A New Approach for Time Series Forecasting: Bayesian Enhanced by Fractional Brownian Motion with Application to Rainfall Series

Cristian Rodriguez Rivero
LIMAC – Department of Electronic Engineering
Universidad Nacional de Córdoba
Córdoba, Argentina

Julian Pucheta
LIMAC – Department of Electronic Engineering
Universidad Nacional de Córdoba
Córdoba, Argentina

Daniel Patiño
INAUT – Institute of Automatics
Universidad Nacional de San Juan
Cordoba, Argentina

Victor Sauchelli
LIMAC – Department of Electronic Engineering
Universidad Nacional de Córdoba
Córdoba, Argentina

Abstract—**A new predictor algorithm based on Bayesian enhanced approach (BEA) for long-term chaotic time series using artificial neural networks (ANN) is presented. The technique based on stochastic models uses Bayesian inference by means of Fractional Brownian Motion as model data and Beta model as prior information. However, the need of experimental data for specifying and estimating causal models has not changed. Indeed, Bayes method provides another way to incorporate prior knowledge in forecasting models; the simplest representations of prior knowledge in forecasting models are hard to beat in many forecasting situations, either because prior knowledge is insufficient to improve on models or because prior knowledge leads to the conclusion that the situation is stable.**

This work contributes with long-term time series prediction, to give forecast horizons up to 18 steps ahead. Thus, the forecasted values and validation data are presented by solutions of benchmark chaotic series such as Mackey-Glass, Lorenz, Henon, Logistic, Rössler, Ikeda, Quadratic one-dimensional map series and monthly cumulative rainfall collected from Despeñaderos, Cordoba, Argentina. The computational results are evaluated against several non-linear ANN predictors proposed before on high roughness series that shows a better performance of Bayesian Enhanced approach in long-term forecasting.

Keywords—long-term prediction; neural networks; Bayesian inference; Fractional Brownian Motion; Hurst parameter

I. INTRODUCTION

Forecasting is based on identifying and estimating through observation and in some instances theory, patterns and/or relationships and then extrapolating or interpolating them in order to predict [1]. Scientists are different than other forecasters in being well aware of uncertainty, by providing probabilistic forecasts, and constantly searching for enhancements using objective feedback [2]. Their overall success rate is improving and their assessment of uncertainty is well calibrated. Finally, weather forecasters have learned that predicting extreme weather events requires different models and skills than those of normal ones [3].

At the same time, they consider such events as an integral part of their job, even though it requires special effort, different models and extra skills to predict them.

The natural phenomena where humans cannot influence their future course, except to a limited extent, do influence with their actions and reactions, changing their future course, making forecasting more difficult but also more challenging [4]. The question regarding why simple forecasting models outperform sophisticated ones is still open. The future is never exactly like the past which means that the accuracy of extrapolative predictions cannot be assured. The crucial question is the extent of accuracy, or inaccuracy of such predictions. Most of the time series in rainfall forecast are influenced by random events and often behave not far from random walks, favoring simple methods that are capable of smoothing such randomness. In long-term forecasts, the accuracy of predictions drops while uncertainty increases. On way used by forecaster is reducing the forecasting errors of predictions by averaging more than one model [5]. The outcome is not only higher accuracy but also a reduction in the size of forecasting errors, with simple averaging being the best way of combining forecasts [6]. The reason is that averaging cancels out the errors of individuals and/or models and in doing so eliminates the noise from the pattern and improve accuracy.

Rossi, Allenby, and McCulloch (2006) argued that there are really no other approaches except the Bayesian approach which can provide a unified treatment of inference and decision as well as properly accounting for parameter and model uncertainty. The Bayesian approach allows researchers to cope with complex problems. However, the Bayesian inference provides answers conditional on the observed data rather than based on distribution of test statistics over imaginary samples not observed. Even though the Bayesian approach has decent benefits, it has some trivial costs including formulation of prior, requirement of a likelihood function, and computation of various integrals required in Bayesian paradigm [7]. The

advancement of computation make complicated integrals become possible. However, choosing an appropriate or objective prior has been an issue in the Bayesian approach [8]. Thus, investigators are facing a practical problem with little information in the real-world situations and should not neglect sources of information outside of the current data set [9].

In this article, the major advantage of the proposed BEA technique is that the complexity does not increase with an increasing number of inputs. The solutions can easily be generalized to the problem of uncertain (noisy) inputs, such as Bayesian inference [10] against other generalized approaches [11]. Here the filters in comparison are based on non-linear stochastic auto-regressive moving average (NAR) models such as Bayesian approach [12] and Neural-Network Modified [13] [14], implemented by ANN.

