PREDICTION OF THE LEVEL OF PARAGUAY RIVER USING NEURAL NETWORKS[1]

LI WEIGANG[2], LEONARDO DEANE DE ABREU SÁ[3], GERALDO PEREIRA GALVÃO and RUTE MARIA BEVILAQUA[4]

ABSTRACT - Backpropagation neural networks are implemented for prediction of the level of Paraguay River at Ladário city, MS. Using 274 monthly mean values, the trained network predicts the levels of the four next months with relative errors smaller than 17%. For some special points, the prediction results also show that the neural network method seems to be useful to predict time series related to phenomena influenced by complex climatic and geophysical processes, and it does not deal directly with causal relationships involved in the phenomena studied. A discussion about the variability of the estimation errors for different predicted data is carried out here.

Index terms: climate, geophysics.

PREVISÃO DOS NÍVEIS DO RIO PARAGUAI USANDO REDES NEURAIS

RESUMO - Redes neurais com retropropagação são usadas para fornecer a previsão do nível do rio Paraguai em Ladário, MS. Usando 274 valores médios mensais, a rede neural treinada prevê o nível dos quatro meses seguintes com erros relativos menores do que 17%. Para alguns pontos especiais, os resultados da previsão também mostram que o método da rede neural parece ser útil para prever séries temporais relacionadas a fenômenos influenciados por processos climáticos e geofísicos complexos. Isto sem que se trate diretamente das relações causais envolvidas nos fenômenos estudados. Também é efetuada uma discussão sobre a variabilidade dos erros de estimação para diferentes valores previstos.

Termos para indexação: clima, geofísica.

INTRODUCTION

Prediction of nonlinear time series is an interesting recent application of neural networks. There are a number of prediction methods available for this kind of problem, such as Polynomials, Rational polynomials, Neural Networks and Radial basis functions (Casdagli, 1989; Kim & Stringer, 1992). Neural networks were found to be useful and competitive with the best recent nonlinear approximation methods (Lapedes & Farber, 1987; Gallant & White, 1992; Gershenfeld & Weigend, 1993; Li et al., 1996; Marzban & Stumpf, 1996).

In this paper, a Backpropagation Neural Network (BPNN), one of the most important developments in neurocomputing (Rumelhart & McClelland, 1986; Hecht-Nielsen, 1990), is implemented for the prediction of the level of Paraguay River, at Ladário city, MS. This river, with a length of 2,550 km, rises in the Mato Grosso region of Brazil and runs Southward between highlands at the West and the Brazilian plateau at the East. Its basin, with an area of approximately 150,000 km^2, consists of a series of huge alluvial plains drained by a complex network of rivers interspersed with marshes, in a region called Pantanal. In this region, many areas suffer a succession of droughts and severe floods with their obvious economic and social consequences.

The levels of Paraguay River are influenced by several different factors from micro to macro scales. Therefore, predicting the level of the Paraguay River

[1] Accepted for publication on March 2, 1998.

[2] Eng. Civil, Dr., Dep. Ciência da Computação, Universidade de Brasília, Caixa Postal 4466, CEP 70919-970 Brasília, DF, Brazil. E-mail: weigang@cic.unb.br

[3] Físico, Dr., Instituto de Pesquisas Espaciais (INPE). Av. dos Astronautas, 1578, Caixa Postal 515, CEP 12201-970 São José dos Campos, SP. E-mail: leo@met.inpe.br

[4] Físico, M.Sc., INPE.

with convenient antecedence (and so estimating the area to be flooded) is a relevant scientific, social and economic goal. Neural networks can treat all these factors simultaneously, which was an incentive to apply this method to complex time series such as those of river levels. The daily data of the series studied in this work were collected from January 1900 to June 1995 (the monthly mean data of the level of Paraguay River is shown in Fig. 1).

In the following sections, firstly, the basic concept of neural networks and its application for prediction are shown. Secondly, the neural networks to predict the next four months data of Paraguay River levels are applied. Finally, a brief discussion of the results and the conclusions are given.

