A Feed-Forward Neural Networks-Based Nonlinear Autoregressive Model for Forecasting Time Series

Modelo Auto Regresivo no Lineal Basado en Redes Neuronales Multicapa para Pronóstico de Series Temporales

Julián A. Pucheta[1], Cristian M. Rodríguez Rivero[1], Martín R. Herrera[2], Carlos A. Salas[2], H. Daniel Patiño[3] and Benjamín R. Kuchen[3]

[1]Mathematics Research Laboratory Applied to Control, Departments of Electrical and Electromechanical Engineering, Faculty of Exact, Physical and Natural Sciences, National University of Córdoba, Córdoba, Argentina.
{julian.pucheta, cristian.rodriguezrivero}@gmail.com
[2]Departments of Electrical Engineering, Faculty of Sciences and Applied Technologies, National University of Catamarca, Catamarca, Argentina.
{martincitohache, calberto.salas}@gmail.com
[3]Institute of Automatics Faculty of Engineering-National University of San Juan, San Juan, Argentina.
{dpatino, bkuchen}@unsj.edu.ar

Article received on January 15, 2010; accepted on October 08, 2010

Abstract. In this work a feed-forward NN based NAR model for forecasting time series is presented. The learning rule used to adjust the NN weights is based on the Levenberg-Marquardt method. In function of the long or short term stochastic dependence of the time series, we propose an on-line heuristic law to set the training process and to modify the NN topology. The approach is tested over five time series obtained from samples of the Mackey-Glass delay differential equations and from monthly cumulative rainfall. Three sets of parameters for MG solution were used, whereas the monthly cumulative rainfall belongs to two different sites and times period, La Perla 1962-1971 and Santa Francisca 200-2010, both located at Córdoba, Argentina. The approach performance presented is shown by forecasting the 18 future values from each time series simulated by a Monte Carlo of 500 trials with fractional Gaussian noise to specify the variance.

Keywords. Neural networks, time series forecast, Hurst's parameter, Mackey-Glass.

Resumen. Se presenta un modelo auto-regresivo no lineal (ARN) basado en redes neuronales para el pronóstico de series temporales. La regla de aprendizaje para ajustar los parámetros de la red neuronal (RN) está basado en el método Levenberg-Marquardt en función de la dependencia estocástica de la serie temporal, proponemos una ley heurística que ajusta el proceso de aprendizaje y modifica la topología de la RN. Esta propuesta es experimentada sobre cinco series temporales. Tres son obtenidas de la ecuación de Mackey-Glass (MG) en un intervalo de tiempo. Las dos restantes son series históricas de lluvia acumulada mensualmente pertenecientes a dos lugares y tiempos diferentes, La Perla 1962-1971 y Santa Francisca 2000-2010, Córdoba, Argentina. El desempeño del esquema se muestra a través del pronóstico de 18 valores de cada serie temporal, donde el pronóstico fue simulado mediante Monte Carlo con de 500 realizaciones con ruido Gaussiano fraccionario para especificar la varianza.

Palabras Clave: Redes neuronales, pronóstico de series temporales, parámetro de Hurst, ecuación Mackey-Glass.

1 Introduction

In order to use and model time series useful for control problems from agricultural activities such as the availability of estimated scenarios for water predictability (Liu and Lee, 1999 Masulli *et al*, 2001), seedling growth (Pucheta *et al*., 2007a; Guo *et al*., 2009) and decision-making , natural phenomena prediction is a challenging topic. In this work, the proposed approach is based on the classical NAR filter using time-lagged feed-forward networks, in which the data from the MG benchmark equation and Monthly Cumulative Rainfall time series are used to forecast the next 18 values. These forecasted data are simulated by a Monte Carlo method (Bishop, 2006). The number of filter parameters is put in function of the roughness of the time series, i.e. the error between the smoothness of the time series data and the forecasted modifies the number of the filter parameters.

1.1. The Neural Network Approach

Motivations that have led for this study follows the closed-loop control scheme (Pucheta *et al*, 2007a). The controller considers future meteorological conditions for designing the control law as shown Fig.1. In that sense, the controller takes into consideration the actual state of the crop by a state observer and the meteorological variables, referred by $\boldsymbol{x}(k)$ and R_o, respectively. However, in this paper only the controller's portion concerning with the prediction system is presented by using a benchmark and rainfall time series.

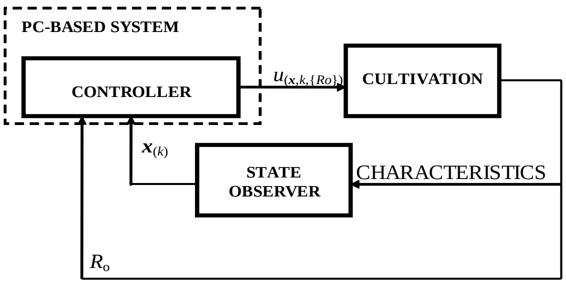


Fig. 1. Closed-loop PC-based control approach

The main contribution of this work is on the learning process, which employs the Levenberg-Marquardt rule and considers the long or short term stochastic dependence of passed values of the time series to adjust at each time-stage the number of patterns, the number of iterations, and the length of the tapped-delay line, in function of the Hurst's value, *H* of the time series. According to the stochastic characteristics of each series, *H* can be greater or smaller than 0.5, which means that each series tends to present long or short term dependence, respectively. In order to adjust the design parameters and show the performance of the proposed prediction model, solutions for the MG equation and Rainfall series are used. The NN-based nonlinear filter is applied to the time series obtained from MG and Monthly Cumulative Rainfall to forecast the next 18 values.

1.2. Samples of MG equation

Samples of MG equation are used to model natural phenomena and have been implemented by different authors to perform comparisons between different techniques for forecasting and regression models (Velásquez Henao and Dyna Red, 2004). This paper propose an algorithm to predict values of time series taken from the solution of the MG equation (Glass and Mackey, 1988) and Rainfall time series from La Perla (Pucheta *et al*, 2009a) and Santa Francisca, South of Cordoba.

The MG equation is explained by the time delay differential equation defined as,

$$\dot{y}(t)=\frac{\alpha y(t-\tau)}{1+y^{c}(t-\tau)}-\beta y(t) \tag{1}$$

where α, β, and c are parameters and τ is the delay time. According as τ increases, the solution turns from periodic to chaotic. Equation (1) is solved by a standard fourth order Runge-Kutta integration step, and the series to forecast is formed by sampling values with a given time interval.

