SOCIETY FOR SCIENCE AND EDUCATION
UNITED KINGDOM
TMLAI TRANSACTIONS ON
MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE
Volume 5 No. 5
ISSN 2054-7390

Using Artificial Intelligence to Predict Animal Behaviour in Food Webs

[1]Tahir Aduragba, [2]Abdulkadir Ahmed, [3]Bilikisu Ojuolape, [4]Ayodeji Ajani and [5]Yinka Adedoyin
[1234]College of Engineering and Technology, Kwara State University, Ilorin, Kwara State, Nigeria.
[5]Faculty of Communication and Information Science, University of Ilorin, Ilorin, Kwara State, Nigeria.
tahirlanre@hotmail.com; abdulkadiradekunle@live.com; bilkisujimada@yahoo.co.uk;
ajaniaa@gmail.com; yinkaadedoyin22@yahoo.com

ABSTRACT

Overfishing of species in the marine life has caused oceans to become deserts at a fast pace. The population of specific species such as Cod and Haddock has reduced over the years. This has affected countries that hugely depend on them as a source of food. This study used Dynamic Bayesian Network (DBN) to predict animal behaviour in a food web. Two independent biomass surveys from the North Sea were used to learn predictive models and test them on the Northern Gulf Ocean. The resulting predictive model is expected to unveil useful information about what affects the population of fishes in the Northern Gulf Ocean. In addition, the predictive model was used to make predictions into the future about the effects of tampering with the population of specific species of fish in the same region. The focus was on the Cod species in the George's Bank in relationship to species network in their food web. Looking at their biomass states and the effects it has on the hidden dependence when there is a change in their biomass states. Also, the different predictive models were used to evaluate species in the George's Bank based on their performance. The result from the experiment shows that there is a hidden dependence, which is responsible for the collapse of species (Cod); due to the temperature or salinity of the ocean.

Keywords: Fisheries management, data mining, Bayesian network and hidden Markov model

1 Introduction

"About 20% of the world's population derives at least one-fifth of its animal protein intake from fish, and some small island states depend almost exclusively on fish" [1]. Overfishing of species in the marine life has made oceans to become deserts quickly and it has also caused loss of some species as well as entire ecosystem. Due to this, the overall ecological unity of the oceans is under stress and at risk of collapse. The ocean is at risk of losing a valuable food source many depend upon for social, economical or dietary reasons [2]. The over-exploitation and mismanagement of fisheries has already led to some spectacular fisheries collapsing. For example, the Cod fishery of Newfoundland, Canada collapsed in 1992, leading to the loss of some 40,000 jobs in the industry. Further, the Cod stocks in the North Sea and Baltic Sea are now heading the same way and are close to complete collapse [3]. Per Darwin's description of tangled bank, he emphasized that tempering with the population of one species of fish can cause surprising and dramatic changes in the population of others [4]. Overfishing of a specific fish has caused increase in the population of their competitors because there is less consumer-resource competition between the

DOI: 10.14738/tmlai.55.3568
Publication Date: 27th August, 2017
URL: http://dx.doi.org/10.14738/tmlai.55.3568

remaining fishes in the ecosystem. Per the Sequential Population Analysis (SPA), total population of Cod in the Northern Gulf declined from 539 million in 1980 to 31 million in 1994, then slowly increasing to attain 58 million individuals in 2010 [5]. Cods population in this region remains depleted and the reasons behind this remain disputed [6]. Although in the North Sea such collapse in Cod population occurred, the population tends to be recovering over the years [7]. The difference between the recoveries of Cod population in these regions remains unrevealed. Consequently, ban on fishing Cod species from the Northern Gulf Ocean has been introduced but this has not helped in the recovery of these species. However, in this changing world, understanding ecosystem stability and fragility is of growing importance—yet to do so there must be an understanding of the networks that forms the systems [8].

