PUC RIO

ISSN
Number 4 | February 2013

COMPUTATIONAL INTELLIGENCE SERIES

A Stochastic Model based on Neural Networks

Luciana C. D. Campos

Marley M. B. R. Vellasco

Juan G. L. Lazo

Laboratório Inteligência Computacional Aplicada

COMPUTATIONAL INTELLIGENCE SERIES

Number 4 | February 2013

A Stochastic Model based on Neural Networks

Luciana C. D. Campos

Marley M. B. R. Vellasco

Juan G. L. Lazo

CREDITS
Publisher:
MAXWELL / LAMBDA-DEE
Sistema Maxwell / Laboratório de Automação de Museus, Bibliotecas Digitais e Arquivos
http://www.maxwell.lambda.ele.puc-rio.br/
Organizer:
Marley Maria Bernardes Rebuzzi Vellasco
Cover:
Ana Cristina Costa Ribeiro

A Stochastic Model based on Neural Networks

Luciana C. D. Campos and Marley M. B. R. Vellasco and Juan G. L. Lazo

Abstract— **This paper presents the proposal of a generic model of stochastic process based on neural networks, called Neural Stochastic Process (NSP). The proposed model can be applied to problems involving phenomena of stochastic behavior and / or periodic features. Through the NSP's neural networks it is possible to capture the historical series' behavior of these phenomena without requiring any a priori information about the series, as well as to generate synthetic time series with the same probabilities as the historical series. The NSP was applied to the treatment of monthly inflows series and the results indicate that the generated synthetic series exhibit statistical characteristics similar to historical series.**

I. Introduction

Many real world problems have complex characteristics, such as nonlinear and chaotic behavior, requiring models fitted to capture their true characteristics in order to obtain an appropriate treatment [1], [2], [3], [4]. However, existing models are generally limited to specific problems, either because they are linear models (whose application to non linear problems leads to inconsistent or inadequate solutions), because they require a complex formulation, or even because they depend on some a priori assumptions, requiring a detailed knowledge of the problem, which is not often available [5], [6], [7], [8].

This motivated the development of a generic model of stochastic process based on neural networks to be applied to a broader range of problems, involving phenomena with stochastic behavior and/or presenting regular features of their probabilistic properties, as mean and variance, among others. This inherently non-linear model is called Neural Stochastic Process (NSP), which captures the historical series behavior to generate synthetic time series, equally likely to the historical series, applicable to the solution of different kinds of problems as those involving climatic or economic phenomena.

The NSP uses Multilayer Perceptron (MLP) neural networks [9], each with a single hidden layer, which are trained using the supervised learning algorithm Levenberg-Marquardt [10], a variation of the back propagation algorithm [9], [11]. As MLP neural networks are universal approximators [9], [12], they have been largely applied to the study of time series forecasting [13], [15], [17], [19].

The use of MLP in time series forecasting involves learning the past behavior of the series and using this knowledge to predict a specific point in the future or, if multi-step process is being used, a single sequence of points. However, in some problems, a single sequency of values is not enough to fully map the intrinsic uncertainties related to the time series. These intrinsic uncertainties can be captured by a stochastic model capable of producing synthetic series, different from the historical series but equally probable. The use of a stochastic model allows the information contained in the time series to be more completely extracted, allowing the evaluation of relevant risks and uncertainties. Therefore, the objective of the proposed Neural Stochastic Process is not to generate a single sequence of predicted points, but a set of scenarios of synthetic series also probable to the analyzed time series.

The proposed model was applied to the analysis of seasonal hydrological series with monthly intervals and the results showed that the NSP is able to generate synthetic series with similar characteristics to the historical series.

This paper is organized as follows: Section II presents the description of the generic NSP, discussing its formulation as well as the generation process of synthetic series. Section III describes the NSP modeling specifically applied to generate synthetic series of monthly inflows, and Section IV presents the results obtained in this specific case study. Finally, Section V presents the conclusions of the paper.

II. NSP Model Description

As mentioned before, the proposed generic model for a stochastic process, NSP, is based on artificial neural networks. Through the neural networks it is possible to capture the characteristics of the time series without making a priori assumptions about the series' behavior or performing any kind of decomposition, such as leaving it stationary or removing certain characteristics as cycle and tendency. To accomplish that, the neural networks must have a short-term memory, which is performed by the well known "windowing" techniques [13], [16]. This technique consists of introducing memory into the hidden layer neurons, providing past values from the analyzed time series.

