BIOINFORMATICS AND STATISTICS IN FISHERIES RESEARCH

VOL-III

EDITORS

AJIT KUMAR ROY

NIRANJAN SARANGI

NEURAL NETWORK AND ITS APPLICATION IN AQUACULTURE

M. Balakrishnan, K. Meena, Satyanarayan Sethi & Aditya N. Sarangi

Central Agricultural Research Institute, Port Blair- 744101, Andman & Nicobar Islands

Introduction: In the past decades the researchers and managers have used empirical statistical models or complicated mathematical models predict the consequences of management regimes or actions, and to assist in decision-making. These models are expressed as mathematical equations. However, some decision-making processes contain qualitative components that do not lend themselves to being integrated into mathematical equations. As [1] point out, decision making in natural resources often leads to complexities beyond the reach of empirical statistical techniques, and requires approaches that are sometimes more heuristic than algorithmic. In many cases, statistical models cannot give accurate results and to solve the more complex problems in fish yield prediction.

The application of artificial intelligence (AI) in agricultural and natural resources management started with the development of expert systems for problem solving and decision-making [2]. ANN is an attempt to mimic the computational process of the brains that consists of a number of interconnected processing element (or) "neurons" [3]. ¹Usually at one end of the network data fed in and at the other end, the network produces outputs. In between there are layers of neurons that may or may not are depending on their current state, their inputs their earlier layers and their own properties [4]. An interest in the use of artificial neural networks (ANN), are distributed and parallel computational algorithm, which have been efficiently used, in various practical problems. ANN has also begun to emerge as an alternative approach for modeling nonlinear and complex phenomena in Fisheries science. However artificial neural network (ANN) is an aspect of computational intelligence research that has been recently shown efficiency in describing non-linear biological systems [4].

Important Features of ANN: There are different neural network models developed since the first prototype neural network was proposed in 1943 [5]. Neuron science was influential in development of neural networks, but physiologists and engineers also contributed of neural network systems- [6] created a great deal of interest when he designed and developed the perceptron. One of the most widely used neural networks in yield prediction is the back-propagation feed-forward network (also referred as the Multi-Layer Perceptions (MLP) networks are widely used due to their simplicity and excellent performance [7]. As one of many possible examples, this section provides a brief introduction to the major features of the BPN & MLP.

Neural network predictor : All the neural network architectures used in this paper are feed-forward networks that learn by adjusting their weights according to the back-propagation algorithm as described in [8]. The number of neurons in the input layer is set equal to the number of bits in the binary inputs strings for the data set used. Each neuron of the input layer is set to either 0 or 1 depending on the value in the input string that it corresponds to. The values then propagate forward through the network until a pattern appears on the neurons of the output layer. The differences between them are propagated back through the network and the connection weights between the layers are adjusted. The threshold function used in the neurons is the unipolar sigmoid function given by,

$$f(x) = 1/1 + \exp(-x)$$

The training patterns are input to the network in random order, to prevent the network from being skewed towards the last pattern presented. After each pattern has been input to the network, the difference between the network and desired output is used to update the weights on the network connections. The change in weights is given by

$$\Delta W_{ij}^{k+1} = c(d_i - o_i) o_i(1 - o_i) + m\Delta w_{ij}^k$$

Where W_{ij} is the connection weight between neuron j on one layer with neuron I on the previous layer and d_i and o_i are the desired and actual network output, respectively, for bit i. The constants c and m denote the learning rate and momentum of the network and they govern how fast the weights will change at each step and how much that change will depend on the previous weight update. Training continues until the sum of squares (S.S) error is less than some critical value (ecrit). The value of Sum of square is calculated from

$$S.S = {}_p\sum_i (d_i - o_i)^2$$

Where P is the total number of patterns presented. In the work presented here, 'ecrit' was chosen to have a value of 1.0. [8]

After training has been completed the learning is switched off and previously unseen test input patterns are presented to the network. These input patterns are propagated forwards to obtain an output. This output is the prediction of the following 24h minimum temperature category and, in the real application, would be used as an indicator as to whether frost prevention measures should be taken. In this investigation, the network output is compared to the output string associated with the input, to measure the network's performance.

Learning and Training: Learning and training are fundamental to nearly all neural networks. Training is the procedure by which the network learns; learning is the end result of that procedure. Learning consists of making systematic changes to the weights to improve the network's response performance to acceptable levels. The networks learn by adjusting the

weights connecting the layers. The network starts by finding linear relationships between the inputs and the output. Weight values are assigned to the links between the input and output neurons. Once those relationships are found, neurons are added to the hidden layer so that nonlinear relationships can be found. The aim of training is to find a set of weights that will minimize error. During training, the output predicted by the network (Y (t)) is compared with the actual (desired) output (K (t)), and the mean squared error (MSE) between the two is calculated. The error function at time, $\Delta(t)$, is given by:

$$\Delta(t) = \frac{1}{2} ? (Y (t)-K (t))$$

The learning algorithm modified the weights associated with each PE such that the system minimizes the error between the target output and the network's actual output [9]. The back-propagation algorithm is the most computationally straightforward algorithm for training the MLP. More detailed explanation is available in most neural network text books [10]

