ecology model, however, the SSE difference was still ≤5.0 between the HDBN model and the model variant, except for razorbill. This suggests that breeding success might be influenced by more complex factors than some of the other ecosystem components that were better predicted by the HDBN model alone. This could also suggest that seabirds might be more prone to variability, thus, highlighting them as potentially more sensitive indicators to pressures.

Interestingly, once the variance was added to the CPDs, the models were better able to capture some of the specific yearly variations in the time series. It seems that the variance adds to the overall model performance by specifically capturing the inter-annual variability in the time series. The potential explanation for this is that in some cases where variables are inherently noisy or more prone to variability (e.g., seabird breeding success), the more complex Bayesian network model applied here can help prevent overfitting by modelling the noise appropriately with higher variance (López et al., 2022; Ramampiandra et al., 2023). This prevents the model from becoming overly influenced by outliers or fluctuations in the data, resulting in more accurate predictions. This was particularly the case for guillemot for which the HDBN + VAR + AC model was the only model that was able to capture the steep decline in breeding success in 2006. Overall, the HDBN + VAR + AC model was able to capture the trends and inter-annual variations with high accuracy, including some specific yearly variations. The accurate performance of this model is also likely due to the inclusion of a rolling window that has been previously discussed as successful in detecting an impending regime shift in ecosystem time series (Carpenter et al., 2014). The successful performance of a dynamic hidden BN model that was used in combination with variance and autocorrelation has been previously demonstrated, but in the context of detecting early-warning signals of functional changes in fisheries across a range of geographic regions (Trifonova et al., 2014).

5.2. Hidden variable

A hidden variable was used in this study to learn and therefore, represent the ecosystem state, and specifically capture any changes in the ecosystem interactions that lead to changes in state. We compared the hidden variable from the HDBN model to the hidden variable from the HDBN + VAR+ AC model to identify to what extent including the statistical metrics impacts the expected value of the hidden variable and therefore, the expected ecosystem state.

Both hidden variables modelled a change in the ecosystem state after 2010, with the hidden variable from the HDBN + VAR + AC modelling the change in state a year earlier in 2009, allowing detection of early-warning signals of functional change across different geographic regions (Trifonova et al., 2014). This change in the ecosystem state has been discussed previously for much larger ecosystem regions and it was thought to be due to changes in the bottom temperature (e.g., cold-water anomalies: Gonzalez-Pola et al., 2019; Trifonova et al., 2021), with the mechanism behind being the interplay between the physical indicators (temperature and mixing) and productivity (Capuzzo et al., 2018). Here, we were able to add further insight on the potential extent of the functional change, even at the smaller spatial scale in this study. Most importantly, with the now added uncertainty (i.e., variance) to the model, we were able to detect such changes earlier in the time series. The hidden variable from the HDBN + VAR+ AC model identified an additional change in the ecosystem state (i.e., 2013–2015) whilst both hidden variables modelled another change in state in 2018. The period 2015–2018 has been previously identified as a period of change with relatively low values in net primary production, most likely attributed to changes in mixing (Capuzzo et al., 2018; Trifonova et al., 2021). Similarly, with the added results from this study, and specifically, once uncertainty was included in the model, we were able to further identify the timing of a second potentially important period in the context of reconstructing the ecosystem dynamics. These results highlight that the use of a hidden variable when modelling ecosystem change is potentially useful in providing insights on the underlying dynamics and patterns in terms of ecological stability and resilience that can contribute towards the general advice on potential response of the system to pressure. Indeed, it is by examining the learned ecosystem state that allows us to conclude whether the environment is in a desirable (predictable) or less desirable state and during which years the state is desirable. Thus, the hidden variable, once set up and updated with rather low effort, could potentially be used to check for possible new changes in the underlying ecosystem dynamics, indicative of major changes in the ecosystem, which could be further investigated (Uusitalo et al., 2018).

