ISSN 2176-5480

22nd International Congress of Mechanical Engineering (COBEM 2013)
November 3-7, 2013, Ribeirão Preto, SP, Brazil

TIME SERIES ANALYSIS OF A RESERVOIR HYDROLOGIC BALANCE

Flavio M. Viola, fmviola@gmail.com
Marcelo A. Savi, savi@mecanica.ufrj.br
Universidade Federal do Rio de Janeiro/COPPE – Department of Mechanical Engineering
21.941.972 – Rio de Janeiro – RJ, P.O. Box 68.503 – Brazil
Antonio C.P. Brasil Jr, brasiljr@unb.br
Universidade de Brasília – Faculdade de Tecnologia
70.910.000 – Brasília – DF - Brazil

Abstract. *Hydroelectric power plants are essential in Brazilian energy matrix being considered one of the cleaner electric generator processes. The water balance of a reservoir can be obtained through an understanding of the behavior of its main hydrological variables: inflow, released through the spillway, turbine discharge and the losses (evaporation and percolation). In general, these variables are measured and stored in time series. Through analysis of these series it is possible to design a model, allowing the predictions of the future dynamics of the reservoir. These series, in the referring to its dynamic, present stationary and non-stationary behavior, being necessary to use specific techniques to analyze them. This article presents an analysis of hydrological time series evaluating predictions for future behavior. Nonlinear tools are employed with this aim establishing state space reconstruction and time series prediction. The method of delay coordinates is employed for the state space reconstruction and the delay parameters are evaluated using the method of average mutual information and the method of false nearest neighbors. The simple nonlinear prediction is employed to model the time series evaluating the prediction of future values. This approach is verified considering known parts of the time series from Tucuruí hydrologic reservoir and afterwards, results are extrapolated for future values.*

Keywords: *Hydrological modeling, time series analysis, nonlinear dynamics, simple nonlinear prediction*

1. INTRODUCTION

Hydroelectric plants are essentials in Brazilian energy matrix being considered one of the cleaner electric generator processes. The water balance of a reservoir can be obtained through an understanding of the behavior of its main hydrological variables: inflow, released through the spillway, turbine discharge and the losses (evaporation and percolation). In general, these variables are measured and stored in time series.

Several research efforts have been dedicated to the study of hydrological models. Different methods have been applied allowing the predictions of future values. Among these methods, it is important to highlight (Ashu & Avadhnam, 2007): autoregressive methods (AR), auto-regressive moving average method (ARMA), auto-regressive integrated moving average method (ARIMA), autoregressive moving average with exogenous inputs method (ARMAX). Moreover, Karunasinghe & Liong (2006) used nonlinear prediction through artificial neural network and Cheng *et al.* (2008) combined dynamic interpolation into multilayer adaptive time-delay neural network for long-term hydrologic prediction. Hong (2012) presented a brief review of other hydrological model methods.

This paper deals with the time series analysis related to the hydrologic reservoir modeling. Hydrological variables are elected to be representative of the system dynamics and time series analysis is applied to predict future values of this variable. The method of delay coordinates is employed for the state space reconstruction and the delay parameters are evaluated using the method of average mutual information and the method of false nearest neighbors. The simple nonlinear prediction is employed evaluating the prediction of future values. This approach is verified considering known parts of the time series and afterwards, results are extrapolated for future values. Nonlinear time series analysis employs the TISEAN package (Hegger *et al.*, 1999).

2. RESERVOIR MODEL

Tucuruí is a hydraulic power plant in the state of Pará, the Northern region of Brazil, constructed by Eletronorte S.A. It extends from $49^o20'$ W to 50^o W and from $3^o45'$ S to 5^o S. The total area at maximum water level is 2,430 km^2. The reservoir has a maximum depth of 72 m, with average stream flow of 11,000 m^3/s (Deus *et al.*, 2013).

In order to give an idea concerning the Tucuruí reservoir behavior, five different time series are presented: affluent, turbine discharge, spillway discharge, reservoir level and downstream, from 1985 to 2011. The Eletronorte S.A. has daily time series evolution since 1984. These series have dynamical behavior that can be classified as stationary or non-stationary, and each one of them needs a proper treatment with appropriate techniques. This article presents an analysis of hydrological time series evaluating predictions for future behavior. Figure 1 shows the five time series related to Tucuruí reservoir.

