Prediction of a complex system with few data: Evaluation of the effect of model structure and amount of data with dynamic bayesian network models

Ana D. Maldonado, Laura Uusitalo, Alan Tucker, Thorsten Blenckner, Pedro A. Aguilera, Antonio Salmerón

Published in:
Environmental Modelling & Software

DOI (link to publication from publisher):
https://doi.org/10.1016/j.envsoft.2019.04.011

Publication date:
2019

Document version:
Accepted author manuscript, peer reviewed version

Citation for published version:
Maldonado, A. D., Uusitalo, L., Tucker, A., Blenckner, T., Aguilera, P. A., & Salmerón, A. (2019). Prediction of a complex system with few data: evaluation of the effect of model structure and amount of data with dynamic bayesian network models. Environmental modelling & software, 118, 281-297. https://doi.org/10.1016/j.envsoft.2019.04.011

Prediction of a complex system with few data: Evaluation of the effect of model structure and amount of data with dynamic Bayesian network models

A. D. Maldonado[a,*], L. Uusitalo[b], A. Tucker[c], T. Blenckner[d], P. A. Aguilera[e], A. Salmerón[a,f]

[a] *University of Almería, Department of Mathematics, Almería, Spain*
[b] *Finnish Environment Institute, Marine Research Centre, Helsinki, Finland*
[c] *Brunel University London, Department of Computer Science, Uxbridge, United Kingdom*
[d] *Stockholm University, Stockholm Resilience Centre, Stockholm, Sweden*
[e] *University of Almería, Department of Biology and Geology, Almeria, Spain*
[f] *University of Almería, Center for the Development and Transfer of Mathematical Research to Industry (CDTIME), Almería, Spain*

Abstract

A major challenge in environmental modeling is to identify structural changes in the ecosystem across time, i.e., changes in the underlying process that generates the data. In this paper, we analyze the Baltic Sea food web in order to 1) examine potential unobserved processes that could affect the ecosystem and 2) make predictions on some variables of interest. To do so, dynamic Bayesian networks with different setups of hidden variables (HVs) were built and validated applying two techniques: rolling-origin and rolling-window. Moreover, two statistical inference approaches were compared at regime shift detection: fully Bayesian and Maximum Likelihood Estimation. Our results confirm that, from the predictive accuracy point of view, more data help to improve the predictions whereas the different setups of HVs did not make a critical difference in the predictions. Finally, the different HVs picked up patterns in the data, which revealed changes in different parts of the ecosystem.

Keywords: Baltic Sea, ecosystem model, model comparison, regime shift, structural change, hidden variable

*Corresponding author
Email addresses: ana.d.maldonado@ual.es (A. D. Maldonado), Laura.Uusitalo@ymparisto.fi (L. Uusitalo), Allan.Tucker@brunel.ac.uk (A. Tucker), thorsten.blenckner@su.se (T. Blenckner), aguilera@ual.es (P. A. Aguilera), antonio.salmeron@ual.es (A. Salmerón)

1. Introduction

Ecosystems are constantly changing in response to both gradual and abrupt natural and human-induced changes, such as changes in climate, land-use, or new species arrival. In some cases, these major changes can lead to abrupt shifts, i.e., regime shifts, affecting the structure and function of ecosystem dynamics [1, 2] that are often costly and hard to reverse [3]. For instance, the Central Baltic Sea has undergone at least two regime shifts. One was induced by constant nutrient loading from land, causing a change from an oligotrophic to a eutrophic state in the 1960s, resulting in harmful algal blooms and anoxic bottoms [4]. Another regime shift occurred in the late 1980s [5, 6] and was induced by overfishing of cod (*Gadus morhua*) and climate, causing a shift toward a sprat (*Sprattus sprattus*) dominated state [6], which affected the magnitude of food web processes [7]. The latter regime shift shows a clear indication of a hysteresis effect as, even after the reduction of cod fishing, the cod biomass could not be recovered [8, 9]. Such non-stationary changes in ecosystem dynamics pose a challenge to ecosystem modelers and data analysts, since it may be that the functional forms describing the relationships between the variables change.

Bayesian networks (BNs), which belong to the probabilistic graphical models, are compact representations of the joint probability distribution over a set of variables whose independence relations are encoded by the structure of an underlying directed acyclic graph [10]. Since BNs do not explicitly model changes over time, a dynamic approach could represent more realistic modeling [11]. Dynamic BNs extend the concept of BN by explicitly modeling change over time, i.e., they allow the representation of the relationship between variables at successive time steps [12]. It is normally assumed that the model structure is the same in each time step and the parameters over time do not change, i.e., the model is assumed to be time-invariant. However, it is possible to add hidden

nodes to represent non-stationary processes [13]. Hidden variables are unobserved variables that might represent relevant processes in the system that can help explain some observed variables of interest [14]. Generally speaking, the observed variables are not the only ones that affect the system, i.e., there are a number of unobserved variables and processes that could have an effect but have not been identified, or no data are available [15]. The value of the HVs can be inferred from the observed variables linked to them so that a change in the HV pattern reflects a change in the system. Therefore, dynamic models with hidden variables are one way of trying to find the signal of change amongst the multiple ecosystem variables and their interactions [16, 15].

The work by [15] evaluated the potential of DBNs with hidden variables in the regime shift analysis on the Baltic Sea food web by linking different configurations of hidden variables to the core structure of the DBN. In their study, they built three versions of the model, which differed in the hidden variable setup, with the core structure being the same, in order to analyze the pattern of the different HVs. This work allowed to investigate whether or not the pattern of the HVs reflecting specific parts of the system (such as the fish dynamics) could be separated from the global ecosystem dynamic. They found out that the different model setups showed the same general patterns. They discussed the relative scarcity of data but did not assess how much data is needed in order to implement this kind of model, or whether some setup of HVs performs better predictions than others. Our new study extends this work by answering the aforementioned questions and considering new configurations of hidden variables. On the one hand, we analyzed the amount of data needed to discover the HV pattern by fitting the models multiple times with an increasing amount of data. In addition, we compared the ability of two different statistical inference approaches to detect changes in the ecosystem: Bayesian and Maximum Likelihood Estimation. On the other hand, we explored the predictive power of our expert-based structure models and compared the results with the ones obtained by a fixed structure model. Moreover, we examined the model's predictive performance across time by i) increasing the sample size, in order to detect if any

of the models constantly outperforms the others, and ii) discarding old data as new information is available, in order to determine either if the variations in the model's performance are due to the amount of data available or if some parts of the time series are easier to predict than others.

2. Material and Methods

2.1. Bayesian networks and dynamic Bayesian networks

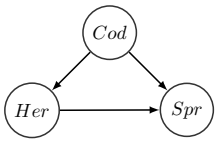
A Bayesian network (BN) is a statistical multivariate model for a set of variables $\mathbf{X} = \{X_1, \ldots, X_n\}$, which is defined in terms of two components:

- Qualitative component: A directed acyclic graph (DAG) where each vertex represents one of the variables in the model, and so that the presence of an edge linking two variables indicates the existence of statistical dependence between them.
- Quantitative component: A conditional distribution $p(x_i|pa(x_i))$ for each variable X_i, $i = 1, \ldots, n$ given its parents in the graph, denoted as $pa(X_i)$.

The joint distribution of the variables in the network is therefore represented in a factorized way as

$$p(X) = \prod_{i=1}^{n} p(x_i|pa(x_i)), \qquad \forall x_1, \ldots, x_n \in \Omega_{X_1,\ldots,X_n}, \tag{1}$$

where $X = x_1, \ldots, x_n$, Ω_{X_i} represents the set of all possible values of variable x_i and $pa(x_i)$ denotes an instantiation of the parents of X_i. Figure 1 shows an example of a DAG and its joint probability distribution.



$$P(Cod, Her, Spr) = P(Cod)P(Her|Cod)P(Spr|Cod, Her)$$

Figure 1: Example of Bayesian network: DAG (left) and joint probability distribution (right).

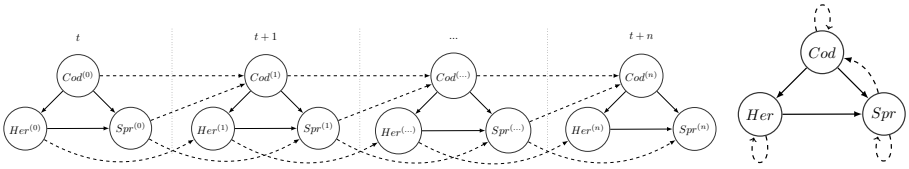
Dynamic Bayesian networks (DBN) extend the concept of BNs by relating variables across time. DBNs are defined as a pair (B_1, B_{2T}) [17], where B_1 is

a classical BN representing the first time slice ($t = 0$) and B_{2T} represents the transition model, i.e, a two-slice DBN for $t > 0$. Therefore, the joint probability distribution of $t = 0$ is the same as Equation (1), whereas the joint probability distribution of the following time slices ($t > 0$) is

$$p(X^{(t)}|X^{(t-1)}) = \prod_{i=1}^{n} p(x_i^{(t)}|pa(x_i^{(t)})), \tag{2}$$

where $x_i^{(t)}$ is the i^{th} node at time t and $pa(x_i^{(t)})$ are the parents of $x_i^{(t)}$ in the graph.

The parents ($pa(x_i^{(t)})$) of a node ($x_i^{(t)}$) can either be in the same time slice or in the previous one, assuming a first-order Markov process. We are making two other assumptions: 1) $X^{(t-1)} \perp X^{(t+1)}|X^{(t)}$, the Markov property, i.e, the future is independent of the past given the present; and 2) the transition processes are time-invariant, i.e. the transition functions are the same for all time slices. Figure 2 shows an example of unrolled DBN.



(a) Unrolled DBN

(b) Unconventional notation used in this paper

Figure 2: Example of DAG of a DBN. Dependencies within a time slice are represented by solid edges whereas dependencies between time slices are represented by dashed edges. The joint probability distribution of the transition model for this DAG can be written as $P(Cod^t, Her^t, Spr^t|Cod^{t-1}, Her^{t-1}, Spr^{t-1}) = P(Cod^t|Cod^{t-1}, Spr^{t-1})P(Her^t|Cod^t, Her^{t-1})P(Spr^t|Cod^t, Her^t, Spr^{t-1})$.

2.2. Parameter learning

The dataset used in this paper contains only continuous variables, which were parameterized using linear Gaussian distributions. Parameters can be learned from data using the Expectation-Maximization (EM) algorithm [18]. The goal of the EM algorithm is to find the maximum likelihood estimate of the model parameters when the data have missing values. The EM algorithm iteratively finds (E-step) and maximizes (M-step) a current approximation to the log-likelihood function of the parameters of a model [19]. The algorithm must be initialized with an initial value of the parameters. In this work, the initial random values of the parameters were drawn from the standard normal distribution, N(0,1). Then, the E-step computes the expected sufficient statistics (mean and variance) using the current parameter values and the observed data. Afterward, the M-step maximizes the log-likelihood of the parameters given the sufficient statistics, obtaining an updated value of the parameter estimate. These two steps iterate until convergence; in our case, until the difference between consecutive log-likelihoods was small enough. Since the EM algorithm can get stuck in a local optimum, it was run 100 times for each model, retaining the model with the highest log-likelihood value for further analyses.

