Proceedings of V Brazilian Conference on Neural Networks - V Congresso Brasileiro de Redes Neurais
pp. 163-168, Abril 2-5, 2001 - Rio de Janeiro - RJ - Brazil

Multivariate Modelling of Water Resources Time Series Using Constructive Neural Networks

Mêuser Valença[1], Teresa Ludermir[2]
[1]Chesf/UNIVERSO-Recife
Rua Gaspar Peres 427/104, CEP 50670-350,CDU
[2]Universidade Federal de Pernambuco
Cx. Postal 7851 CEP 50732-970
e-mails: meuser@elogica.com.br, tbl@cin.ufpe.br

Abstract

This paper presents a constructive neural network model for daily streamflow forecasting. The Surface water hydrology is basic to the design and operation of the reservoir. A good example is the operation of a reservoir with an uncontrolled inflow but having a means of regulating the outflow. If information on the nature of the inflow is determinable in advance, then the reservoir can be operated by some decision rule to minimize downstream flood damage. For this reasons, several companies in the Brazilian Electrical Sector use the linear time-series models such as PARMA (Periodic Auto regressive Moving Average) models developed by Box-Jenkins. This paper provides for river flow prediction a numerical comparison between neural networks, called non-linear sigmoidal regression Blocks networks (NSRBN), and PARMA models. The NSRBN model approach is shown to provide better representation of the daily average water inflow forecasting, than the models based on Box-Jenkins method, currently in use on the Brazilian Electrical Sector.

1. Introduction

Models Streamflow forecasting are an essential requirement for solving a wide range of scientific and/or management tasks. Conceptual models [1][2] offer one possible forecasting method, but such tools are often considered to be too complex for most practical implementations. While conceptual models are of importance in the understanding of hydrologic processes, there are many practical situations such as streamflow forecasting where the main concern is with making accurate predictions at specific watershed locations. In such situation, the linear time series models such as PARMA [3][4][5] models have been used.

The work presented here aims to develop alternative models to the forecast of daily average water inflow for the Boa Esperança Hydroelectric power plant, part of the Chesf (Companhia Hidrelétrica do São Francisco) system. This dam is located at Parnaíba River, in the borderline between Maranhão and Piauí, two Brazilian States.

We propose the use of constructive networks as an alternative to daily average water inflow forecasting to this dam [6][7][8][9]. We evaluate the results obtained by the use of NSRBN against results from the applications of the traditional Box-Jenkins models.

Section 2 brings an overview of the NSRBN algorithms, followed by a brief presentation of PARMA model in the section 3. Section 4 shows an evaluation of the obtained results. Finally, section 5 concludes the paper.

2. NSBRN -NON-LINEAR SIGMOIDAL REGRESSION BLOCKS NETWORKS

2.1. Network Architecture

The goal here are to present practical methods to realize compact networks using a model with hidden units with sigmoidal blocks activation functions [9]. The activation function is

$$f(x) = (\sigma_{net(h)} + \theta_h) \tag{1}$$

where:

h is the order of the block (= number of hidden units);

θ_h is bias and $\sigma_{net(h)}$ is the hyperbolic tangent function.

However the analyses of the are not restricted to hyperbolic tangent function.

The first design step is to divide f (x) up into blocks of equal-degree terms, as in Figure 1. That is

$$f(x) = \tilde{f}_1(x) + \tilde{f}_2(x) + \ldots + \tilde{f}_d(x) \tag{2}$$

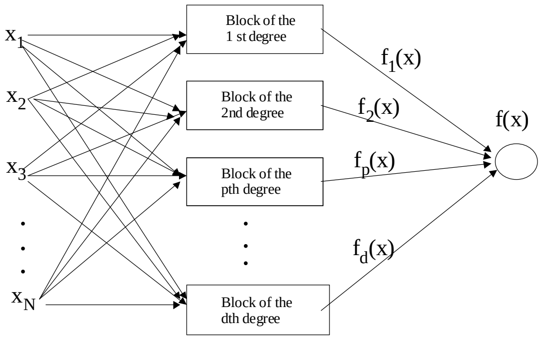

Figure 1- NSRBN (network architecture)

The block approach is to realize all terms in $\tilde{f}_p(x)$ functions at the same time, as in Figure 2.