ISSN 2176-5480

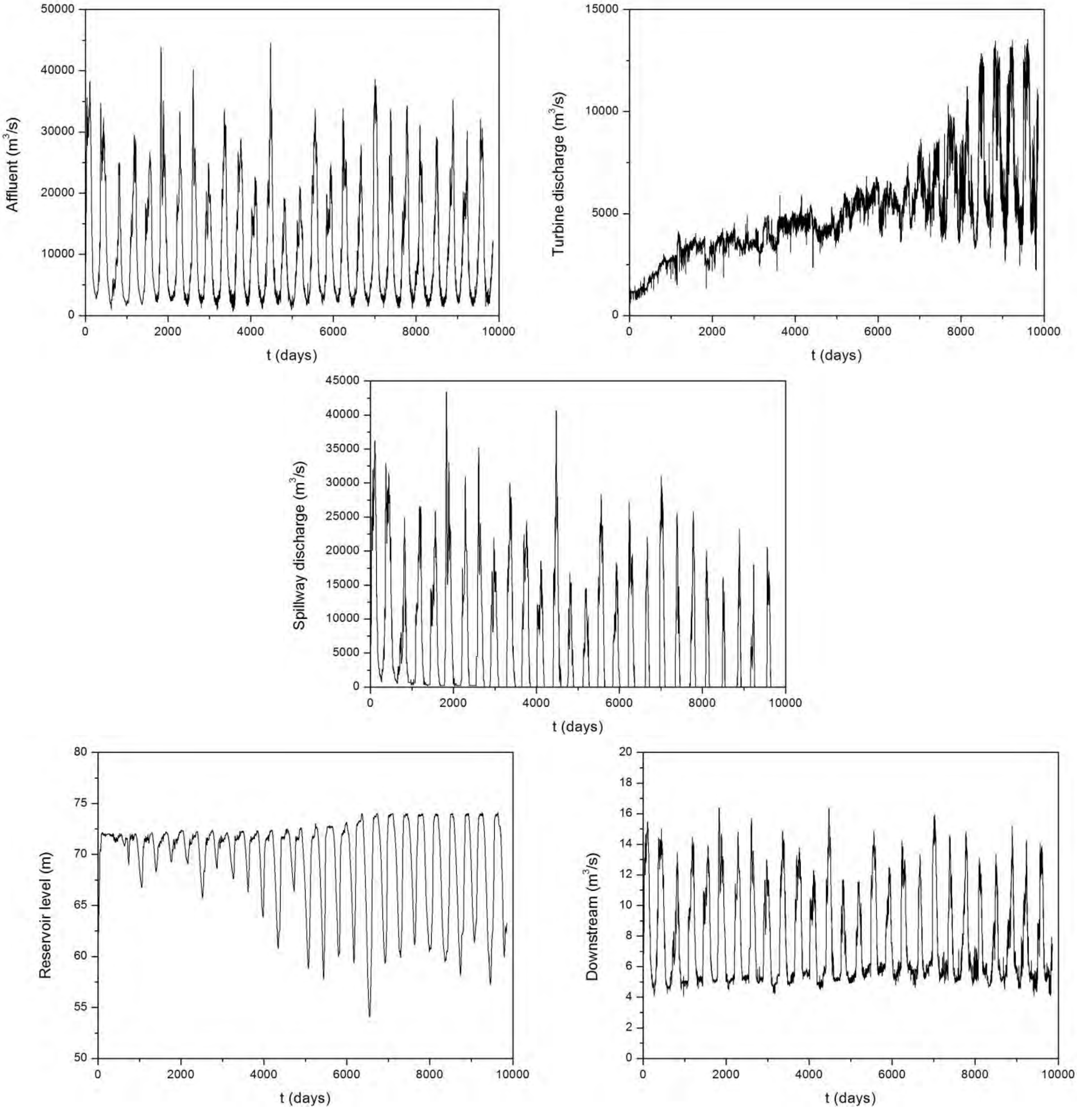


Figure 1. Hydrologic time series from Tucuruí hydroelectric reservoir (1985-2011)

3. TIME SERIES ANALYSIS

The basic idea of the state space reconstruction is that a signal contains information about unobserved state variables that can be used to predict the present state (Savi, 2006). Therefore, a scalar time series, S_n, may be used to construct a vector time series that is equivalent to the original dynamics from a topological point of view. The state space reconstruction needs to form a coordinate system to capture the structure of orbits in state space, which could be done using lagged variables, $S_{n+\tau}$, where τ is the time delay. Then, it is possible to use a collection of time delays to create a vector in a D_e-dimensional space,

$$U(t) = \{S_n, S_{n+\tau}, \ldots, S_{n+(D_e-1)\tau}\}^T \tag{1}$$

ISSN 2176-5480

The mutual information method (Fraser & Swinney, 1986) is a good alternative to evaluate the time delay, τ. The determination of embedding dimension, D_e, on the other hand, may be evaluated from the method of the false nearest neighbors (Kennel *et al.*, 1992). This reconstructed space can be used for the forecast and the simple nonlinear prediction is a good alternative for this aim. The forthcoming sections present a brief discussion of each of the employed methods.

3.1 - Method of Average Mutual Information

Fraser & Swinney (1986) establishes that the time delay τ corresponds to the first local minimum of the average mutual information function $I(\tau)$, which is defined as follows,

$$I(\tau) = \sum_{n=1}^{N-\tau} \Gamma(S_n, S_{n+\tau}) \log_2 \left[\frac{\Gamma(S_n, S_{n+\tau})}{\Gamma(S_n)\, \Gamma(S_{n+\tau})} \right] \tag{2}$$

where $\Gamma(S_n)$ is the probability of the measure S_n, $\Gamma(S_{n+\tau})$ is the probability of the measure $S_{n+\tau}$, and $\Gamma(S_n, S_{n+\tau})$ is the joint probability of the measure of S_n and $S_{n+\tau}$. When the measures S_n and $S_{n+\tau}$ are completely independent, $I(\tau) = 0$. On the other hand, when S_n and $S_{n+\tau}$ are equal, $I(\tau)$ is maximum. Therefore, the analysis of the $I(\tau)$ curve allows one to determine the best time delay to be used in the state space reconstruction.

3.2 - Method of the False Nearest Neighbors

The method of the false nearest neighbors was originally developed for determining the number of time delay coordinates needed to recreate autonomous dynamics, but it is extended to examine the problem of determining the proper embedding dimension. In an embedding dimension that is too small to unfold the attractor, not all points that lie close to one another will be neighbors because of the dynamics. Some will actually be far from each other and simply appear as neighbors because the geometric structure of the attractor has been projected down onto a smaller space (Kennel *et al.*, 1992).

In order to use the method of the false nearest neighbors, a D-dimensional space is considered where the point U_n has r-th nearest neighbors, U_n^r. The square of the Euclidean distance between these points is,

$$r_D^2(n,r) = \sum_{k=0}^{D-1} \left[S_{n+k\tau} - S_{n+k\tau}^r \right]^2 \tag{3}$$

Now, going from dimension D to D+1 by time delay, there is a new coordinate system and, as a consequence, a new distance between U_n and U_n^r. When these distances change from one dimension to another, these are false neighbors. A proper space dimension may be obtained when there are no false neighbors after a dimension increase.

3.3 – Prediction

Prediction is a particular application related to system modeling that has the objective of estimating future values from a known time series, called past, $S_n, n = 1,\ldots,N$. Therefore, it is necessary to estimate future time series, employing some prediction technique that results in an estimated series: $P_{N+1}, P_{N+2}, \ldots, P_{N+p}$. Figure 2 shows a schematic plot related to the prediction problem. A verification procedure can be performed using known parts of the series and establishing a comparison between estimated values with future values associated with the original series in order to establish prediction accuracy (Viola *et al.*, 2010).

ISSN 2176-5480

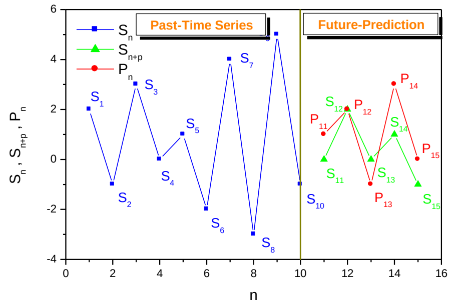


Figure 2. Time series prediction.