In the basic case, where the series being examined are non periodic, the NSP is formed by a single stochastic component (SC). A SC is composed of a neural network (with a temporal window in the input layer, containing the past terms of the series to be modeled), and a probability distribution function. The adjustment of this neural network is performed in its training process, where the difference between the desired output and the output provided by the network provides a series of residues. The probability distribution function is obtained using this series of residues, and its function is to provide the random values that are summed to the output of the neural network. These random values are the stochastic

Luciana Conceição Dias Campos, Marley Maria B. R. Vellasco and Juan G. Lazo Lazo are with the Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil (email: {lcampos, marley, juan}@ele.puc-rio.br).

Luciana Conceição Dias Campos is with the Department of Computer Science, Federal University of Juiz de Fora, Juiz de Fora, Minas Gerais, Brazil (email: luciana.campos@ufjf.edu.br).

part of the model that allows the generation of scenarios of synthetic series, as illustrated in Figure 1.

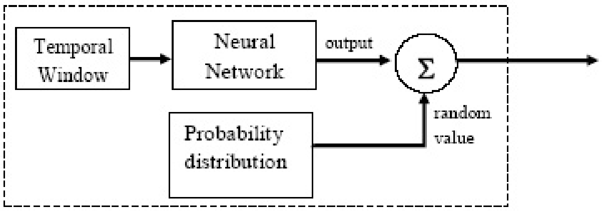


Fig. 1. Component Stochastic of NSP

For periodic time series, on the other hand, the NSP's parameters must be adequate to the time series interval, as well as its period. For this reason, the NSP is modeled with a specific stochastic component (SC) for each period of the time series. For example, in a time series with a monthly period, the NSP is composed of 12 SCs - one for each month.

A. Synthetic series generated through the NSP

Let $Z(t)$ be one time series with s seasonal period and n simultaneous observations in all periods. The time index t is then described by equation 1:

$$t = (r-1) \cdot s + m \qquad (1)$$

where

- $r = 1 \dots n$ is the observation number at each period of the series;
- $m = 1 \dots s$ corresponds to a period of the time series;
- $s \in \mathbb{N}$ is the total number of periods;
- $n \cdot s$ is the size of the observed series.

For example, in a monthly time series, r is the year, m corresponds to the month, $s = 12$ and n is the total number of years considered by the series.

As mentioned in the previous section, to model a periodic time series the NSP is composed of s stochastic components (SCs), one for each period m of the series, as illustrated by the NSP's block diagram on Figure 2.

As can be observed from Figure 2, when NSP is modeling a periodic time series, there is a concatenation among the stochastic components where the value of the time series provided by the SC for a given period is an input of the neural network's temporal window of the following period SC. It must be pointed out that at the beginning of the synthetic time series generating process, it is necessary to provide the first values of past terms for setting up the time windows of NSP's neural networks. These initial values are taken directly from the historical series.

All synthetic series are generated using the same set of initial values taken from the historical series. What differentiates the series of each scenario is the random value, derived from the probability distribution function, which is added to the output given by the neural network, in each stochastic component. Figure 3 illustrates a set of scenarios of synthetic series generated by the NSP.