The paper is organized as follows: Section II presents the data series as an important case of study, such as well-known chaotic time series. Section III provides the Bayesian Enhanced approach as a method using fractional Brownian motion for obtaining optimal network model. In Section IV, the proposed prediction method is highlighted by showing the performance of the proposed algorithm detailing experimental setup, results and analysis, with Section V providing some discussions and concluding remarks.

II. DATA TREATMENT

A. Overview on fBm

The fractional Brownian motion, which provides a suitable generalization of the Brownian motion, is one of the simplest stochastic processes exhibiting long-range dependence [15] [16]. It has been used as a modeling tool. The following demonstrates the stochastic integral representation of fractional Brownian motion [17]. The process is as follows,

$$B(t) = \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} \int \left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right) dB(s)$$
(1)

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

where $B(t)$ is a standard Brownian motion and Γ refers to the gamma function, is a *fBm* with 0<H<1. The constant $1/\Gamma(H+1/2)$ in the following computation is dropped for the sake of simplicity. According to the definition, a fractional Brownian motion (B(H)(t)) t>0 of Hurst parameter H is a continuous and centered Gaussian process with covariance function,

$$B^{H}(s) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$
(2)

Therefore, $B(t)$ is a fBm of Hurst index *H*. The fBm is divided into three different families corresponding to 0<*H*<1/2, *H*=1/2, 1/2<*H*<1, respectively. The basic feature of *fBm* is that the span of independence between their increments can be infinite [18]. As the Hurst parameter *H* governs the fractal

dimension of the fractional Brownian motion, its regularity and the long-memory behavior of its increments, the estimation of H is an important but difficult task which has led to very vast literature [19].

In this work, the H index is measured by wavelet method [20] [21].

B. Overview on Benchmark Time Series

The standard non-parametric approaches presented in this article are based on stochastic techniques that assume non-linear relationship among data that reproduce the benchmark chaotic time series and rainfall data only in statistical sense. Although there are many situations when accurate forecasting is impossible, there are many others where predictions can provide useful information to improve our decisions and gain from effective action. Weather forecasts, made several times a day, in hundreds of thousands of locations around the world, are an example, as it is proposed in this work. The rainfall dataset used is from Despeñaderos located at Cordoba, province of Argentina (-31.824703;-64.289692) and the collection date is from year 2000 to 2014 as shown in Fig.1.
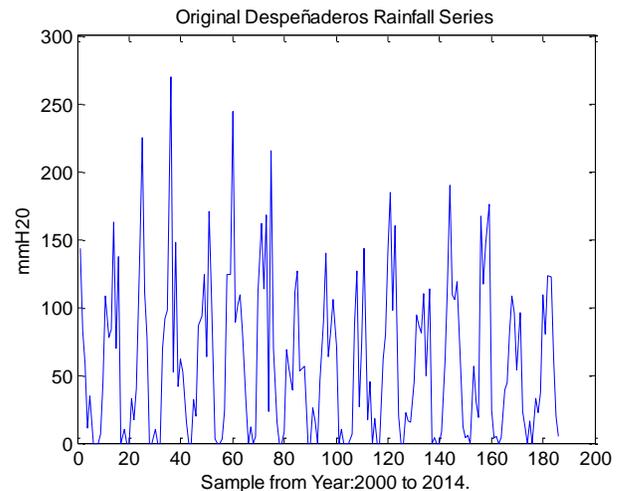
Fig. 1. Original Rainfall times series from Despeñaderos, Cordoba, Argentina

The rest of the benchmark time series are presented in the following subsections.

C. The Mackay-Glass Chaotic Time Series

The dataset ensemble is by sampling the Mackay-Glass (MG) equations [22] defined by:

$$\dot{x}(t) = \frac{ax(t-\tau)}{1 - x(t-\tau)^c} - bx(t)$$
(3)

with *a, b, c*, τ setting parameters shown as follows in Table I.

TABLE I. PARAMETERS TO GENERATE MG TIME SERIES

Series No.	β	α	c	τ	H
MG17	0.2	0.1	10	17	0.980
MG30	0.2	0.1	10	30	1

D. The Logistic Chaotic Time Series

The logistic series (LOG) is defined by:

$$x(t+1) = ax(t)[1-x(t)]. \qquad (4)$$

When a=4, the iterates of (4) perform a chaotic time series [23].

TABLE II.   PARAMETERS TO GENERATE LOG TIME SERIES

Series No.	N	a	X_0	H
LOG01	50	4	0.1	0.151
LOG03	50	4	0.3	0.100

E. The Henon Chaotic Time Series

The Henon chaotic time series can be constructed by following (5), however, it presents many aspects of dynamical behavior of more complicated chaotic systems [24].

$$x(t+1) = b+1-ax^2(t) \qquad (5)$$

When generating data for our experiments, a and b are set as shown in Table III. These same parameters are used in both [25].