In general, techniques for time series prediction may be classified in linear and nonlinear methods. Other classification reported in literature considers local and global methods. An overview of the main aspects related to nonlinear time series analysis and prediction is provided in the following references: Kantz & Schreiber (1997); Abarbanel (1995); Casdagli (1989); Schreiber (1999); Weigend & Gershenfeld (1994); Pinto & Savi (2003).

Simple nonlinear prediction is based on the state space reconstruction. After the reconstruction, in order to predict a time instant Δn (Δn = 1,..., p) ahead N, it is necessary to define a parameter ε that is related to the size of the neighborhood $V_\varepsilon(U_N)$ around point U_N. Therefore, for all points U_n closer than ε to U_N $(U_n \in V_\varepsilon(U_N))$ look up the individual prediction $S_{n+\Delta n}$. The prediction $P_{N+\Delta n}$ is then calculated from the average of the individual predictions $S_{n+\Delta n}$.

$$P_{N+\Delta n} = \frac{1}{|V_\varepsilon(U_N)|_{U_n \in V_\varepsilon(U_N)}} \sum S_{n+\Delta n} \tag{4}$$

where $|V_\varepsilon(U_N)|$ denotes the number of elements of the neighborhood $V_\varepsilon(U_N)$. Figure 3 presents a schematic representation of the simple nonlinear prediction applied to a time series with 10 elements and D_e = 2. For a parameter ε, points U_2, U_4, U_5, U_7 and U_8 are inside the neighborhood and hence, the first prediction, P_{11}, is evaluated from the average of these values.

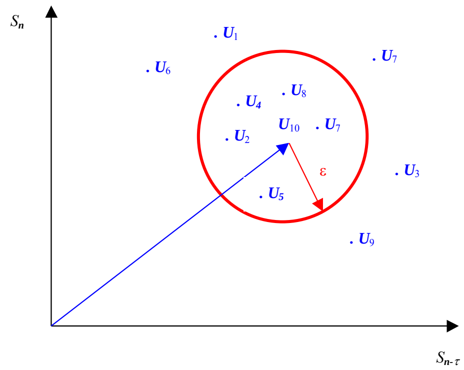


Figure 3. Simple nonlinear prediction.

4. PREDICTION ANALYSIS OF THE HYDROLOGIC BALANCE OF THE RESERVOIR

Hydrologic time series measured in Tucuruí hydroelectric is used for the analysis developed in this work. This series has 9,861 data points corresponding to 27 years (1985 to 2011). Initially, a verification procedure is carried out considering two different situations defined by distinct parts of the series: 1985 to 2006 (22 years) performing the prediction from 2007 to 2011 (5 years) and 2005 to 2009 (5 years) performing the prediction from 2010 to 2011 (2 years). These choices are made considering the number of data points and because the turbine discharge series has a non-stationary behavior.

In order to establish the model verification, predicted results are compared with time series and two errors are defined: the average error and the daily error (Viola *et al.*, 2010). The average error is defined as follows:

$$\bar{\mathrm{E}} = \frac{|\bar{S} - \bar{P}|}{\bar{S}} \tag{5}$$

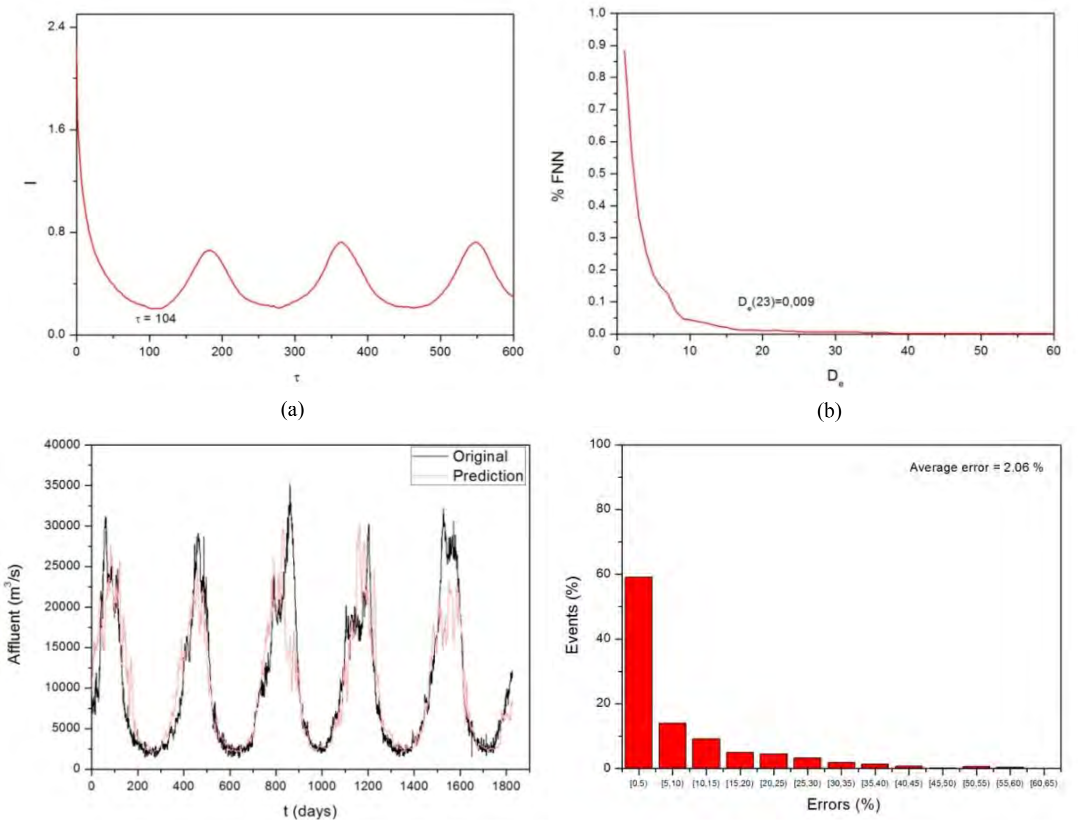
ISSN 2176-5480

where $\overline{S}$ is the average of the time series and $\overline{P}$ is the average of the prediction evaluated during the same period. On the other hand, the daily error is defined by the expression:

$$E_n^D = \frac{|S_n - P_n|}{(H_{\max} - H_{\min})} \tag{6}$$

where $H_{\max}$ and $H_{\min}$ are, respectively, the maximum and minimum of the time series.

