# Water Inflow Forecasting using the Echo State Network: a Brazilian Case Study

Rodrigo Sacchi, Mustafa C. Ozturk, José C. Príncipe, Adriano A. F. M. Carneiro and Ivan N. da Silva

*Abstract* – **A type of recurrent neural network has been proposed by H. Jaeger. This model, called Echo State Network (ESN), possesses a highly interconnected and recurrent topology of nonlinear processing elements, which constitutes a "reservoir of rich dynamics" and contains information about the history of input or/and output patterns. The interesting property of ESN is that only the memoryless readout is trained, whereas the recurrent topology has fixed connection weights. This reduces the complexity of recurrent neural network training to simple linear regression while preserving a recurrent topology. In this paper, the ESN is used to forecast hydropower plant reservoir water inflow, which is a fundamental information to the hydrothermal power system operation planning. A database of average monthly water inflows of Furnas plant, one of the Brazilian hydropower plants, was used as source of training and test data. The performance of the ESN is compared with SONARX network, RBF network and ANFIS model. The results show that the Echo State Network provides pretty good results for one-step ahead water inflow forecasting, providing a valuable information for the system operator.**

## I. INTRODUCTION

THIS paper presents and applies a constructive learning algorithm for recurrent neural networks, which modifies only the weights to output units in order to achieve the learning task.

Under certain conditions, the activation state $\mathbf{x}(n)$ of a recurrent neural network (RNN) is a function of the input history $\mathbf{u}(n)$; $\mathbf{u}(n\text{-}1)$; … presented to the network, i.e., there exists a function $\mathbf{E}$ such that $\mathbf{x}(n) = \mathbf{E}(\mathbf{u}(n); \mathbf{u}(n\text{-}1);…)$. Metaphorically, the state $\mathbf{x}(n)$ can be understood as an "echo" of the input history [1].

At [1], Jaeger investigated what can be gained when RNN states are understood as echo states, specifically, under which conditions echo states arise and describes how RNNs can be trained, exploiting echo states.

The Echo State Network (ESN) is used to forecast hydropower plant reservoir water inflow, an essential problem in electric power system operation and planning.

R. Sacchi, A. A. F. M. Carneiro and I. N. da Silva are with the Department of Electrical Engineering, Engineering School of São Carlos, University of São Paulo, São Carlos, SP 13.566-590 BRAZIL.
(e-mail: { sacchi , adriano , insilva }@sel.eesc.usp.br).
M. C. Ozturk and J. C Principe are with the Computational NeuroEngineering Laboratory, Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA.
(e-mail: { can , principe }@cnel.ufl.edu).

Long term hydrothermal operative planning is a complex problem due to various aspects of the modeling involved, including the randomness of inflows into the hydropower plants, the interconnection of hydropower plants located in a cascade, and the nonlinearity of hydro production and thermal cost functions.

A promising approach to this problem combines a deterministic optimization of the hydrothermal scheduling problem with inflows furnished by a forecasting model [2].

At each stage of the planning period this feedback control scheme determines an optimal decision based on the current forecast of future values, and this decision is utilized until a new forecast becomes available. When a new forecast becomes available, a new optimal decision is determined within the framework of partial open-loop feedback control.

## II. ECHO STATE NETWORK

Recurrent neural networks (RNN) are examples of neural-based computation models that handle time without the need for pre-processing delay lines. RNNs have recurrent connections between the processing elements (PEs) creating internally the memory required to store the history of the input patterns. RNNs have been widely used in many applications such as system identification and control of dynamical systems [3]-[5].

The main problem with the RNNs is the difficulty to adapt the system weights. The training algorithms for RNNs usually suffer from a variety of problems: computational complexity resulting in slow training, complex performances surfaces, the possibility of instability, and the decay of gradients through the topology and time [6].

Recently, a new utilization of recurrent network topologies that aims at addressing the difficulties in RNN training has been introduced by Jaeger under the name of echo state networks (ESN) [1][7].