Thus, samples of MG time series with a random-like behavior is obtained, and the long-term behavior changes thoroughly by changing the initial conditions. Furthermore, by setting the parameter β between 0.1 and 0.9 the stochastic dependence of the deterministic time series obtained varies according to its roughness.

1.3. fBm overview

Due to the random process of the time series, it is proposed to use the Hurst's parameter in the learning process to modify on-line the number of patterns, the number of iterations, and the number of filter inputs. This H tells us about the roughness of a signal, and also to determine its stochastic dependence. The definition of the Hurst's parameter appears in the Brownian motion from generalize the integral to a fractional one. The Fractional Brownian Motion (fBm) is defined in the pioneering work by Mandelbrot (Mandelbrot, 1983) through its stochastic representation

$$B_H(t)=\frac{1}{\Gamma\left(H+\frac{1}{2}\right)}\left(\int_{-\infty}^{0}\left((t-s)^{H-\frac{1}{2}}-(-s)^{H-\frac{1}{2}}\right)dB(s)+\int_{0}^{t}(t-s)^{H-\frac{1}{2}}dB(s)\right) \tag{2}$$

where, $\Gamma(\cdot)$ represents the Gamma function

$$\Gamma(\alpha)=\int_{0}^{\infty}x^{\alpha-1}e^{-x}dx \tag{3}$$

and $0<H<1$ is called the Hurst parameter. The integrator *B* is a stochastic process, ordinary Brownian motion. Note, that *B* is recovered by taking H=1/2 in (2). Here, it is assumed that B is defined on some probability space (Ω, F, P), where Ω, F and P are the sample space, the sigma algebra (event space) and the probability measure respectively. Thus, a fBm is a continuous-time Gaussian process depending on the so-called Hurst parameter 0<H<1. The ordinary Brownian motion is generalized to H=0.5, and whose derivative is the white noise.

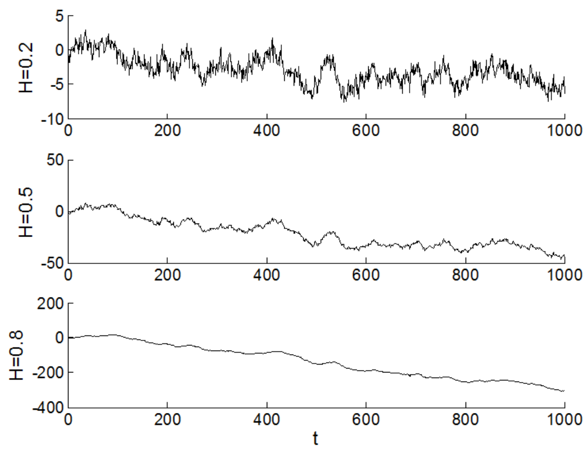


Fig. 2. Three sample path from fBm

The fBm is self-similar in distribution and the variance of the increments is defined by

$$Var\left(B_H(t)-B_H(s)\right)=\mathrm{v}\left|t-s\right|^{2H} \tag{4}$$

where, v is a positive constant.

This special form of the variance increments suggests various ways to estimate the parameter H. In fact, there are different methods for computing the parameter H associated to Brownian motion (Bardet *et al*, 2003; Dieker, 2004; Istas and Lang, 1994). In this paper, the algorithm uses a wavelet-based method for estimating H from a trace path of the fBm with parameter H (Abry *et al*, 2003; Dieker, 2004; Flandrin, 1992). Three trace paths from fBm with different values of H are shown in Fig. 2., where the difference in the velocity and the amount of its increments can be noted.

2 Problem statement

The classical prediction problem may be formulated as follow. Given past values of a process that are uniformly spaced in time, as shown by $x(n\text{-}T)$, $x(n\text{-}2T)$, . . . , $x(n\text{-}mT)$, where T is the sampling period and m is the prediction order. It is desired to predict the present value $x(n)$ of such process. Therefore, obtain the best prediction (in some sense) of the present values from a random (or pseudo-random) time series is desired. The predictor system may be implemented using either an autoregressive model-based nonlinear adaptive. The NNs are used as a nonlinear model building; in the sense that smaller the prediction error is in a statistical sense, the better the NN serves as model of the underlying physical process responsible for generating the data. In this work, time lagged feed-forward networks are used.

The present value of the time series is used as the desired response for the adaptive filter, and the past values of the signal supply as input of the adaptive filter. Then, the adaptive filter output will be the one-step prediction signal. In

Fig.3 the block diagram of the nonlinear prediction scheme based on a NN filter is shown. Here, a prediction device is designed such that starting from a given sequence $\{x_n\}$ at time n corresponding to a time series it can be obtained the best prediction $\{x_e\}$ for the following 18 values sequence. Hence, it is proposed a predictor filter with an input vector l_x, which is obtained by applying the delay operator, Z^{-1}, to the sequence $\{x_n\}$. Then, the filter output will generate x_e as the next value, that will be equal to the present value x_n. So, the prediction error at time k can be evaluated as

$$e(k)=x_n(k)-x_e(k) \tag{5}$$

which is used for the learning rule to adjust the NN weights.

The coefficients of the nonlinear filter are adjusted on-line in the learning process, by considering a criterion that modifies at each pass of the time series the number of patterns, the number of iterations and the length of the tapped-delay line, in function of the Hurst's value H calculated from the time series. According to the stochastic behavior of the series, H can be greater or smaller than 0.5, which means that the series tends to present long or short term dependence, respectively (Pucheta *et al*, 2007b).

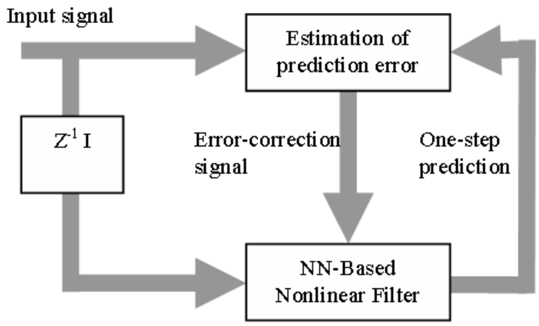


Fig.3. Block diagram of the nonlinear prediction

3 PROPOSED APPROACH

3.1NN-Based NAR Model

Some results had been obtained from a linear autoregressive approach, which are detailed on (Pucheta *et al*, 2009a). These results were promising and deserve to be improved by more sophisticated filters. Here, a NN-based NAR filter model (Haykin, 1999; Mozer, 1993; Zhang *et al*, 1998) is proposed. The NN used is a time lagged feed-forward networks type. The NN topology consists of l_x inputs, one hidden layer of H_o neurons, and one output neuron as shown Fig. 4. The learning rule used in the learning process is based on the Levenberg-Marquardt method (Bishop, 1995).

