BioOne COMPLETE

Predicting population fluctuations with artificial neural networks

Authors: Lindström, Jan, Kokko, Hanna, Ranta, Esa, and Lindén, Harto

Source: Wildlife Biology, 4(1) : 47-53

Published By: Nordic Board for Wildlife Research

URL: https://doi.org/10.2981/wlb.1998.009

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

RENDEZ-VOUS

Rendez-vous is a forum for promoting discussions among and between scientists and other professionals. New ideas and questions raised may be merely scientific presented in a scientific way, or they may be literary and political contributions to the environmental discussion. Rendez-vous articles typically deal with enthusiastic ideas and expressions of opinion which may lack firm data basis.

Predicting population fluctuations with artificial neural networks

Jan Lindström, Hanna Kokko, Esa Ranta & Harto Lindén

Lindström, J., Kokko, H., Ranta, E. & Lindén, H. 1998: Predicting population fluctuations with artificial neural networks. - Wildl. Biol. 4: 47-53.

Successful predictions of population fluctuations are valuable in game management, as population estimates are instrumental in increasing the time available for management decisions. However, finding a population model which produces predictions accurate enough to be used for management purposes is often precluded due to scarcity and noisiness of population data. Using two long-term population data sets, 1964-1984 data on Finnish grouse (*Tetrao urogallus, T. tetrix* and *Bonasa bonasia*) and 1914-1950 data on coloured fox *Vulpes fulva* from Canada, we demonstrate the use and power of an artificial neural network in predicting population fluctuations. The performance of an artificial neural network model is compared to two benchmark forecasts: time series mean and the previous data value. Unfortunate as it is, in practise management decisions often have to be made with limited data. Therefore, a notable advantage of neural network modelling is the forecast accuracy even in cases when the time series available are short and noisy, and the processes underlying population fluctuations are not fully understood.

Key words: artificial neural network, forecasting, population dynamics, wildlife management

Jan Lindström, Hanna Kokko & Esa Ranta, Department of Ecology and Systematics, Division of Population Biology, P.O. Box 17 (Arkadiankatu 7), FIN-00014 University of Helsinki, Finland*
Harto Lindén, Finnish Game and Fisheries Research Institute, P.O. Box 202, FIN-00151 Helsinki, Finland

**Present address: University of Cambridge, Department of Zoology, Downing Street, Cambridge CB2 3EJ, UK,*
e-mail: J.LINDSTROM@zoo.cam.ac.uk

Received 24 March 1997, accepted 12 November 1997

Associate Editor: Nigel G. Yoccoz

The essence of the theory on applied population dynamics can be rephrased as: given the past, what can be said about the future? Forecasting future events is of special importance in many branches of

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

natural sciences aiming at sustainable use of natural resources. For instance, in Finland, there are long traditions in monitoring game populations (e.g. Lindén, Helle, Helle & Wikman 1996). Consequently, various management organisations are accustomed to utilising long-term monitoring data to derive predictions of the population numbers in the next hunting season to guarantee timely decisions in hunting regulations. In Finland, the grouse populations of the autumn have already been predicted in the spring for several years (Lindén, Laurila & Wikman 1990). Besides benefiting game harvesting, precise enough forecasts also provide time for decision making, e.g. in pest control (Entwistle & Dixon 1986, Paton 1986) and fisheries biology (Crecco, Savoy & Whitworth 1986, Cochrane & Hutchings 1995). Here we restrict the discussion of forecasting to point forecasts one step ahead, and do not attempt to identify and predict long-term trends (see Ives 1995, Link & Sauer 1996). Thus, our approach here applies to tasks such as developing annual game management recommendations more than to conservational issues, where long-term predictions are of interest.