MATERIAL AND METHODS

For a long time, the linear models were the general theoretical framework for time series analysis and prediction. However, there are many cases for which linear models are inadequate for accurate predictions as in the case of forecasting the levels of the Paraguay River, particularly when the data has a broadband power spectrum that cannot be well modelled by a linear approximation (Gershenfeld & Weigend, 1993). In such situations, more complex nonlinear function representations need to be applied to the data, like those of machine learning, typified by neural networks, that can adaptively explore a large space of potential models. In this way, a Backpropagation Neural Network (BPNN) is implemented in the Stuttgart Neural Network Simulator (SNNS) developed at the Institute for Parallel and Distributed High Performance Systems (IPVR) at the University of Stuttgart (Zell et al., 1995).

A neural network is an interconnected network of simple processing elements. Communication between processing elements occurs along paths of variable connection weights. By changing the values of these connection strengths (weights) the network can collectively produce complex overall behavior (Welstead, 1994). The network has three layers: an input layer, an output layer, and a layer in the middle, not connected directly to the input or output, called the hidden layer. The output layer has a single output x(t), the input values

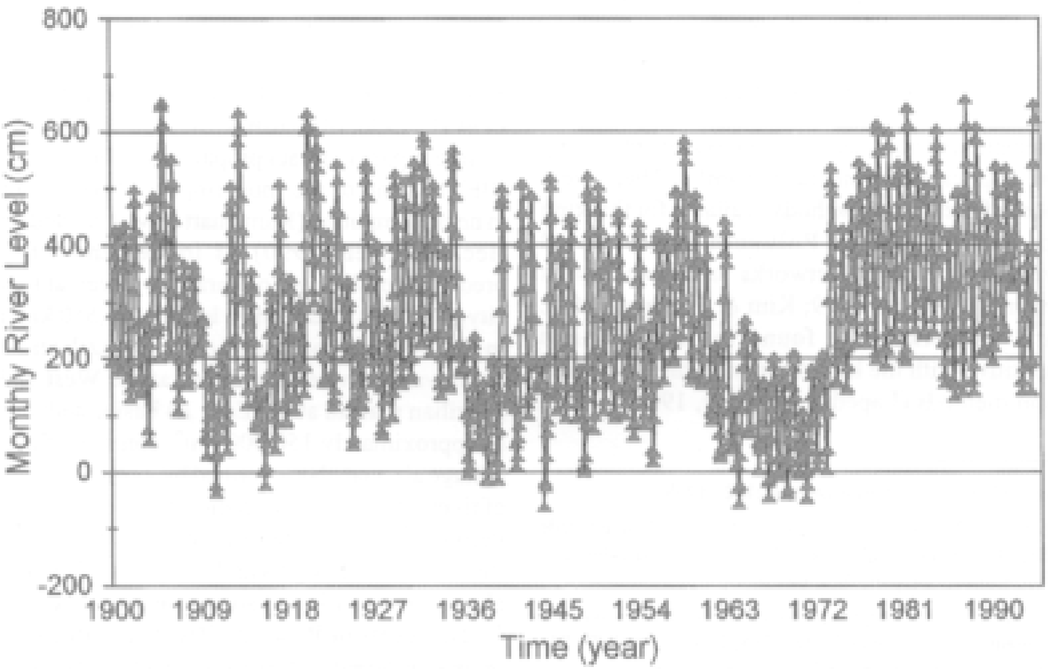

FIG. 1. The monthly mean data of the level of Paraguay River.