2.3. The Baltic Sea food web model

2.3.1. Data description

The data originate from the Gotland Basin in the central Baltic Sea (Figure 3), covering the 38-year period from 1975 to 2012 (Table 1). They were obtained from different sources: the fish data are derived from the fish stock assessment model called Virtual Population Analysis [20], tuned using the Extended Survival Analysis (XSA) method [21]. This model uses both fish catch and fish survey data from multiple years as input data and fits an age-structured fish stock model that accounts for the amount and mean weight of fish in each year class. It estimates the total biomass of the spawning stock as well as the number of fish in the next year class (recruitment). The reproductive volume of cod, i.e., the volume of water in which the environmental factors (in particular,

salinity and oxygen concentration) are such that cod eggs survive, is based on a spatial interpolation model [22]. Water temperature, zooplankton, and chlorophyll variables are based on direct measurements taken during field sampling campaigns. The data are highly variable, both temporally and spatially, and therefore the data tend to be noisy, as the observation may vary considerably depending on the sampling date and the exact location of the sampling. In order to avoid numerical instability, the data were log-transformed and standardized to mean 0 and standard deviation 1.

Table 1: Summary of the observed variables used to build the models.

Variable	**Description (units)**
FCod [a]	Fishing-induce mortality of cod (instantaneous fishing mortality)
Cod2y [a]	Abundance of juvenile cod, age 2 (number of individuals, in thousands)
SSBCod [a]	Spawning stock biomass of cod (metric tonnes)
FHer [a]	Fishing-induce mortality of herring (instantaneous fishing mortality)
Her1y [a]	Abundance of juvenile herring, age 1 (number of individuals, in thousands)
SSBHer [a]	Spawning stock biomass of herring (metric tonnes)
FSpr [a]	Fishing-induce mortality of sprat (instantaneous fishing mortality)
Spr1y [a]	Abundance of juvenile sprat, age 1 (number of individuals, in thousands)
SSBSpr [a]	Spawning stock biomass of sprat (metric tonnes)
Ac [b]	Biomass of zooplankton genus *Acartia*, measured in spring (mg/m^3)
Tem [b]	Biomass of zooplankton genus *Temora*, measured in spring (mg/m^3)
Ps [b]	Biomass of zooplankton genus *Pseudocalanus*, measured in summer (mg/m^3)
Chla [c]	Chlorophyll *a* concentration in water, measured in spring (mg/m^3)
Tspring [d]	Sea surface temperature in spring (°C)
Tsum [d]	Sea surface temperature in summer (°C)
RV [e]	Reproductive volume of cod, i.e., volume of water in which the cod eggs can survive (km^3)

Data source: [a] ICES stock assessment, XSA model [23]; [b] Quantitative Juday net sampling from 0-100 m water layer; [c] Quantitative measurement from water sample 0-10 m water layer; [d] Measurement from water sample; [e] Spatial interpolation model [22].

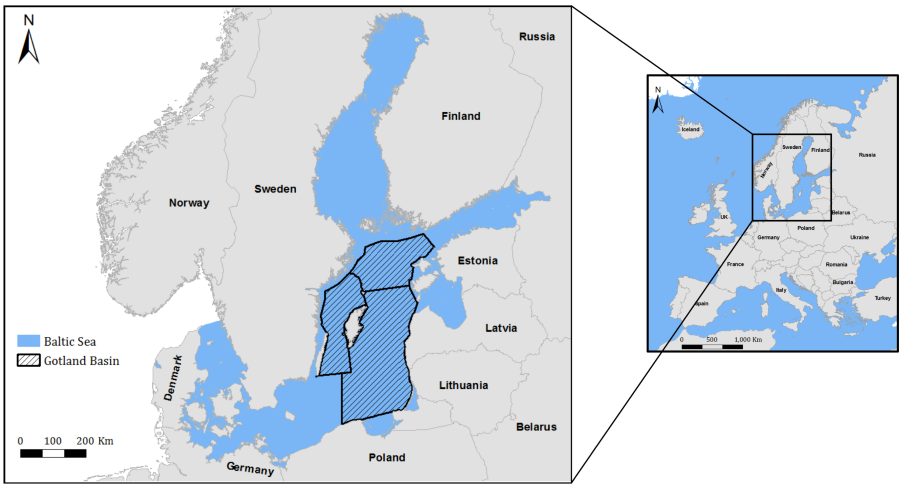


Figure 3: Location of the study area, the Gotland basin (central Baltic Sea, Europe). Source: Map created from layers *'Europe'* (downloaded from http://tapiquen-sig.jimdo.com) and *'HELCOM subbasins with coastal and offshore division 2018'* (downloaded from https://maps.helcom.fi), using the ArcGis software by Esri.

2.3.2. Expert knowledge-based structure

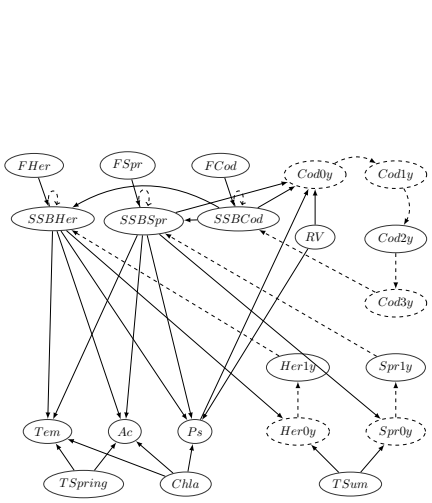
BNs can be used to encode expert knowledge over a domain [24, 25]. The graphical representation provided by BNs makes them a transparent tool, which is especially useful when expert elicitation is required [26, 27]. A Gotland Basin food web model structure was built based on expert elicitation (Figure 4a), including the key components of the food web and their interactions. Some of the variables in this model are only linked to variables in the contiguous time slices, which is essential to transmit the temporal dynamic. Five non-observed variables, representing juvenile fish stages, were included to act as placeholders of the fish stage from birth to maturation. These unobserved variables include 0-year old herring, sprat and cod and one-year old and three-year old cod. Herring and sprat mature and join the spawning stock at the age two and cod at age four [28].

The dynamic model is defined in one-year time steps. The fish related variables have temporal dependencies across time steps: the spawning stock

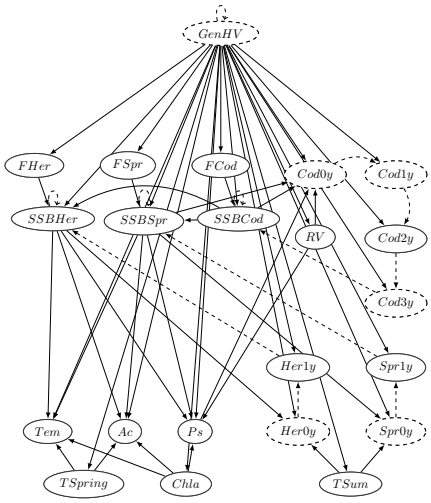
sizes (SSB) of the three fish species are autoregressive, as they consist of individual fish that live for multiple years and therefore their value in one time-slice depends on their value in the previous one. The juvenile fish are modeled separately: they exhibit temporal dependency so that the k-year old fish ($k = \{0, \ldots, 3\}$ for cod and $k = \{0, 1\}$ for herring and sprat) are k+1 years old in the following time slice until they reach maturation and join the spawning stock (SSB). The remaining variables are assumed not to directly depend on variables in the preceding year, though they may have temporal autocorrelation due to the fact that variables affecting them are temporally autocorrelated.

The variables described above form the core structure of the expert knowledge-based model structure (Figure 4a). However, it is clear that there are other variables and processes that could affect the ecosystem dynamics but have not been identified or no data are available. For this reason, models with different configurations of generic hidden variables (HVs) were built:

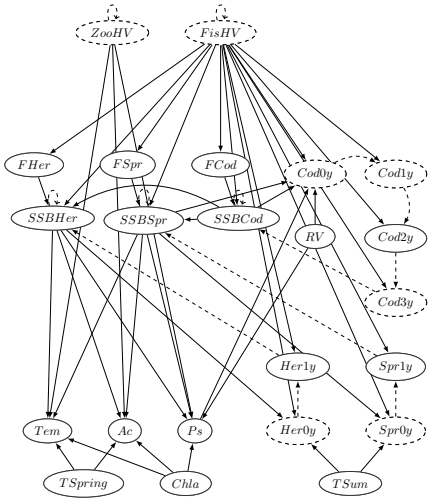
- a model with no general HVs (M0, Figure 4a);
- a model with one HV (M1, *GenHV*), linked to all other variables in each time slice and to itself in the consecutive time slices (Figure 4b);
- a model with two semi-generic HVs (M2): one linked to all fish-related variables (*FisHV*) and another linked to all zooplankton-related variables (*ZooHV*) in each time slice and to themselves in the consecutive time slices (Figure 4c); and
- a model (M2r) including *FisHV* and *ZooHV* (Figure 4d), as in M2 but reversing the links between the variables *SSBHer* and *SSBSpr* and the zooplankton variables. This was done for food web reasons (energy flows from phytoplankton to zooplankton to fish) and for computer science reasons (in model M2, variables *SSBHer* and *SSBSpr* and hidden variable *ZooHV* are conditionally dependent given the zooplankton variables, due to the v-structure; by reversing the links, they are conditionally independent given the zooplankton variables).



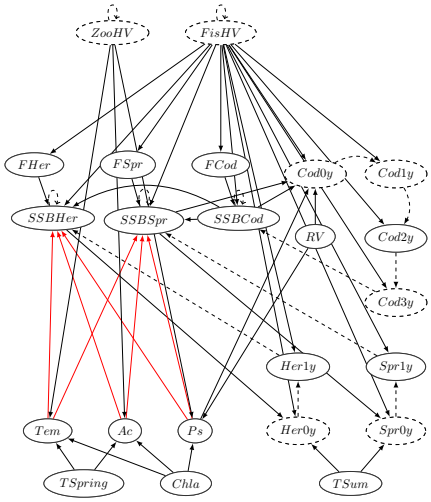
(a) Model M0: no HV.



(b) Model M1: 1 generic HV (*GenHV*).



(c) Model M2: 2 HVs (*FisHV* and *ZooHV*).



(d) Model M2r: 2 HVs (*FisHV* and *ZooHV*).

Figure 4: Expert knowledge-based model structure showing different configurations of HVs. Solid nodes represent observed variables whereas dashed nodes indicate unobserved variables. Solid edges represent direct dependencies between variables within a time slice whereas dashed edges represent direct dependencies between two time slices. SSBHer, SSBSpr, SSBCod, GenHV, ZooHV and FisHV are autoregressive variables.