In the simplest cases the future population size, i.e. the forecast, can be obtained by projecting the present state of the population into the future according to a specified model. This can be done, e.g. by multiplying the current population size by its estimated growth rate, or by building a projection matrix with the survival and fecundity values of the individuals in the population (May 1976, Caswell 1989). In reality, however, this approach often poses several difficulties. Selecting an appropriate population model is not necessarily a straightforward task (Pascual, Kareiva & Hilborn 1997) and parameter estimation in practice may also prove to be dreadfully difficult especially with a large set of parameters and scarce data in the presence of noise. Therefore, in order to attain a more general and flexible forecasting system, population managers may make use of forecasting methodology derived from the time series statistics.

Time series models are of special interest in forecasting, since their implementation is clear: the only information they use is the past time series of population size observations (Box, Jenkins & Reinsel 1994). Thus, despite not readily offering biologically relevant interpretations for each parameter value, they have sometimes been used in population studies (e.g. Moran 1953, Mendelssohn 1980, Mendelssohn & Cury 1987, Jeffries, Keller & Hale 1989, Meltzer & Norval 1992). In some cases population data have been used together with external variables (Howard & Dixon 1990, Bautista, Alonso & Alonso 1992).

However, a common problem for traditional Box-Jenkins type of forecasting approaches as well as for econometric modelling (e.g. Pindyck & Rubinfeld 1991) is that ecological time series do not generally match the theoretical preconditions of time series analysis. For instance, they seldom have time invariant mean and variance (second-order stationarity; Diggle 1990, Chatfield 1996), or cannot be conveniently transformed to reach this assumption. The inherent features of population records, viz. short and noisy time series, often further hamper usage of conventional time series analysis. Additionally, though Box-Jenkins forecasting does not build on any specific population growth model, the range of possible population processes to be successfully modelled with time series models is limited (Tong 1990). Clearly, a more flexible tool would be of interest.

Notable progress in time series analysis has been done in the domain of non-linear time series models, which provide a flexible tool for describing complex processes. However, these model types, such as threshold autoregressive models (TAR), autoregressive conditionally heteroscedastic models (ARCH) and its generalised version GARCH, do not generally improve the accuracy of the actual point forecast over the older methods. Instead, their credit in this task is more due to producing better prediction intervals and thus better assessment of the risk involved in the forecast (Davies, Pemberton & Petrucelli 1988, Chatfield 1996).

In this paper, our aim is to focus on the possibilities and capability of one recent alternative tool, the artificial neural network technique (e.g. Carling 1992, Haykin 1994), in forecasting ecological time series. The power of neural networks in predicting time series is based on their flexible function approximation. We shall demonstrate the performance of the neural network approach with long-term grouse data from Finland (Lindén 1989, Lindström, Ranta, Kaitala & Lindèn 1995) and compare it to two simple benchmarks, time series mean and the previous data value (e.g. Casti 1993). Additionally, a different data set on coloured fox *Vulpes fulva* fur returns in Canada (Keith 1963), is used to show that the neural network algorithm like the one developed for grouse can also be generalised to a broader usage.

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Material and methods

Material

The Finnish data are population records of capercaillie *Tetrao urogallus*, black grouse *Tetrao tetrix* and hazel grouse *Bonasa bonasia* dynamics in 11 provinces in Finland (Lindén 1989, Lindström et al. 1995) collected during a 21-year period from 1964 to 1984. The coloured fox data are fur returns before 1951 from the following seven Canadian provinces: British Columbia (32 years), Alberta (32 years), Saskatchewan (37 years), Manitoba (32 years), Ontario (32 years), New Brunswick (27 years) and Nova Scotia (32 years, Keith 1963).

To enhance the forecasting possibilities of the neural network system, the Finnish grouse data were detrended by regressing all the log-transformed time series against time and scoring the residuals. This was done to make the data more stationary, since the grouse fluctuations in Finland show a clear decreasing trend over the time span covered (Lindén 1989, Lindström et al. 1995). The coloured fox data did not show any profound trend, but the average densities and variances differed much. Therefore, they were standardised to zero mean and unit variance (Sokal & Rohlf 1995). Pre-processing the data is often a necessity in ordinary time series modelling (Diggle 1990, Box et al. 1994, Chatfield 1996), and vital in many cases in neural network modelling as well (Smith 1993, Azoff 1994). Worsened performance was also found in our initial explorations while training the neural network without pre-processing the data. Note, however, that despite the pre-processing the forecasted values can easily be transformed back to the original scale.