TABLE III.   PARAMETERS TO GENERATE HEN TIME SERIES

Series No.	N	a	b	X_0	Y_0	H
HEN01	120	1.4	0.3	0	0	0.187
HEN03	120	1.3	0.22	0	0	0.833

F. The Lorenz Chaotic Time Series

Lorenz found three ordinary differential equations which closely approximate a model for thermal convection [26]. These equations have also become a popular benchmark for testing non-linear predictors. The Lorenz model is given by the equations (4), the data is derived from the Lorenz system, which is given by three time-delay differential systems

$$\begin{cases} \dfrac{dx}{dt} = a(y-x), \\ \dfrac{dy}{dt} = bx - y - xz, \\ \dfrac{dz}{dt} = xy - bz \end{cases} \qquad (4)$$

A typical choice for the parameter values are as a = 10, b = 28, and c = 8/3. In this case, the system is chaotic. The data set is constructed by using four-order Runge–Kutta method with the initial value as is shown in Table IV for LOR01 and LOR03 series. The step size is chosen as 0.01, respectively. These sets of parameters are commonly used in generating the Lorenz system because exhibits deterministic chaos.

TABLE IV.   PARAMETERS TO GENERATE ROS TIME SERIES

Series No.	n	X(0)	Y(0)	Z(0)	H
LOR01	120	12	9	2	0.158
LOR03	120	0.1	0.1	2	0.080

G. The Rössler Chaotic Time Series

In this example, the data is derived from the Rössler system [27], which is given by three time-delay differential systems. The data set is constructed by using four-order Runge–Kutta method with the initial value as shown in Table V, and the step size is chosen as 0.01, respectively.

$$\begin{cases} \dfrac{dx}{dx} = -y(t) - z(t), \\ \dfrac{dy}{dx} = x(t) + ay(t), \\ \dfrac{dz}{dx} = b + z(t)(x(t) - c) \end{cases} \qquad (4)$$

TABLE V.   PARAMETERS TO GENERATE LOR TIME SERIES

Series No.	n	a	b	c	X(0)	Y(0)	Z(0)	H
ROS01	120	0.2	0.2	5.7	12	9	2	0.203
ROS03	120	0.42	0.42	0.42	0.1	0.1	2	0.274

H. The Ikeda Chaotic Time Series

Before describing the reconstruction, I introduce the system which will be used to generate most of the time series described herein, namely the Ikeda map [28]. The Ikeda map is given as follows:

$$\begin{aligned} \dot{x} &= 1 - \mu[x\cos(t) - y\sin(t)], \\ \dot{y} &= \mu[x\cos(t) + y\sin(t)] \end{aligned} \qquad (7)$$

where $t = 1/(1 + x^2 + y^2)$. This system displays chaotic behavior over a range of values for the parameter μ including the values chosen here.

TABLE VI.   PARAMETERS TO GENERATE IK TIME SERIES

Series No.	n	μ	X(0)	Y(0)	H
IK01	120	0.9	0.5	0.5	0.029
IK03	120	0.8	0.9	0.6	0.060

I. The Quadratic Chaotic Time Series

The quadratic map is defined by the equation

$$x(t+1) = \mu x(t)(1 - x(t)) \qquad (7)$$

If this mapping is iterated by μ=4, starting with a random number in the interval between 0 and 1, then different behavior is obtained dramatically depending upon the initial value of *x*. Initial values of *x* which are quite close together can have dramatically different iterates.

This unpredictability or sensitive dependence on initial conditions is a property familiar in displaying chaotic behavior [29] over a range of values for the parameter μ including the values chosen here.

TABLE VII.   PARAMETERS TO GENERATE QUA TIME SERIES

Series No.	n	c	X(0)	H
QUA01	120	1.95	0.5	0.31
QUA03	120	0.8	0.9	0.011

120 samples are for the selected time series, the first 102 values are used for training and the remaining 18 values are kept for validation and test data. The long-term behavior changes thoroughly by changing the initial conditions to obtain the stochastic dependence of the deterministic time series according to its roughness assessed by the H parameter. Then, extra 18 testing data are used to measure the prediction performance.

Fig. 2.   Bloxplot diagram of the benchmark chaotic time series

the information driven as likelihood function from fractional Brownian, provided by neural networks weights from observed data in order to generate point and interval forecasts by combining all the information and sources of uncertainty into a predictive distribution for the future values.

A weight vector *w* defines a mapping from an input vector *x* to a predicted output vector $\hat{y}$ given by $\hat{y} = f(x, w)$. Assuming a fractional Brownian model, the conditional probability distribution for the output given the input vector l_x is a follows:

$$P(D/w, \mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{R}|^{\frac{1}{2}}} \exp\left(-\frac{\left(y(x_n; w) - t_n\right)^T \mathbf{R}\left(y(x_n; w) - t_n\right)}{2}\right), \quad (5)$$

$$P(t/H) = \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{R}|^{\frac{1}{2}}} \exp\left(-\frac{t^T \mathbf{R} t}{2}\right), \quad (6)$$

The application of the regression problem involving the correspond neural network function *y(x,w)* and the data set consisting of *N* pairs, input vector l_x and targets t_n (*n=1,....,N*).