6. Conclusion

The dynamic Bayesian network approach is a promising method to analyse complex ecosystem-level interactions, and it may help reveal underlying ecological patterns. Here, we demonstrated that the applied hidden dynamic Bayesian network model can handle the increased complexity by accounting for uncertainty (i.e., variance) in the conditional probability distributions (CPDs) of selected physical, biogeochemical and biological indicators. Indeed, model performance was improved for 77 % of the higher trophic level species (fish, seabirds and marine mammals) once variance was included in the CPDs. Therefore, models that account for additional sources of variation seem to better reflect on the underlying ecosystem-level dynamics. These results provide real insights into the characteristics of the study region, which is a changeable ecosystem state prone to variability, and pave the route for better understanding of the ecosystem structure and function under different pressures. Most importantly, the methodology provides an effective baseline that can be used within marine spatial planning considerations of the relevant implications of future climate change versus anthropogenic impacts (e.g., offshore large-scale wind developments). Our results show temporally specific ecological interactions that indicate a regional relationship of ecosystem components and their habitat with the mechanisms varying from bottom-up (e.g., primary production) through to top-down (e.g., fisheries). We were able to identify the consistent drivers and illuminate the likely mechanisms that led to consistently accurate model predictions. However, it must be noted that perfect reconstruction is unlikely due to the noisy input data and complex ecological process involved in generating such data (Faisal et al., 2010). However, our findings complement more traditional mechanistic (Heath et al., 2021) and statistical (Lynam et al., 2017) approaches; and have extended our knowledge into the ecosystem-level understanding of this North Sea region and its ecological structure and stability. Further, the success of applying the HDBN + VAR+ AC model highlights the usefulness of the rolling window approach in combination with the use of statistical metrics in characterizing the temporal dynamics of this region, specifically improving predictive performance in capturing the inter-annual variability in the time series. The two hidden variables successfully modelled changes in the ecosystem state, one attributed to cold-water anomalies and a second one attributed to the interplay between the physical indicators (temperature and mixing) and productivity. The hidden variable from the HDBN + VAR+ AC model was able to capture these changes earlier than the hidden variable from the HDBN model.

Future work can use the methods shown here with the hidden dynamic Bayesian network model and with added variance in the CPDs from selected physical and biological indicators to produce a range of "what-if?" scenarios to better understand the combined ecosystem-level effects of offshore large-scale wind developments, climate change and fisheries displacement. Such approaches will be useful to guide what habitats/species are more representative of what disturbances and what management decisions are required to steer towards more ecologically sustainable conditions under the influence of future changes (Trifonova

and Scott, 2024). These types of outputs can be used to assess the cumulative effects across a range of trophic species to support the development of evidence-based policy and marine management.

Funding

This work fall under PELAgIO funded by the Natural Environment Research Council (NE/X008835/1): https://ecowind.uk/projects/pelagio.

CRediT authorship contribution statement

Neda Trifonova: Methodology, Formal analysis, Data curation, Conceptualization. **Juliane Wihsgott:** Writing – review & editing. **Beth Scott:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would also like to thank the following people for providing data to this study: Debbie Russel, James Waggitt, Amy Walker, Hannah Fougner, Andrew Logie, Mirko Hauswirth, Tim Dunn, Maria Pagla, Oliver Boisseau, Jared Wilson, Jen Graham, Kate Abbott, Sally Hamilton, Alex Banks, Phil Hammond, Peter Evans, Chelsea Bradbury, Paul Thompson, Nele Markones, Alice Walters, Andrea Salkeld. For detailed information on their organizations and contacts, please refer to the SI.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2025.103510.

Data availability

The sources (i.e., public links) for the data used in this study are shown in Table 1. Data were directly downloaded from the public links provided, except for the zooplankton data for which a data request process was needed, please see here: https://www.cprsurvey.org/data/our-data/. The harbour porpoise data is not publicly available. For specific request regarding access to the harbour porpoise data, please refer to the organizations involved in the collection of the data, provided in the SI. The source code is available in the SI. The BNT toolbox with build in functions to reproduce the work is available here: https://github.com/bayesnet/bnt.