Initially, a time series corresponding to 22 years corresponding to 8,035 data points (from 1985 to 2006) is of concern. The analysis starts by evaluating delay parameters for affluent time series. Figure 4 (upper part) presents average mutual information and false nearest neighbors analysis. From these, it is possible to conclude that time delay is τ = 104 and embedding dimension is D_e = 23. Afterwards, simple nonlinear prediction is employed to model the series, predicting future values from 2007 to 2011 (5 years). Figure 4 (lower part) presents the original time series together with the prediction made by the simple nonlinear prediction and an error histogram that shows the distribution of events related to daily error between the series and the prediction. Results show a good agreement between the original and the predicted series and it is important to notice that both series have average values that are close (respectively, 10,297.02 m^3/s and 10,084.03 m^3/s representing a difference of 2.06 %).



(a) (b) (c) (d)

Figure 4. Affluent prediction (time series from 1985 to 2006). (a) Average mutual information; (b) embedding dimension; (c) comparison between original time series and prediction; (d) error analysis.

The same analysis is applied to the other time series (spillway discharge, reservoir level, downstream and turbine discharge), and results are showed in Figures 5-8, respectively. Figure 5 presents results to spillway discharge time series. It is possible to conclude that time delay is τ = 96 and embedding dimension is D_e = 57. Results show a good agreement in the distribution of events related to daily error between the series and the prediction. Nevertheless, the average value for the time series is 2,486.28 m^3/s while the prediction has a value of 4,794.99 m^3/s, representing a difference of 92.85 %. This discrepancy should be better investigated. Figure 6 presents results to reservoir level time series. From these, it is possible to conclude that time delay is τ = 107 and embedding dimension is D_e = 57. Results show a good agreement between the original and the predicted series and it is important to notice that both series has

ISSN 2176-5480

average values that are very close (respectively, 67.77 m and 68.07 m representing a difference of 0.45 %). Figure 7 presents the results to downstream time series. For this time series, time delay is $\tau = 99$ and embedding dimension is $D_e = 31$. Once again, results present a good agreement between the original and the predicted series presenting average values that are close (respectively, 7.26 m^3/s and 7.03 m^3/s representing a difference of 3.16 %).

Figure 8 presents results of the turbine discharge time series. The time series analysis indicates that the estimation of the time delay is difficult and the value $\tau = 172$ is not accurate. Besides, the embedding dimension is $D_e = 40$. Results do not present a good agreement between the original and the predicted series. It is important to notice that both errors (daily and average) confirm these results. The original and the predicted series have average values that are not close (respectively, 7,778.59 m3/s and 6,807.13 m3/s representing a difference of 12.49 %).

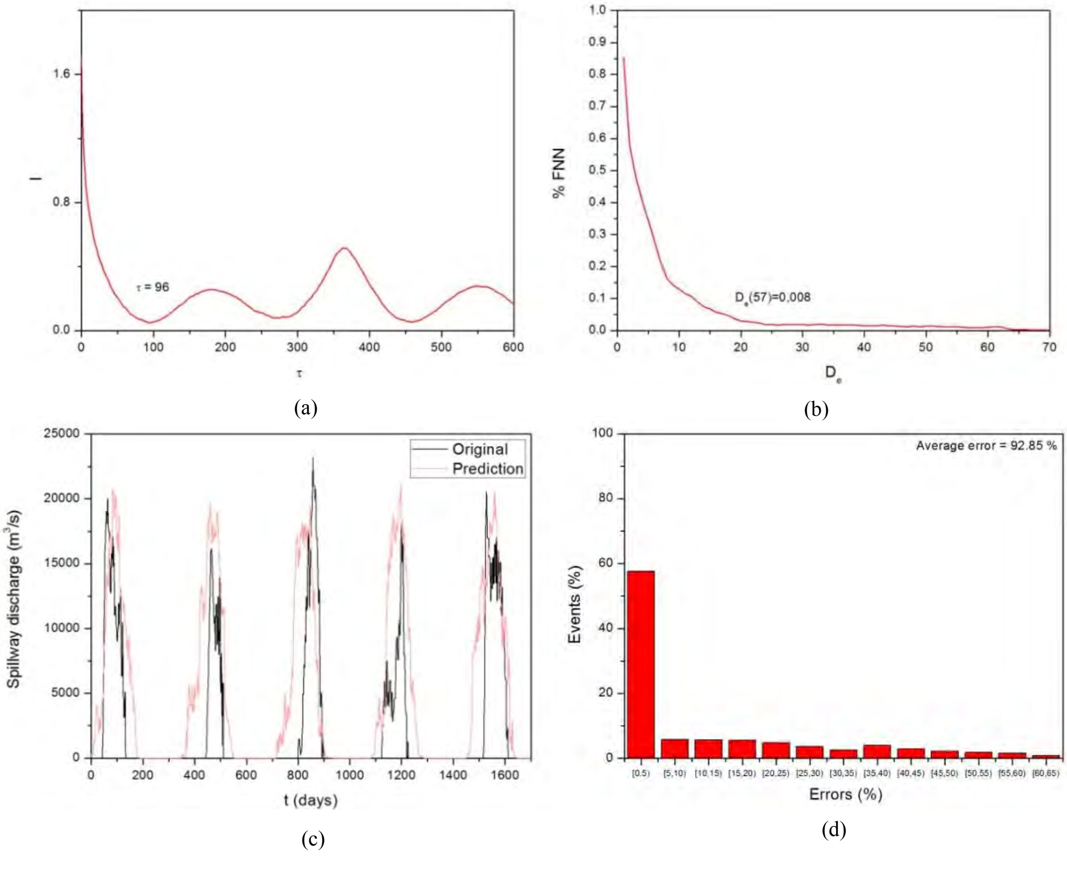


Figure 5. Spillway discharge prediction (time series from 1985 to 2006). (a) Average mutual information; (b) embedding dimension; (c) comparison between original time series and prediction; (d) error analysis.