To complete the Bayesian enhanced approach for this work [30], prior information for the network is required. The beta distribution is chosen for this purpose. The Beta density function is a very versatile way to represent outcomes like proportions or probabilities. It is defined on the continuum between 0 and 1. There are two parameters α and β which work together to determine if the distribution has a mode in the interior of the unit interval and whether it is symmetrical. This is a probability model which describes the knowledge gained after observing a set of data. It is proposed to use fractional Brownian, where *H* is the Hurst parameter, $t_n = (t_1, t_2, \ldots, t_n)$ follows a standard *fBm* and

$$R_{ij} = \frac{1}{2}\left(i^{2h} + j^{2h} - |i - j|^{2h}\right) \quad (7)$$

assuming that the expected scale of the weights is given by *w* set by hand. The Beta prior distribution for H is

$$P(h) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} h^{\alpha-1} (1-h)^{\beta-1}, \quad (8)$$

The full probability model is derived from the product,

$$P(w/h) = \frac{1}{(2\pi)^{\frac{n}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{w^T R w}{2}\right) \frac{\beta^\alpha}{\Gamma(\alpha)} h^{-\alpha-1} \exp\left(-\frac{\beta}{h}\right), \quad (9)$$

Then, the posterior distribution is as follows,

$$P(h/w) = \frac{\frac{1}{(2\pi)^{\frac{n}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{w^T R w}{2}\right) P(h)}{\int_0^1 \frac{1}{(2\pi)^{\frac{n}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{w^T R w}{2}\right) P(h)}, \quad (10)$$

Metropolis-Hasting algorithm was used for computation with a starting value of 0.1. The number of iterations was set to be 10,000. Monte Carlo Error was used to examine the convergence. This was carried out considering that the network function *f(xₙ+1,w)* is approximately linear with respect to *w* in the vicinity of this mode, in fact, the predictive distribution for y_{n+1} will be another multivariate Gaussian. This was carried out considering that the network function *f(xₙ+1,w)* is

III.   BAYESIAN ENHANCED APPROACH

This section presents a new method for issuing time series forecasting by focusing on three aspects: the formalization of one-step forecasting problems as supervised learning tasks, the discussion of modeling with Bayes inference techniques as an effective tool for dealing with temporal data and the key of the forecasting strategy when multiple-step-ahead is used for forecasting.

The increasing availability of large amounts of historical data and the need of performing accurate forecasting of future behavior in several scientific and applied domains demands the definition of robust and efficient techniques able to infer from observations the stochastic dependency between past and future. The forecasting domain has historically been influenced by linear statistical methods such as ARIMA models. More recently, machine learning models have drawn attention and have established themselves as serious contenders to classical statistical models in the forecasting community.

In this research, the Bayes assumption is used to update a prior distribution into a posterior distribution by incorporating

approximately linear with respect to *w* in the vicinity of this mode, in fact, the predictive distribution for y_{n+1} will be another multivariate Gaussian.

IV. PREDICTION RESULTS

The simulation results in different order approximations and time periods are presented in the following Table VIII. The performance of the comparison is measured by the Symmetric Mean Absolute Percent Error (SMAPE) and Root Mean Square Error (RMSE) proposed in the most of metric evaluations [31], defined by

$$SMAPE_s = \frac{1}{n}\sum_{t=1}^{n} \frac{|X_t - F_t|}{(|X_t| + |F_t|)/2} \cdot 100 \qquad (11)$$

$$RMSE_s = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(X_t - F_t)^2} \qquad (12)$$