The ESN idea divides the RNN architecture into two separate parts: a recurrent topology of nonlinear PEs, called the dynamical reservoir, and a memoryless linear network, called the readout. The state of the dynamical reservoir is called "echo states" and they contain information about the history of input patterns. The outputs of the internal PEs are fed to a readout network that reads the reservoir and produces the network output. The interesting property of ESNs is that only the memoryless linear readout is trained whereas the recurrent topology has fixed connections.

Consider the recurrent discrete-time neural network given

in Figure 1 with *M* input units, *N* internal PEs and *L* output units. The value of the input unit at time *n* is u(*n*), of internal units are x(*n*), and of output units are y(n).
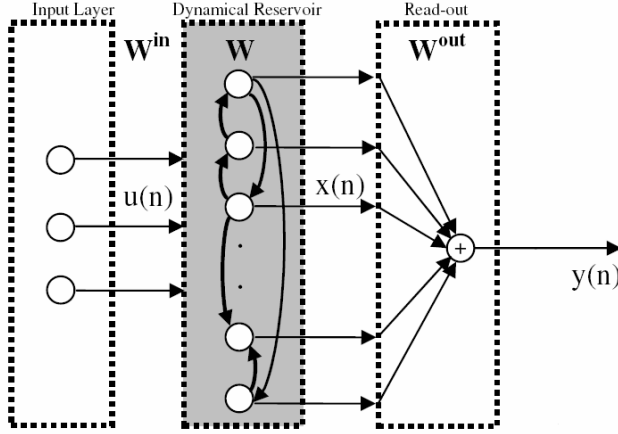


Fig. 1 - Diagram of an echo state network (ESN)

The activation of the internal PEs (echo states) is updated according to:

$$\mathbf{x}(n+1) = \mathbf{f}(\mathbf{W^{in}u}(n+1) + \mathbf{Wx}(n)) \qquad (1)$$

where $\mathbf{x}(n)$ is the echo state vector, $\mathbf{W}$ is the recurrent connection weight matrix, $\mathbf{u}(n)$ is the input signal, $\mathbf{W^{in}}$ are the weights between the input and the internal PEs. Here, $\mathbf{f}$ represents the activation function of the internal units, which is usually a hyperbolic tangent function.

The echo state condition is defined in terms of the spectral radius of the reservoir's weight matrix ($|\mathbf{W}|<1$), which relates to the condition under which the recurrent network states are strongly coupled with the input history. Although this condition is very useful in terms of defining the region of parameters resulting in echo states, it does not specify a sufficiently accurate design principle to construct ESNs for function approximation.

The output of the ESN is computed according to:

$$\mathbf{y}(n) = \mathbf{W}^{out}\mathbf{x}(n), \qquad (2)$$

For a given input signal, the echo states can be computed using equation (1). The optimal output weight matrix, $\mathbf{W}^{linear}$ in the mean square error (MSE) sense can be analytically computed by:

$$W^{out} = E[\mathbf{xx}^T]^{-1} E[\mathbf{xd}] \cong (\frac{1}{P}\sum_n \mathbf{x}(n)\mathbf{x}(n)^T)^{-1}(\frac{1}{P}\sum_n \mathbf{x}(n)\mathbf{d}(n)).$$
$$(3)$$

here $E[.]$, $P$, $\mathbf{x}$ and $\mathbf{d}$ denote the expected value operator, the number of input data, the input and the desired target signals, respectively.

For the ESN experiment we submit in this paper, the input weight matrix $\mathbf{W^{in}}$ was fully connected and took on values of $\pm 1$ with equal probability. The recurrent connection matrix $\mathbf{W}$ generated from a zero mean random Gaussian distribution. Output feedback and the direct connection from inputs to the output mapper were disconnected.

## III. THE CASE STUDY: FORECASTING WATER INFLOW

Analysis and forecast of water inflow series are of utmost importance in the operation planning of hydropower systems. One of the greatest difficulties to forecast is the nonstationary nature of the streamflow series due to wet and dry periods of the year.

Forecasting models based on Box & Jenkins methodology has been largely used in streamflow forecasting problems [8],[9]. However, those parametric auto-regressive models assume linear relationship between the series values [10].