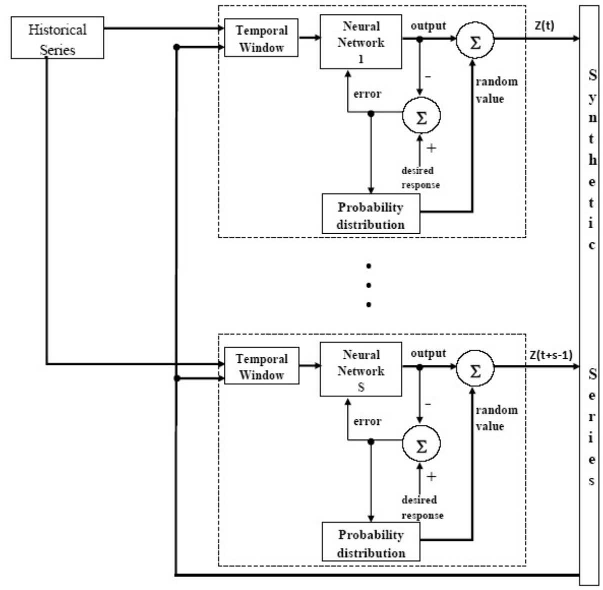


Fig. 2. Generation of synthetic series in NSP

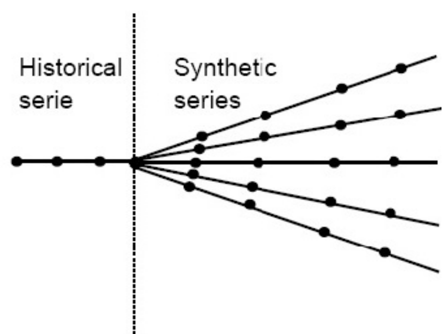


Fig. 3. Synthetic series generated by NSP

B. NSP Formulation

The number of past terms used in the input window of SC for period m is called order and is represented by p_m. Thus, to obtain a value for the series at time t, $Z(t)$, the NSP accesses the corresponding SC for period m, and its neural network receives the terms: $Z(t-1), Z(t-2), \dots, Z(t-p_m)$. Additionally, to reinforce the learning of the series' periodic behavior, the time series value at the previous period, corresponding to $Z(t-s)$, is also added at the input.

The first past terms used are obtained from the historical data. After, the past terms are obtained from the synthetic serie that are being generated.

Figure 4 shows, in details, a neuron belonging to the hidden layer of the neural network of order p_m, whose output is given by equation 2.

$$y_i = \varphi\left(\omega_{i,0} \cdot Z(t-s) + \left(\sum_{j=1}^{p_m} \omega_{i,j} \cdot Z(t-j)\right) + \theta_i\right) \qquad (2)$$

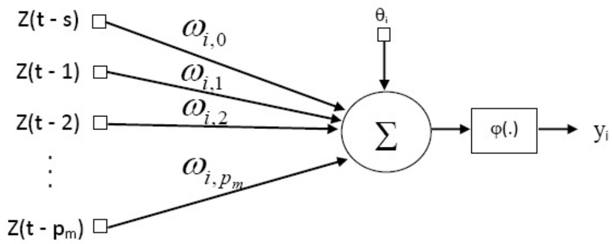


Fig. 4. Neuron belonging to a hidden layer of a NSP's network of order p_m

where φ is the activation function of neuron i, $\omega_{i,j}$ is the synaptic weight of the connection between the input j and neuron i and θ_i is the neuron bias.

Assuming that the neural network of order p_m has l_m neurons in the hidden layer, the output neuron is shown in Figure 5, with its output calculated by equation 3.

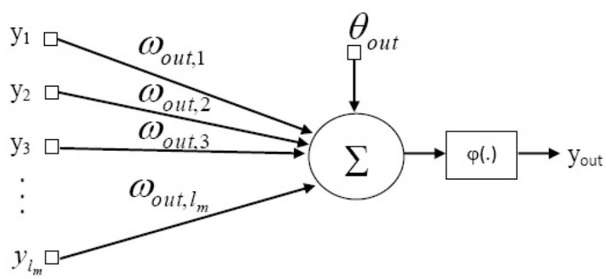


Fig. 5. Output Neuron of the NSP's neural network with l_m neurons in the hidden layer

$$y_{out} = \varphi_{out}\left(\sum_{i=1}^{l_m} \omega_{out,i} \cdot y_i + \theta_{out}\right) \tag{3}$$

where φ_{out} is the activation function of output layer's neuron, represented by out, $\omega_{out,i}$ is the synaptic weight between input i (corresponding to the hidden layer's output neuron i) and the neuron out and θ_{out} is the output neuron bias.

As illustrated in Figure 2, the SC's output is the sum of the neural network's output with a random value from the probability distribution function of neural network's errors obtained from the difference between the desired output and neural network's output (y_{out}). Thus, the value of one synthetic serie $\bar{Z}(t)$ whose time index t is described by equation 1, is generated using equation 4.

$$\bar{Z}(t) = y_{out} + \alpha(t) \tag{4}$$

where $\alpha(t)$ is the random value from the probability distribution function of the neural network's errors of the SC for the corresponding period m.