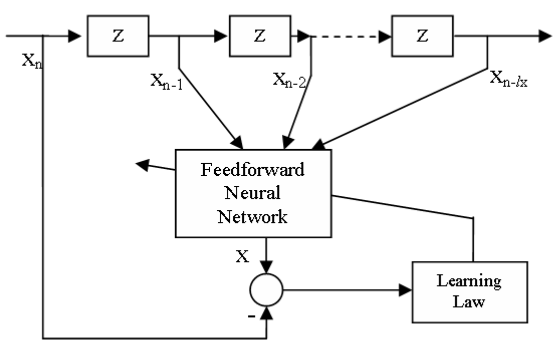


Fig. 4. Neural Network-based nonlinear predictor filter. The one-step delay operator is denoted by Z

However, if the time series is smoother or rougher then the tuning algorithm may change in order to fit the time series. So, the learning rule modifies the number of patterns and the number of iterations at each time-stage according to the Hurst's parameter *H*, which gives short or long term dependence of the sequence $\{x_n\}$. From a practical standpoint, it gives the roughness of the time series.

In order to predict the sequence $\{x_e\}$ one-step ahead, the first delay is taken from the tapped-line x_n and used as input. Therefore, the output prediction can be denoted by

$$x_e(n+1) = F_p\left(Z^{-1}I(\{x_n\})\right) \tag{6}$$

where, F_p is the nonlinear predictor filter operator, and $x_e(n+1)$ the output prediction at $n+1$.

3.2 Proposed Learning Process

The NN's weights are tuned by means of the Levenberg-Marquardt rule, which considers the long or short term stochastic dependence of the time series measured by the Hurst's parameter H. The proposed learning approach consists on changing the number of patterns, the filter's length and the number of iterations in function of the parameter H for each corresponding time series. The learning process is performed using a batch model. In this case the weight updating is performed after the presentation of all training examples, which forms an epoch. The pairs of the used input-output patterns are

$$(x_i, y_i) \quad \mathrm{i} = 1,2,\ldots,N_p \tag{7}$$

where, x_i and y_i are the corresponding input and output pattern respectively, and N_p is the number of input-output patterns presented at each epoch. Here, the input vector is define as

$$X_i = \mathrm{Z}^{-1}\mathrm{I}(\{x_i\}) \tag{8}$$

and its corresponding output vector as

$$Y_i = x_i. \tag{9}$$

Furthermore, the index *i* is within the range of N_p given by

$$l_x \le N_p \le 2 \cdot l_x \tag{10}$$

where l_x is the dimension of the input vector.

In addition, through each epoch the number of iterations performed i_t is given by

$$l_x \le i_t \le 2 \cdot l_x. \quad (11)$$

The proposed criterion to modify the pair (i_t, N_p) is given by the statistical dependence of the time series $\{x_n\}$, supposing that is a fBm. The dependence is evaluated by the Hurst's parameter H, which is computed using a wavelet-based method (Abry *et al*, 2003; Flandrin, 1992).

Thus, a heuristic adjustment for the pair (i_t, N_p) in function of H according to the membership functions shown in

Fig. Fig.5 is proposed.

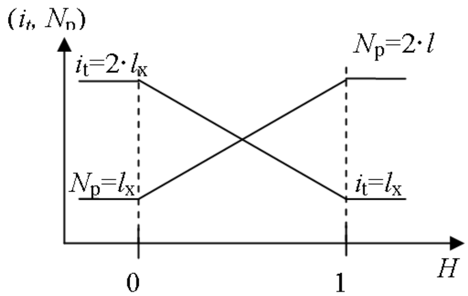


Fig.5. Heuristic adjustment of (i_t, N_p) in terms of H after each epoch

Finally, after each pass the number of inputs of the nonlinear filter is tuned —that is the length of tapped-delay line, according to the following heuristic criterion. After the training process is completed, both sequences —$\{x_n\}$ and $\{\{x_n\},\{x_e\}\}$, should have the same H parameter. If the error between $H(\{x_n\})$ and $H(\{\{x_n\},\{x_e\}\})$ is grater than a threshold parameter θ the value of l_x is increased (or decreased), according to $l_x \pm 1$. Explicitly,

$$l_x = l_x + sign(\theta). \quad (12)$$

Here, the threshold θ was set about 1%.

4 Main results

4.1 Generations of the data series from MG equations and cumulative rainfall

Data time series are obtained from the MG equations (1) with the parameters shown in Table 1, by fixing $\tau=100$ and $\alpha=20$. This collection of coefficient was chosen for generating time series whose H parameters varies between 0 and 1.

In addition, two time series extracted from meteorological data are used to forecast. Such data corresponds to monthly cumulative rainfall data measured at La Perla (-31.4309, -64.3322) and Santa Francisca (-31.8670, -64.3655) Córdoba, Argentina. The criterion for selecting these particular time series was the roughness, so the observation time for each time series does not match. Note that the observation period of La Perla time series comprises from January 1962 to December 1971, whereas the associated to Santa Francisca comprises from January 2000 to May 2010.

Table 1. Parameters for generating the times series

Time Series No.	Type	Characteristic	H
1	MG	β =0.32	0.96
2	MG	β =0.8	0.47
3	MG	β =0.85	0.262
4	Meteorological	La Perla	0.1353
5	Meteorological	Santa Francisca	0.0205

4.2 Set-up of Model

The initial conditions for the filter and learning algorithm are shown in Table 2. Note that the first number of hidden neurons and iteration are set in function of the input number. These conditions of the learning algorithm were used for forecasting the time series, whose sizes have a length of 102 values each.

Table 2. Initial condition of the parameters

Variable	Initial Condition
l_x	7
H_o	9
i_t	105
H	0.5

4.3 Measure of the Performance

In order to test the proposed design procedure of the NN -based nonlinear predictor, an experiment with time series obtained from the MG solution and cumulative Rainfall time series was performed.

The performance of the filter is evaluated using the Symmetric Mean Absolute Percent Error SMAPE) proposed in the most of metric evaluation, defined by

$$SMAPE_s = \frac{1}{n}\sum_{t=1}^{n}\frac{|X_t - F_t|}{(X_t + F_t + 0.005)/2}\cdot 100 \qquad (13)$$

where, t is the time observation, n is the test set size, s each time series, Xt and Ft are the actual and the forecast time series values at time t respectively. The SMAPE of each series s calculates the symmetric absolute error in percent between the actual Xt and its corresponding forecast Ft value, across all observations t of the test set of size n for each time series s.