A large amount of data is collected and stored in storage devices every day from business, science and engineering and almost every facet of our daily life. This explosive growth of available data volume is a result of the automation of our society and the progress achieved in developing powerful data collection and storage tools. Therefore, there is desperate need for powerful tools to automatically reveal valuable information from the huge amount of data and to transform such data into organized knowledge. This necessity has led to the birth of data mining [9]. Data mining can be described as the process of discovering patterns in data that are meaningful and can be used advantageously [10]. Although, data collected from ecological sources has been less explorative compared to data collected from other sources mentioned above [11]. In this paper, we apply data mining tools including dynamic Bayesian network and Hidden Markov models to fisheries data to identify species that perform similar functional roles both in the George's Bank and North Sea and these species are used to predict functional collapse in their respective fishing communities.

Bayesian Network (BN) is an exceptional case of a broader class called graphical models, where nodes represent random variables, and the conditional independence assumption is represented by the lack of arcs. BNs do not deal with time, however the Dynamic Bayesian Network (DBN) does by representing how these random variables evolve over time [12]. DBNs are devised to model probability distributions over a sequence of random variables to manage sequenced observations that are propagated by some fundamental hidden states that evolve in time [13]. DBNs consist of two networks. One of them represents the prior probabilities of all variables in the network in the initial time slice $(i.e\ t = 0)$. This is known as the prior network. While the other one represents the probabilities of all the variables in all other time slices $(i.e\ t = 1,2, \dots n)$. This is known as transition network [14].

From the graph below (figure I), X_3 is independent of X_1 given X_2. This explains the first order Markov property that the future is independent of the past given the present. Hence, this graph entails the basic concept of HMM. HMMs, Kalman Filters, Vector Quantization, etc. are all variants of DBNs [15]. This project will implement Hidden Markov Models (HMM) as DBNs. The goal of implementing an HMM as a DBN is to infer the hidden state given the observation sequence, which can be represented as $P(X_t = i|O_{1:t})$. Implementing HMM based models as Dynamic Bayesian Networks (DBN) facilitates compact representation as well as additional flexibility regarding the model structure [14].

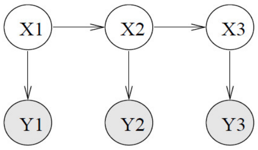


Figure I - Representation of Dynamic Bayesian Network. Observed nodes are shaded, whereas hidden nodes are not shaded

Tucker and Duplisea (2013) used bioinformatics techniques to exploit functional equivalence between different fisheries datasets and used the identified species in conjunction with a dynamic model that uses latent variables to predict functional collapse (and future biomass). The latent variable was used partially to represent something's external to the fish community such as oceanographic conditions. They explored this further by using data of likely factors such as temperature, nutrients and fishing mortality. The result suggests that changes in conditions external to the fish community may be responsible for collapse in GB and ESS.

2 Material and Methods

2.1 Data Source

To conduct the research the data about species biomass and food webs in the studied oceans were used. The information about the species and their food webs in the George's Bank Ocean were provided by the Department for Fisheries and Ocean in Canada while the data sets for the North Sea were provided by the International Council for the Exploitation of Sea. In the George's Bank, the data set contains biomass surveys of thirty-nine species from 1963 - 2008 and for the North Sea biomass surveys of forty-four species from 1967 - 2009 were provided in the data sets. This data is all recorded as typical continuous values. Information about the food web was collected by examination of the species stomach content and recording the fishes that were found in their stomach. This was done for the species in both oceans. Although there are many species, only an overlap of species will be used to carry out the various experiments. These species include: Cod, Herring, Cusk, Mackerel, Spiny dogfish and Red Hake. The choice of these species was based on their trophic relationships to one another. For example, Herring was chosen because it is a prey of Cod. However, the data sets contain missing values, so it was decided to interpolate the data to fit in missing values in the data sets. Although, DBNs are capable of handling missing values from the data sets. Furthermore, the data will also be discretized into simple discrete values.

2.2 Experiments

The experiments undertaken in this paper involve discretization, inference and prediction. To represent the continuous values into a small number of finite values, the data can be discretized. Discretization of real data into a typically small number of finite values is often required by Bayesian Net apps [16]. The biomass surveys are discretized into two qualitative states as low and high. These states are interpreted as the relative weight of species at time slice. For example, if the relative weight of Cod species is 1 at a time, then it is in a low state. However, sometimes they are discretized into three qualitative states low, medium and high to test how they affect the predictions. Number 1 represents low, number 2 medium and number 3 high. Table 1 shows the hypothetical relative weight of species.