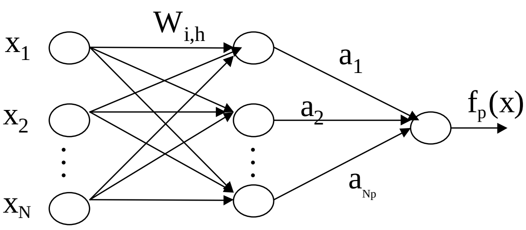

Figure 2- Block of degree p architecture

The input **x** is an N dimensional vector and x_i is the i-th component of **x.** The inputs are weighted and fed to a layer of h hidden units, where h is the order of the block [10]. Let $f_p(x)$ be the output of the block of degree p. Then,

$$f_p(x)=a_1(f(net(1)))^1+a_2(f(net(2)))^2+a_3(f(net(3)))^3 +...+a_p(f(net(p)))^p \tag{3}$$

and a_h is the weight between h-th hidden unit to output unit and h=1,2,...,p.

A non-linear sigmoidal regression blocks networks (NSRBN) is defined as a feedforward network based on Eq. 3. A NSRBN [9][11] is a polynomial function that can be represented as

$$f(x)=\sigma_{net(o)}(\sum_{p=1}^{d} f_p(x)) \tag{4}$$

2.2. Constructive learning method

The constructive algorithm for NSRBN is based on the constructive learning method. The goal here is to present a practical method to realize NSRBN using blocks of the sigmoidal function. The constructive learning algorithm proceeds as follows. We denote p an algorithmic step at which f_k is added to the network. Therefore an unknown function f is successively approximated by

$$f(x)=\sigma_{net(o)}(\sum_{p=0}^{d-1} f_p(x)+f_d(x)) \tag{5}$$

where weights in $f_{k-1}(x)$ are frozen once the k-th degree $f_k(x)$ is added.

2.3. Why d hidden units?

Zhang-Yu and Jia [12] give a direct proof of approximation property of a multi-layer perceptron with one hidden layer (MLP1) in single input-output situation, and find out the relationship between the order of the best approximation polynomial and the number N of hidden units of a MLP1 for a given function f(x) and a finite N. Based in this proof, we formulated a more general and efficient NSRBN.

2.3.1. Relationship between the order of the best approximation polynomial and the number N of hidden units of a MLP1:

Suppose $f(x) \in C_0$ [-a,a], and $f(x)=f_0(x)+f_e(x)$, where $f_0(x)$ is the odd part of f(x) and $f_e(x)$ is its even part. Let the output of a MLP1 with N_0 be $P_0(x)$, N_0 is fixed, hidden units. So, the output $P_0(x)$ can approximate $f_0(x)$ in the degree of accuracy better than its $2(2N_0 - 1)$ order Best Polynomial Approximation (BAP).

Let the output of a MLP1 with $2N_e$ be $P_e(x)$, N_e is fixed, hidden units. So, the output $P_e(x)$ can approximate $f_e(x)$ in the degree of accuracy better than its $2(3N_e - 1)$ order best polynomial approximation (BAP). Using a MLP1 with $N = N_0 + N_e$ hidden units to approximate f(x), combined approximation, we get its output: $P(x)= P_0(x)+ P_e(x)$ where

$$P(x)=\sum_{n=0}^{2(6N/7)-2} a_n x^n \tag{6}$$

The original form of the proof [12] is more complicated than the one presented here. Since we are concerned with the existence of a representation of multivariate polynomials in terms of NSRBN, a simpler statement is adopted.

From the analysis above, it is clear that to approximate an even function, we need many double hidden units to combine some even base functions, thus the freedoms of parameters are lost, and this case may affect the approximation of the MLP1. So we propose a NSRBN in which some of hidden units have even blocks activation function and the rest have odd blocks activation function. This kind of hybrid activation function model may have stronger nonlinear mapping ability.

2.4 Neurobiological plausibility of NSRBN

Are NSBRN computing structures even remotely neurobiologically plausible? In neurophysiology, the possibility that dendritic computations could include local multiplicative nonlinearities is widely accepted [13]. Indeed, Durbin and Rumelhart [14] observe that there is a natural neurobiological interpretation of a combination of product and summing units in terms of a single neuron or a group of neurons. Mel has recently proposed clusteron as an abstraction for a complex neuron that can extract higher order statistics from input stimuli [15]. In his model, a dendritic tree receives weighted synaptic contacts from a set of afferent axons. Each synaptic contact is given by a product of direct stimulus intensity and a weighted sum of neighborhood activity. We note that this description translates to an NSRBN, which can be considered as a mathematical abstraction of the clusteron model.

3. The PARMA Model

Let us consider the original periodic series $x_{v,\tau}$, where v denotes the year, $\tau = 1,...,\omega$ and ω is the number of time intervals in the year. Assuming that the distribution of the series is skewed, an appropriated transformation can be used to transform $x_{v,\tau}$ to the normal series $y_{v,\tau}$.