ISSN 2176-5480

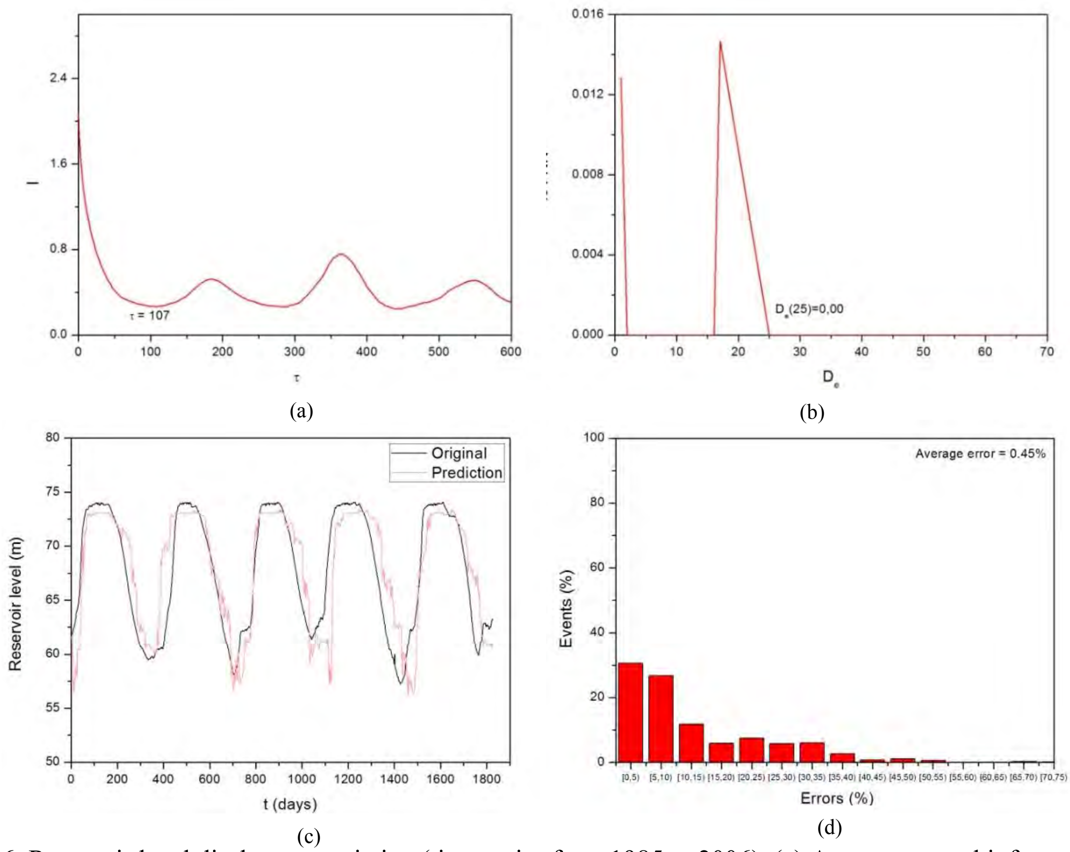


Figure 6. Reservoir level discharge prediction (time series from 1985 to 2006). (a) Average mutual information; (b) embedding dimension; (c) comparison between original time series and prediction; (d) error analysis.

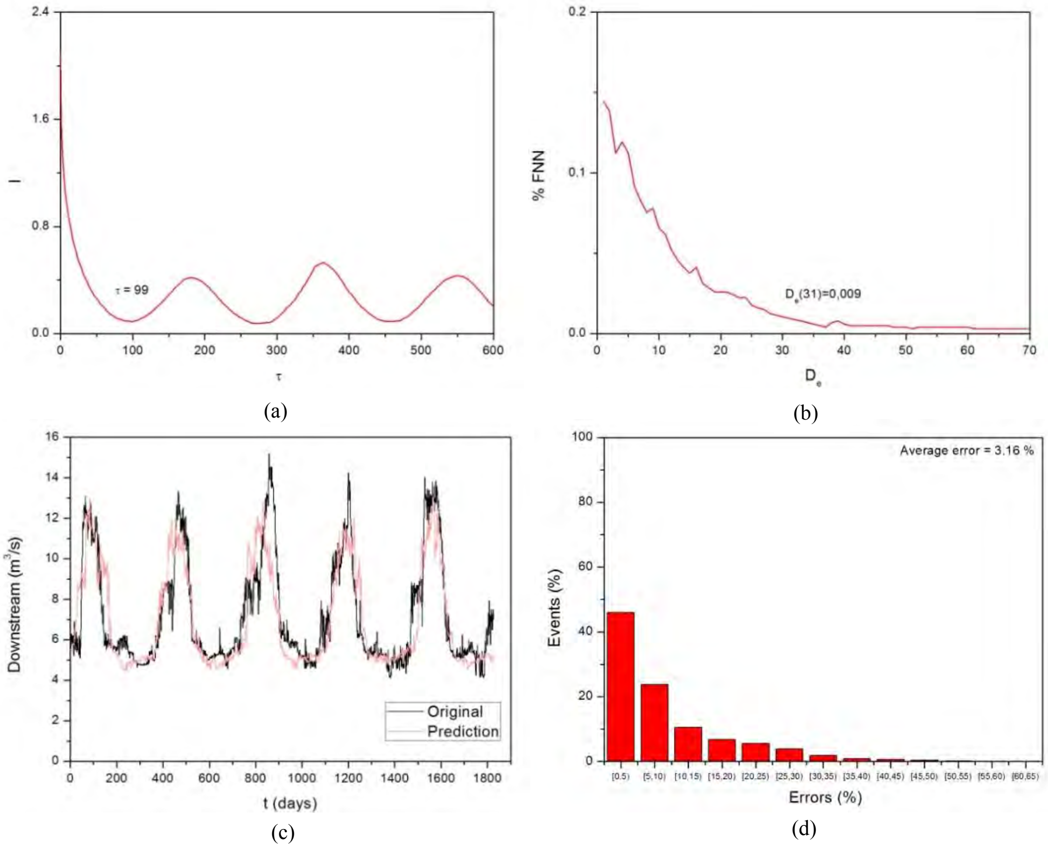


Figure 7. Downstream prediction (time series from 1985 to 2006). (a) Average mutual information; (b) embedding dimension; (c) comparison between original time series and prediction; (d) error analysis.