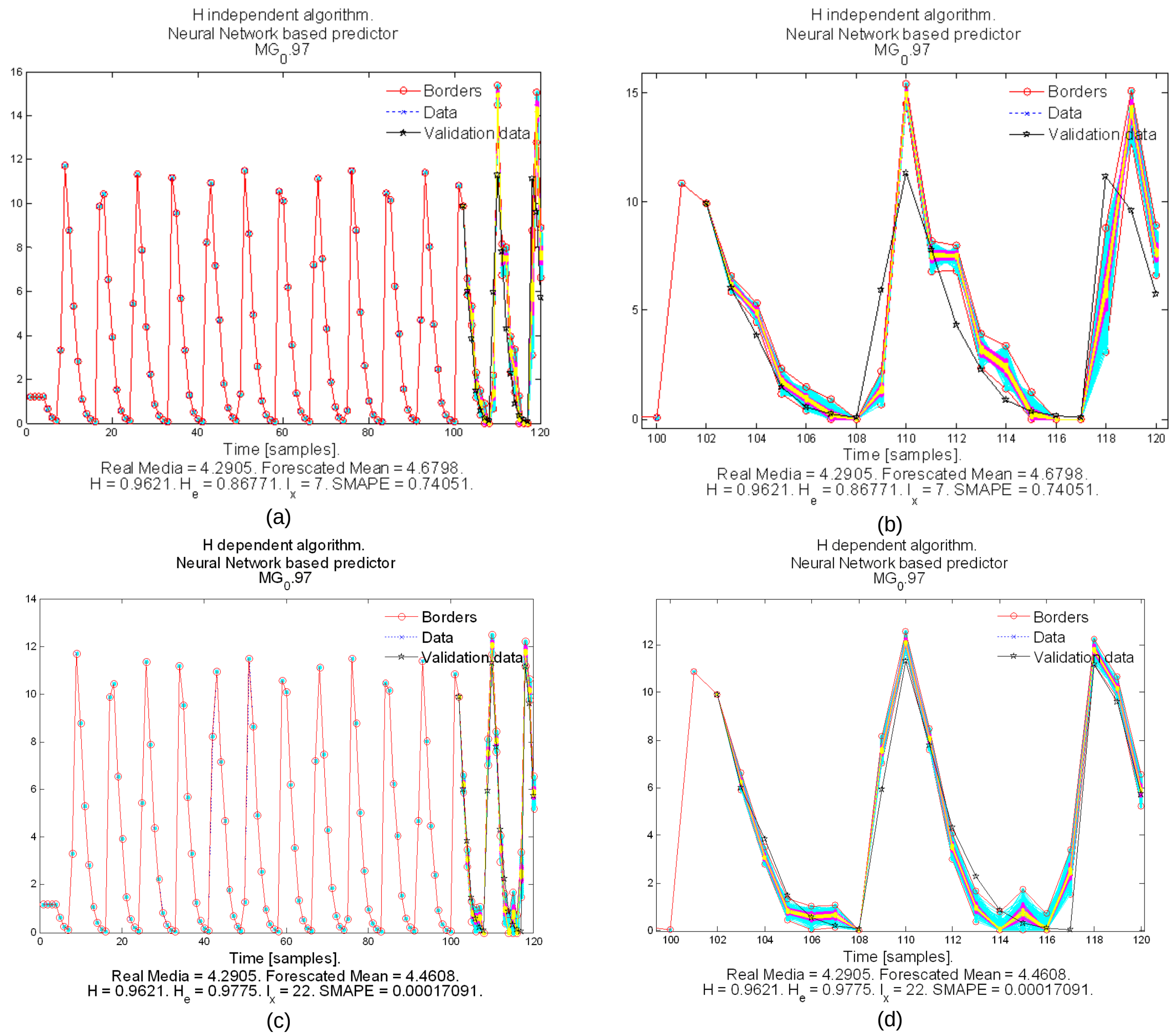


Fig.6. Forecast for MG time series No 1 from Table 1

4.4 Forecasting Results

Each time series are composed either by sampling the MG solutions or by performing the monthly cumulative rainfall from two geographical sites. However, there are two classes of data sets employed. One is used for the algorithm in order to give the forecast one step ahead used to compare whether the forecast is acceptable or not in which the 18 last values can be used to validate the performance of the prediction system.