Table 1: Hypothetical relative weight for species

States	Low	High
Relative weight	1	2

An important aspect of this paper is trying to predict the behaviour of species over time. To do this, $P(X_{t+h}|y_{1:t})$ needs to be computed where h is how far to predict. Once the future hidden state is known, the information about the future observation state can be retrieved by marginalizing out X_{t+h} as shown in equation 2.1.

$$P(Y_{t+h} = y|y_{1:t}) = \sum_x P(Y_{t+h} = y|X_{t+h} = x)P(X_{t+h} = x|y_{1:t}) \tag{1}$$

The goal of inference can be achieved by using the Junction Tree algorithm. The simplest approach to applying junction tree algorithm to DBNs for inference is by unrolling the DBN for “T” slices and applying the algorithm to each static Bayes net. The disadvantage of this approach is that it takes too long for algorithm to run out of memory if the unrolled DBN becomes huge, i.e. there are many time steps. This is the approach used for the all experiments because it allows relative weight of species over a period to be entered in form of evidence. It was used because the number of time steps involved is not much. However, a better approach to applying Junction Tree algorithm to a DBN is by representing the DBN using only the first two time slices of a process. This approach is based on the first order Markov property that the current time slice is only dependent on the preceding time slice and not any previous time slices. It is much faster than unrolling the DBN into several time slices. The Junction Tree algorithm applied to DBNs for inference can be outlined as follows [17]:

1. Initialization
    A. On initialization, Junction Trees, J_1 and J_t are created
        I. J_1 is the junction tree for the initial time slice and is created from time slice 1 of the 2TBN
        II. J_t is the junction tree for each subsequent time slice and is created from time slice 2 of the 2TBN and the outgoing interface of time slice 1
    B. Time is initialized to 0
2. Queries
    A. Marginals of nodes at the current time slice can be queried
        I. If current time = 0, queries are performed on “_1” nodes in J_1
        II. If current time > 0, queries are performed on “_2” nodes in J_t
3. Evidence Application
    A. Evidence can be applied to any node in the current time slice
        I. If current time = 0, evidence is applied to “_1” nodes in J_1
        II. If current time > 0, evidence is applied to “_2” nodes in J_t
4. Advance
    A. Increment time counter
    B. Use outgoing interface from active time slice to do inference in next time slice
        I. Since the outgoing interface d-separates the past from the future, this ensures that when inference is done in the next time slice it takes everything that has occurred “so far” into account.

The detailed biomass state prediction procedure is as follows:

Step 1: Collect data about biomass surveys of species to use in building the model.

Step 2: Discretize data sets into qualitative states. For example, into two states: low and high.

Step 3: Build the DBN from the collected data and train it.

Step 4: After training the model, the inference model can be used to compute the future states, $P(X_{(t+h)}|y(1:t))$ using equation (2.1)

Step 5: Find the value of $maxP(X_{(t+h)}|y(1:t))$ using equation 2.1 and return as the most probable state.

The first experiment involves simulating population of species over a certain period. Essentially, specific states of biomass of species are entered for certain annual periods and predictions over certain annual periods are made based on the built model as to how these populations behave over certain annual periods. These states are entered in the form of evidence. For example, the biomass state of Cod is set to be low and the remaining species used to build the model are set to be high over a five-year period. So, the behaviour given overtime can monitor this information. Experiment 1 tries to predict if Cod species can recover in the George's Bank and identifies the conditions that are responsible for this behaviour. Throughout this experiment, four species are used to build the model which includes Cod and Haddock. However, the other two species used have a kind of trophic impact be it negative or positive on them.