III. Case Study: Generating Monthly Inflows Scenarios

The Brazilian National Interconnected System (NIS) is a coordination and control system, composed by the companies from four different regions: South, Southeast/Midwest, Northeast and part of the North, which congregates the electricity production and transmission system. The Brazilian NIS is, a very large system, composed maily by hydrolelectric plants [20].

Currently, the NIS is segmented into four subsystems corresponding to the interconnected systems: South, Southeast/Midwest, Northeast and North. Since it is a huge system, for medium and long term planning an aggregation occurs from the plants reservoirs into power equivalent reservoirs, one for each subsystem. There is also the aggregation of inflows to energy plants in Affluent Natural Energy (ANE), which correspond to the estimate of the energy that can be generated with all the inflows to each reservoir from that equivalent reservoir, under a given operational policy [21], [22].

Each ANE is a non stationary series, due to periods of flooding and dry season in the year, and seasonal with periods of 12 months, which, generally, exhibit periodic correlations [23]. In the context of the energetic operation planning of the Brazilian hydrothermal system, the generation and analysis of different scenarios is crucial to obtain acceptable risk rates in the future [21]. The use of the only scenario available in practice, which is the record of observed inflows (called historical series), is insufficient to provide a broader analysis of the possible long term inflows scenarios. Through the creation of plausible scenarios, it is possible to reproduce the basic features of the historical data, gathering the real information of this series and allowing the assessment of risks and uncertainties pertaining to a hydroelectric system. Thus, the NSP has been applied to treat ANE series uncertainties, generating a set of synthetic series.

The modeling of neural stochastic processes for the treatment of ANE series is composed of four systems NSP(p, l), one for each subsystem of the NIS. As the series of ANE are seasonal with periods of 12 months, each NSP(p, l) consists of 12 SCs, one for each month, as can be observed from Figure 6. Therefore:

- $p = p_1 \ldots p_{12}$: vector with the neural network's order of each SC;
- $l = l_1 \ldots l_{12}$: vector with the number of neurons in the neural network's hidden layer of each SC.

The order p_m of the neural network from the SC of m period can vary from 1 to 11, which corresponds to the maximum number of past terms allowed to keep the annual seasonality. From preliminary experiments it was observed that some months usually need smaller orders than others. Therefore, four types of order for each neural network were evaluated: 3, 6, 9 and 11 months lags. For each defined order, 1 to 20 hidden nodes were tested, resulting in a total of 80 configurations to be evaluated for each month m of each NIS subsystem. In addition, each of these 80 neural networks' configurations were trained 10 times with different synaptic weights initializations.

To adjust each NSP(p, l), an historic of ANE values was used in this work. There is an historic of ANE values, from

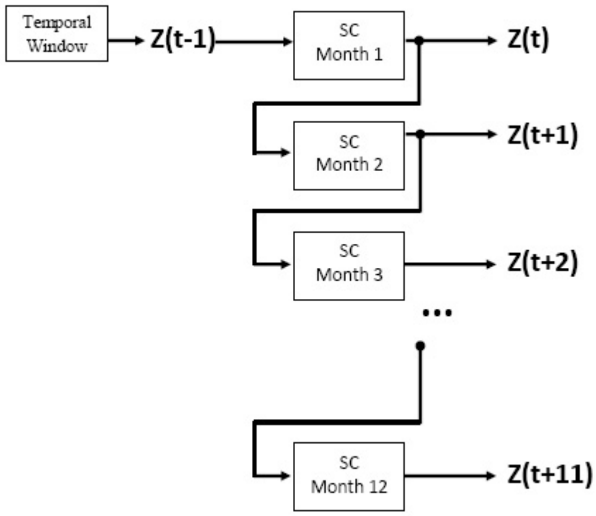


Fig. 6. Concatenation among the 12 NSP's stochastic components

1931 to 2005, for each NIS subsystem. This historic data was divided into two subsets: training and validation sets. Each validation set weights was used to determine the optimum number of training cycles, in the early stopping process to avoid overfitting, as well as to verify the best number of hidden neurons. The validation stage of the NSP neural networks consist of the generation of a series with k years, composed by the concatenation of the s neural networks' outputs. Therefore, the k last years in the historical data has been used to compose the validation set, while all prior years define the input-output training subset. In this study k is equal to five, since the current NIS planning system is performed with synthetic time series for the next five years. Therefore, the validation set is composed of the last five years of the historical data.