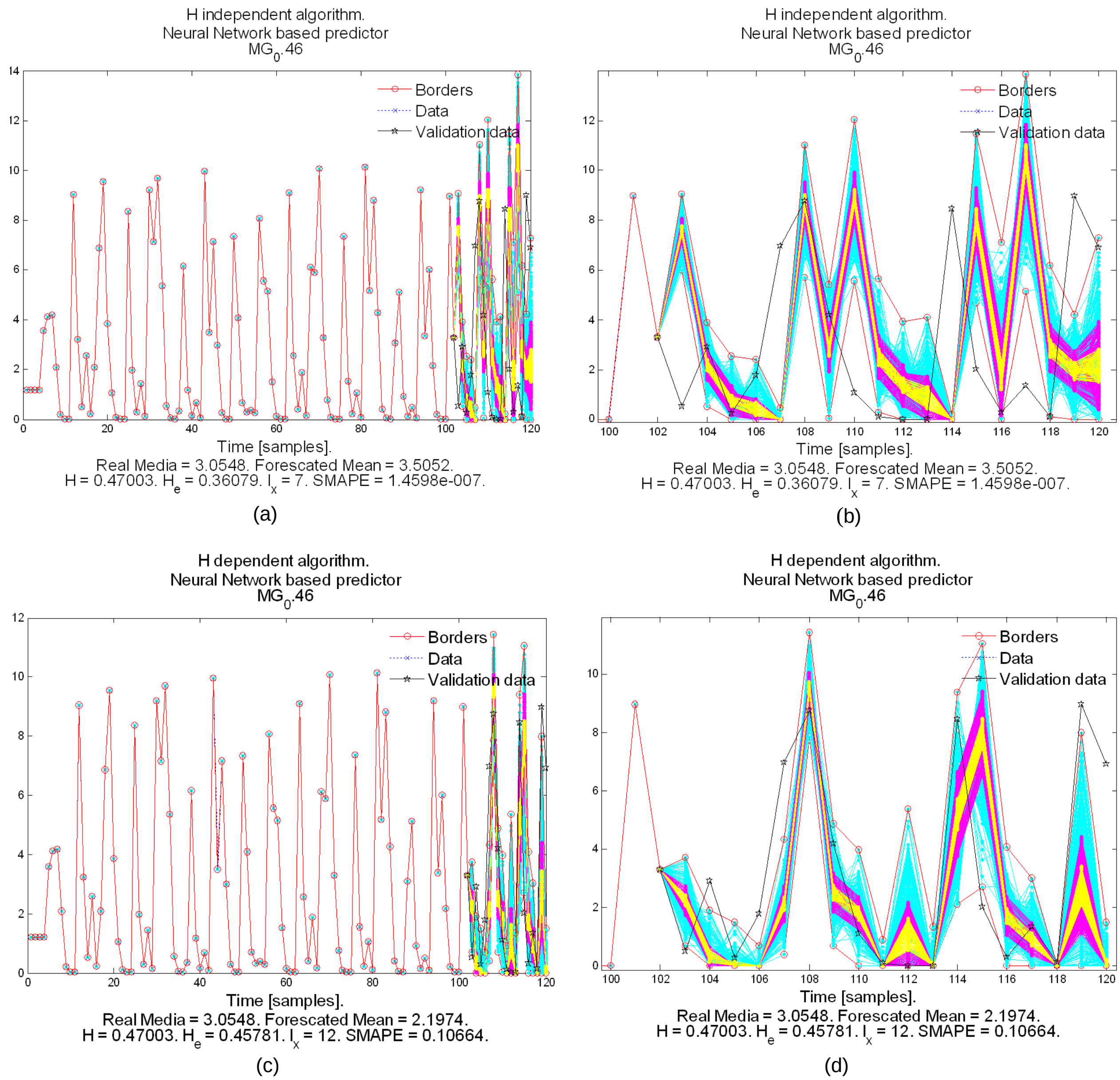


Fig.7. Forecast for MG time series No 2 from Table 1

The Monte Carlo method was employed to forecast the next 18 values with an associated variance. Here it was performed an ensemble of 500 trials with a fractional Gaussian noise sequence of zero mean and variance of 0.11.

The fractional noise was generated by the Hosking method (Dieker, 2004) with the H parameter estimated from the data time series.

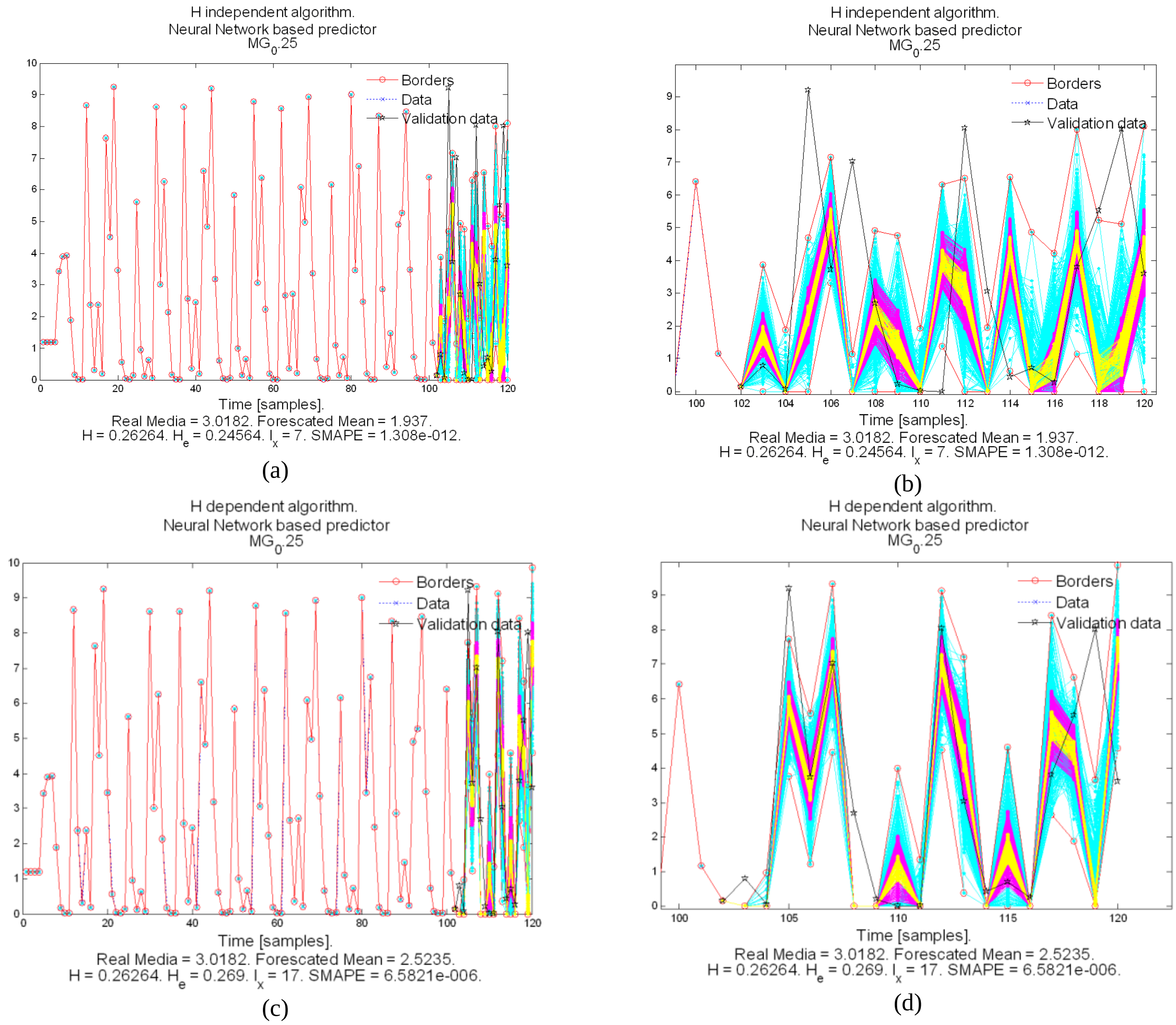


Fig.8. Forecast for MG time series No 3 from Table 1

The following figures yield the results of the mean and the variance of 500 trials of the forecasted 18 values. Such outcomes for one (30%) and two (69%) sigma are shown in Fig. 6, Fig. 7, Fig. 8, Fig. 10, and Fig. 11 for each case detailed in Table 1, respectively. Each figure shows the performance of an invariant NN predictor filter and the one of the H dependent filter proposed here. In Fig.10.a the estimated *H* for the forecasted time series results negative. This given value is meaningless because of the fact that *H* varies between 0 and 1 and the forecasted time series results too rough for the estimation algorithm (Abry *et al*, 2003; Flandrin, 1992). In each figure, the legend “Forecasted” denotes the values obtained by Eq. (6), the legend “Data” denotes the available data set, and the legend “Real” denotes de actual values (not available experimentally) used here for verification purposes only. The obtained time series has a mean value, denoted at the foot of the figure by “Forecasted Mean”, whereas the “Real Mean” although it is not available at time 102. This procedure is repeated 500 times for each time series.