2.3.3. Naive Bayes structure

In order to explore how well a simple approach would perform, we built a Naive Bayes model with a HV as the class. Naive Bayes (NB) is a BN with a fixed structure in which one variable, the class C, is the parent of all remaining variables, $X_1, \ldots, X_n$, which are independent to each other given C. The strong assumption of independence behind NB models is somehow compensated by the reduction on the number of parameters to be estimated from data, since in this case, it holds that

$$p(c|x_1, \ldots, x_n) \propto p(c) \prod_{i=1}^{n} p(x_i|c) \ , \tag{3}$$

which means that, instead of one n-dimensional conditional density, n one-dimensional conditional densities have to be estimated.

Variables included in Table 1 as well as a generic hidden variable ($GenHV$) were used to build the model. The HV takes the place of the class, which is the only autoregressive variable in the model (Figure 5). The dynamic model is defined in one-year time steps, where the links between time slices only correspond to the autoregressive hidden variable, $GenHV$.

In what follows we derive the explicit equation of the expected value of the HV for model NB. The equations for the rest of the models are derived in a similar way. For short, let's denote the variables in model NB (see Figure 5) as C=GenHV, X_1=FHer, X_2=FSpr, X_3=FCod, X_4=SSBHer, X_5=SSBSpr, X_6=SSBCod, X_7=Her1y, X_8=Spr1y, X_9=Cod2y, X_{10}=Tem, X_{11}=Ac, X_{12}=Ps, X_{13}=RV, X_{14}=TSpring, X_{15}=Chla, X_{16}=Tsum. Let $\mathbf{X}^{(t)} = (X_1^{(t)}, \ldots, X_{16}^{(t)})$ denote the observed variables at time step t. Let $\mathbf{x}^{(t)} = (x_1^{(t)}, \ldots, x_{16}^{(t)})$ be any of the possible configurations of the observed variables at time step t. Our goal is to model the expected value of the hidden variable at time step t given all the previous observations, i.e.

$$E[C^{(t)}|\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(t)}] = \int c^{(t)} p(c^{(t)}|\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(t)}) dc^{(t)}, \tag{4}$$

where

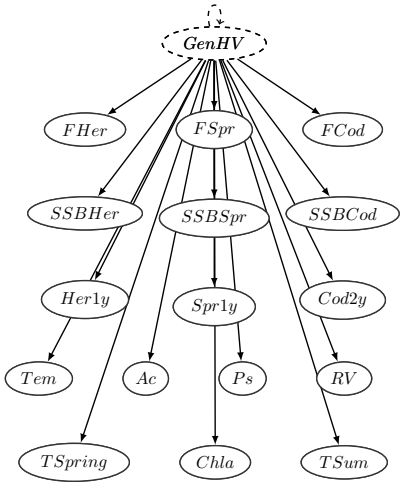


Figure 5: Naive Bayes model. Solid nodes represent observed variables whereas dashed nodes indicate unobserved variables. Solid edges represent direct dependencies between variables within a time slice whereas dashed edges represent direct dependencies between two time slices. *GenHV* is the only autoregressive variable.

$$
\begin{aligned}
p(c^{(t)}|\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(t)}) &= \int p(c^{(1)},\ldots,c^{(t)}|\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(t)})dc^{(1)}\ldots dc^{(t-1)} \\
&= \frac{1}{Z}\int p(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(t)}|c^{(1)},\ldots,c^{(t)})p(c^{(1)},\ldots,c^{(t)})dc^{(1)}\ldots dc^{(t-1)} \\
&= \frac{1}{Z}\int p(c^{(1)})\prod_{i=1}^{t}\left(p(\mathbf{x}^{(i)}|c^{(i)})p(c^{(i)}|c^{(i-1)})\right)dc^{(1)}\ldots dc^{(t-1)} \\
&= \frac{1}{Z}\int p(c^{(1)})\prod_{i=1}^{t}\left(p(c^{(i)}|c^{(i-1)})\prod_{j=1}^{16}p(x_j^{(i)}|c^{(i)})\right)dc^{(1)}\ldots dc^{(t-1)},
\end{aligned}
\tag{5}
$$

with Z being a normalization constant equal to $p(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(t)})$. Each density $p(c^{(i)}|c^{(i-1)})$ is a Gaussian density for $c^{(i)}$ with mean equal to a linear function of $c^{(i-1)}$, and each $p(x_j^{(i)}|c^{(i)})$ is a Gaussian density for $x_j^{(i)}$ with a linear function of $c^{(i)}$ as mean.

2.4. Approach comparison of regime shift detection

As an alternative to the already proposed method to detect unobserved processes, we also followed a fully Bayesian approach, i.e., including the parameters

as random variables and then updating the model by making probabilistic inference (belief update) for each record in the dataset. The goal of this comparison was to explore the ability of these two approaches at detecting changes in the ecosystem. In order to build a fully Bayesian model, we used the Naive Bayes concept drift detector (NBCD) algorithm [29] from the R package `ramidst`, which is an R interface to the AMIDST toolbox [30] written in Java. The NBCD relies on the variational Bayes framework (a class of approximation methods) for the inference and learning tasks, and uses the NB structure as the base model, with a variable of interest being the class and the remaining being predictive variables. Then, a hidden variable (H) is added and linked to the predictive variables in the model. The specific inference method used in our experiments was the streaming variational Bayes (SVB) algorithm [31].

On the other hand, we built a model with the same DAG structure as the NBCD, but using the EM algorithm to learn the parameters, as implemented in the `learn_params_dbn_em` function of the Bayes Net Toolbox in `MATLAB`. Since we are interested in six variables (SSBCod, SSBSpr, SSBHer, Cod2y, Spr1y and Her1y), we built six models using the NBCD algorithm and six using the EM algorithm. Figure 6 shows the DAG used to build the model with Her1y as the class variable. Note that the DAG is analogous for the remaining variables of interest. We will analyze the hidden variable H for each model. For the sake of clarity, we will refer to this hidden variable as H_{EM} if it was learned using the EM algorithm or H_B if it was learned following the Bayesian approach. In the case of no ambiguity or no need to distinguish between them, H will be used instead.

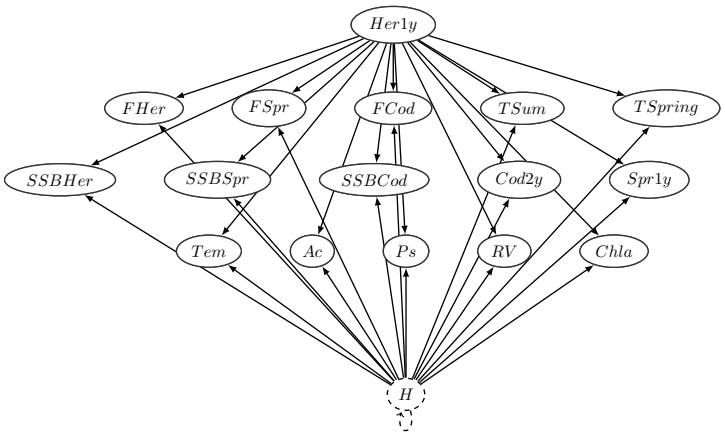


Figure 6: Naive Bayes model with one hidden variable and Her1y being the class. Solid nodes represent observed variables whereas dashed nodes indicate unobserved variables. Solid edges represent direct dependencies between variables within a time slice whereas dashed edges represent direct dependencies between two time slices.

In all the models described here and in the previous sections, the autocorrelated hidden variables try to capture the evolution of the model uncertainty over time. Furthermore, the Bayesian approach used in model NBCD takes into account the possible lack of independence in the data to some extent by updating the parameters each time a new data item arrives. In other words, the posterior distribution in a time step becomes the prior distribution in the next time step. Notice, however, that we only consider discrete time steps, as can be seen in Equation (4). Finer granularity can be achieved by adopting a continuous time approach [32] in which time becomes a variable into the model equations.

2.5. Model validation

In order to validate the predictive performance of the aforementioned models, five-step ahead predictions were carried out. Two cross-validation (CV) techniques to train and test the models were used: rolling-origin and rolling-window [33]. In the first one, the sample size increases in each fold with respect to the origin (last observation of the training set) (Figure 7a), i.e., the data from

the test set move to the train set sequentially and the model is recalibrated. In the second one, the sample size is kept constant in each fold (except for the first one due to the total amount of data), moving the training set as a window across time (Figure 7b), i.e., data from the beginning of the time series are discarded as new data are available.

The reason for doing this was to explore whether the variation in the model's performance across time is due to the amount of data available or because some parts of the time series are easier to predict than others. The goodness of the predictions was evaluated using the log-likelihood of the observations given the predicted values. We compared the log-likelihoods of the pairs rolling-origin, rolling-window of each model version using the Wilcoxon Signed Rank test.

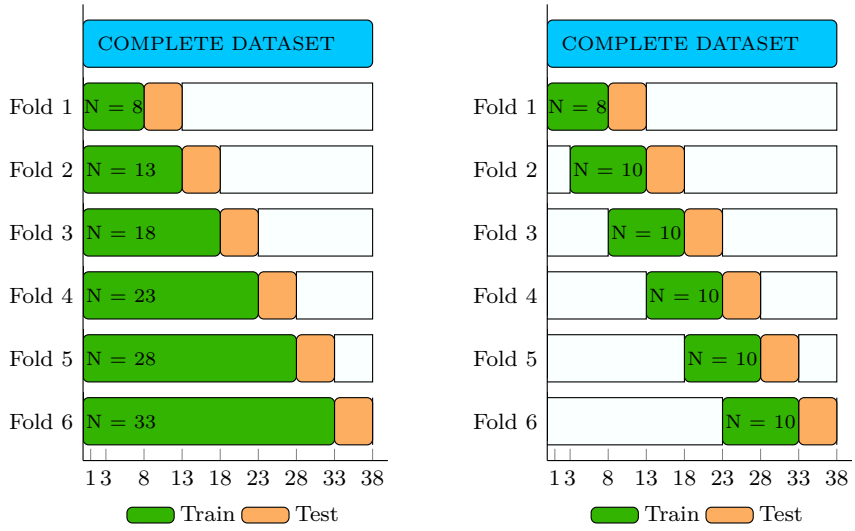
Moreover, for the rolling-origin case, the log-likelihoods of the five-year prediction of different sets of models (M0, M1, M2, M2r, NB) were compared using either Wilcoxon Signed Rank Test (for two groups) or Friedman Test with maxT statistic [34] (for more than two groups). In the case of applying Friedman Test, in those cases where significant differences were found, we deployed Wilcoxon–Nemenyi–McDonald–Thompson's post-hoc test [35].

3. Results

3.1. Regime shift detection

Models with different configurations of HVs were built in order to detect processes that could affect different parts of the system. In particular, three HVs were explored: a *general* HV (GenHV), a *fish* HV (FisHV) and a *zooplankton* HV (ZooHV). To analyze the HV patterns, we fit the models following the rolling-origin approach as shown in Figure 7a with the difference that, for each fold, both train and test sets were used to learn the models. Figure 8 shows the evolution of the expected value of the HVs in each model (M1, M2, M2r and NB) for each fold, computed as explained in Equation (4).