References

Aderhold, A., Husmeier, D., Lennon, J.J., Beale, C.M., Smith, V.A., 2012. Hierarchical Bayesian models in ecology: reconstructing species interaction networks from non-homogeneous species abundance data. Eco. Inform. 11, 55–64.
Barange, M., Merino, G., Blanchard, J.L., Scholtens, J., Harle, J., Allison, E.H., Allen, J.I., Holt, J., Jennings, S., 2014. Impacts of climate change on marine ecosystem production in societies dependent on fisheries. Nat. Clim. Chang. 4 (3), 211–216.
Beaugrand, G., 2004. Monitoring marine plankton ecosystems. I: description of an ecosystem approach based on plankton indicators. Mar. Ecol. Prog. Ser. 269, 69–81.
Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X., Chaigneau, A., Vargas, G., Mousseigne, A., Fablet, R., 2014. Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds. Nat. Commun. 5 (1), 5239.
Bilmes, J.A., 1998. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. Int. Comp. Sci. Inst. 4 (510), 126.
Blanchard, J.L., Jennings, S., Holmes, R., Harle, J., Merino, G., Allen, J.I., Holt, J., Dulvy, N.K., Barange, M., 2012. Potential consequences of climate change for primary production and fish production in large marine ecosystems. Philosoph. Trans. Royal Soc. B Biol. Sci. 367 (1605), 2979–2989.
Blenckner, T., Llope, M., Möllmann, C., Voss, R., Quaas, M.F., Casini, M., Lindegren, M., Folke, C., Chr. Stenseth, N., 2015. Climate and fishing steer ecosystem regeneration to uncertain economic futures. Proc. R. Soc. B Biol. Sci. 282 (1803), 20142809.
Boon, A., Caires, S., Wijnant, I.L., Verzijlbergh, R., Zijl, F., Schouten, J.J., Kooten, T., 2018. The assessment of system effects of large-scale implementation of offshore wind in the southern North Sea. no. June, p. 61.
Boyd, C., Castillo, R., Hunt Jr., G.L., Punt, A.E., VanBlaricom, G.R., Weimerskirch, H., Bertrand, S., 2015. Predictive modelling of habitat selection by marine predators with respect to the abundance and depth distribution of pelagic prey. J. Anim. Ecol. 84 (6), 1575–1588.
Capuzzo, E., Lynam, C.P., Barry, J., Stephens, D., Forster, R.M., Greenwood, N., McQuatters-Gollop, A., Silva, T., van Leeuwen, S.M., Engelhard, G.H., 2018. A decline in primary production in the North Sea over 25 years, associated with reductions in zooplankton abundance and fish stock recruitment. Glob. Chang. Biol. 24 (1), e352–e364.
Carpenter, S.R., Cole, J.J., Pace, M.L., Batt, R., Brock, W.A., Cline, T., Coloso, J., Hodgson, J.R., Kitchell, J.F., Seekell, D.A., Smith, L., 2011. Early warnings of regime shifts: a whole-ecosystem experiment. Science 332 (6033), 1079–1082.
Carpenter, S.R., Brock, W.A., Cole, J.J., Pace, M.L., 2014. A new approach for rapid detection of nearby thresholds in ecosystem time series. Oikos 123 (3), 290–297.
Carroll, M.J., Butler, A., Owen, E., Ewing, S.R., Cole, T., Green, J.A., Soanes, L.M., Arnould, J.P., Newton, S.F., Baer, J., Daunt, F., 2015. Effects of sea temperature and stratification changes on seabird breeding success. Clim. Res. 66 (1), 75–89.
Chassot, E., Mélin, F., Le Pape, O., Gascuel, D., 2007. Bottom-up control regulates fisheries production at the scale of eco-regions in European seas. Mar. Ecol. Prog. Ser. 343, 45–55.
Chavez-Rosales, S., Palka, D.L., Garrison, L.P., Josephson, E.A., 2019. Environmental predictors of habitat suitability and occurrence of cetaceans in the western North Atlantic Ocean. Sci. Rep. 9 (1), 5833.
Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environ. Model Softw. 37, 134–145.