4.5 Comparative Results

The performance of the stochastic NN-based predictor filter is evaluated through the SMAPE index —Eq. (13), shown in Table 3 along the time series from MG solutions and Monthly Cumulative Rainfall time series. The comparison between both the deterministic approach (Pucheta *et al*, 2009b) and the present forecasted time series is shown in Fig. 9. The SMAPE index for each time series is obtained by a deterministic NN-based filter, which uses the Levenberg–Marquardt algorithm with fixed parameters (i_t, N_p). This result gotten through the predictor filter, an ensemble of 500 trials were performed between five sets of time sires in order to give the mean value of the times series, in which it is defined an upper and bottom range of percentage of possibility where the predicted value is validated as useful value. In addition, the figures of the SMAPE obtained by the stochastic NN-based filter proposed here are also shown in Fig. 9. Thus, the legend “Traditional” refers to the first filter and the legend “Modified” refers to the second one.

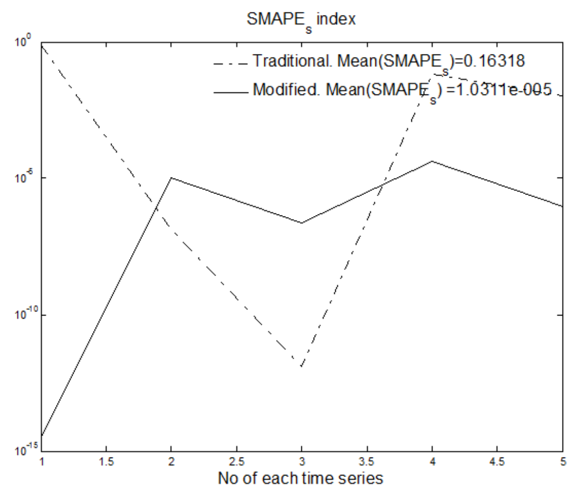


Fig.9. The SMAPE index applied over the five time series, by each tuning algorithm

Table 3. Figures obtained by the proposed approach and the traditional algorithm

			Modified			Traditional		
Series No.	**H**	**Real mean**	**H_e**	**Mean Forecasted**	**SMAPE**	**H_e**	**Mean Forecasted**	**SMAPE**
1	0.9621	0.214	0.857	0.184	$3.388\ 10^{-15}$	0.8677	4.679	0.7405
2	0.47	0.152	0.339	0.18	$1.0411\ 10^{-5}$	0.3607	3.505	$1.4598\ 10^{-7}$
3	0.262	0.15	0.258	0.123	$2.34\ 10^{-7}$	0.2456	1.937	$1.308\ 10^{-12}$
4	0.13538	50.84	0.139	46.28	$4\ 10^{-5}$	-0.108	104.52	0.0655
5	0.02055	65.21	0.088	80.7	$9.1\ 10^{-7}$	0.073412	63.633	0.0099

In Fig. 9 the value of SMAPEs is indicated for both filters. It can be noted the improvement since the SMAPE index diminishes from 0.16318 to 0.000010311 which means an improvement of fourth order, averaging over the five time series.

5 Discussion

The assessment of the experimental results has been obtained by comparing the performance of the proposed filter against the classic filter, both based on NN. Although the difference between both filters only resides in the adjustment algorithm, the coefficients each filter has, they performed different behaviors. In the five analyzed cases, the generation of 18 future values was made by each algorithm. The same initial parameters were used for each algorithm, although such parameters and filter structure are changed by the proposed algorithm, they are not modified by the classic algorithm. In the adjustment algorithm for the proposed filter, the coefficients and the structure of the filter are tuned by considering their stochastic dependency. It can be noted that in each of the figures — Fig. 6 to Fig.11— the computed value of the Hurst parameter is denoted by H_e or H when it is taken either from the Forecasted time series or from the Data series, respectively, since the Real (future time series) are unknown. Index SMAPE is computed between the complete Real series (it includes the series Data) and the Forecasted one, as indicates the Ec. (13) for each filter. Note that the forecast improvement is not over any given time series, which results from the use of a stochastic characteristic for generates a deterministic result, such as a prediction.