The three HVs (GenHV, FisHV, ZooHV) showed different patterns among them and similar to themselves (except for ZooHV) in the other models, i.e.,



(a) Rolling-origin, increasing N_{train} by 5.

(b) Rolling-window.

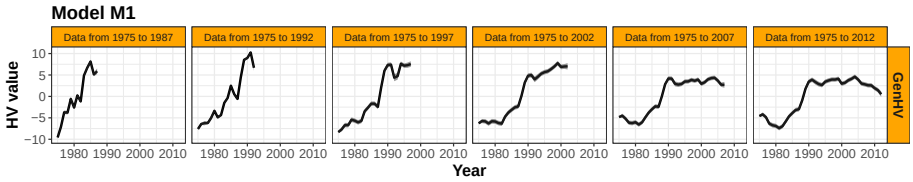
Figure 7: Sample size (N) used to train and test the models. $N_{test} = 5$.

GenHV shows similar behavior in models M1 and NB, and so does FisHV in models M2 and M2r. Regarding the last Fold, in which the complete dataset is used to fit the models, GenHV shows a decrease at the very beginning of the time series in both models (M1 and NB), followed by an increase from the period 1981-1991, point at which they stabilize till the end of the time series, when they start decreasing again. On the other hand, FisHV decreases at the beginning of the time series, then increases from 1978 to 1998 and finally decreases to its first values.

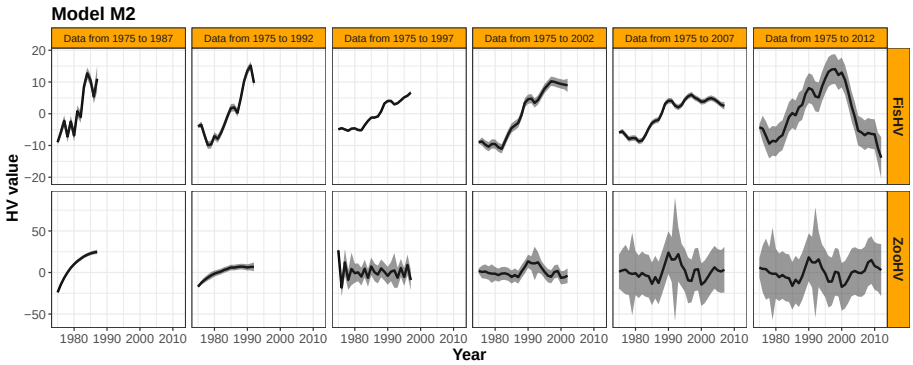
Of special interest is variable ZooHV, which shows a different pattern in models M2 and M2r, with the former having a zigzag trend and higher variance and the latter showing two stable periods separated by an abrupt increase from the mid-1980s till 1990s. The difference between ZooHV-M2 and ZooHV-M2r is a consequence of reversing the links between the zooplankton variables (Tem, Ac and Ps, which will be referred to as **Z**) and the SSBHer and SSBSpr variables (which will be referred to as **S**). By inverting these links, we transformed the v-structure involving these variables (i.e., $\mathbf{S} \rightarrow \mathbf{Z} \leftarrow$ ZooHV) to a serial connection

(i.e., $\mathbf{S} \leftarrow \mathbf{Z} \leftarrow$ ZooHV), so that the $\mathbf{S}$ variables and ZooHV are conditionally independent given $\mathbf{Z}$. In other words, given $\mathbf{Z}$, new information about $\mathbf{S}$ does not influence the ZooHV variable. On the other hand, variables TSpring and Chla are also involved in the v-structure, therefore given any of the variables in $\mathbf{Z}$ is observed, TSpring, Chla and ZooHV become conditionally dependent, in both M2 and M2r models.

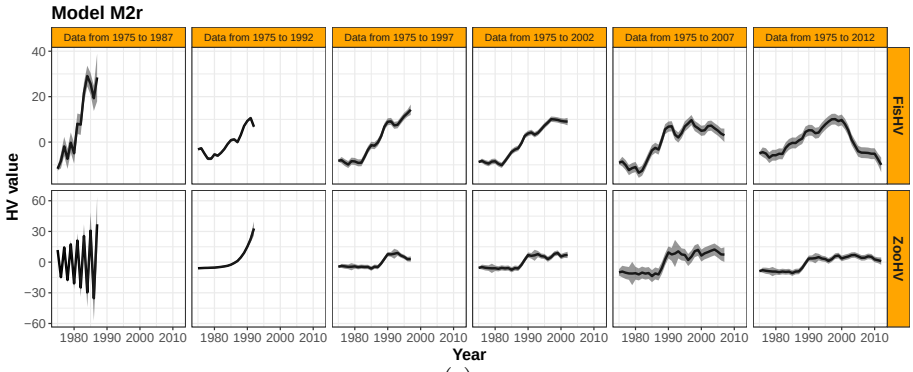
Since the size of the training set is increased by five in each fold, it is not surprising that the first and last folds show very different patterns. The fold at which the pattern of the HV is revealed depends on the specific HV. For instance, the abrupt increase of variable ZooHV (model M2r) can be seen from Fold 2 onwards (i. e., with sample size $N = 13$) whereas the stabilization phase can be observed from Fold 3. On the other hand, variable GenHV (model M1 and NB) takes a bit longer to reveal its final pattern, which can be seen from Fold 5 onwards. Finally, variable FisHV has a similar pattern to GenHV until Fold 5. Unlike GenHV, variable FisHV experiments a drastic decrease from Fold 5 to 6. It appears that variable GenHV could be capturing both FisHV and ZooHV trends, combining them in a smoother pattern.



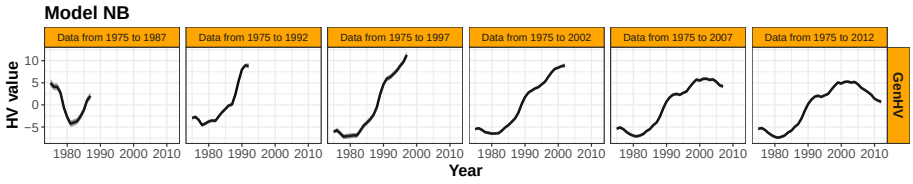
(a)



(b)



(c)



(d)

Figure 8: Evolution of the expected value of the hidden variables, increasing the size of the training set by five. Fold 1: *Data from 1975 to 1987*; Fold 2: *Data from 1975 to 1992*; Fold 3: *Data from 1975 to 1997*; Fold 4: *Data from 1975 to 2002*; Fold 5: *Data from 1975 to 2007*; Fold 6: *Data from 1975 to 2012*.

3.1.1. *Approach comparison of regime shift*

The comparison of regime shift detection between the EM algorithm and the fully Bayesian approach can be seen in Figure 9. Each individual plot shows the observed variable of interest (i.e., the *class* variable in the model) and the expected value of the hidden variable H following the EM (H_{EM}) and Bayesian (H_B) approaches.

In general, the Bayesian approach picks up the changes in the observed variables at the end of the time series. The main reason for this peculiar performance is the lack of data (it should be kept in mind that our dataset consists of only 38 data points), i.e., what we see at the beginning of the time series is the flat prior, which rules out the data likelihood until sufficient data is explored. On the other hand, the EM approach captures the trend of the observed variables earlier than the Bayesian. It can be seen that H_{EM} picks up the pattern of the observed variable acting as the *class*, even when they are noisy, as in the case of Spr1y and Her1y.

In both approaches, a change can be seen near the year 2000 in all models, which coincides to the abrupt change seen in variable FisHV in models M2 and M2r (Figures 8b and 8c) when the entire dataset is used to train the models (fold 6). Moreover, for models where Spr1y and Her1y are the *class* variable, H_{EM} shows a change near the year 1980, which coincides to the first drift identified in variable GenHV in models M1 and NB (Figures 8a and 8d) when the entire dataset is used to train the models (Fold 6).

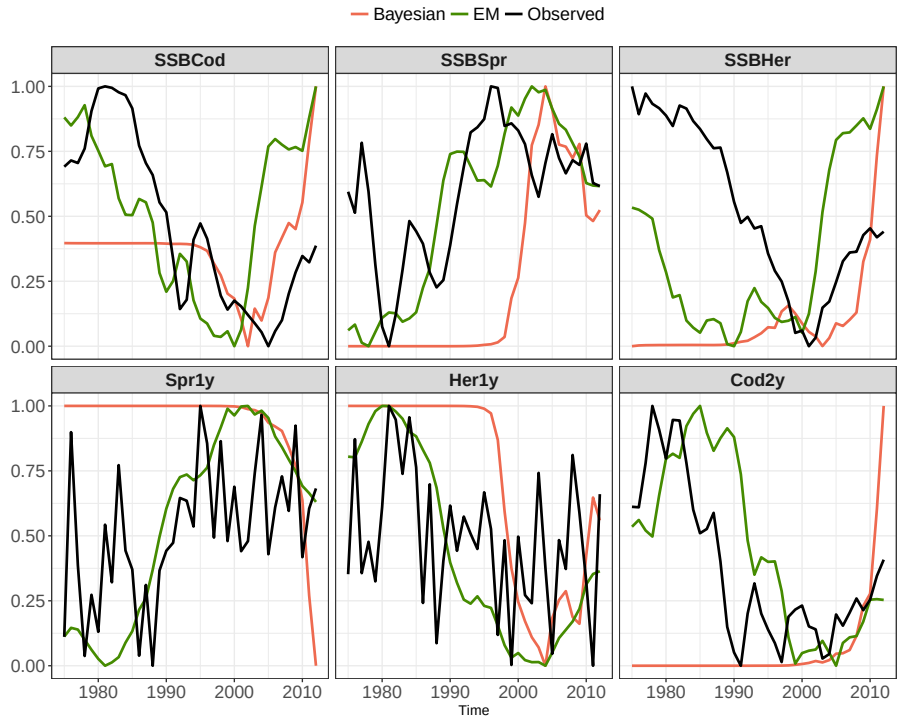


Figure 9: Evolution of the expected value of the hidden variable H for each model obtained as indicated in Section 2.5. The black line shows the observed data. For visual purposes, the data have been rescaled to interval [0,1].

3.2. Predictive performance

The time series of the observed and predicted values for each variable of interest, model and Fold of both cross-validation techniques are shown in Figures 14a to 14f in Appendix A. From these predictions, we computed the performance of each model for each variable of interest in terms of its log-likelihood (Figure 10b). Furthermore, for the rolling-origin case only, we performed some hypothesis testing, which will be further discussed in the upcoming sections.