Artificial neural network

An artificial neural network transforms the values of an input matrix **P** to values of an output matrix **T**. The actual neural network architecture between these two matrices largely depends on the specific problem. In time series forecasting an efficient design is a two-layer network, where the transfer function, **F**, is sigmoidal in the first one and linear in the second one (Smith 1993, Azoff 1994). In the network architecture, each element of the input matrix **P** is connected to each input neuron. Elements of **P** are weighed with layer-specific values, $\mathbf{W}_i$, and shifted by a layer-specific bias value $\mathbf{b}_i$. The results are then transformed by the layer functions of the corresponding level of network layers, in which form they reach the second layer. After the corresponding procedure in the second layer the output vector, **T**, is reached as:

$$\mathbf{T} = \mathbf{F}_2\,(\mathbf{W}_2\mathbf{F}_1(\mathbf{W}_1\mathbf{P} + \mathbf{b}_1) + \mathbf{b}_2) \qquad (1).$$

During the training of the network it simulates the behaviour of the data, and 'learns' by changing the values of the weight matrix, **W**, and comparing the training output matrix $\mathbf{T}_t$ to the target output matrix **T**.

Using as efficient a tool as a two-layer non-linear neural network, finding an excellent fit for the model poses no problem. Too precise learning in the presence of noise, however, prevents the generalisation of the network, i.e. using the trained network for forecasting the future values of new, independent data sets. This phenomenon is called overfitting (e.g. Smith 1993). Consequently, the art of building a successful network for forecasting purposes is to find a balance between finding a fit which is sufficiently accurate and simultaneously retaining the capacity for general solutions.

When the network is trained, increasing the number of learning epochs improves both the fit and forecast of the network to a certain limit, after which the fit gets still closer but the forecast deteriorates. This is where overfitting begins (Smith 1993). Overfitting originates from the network attempting to model noise present in the data, and is thus linked to a general rule of forecasting stating that the model providing the best fit is seldom the best one in forecasting (Pindyck & Rubinfeld 1991). This corresponds to a situation where one can find a perfect match between any data and an adequately high order polynomial only to have no success whatsoever in predicting future values.

One notable advantage in using a neural network in forecasting is that it can form the forecasting rules in an extremely flexible and efficient way. When there are data from the same process, i.e. population data on the same species from different locations as is the case of all species mentioned in this article, the input matrix, **P**, can be constructed so that it includes all the information available of the process. This is achieved by splitting the time series data according to the desired time lag used in the model, and combining all the original time series as follows. Suppose we have two time series $\{X(1),...,X(k)\}$ and $\{Y(1),...,Y(n)\}$, and the desired maximum time lag used for forecasting is d. Then the first column, $\mathbf{p}_X(t)$, of the input matrix, **P**, and its corresponding counterpart, $\mathbf{t}_X(t)$, of target vector, **T**, are:

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

$$\mathbf{P}_x(t) = \begin{bmatrix} X(t-d) \\ X(t-d+1) \\ \vdots \\ X(t) \end{bmatrix} \qquad (2a)$$

$$\mathbf{t}_x(t) = \begin{bmatrix} X(t+1) \end{bmatrix} \qquad (2b).$$

Vectors $\mathbf{p}_Y(t)$ and $\mathbf{t}_Y(t)$ are formed correspondingly. The whole input matrix, **P**, and target vector, **T**, will comprise a combination of inputs $\mathbf{p}_X$ and $\mathbf{p}_y$, and targets $\mathbf{t}_X$ and $\mathbf{t}_y$:

(3a)

$$\mathbf{P}=[\mathbf{p}_x(d+1),\mathbf{p}_x(d+2),...,\mathbf{p}_x(k),\mathbf{p}_Y(d+1),\mathbf{p}_Y(d+2),...,\mathbf{p}_y(n)]$$

(3b)

$$\mathbf{T}=[\mathbf{t}_x(d+1),\mathbf{t}_x(d+2),...,\mathbf{t}_x(k),\mathbf{t}_Y(d+1),\mathbf{t}_Y(d+2),...,\mathbf{t}_Y(n)]$$

This procedure, called cross-sectioning (Chakraborty, Mehrotra, Mohan & Ranka 1992, Smith 1993), was done for all species to achieve a maximal set of learning examples for training the neural network.

Neural network performance

Forecasting performance of the neural network was compared to two traditional benchmarks: 1) predicting the future value to equal the present, and 2) setting the prediction equal to the mean of the observed values so far. These benchmarks are suitable here since realisations of most stochastic processes, such as observed population dynamics in time, tend to aggregate near the mean and they are also temporally correlated (Turchin 1990, Royama 1992). Thus, these benchmarks show the performance of simple and extremely conservative forecasting methods.

The comparison was done so that for the grouse data, we initially optimised the time lag and learning rate of the neural network using cross-sectioned data. The optimisation was performed by seeking the time lag and learning rate which minimise the forecast error (sum of squared errors, SSE) for 25 randomly chosen data points for each parameter combination and species. Then we tested the performance of this optimised network against the benchmarks. The forecasting was done both in the optimisation and test phases by randomly removing one column from the **P** matrix (equation 3a) and its corresponding target value from the **T** vector (equation 2b), and letting the network start learning from the beginning with the earlier defined time lag, learning rate and newly constructed **P** and **T**. The obtained network parameters were then used to forecast the excluded target value by entering the corresponding input vector, **p**, without the target vector, **t**, into the network. This was repeated 100 times for each species, capercaillie, black grouse and hazel grouse. In each repeat, the SSE of the forecast error was scored, as well as the SSE of the forecast error using the benchmark forecasts.