Then the periodic PARMA [5] model for $y_{v,\tau}$ can be written as

$$y_{v,\tau} = \mu_\tau + \sigma_\tau . z_{v,\tau} \tag{7}$$

where μ_τ and σ_τ are the periodic mean and periodic standard deviation and $z_{v,\tau}$ may be represented by an PARMA model .

The PARMA(p,q) model with time-varying coefficients as

$$z_{v,\tau} = \sum_{j=1}^{p} \phi_{j,\tau} . z_{v,\tau-j} - \sum_{i=1}^{q} \theta_{i,\tau} . \varepsilon_{v,\tau-i} + \varepsilon_{v,\tau} \tag{8}$$

where $\phi_{j,\tau}$ and $\theta_{i,\tau}$ are time varying autoregressive and moving average coefficients, respectively, and $\varepsilon_{v,\tau}$ is an independent and identically distributed normal random variable.

4. Evaluation of the obtained results

Our experiment is being undertaken on the basis of data from the Boa Esperança hydroelectric power plant, which constitutes a natural borderline between the Brazilian States of Maranhão and Piauí. Figure 3 shows the Parnaíba River basin with the location of this dam, as well as its River flow stations, which are controlled by CHESF.

The River flow stations used here are: Alto Parnaíba, Ribeiro Gonçalves, Balsas, São Félix de Balsas e Boa Esperança. These stations present measures of daily average flow since 1966.

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 [16][17][18].

The values measured from 1966 to 1990 (which corresponds to a total of 8797 values of daily flow) were used to train of the statistical and test the NSRBN and to the estimation of parameters the statistical models. Values measured from 1991 to 1997 were used to evaluate the performance of the Box-Jenkins model and the NSRBN (which amounts to a set of 2435 daily flow values).

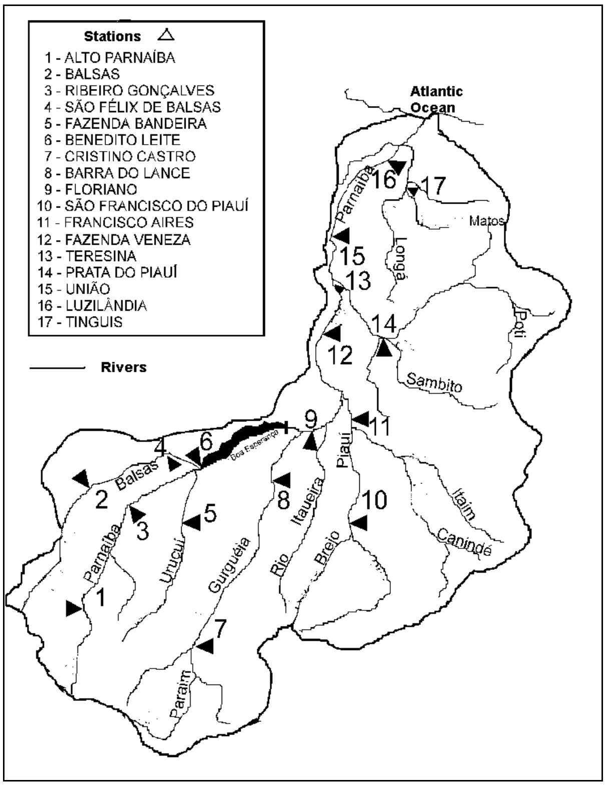

Figure 3 – Parnaíba River Basin

In the evaluation of the performance of the NSRBN and of the PARMA(p(i),q(i)) model (Multivariate ARMA models) we used the absolute average error (AAE) (Equation 9), the absolute average percentual error (AAPE) (Equation 10) and the forecast standard error (FSE) (Equation 11).

$$AAE = \frac{1}{N} * \sum_{i=1}^{N} [\,| Zp - Zo |\,] \qquad (9)$$

$$\text{AAPE}(\%) = \frac{1}{\text{N}} * \{ \sum_{i=1}^{N} [\,| Zp - Zo | / Zo \} * 100 \qquad (10)$$

$$\text{FSE} = [\frac{1}{\text{N}} * \sum_{i=1}^{N} (Zp - Zo)^2]^{0.5} \qquad (11)$$

where : Z_p – forecast value ; Z_o – measured value and N – number of values.

Among the tested multivariate models, we highlight the one which provided better results. This model uses as inputs for the forecast of daily average water inflow from the Boa Esperança dam: the last four values from the station of Alto Parnaíba (ATP), Balsas(BLS), São Félix de Balsas(SFB) e Ribeiro Gonçalves(RBG) and three values from the variable itself, the Boa Esperança station.