ISSN 2176-5480

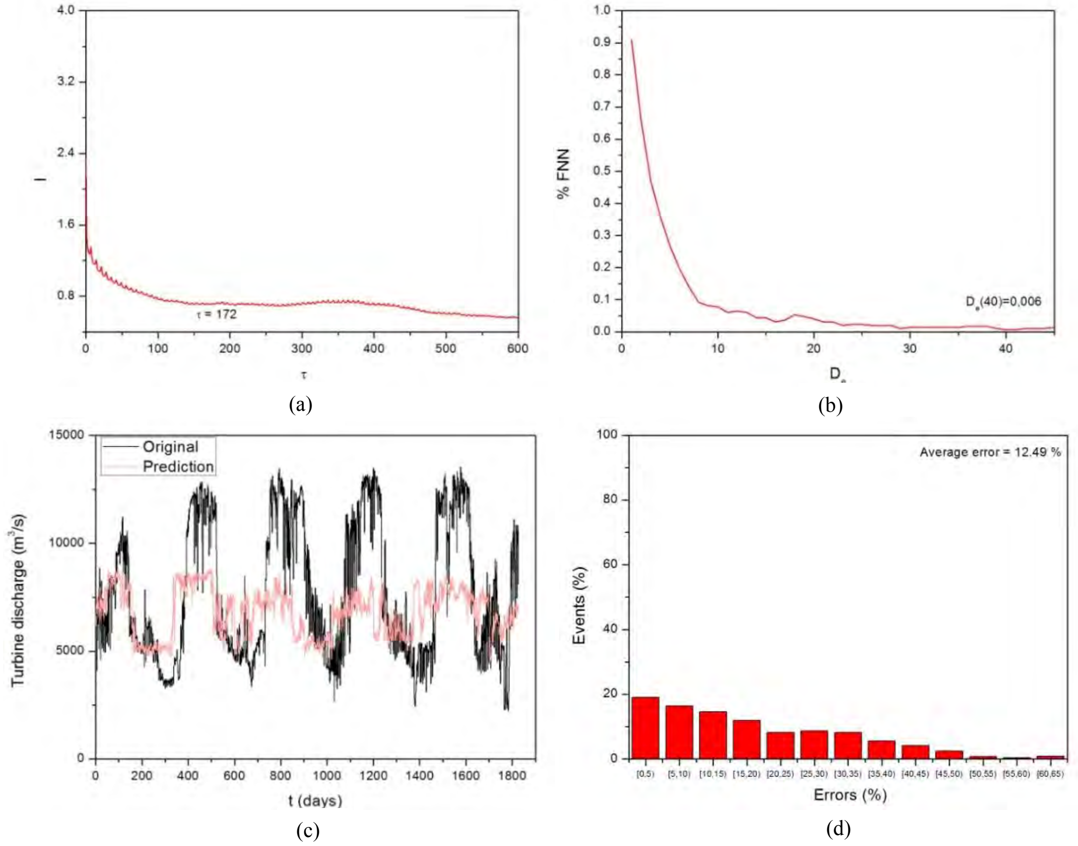


Figure 8. Turbine discharge prediction (time series from 1985 to 2006). (a) Average mutual information; (b) embedding dimension; (c) comparison between original time series and prediction; (d) error analysis.

Since predictions of the turbine discharge time series do not present good results, it is important to look for alternatives approaches to this aim. The main reason for the bad results is the non-stationary behavior of the time series (Coulibaly & Baldwin, 2005; Aguirre, 2007). Chen & Rao (2002) suggested that the non-stationary series may be partitioned into stationary segments. Therefore, this time series is split in stationary periods. Under this assumption, it is considered a known series from 2005 to 2009 (5 years, with 1,826 data points) in order to perform the prediction from 2010 to 2011 (2 years). Figure 9 shows the turbine discharge time series for this period.

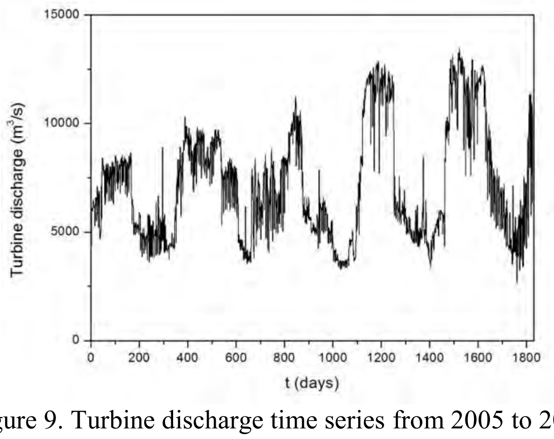


Figure 9. Turbine discharge time series from 2005 to 2009.

The analysis starts by evaluating delay parameters for affluent time series. Figure 10 (upper part) presents average mutual information and false nearest neighbors analyses. From these, it is possible to conclude that time delay is $\tau = 109$ and embedding dimension is $D_e = 15$. Afterwards, simple nonlinear prediction is employed to model the series, predicting future values from 2010 to 2011 (2 years). Results show a good agreement between the original and the

ISSN 2176-5480

predicted series and it is important to notice that both series has average values that are close (respectively, 8,215.01 m^3/s and 8,297.01 m^3/s representing a difference of 1 %).