Cheung, W.W., Bruggeman, J., Butenschön, M., 2019. Projected changes in global and national potential marine fisheries catch under climate change scenarios in the twenty-first century. Impacts Clim. Change Fish. Aquac. 63.
Christensen, V., Walters, C.J., 2004. Ecopath with Ecosim: methods, capabilities and limitations. Ecol. Model. 172 (2–4), 109–139.
Cox, S.L., Embling, C.B., Hosegood, P.J., Votier, S.C., Ingram, S.N., 2018. Oceanographic drivers of marine mammal and seabird habitat-use across shelf-seas: a guide to key features and recommendations for future research and conservation management. Estuar. Coast. Shelf Sci. 212, 294–310.
Daewel, U., Akhtar, N., Christiansen, N., Schrum, C., 2022. Offshore wind farms are projected to impact primary production and bottom water deoxygenation in the North Sea. Commun. Earth Environ. 3 (1), 292.
De Boer, G.J., Pietrzak, J.D., Winterwerp, J.C., 2008. Using the potential energy anomaly equation to investigate tidal straining and advection of stratification in a region of freshwater influence. Ocean Model 22 (1–2), 1–11.
De Dominicis, M., Wolf, J., O'Hara Murray, R., 2018. Comparative effects of climate change and tidal stream energy extraction in a shelf sea. J. Geophys. Res. Oceans 123 (7), 5041–5067.
Doney, S.C., Ruckelshaus, M., Duffy, J.E., Barry, J.P., Chan, F., English, C.A., Galindo, H. M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., 2012. Climate change impacts on marine ecosystems. Annu. Rev. Mar. Sci. 4 (2012), 11–37.
Dorrell, R.M., Lloyd, C.J., Lincoln, B.J., Rippeth, T.P., Taylor, J.R., Caulfield, C.P., Sharples, J., 2022. Anthropogenic mixing in seasonally stratified shelf seas by offshore wind farm infrastructure. Front. Mar. Sci. 9, 1–25.
Dunne, J.A., Williams, R.J., Martinez, N.D., 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecol. Lett. 5 (4), 558–567.
Edwards, M., Atkinson, A., Bresnan, E., Helaouet, P., McQuatters-Gollop, A., Ostle, C., Pitois, S., Widdicombe, C., 2020. Plankton, jellyfish and climate in the north-East Atlantic. MCCIP Sci. Rev. 2020, 322–353.
Elliott, M., O'reilly, M.G., Taylor, C.J.L., 1990. The forth estuary: a nursery and overwintering area for North Sea fishes. Hydrobiologia 195, 89–103.
Evans, P., Waggitt, J., 2020. Impacts of Climate Change on Marine Mammals, Relevant to the Coastal and Marine Environment around the UK.
Faisal, A., Dondelinger, F., Husmeier, D., Beale, C.M., 2010. Inferring species interaction networks from species abundance data: a comparative evaluation of various statistical and machine learning methods. Eco. Inform. 5 (6), 451–464.
Fernandes, J.A., Cheung, W.W., Jennings, S., Butenschön, M., de Mora, L., Frölicher, T.L., Barange, M., Grant, A., 2013. Modelling the effects of climate change on the distribution and production of marine fishes: accounting for trophic interactions in a dynamic bioclimate envelope model. Glob. Chang. Biol. 19 (8), 2596–2607.
Friedman, N., Goldszmidt, M., Wyner, A., 1999. Data Analysis with Bayesian Networks: A Bootstrap Approach. Morgan Kaufmann Publishers Inc, pp. 196–205.
García Molinos, J., Halpern, B.S., Schoeman, D.S., Brown, C.J., Kiessling, W., Moore, P.J., Pandolfi, J.M., Poloczanska, E.S., Richardson, A.J., Burrows, M.T., 2016. Climate velocity and the future global redistribution of marine biodiversity. Nat. Clim. Chang. 6 (1), 83–88.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian data analysis. Chapman and Hall/CRC.
Gonzalez-Pola, C., Larsen, K.M.H., Fratantoni, P., Beszczynska-Moller, A., 2019. ICES report on ocean climate 2018. In: ICES Cooperative Research Report, 349.
Harris and Wanless, 1998. Isle of May Seabird Studies in 1998, 282. JNCC Report.