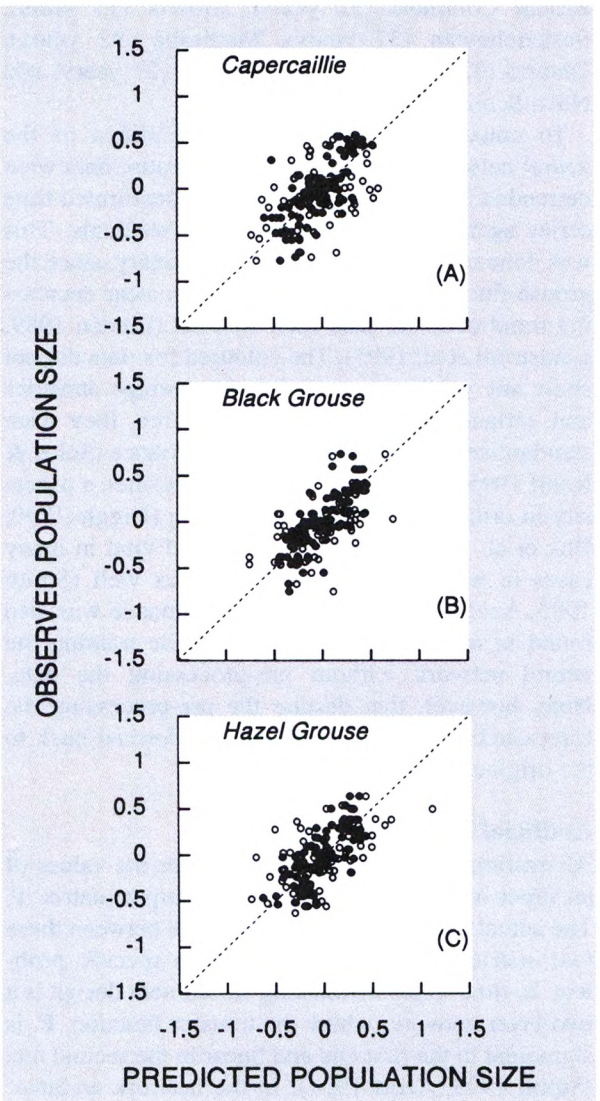


Figure 1. A sample of 100 predicted and observed population estimates for capercaillie (A), black grouse (B) and hazel grouse (C). Black dots refer to forecasts made by generalised neural network and open circles show the previous data values, which were used as a benchmark for the performance of neural network predictions. These provided slightly better forecasts than time series mean (see text).

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Results

Optimisation of the artificial neural network resulted in a learning rate of 0.007 for both data sets, and the optimal time lag for forecasting was 10 for the grouse species and 7 for the coloured fox. With these choices, the forecast error (SSE) was found to be smallest for 50 test runs. After this optimisation was completed, the network forecasting performance was tested against the two benchmarks.

Both benchmark forecasts and neural network forecast matched the observed grouse population size to a high degree (Fig. 1) as would be expected of a good forecasting method. However, the generalised neural network outperformed the benchmarks in every species (see below, binomial test: event E_i = 'forecasting method i has lower SSE than the corresponding benchmark method', H_0: $P(E_i) = 0.5$, H_1: $P(E_i) > 0.5$).

The benchmark using the previous data value turned out to yield slightly smaller forecast errors than the time series mean (Table 1).

Also, optimising and training in a similar manner

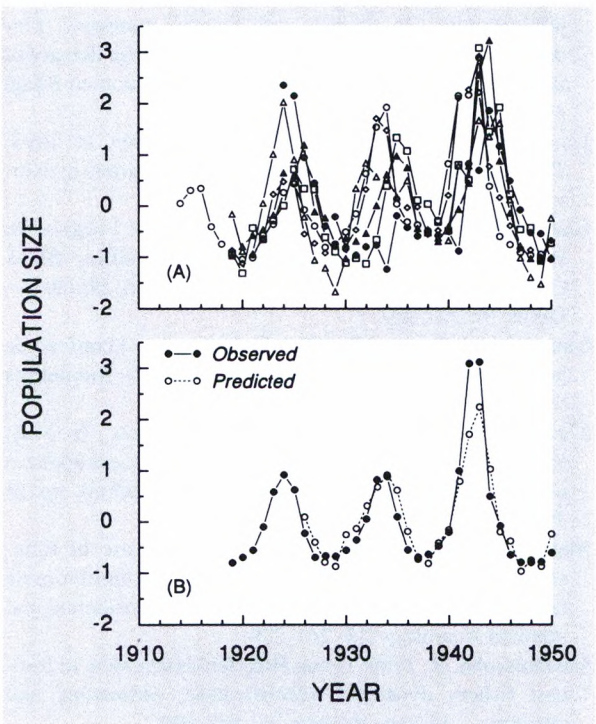


Figure 2. An example of neural network generalisation. (A) shows the coloured fox (Keith 1963) fur returns for six Canadian provinces. (B) shows the corresponding figures for the seventh province, not used in the training of the neural network, and the prediction obtained by the neural network model.

Table 1. P-values of binomial tests for comparing SSE of neural network performance to two benchmark forecasts (see text for details).

	Time series mean	Previous value
Capercaillie	0.0033	0.0004
Black grouse	0.0004	0.0018
Hazel grouse	0.0009	0.0284

the same network structure with data from the coloured fox, we were able to produce a rather well generalisable neural network resulting in a successful forecast performance for the coloured fox data as well (Fig. 2).

Discussion

Artificial neural networks provide a very powerful tool in cases where forecasts are required but the underlying process is unknown or only partially understood. An obvious drawback is the complexity of its usage. The most demanding phase in using artificial neural networks is the optimisation of the network training parameters: learning rate and the number of learning epochs. This is due to the non-linearity of the network, which on the other hand renders it possible to find a signal even when hidden in a rather noisy data set. This is why a neural network forecasting is always a data-driven process without much possibility of giving general rules for optimal parameter choice (Smith 1993, Azoff 1994).

