Multiplicative-Additive Neural Networks with Active Neurons

Mêuser Valença[1] and Teresa Ludermir[2]
{mjsv, tbl}@di.ufpe.br
Companhia Hidro-Elétrica do São Francisco[1]
Universidade Federal de Pernambuco – UFPE, Brazil[2]

Abstract

An Artificial Neural Network is a flexible mathematical structure which is capable of identifying complex nonlinear relationships between input and output data sets. Such Neural Networks have been characterized by passive neurons that are not able to select and estimate their own inputs. In a new approach, which corresponds in a better way to the actions of human nervous system, the connections between several neurons are not fixed but change in dependence on the neurons themselves. This paper is concerned with the applications of the self-organization multiplicative–additive algorithm with active neurons to prediction models of river flow. The nonlinear multiplicative–additive model (MAM) approach is shown to provide better representation of the weekend average water inflow forecasting, in comparison to the models based on Box-Jenkins method, currently in use on the Brazilian Electrical Sector.

1. Introduction

Models for river flow forecast are a fundamental tool in water resource studies, since they are in charge of establishing future reservoir water inflow. These predictions are of central importance in the planning of a water resource system, being responsible for the optimization of the system as a whole.

Among the tradition techniques used for this purpose, we highlight conceptual models for simulation and the linear time-series models, such as ARIMA (Auto regressive Integrated Moving Average) models, developed by Box-Jenkins, 1976 [1]. Conceptual models are designed to represent the general internal sub-processes and physical mechanisms which govern the hydrologic cycle. While these models ignore simple aspects such as spatially distributed and time-varying, they attempt to incorporate realistic representations of the major nonlinearities inherent in the rainfall- runoff relationships. However, the implementation and calibration of such a model can typically present various difficulties, requiring sophisticated mathematical tool, significant amounts of calibration data and some degree of expertise and experience with respect to the model.

For this reasons, several companies in the Brazilian Electrical Sector use the linear time-series models such as ARMA (AutoRegressive Moving Average) models developed by Box-Jenkins [1]. These models are relatively easy to develop and implement and they have been found to provide satisfactory predictions in many applications.

The work presented here aims to develop alternative models to the forecast of daily average water inflow of the Sobradinho Hydroelectric power plant, part of the CHESF (Companhia Hidrelétrica do São Francisco) system. This dam is located at São Francisco River, in Northeast of Brazil. We propose the use of multiplicative-additive algorithm with actives neurons as an alternative to weekend average water inflow forecasting to this dam.

We evaluate the results obtained by the use of self-organizing of nets with actives neurons [2] (GMDH-Group Method of Data Handling) [3][4] against results from the applications of the traditional Box-Jenkins models.

Section 2 brings an overview of the multiplicative-additive algorithms, followed by a brief presentation of Box-Jenkins in the section 3. In section 4, we describe our simulation results for the method. Finally, section 5 brings conclusions.

2. Multiplicative Additive Algorithm

The algorithms of Group Method of data Handling (GMDH) combines the best of both statistics and neural networks. GMDH creates adaptive models from data in the form of networks of active neurons in an evolutionary

0-7803-5529-6/99/$10.00 ©1999 IEEE

fashion of repetitive generation of populations of alternative models of growing complexity, and corresponding model validation and selection, until an optimal complex model has been created. Nets with active neurons using GMDH algorithm should be applied to rise up accuracy of short-term forecasts and therefore, to increase the lead time of step-by-step longterm forecast. Active neurons choose their inputs by the process of structure self-organization and, therefore, solve the difficult problem of neuronet self-organization.

When the basic inductive algorithms, where the variables have integer powers [5], do not lead to unbiased and accurate predictions, it is necessary to shift the solution space to another region of functional space; for example, to the region of polynomials with other than integer powers of generalized arguments. This is possible with the following multiplicative-additive algorithm. First of all, on has to choose certain multiplicative-additive models with optimal complexity on the basis of the external criteria. An original model is represented in the form of a product of given arguments with unknown powers;

$$y_i = a_{0i} x_1^{k1(i)} x_2^{k2(i)} x_3^{k3(i)} \ldots x_m^{km(i)} \tag{1}$$

This can be rewritten in the following form by taking logarithms on both sides:

$$\ln(y) = \ln(a_0) + k_1 \ln(x_1) + k_2 \ln(x_2) + \ldots + k_m \ln(x_m) \tag{2}$$

Using the original data table of the quantities y, x_1, x_2, ..., x_m, a new data table for the variables with the algorithms values can be set up. Data is separated into training, testing, and examining sets. Several partial, but best, models can be chosen by using the inductive learning algorithm multilayer [6] with the combined criteria of "minimum-bias plus prediction".

At the second level, to obtain the generalized multiplicative-additive model, we combine the selected multiplicative models into a single complete polynomial as:

$$Y = b_0 + b_1 . y_1 + b_2 . y_2 + \ldots + b_n . y_n \tag{3}$$

where Y is the desired output of the process; y_j j=1, ..., n are the estimated outputs of the selected multiplicative models; and b_j , j=1,...,n are the coefficients. The combinatorial algorithm enables us to obtain a unique optimal model. This model can be rewritten in terms of the original input variables.