While conducting the first experiment, predictions were made. To determine how accurate these predictions are, cross validation was used to evaluate the built predictive model. The second experiment involves performing cross-validation to ascertain how accurate the predictions are. Cross-validation is a statistical method of evaluating and comparing learning algorithm by dividing data into two segments: one use to train a model and the other used to validate the model. There are several methods that can be used for cross-validation but in this study, 10-fold cross-validation was used because it is the most common method in data mining and machine learning [18]. While performing cross-validation, certain performance measures can be used. This experiment used sensitivity analysis in the form of confusion matrix to effectively compare the performance of predictive models. A confusion matrix is a visualization tool used to determine the predictive capability of a model. Each row of the matrix represents the actual value while the column of the matrix represents the predicted value. From this, the accuracy of prediction can be determined by estimating the proportion of the total number of correct predictions. The equation is given below:

$$\text{Accuracy, A = correct / (correct + incorrect)} \quad (2)$$

The cross-validation estimation of the overall accuracy can be defined as,

$$CVA = \frac{1}{k}\sum_{i=1}^{k} A_i \quad (3)$$

Where, CVA is the average accuracy at each fold, k is the number of fold and A is the accuracy at each fold.

The third experiment involves testing the learnt model on an unseen data. This is similar to the second experiment however; the difference is that the model is tested on an independent data. In this case, several species from the North Sea were used to build a predictive model and this model was tested on the same species from the George's Bank. The main aim of this experiment is to see if the conditions that helped Cod species recover in the North Sea can be applied to the Cod species in the George's Bank.

The final experiment uses information from the food web to build a model using different species and trying to predict the behaviour over their population over time by simulating their populations over certain periods. Different species were used in building a model based on the information about their trophic impact on each from the food web. For example, predators or prey of Cod are switched to see how Cod relates to them in different conditions.

3 Results and Discussion

In the first experiment, an initial model was built using Cod, Haddock, Herring and Cusk from the George's Bank. These species were chosen based on the information that Cusk is the predator of Cod and Haddock, Cod is the predator of Haddock and Haddock is the predator of Herring. The entire species nodes were represented in the two usual states (low and high). After building the model it was decided to enter evidence that for the first-time slices, Cod is set to be low. This is because the major interest is trying to simulate recovery of Cod. However, the biomass states of other species are randomly entered. In this case, the states of other species are set to be high.

Figure 2: The prediction results of the biomass state of four species

Given these evidences, the model produced the result in figure II. It is expected for the biomass state of Cod to be high because the biomass state of Haddock (prey of Cod) is also high. The biomass state of Cod increases almost immediately. After certain time slices, the biomass state of Cod dropped to a low state. Notes were taken on the hidden states at each of the time slices. From figure III, it was discovered that there was change of state in the hidden dependence immediately the biomass state of Cod dropped. This result suggests that there is a hidden dependence, which is involved with the collapse of Cod biomass state in the George's bank.

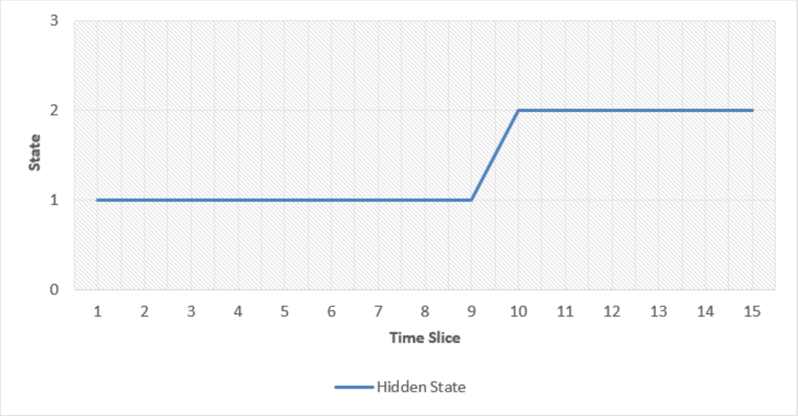


Figure 3: The plot of the hidden dependence

Based on this network, evidences were given that the biomass state of Cod is high, the biomass state of Haddock is high and the biomass state of Cusk (the predator of Cod) is low. It is expected that Cod should remain high due to high biomass of its prey (Haddock) and low biomass of its predator (Cusk). Given this condition, the Cod species remained high for a long number of time slices but later dropped after twelve time slices. (See figure IV)

Figure 4: The prediction results of the biomass state of four species

A new model was constructed but this time the network excluded the predator of Cod. It was replaced with a prey of Cod. The prey of Cod used is Spiny dogfish. Evidences were entered that Cod species are in a high state as well as other species in the network for five time slices. This favours the Cod species, as they remained in a high state most of the time; also, there is no change of state in the hidden dependence. (See figure V)