Once all neural networks have been trained, the neural network of month m that provides the best performance for the validation set is selected to form the m-th SC of the NSP. As the validation set is composed of only five years, it was also decided to verify the performance of the trained neural networks' over a larger number of data patterns, composed of the unification of training and validation sets. The chosen performance measure is the mean absolute percentage error (MAPE) [24], which is widely used to validate time series models, calculated by equation 5.

$$MAPE = \frac{1}{x} \cdot \sum_{k=1}^{x} \left| \frac{\bar{Z}(k) - C((k-1) \cdot s + m)}{\bar{Z}(k)} \right| \tag{5}$$

where m is the month, x is the total number of patterns of month m in the series, k is the element index, $s = 12$ is the amount of months, $Z(k)$ is the k-th unified pattern's desired output of month m, C is a series created with the neural networks' outputs.

Once the neural network that will compose the m-th SC of the NSP is selected, the error series obtained from the difference between the neural network's output and the training set outputs is used to calculate a theoretical probability distribution function for the composition of that SC.

IV. Simulation Results

Table I below presents, for each NIS's subsystem, the final configuration of each neural network that integrates the NSP. Table I also presents the order and the number of neurons in the hidden layer of each neural network in the NSP of each subsystem.

TABLE I
Configurations of Selected Neural Networks

Southeast-Midwest Sub-system			South Sub-system		
Month	Lags	# Neurons	Month	Lags	# Neurons
1	3	19	1	6	3
2	9	20	2	9	17
3	6	15	3	9	1
4	11	7	4	11	3
5	3	10	5	3	19
6	3	10	6	6	10
7	9	3	7	3	14
8	6	16	8	6	1
9	11	16	9	11	2
10	3	19	10	3	3
11	3	12	11	11	3
12	6	9	12	6	1
Northeast Sub-system			**North Sub-system**		
Month	Lags	# Neurons	Month	Lags	# Neurons
1	11	9	1	11	11
2	3	6	2	9	14
3	11	6	3	6	4
4	3	13	4	11	15
5	3	2	5	9	16
6	11	19	6	11	3
7	6	1	7	6	1
8	11	1	8	11	8
9	9	3	9	9	17
10	6	1	10	6	19
11	9	5	11	9	18
12	9	7	12	9	15

With each final NSP (p, l) of Table I, 200 synthetic series of 5 years (60 months) were generated. In order to analyze the quality of these generated synthetic series, statistical tests [25], [26] were performed to verify whether the synthetic series are also equally probable to ANE series.

The *t*-Test was applied to evaluate whether the mean of the synthetic series is statistically equal to the historical mean. The hypothesis is that the average of 200 values of each month in each year is statistically equal to the historical mean for the corresponding month. In each test run, a p-value is obtained and if that p-value is above the significance level of 5%, that hypothesis is accepted. The percentage of 60 p-values (one for each month of the five-year series) above the level of significance indicates the performance of NSP to generate synthetic series, informing how well the model would reproduce the first moments of the series.

Analogously, the Levene test was also applied to evaluate whether the variance of each period of the scenarios is statistically equal to the historical variance of the corresponding month, as well as the Kolmogorov-Smirnov (KS) test to verify if the scenarios come from the same probability distribution of the historic, indicating that the model reproduces correctly the behavior of the historical series.

As the ANE's historic was divided into two subsets, one for training and the other for validation of the NSP's neural networks, the synthetic series assessment was carried out with these two sets. Table II shows the obtained results on the goodness of fit test. As can be noted, all NSPs fitted with the historical series of the four subsystems provided, with a performance over 50 %. The best fitting occurred with the models of Southeast-Midwest and South subsystems. The Northeast system resulted in a higher adhesion in tests with the Training historical data than with the Validation data. The North model got better compliance with the Levene's test and lower adherence to KS tests. This difference of NSP performance occurs because the influence of ANE's behavior in each subsystem. Each subsystem of NIS correspond a different regions on Brazil, so that its ANE's serie is affected for a specific climate phenomena.