Applications in Aquaculture

Finfish and Shell fish disease forecasting using ANN : A database was probed with artificial neural network (ANN) and multivariate logistic regression (MLR) models to investigate the efficacy of predicting PCR-identified human Adenovirus (ADV), Norwalk-like virus (NLV), and enter virus (EV) presence or absence in shellfish harvested from diverse countries in Europe (Spain, Sweden, Greece, and the United Kingdom). The relative importance of numerical and heuristic input variables to the ANN model for each country and for the combined data was analyzed with a newly defined relative strength effect, which illuminated the importance of bacteriophages as potential viral indicators [11]. The results of this analysis showed that ANN models predicted all types of viral presence and absence in shellfish with better precision than MLR models for a multicountry database. For overall presence/absence classification accuracy, ANN modeling had a performance rate of 95.9%, 98.9%, and 95.7% versus 60.5%, 75.0%, and 64.6% for the MLR for ADV, NLV, and EV, respectively. The selectivity (prediction of viral negatives) was greater than the sensitivity (prediction of viral positives) for both models and with all virus types, with the ANN model performing with greater sensitivity than the MLR. ANN models were able to illuminate site-specific relationships between microbial indicators chosen as model inputs and human virus presence. A validation study on ADV demonstrated that the MLR and ANN models differed in sensitivity and selectivity, with the ANN model correctly identifying ADV presence with greater precision.

Quality control of finfish & Shellfishes: When fish die, bacteria or the enzymes they produce invade the flesh of fish. This process produces toxic compounds in the fish and the fish becomes spoiled. Fourier Transform Infrared spectroscopy (FT-IR) allows chemical based non-destructive discrimination of intact microbial cells, and produces complex

biochemical fingerprints which are reproducible and distinct for different accumulation of byproducts of spoilage. Data obtained through FT-IR spectroscopic analysis can be used to train an artificial neural network (ANN) for the development of an ANN based FT-IR Screening System for fish-quality inspection [12].

Eco-path modeling of river systems: Human activities have severely deteriorated the river systems, and many functions such as drinking water supply, ûshing, etc. are threatened. Because the restoration of these river systems entails drastic social (e.g. change in habits with regard to water use and discharge, urban planning) and economical (e.g. investment in nature restoration, wastewater treatment system installation) consequences, the decisions should be taken with enough forethought. Ecosystem models could therefore act as interesting tools to support decision-making in river restoration management. In particular models that can predict the habitat requirements of organisms are of considerable importance to ensure that the planned actions have the desired effects on the aquatic ecosystems [13]. It was shown that machine learning techniques such as Artificial Neural Networks (ANNs) basically mimic aspects of biological information processing for data modeling and could be useful in ecology.

Predicts Antibiotic residues in the Ecosystem: Fish antibiotics which are an important group of pharmaceuticals used in fish farming to treat infections and, until recently, most of them have been exposed to the environment with very little attention. Information about the environmental behaviour and the description of the environmental fate of medical substances are difficult or expensive to obtain [14].

Fish Yield Map Modeling: Fish yield simulation is no longer restricted to predicting performance at a point. Models for pond dynamics and models for machinery, such as irrigation systems and harvesting equipment, allow us to predict the variability in fish production performance in the pond. Yield maps are now one of the most common sources of map data available to farmers for their fields. GPS units and ANN are a relatively low-cost addition to a new combine, and examples exist where farmers have used the yield/production maps to discover expensive production problems.

Satellite Image Processing: During 90's there have been considerable increases in both the availability of large remotely sensed data and the use of neural networks. This provides the opportunity to test the ability of neural networks, in particular the feed-forward back-propagation multi-layer perceptron, and compare the performance of particular neural networks with other traditional methods of satellite image processing The ANN approach has also been used to retrieve the correlation lengths and variance from rough surface to reconstruct the snow parameters to estimate the total culture area and also to estimate the production and productivity of water bodies.

Spatial modeling with GIS : Recent research has shown that coupling ANN with GIS has significantly improved the modeling capabilities of GIS for spatial decision making In the pioneering work by the successfully strengthened the spatial data modeling capabilities of GIS for aquaculture land suitability analysis by integrating a neural network into a GIS environment.

Utilities, Value and Future perspective : In general, ANN technology mimics the brain's own reasoning process. The advantage of ANN for yield prediction has been motivated by the realization that the human brain is very efficient at processing large quantities of data from a variety of different sources, and making decisions in a complex environment. As humans apply knowledge gained from past experience to new problems or situations, a neural network takes previously solved examples to build a system of "neurons" that makes new decision, classifications, and predictions accurately and rapidly. In particular, the ANN approach shows advantages over statistical modeling approaches traditionally used to study natural systems [1], [7]; Artificial Neural Network's

➢ They deal with the non-linearity's in the world in which we live

➢ They handle noisy or missing data.

➢ They create their own relationship amongst information and no equation. They do not require a *priori* knowledge of the underlying process or assumptions of the structure of the target function.

➢ They can work faster than other when the problem is extremely complex or large numbers of variables or parameters. The neural network can develop its own weighting scheme based on relationships between the variables, thus reducing the requirement that user provide all known information about a problem.