Figure 5: The prediction results of the biomass state of four species

Following the initial model, where the network comprised of the prey of Cod, a new model was introduced to include a predator of Cod, Red Hake. The evidence entered is that the biomass of Cod is low, the biomass of Haddock is high, as well as Spiny dogfish and the biomass of Red Hake is low over five time slices. Interestingly, Cod rises immediately, however after two time slices it drops into a low state. Moreover, all species in the model, dropped immediately, while the hidden dependencies were changing. This suggests a possible correlation between the biomass of Cod and the hidden dependencies. The result is shown in the figures VI & VII.

Figure 6: The prediction results of the biomass state of four species

Figure 7: The plot of the hidden dependence

In a completely new model, evidence was entered that the biomass state of Cod and Herring were both high, with Haddock and Cusk to being low, for the first five time slices. There was no recovery of Cod in this network, where changes to the hidden states were again found.

The models used for prediction were evaluated based on the performance measures described earlier using equations 2.2 and 2.3. The results were achieved by using ten-fold cross-validation for each model. Each model was built using the same species. For the second experiment, the models were built using Cod, Haddock, Herring and Cusk species from the North Sea and the observation sequence is represented using two states (low and high). These built models are used to predict the biomass states of Cod species over time and the results are validated using the known biomass of Cod species from the George's Bank Ocean. Table 2 shows the complete set of results for each fold with the overall accuracy calculated based on the average results obtained for each fold. The correctly classified states in the datasets are in yellow and the incorrect are in white. On average, it was discovered that the Auto-Regressive HMM model achieved an accuracy of 0.8150 while the DBN achieved an accuracy of 0.7900. In addition, the standard deviation of the accuracy for the Auto-Regressive HMM model is lower than that of the DBNs. This implies that the Auto-Regressive HMM model show more consistency while predicting. Overall, the Auto-Regressive HMM model performed best of both the models used for this experiment. Confusion matrices were also created based on the total number of correct and incorrect classifications (after the 10th fold) to illustrate the performance of these models.

Table 2: Details of accuracy and confusion matrix at each fold of 10-fold Cross-Validation

Fold No.	Auto regressive HMM			Dynamic Bayesian Network		
	Confusion Matrix		Accuracy	Confusion Matrix		Accuracy
1	0	0	1.0000	0	0	1.0000
	0	4		0	4	
2	0	1	0.6000	0	1	0.8000
	1	3		0	4	
3	2	1	0.6000	0	3	0.4000
	1	1		0	2	
4	0	0	1.0000	0	0	1.0000
	0	5		0	5	
5	0	1	0.8000	0	1	0.8000
	0	4		0	4	
6	5	0	1.0000	5	0	1.0000
	0	0		0	0	
7	1	1	0.4000	1	1	0.4000
	2	1		2	1	
8	2	1	0.7500	3	0	0.7500
	0	1		1	0	
9	4	0	1.0000	4	0	1.0000
	0	0		0	0	
10	4	0	1.0000	3	1	0.7500
	0	0		0	0	
Mean			0.8150			0.7900
St Dev.			0.2212			0.2319

While defining DBNs, assumptions are made as to the number of states used to represent the observation sequence of the model. These assumptions are made based on the observed data sets. For this project, all models were built using two states based on the biomass for all species (Low and High). However, it was decided to increase the number of states to three (low, medium and high) and see what effects it has on the accuracies of the models. Same information except for the number of states, which was increased to three, was used to perform 10-fold cross-validation.

As shown in table 3, the accuracies for both the Auto-Regressive HMM model and the DBN dropped to 0.6250 and 0.5870 respectively. Increasing the number of states does not increase performance in terms of accuracy rather it reduces it. Hence, it is concluded that the models should be built using small groups of data sets to achieve better performance.