TABLE II
GOODNESS OF FIT TEST RESULTS

Sub-system	Historic					
	Training			Validation		
Test	t	Levene	K-S	t	Levene	K-S
Southeast-Midwest	94%	95%	94%	97%	100%	97%
South	75%	90%	90%	95%	99%	99%
Northeast	97%	99%	94%	69%	64%	54%
North	85%	90%	64%	74%	99%	67%

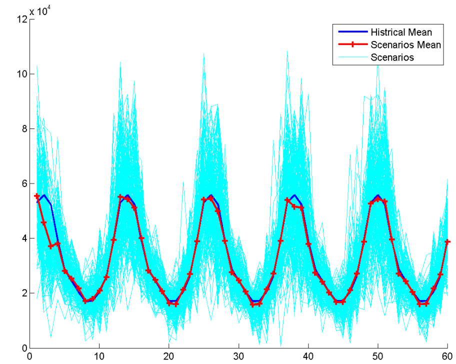
Figures 7 and 8 show the envelope of the 200 scenarios (synthetic series) with 5 years of ANE-generated by the complete NSP, including the curve of the average for each month in each year of synthetic series and also the curve of the monthly average from the historical set replicated in 5 years, in order to analyze the behavior of this average's series generated in all 60 months. This was accomplished for the two historical sets in each subsystem.

It is observed that, in all subsystems, the average of the historical subsets are inside the envelopment scenarios. In overall, the envelopes scenarios had followed the performance of ANE historical series.

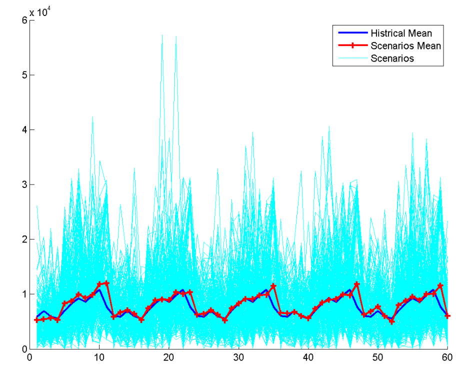
V. CONCLUSIONS

The objective of this study was to develop a new general stochastic process model, intrinsically non-linear, which can be applied to a range of problems with stochastic behavior phenomena and/or with periodic characteristics of their properties.

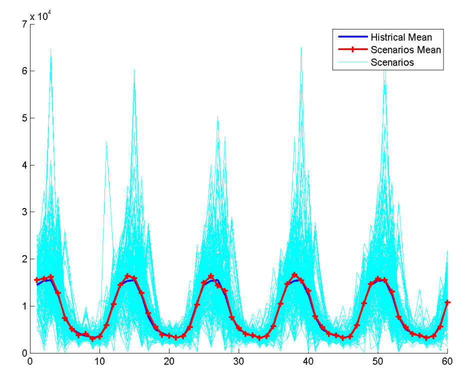
The proposed model, called NSP, was based on neural networks. Through the neural networks, the NSP is able to identify and assimilate characteristics of historical time series, such as seasonality, periodicity and trend, without requiring any a priori information about the series. The goal of the NSP is to generate scenarios with synthetic time series equally probable to the analyzed historical series, addressing any period of time for the necessary amount.