where *t* is the observation time, *n* is the size of the test set, **s** is each time series, X_t and F_t are the actual and the forecasted time series values at time *t* respectively. The SMAPE and RMSE of each series **s** calculate the error in percent between the actual X_t and its corresponding forecast value F_t, across all observations *t* of the test set of size *n* for each time series s.
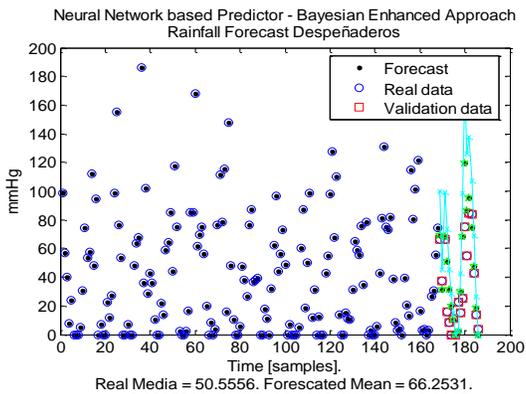
Fig. 3. Despeñaderos Rainfall time series Neural-Network Bayesian enhanced approach (BEA)
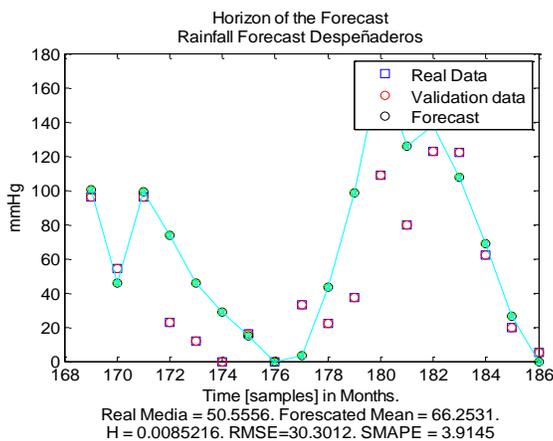
Fig. 4. Despeñaderos Rainfall time series Neural-Network Bayesian enhanced approach (BEA)
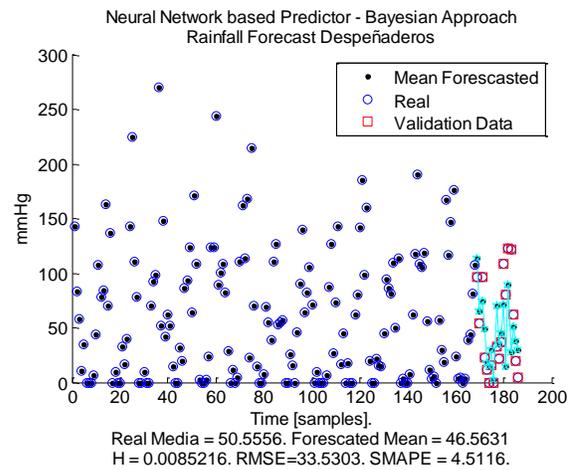
Fig. 5. Despeñaderos Rainfall time series Neural-Network Bayesian approach (BA)
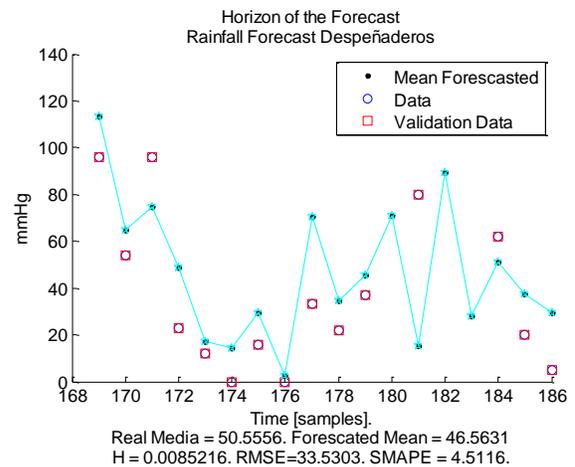
Fig. 6. Despeñaderos Rainfall time series Neural-Network Bayesian approach (BA)
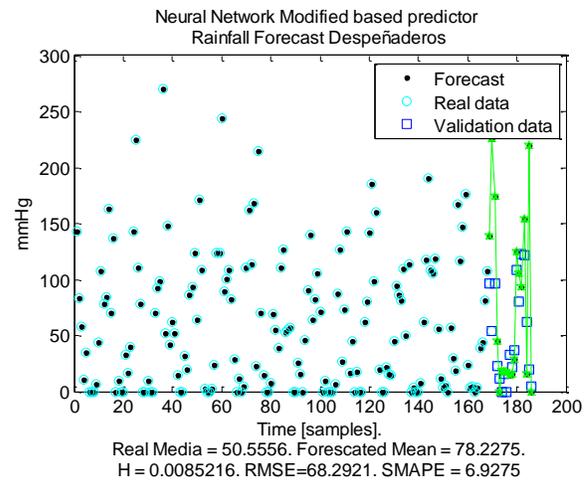
Fig. 7. Despeñaderos Rainfall time series Neural-Network Modified approach (NN-Mod)