In the last years, Artificial Neural Networks (ANN) have been suggested for time series analysis due to their ability to deal with nonlinear input-output relationships [6],[11]. MultiLayer Perceptrons (MLP) with a backpropagation algorithm, the most popular of them, have been applied to forecast streamflow with promising results [12]. Afterward, other types of ANN have also been used for this problem with the same success [13]-[16].

In this problem, the use of ANN usually require considerable modeling effort, since they are dealing with a very difficult real world problem. Consequently, it implies in a task to estimate a great number of parameters. Moreover, another great problem in forecasting applications is the definition of the "optimal" embedding that defines the most important previous measured samples which must be considered as input data of the predictor.

Therefore, the ESN appears to deal with those characteristics. This is because one of interesting property of ESN is that only the memoryless readout is trained, whereas the recurrent topology has fixed connection weights. This reduces the complexity of recurrent neural network training to simple linear regression while preserving a recurrent topology which is proven to be very powerful in neural networks literature [1][7]. Moreover, since the system has internal built-in memory resulting from the feedback connections, it is not necessary to embed the input signal before further processing. The training of the linear readout weights can be done analytically; hence very fast.

In this paper, the ESN works as a one-step-ahead predictor of the monthly average hydropower plant reservoir water inflow. The performance of this approach is evaluated and compared to the Self-Organizing Nonlinear Auto-Regressive model with eXogenous input (SONARX) model [17], the Radial Basis Function (SONARX-RBF) network [16] and the Adaptive Neuro-Fuzzy Inference System (ANFIS) model [18].

*A. Results*

In this case study, the Furnas hydroelectric power plant (HPP) was chosen because of its relative position on the cascade, an upstream reservoir, consequently, without whatever operative influence.

A monthly water inflow historic data, covering the period of 1931-1994, have been available to this study. A stretch of 5 (five) years was selected as testing data, which corresponds to the period of 1972-1976. Then, the whole historic period without theses testing samples, 1972-1976, was taken as training data.

Usually, most of the predictor models, including the models that ESN is compared with, require the previous choice of the "optimal" embedding that defines the most important previous measured samples which must be considered as input data.

Here, we chose for all predictor models, except for ESN, 3 (three) options of embedding. The first option of embedding was chosen as a short time delay memory, defined as:

$$x_1(t) = wi(t)$$
$$x_2(t) = wi(t-1)$$
$$x_3(t) = wi(t-2)$$

where $t = 1,...,T$ denotes the current month, $wi$ the water inflow sample, and $x_i$ the input variable to the predictor model.

The second option of embedding was chosen as a longer time delay memory, which probably contains more useful information; defined as:

$$x_1(t) = wi(t)$$
$$x_2(t) = wi(t-1)$$
$$x_3(t) = wi(t-2)$$
$$x_4(t) = wi(t-3)$$
$$x_5(t) = wi(t-4)$$
$$x_6(t) = wi(t-5)$$

And, the third one was chosen with two short periods, one within the immediate past and another related to the previous year. This was a way to fid the models not only with the monthly information but also with the year season information. It was, then, defined as:

$$x_1(t) = wi(t)$$
$$x_2(t) = wi(t-1)$$
$$x_3(t) = wi(t-2)$$
$$x_4(t) = wi(t-10)$$
$$x_5(t) = wi(t-11)$$
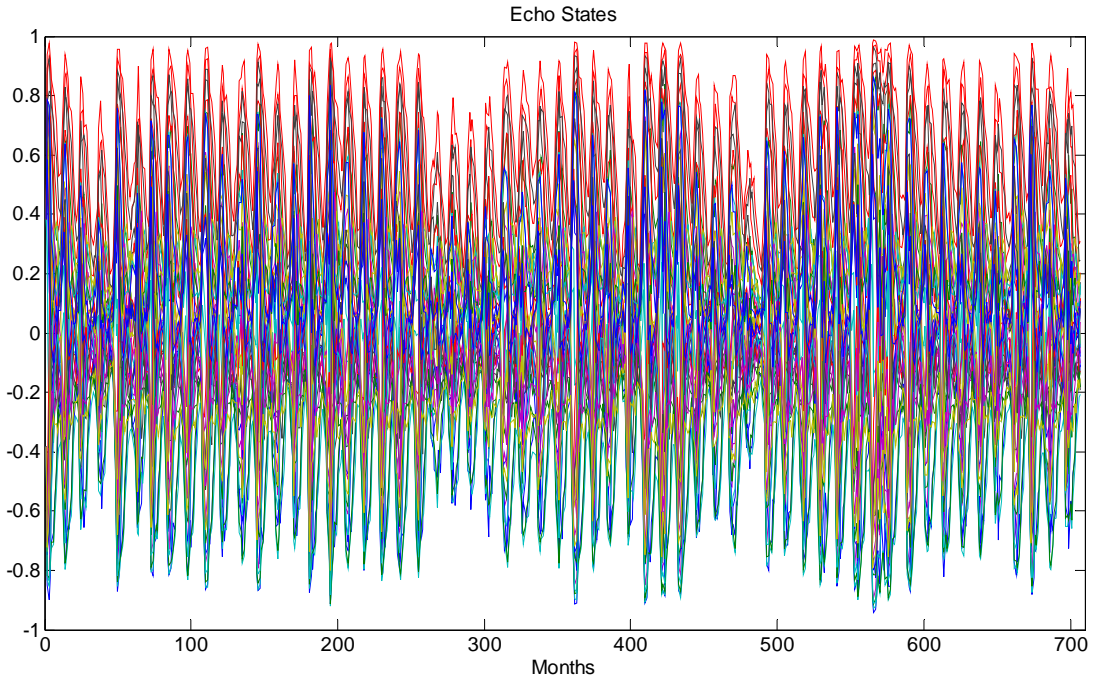$$x_6(t) = wi(t-12)$$



Fig. 2 – "Echo States" of the ESN model trained with monthly water inflow historic samples
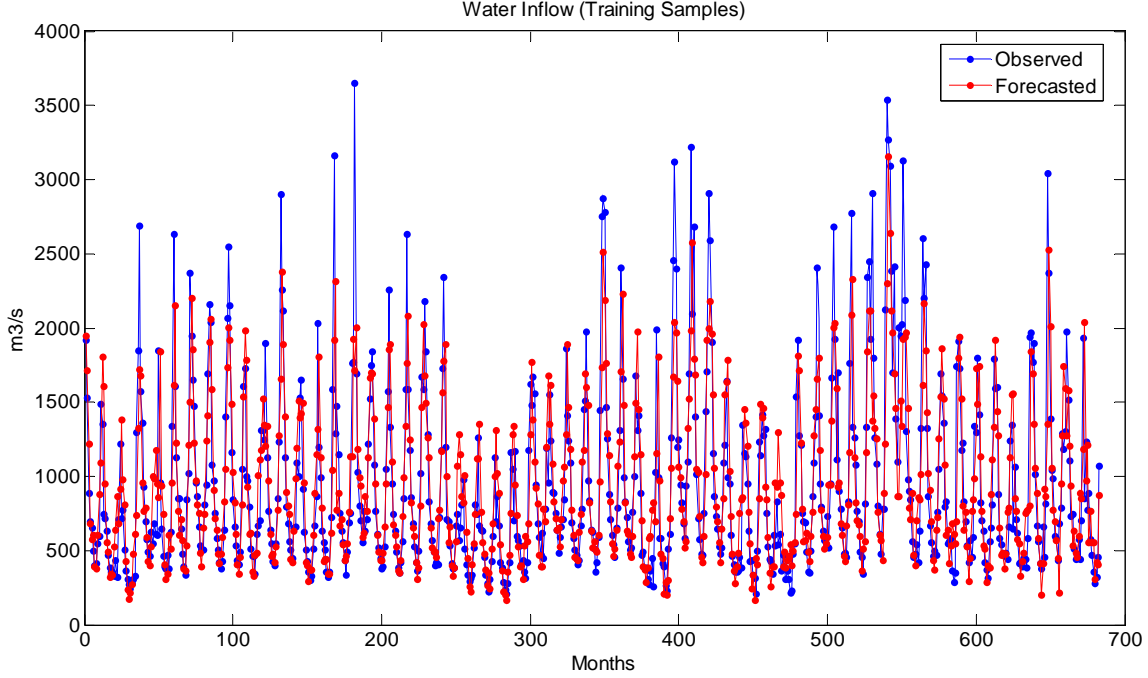
Fig. 3 – Training Patterns: Water Inflow of Furnas Hydroelectric Power Plant

Nevertheless, as mentioned before, the ESN doesn't need to use any of those options of embedding, since ESN has internal built-in memory resulting from the feedback connections.