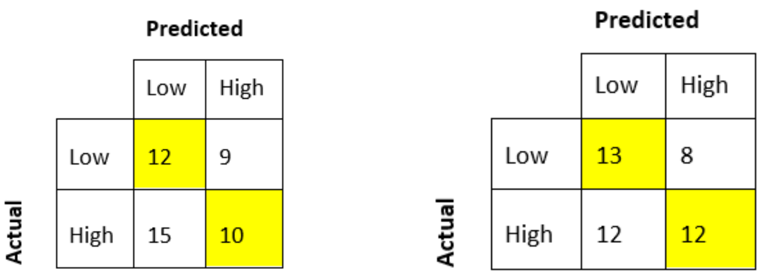
Table 3: Accuracies of three state representations

Auto-Regressive HMM			Dynamic Bayesian Network		
15	0	0	12	3	0
2	3	10	1	6	8
1	5	10	0	7	9
Accuracy = 0.6250			Accuracy = 0.5870		

As earlier described the Cod species suffered overfishing in both the North Sea and the George's bank. However, the Cod species in the North Sea tends to be recovering but it's a different case in the George's bank. One of the experiments conducted in this project is to see how a predictive model learnt from an ocean performs on another independent ocean. The aim of this experiment is to see whether the condition that favored cod species in the North Sea can be applied to Cod species in the George's bank. The network was built using species present in both oceans. This includes Cod Haddock, Spiny Dogfish and Cusk. The two usual biomass states (low and high) were used based on this network.

Table 4 shows the result of the experiment. Analysing the confusion matrix, the accuracy of prediction is 0.50. Also, it tends to predict the high state of biomass for Cod species most of the time. This is expected because this is the situation of Cod in the North Sea. This suggests that the model doesn't fit in and the conditions present in both oceans are different. However, to ascertain this, another network was built using different species. The species used are Haddock, Red Hake, and Spiny Dogfish. The choice of choosing these species was based on the fact that apart from Cod, the other species behave in similar ways in both oceans. From this network, an accuracy of 0.53 was achieved. This showed an improvement from the previous network. However, it doesn't still give an encouraging result to draw a conclusion on this experiment.

Table 4: Confusion Matrices for (a) Auto-regressive HMM and (b) Dynamic Bayesian Network



Actual \ Predicted	Low	High
Low	12	9
High	15	10

(a)

Actual \ Predicted	Low	High
Low	13	8
High	12	12

(b)

The final experiment is similar to the first, where initially there were four species, now it is reduced to two, where the species were switched based upon the food web information. The entire species nodes were represented in the two usual states (low and high). The species Cod and Cusk were used to build the first model. Cusk is a primary predator of Cod. Evidences were entered that the biomass states of Cusk were high for the first time slices, while the biomass state of Cod was set to be low, with a fluctuation between the first 5 times slices. Given the evidence, the biomass state of Cod remained high up to the ninth time slice, then dropped immediately to a low state thereafter, while the biomass state of Cusk remained high, then it dropped after three times slices. However, there was no change in the hidden states until after this occurs. The result of this model is shown in the figures VIII & IX below.