of the time series x(t–1), x(t–2),...x(t–d) are received through d input units, which simply pass the input forwards to the hidden units u_j, j = 1,2,...,q. Each connection performs a linear transformation determined by the connection strength (weight) w_{ij}, so the total input for hidden unit u_j is $\Sigma_{i=1}^{d}\ w_{ij}\ x(t\text{-}i)$. Each unit performs a nonlinear transformation on its total input, producing the output:

$$u_j = \Psi(W_{0j} + \sum_{i=1}^{d} W_{ij} X(t-i)) \qquad (1)$$

The activation function Ψ is the same for all units, but each unit may have its own bias w_{0j}, representing an external input or the neuron's intrinsic activity level. Here, Ψ is a sigmoid function with limiting value 0 and 1 as $u_j \rightarrow -\infty$ and $u_j \rightarrow +\infty$, respectively:

$$\Psi(u_j) = \frac{1}{(1+e^{-u_j})} \qquad (2)$$

The hidden layer outputs u_j are passed along to the single output unit with connection strength v_j, which performs an affine transformation on its total input. Then, the network's output x(t) can be represented as:

$$x(t) = v_0 + \sum_{j=1}^{q} v_j . \Psi . \left(w_{0j} + \sum_{i=1}^{d} w_{ij} . x(t-i) \right) \qquad (3)$$

for *d* inputs and *q* units in the hidden layer.

Training a Neural Network involves the minimization of the mean-square error (MSE) of the outputs of the network:

$$MSE(w) = \frac{1}{N} \sum_{k=1}^{N} [\hat{x}_k(t) - x_k(t,w)]^2 \ cm^2 \qquad (4)$$

where the w is the weight set of w_{ij} of the network, *N* is the number of cases in the training set, and $\hat{X}_k$ (t) and $x_k(t,w)$ are the actual and the predicted values of a single output.

One way to make predictions at various next steps t, t–1, t–2,...(t–iΔt) is to place previously predicted values at the input lines to bootstrap to higher i values (Lapedes & Farber, 1987). After training a network to predict at t, the predicted values can be fed back to the inputs to predict at t+1, t+2,... etc. In this way, unacceptable prediction results should not occur until k is quite large. The training procedure was performed on a workstation SUN-SPARC 10, running at 51 Mhz frequency and 6 Gb disk capacity.

RESULTS AND DISCUSSION

Firstly, a subset of 274 data to predict the river levels from March to June 1995 was used. Then, 135, 571 and 1,142 monthly mean values were used separately to predict the river levels from March to June 1995 too, in order to show the change in the prediction errors. After that, some special points were predict to show the ability of the neural network. Finally, a discussion of the results is made.

Using the 274 monthly mean values (from May 1972 to February 1995), the trained network predicted the next four months. For this size of data, the training usually takes from 30 to 60 minutes, depending on the structure of the neural network (for example, one hidden layer networks with 12 inputs, 48 hidden units and a single output) and the learning rate (in this study, a 0.2 value is used). The monthly mean values were separated into three sets: training set of 260 data, training test set of 14 data and test set or validation set of 4 data. The training set is used to train the network and the training test set is used to evaluate performance of the network. The test set (four monthly mean values), in Table 1 is used to compare the values predicted by the network in order to measure the predictive ability of the network. As shown in Table 1, the relative errors of the predictions of the next four months are less than 17%. As Papoulis (1990) mentions, for nonlinear prediction the mean-square error is a reasonable criterion to evaluate the actual capability of the network to predict the data. For most neural networks, the mean-square error is well defined. The advantage of mean-squared error is that it uniformly weights each training trial error in accordance with the square of magnitude of the error vector [$\hat{x}_k(t) - x_k(t,w)$] in equation (4). In this case, after 2,000 circles of training, the Root Mean Square (RMS) error of the network is calculated as 0.0013. Fig. 2 shows the reconstruction (from May 1972 to February 1995) and predictions (from March 1995 to June 1995) of the monthly mean levels of the River.

TABLE 1. Prediction of the monthly level of the Paraguay River from March to June 1995.

	Mar.	Apr.	May	June
Observations	542.90	649.87	622.16	588.00[1]
Prediction	506	540	541	510
Relative errors	7%	17%	13%	13%

[1] Mean of the first eight daily observations of June.