In addition, as the objective, here, was to obtain a one-step-ahead forecasting, the desired output of all models was defined as:

$$y_d(t) = wi(t+1) \qquad (4)$$

Finally, the predictor models used here could be trained and adjusted: SONARX, SONARX-RBF, ANFIS and Echo State Network.

During both the training and testing periods of the ESN, the entries of **W** (the recurrent connection weight matrix) were scaled to have a spectral radius of 0.8, so that the outputs of echo states could be almost uniformly distributed within their dynamic range.

The dynamic reservoir was initialized empty, that is, with no activation. Considering that after some steps it achieves its steady-state, we have disregarded those first responses, avoiding the initialization contamination.

After some experimental evaluations, the number of hidden layer processing elements was defined as $N = 36$, which presented best results.

Figure 2 shows the resulted "Echo States" from the ESN training with the training data. Each line (color) represents the behavior of the activation level of each PE. We note that ESN states have a wide dynamic range in this problem, when it would be necessary to cover the feature space during the whole training period.

Those variety ESN states represent a set of functional bases constructed dynamically by the input, while the readout simply projects the desired response onto this representation space.

We can see, at Figure 3, the result of the water inflow recovering for the training data. The blue line refers to the observed measures of this time series, and the red one refers to the "forecasted" amount of water inflow.

Based on Figure 3, we can say the predictor model was able to learn and capture most of the behavior variability present on this time series, which seems but is not a easy task for any predictor.

Figure 4 shows the water inflow forecasting to the period of 1972-1976. The ESN predictor presented a great result, where, we can see how similar was both time series. The red line, representing water inflow forecasting, shows a behavior very close to the blue one, observed measures.

Even at the last year, 1976, we can see the ESN predictor was able to forecast a higher, wetter dry season at the middle of that year.

The performance of the forecasting models was evaluated according to 4 (four) error criteria: Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAD), and Mean Percentual Error (MPE), as follows:

$$MSE = \frac{1}{T} \sum_{t=1}^{T} (y(t) - y_d(t))^2 \qquad (5)$$

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y(t)-y_d(t))^2} \qquad (6)$$

$$MAD = \frac{1}{T}\sum_{t=1}^{T}\left|y(t)-y_d(t)\right| \qquad (7)$$

$$MPE = \frac{100}{T}\sum_{t=1}^{T}\frac{\left|y(t)-y_d(t)\right|}{y_d(t)} \qquad (8)$$

where, $T$ is the number of monthly samples, $y(t)$ is the model output, and $y_d(t)$ the desired output.

Table I shows the forecasting errors obtained by the predictors used and compared in this paper: Echo State Network, SONARX, SONARX-RBF and ANFIS models. We can see that ESN, SONARX-RBF and ANFIS performed significantly better than SONARX model, which has an intrinsic error since it is based on Vector-Quantized Temporal Associative Memory (VQTAM). In addition, we note that SONARX-RBF and ANFIS models presented practically the same performance.