Horizon of the Forecast
Rainfall Forecast Despeñaderos

Real Media = 50.5556. Forescated Mean = 78.2275.
H = 0.0085216. RMSE=68.2921. SMAPE = 6.9275
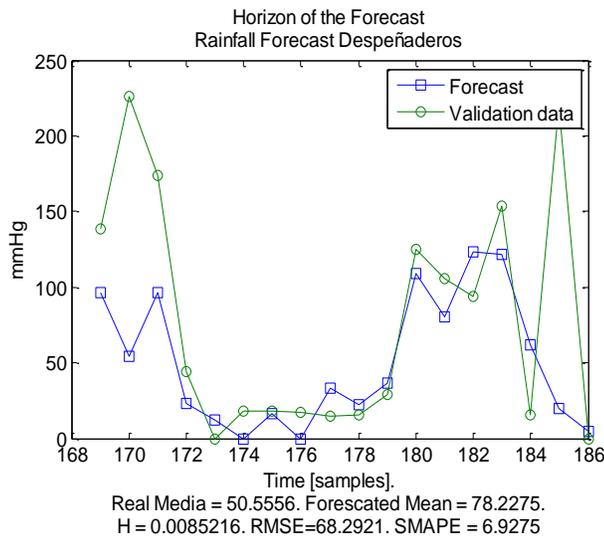
Fig. 8. Despeñaderos Rainfall time series Neural-Network Modified approach (NN-Mod)

The Monte Carlo method was used to forecast the next 18 values from benchmark chaotic and rainfall time series. Such outcomes are shown from Fig.3 up to Fig.8. Here, previous algorithms are used [10] [11] [13] [14] to compare the Bayesian enhanced approach.

Comparisons are preformed between the Bayesian and NAR models by using long-term time series; in this case 15 years of monthly rainfall data (2000-2014) served as the historical data to forecast 2015 and benchmark of chaotic time series proposed in the literature. The results in Table VIII show that Bayesian enhanced approach (BEA) were a bit superior for the lengthy time series, with a SMAPE and RMSE about one half that of the Bayesian approach (BA) and neural network-modified predictor filter (NN-Mod).

The results show that the performances of the BEA with BA and NN-Mod are better than those in term of SMAPE and RMSE, due to the existence of outliers. With this lengthy series, BEA could adequately detect the underlying the relationship among those correlated variables. The simulation results of the BEA methods are compared with the BA and NN-Mod summarized in Table VIII for the benchmark time series. The similarity of the trend of the prediction performance between them is clear, BEMA is slightly better particularly on rainfall time series with reference to BEA and BA approaches.

TABLE VIII.   RESULTS OF THE FORECASTING APPROCHES

Series	Method	Real Mean	Forecasted Mean	SMAPE	RMSE
Rainfall DES	BEA	50.55	66.25	3.91	30.30
	BA	50.55	46.56	4.51	33.53
	NN-Mod.	50.55	78.22	6.92	68.29
MG17	BEA	0.94	0.92	0.045	0.032
	BA	0.94	0.91	0.085	0.036
	NN-Mod.	0.94	0.89	0.32	0.047
MG30	BEA	0.87	0.85	0.027	0.09
	BA	0.87	0.91	0.11	0.017
	NN-Mod.	0.87	0.80	0.32	0.038
ROS01	BEA	0.11	0.10	0.009	0.05
	BA	0.11	0.14	0.017	0.10
	NN-Mod.	0.11	0.20	0.025	0.14
ROS03	BEA	0.0014	0.0026	0.00005	0.0003
	BA	0.0014	0.0030	0.00014	0.0014
	NN-Mod.	0.0014	0.0026	0.00031	0.0018
IK01	BEA	-0.44	-0.28	0.017	0.125
	BA	-0.44	0.008	0.105	0.29
	NN-Mod.	-0.44	0.012	0.136	0.99
IK03	BEA	-0.23	0.027	0.016	0.101
	BA	-0.23	0.053	0.021	0.121
	NN-Mod.	-0.23	0.017	0.159	0.93
LOR01	BEA	33.37	30.37	0.12	7.61
	BA	33.37	28.99	1.42	10.01
	NN-Mod.	33.37	17.79	5.06	32.41
LOR03	BEA	25.30	35.62	2.84	18.30
	BA	25.30	36.74	3.15	20.60
	NN-Mod.	25.30	23.30	4.95	30.08
LOG01	BEA	0.49	0.50	0.059	0.017
	BA	0.49	0.51	0.12	0.043
	NN-Mod.	0.49	0.46	0.35	0.063
LOG03	BEA	0.34	0.38	0.07	0.034
	BA	0.34	0.41	0.12	0.057
	NN-Mod.	0.34	0.47	0.40	0.072
QUA01	BEA	0.16	0.36	0.022	0.128
	BA	0.16	0.46	0.027	0.169
	NN-Mod.	0.16	0.53	0.22	1.32
QUA03	BEA	0.27	0.51	0.031	0.186
	BA	0.27	0.18	0.034	0.231
	NN-Mod.	0.27	0.04	0.27	1.51
HEN01	BEA	0.21	0.20	0.002	0.012
	BA	0.21	0.20	0.0025	0.014
	NN-Mod.	0.21	0.18	0.018	0.117
HEN03	BEA	0.13	0.13	0.0001	0.007
	BA	0.13	0.14	0.0019	0.012
	NN-Mod.	0.13	0.12	0.009	0.061

V. DISCUSSION

This paper reports the results of a comparison of three different forecasting techniques for a class of high roughness long-term time series forecasting. The series were selected regarding the long or short term stochastic dependence of the time series assessed by the Hurst parameter H to give a forecast horizon of 18. The rainfall forecasts obtained between those algorithms are compared with NAR ANN predictor, namely BA and NN-Mod, for a case study on southwestern province of Cordoba. The study analyzed and compared the relative advantages and limitations of each time-series predictor filter technique, used for issuing rainfall and chaotic time series prediction. The discussion of how feed-forward networks can successfully approximate the quantitative changes in the dynamics of the time series data due to changes in the parameter values of the exogenous variables remains open for study, mainly.