Heath, M.R., Speirs, D.C., Thurlbeck, I., Wilson, R.J., 2021. StrathE2E2: an R package for modelling the dynamics of marine food webs and fisheries. Methods Ecol. Evol. 12 (2), 280–287.
Holt, J., Schrum, C., Cannaby, H., Daewel, U., Allen, I., Artioli, Y., Bopp, L., Butenschon, M., Fach, B.A., Harle, J., Pushpadas, D., 2016. Potential impacts of climate change on the primary production of regional seas: a comparative analysis of five European seas. Prog. Oceanogr. 140, 91–115.
Hui, E., Stafford, R., Matthews, I.M., Smith, V.A., 2022. Bayesian networks as a novel tool to enhance interpretability and predictive power of ecological models. Eco. Inform. 68, 101539.
Hunsicker, M.E., Ciannelli, L., Bailey, K.M., Buckel, J.A., Wilson White, J., Link, J.S., Essington, T.E., Gaichas, S., Anderson, T.W., Brodeur, R.D., Chan, K.S., 2011. Functional responses and scaling in predator–prey interactions of marine fishes: contemporary issues and emerging concepts. Ecol. Lett. 14 (12), 1288–1299.
Jiao, Y., 2009. Regime shift in marine ecosystems and implications for fisheries management, a review. Rev. Fish Biol. Fish. 19, 177–191.
Levin, S.A., 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73 (6), 1943–1967.
Link, J.S., Ihde, T.F., Harvey, C.J., Gaichas, S.K., Field, J.C., Brodziak, J.K.T., Townsend, H.M., Peterman, R.M., 2012. Dealing with uncertainty in ecosystem models: the paradox of use for living marine resource management. Prog. Oceanogr. 102, 102–114.
Lotze, H.K., Tittensor, D.P., Bryndum-Buchholz, A., Eddy, T.D., Cheung, W.W., Galbraith, E.D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J.L., Bopp, L., 2019. Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. Proc. Natl. Acad. Sci. 116 (26), 12907–12912.
Lynam, C.P., Llope, M., Möllmann, C., Helaouët, P., Bayliss-Brown, G.A., Stenseth, N.C., 2017. Interaction between top-down and bottom-up control in marine food webs. Proc. Natl. Acad. Sci. 114 (8), 1952–1957.
Milns, I., Beale, C.M., Smith, V.A., 2010. Revealing ecological networks using Bayesian network inference algorithms. Ecology 91 (7), 1892–1899.
Möllmann, C., Müller-Karulis, B., Kornilovs, G., St John, M.A., 2008. Effects of climate and overfishing on zooplankton dynamics and ecosystem structure: regime shifts, trophic cascade, and feedback loops in a simple ecosystem. ICES J. Mar. Sci. 65 (3), 302–310.
Montesinos López, O.A., Montesinos López, A., Crossa, J., 2022. Overfitting, model tuning, and evaluation of prediction performance. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer International Publishing, Cham, pp. 109–139. https://doi.org/10.1007/978-3-030-89010-0_4.
Murphy, K., 2001. The bayes net toolbox for matlab. In: Computing Science and Statistics, 33.
Murphy, K.P., 2002. Dynamic Bayesian Networks: Representation, Inference and Learning. University of California, Berkeley.
Nabney, I.T., Cheng, H.W., 1997. Estimating conditional volatility with neural networks. In: Fourth Internation Conference: Forecasting Financial Markets.
Polis, G.A., Holt, R.D., Menge, B.A., Winemiller, K.O., 1996. Time, space, and life history: Influences on food webs. In: Food Webs: Integration of Patterns & Dynamics. Springer US, Boston, MA, pp. 435–460.
Powley, H.R., Bruggeman, J., Hopkins, J., Smyth, T., Blackford, J., 2020. Sensitivity of shelf sea marine ecosystems to temporal resolution of meteorological forcing. J. Geophys. Res. Oceans 125 (7), e2019JC015922.
Ramampiandra, E.C., Scheidegger, A., Wydler, J., Schuwirth, N., 2023. A comparison of machine learning and statistical species distribution models: quantifying overfitting supports model interpretation. Ecol. Model. 481, 110353.
Sadykova, D., Scott, B.E., De Dominicis, M., Wakelin, S.L., Sadykov, A., Wolf, J., 2017. Bayesian joint models with INLA exploring marine mobile predator–prey and competitor species habitat overlap. Ecol. Evol. 7 (14), 5212–5226.
Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., Van Nes, E.H., Rietkerk, M., Sugihara, G., 2009. Early-warning signals for critical transitions. Nature 461 (7260), 53–59.
Sea Mammal Research Unit, St. Andrews. Special Committee on Seals (SCOS), 2022. Scientific Advice on Matters Related to the Management of Seal Populations: 2022. Available at: https://www.smru.st-andrews.ac.uk/scos/scos-reports/index.html.