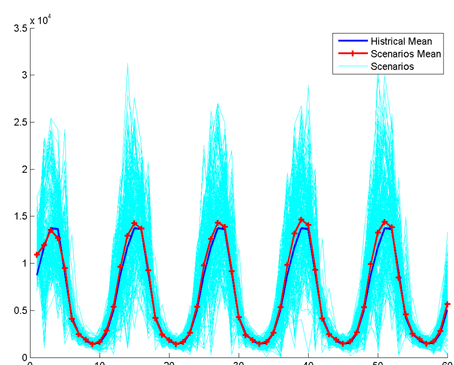
Sub-system Southeast-Midwest



Sub-system South

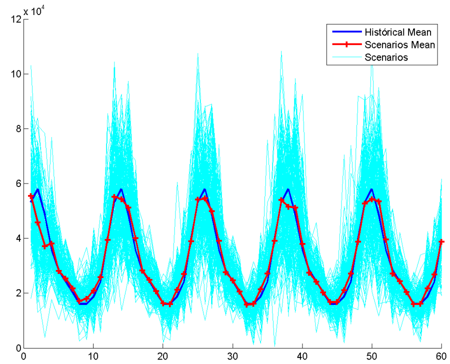


Sub-system Northeast

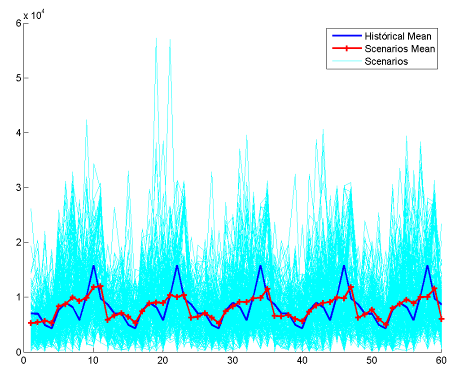


Sub-system North

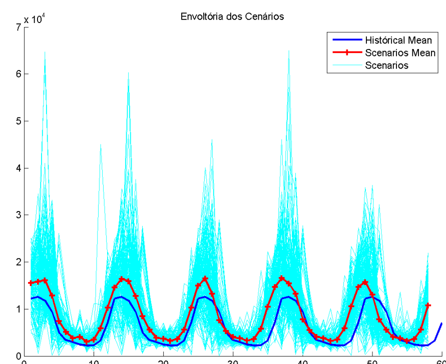
Fig. 7. Envelopment of the scenarios generated in the NSP and the historical average of the training set



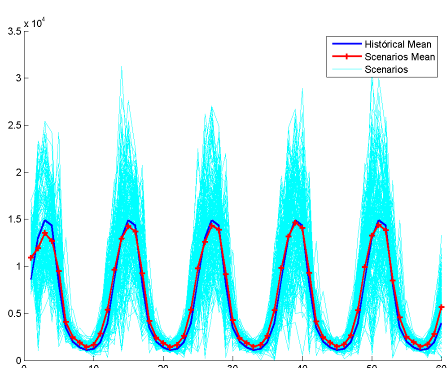
Sub-system Southeast-Midwest



Sub-system South



Sub-system Northeast



Sub-system North

Fig. 8. Envelopment of the scenarios generated in the NSP and the historical average of the validation set

In the case study, the NSP was applied to the uncertainty analysis of the monthly inflows series. A 200 synthetic series of 5 years of Affluent Natural Energy (ANE) were generated for each one of the four sub-systems that comprise the Brazilian National Interconnected System (NIS). In each set of synthetic series, an evaluation was performed with some goodness of fit tests to verify whether they were statistically similar to historical series of ANE. It was found that the synthetic series generated by this NSP showed the best goodness of fit to the set of historical series and the envelopes of these scenarios encompass the ANE historical average. Therefore, it can be concluded that the proposed NSP is able to capture the behavior of the historical series to generate synthetic series with similar characteristics.

REFERENCES

[1] Oliver Knil, "Probability and Stochastic Processes with Applications", *http://www.math.harvard.edu/∼knill/teaching/math*$144_1$994*/probability.pdf*, february, 2011.

[2] Rabi N. Bhattacharya and Edward C. Waymire, "Stochastic Processes with Applications", *Society for Industrial and Applied Mathematics*, Philadephia, PA, 2009, DOI:10.1137/1.9780898718997, isbn = 9780898718997.

[3] Zeev Schuss, "Theory and Applications of Stochastic Processes: An Analytical Approach", *Publisher: Springer*, 2010, ISBN: 1441916040.

[4] H. D. I. Ababarnel, R. Brown, J. L. Sidorowich and L. S. Tsimring, "The analysis of observed chaotic data in physical systems", Rev. Modern Phys, 65(4) (1993), 1331-1392.

[5] N.R. Sanders, "Managing the forecasting function", *Industrial Managment and Data Systems*, vol. 95, number 4, pp. 12-18, 1995.

[6] Simon M. Potter, "Nonlinear time series modelling: an introduction", *Federal Reserve Bank of New York*, Staff Reports, no. 87, http://ideas.repec.org/p/fip/fednsr/87.html, 1999.