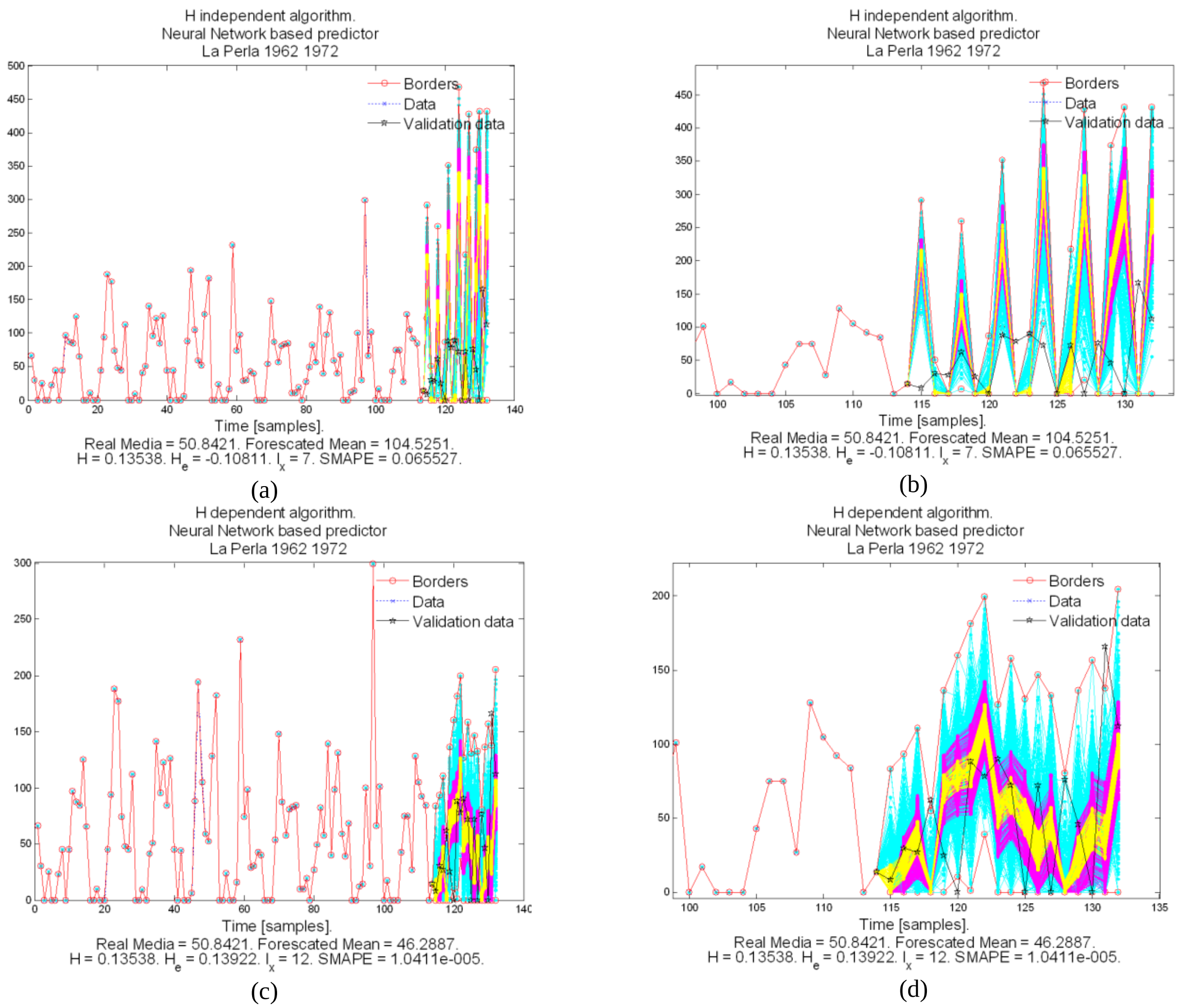


Fig.10. Forecast for Monthly cumulative rainfall time series No 4 from Table 1

6 Conclusions

In this work a feed-forward neural networks based nonlinear autoregressive (NAR) model for forecasting time series was presented. In addition, during forecasting stage the Monte Carlo simulation method with a fractional Gaussian noise was implemented. The learning rule proposed to adjust the NN's weights is based on the Levenberg-Marquardt method. Furthermore, in function of the long or short term stochastic dependence of the time series assessed by the Hurst parameter *H*, an heuristic adaptive law was proposed to update the NN topology at each time-stage, which is the number of input taps, the number of patterns and the number of iterations. The main result shows a good performance of the predictor system applied to time series from several benchmark of MG equations and Monthly Cumulative Rainfall time series, due to similar roughness between the original and the stochastic forecasted time series, evaluated by H and H_e, respectively. This fact encourages us to apply the proposed approach to meteorological time series when observations are taken from a single standpoint.