This model can be represented by the following mathematical equation:

$$\begin{aligned} Q_{UBE} &= \phi_1^1 Q_{UBE-1} + \phi_2^1 Q_{UBE-2} + \phi_3^1 Q_{UBE-3} \\ &+ \phi_1^2 Q_{ATP-1} + \phi_2^2 Q_{ATP-2} + \phi_3^2 Q_{ATP-3} + \\ &\phi_4^2 Q_{ATP-4} + \phi_1^3 Q_{RGB-1} + \phi_2^3 Q_{RGB-2} + \\ &\phi_3^3 Q_{RGB-3} + \phi_4^3 Q_{RBG-4} + \phi_1^4 Q_{BLS-1} + \\ &\phi_2^4 Q_{BLS-2} + \phi_3^4 Q_{BLS-3} + \phi_4^4 Q_{BLS-4} + \\ &\phi_1^5 Q_{SFB-1} + \phi_2^5 Q_{SFB-2} + \phi_3^5 Q_{SFB-3} + \phi_4^5 Q_{SFB-4} \end{aligned} \qquad (12)$$

In this expression, Q represents the daily flow in each station, followed by the already introduced abbreviations . For instance, Q_{UBE} represents the flow to be forecast on a given day in Boa Esperança station; Q_{UBE-1} is the flow in this station on the previous day; Q_{UBE-2} is the flow 2 days before, and so on.

The inputs of the window used for the NSRBN are identical to the one used in the statistical model.

Table 1 shows a comparative study of the better results obtained with Box-Jenkins models and the results obtained with NSRBN.

Table 1 – Comparison between NSRBN and Box-Jenkins

forecasts	Models	AAPE (%)	AAE (m^3/s)	FSE (m^3/s)
1 day ahead	NSRBN	6,3	24	43
	PARMA(p,q)	9,3	42	82
2 day ahead	NSRBN	7,8	35	76
	PARMA(p,q)	11,4	52	89
3 day ahead	NSRBN	10,4	47	83
	PARMA(p,q)	13,5	59	99

The results obtained by the use of the two methods show a clear superiority of the NSRBN in relation to the Box-Jenkins

The Figures 4,5 and 6 shows, for example, the tracking behavior of the predicted values by the NSRBN and the predicted values by the Box-Jenkins models of the Parnaíba River flow in Boa esperança dam for the period from 15/01/92 to 24/02/92 (figure 4), 09/03/97 to 18/04/97 (figure5) and from 31/10/92 to 31/08/93 (figure6).

.

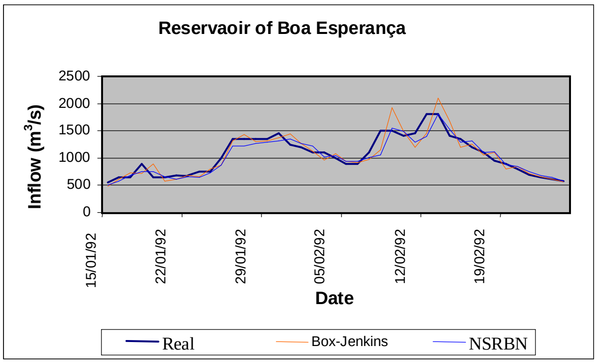

Figure 4 – NSRBN vs Box-Jenkins, 1992 –1 day ahead

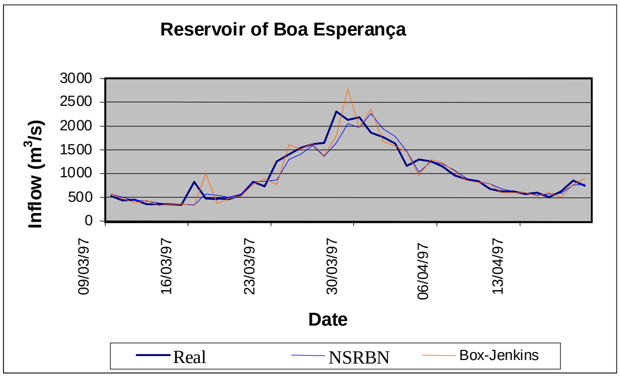

Figure 5 – NSRBN vs Box-Jenkins, 1997 –1 day ahead

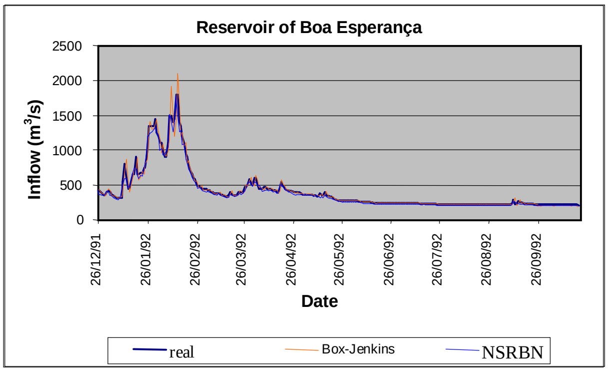

Figure 6 – NSRBN vs Box-Jenkins, 92/93 –1 day ahead

5. Conclusions

The NSRBN is a powerful modelling and prediction of complex linear or non-linear multi-input/multi-output systems.