3. The Box-Jenkins Model

The Box-Jenkins method is one of the most popular time series forecasting methods . The method uses a systematic procedure to select an appropriate model from a rich family of models, namely, ARIMA models. Here AR stands for auto regressive and MA for moving average. AR and MA are themselves time series models. The former model yields a series value x_t as a linear combination of some finite past series values, x_{t-1}, x_{t-2}, ..., x_{t-p}, where p is an integer, plus some random error e_t . The latter gives a series value x_t as a linear combination of some finite past random errors, e_{t-1} , e_{t-2} , e_{t-q} , where q is an integer. P and q are referred as orders of the models. AR and MA models can be combined to form an ARMA model. Most time series are ``non-stationary" meaning that the level or variance of the series change over time. Differencing is often used to remove the trend component of such time series before a ARMA model can be used to describe the series. The ARMA models applied to the differenced series are called integrated models, denoted by ARIMA.

A general ARIMA model has the following form:

$$\phi(B^w)\,\phi(B)\,(1-B^w)^D\,(1-B)^d\, Z_t = \theta(B^w)\,\theta(B)\,\varepsilon_t \tag{4}$$

where $\phi(B)$ and $\theta(B)$ are auto regressive and moving average operators respectively; w is the back shift operator; ε_t is called random error with normal distribution N(0; σ^2) and Z_t is the time series data, transformed if necessary.

The Box-Jenkins method performs forecasting through the following process:

a. Model Identification: The orders of the model are determined.

b. Model Estimation: The linear coefficients of the model are estimated (based on maximum likelihood).

c. Model Validation: Certain diagnostic methods are used to test the suitability of the estimated model. Alternative models may be considered.

4. Simulation Results

Our experiment uses data from the Sobradinho hydroelectric power plant, situated in the São Francisco river in Northeast of Brazil. The values measured from 1929 to 1988 were used to train and test the network, and to estimate the parameters of the statistical models. Values measured from 1989 to 1997 were used to evaluate the performance of the Box-Jenkins model and the multiplicateve-additive model (MAM).

The cross-correlation between inputs and outputs variables, the correlation structure of the output variable and the physical features of the problem were taken into account on the establishment of the structure of the examined models and of the number of neurons in the first layer used in the forecast of future values.

In the evaluation of the performance of the network and of the statistical model we used the absolute average error (AAE) (Equation 5), the absolute average percentual error (AAPE) (Equation 6) and the forecast standard error (SE) (Equation 7).

$$AAE = \frac{1}{N}\left[\sum_{p=1}^{N}\left|Z_p - Z_o\right|\right] \quad (5)$$

$$AAPE = \frac{1}{N}\left[\sum_{p=1}^{N}\left|Z_p - Z_o\right|/Z_o\right].100 \quad (6)$$

$$SE = \frac{1}{N}\left[\sum_{p=1}^{N}(Z_p - Z_o)^2\right]^{0,5} \quad (7)$$

Table 1 below shows a comparative study of the better results obtained with Box-Jenkins models and the results obtained with multiplicative-additive model.

	prediction					
	1 step ahead			2 steps ahead		
Models	AAPE	AAE	SE	AAPE	AAE	SE
MAM	13,3	287	492	21,6	471	852
Box-Jenkins	14,1	304	518	22,7	490	865

Table 1 – Comparison the models

5. Conclusions

The multiplicative-additive algorithms are explicit mathematical models, obtained in a relatively short time on the basis of extremely small samples. This model provides better results than the statistical models with respect to the three errors types mentioned above. This is due to the fact that Multiplicative-additive networks solve the well-known problems of choosing the optimal network architecture by means of an adaptive synthesis (objective choice) of the architecture to provide a parsimonious model for the particular desired function. The statistical models, in general, do not generate such good results.

References

[1] G. E. P. Box & G. M. R. Jenkins. Time Series Analysis-Forecasting and Control. Holden Day. California, 1976.

[2] *A. G.* Ivakhnenko, "Self-Organisation of Neuronet with Active Neurons for Effects of Nuclear Test Explosions Forecasting," System Analysis Modeling Simulation (SAMS*), vol.20, pp.107-116, 1995.*

[3] A. R. Barron, "Predicted squared error: A criterion for automatic model selection, "Self-Organizing Methods in modeling, S. Farlow and Marcel Dekker, Eds. 1984, chap. 4.

[4] Barron, A.R., Barron, R.L. Statistical learning networks*: a* unifying *view*. Proceedings of the 20 th. Symposium Computer Science and Statistics 1988.

[5] M. J. S. Valença and T. B. Ludermir, "Self-Organizing modeling in forecasting daily river flows,"Vth Brazilian Neural Networks Symposium on -Brazilian Computer Society (IEEE), vol. 1, pp. 210-214, 1998.

[6] M. J. S. Valença, " Analysis and design of Neural Networks using Inductive learning methods. Ph. D. thesis in preparation. UFPE-Brazil, 1999.