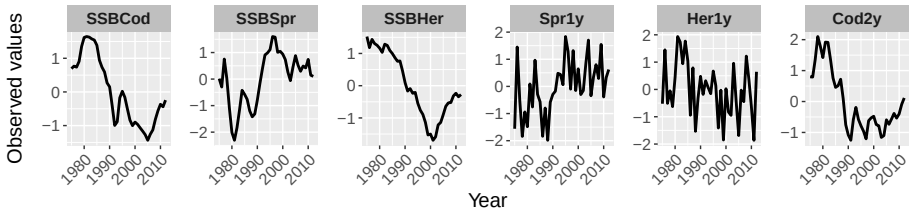
3.2.1. Does the increase of data improve the performance of the models?

In order to detect whether or not more data help to predict the variables of interest, two cross-validation techniques were used to build the models: rolling-origin (RO) and rolling-window (RW). Figure 10b shows the performance of

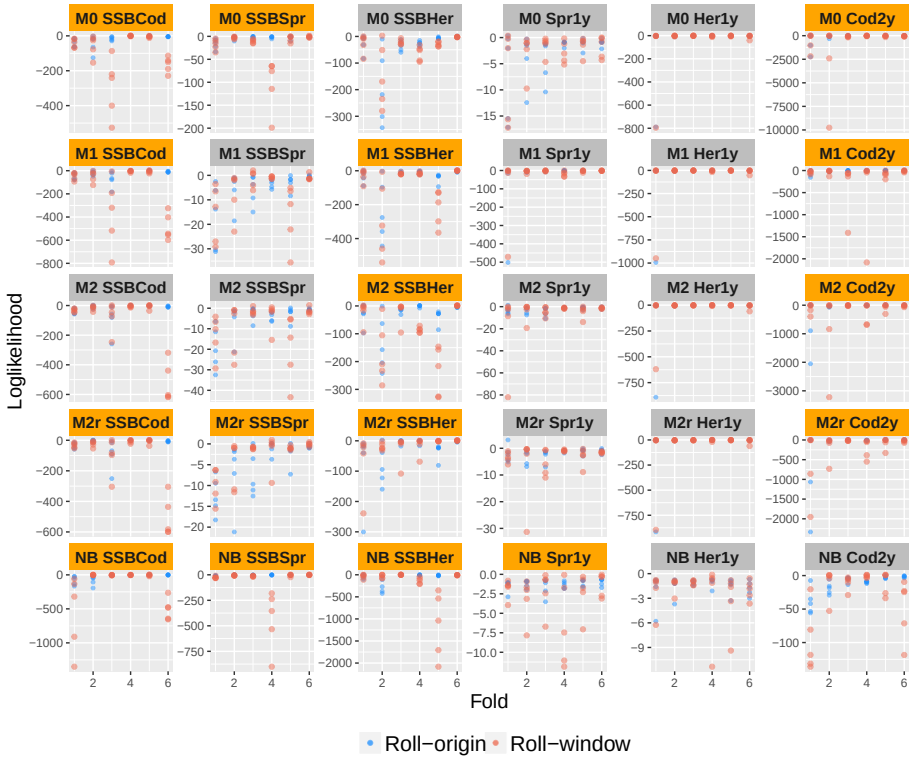
each model for each variable of interest in terms of log-likelihood. The results show that most variables get improvements in their predictions when more data are available to fit the models (p-value <0.05), i.e., when the rolling-origin technique is used. Table 2 shows an overview of the hypothesis testing results. In particular, variable SSBCod (Figure 14a) showed significant differences in all models except the M2 model, with RO performing better than RW in all significant cases; variable SSBSpr (Figure 14b) showed significant differences in 3 out of 5 models, with RO performing better than RW in models M0 and NB and RW performing better in model M2r; variable SSBHer (Figure 14c) showed significant differences in all models except the M0 model, with RO performing better than RW in 3 out of the 4 significant cases (RW performed better in the M2r model); variable Spr1y (Figure 14d) showed significant differences only in the Naive Bayes (NB) model, with RO performing better than RW; variable Her1y (Figure 14e) did not show significant differences in any model; and variable Cod2y (Figure 14f) showed significant differences in all models except the NB, with RO performing better than RW in all significant cases.

Table 2: An overview of the results of the hypothesis testing of the difference between the rolling-origin (RO) and rolling-window (RW) cross-validation techniques. For each variable of interest, the number of models that present 1) no statistically significant difference in terms of log-likelihood, 2) improvements when more data are available (RO), or 3) improvements when old data are dropped (RW), is shown.

	No differences	**RO better**	**RW better**
SSBCod	1	4	0
SSBSpr	2	2	1
SSBHer	1	3	1
Spr1y	4	1	0
Her1y	5	0	0
Cod2y	1	4	0



(a)



(b)

Figure 10: Log-transformed values of the observed variables of interest (a) and log-likelihoods of the predictions obtained from the rolling-origin and rolling-window validation techniques (b). The pairs rolling-origin - rolling-window of each variable and model version were compared using the Wilcoxon signed rank test. Orange labels indicate that the hypothesis test detected significant differences (p-value < 0.05). Figure 10b is arranged so that the five models described in Sections 2.4.2 and 2.4.3 correspond to rows and the six different variables of interest correspond to columns.

3.2.2. Does the introduction of HVs improve the performance of the models?

In order to find out whether or not the introduction of hidden variables improves the predictive performance, we compared the loglikelihood of the predictions of models M0, M1 and M2, with M0 being the model without hidden variables; M1 the model with one generic hidden variable; and M2 the model with two hidden variables. We used the Friedman Test to carry out the comparison. Figure 11 shows the boxplot of the differences between pairs of models. Boxplots outlined with color orange indicate that significant differences were found between the corresponding pair of models. The results of the tests show significant differences only for variables SSBHer and Cod2y, with models M1 and M2 outperforming model M0 in the case of Cod2y, whereas model M2 outperformed model M1 in the case of SSBHer.

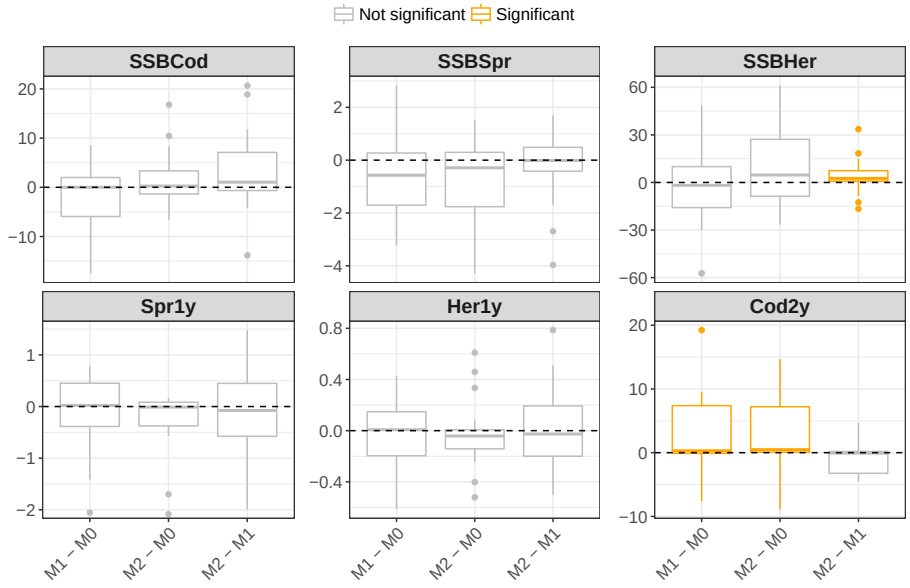


Figure 11: Boxplots of the differences in model performance. Orange outlined boxes indicates that the post-hoc analysis carried out after Friedman test detected statistical significant differences (p-value < 0.05) between the corresponding models. Note that extreme outliers have been removed from the plot.

3.2.3. Does reversing some links improve the performance of the models?

In order to test whether or not reversing the links between variables SSBHer and SSBSpr and the zooplankton variables would make any difference in the models' prediction, we compared the loglikelihoods of the predictions of models M2 and M2r using the Wilcoxon Signed Rank test. Figure 12 shows the boxplots of the differences between the pairs of models, for each variable of interest. The results of the tests show significant differences only for variables SSBHer and Spr1y, with model M2r outperforming model M2 in both cases.

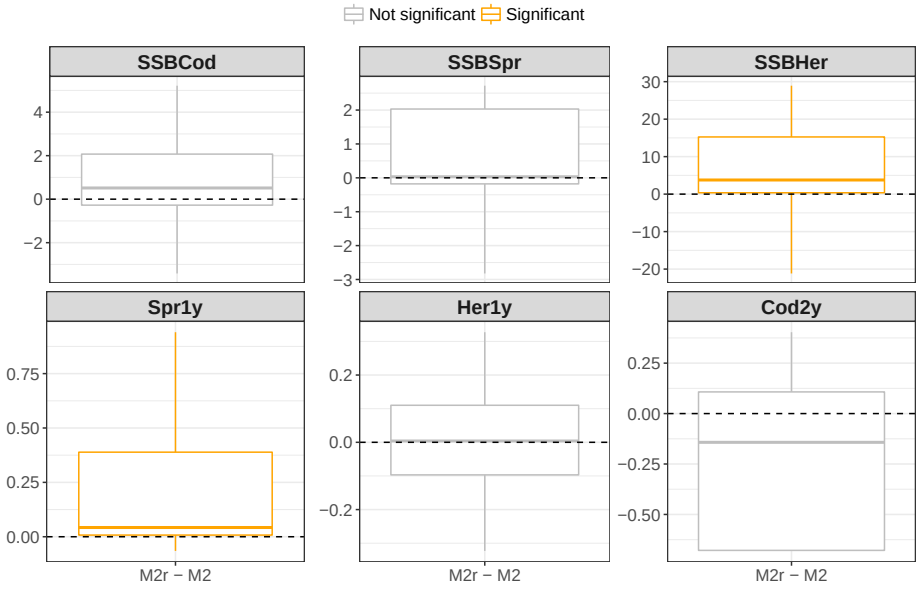


Figure 12: Boxplots of the differences in model performance. Orange outlined boxes indicates that the Wilcoxon Signed Rank test detected statistical significant differences (p-value < 0.05) between the corresponding models. Note that extreme outliers have been removed from the plot.

3.2.4. Does the use of expert-based structures improve the performance of the models?

In order to test whether or not expert-based structures outperform simpler models, such as Naive Bayes, we compared each expert-based model (M0, M1, M2, and M2r) with the Naive Bayes (NB) model using the Wilcoxon Signed

Rank test. Figure 13 shows the boxplots of the differences between the pairs of models, for each variable of interest. The results of the tests show significant differences for models SSBSpr, Spr1y, Her1y and Cod2y. In most cases, NB outperformed the expert-based models. On the other hand, for variable Cod2y, the expert-based models outperformed the NB model, except in the case of the pair M0-NB, where no significant differences were found.