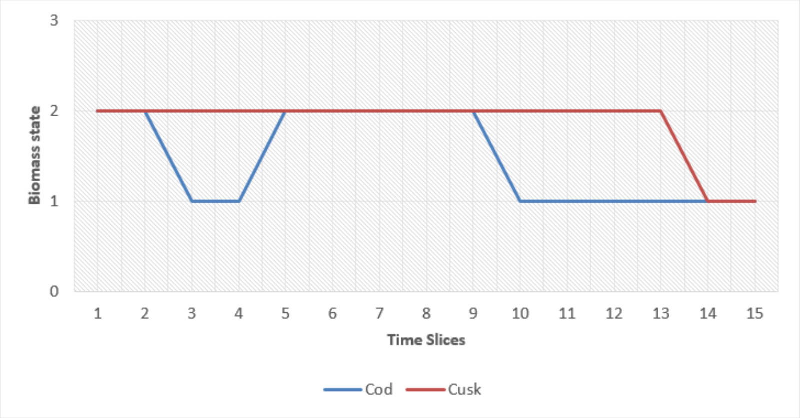


Figure 8: The prediction results of the biomass state of two species

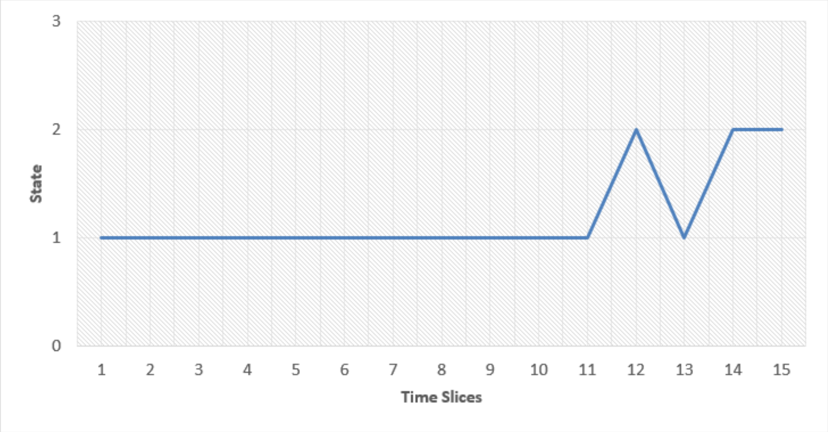


Figure 9: Plot of the hidden dependence

In a second model, using the network of Cod and Haddock (the prey of Cod), the evidence entered that the biomass state of Cod fluctuated between high and low, in respect to the biomass state of Haddock which was relatively high for the first five times slices. Interestingly, after the evidence was entered, the hidden dependences were unchanged for up to ten time slices; even though the biomass state of Cod continued to fluctuate. However, when the state of the hidden dependence changes, the biomass state of Cod, drops and then remains in a low state, which prior to this had consistently been in fluctuation. The result of this model is shown in the figures X & XI below.