Moreover, the ESN model performed slightly better than all other models, even though it presents a significantly simpler, and faster, training algorithm. This fact prove how powerful and promising is this approach.
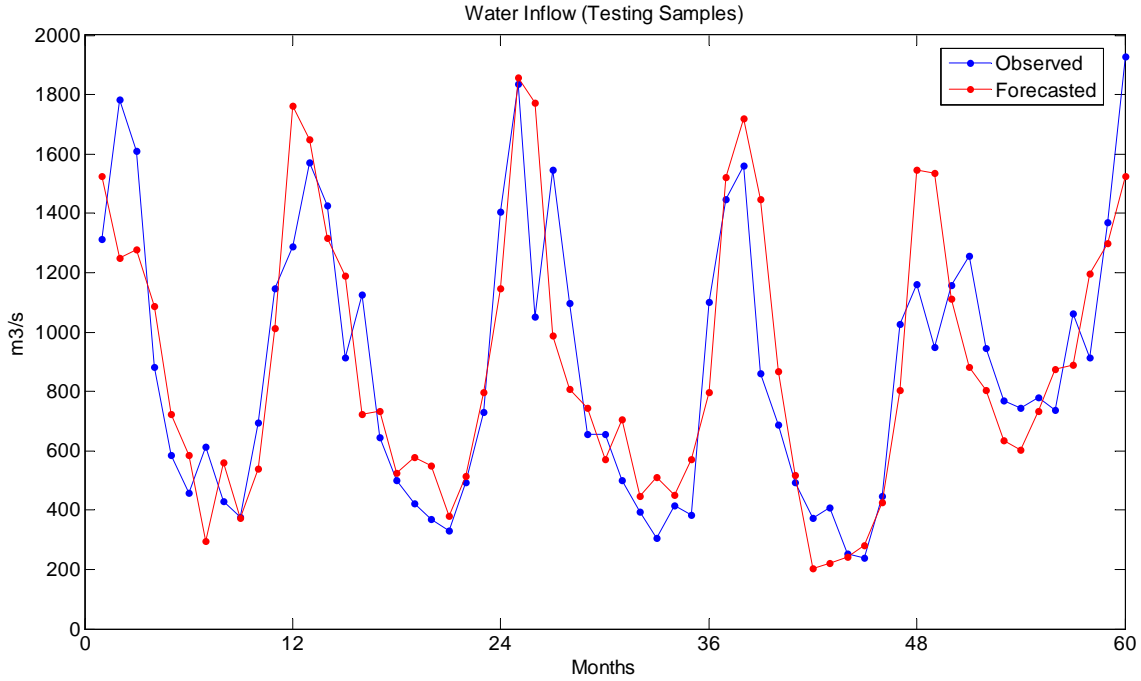


Fig. 4 – Testing Patterns (1972 - 1976): Water Inflow of Furnas Hydroelectric Power Plant

Table I
Forecasting Errors for Furnas Hydroelectric Power Plant

| Models | Embedding | MSE $(x10^4)$ | RMSE | MAD | MPE (%) |
|---|---|---|---|---|---|
| Echo State Network | built-in | 5.97 | 244.45 | 177.41 | 21.04 |
| SONARX | t,t-1,t-2 | 21.78 | 466.73 | 320.13 | 37.94 |
| | t,t-1,t-2,t-3,t-4,t-5 | 12.20 | 349.30 | 253.18 | 28.78 |
| | t,t-1,t-2,t-10,t-11,t-12 | 13.54 | 367.90 | 261.04 | 28.32 |
| SONARX-RBF | t,t-1,t-2 | 9.24 | 303.94 | 233.24 | 29.35 |
| | t,t-1,t-2,t-3,t-4,t-5 | 6.35 | 252.02 | 187.53 | 22.41 |
| | t,t-1,t-2,t-10,t-11,t-12 | 7.13 | 266.94 | 202.56 | 24.35 |
| ANFIS | t,t-1,t-2 | 9.45 | 307.47 | 235.67 | 29.89 |
| | t,t-1,t-2,t-3,t-4,t-5 | 7.07 | 265.97 | 198.52 | 24.27 |
| | t,t-1,t-2,t-10,t-11,t-12 | 6.72 | 259.25 | 205.99 | 25.02 |

## IV. Conclusions

This paper presents a type of recurrent neural network called Echo State Network (ESN), which possesses a highly interconnected and recurrent topology of nonlinear processing elements.

The ESN has two interesting properties: one is that only the memoryless readout is trained, whereas the recurrent topology has fixed connection weights; and another one is that ESN has internal built-in memory resulting from the feedback connections, it is not necessary to embed the input signal before further processing.

Here, the ESN is used to forecast hydropower plant reservoir water inflow, fundamental information to the hydrothermal power system operation planning. A database of average monthly water inflows of Furnas plant, one of the Brazilian hydropower plants, was used as source of training and test data.