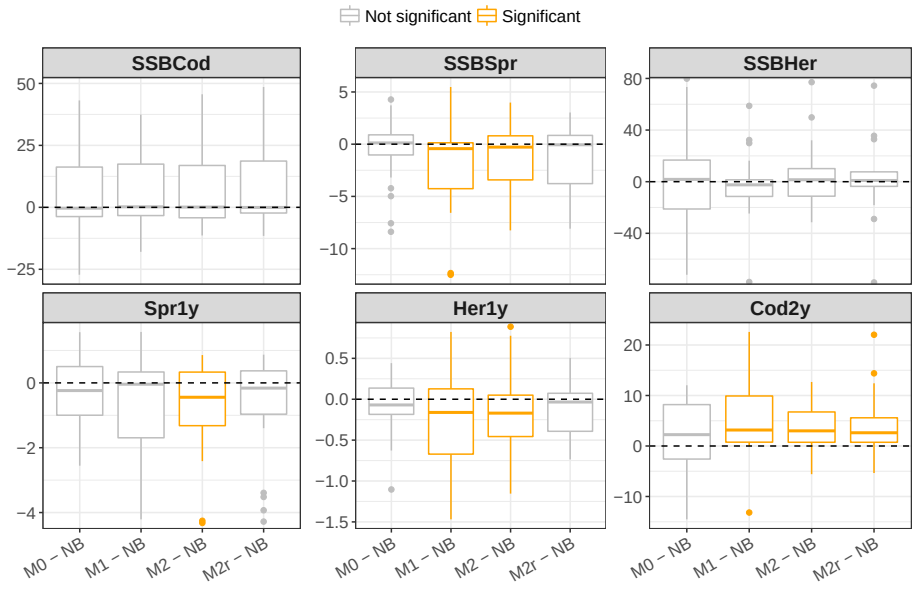


Figure 13: Boxplots of the differences in model performance. Orange outlined boxes indicates that the Wilcoxon Signed Rank test detected statistical significant differences (p-value < 0.05) between the corresponding models. Note that extreme outliers have been removed from the plot.

4. Discussion

In this paper we have analyzed the Baltic Sea food web aiming at 1) detecting changes in its pattern and 2) making reliable predictions on some variables of interest. The hidden variables called GenHV and FisHV (models M1, M2, M2r and NB) showed a clear increase near the year 1980, with the former occurring slightly later. Variable ZooHV (model M2r) shows an abrupt jump, which is

delayed in comparison to the increase of the two other HVs.

Former studies have found that the Baltic Sea has undergone several periods of change, with one of them being described in the 1980s [36]. Many publications have also described a regime shift in the North Sea in the late 80s [37, 38, 39, 40], which is connected to the Baltic Sea through the Danish Straits, where water exchange between the two areas occurs. According to the suggestion of some authors, this regime shift may have been partly caused as a response to a change of the North Atlantic Oscillation (NAO) in winter [5], which led to an increase in the winter and early spring air and water temperature. NAO-induced changes, along with overfishing of cod, triggered a regime shift [36], leading to a period dominated by clupeids, i.e., sprat and herring [4], after 1990.

The regime shift described in the literature coincides with the changes observed in variable ZooHV (in model M2r). This is in in line with the research by [5], where they describe that the abundance of zooplankton varied in accordance with the variations in the NAO. Therefore, it seems reasonable that the model reflects these variations in the NAO. On the other hand, variable FisHV (in model M2 and M2r) shows an increase from the beginning of the 1980s till the beginning of the 2000s, when its tendency drastically changes, towards a decreasing trend. This hidden variable is linked to 14 fish-related variables; therefore the hidden variable is trying to capture the global trend of all these variables. The FisHV could be partly reflecting the indirect change of zooplankton (as they partly predate on those) and partly the change in the large anoxic area [41] along with the clupeid-dominated period. Variable GenHV (models M1 an NB) shows a gradual increase from the beginning of the 1980s till 1990, where a constant period begins. The behavior of this hidden variable could be reflecting the underlying dynamic of the zooplankton and fish variables, as a whole, but more research is needed.

From the predictive accuracy point of view, significant differences among the models were scarcely found. In particular, the similar results for the expert-based models could be regarded as a positive finding in the sense that as long as the model structure is coherent, the details will not make a critical difference

in the predictions. When significant differences are found, our results suggest that models with a higher number of hidden variables are not outperformed (Section 3.2.2). Moreover, reversing some links increased the predictive accuracy of two variables only, with the remaining variables not showing statistically significant differences (Section 3.2.3). Finally, the use of simple models, such as the Naive Bayes, helped with the prediction of some variables, compared to more complex models based on expert elicitation (Section 3.2.4). The reason for that is that simple models need to estimate fewer parameters, which is very convenient when few data are available. On the other hand, improvements were found rather often when comparing the rolling-origin and rolling-window approaches. In most of the significant cases, having more data improved the prediction of the outcomes.

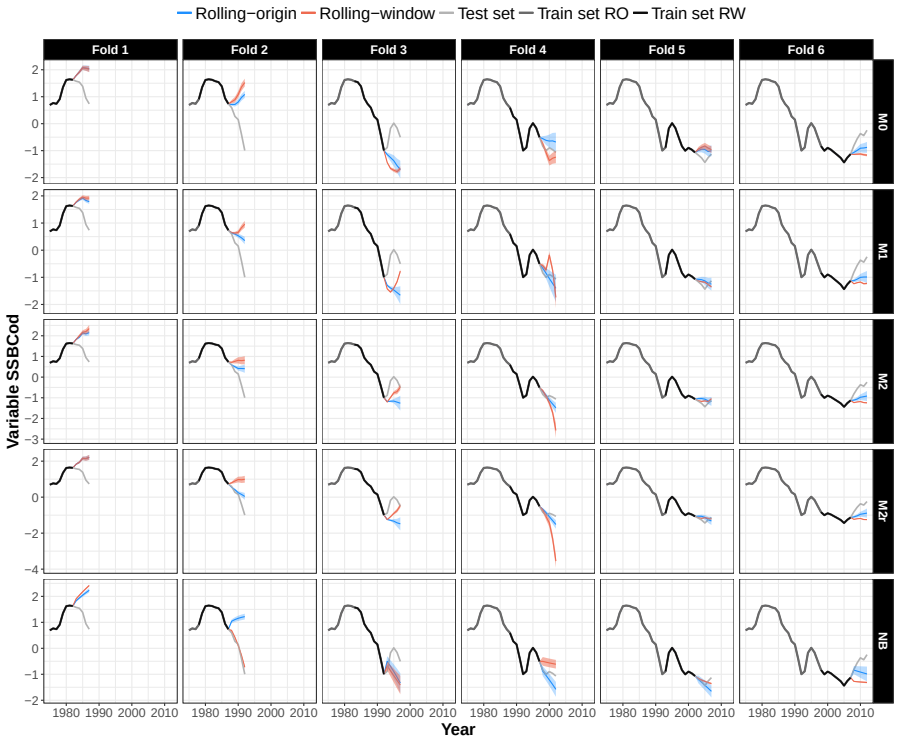
5. Conclusions

The methodology proposed in this paper managed to identify a major regime shift that occurred in the Baltic Sea during the 1980s. The DBN methodology can identify changes in ecosystems where only limited data are available, which is an additional difficulty to cope with. Not surprisingly, the increase in data size improved the models' predictions. Some systems understanding is needed to construct the model structure, but the present experiments show that the exact model setup does not make a critical difference to the results, from the predictive accuracy point of view. However, the use of different HV setups helped reveal changes in the different parts of the ecosystem. Therefore, the presented approach can be a highly useful tool in the study of potentially critical changes in complex ecological interactions.

Appendix A Model predictions

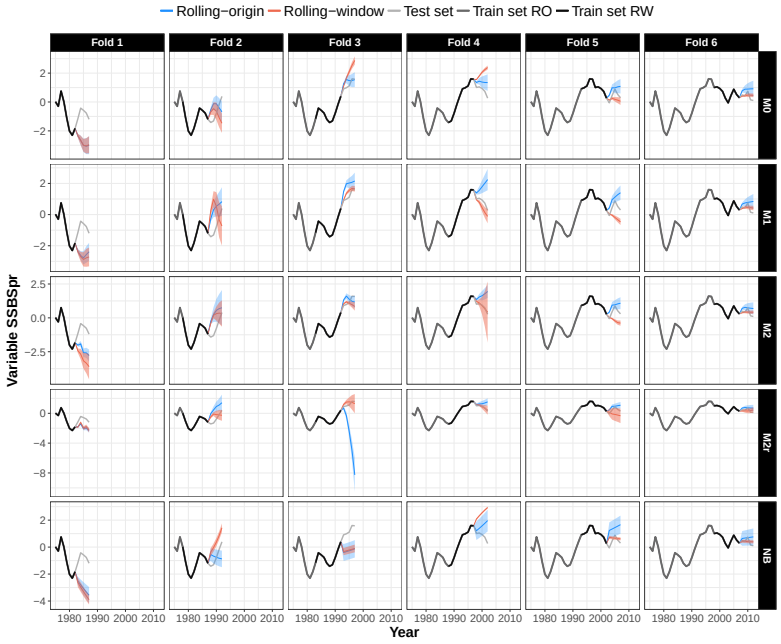
The time series of the observed and predicted values for each variable of interest, model and Fold of both cross-validation techniques are shown in Figures 14a to 14f. From these predictions, we computed the performance of each

model (in terms of log-likelihood) for each variable of interest (Figure 10b). Furthermore, for the rolling-origin case only, we performed some hypothesis testing to compare the predictive performance of the different models (Figures 11 to 13).

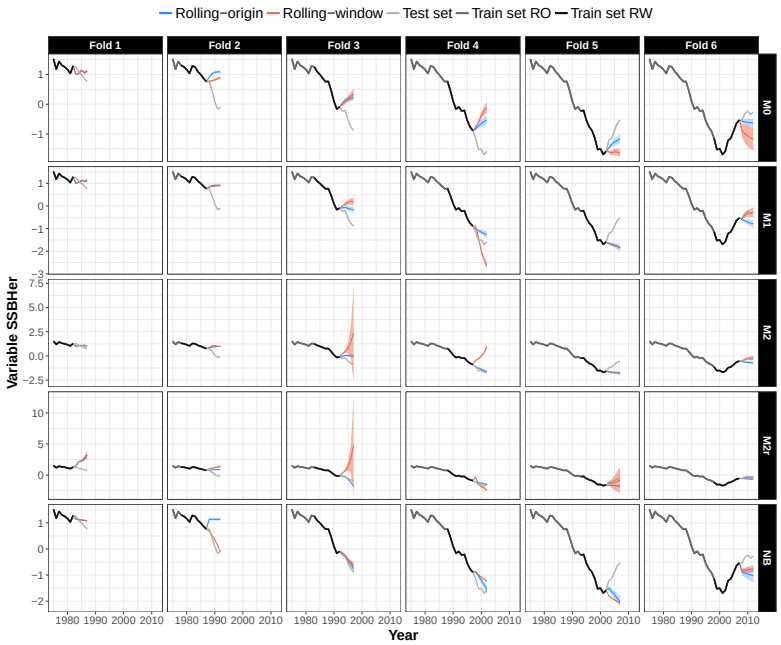


(a) Model predictions for variable SSBCod.