Although the comparison was performed on ANN-based filters, the experimental results confirm that the enhanced Bayesian method can predict chaotic time series more effectively in terms of SMAPE and RMSE indices when is compared with other existing forecasting methods in the literature. However, the wish to preserve the stochastic dependencies constrains all the horizons to be forecasted with the same model structure. Since this constraint could reduce the flexibility of the forecasting approach, a variant of BEA approach is still open. Fig.10 and Fig.11 shows the evolution of the SMAPE and RMSE indices for BEA, BA and NN-Mod filter, which use the H parameter to adjust heuristically either structure of the net or parameters of the learning rule.
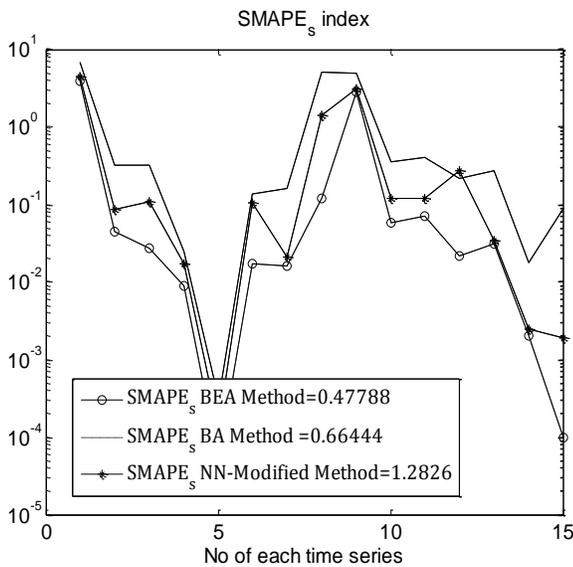
Fig. 9. The SMAPE index applied over the 15 time series
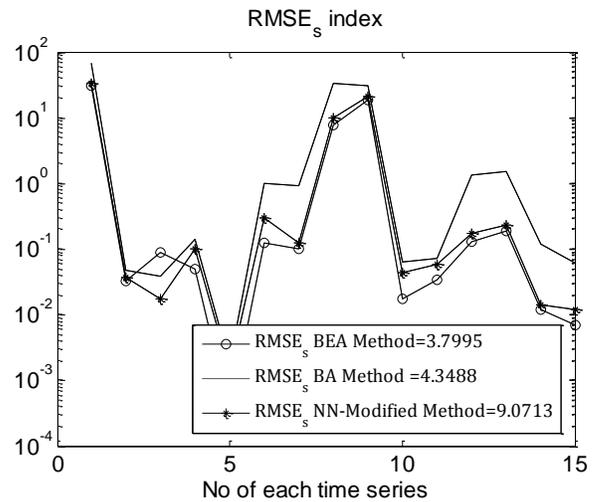
Fig. 10. The RMSE index applied over the 15 time series

VI. CONCLUSIONS

A new approach for time series forecasting: Bayesian enhanced by fractional Brownian motion with application to rainfall series is presented. Building effective predictors form historical data demands computational and statistical methods for inferring dependencies between past and long-term future values of observed values as well as appropriate strategies to deal with longer horizons [32]. This work showed and discussed the BEA supervised learning technique to deal with long-term forecasting problems. In particular the *fBm* model assumption is stressed in Bayes inference by local learning approximators in dealing with important issued in forecasting, like nonlinearity and nonstationarity. The main results show a good performance in term of SMAPE and RMSE indices of the predictor system based on Bayesian enhanced approach, particularly on rainfall time series from a geographical observation point, such as Despeñaderos, Cordoba, Argentina.

Future research should be concerned with the extension of these techniques to some recent directions in big data and the application to spatiotemporal tasks. These results encouraged us to continue working on new machine learning algorithms using novel forecasting methods.

REFERENCES

[1] Armstrong, J.S., Robert Fildes (2006). Making progress in forecasting, International Journal of Forecasting,Volume 22, No. 3.

[2] Green, K. C. and Armstrong, J. S. (2015)Journal of Business Research, Volume 68, Issue 8, Pages 1657-1818 (August 2015) "Special Issue on Simple Versus Complex Forecasting".