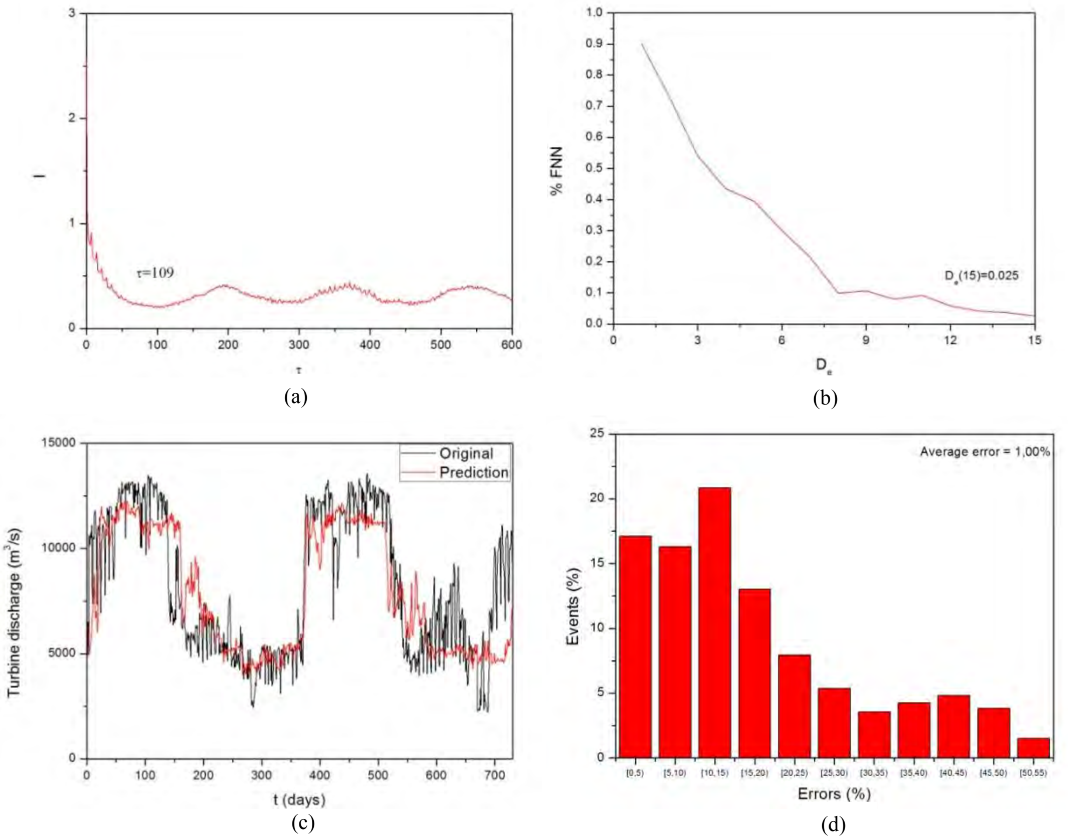


Figure 10. Turbine discharge prediction (time series from 2005 to 2009). (a) Average mutual information; (b) embedding dimension; (c) comparison between original time series and prediction; (d) error analysis.

Since the proposed procedures have captured the general behavior of the time series evolution, we are encouraged to use this approach to make predictions for future values. Therefore, we use a 27 years time series, from 1985 to 2011, with 9,861 data points, establishing a simple nonlinear prediction of 10 years (from 2012 to 2021) for all time series. By establishing a linear fit it is possible to observe that the original series and the predicted series showed the same trend behavior Figure 11a to 11d, except the turbine discharge time series shown in Figure 11e. Afterwards, for turbine discharge prediction we use a 7 years time series, from 2005 to 2011, with 2,556 data points, establishing a simple nonlinear prediction of 10 years (from 2012 to 2021). Therefore, by establishing a linear fit is possible to observe that the original series and the predicted series showed the same behavior, as shown in Figure 12.

ISSN 2176-5480

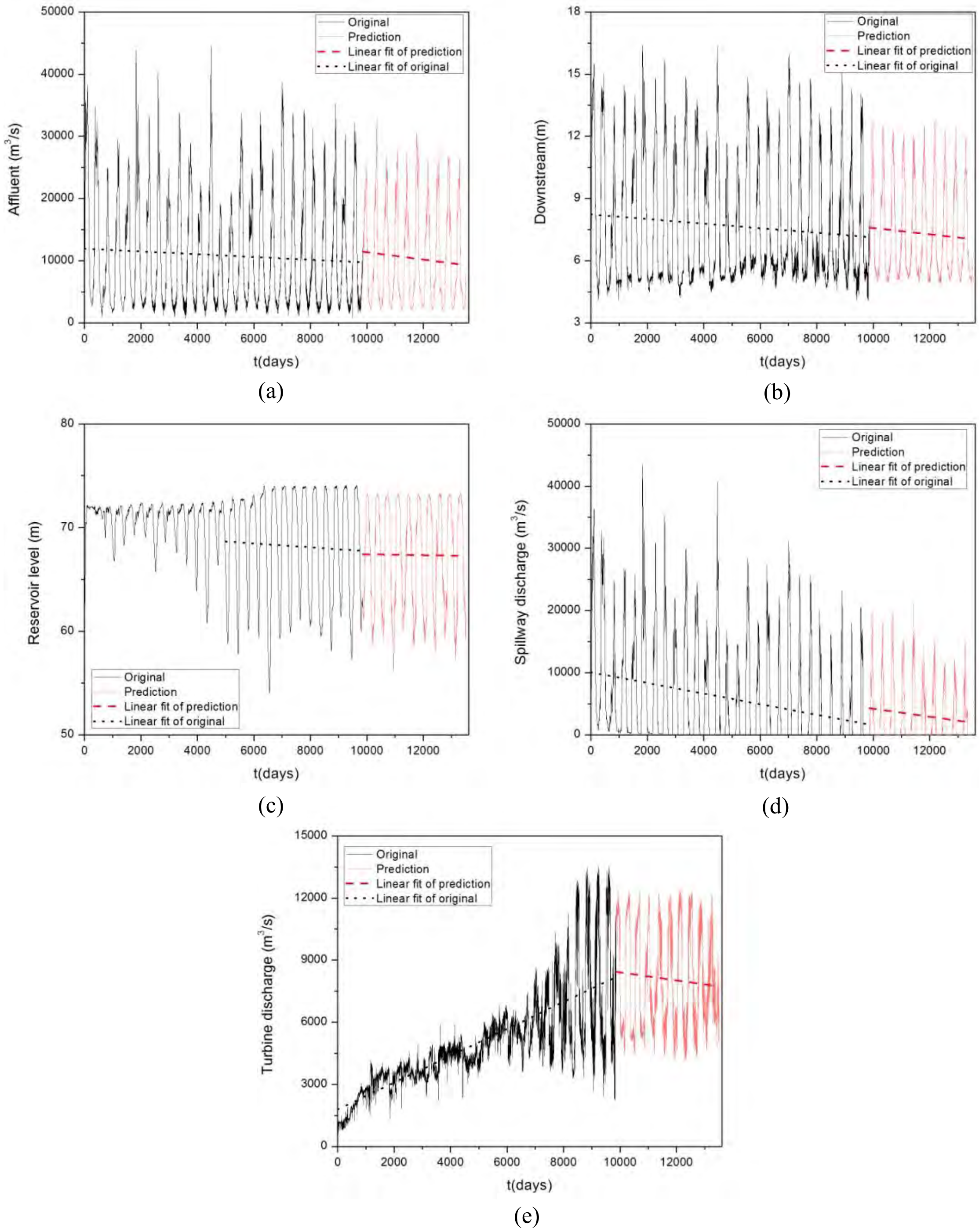


Figure 11. Prediction from 2012 to 2021 and linear fit.