➢ They provide general solutions with good predictive more accurate than other statistical techniques, particularly when the problem or task addressed is either poorly defined or misunderstood, and observations of the process may be difficult or impossible to perform using incomplete data;

➢ Once trained, the nets can be used to analyze new conditions and provide suggested solutions. The ability of the net to learn complex relationships and the capability of including both qualitative as well as quantitative data makes the neural net approach a very flexible and powerful tool.

However, it is equally important to understand the basic problems of neural networks. There are three issues to be aware of, particular for those who are new to the use of neural networks. ANN is usually treated as a "black-box", with which the weights are uninterruptible

due to presence of hidden layers and the nonlinearity of the activation function. Neural nets are not self-explanatory; there are no standard tests can measure the degree of variability in the outputs explained by certain inputs or the significance level of the predictions. This is one of reasons that forest managers are less likely to use ANN when a more familiar and better-understood procedure such as a regression analysis is available .On other hand, the research on ANN applications has been limited compared to other AI techniques (e.g., Knowledge based System) in natural resource management [2]. There is an urgent need to widely recognize the potential uses of ANN as an alternative tool in fisheries field.

Conclusion: In this research an artificial neural network was employed to model the relation ship between yield and the factors influencing yield. ANN techniques have been proven as a useful tool for predicting, classifying, and approximating functions in different fields, and are finding a wide range of applications in yield prediction and stock prediction. The utility of the ANN approach are apparent in applications (i) It is entirely numerical method as opposed to the implicit symbolism in the system of rules where the problem addressed may be either poorly understood, or observations of the process may be difficult to carry out using noisy or incomplete data; and (ii) It does not demands a previous knowledge of the system that we wish to predict and it is characterized when the problem is extremely complex, particularly when dealing with nonlinear. Systems, where traditional statistical techniques or mathematical models cannot to be formulated.

However, the well-known disadvantage of the neural network methods are i) It does not contributes knowledge about the background of the problem, while rules based system are more easily comprehensible. ii) In the majority of cases it does not permits an incremental training and including uninterruptible black box components, numerous training time, possible data over fitting and unsuited to such task (eg. Calculating Pay roll) we should balance its strengths against limitations when compared to traditional statistical techniques. The discipline of ANN not a panacea and still it is immature. It will not replace traditional quantitative techniques completely.

References

1. Gimblett, R. H. and G. L. Ball, 1995. Neural network architectures for monitoring and simulating changes in forest resources management. *AI Applications* **9**:103123.

2. Coulson, R. N., L. J. Folse and D. K. Loh, 1987. Artificial intelligence and natural resource management. *Science* **237**:262-267.

3. Pham, D. T., X. Liu, 1995. Neural networks for identification, prediction and control. Springer-Verlag London Limited, Great Britain. 238 p.

Brown, D., P. Rothery, 1993. Models in biology: Mathematics, statistics and computing. John Wiley and Sons Ltd., England. 688 p.

George, R. K., S. Rammohan, M. S. Kulshretha, A. M. Shekh and H. Jaita, 2001. Prediction of soil temperature using neural networks. *Journal of Agrometeorology* 3(1&2):169-173. India

McCulloch, W.C and W. Pitts, 1943. A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics* 5:115-133.

Atkinson, P. M. and A. R. Tatnall, 1997. Introduction: Neural networks in remote sensing, *Int. J. Remote Sensing* 18:699-709.

Zurada, J. M., 1998. Introduction Artificial Neural Systems. Jaico publishing house, New Delhi, India.

Changhui Peng1 and Xuezhi Wen2, 1999. Recent Applications of Artificial Neural Networks in Forest Resource Management: An Overview In: Environmental Decision Support Systems and Artificial Intelligence, Ulises Corté and Miquel Sànchez-Marrè, Cochairs (eds.).pp.15-22.

Technical Report .1999.WS-99-07, AAAI Press, Menlo Park, CA,

Gail Brion, Chandramouli Viswanathan, T. R. Neelakantan, Srinivasa Lingireddy, Rosina Girones, David Lees, Annika Allard and Apostolos Vantarakis, 2005. Artiûcial Neural Network Prediction of Viruses in Sheliûsh. Applied and Environmental Microbiology, Sept. 2005, 71(9):5244-5253.

Singhal, J. Zhang, H. E. Michel and B. R. Singh, .2002. Fish-Quality Analysis using Artificial Neural Networks and Spectroscopic Data. Proceedings of International Conference on Artificial Intelligence and Soft Computing ASC Banff, Canada P.357-102.

Andy P., Dedecker, Peter L.M. Goethals, Wim Gabriels, Niels De Pauw, 2004. Optimization of Artiûcial Neural Network (ANN) model design for prediction of macro invertebrates in the Zwalm river basin (Flanders, Belgium). Ecological Modeling 174:161-173.

Maran, E., M., NoviÄ, P. Barbieri, J. Zupan, 2004. Application of counter propagation artificial neural network for modeling properties of fish antibiotics. SAR and QSAR in *Environmental Research*, 15(5-6):469-480(12).