The results obtained in the evaluation of the performance of NSRBN were better than the results obtained with statistical models with respect to the three types of errors. NSRBN provides good results because the well-known problems of an optimal (subjective) choice of the neural network architecture are solved in the NSRBN algorithms 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 those good results.

The NSRBN models described in this work are in use in Chesf for the forecasting river flow from one day to seven days ahead.

References

[1] U. S. Army Corps of Engineers Report. Simulation of Flood Control and Conservation Systems. Users Manual, HEC-Hydrologic Engineering Center, 1976.

[2] N. Crawford and R. K. Linsley . Digital simulation in hydrology : Stanford Watershed Model IV. Technical Report 39, Stanford University , Califórnia, 1966.

[3] B. M. Fiering. Streamflow Syntehsis. Harvard University Press, Cambridge, Massachusetts, 1967.

[4] J. D. Salas, J. W. Delleur, V. Yevjevich and W. L. Lane. Applied Modeling of Hydrologic Time Series. Water Resources Publications, Colorado, 1980.

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

[6] M. J. S. Valença and T. B. Ludermir. Multiplicative-Additive Neural Networks with Active Neurons. International Joint Conference on Neural Networks (IJCNN), IEEE, Book of Summaries (2073), Washington, DC, July 1999.

[7] M. J. S. Valença and T. B. Ludermir. Self-organization Sigmoidal Blocks Networks. International Joint Conference on Neural Networks (IJCNN), IEEE, Book of Summaries (2086), Washington, DC, July 1999.

[8] M. J. S. Valença and T. B. Ludermir. Uma Rede Neural Construtiva com Atualização dinâmica dos Pesos. IV Congresso Brasileiro de Redes Neurais, pages 114-117, ITA, São José dos Campos, Julho 1999.

[9] M. J. S. Valença. Analysis and Design of the constructive neural networks for complex systems modeling (in portuguese). Ph.D Thesis, UFPE, Brazil, 1999.

[10] B. J. T. Morgan. The cubic logistic model for quantal assay data. Applied Statistics, 34: 105-113.

[11] T. Chen and H. Chen. Universal Approximation to Nonlinear Operators by Neural Networks with Arbitrary Activation Functions and Its Applications to Dynamical Systems. IEEE Transactions on Neural Networks, **6**(4): 911-917, July 1995.

[12] Z. Wei, Y. Yinglin and J. Qing. Approximation Property of Multi-Layer Neural Network (MLNN) and Its Application in Nonlinear Simulation. In international Joint Conference on Neural Networks, **1**: 171-176, 1991.

[13] B. W. Mel. and C. Koch. Sigma-pi learning: On radial basis functions and cortical associative learning. In advances in Neural Information Processing Systems 2, D. S. Touretzky. ED. San

Marco, CA: Morgan-Kaufmann ,pages 474-481, San Marco, 1990.

[14] R. Durbin and D. E. Rumelhart. Product units: A computationally powerful and biologically plausible extension to backpropagation networks. Neural Computation, vol. 1, pages 133-142, 1989.

[15] B. W. Mel. The clusteron: Toward a simple abstraction for a complex neuron. In advances in Neural Information Processing Systems 4, S. J. Hanson, J. E. Moody, and R. P. Lippmann, Eds. San Marco, CA: Morgan-Kaufmann ,pages 35-42, San Marco, 1992.

[16] D.T. Pham and X. Liu,. Neural Networks for Identification, Prediction and Contro*l*. Springer-Verlag, 1995.

[17] M. J. S. Valença and T. B. Ludermir. Self-organization Neurons Blocks Networks. International Conference on Computational Intelligence and Multimedia Applications (ICCIMA), IEEE, pages 60-64, New Delhi, India, September 1999.

[18] M. J. S. Valença and T. B. Ludermir. Uma nova rede neural polinomial com aplicação na previsão de vazões. V Simpósio Brasileiro de Redes Neurais, . **2**: 273-278, Belo Horizonte 1998.