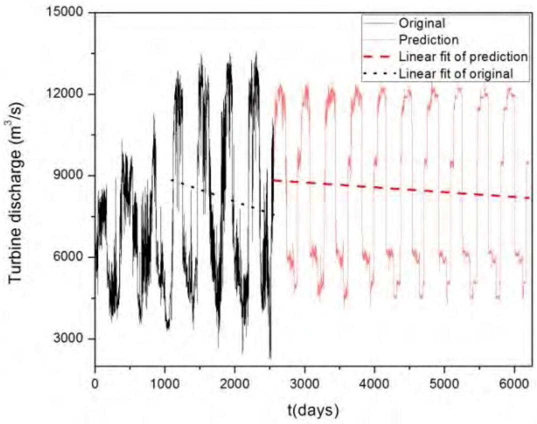


Figure 12. Prediction from 2012 to 2021 and linear fit.

5. CONCLUSIONS

This paper deals with the nonlinear time series analysis related to hydrologic reservoir. Hydrologic time series from Tucuruí reservoir is employed in order to establish a prediction model. State space reconstruction is done using the method of delay coordinates and delay parameters, time delay and embedding dimension, are respectively calculated by the method of average mutual information and the method of false nearest neighbors. Prediction is performed using the simple nonlinear prediction technique. In order to establish the model verification, predicted results are compared with time series by the average error and the daily error. Different number of data points is employed for model verification. Results show that the method captures the general behavior of the stationary time series. The non-stationary time series needs to be partitioned into stationary segments. After this verification, the procedure is employed to establish prediction of future values. In this regard, 10 years forecast is performed evaluating the hydrologic reservoir until 2021. By establishing a linear fit it is possible to observe that the original time series and the predicted series showed the same behavior. The authors agree that the nonlinear tools employed in this work can be useful for the analysis of hydrologic reservoir.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Eletronorte, Brazilian Research Agencies CNPq, CAPES and FAPERJ and through the INCT-EIE (National Institute of Science and Technology - Smart Structures in Engineering) the CNPq and FAPEMIG. The Air Force Office of Scientific Research (AFOSR) is also acknowledged.

7. REFERENCES

Abarbanel, H. D. I. (1995), "Tools for analyzing observed chaotic data", in Guran, A. & Inman, D. J., editors, *Smart Structures, Nonlinear Dynamics, and Control*, Prentice Hall, U.S.A., pp. 1 – 86.

Ashu, J. & Avadhnam, M. K. (2007), "Hybrid neural network models for hydrologic time series forecasting", *Applied Soft Computing*, v.7, n.2, pp.585-592.

Aguirre, L.A., (2007), "*Introdução à identificação de sistemas*", Editora UFMG, Brasil.

Casdagli, M. (1989), "Nonlinear prediction of chaotic time series", *Physica D*, Vol. 35, pp. 335 – 356.

Chen, H.L. & Rao, A.R. (2002), "Testing hydrologic time series for stationarity", *Journal of Hydrologic Engineering*, v.7, n.2, pp. 129-136.

Cheng, C. T., Xie, J.X, Chau, K. W. & Layeghifard, M. (2008), "A new indirect multi-step-ahead prediction model for a long-term hydrologic prediction", *Journal of hydrology*, v.361, n.1, pp.118-130.

Coulibaly, P. & Baldwin, C.K (2005),"Nonstationary hydrological time series forecasting using nonlinear dynamic methods", *Journal of Hydrology*,v.307,n.1,pp.164-174.

Deus, R., Brito, D., Mateus, M., Kenov, I., Fornaro, A., Neves, R. & Alves, C.N. (2013), "Impact evaluation of a pisciculture in the Tucuruí reservoir (Pará, Brazil) using a two-dimensional water quality model Original Research Article", *Journal of Hydrology*, v.487, n.22, pp.1-12.

ISSN 2176-5480

Fraser, A. M. & Swinney, H. L. (1986), "Independent coordinates for strange attractors from mutual information", *Physical Review A*, v.33, pp.1134–1140.

Hegger, R., Kantz, H. & Schreiber, T. (1999), "Practical implementation of nonlinear time series methods: The TISEAN package", *Chaos*, v.9, n.2, pp.413–435.

Hong, Y.S.T. (2012), "Dynamic nonlinear state-space model with a neural network via improved sequential learning algorithm for an online real-time hydrological modeling", *Journal of Hydrology*, v.468–469, pp.11–21

Kantz, H. & Schreiber, T. (1997), "*Nonlinear time series analysis*", Cambridge University Press, United Kingdom.

Karunasinghe, D.S.K. & Liong, S.Y. (2006), "Chaotic time series prediction with a global model: Artificial neural network", *Journal of Hydrology*, v.323, n.1, pp. 92–105.

Kennel, M. B., Brown, R. & Abarbanel H. D. I. (1992), "Determining embedding dimension from phase-space reconstruction using a geometrical construction", *Physical Review A*, v.25, n.6, pp.3403–3411.

Pinto, E.G.F. & Savi, M.A. (2003), "Nonlinear prediction of time series obtained from an experimental pendulum", *Current Topics in Acoustical Research – Research Trends*, v.3, pp.151-162.

Savi, M. A. (2006), "*Nonlinear dynamics and chaos*", Editora E-papers (in portuguese).

Schreiber, T. (1999), "Interdiciplinary application of nonlinear time series methods", *Physics Reports*, v.308, pp.1–64.

Viola, F. M., Paiva, S. L. D. & Savi, M. A., 2010 "Analysis of the global warming dynamics from temperature time series", *Ecological Modelling*, v.221, n.16, pp.1964-1978, 2010. doi: 10.1016/j.ecolmodel.2010.05.00.

Weigend, A. S. & Gershenfeld, N. A. (1994), "*Time series prediction: forecasting the future and understanding the past*", Perseus Books.

8. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.