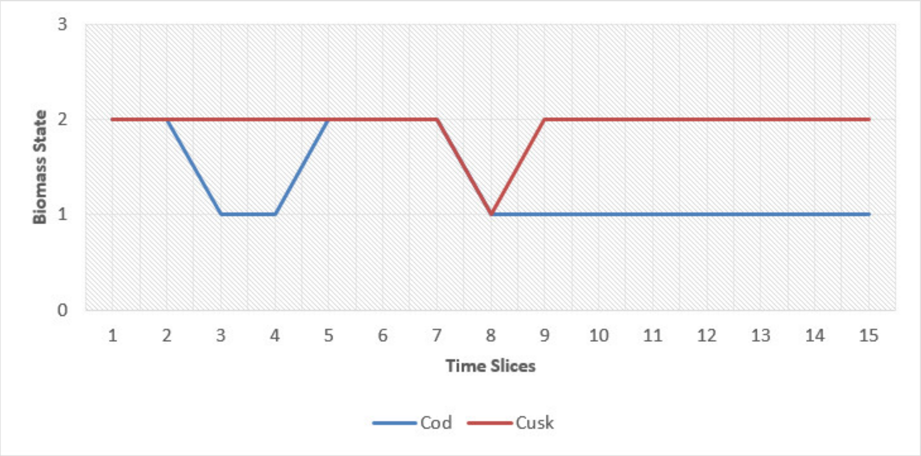


Figure 10: The prediction results of the biomass state of two species

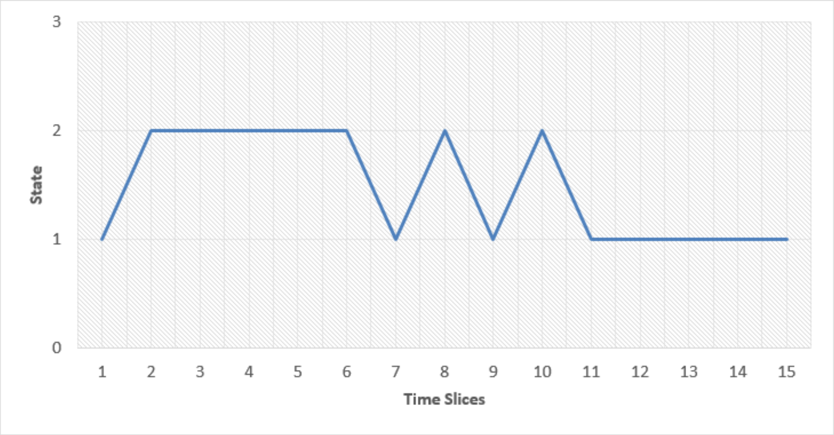


Figure 11: Plot of the hidden dependence

4 Conclusion

This paper looks at the potential correlation between the biomass state of Cod and the changing or unchanging state of the hidden dependencies. Overall, the Cod species did not recover, however it was discovered that there is a change in the state of the hidden dependencies when the biomass state of Cod drops, suggesting that there might be other factors affecting the Cod population. Species such Cod have suffered overfishing, which appear to cause danger to their biomass state, however it is not a conclusive reason for their inability to make a recovery, as they are no longer being fished. New experimentation suggests that the biomass state of the Cod and other network species such as Haddock (the prey of Cod) and Cusk (a predator of Cod), while observing the state of the hidden dependencies, may all be affected by externalities which remain unknown. However, this study investigated this, using data mining techniques to model dependencies of fishes in marine life and it was discovered that there is a hidden dependence, which is responsible for the collapse of species (in this case Cod); possibly the temperature or salinity of the ocean that has affected the species.

REFERENCES

[1] World health Organisation (2011). *Global and regional food consumption patterns and trends.* Available: http://www.who.int/nutrition/topics/3_foodconsumption/en/index.html

[2] Overfishing.org (2011) *Why overfishing is a problem.* Available at: http://overfishing.org/pages/why_is_overfishing_a_problem.php

[3] Greenpeace International (2010). *Overfishing. [Online]* Available at: http://www.greenpeace.org/international/en/campaigns/oceans/overfishing/

[4] Ali, F., Frank, D., Dirk H., Colin M. (2010). Inferring Species Interaction Networks from Species Abundance Data: A comparative evaluation of various statistical and machine learning methods. *Ecological Informatics*, 5 (6), 451-464.

[5] DFO (2010). Assessment of Cod stock in the northern Gulf of St. Lawrence (3Pn, 4RS) in 2009. *DFO Canada Science Advisory Secretariat Science Advisory Report 2010/011*

[6] DFO-SCIENCE (2009) *What's Holding Back the Cod Recovery?* [Online] Available at: http://www.dfo-mpo.gc.ca/science/Publications/article/2006/01-11-2006-eng.htm

[7] The Independent (2007). *North Sea cod population recovering, scientists say*. [Online] Available: http://www.independent.co.uk/environment/nature/north-sea-cod-population-recovering-scientists-say-397350.html

[8] Dunne, J. A., Richard J. W, and Martinez N. D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters* Vol. 5:558–567

[9] Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.

[10] Witten I.H. & Frank E. (2005) *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed. San Francisco: Morgan Kaufmann.

[11] Tucker, A. and Duplisea, D. (2012). Bioinformatics tools in predictive ecology: applications to fisheries. *Philosophical Transactions of the Royal Society, Biological Science* 367(1586), 279-290

[12] Kevin Murphy (2002). Dynamic *Bayesian Network: Representation, Inference, and Learning*. PhD Dissertation, University of California, Berkeley.

[13] Thomas Dean and Keiji Kanazawa (1989). A model for reasoning about persistence and causation. *Artificial Intelligence*, 93(1-2), 1-27.

[14] Fatih C, Ratna B C (2005). Dynamic Bayesian Networks for Machine Diagnostics: Hierarchical Hidden Markov Models vs. Competitive Learning. *Proceedings of International Joint Conference on Neural Networks.* Vol. 3. pp 1752 – 1757.

[15] Smyth, P. (1997). Belief networks, hidden Markov models, and Markov random fields: a unifying view, *Pattern Recognition Letters* 18(11-13) 1261-1268.

[16] Friedman, N. and Goldszmidt, M. (1996) Discretization of continuous attributes while learning Bayesian networks. *Proceedings Of the 13th International Conference on Machine Learning*, 157-165.

[17] Gimpel K (2005). Junction Tree Algorithm for Inference in DBNs. School of Computer Science, CMU, United States.

[18] Payam R, Lei T, Huan L (2008). *Cross-Validation*. Arizona State University, Arizona, USA.