The results of the ESN forecasting were compared with SONARX network, RBF network and ANFIS model. They showed that the ESN performed significantly better than SONARX, and slightly better than SONARX-RBF and ANFIS, even though it presents a significantly simpler, and faster, training algorithm.

Further work shall consider any type of preprocessing of the data. It could transform the original time series in a simpler and easier one, so that we could achieve even lower forecasting errors.

## Acknowledgment

## References

[1] H. Jaeger, "The echo state approach to analyzing and training recurrent neural networks". Technical Report GMD Report 148, *German National Research Center for Information Technology*, 2001.

[2] L. Martinez and S. Soares, "Comparison Between Closed-Loop and Partial Open-Loop Feedback Control Policies in Long Term Hydrothermal Scheduling", *IEEE Transactions on Power Systems*, vol. 17, no. 2, may 2002

[3] G. Kechriotis, E. Zervas, and E. S. Manolakos, "Using recurrent neural networks for adaptive communication channel equalization", *IEEE Transactions on Neural Networks*, 5(2):267–278, 1994.

[4] G. V. Puskorius and L. A. Feldkamp, "Neurocontrol of nonlinear dynamical systems with kalman filter trained recurrent networks". *IEEE Transactions on Neural Networks*, 5(2):279–297, 1994.

[5] A. Delgado, C. Kambhampati and K. Warwick, "Dynamic recurrent neural network for system identification and control". *IEEE Proceedings of Control Theory and Applications*, 142(4):307–314, 1995.

[6] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall, NJ, second edition, 1998.

[7] H. Jaeger and H. Hass, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication", *Science*, 304(5667):78–80, 2004.

[8] J. D. Salas, D. C. Boes, and R. A. Smith, "Estimation for arma models with seasonal parameters," *Water Resources Res.*, vol. 18, no. 4, pp. 1006–1010, 1982.

[9] A. I. MecLeod, "Diagnostic checking of periodic autoregression," *J. Time Series Anal.*, vol. 15, no. 2, pp. 221–223, 1994.

[10] G. E. P. Box, G. Jenkins, and G. C. Reinsel, *Time Series Analysis, Forecasting, and Control*. San Francisco, CA: Prentice-Hall, 1994.

[11] A. S. Weigend, D. E. Rumelhart, and B. A. Huberman, "Generalization by weight-elimination applied to currency exchange rate prediction," in *Proc. IEEE Int. Joint Conf. Neural Networks*, vol. I, pp. 837–841, 1991.

[12] A. F. Atiya, S. M. El-Shoura, S. I. Shaheen, and M. S. El-Sherif, "A comparison between neural-network forecasting techniques – Case study: River flow forecasting," *IEEE Tran. Neural Networks*, vol. 10, pp. 402–409, Mar. 1999.

[13] P. Coulibaly and F. Anctil, "Real-Time Short-Term Natural Water Inflows Forecasting Using Recurrent Neural Networks", *International Joint Conference on Neural Networks - IJCNN '99*, 1999.

[14] M. Figueiredo, R. Ballini, S. Soares, M. Andrade, and F. Gomide, "Learning Algorithms for a Class of Neurofuzzy Network and Application", *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews*, vol. 34, no. 3, Aug. 2004.

[15] M. Valença, T. Lurdermir and A. Valença, "River Flow Forecasting for Reservoir management through Neural Networks", *Proceedings of the Fifth International Conference on Hybrid Intelligent Systems (HIS'05)*, 2005.

[16] R. Sacchi, A. A. F. M. Carneiro and A. F. R. Araújo, "A RBF network trained by the SONARX model and applied to obtain the operation policies of the hydropower systems", *Brazilian Symposium on Neural Networks - SBRN*, Brazil, 2004.

[17] G. A. Barreto and A. F. R. Araújo, "A self-organizing NARX network and its application to prediction of chaotic time series". *International Joint Conference on Neural Networks - IJCNN '01*, Washington D.C., USA, 2001.

[18] J.-S. R. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems." *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no03: pp. 665-685, 1993.