Simoen, E., De Roeck, G., Lombaert, G., 2015. Dealing with uncertainty in model updating for damage assessment: a review. Mech. Syst. Signal Process. 56, 123–149.
Simpson, J.H., Bowers, D., 1981. Models of stratification and frontal movement in shelf seas. Deep Sea Res Part A 28 (7), 727–738.
Simpson, S.D., Jennings, S., Johnson, M.P., Blanchard, J.L., Schön, P.J., Sims, D.W., Genner, M.J., 2011. Continental shelf-wide response of a fish assemblage to rapid warming of the sea. Curr. Biol. 21 (18), 1565–1570.
Sivia, D., Skilling, J., 2006. Data Analysis: A Bayesian Tutorial. OUP Oxford.
Trifonova, N.I., Scott, B.E., 2024. Ecosystem indicators: predicting population responses to combined climate and anthropogenic changes in shallow seas. Ecography 2024 (3), e06925.
Trifonova, N., Duplisea, D., Kenny, A., Maxwell, D., Tucker, A., 2014. Incorporating regime metrics into latent variable dynamic models to detect early-warning signals of functional changes in fisheries ecology. In: Discovery Science: 17th International Conference, DS 2014, Bled, Slovenia, October 8-10, 2014. Proceedings 17. Springer International Publishing, pp. 301–312.
Trifonova, N., Kenny, A., Maxwell, D., Duplisea, D., Fernandes, J., Tucker, A., 2015. Spatio-temporal Bayesian network models with latent variables for revealing trophic dynamics and functional networks in fisheries ecology. Eco. Inform. 30, 142–158.
Trifonova, N., Maxwell, D., Pinnegar, J., Kenny, A., Tucker, A., 2017. Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model. ICES J. Mar. Sci. 74 (5), 1334–1343.
Trifonova, N.I., Scott, B.E., De Dominicis, M., Waggitt, J.J., Wolf, J., 2021. Bayesian network modelling provides spatial and temporal understanding of ecosystem dynamics within shallow shelf seas. Ecol. Indic. 129, 107997.
Trifonova, N., Scott, B., De Dominicis, M., Wolf, J., 2022. Use of our future seas: relevance of spatial and temporal scale for physical and biological indicators. Front. Mar. Sci. 8, 769680.
Tucker, A., Duplisea, D., 2012. Bioinformatics tools in predictive ecology: applications to fisheries. Philos. Trans. R. Soc. B 367 (1586), 279–290.
Tucker, A., Liu, X., 2004. A Bayesian network approach to explaining time series with changing structure. Intellig. Data Analys. 8 (5), 469–480.
Uusitalo, L., 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. 203 (3–4), 312–318.
Uusitalo, L., Tomczak, M.T., Müller-Karulis, B., Putnis, I., Trifonova, N., Tucker, A., 2018. Hidden variables in a Dynamic Bayesian network identify ecosystem level change. Eco. Inform. 45, 9–15.
van der Molen, J., Smith, H.C., Lepper, P., Limpenny, S., Rees, J., 2014. Predicting the large-scale consequences of offshore wind turbine array development on a North Sea ecosystem. Cont. Shelf Res. 85, 60–72.
Van Leeuwen, S.M., Le Quesne, W.F., Parker, E.R., 2016. Potential future fisheries yields in shelf waters: a model study of the effects of climate change and ocean acidification. Biogeosciences 13 (2), 441–454.
Waggitt, J.J., Evans, P.G., Andrade, J., Banks, A.N., Boisseau, O., Bolton, M., Bradbury, G., Brereton, T., Camphuysen, C.J., Durinck, J., Felce, T., 2020. Distribution maps of cetacean and seabird populations in the North-East Atlantic. J. Appl. Ecol. 57 (2), 253–269.
Wakefield, E.D., Owen, E., Baer, J., Carroll, M.J., Daunt, F., Dodd, S.G., Green, J.A., Guilford, T., Mavor, R.A., Miller, P.I., Newell, M.A., 2017. Breeding density, fine-scale tracking, and large-scale modeling reveal the regional distribution of four seabird species. Ecol. Appl. 27 (7), 2074–2091.
Wakelin, S.L., Artioli, Y., Butenschön, M., Allen, J.I., Holt, J.T., 2015. Modelling the combined impacts of climate change and direct anthropogenic drivers on the ecosystem of the northwest European continental shelf. J. Mar. Syst. 152, 51–63.
Wiens, J.A., 1989. Spatial scaling in ecology. Funct. Ecol. 3 (4), 385–397.
Wiens, J.A., 1990. On the use of grain and grain size in ecology. Funct. Ecol. 4 (5).
Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84 (6), 1382–1394.
Yıldırım, H.C., Liaw, P.K., 2024. Data-driven conditional probability to predict fatigue properties of multi-principal element alloys (MPEAs). Comput. Methods Appl. Mech. Eng. 432, 117358.
Zampollo, A., Murray, R.O.H., Gallego, A., Scott, B., 2025. Does the oceanographic response to wind farm wind-wakes affect the spring phytoplankton bloom? Prog. Oceanogr. 237, 103512.