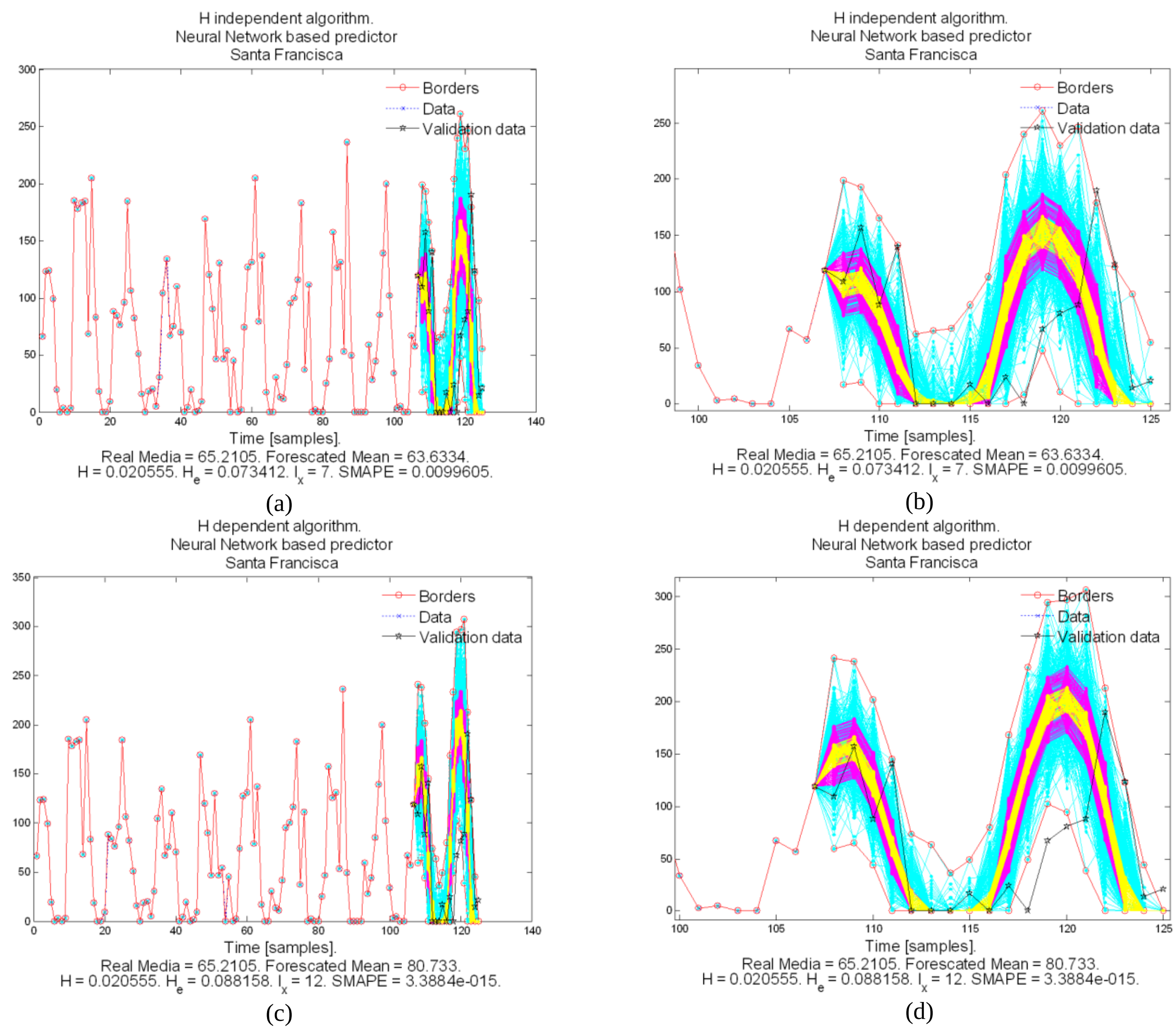


Fig.11. Forecast for Monthly cumulative rainfall time series No 5 from Table

Acknowledgments

This paper was supported by National University of Córdoba (Secyt UNC 69/08), National University of San Juan (UNSJ), National Agency for Scientific and Technological Promotion (ANPCyT) under grant PAV 076, PICT/04 No. 21592 and PICT-2007-00526. The authors want to thank Carlos Bossio (Coop. Huinca Renancó), Ronald del Águila (LIADE) and Eduardo Carreño (Santa Francisca) for their help.

References

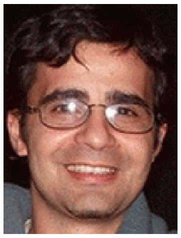
1. **Abry, P., Flandrin, P., Taqqu, M.S. & Veitch, D. (2003).** Self-similarity and long-range dependence through the wavelet lens In P. Doukhan, G. Oppenheim & M. Taqqu (Eds.), *Theory and applications of long-range dependence* (527-556**).** Boston: Birkhäuser.
2. **Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S. & Taqqu, M.S. (2003).** Semi-parametric estimation of the long-range dependence parameter: a survey. In P. Doukhan, G. Oppenheim & M. Taqqu (Eds.), *Theory and applications of long-range dependence.* (557-577) Boston, MA: Birkhäuser.
3. **Bishop, C.M. (1995).** *Neural Networks for Pattern Recognition*. New York: Oxford University Press.
4. **Bishop, C.M. (2006).** *Pattern Recognition and Machine Learning*. New York: Springer.
5. **Dieker, T. (2004).** *Simulation of fractional Brownian motion*. MSc thesis, University of Twente, Amsterdam, The Netherlands.
6. **Flandrin, P. (1992).** Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Transactions on Information Theory,* 38(2), 910-917.
7. **Glass L., & Mackey M.C. (1998).** *From Clocks to Chaos, The Rhythms of Life*. Princeton, NJ: Princeton University Press.
8. **Guo, W.W., Li L.D. & Whymark G. (2009).** Statistics and neural networks for approaching nonlinear relations between wheat plantation and production in queensland of Australia. *IEEE International Conference on Industrial Technology,* Gippsland, Australia, 1-4.
9. **Haykin, S. (1999).** *Neural Networks: A comprehensive Foundation (2nd Ed)*. Upper Saddle River, N.J.: Prentice Hall.
10. **Istas, J. & G. Lang. (1997).** Quadratic variations and estimation of the local Hölder index of a Gaussian process. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, v 33 (4), 407-436.
11. **Liu, J.N.K. & Lee, R.S.T (1999).** Rainfall forecasting from multiple point sources using neural networks. *International Conference on Systems, Man, and Cybernetics,* Tokyo, Japan, Vol 3, 429-434.
12. **Mandelbrot, B.B. (1983).** *The Fractal Geometry of Nature*. New York: W.H. Freeman.
13. **Masulli, F., Baratta, D., Cicione, G. & Studer, L. (2001).** Daily rainfall forecasting using an ensemble technique based on singular spectrum analysis. *International Joint Conference on Neural Networks*, Washington, D.C, USA, 263-268.
14. **Mozer, M.C. (1994).** Neural net architectures for temporal sequence processing. In A. S. Weigend & N. A. Gershenfeld (Eds.), *Time series prediction: Forecasting the future and understanding the past* (243–264). Redwood Reading, MA: Addison-Wesley Publishing.

Julián Antonio Pucheta

Received the Electrical Engineering degree from National Technological University - Cordoba Regional Faculty, Argentina, the M.S. and Ph.D. degrees from National University of San Juan, Argentina, in 1999, 2002 and 2006, respectively. He is currently Professor at the Laboratory of research in mathematics applied to control, National University of Cordoba, Argentina. His research interests include stochastic and optimal control, time series forecast and machine learning.

Cristian Rodriguez Rivero

Received the Electrical Engineering degree from Faculty of Exact, Physical and Natural Sciences at National University of Cordoba, Argentina, in 2007, and he is currently pursuing the Ph.D. degree in electrical engineering at University of Cordoba, Argentina, under the supervision of Dr. Pucheta and Grant PRH National Agency for Scientific and Technological Promotion (ANPCyT), in the field of automatic control for slow dynamics processes, stochastic control, optimization and time series forecast. He joined the Laboratory of Research in Applied Mathematics to Control in 2009.



Rafael Martín Herrera

Received the Electrical Engineering degree in 2007 from Faculty of Exact, Physical and Natural Sciences at National University of Cordoba, Argentina. He is Associated Professor and researcher from Faculty of Applied Sciences and Technology at National University of Catamarca. His research interests include control systems and automation.

Carlos Alberto Salas

Received the Electrical Engineering degree in 1995 from Faculty of Exact, Physical and Natural Sciences at National University of Cordoba, Argentina. He is researcher Professor at Faculty of Technology, National University of Catamarca. His interests include automation and industrial process control.

H. Daniel Patiño

Received the Electrical Engineer and Ph.D. degrees from National University of Juan, Argentina, in 1986 and 1995 respectively. He is currently Professor at Institute of Automatic, Faculty of Engineering, National University of San Juan, Argentina and Head of Laboratory of Advanced Intelligent Systems. His research interests include applied computational intelligence, adaptive neuro-dynamic programming, aerial robotics and advanced avionics.

Benjamín Rafael Kuchen

Received the Electrical Engineer degree from Catholic University of Cordoba, Cordoba, Argentina, and the Ph.D. degree from Rheinisch-Westfälischen Technischen Hochscule Aachen (RWTH Aachen), Germany, in 1968 and 1974, respectively. He is currently Professor and Principal at National University of San Juan. He is also an independent researcher of National Council for scientific and technical research of Argentina (CONICET). His interests include automation, robotics and industrial process control.