As to practical applications, the neural network methodology - like any other forecasting method - works better the more data that are available. Especially important are extreme values: if they are missing in the training data, they cannot be reliably forecasted either. The coloured fox results also show that although there is a considerable amount of synchrony in the Finnish grouse data (Ranta, Lindström & Lindén 1995, Lindström, Ranta & Lindén 1996), this is not necessary in order to obtain good results with cross-sectioning of the data (see Fig. 2).

Neural network does not assume anything about the underlying process - alas - neither does it provide information about it. Herein lies its strength as well as its limitations: forecasts will not be destroyed because of an unfortunate choice of a population model, but validation of ecological hypotheses about the data underlying population processes are also excluded. The sole aim of the method described here has been to forecast future population numbers.

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

However, this is a task of utmost importance in making management decisions based on limited knowledge. In those cases, an artificial neural network provides a potential tool. We would like to emphasise, however, that in any particular situation the choice of the 'best' method for forecasting depends on the data and the objectives of the study. In most cases it may be advisable to have a range of models at hand and choose an appropriate one among them.

Acknowledgements - this study was funded by the Academy of Finland and Oskar Öflund Foundation. We thank Chris Chatfield for statistical advice, Peter Hudson for comments and Nigel Yoccoz and two anonymous referees for their constructive criticism.

References

Azoff, E.M. 1994: Neural network time series forecasting of financial markets. - John Wiley & Sons, Chichester, 196 pp.

Bautista, L.M., Alonso, J.C. & Alonso, J.A. 1992: A 20-year study of wintering common crane fluctuations using time series analysis. - Journal of Wildlife Management 56: 563-572.

Box, G.E.P., Jenkins, G.M. & Reinsel G.C. 1994: Time series analysis: forecasting and control. - Prentice Hall, Englewood Cliffs, 598 pp.

Carling, A. 1992: Introducing neural networks. - Sigma Press, Wilmslow, 338 pp.

Casti, J.L. 1993: Searching for certainty: what science can know about the future. - Abacus, London, 320 pp.

Caswell, H. 1989: Matrix population models. - Sinauer Associates, Sunderland, 328 pp.

Chakraborty, K., Mehrotra, K., Mohan, C.K. & Ranka, S. 1992: Forecasting the behavior of multivariate time series using neural networks. - Neural Networks 5: 961-970.

Chatfield, C. 1996: The analysis of time series. An introduction. - Chapman & Hall, London, 283 pp.

Cochrane, K.L. & Hutchings, L. 1995: A structured approach to using biological and environmental parameters to forecast anchovy recruitment. - Fisheries Oceanography 4: 102-127.

Crecco, V., Savoy, T. & Whitworth, W. 1986: Effects of density-dependent and climatic factors on American shad, Alosa sapidissima, recruitment: A predictive approach. - Canadian Journal of Fisheries and Aquatic Sciences 43: 457-463.

Davies, N., Pemberton, J. & Petrucelli, J.D. 1988: An automatic procedure for identification, estimation and forecasting univariate self-exciting threshold autoregressive models. - The Statistician 37: 199-204.

Diggle, P. 1990: Time series. A biostatistical introduction. - Clarendon Press, Oxford, 257 pp.

Entwistle, J.C. & Dixon, A.F.G. 1986: Short-term forecasting of peak population density of the grain aphid (Sitobion avenae) on wheat. - Annals of Applied Biology 109: 215-222.

Haykin, S. 1994: Neural networks. - Macmillan College Publishing Company, New York, 696 pp.

Howard, M.T. & Dixon, A.F.G. 1990: Forecasting of peak population density of the rose grain aphid Metopolophium dirhodum on wheat. - Annals of Applied Biology 117: 9-19.