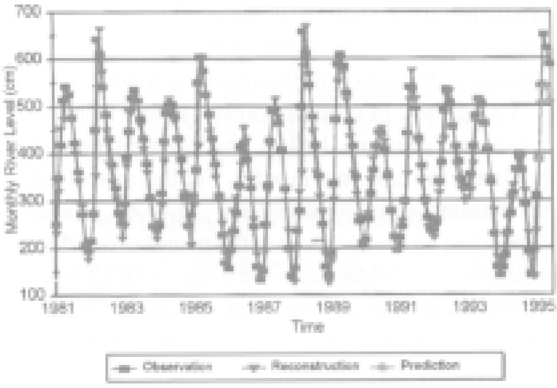

FIG. 2. Reconstruction (from January 1981 to February 1995) and prediction (March to June 1995) of the monthly level of Paraguay River.

For further evaluation of the precision of the predictions of neural network, other subsets of data are used to train the networks to predict some special points: these are from January to April 1948, October 1930 to January 1931, May to August 1959 and October 1972 to January 1973, which display strong deviation from the average behaviour. Table 2 shows the results of the predictions. To predict the data January to April 1948, 564 data are used to train the network. Then, the network predicted the levels of January, February and March 1948 with relative errors of 30% (or less). To predict the data from October 1930 to January 1931, 344 data are used to train the network and the relative errors of the next four months prediction are less than 8%. To predict the data from May to August 1959, a subset of data is used to train the network with sizes 700. The corresponding relative errors are less than 11%. As Fig. 1 shows, the variability of the data from 1970 to 1975 is very strong. The monthly data from January 1900 to September 1971 is used to train the network in order to reconstruct and predict the sharp changes in data from October, 1971 to April, 1972. Fig. 3 shows the reconstruction (from January, 1971 to September, 1971) and the prediction (from October, 1971 to April, 1972). The prediction results show that the neural network detects adequately only the tendency of the monthly mean data during this period. But the relative errors of predictions from October, 1972 to January, 1973 are between 33% and 70% (Table 2).

In some other situations, data records are not complete, or the amount of available data is not enough. So, a question may be proposed: how many data are necessary for prediction using this Neural Network? This is an open problem in the neural network field. To study it the networks were used to test a group of data with 135, 274, 571 and 1,142 of our data. In order to have the same condition, all of these sets were used to train the networks for 2,000 circles. The trained networks were applied to predict the data from March to June, 1995. Table 3 shows the results. One can see that the subset of 274 data gives the best prediction results.

The results show that the ability of the neural network to predict depends on the period chosen for prediction. After realizing this, it is important to make remarks about some characteristics of the variability of the level of a river like Paraguay River to find possible physical explanations for the results found.

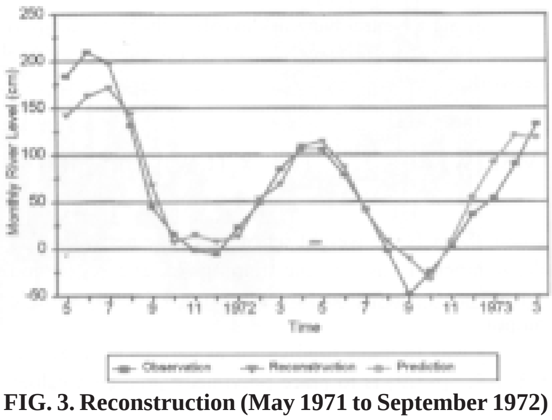

FIG. 3. Reconstruction (May 1971 to September 1972) and prediction (October 1972 to April 1973) of the monthly level of Paraguay River. This Figure shows the capability of the Neural Network for reconstruction and prediction of the tendency of the variation of data.

TABLE 2. Prediction of some special points.