[7] P. Bougerol and N. Picard, "Strict stationarity of generalized autoregessive processes", *Annals of Probability*, 20, 1714-1730, 1992.

[8] N. A. Gershenfeld and A. S. Weigend, "The future of time series, in:", A. N. Gershenfeld and A. S.Weigen (eds), *Time Series Prediction: Forecasting the Future and Understanding the Past*, pp. 1-70, 1993.

[9] S. Haykin, "Neural Networks: A Comprehensive Foundation", *Prentice Hall*, 2 edition , 1998, ISBN-10:013273350.

[10] M. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm", *IEEE Transactions on Neural Networks*, november, 1994.

[11] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by backpropagating errors", *Nature (London)*, 323, pp. 533-536, 1986.

[12] M. K. Hornik and H. White Stinchcombe, "Multilayer feedforward networks are universal approximators ", *Neural networks*, 2(5), pp. 259-366, 1989.

[13] Eric A. Plummer, "Time Series Forecasting with Feed-forward Neural Networks: guidelines and limitations", *Department of Computer Science and The Graduate School of The University of Wyoming*, Master of Science in Computer Science, Laramie, Wyoming, July, 2000.

[14] V. R. Vemuri and R. D. Rodgers, "Artificial Neural Networks: Forecasting Time Series", *IEEE Computer Society Press*, 1994.

[15] R.B.C. Prudêncio, "Projeto Híbrido de Redes Neurais", *Federal University of Pernambuco, Msc.Computer Sciences*, Recife, PE, February 2002.{in portuguese}

[16] R. J. Frank and N. Davey and S. P. Hunt, "Time Series Prediction and Neural Networks." *J. Intell. Robotics Syst., Kluwer Academic Publishers*, Hingham, MA, USA, vol. 31, pp. 91–103, May, 2001.

[17] M.A.A. Fayal, "Previsão de Vazão por Redes Neurais Artificiais e Transformada Wavelet," *Pontifical Catholic University of Rio de Janeiro, Msc. Dissertation, Electrical Engineering Dep*, February 2008.{in portuguese}

[18] A. L. F. Batista, "Modelos de Séries Temporais e Redes Neurais Artificiais na Previsão de Vazão", *Federal University of Lavras, MSc. System Engineering*, Lavras, Minas Gerais, 2009.{in portuguese}

[19] H. Zhao, "A chaotic time series prediction based on neural network: Evidence from the Shanghai Composite index in China", *Test and Measurement, 2009. ICTM '09. International Conference on*, vol. 2, pp. 382 -385, 5-6 2009, doi = 10.1109/ICTM.2009.5413024.

[20] "ONS - Brazilian Operator of the Electric System", *http://www.ons.com.br*, 2011.

[21] M. V. F. Pereira and L. M. V. G. Pinto, "Operation planning of large-scale hydroelectric systems", *Proc. 8th Power System Computation Cont.*, 1984.

[22] K. W Hipel and A. I. McLeod, "Time Series Modelling of Water Resources and Environmental Systems", *Elsevier*, Amsterdam, The Netherlands, 1994, isbn = 0-444-89270-2.

[23] H. A. Thomas and M. B. Fierinig, "Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation", *Design of Water Resource Systems*, Maass, A.; Hufschmidt, M. M.; Dorfman, R.; Thomas, H. A.; Marglin, S. A.; Fair, G. M., Harvard University Press, Mass, Cambridge, 459-493, 1962.

[24] Z. Tang, C. Almeida, P. A. Fishiwick, "Time series forecasting using neural networks vs Box-Jenkins methodology," *Simulation*, vol. 57, no. 5, pp. 303-310, 1991.

[25] D. C. Montgomery and G. C. Runger, "Applied Probability and Statistics for Engineers", *Arizona State University*, 4th Edition, 2007, ISBN 978-0-471-74589-1.

[26] M. Ross Sheldon, "Introduction To Probability And Statistics For Engineers And Scientists", *ELSEVIER SCIENCE & TECHNOLOGY*, 2009, ISBN 9780123704832.

[27] M.P. Kennedy and L.O. Chua, "Neural networks for nonlinear programming," *IEEE Transactions on Circuits and Systems*, vol. 35, no. 5, pp. 554-562, May. 1988.