Ives, A.R. 1995: Predicting the response of populations to environmental change. - Ecology 76: 926-941.

Jeffries, P., Keller, A. & Hale, S. 1989. Predicting winter flounder (Pseudopleuronectes americanus) catches by time series analysis. - Canadian Journal of Fisheries and Aquatic Sciences 46: 650-659.

Keith, L.B. 1963: Wildlife's ten-year cycle. - The University of Wisconsin Press, Madison, 201 pp.

Lindén, H. 1989: Characteristics of tetraonid cycles in Finland. - Finnish Game Research 46: 34-42.

Lindén, H., Helle, E., Helle, P. & Wikman, M. 1996: Wildlife triangle scheme in Finland: methods and aims for monitoring wildlife populations. - Finnish Game Research 49: 4-11.

Lindén, H., Laurila, A. & Wikman, M. 1990: Metson ja teeren kannanvaihtelujen ennustettavuus - varovaista jälkiviisautta. (In Finnish with English summary: Predictability of the capercaillie and black grouse density of the next year - wisdom after the event) - Suomen Riista 36: 82-88.

Lindström, J., Ranta, E., Kaitala, V. & Lindén, H. 1995: The clockwork of Finnish tetraonid population dynamics. - Oikos 74: 185-194.

Lindström, J., Ranta, E. & Lindén, H. 1996: Large-scale synchrony in the dynamics of Capercaillie, Black Grouse and Hazel Grouse populations in Finland. - Oikos 76: 221-227.

Link, W.A. & Sauer, J.R. 1996: Estimation and confidence interval for empirical mixing distribution. - Biometrics 51: 810-821.

May, R.M. 1976: Models for single populations. - In: May, R.M. (Ed.); Theoretical ecology. Principles and applications. Blackwell Scientific Publications, Oxford, pp. 5-29.

Meltzer, M.I. & Norval, R.A.I. 1992: The use of time-series analysis to forecast bont tick (Amblyomma hebraeum) infestations in Zimbabve. - Experimental and Applied Acarology 13: 261-279.

Mendelssohn, R. 1980: Using Box-Jenkins models to forecast fishery dynamics: identification, estimation, and checking. - Fishery Bulletin 78: 887-896.

Mendelssohn, R. & Cury, P. 1987: Fluctuations of a fortnightly abundance index of the Ivoirian Coastal pelagic species and associated environmental conditions. - Canadian Journal of Fisheries and Aquatic Sciences 44: 408-421.

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Moran, P.A.P. 1953: The statistical analysis of the Canadian lynx cycle. I. Structure and prediction. - Australian Journal of Zoology 1: 163-173.

Pascual, M.A., Kareiva, P. & Hilborn, R. 1997: The influence of model structure on conclusions about the viability and harvesting of Serengeti wildebeest. - Conservation Biology 11: 966-976.

Paton, G. 1986: A matrix modelling approach to population growth systems involving multiple time delays. - Ecological Modelling 34: 197-216.

Pindyck, R. & Rubinfeld, D. 1991: Econometric models and economic forecasts. - McGraw-Hill, New York, 596 pp.

Ranta, E., Lindström, J. & Lindén, H. 1995: Synchrony in tetraonid population dynamics. - Journal of Animal Ecology 64: 767-776.

Royama, T. 1992: Analytical population dynamics. - Chapman & Hall, London, 371 pp.

Smith, M. 1993: Neural networks for statistical modeling. - Van Nostrand Reinhold, USA, 235 pp.

Sokal, R.R. & Rohlf, F.J. 1995: Biometry. - W.H. Freeman and company, New York, 887 pp.

Tong, H. 1990: Non-linear time series. A dynamical system approach. - Clarendon Press, Oxford, 564 pp.

Turchin, P. 1990: Rarity of density dependence or population regulation with lags? - Nature 344: 660-663.

Downloaded From: https://complete.bioone.org/journals/Wildlife-Biology on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use