Figure 14: Model predictions (expected value) and observed values used to fit the models, following the rolling-origin (RO) and rolling-window (RW) cross-validation techniques as shown in Figure 7. Black, dark gray and light gray lines correspond to data belonging to the train set of the RW approach, train set of the RO approach and test set, respectively. Red and blue lines correspond to the predictions obtained following the RW and RO techniques, respectively. The shaded bands correspond to the standard deviation.

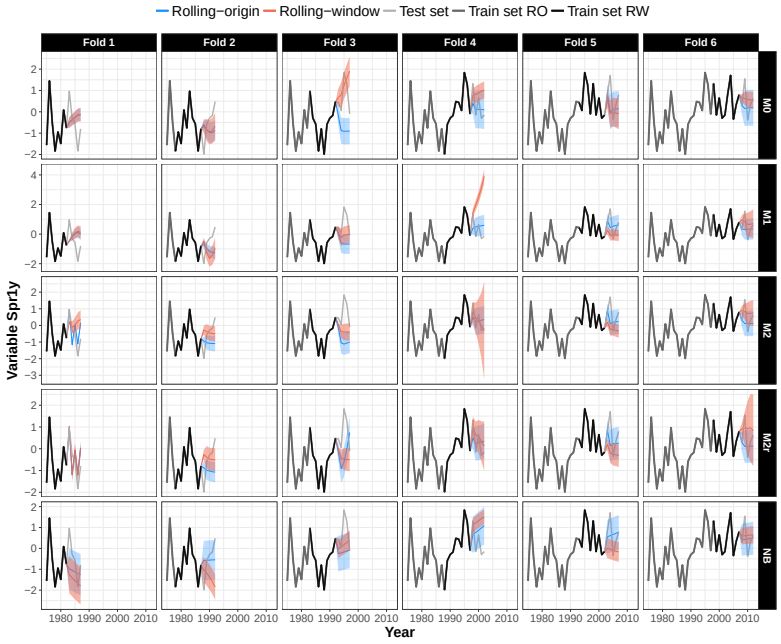


(b) Model predictions for variable SSBSpr.

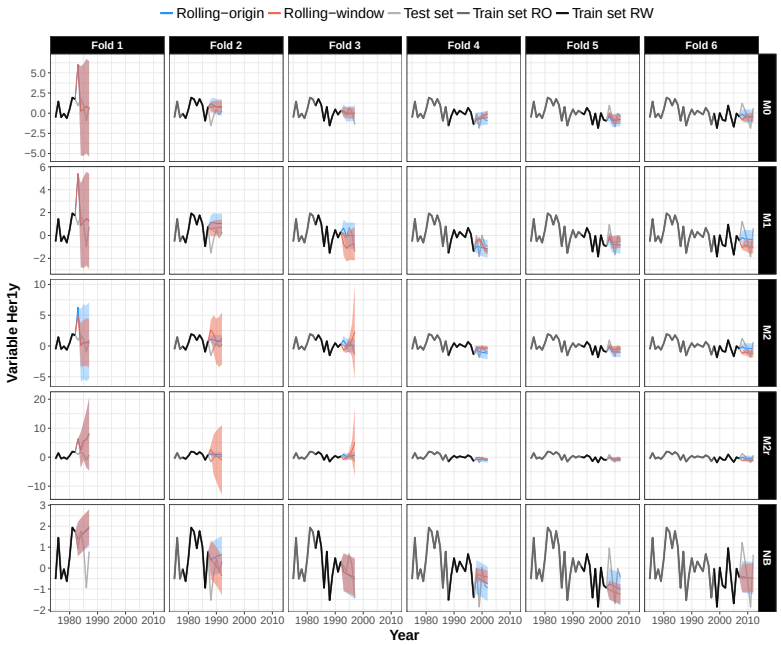


(c) Model predictions for variable SSBHer.

Figure 14: (Cont.) Model predictions and observed values.

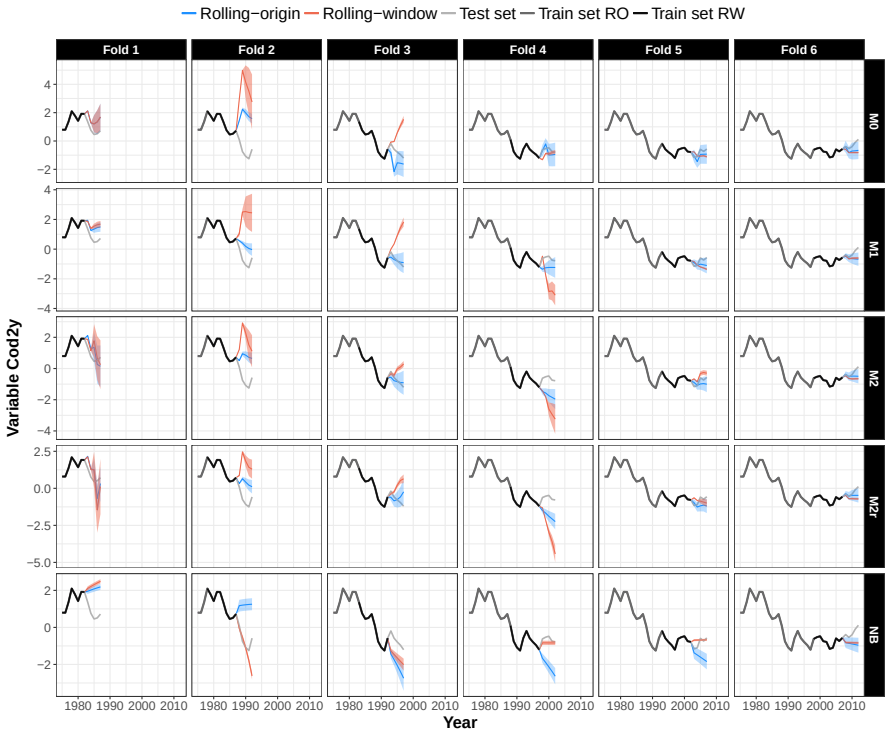


(d) Model predictions for variable Spr1y.



(e) Model predictions for variable Her1y.

Figure 14: (Cont.) Model predictions and observed values.



(f) Model predictions for variable Cod2y.

Figure 14: (Cont.) Model predictions and observed values.

Appendix B Code

The code necessary to run the experiments is available as *Supplementary material*. In particular, *Supplementary material* contains two `.m` files with the `MATLAB` code, one `.tar.gz` file with the `ramidst R` package, and four files containing simulated data (three `.csv` and one `.arff`). As the real dataset used in this work is protected and cannot be published, we provide simulated data to run the code.

Most of the experiments carried out in this work were performed using the `Bayes Net Toolbox` (BNT) package in `MATLAB`. We have created a wrapper function, `DBN_foodWebModel`, to reproduce our experiments. We provide some examples of how to use this function for the simulated dataset. The results of

running this function are stored in `.txt` files. On the one hand, the expected value of the hidden variables can be obtained by running the following piece of code:

```
%% ARGUMENTS OF THE DBN_foodWeb FUNCTION:
% 1. the csv file
% 2. the number of HVs
% 3. the goal: 'learning' or 'inference'
% 4. the core structure: 'normal', 'reverse', 'NB'
% 5. the model version: 'M0', 'M1', 'M2', 'M2r'.
% 6. the number of years we want to predict
% OPTIONAL ARGUMENTS
% CV. Type of cross-validation: 'rol-wind' or 'rol-org'

%% COMPUTE EXPECTED VALUE OF HIDDEN VARIABLES
% Model M1
DBN_foodWeb('simulated.csv', 1, 'learning','normal', 'M1', 0, ...
     'CV','rol-org');

% Model M2
DBN_foodWeb('simulated.csv', 2, 'learning','normal', 'M2', 0, ...
     'CV','rol-org');

% Model M2r
DBN_foodWeb('simulated_rev.csv', 2, 'learning','reverse', 'M2r', 0, ...
     'CV','rol-org');

% Model NB
DBN_foodWeb('simulated_NB.csv', 1, 'learning','NB', 'M1', 0, ...
     'CV','rol-org');
```

On the other hand, the predictions of the target variables can be obtained running the following piece of code:

```
%% RUN PREDICTIONS
```

```
% Rolling-origin
% Model M0
DBN_foodWeb('simulated.csv', 0, 'inference','normal', 'M0', 5, ...
      'CV','rol-org');

% Model M1
DBN_foodWeb('simulated.csv', 1, 'inference','normal', 'M1', 5, ...
      'CV','rol-org');

%Model M2
DBN_foodWeb('simulated.csv', 2, 'inference','normal', 'M2', 5, ...
      'CV','rol-org');

% Model M2r
DBN_foodWeb('simulated_rev.csv', 2, 'inference','reverse', 'M2r', 5, ...
     'CV','rol-org');

% Model NB
DBN_foodWeb('simulated_NB.csv', 1, 'inference','NB', 'M1', 5, ...
      'CV','rol-org');
```

Note that the given examples provide the predictions for the rolling-origin case; to obtain the results for the rolling-window approach, the argument 'CV' must be changed from 'rol-org' to 'rol-wind'. Also, note that it is not necessary to specify which are the target variables since the function is specifically programmed to return the values of these variables. Nevertheless, this function can be easily adapted to other datasets and model structures.

To compare the variational Bayes and the EM algorithms (for the regime shift detection), we used the `ramidst R` package and the `BNT MATLAB` package. We have created another wrapper function, `NBCDD`, for the experiments carried out in `MATLAB`. An example of how to use this function is provided. To illustrate the example, variable `Cod2y` is used as the class variable. Note that we computed the expected value of the hidden variable (H) for 6 different models, i. e., using

6 different class variables.

```
%% ARGUMENTS OF THE NBCDD FUNCTION:
% 1. the class variable: 'SSBCod', 'SSBSpr', 'SSBHer',
%   'Spr1y', 'Her1y', 'Cod2y'
% 2. the csv file
% 3. the first row to read in the dataset
% 4. the last row to read in the dataset

% An example is given for variable 'Cod2y'
NBCDD('Cod2y', 'simulated_NB.csv', 1, 38)
```

An example of how to use the `nb_concept_drift_detector_from_stream` function of `ramidst R` package is also provided.

```
# Import dataset in .arff format
df <- amidst_data_stream("simulated_NB.arff")

# Specify the window size and position of the class variable
# in the dataset (starting from 0).
# For instance, variable Cod2y is the variable in position 15.
w_size = 1L
class_var = 15L
results <- nb_concept_drift_detector_from_stream(df,
       class_index = class_var,
       window_size = w_size)

# 'results' is a list containing the expected value of the hidden
# variable (H), when variable Cod2y is chosen as the class
# variable in the naive Bayes structure.
```

Acknowledgments

This work has been partly supported by the Spanish Ministry of Economy and Competitiveness, through project TIN2016-77902-C3-3-P. In addition, this

work has partially resulted from the BONUS BLUEWEBS project which has received funding from BONUS (Art 185), funded jointly by the EU, the Academy of Finland, Projektträger Jülich (PtJ), Germany, the State Education Development Agency of Latvia, the National Centre for Research and Development, Poland, and the Swedish Research Council Formas. A.D. Maldonado was supported by the Spanish Ministry of Education, Culture and Sport through an FPU research grant, FPU2013/00547, and a research visit grant, EST16/00723.