Data used for training neural network	MSE[1] of trained network	Relative error of monthly prediction (%)			
564 (01/1900-12/1947)	0.00090	24 (Jan./1947)	3 (Feb.)	28 (Mar.)	60 (April)
344 (01/1900-09/1930)	0.00124	5 (Oct./1930)	7 (Nov.)	8 (Dec.)	4 (Jan./1931)
700 (01/1900-04/1959)	0.00101	8 (May/1959)	10 (June)	11 (July)	8 (Aug./1959)
861 (01/1900-09/1972)	0.00079	33 (Oct./1972)	- (Nov.)	47 (Dec./1972)	70 (Jan./1973)

[1] MSE in this specific Table is calculated with data to the interval [0.15, 0.85].

TABLE 3. Comparison of the prediction from different training data set.

Data used for training neural network	MSE[1] of trained network	Relative error of monthly prediction (%)			
		March	April	May	June
135 (04/1983 - 02/1995)	0.00157	16	22	16	14
274 (05/1972 - 02/1995)	0.00130	7	17	13	13
571(05/1972 - 02/1995	0.00109	17	24	18	16
1142 (01/1900 - 02/1995)	0.00096	16	22	13	11

[1] MSE in this specific Table is calculated with data to the interval [0.15, 0.85].

In the first place, it is necessary to ask why the river can be predictable, that is, why the data series of the river would have persistence. Why would this data series contain a memory that would permit prediction of the future based on information contained in the past? That introduces the problem of what causes the variation in the level of the river. To answer this question one must take into account that there must be a correspondence between the precipitation that falls on the drainage basin of a river and its flow. Although this relationship is probably nonlinear, it is reasonable to expect that strong anomalies in the precipitation are related to anomalies of the same tendency in the behaviour of a river (Marengo, 1995). Therefore, the flow of the river, and so its height, constitutes one of the most robust integrators of the long-term hydrologic properties of its drainage basin (Richey et al., 1989).

In this way, the effects of persistence or of memory found in the river level should, at least partly, be due to characteristics of persistence of anomalies existing in the factors, which determine the constitution of the hydrologic cycle. Some of such factors are determined by the general atmospheric circulation in their diverse scales (Trenberth, 1985) and their anomalies.

However, some external factors to the ocean-atmosphere system can also eventually influence the variability in the hydrologic cycle. Their action on the hydrologic cycle can occur before one could detect it from the information contained in the memory of the time series. This appears to be the case of the modifications introduced by the volcanic eruptions, whose influence on the radiation budget over large hemispheric areas seems to be appreciable (Robock & Mao, 1995).

It is noteworthy that the period between 1960 and 1970, for which the predictions made by the neural network were not satisfactory, corresponded to the highest values of most of the indexes of global volcanism in the South Hemisphere elaborated by Robock & Free (1995).

The hypothesis is therefore formulate that the decrease in the level of Paraguay River in this period is related to a period of reduction in the flux of solar energy incident to the surface with its consequences on the energy budget and on the hydrological cycle. This decrease was probably caused by the existence of a great number of volcanic aerosols in the atmosphere of the South Hemisphere. These aerosols could introduce alterations in the hydrologic cycle of the Pantanal region in this period.

In such a situation the neural network would not be able to make good prediction because the time-series of the levels of Paraguay River does not contain past information that makes it possible to predict modifications introduced by volcanoes. These physical factors are external to the ocean-atmosphere system. When the variations on the river level are introduced by the factors internal to the ocean-atmosphere system, they have considerable persistence. Then, the predictions of the neural networks are acceptable. That happens, for example, in the prediction of May to August 1959 and March to June 1995.

CONCLUSIONS

1. The trained networks predict the levels of the Paraguay River at Ladário, MS, for the four next months with relative errors smaller than 17%.

2. The ability of prediction of the neural network varies accordingly to the prediction period chosen.

ACKNOWLEDGMENTS

To CNPq (contract 300585/94.2.NV) and FAPESP (contract 93/2715-1), for supporting this work; to Alexandre Guirland Nowosad and Dr. Daniel Nordemann for their help in the review of this text; to Empresa Brasileira de Pesquisa Agropecuária and Departamento Nacional de Água e Energia Elétrica, for supplying the data of the levels of Paraguay River.