References

[1] M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems, Nature 413 (6856) (2001) 591.

[2] A. Conversi, V. Dakos, A. Gardmark, S. Ling, C. Folke, P. J. Mumby, C. Greene, M. Edwards, T. Blenckner, M. Casini, et al., A holistic view of marine regime shifts, Philosophical Transactions of the Royal Society B: Biological Sciences 370 (1659) (2014) 20130279–20130279. doi:10.1098/rstb.2013.0279.

[3] K. A. Selkoe, T. Blenckner, M. R. Caldwell, L. B. Crowder, A. L. Erickson, T. E. Essington, J. A. Estes, R. M. Fujita, B. S. Halpern, M. E. Hunsicker, et al., Principles for managing marine ecosystems prone to tipping points, Ecosystem Health and Sustainability 1 (5) (2015) 1–18. doi:10.1890/ehs14-0024.1.

[4] H. Österblom, S. Hansson, U. Larsson, O. Hjerne, F. Wulff, R. Elmgren, C. Folke, S. R. Centre, Human-induced Trophic Cascades and Ecological Regime Shifts in the Baltic Sea, Ecosystems 10(6) (2007) 877 – 889.

[5] J. Alheit, C. Möllmann, J. Dutz, G. Kornilovs, P. Loewe, V. Mohrholz, N. Wasmund, Synchronous ecological regime shifts in the central Baltic and the North Sea in the late 1980s, ICES Journal of Marine Science: Journal du Conseil 62 (2005) 1205–1215.

[6] C. Möllmann, R. Diekmann, B. Müller-Karulis, G. Kornilovs, M. Plikshs, P. Axe, Reorganization of a large marine ecosystem due to atmospheric and anthropogenic pressure: a discontinuous regime shift in the Central Baltic Sea, Global Change Biology 15 (2009) 1377–1393.

[7] J. Yletyinen, O. Bodin, B. Weigel, M. C. Nordström, E. Bonsdorff, T. Blenckner, Regime shifts in marine communities: a complex systems perspective on food web dynamics, Proceedings of the Royal Society B: Biological Sciences 283 (1825) (2016) 20152569. doi:10.1098/rspb.2015.2569.

[8] T. Blenckner, M. Llope, C. Mollmann, R. Voss, M. F. Quaas, M. Casini, M. Lindegren, C. Folke, N. C. Stenseth, Climate and fishing steer ecosystem regeneration to uncertain economic futures, Proceedings of the Royal Society B: Biological Sciences 282 (1803) (2015) 20142809–20142809. doi:10.1098/rspb.2014.2809.

[9] M. Casini, F. Käll, M. Hansson, M. Plikshs, T. Baranova, O. Karlsson, K. Lundström, S. Neuenfeldt, A. Gårdmark, J. Hjelm, et al., Hypoxic areas, density-dependence and food limitation drive the body condition of a heavily exploited marine fish predator, Royal Society Open Science 3 (10) (2016) 160416. doi:10.1098/rsos.160416.

[10] J. Pearl, Probabilistic reasoning in intelligent systems, Morgan-Kaufmann (San Mateo), 1988.

[11] L. Uusitalo, Advantages and challenges of Bayesian networks in environmental modelling, Ecological Modelling 203 (2007) 312–318. doi:10.1016/j.ecolmodel.2006.11.033.

[12] K. B. Korb, A. E. Nicholson, Bayesian artificial intelligence, CRC Press/Chapman & Hall Book, 2011.

[13] K. Murphy, Machine Learning: A Probabilistic Perspective, Adaptive computation and machine learning, MIT Press, 2012.
URL https://books.google.co.in/books?id=NZP6AQAAQBAJ

[14] N. Trifonova, A. Kenny, D. Maxwell, D. Duplisea, J. Fernandes, A. Tucker, Spatio-temporal Bayesian network models with latent variables for revealing trophic dynamics and functional networks in fisheries ecology, Ecological Informatics 30 (2015) 142–158.

[15] L. Uusitalo, M. T. Tomczak, B. Müller-Karulis, I. Putnis, N. Trifonova, A. Tucker, Hidden variables in a Dynamic Bayesian Network identify ecosystem level change, Ecological Informatics 45 (2018) 9–15.

[16] N. Trifonova, D. Maxwell, J. Pinnegar, A. Kenny, A. Tucker, Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model, ICES Journal of Marine Science 74 (2017) 1334–1343.

[17] K. P. Murphy, Dynamic Bayesian Networks: Representation, Inference and Learning, Ph.D. thesis, University of California, Berkeley (2002).

[18] S. L. Lauritzen, The EM algorithm for graphical association models with missing data, Computational Statistics and Data Analysis 19 (1995) 191 – 201.

[19] C. Liu, ML Estimation of the MultivariatetDistribution and the EM Algorithm, Journal of Multivariate Analysis 63 (2) (1997) 296 – 312. doi:https://doi.org/10.1006/jmva.1997.1703.

[20] R. Hilborn, C. Walters (Eds.), Quantitative Fisheries Stock Assessment. Choice, Dynamics and Uncertainty, Springer US, 1992. doi:10.1007/978-1-4615-3598-0.

[21] C. Darby, S. Flatman, G. B. D. of Fisheries Research, Virtual Population Analysis: Version 3.1 (Windows/DOS) User Guide, Information technology series, Great Britain, Ministry of Agriculture, Fisheries and Food, Directorate of Fisheries Research, 1994.
URL https://books.google.es/books?id=HzzKcQAACAAJ

[22] F. W. Köster, B. Huwer, H.-H. Hinrichsen, V. Neumann, A. Makarchouk, M. Eero, B. V. Dewitz, K. Hüssy, J. Tomkiewicz, P. Margonski, A. Temming, J.-P. Hermann, D. Oester-wind, J. Dierking, P. Kotterba, M. Plikshs, Eastern Baltic cod recruitment revisited—dynamics and impacting factors, ICES Journal of Marine Science 74-1 (2016) 3 – 19. doi:10.1093/icesjms/fsw172.

[23] ICES, Report of the Baltic Fisheries Assessment Working Group (WGBFAS), Tech. Rep. ICES CM 2013/ACOM:10 747pp, ICES Headquarters, Copenhagen (10 - 17 April 2013).

[24] J. Pearl, Fusion, propagation and structuring in belief networks, Artificial Intelligence 29 (1986) 241 – 288.

[25] L. Uusitalo, S. Kuikka, A. Romakkaniemi, Estimation of Atlantic salmon smolt carrying capacity of rivers using expert knowledge, Journal of Marine Science 62 (2005) 708–722.

[26] D. Landuyt, S. Broekx, R. D'hondt, G. Engelen, J. Aertsens, P. Goethals, A review of bayesian belief networks in ecosystem service modelling., Environmental Modelling & Software 46 (2013) 1 – 11. doi:10.1016/j.envsoft.2013.03.011.

[27] P. A. Aguilera, A. Fernández, R. Fernández, R. Rumí, A. Salmerón, Bayesian networks in environmental modelling, Environmental Modelling & Software 26 (2011) 1376–1388. doi:10.1016/j.envsoft.2011.06.004.

[28] ICES, Report of the Baltic Fisheries Assessment Working Group (WGBFAS), Tech. Rep. ICES CM 2016/ACOM:11 594pp, ICES Headquarters, Copenhagen (12 - 19 April 2013).

[29] H. Borchani, A. M. Martínez, A. R. Masegosa, H. Langseth, T. D. Nielsen, A. Salmerón, A. Fernández, A. L. Madsen, R. Sáez, Modeling concept drift: A probabilistic graphical model based approach, in: E. Fromont,

T. De Bie, M. van Leeuwen (Eds.), Advances in Intelligent Data Analysis XIV, Springer International Publishing, Cham, 2015, pp. 72–83.

[30] A. R. Masegosa, A. M. Martínez, D. Ramos-López, R. Cabañas, A. Salmerón, H. Langseth, T. D. Nielsen, A. L. Madsen, AMIDST: A Java toolbox for scalable probabilistic machine learning, Knowledge-Based Systems 163 (2019) 595–597.

[31] T. Broderick, N. Boyd, A. Wibisono, A. C. Wilson, M. I. Jordan, Streaming Variational Bayes, in: Advances in Neural Information Processing Systems, 2013, pp. 1727–1735.

[32] P. Reichert, J. Mieleitner, Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters, Water Resources Research 45 (2009) W10402.

[33] C. Bergmeir, J. M. Benítez, On the use of cross-validation for time series predictor evaluation, Information Sciences 191 (2012) 192–213.

[34] T. Hothorn, K. Hornik, M. van de Wiel, A. Zeileis, Implementing a class of permutation tests: The coin package, Journal of Statistical Software 28 (2008) 1–23. doi:10.18637/jss.v028.i08.

[35] M. Hollander, D. A. Wolfe, Nonparametric Statistical Methods, 2nd Edition, Wiley, 1999.

[36] L. Shannon, A. Jarre, F. Schwing, Regime shifts, ecological aspects, in: J. H. Steele (Ed.), Encyclopedia of Ocean Sciences (Second Edition), second edition Edition, Academic Press, Oxford, 2009, pp. 699 – 708. doi:10.1016/B978-012374473-9.00720-7.

[37] G. Beaugrand, The North Sea regime shift: Evidence, causes, mechanisms and consequences, Progress in Oceanography 60 (2) (2004) 245 – 262, regime shifts in the ocean. Reconciling observations and theory. doi:https://doi.org/10.1016/j.pocean.2004.02.018.

[38] M. Weijerman, H. Lindeboom, A. F. Zuur, Regime shifts in marine ecosystems of the North Sea and Wadden Sea, Marine Ecology Progress Series 298 (2005) 21 – 39.

[39] B. deYoung, M. Barange, G. Beaugrand, R. Harris, R. I. Perry, M. Scheffer, F. Werner, Regime shifts in marine ecosystems: detection, prediction and management, Trends in Ecology and Evolution 23 (2008) 402 – 409.

[40] J. W. Dippner, C. Möller, J. Hänninen, Regime shifts in North Sea and Baltic Sea: A comparison, Journal of Marine Systems 105-108 (2012) 115 – 122. doi:https://doi.org/10.1016/j.jmarsys.2012.07.001.

[41] J. Carstensen, J. H. Andersen, B. G. Gustafsson, D. J. Conley, Deoxygenation of the baltic sea during the last century, Proceedings of the National Academy of Sciences 111 (15) (2014) 5628–5633.