REFERENCES

CASDAGLI, M. Nonlinear Prediction of Chaotic Time Series. **Physica D: Nonlinear Phenomena**, North Holland, v.35, p.335-356, 1989.

GALLANT, A.R.; WHITE, H. On learning the derivatives of an unknown mapping with multilayer feedforward networks. **Neural Networks**, v.5, p.129-138, 1992.

GERSHENFELD, N. A.; WEIGEND, A.S. The future of Time Series: Learning and Understanding. In: WEIGEND, A.S.; GERSHENFELD, N.A. (Eds.). **Time Series Prediction**: forecasting the future and understanding the past. Reading, MA: Addison--Wesley, 1993. p.1-70.

HECHT-NIELSEN, R. **Neurocomputing**. Reading, MA: Addison-Wesley, 1990. 433p.

KIM, J. H.; STRINGER, J. **Applied chaos**. New York: John Wiley & Sons, 1992. 546p.

LAPEDES, A.; FARBER R. **Nonlinear signal processing using neural networks**: prediction and signal modeling. Los Alamos: Los Alamos National Laboratory, 1987. Technical Report LA-UR-87-2662.

LI, W.G.; SÁ, L.D.A.; PRASAD, G.S.S.D.; NOWOSAD, A.G; BOLZAN, M.J.A.; CHIANG, E.S.M. Neural networks adaptive wavelets for prediction of the northeastern Brazil monthly rainfall anomalies time series. In: ROGERS, S.K.; RUCK, D.W. (Eds.). **Applications and Science of Artificial Neural Networks II**. Bellingham: SPIE, 1996. p.175-187. (SPIE Proceedings Series, 2760.)

MARZBAN, C.; STUMPF, G.J. A neural network for tornado prediction based on Doppler-Derived attributes, **Journal of Applied Meteorology**, v.35, p.617-626, May 1996.

MARENGO, J.A. Variations and change in South American streamflow. **Climatic Change**, Dordrecht, v.31, n.1, p.99-117, Sept. 1995.

PAPOULIS, A. **Probability and Statistics**. Englewood Cliffs, NJ: Prentice Hall, 1990. 583p.

RICHEY, J.E.; NOBRE, C.A.; DESER, C. Amazon River discharge and climate variability: 1903 to 1985. **Science**, Washington, v.246, n.4926, p.101-103, Oct.1989.

ROBOCK, A.; FREE, M.P. Ice cores as an index of global volcanism from 1850 to the present. **Journal of Geophysical Research**, Washington, v.100, n.D6, p.11549-11568, June 1995.

ROBOCK, A.; MAO, J. The volcanic signal in surface temperature observations. **Journal of Climate**, Boston, v.8, n.5, p.1086-1103, May 1995.

RUMELHART, D. E.; MCCLELLAND, J. L. **Parallel distributed processing**: explorations in the microstructure of cognition. Cambridge: MIT Press, 1986. v.1, 548p.

TRENBERTH, K.E. Persistence of daily geopotential heights over the Southern hemisphere. **Monthly Weather Review**, Boston, v.113, n.1, p.38-53, Jan. 1985.

WELSTEAD, S.T., **Neural network and fuzzy logic applications in C/C^{++}.** New York: John Wiley & Sons, 1994. 494p.

ZELL, A.; MAMIER, G.; MACHE, M.V.N.M.; HUBNER,R.; HERRMANN, S.D.K.U.; SOYEZ, T.; SOMMER M.S.T.S.; HATZIGEORGIOU, A.; SCHREINER, D.P.T.; KETT, B.; CLEMENTE, G. **Stuttgart Neural Network Simulator - SNNS:** User Manual, version 4.0. Stuttgart: University of Stuttgart, 1995. (Report Nr. 6/95).