[3] Buchanan, M. (2013). Forecast: What Physics, Meteorology, and Natural Sciences Can Teach us about Economics, New York: Bloomsbury.

[4] Fildes, R., and Makridakis, S., (1995). "The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting", International Statistical Review, Vol. 63, no. 3, p. 289-308.

[5] Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting,* No. 5, p. 559–583.

[6] Makridakis, S., and Winkler, R. (1983). "Average of Forecasts: Some Empirical Results", Management Science, Vol. 29, No. 9, p. 987-996.

[7] Rossi, P. E., Allenby G. M., & McCulloch, R. (2006). Bayesian statistics and marketing. Hoboken, NJ: John Wiley & Sons.

[8] Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). Bayesian data analysis (2nd ed.). New York, NY: Chapman & Hall/CRC.

[9] Chatfield, C. (2004). The analysis of time series: An introduction (6th ed.). London, England: Chapman and Hall.

[10] Cristian M. Rodríguez Rivero, Julián A. Pucheta. FORECASTING RAINFALL TIME SERIES WITH STOCHASTIC OUTPUT APPROXIMATED BY NEURAL NETWORKS BAYESIAN APPROACH. International Journal of Advanced Computer Science and Applications (IJACSA).West Yorkshire: The Science and Information (SAI) Organization. 2014 vol.5 n°6. p 145 - 150. ISSN 2156-5570.

[11] Julián Pucheta, C. Rodríguez Rivero, Martín Herrera, Carlos Salas, Víctor Sauchelli, H. Daniel Patiño, Chapter: NON-PARAMETRIC METHODS FOR FORECASTING TIME SERIES FROM CUMULATIVE MONTHLY RAINFALL .Book Rainfall: Behavior, Forecasting and Distribution. Editors: Olga E. Martín and Tricia M. Roberts, Nova Science Publishers, Inc.   ISBN: 978-1-62081-551-9 (2012).

[12] Cristian M. Rodríguez Rivero, Julián A. Pucheta, Martín R. Herrera, Victor Sauchelli, Sergio Laboret. Time Series Forecasting Using Bayesian Method: Application to Cumulative Rainfall, (Pronóstico de Series Temporales usando inferencia Bayesiana: aplicación a series de lluvia de agua acumulada). ISSN 1548-0992. Pp. 359 364. IEEE LATIN AMERICA TRANSACTIONS, VOL. 11, NO. 1, FEB. 2013.

[13] Pucheta, J., Patiño, H.D., Kuchen, B. (2007). Neural Networks-Based Time Series Prediction Using Long and Short Term Dependence in the Learning Process. In proc. of the 2007 International Symposium on Forecasting, New York, USA.

[14] Pucheta, J., Patiño, H. D., Kuchen, B. "A Statistically Dependent Approach for the Monthly Rainfall Forecast from One Point Observations". In Proc. of the Second IFIP Conference on Computer and Computing Technologies in Agriculture (CCTA2008) October 18-20, 2008. Beijing, China. (2008).

[15] Bishwal, J. P. N. (2003). Maximum likelihood estimation in partially observed stochastic differential system driven by a fractional Brownian motion. Stochastic Analysis and Applications, 21(5), 995-1007.

[16] Breton, A. L. (1998). Filtering and parameter estimation in a simple linear system driven by a fractional Brownian motion. Statistics and Probability Letters, 38, 263-274.

[17] Biagini, F., Hu, Y., Øksendal, B., & Zhang, T. (2010). Stochastic calculus for fractional Brownian motion and applications. Berlin: Springer.

[18] Mandelbrot, B. B., & Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. SIAM Review, 10, 422-437.

[19] Achard, S., & Coeurjolly, J. F. (2010). Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise. Statistical Surveys, 4, 117-147.

[20] Abry, P.; P. Flandrin, M.S. Taqqu, D. Veitch., Self-similarity and long-range dependence through the wavelet lens. Theory and applications of long-range dependence, Birkhäuser, pp. 527-556. 2003.

[21]   Bardet, J.-M.; G. Lang, G. Oppenheim, A. Philippe, S. Stoev, M.S. Taqqu. Semi-parametric estimation of the long-range dependence parameter: a survey. Theory and applications of long-range dependence, Birkhäuser, pp. 557-577. 2003.

[22] M. Mackey and L. Glass, \Oscillation and chaos in physiological control systems," Science, vol. 197, p. 287, 1977.

[23] A. Lapedes and F. Farber. Nonlinear signal processing using neural networks: Prediction and System modelling. LA-UR 87-2662. Los Alamos National Laboratory, 1987.

[24] H. Abarbanel, Analysis of Observed Chaotic Data. 1996.

[25] P. Cortez, M. Rocha, and J. Neves, \Evolving time series forecasting neural network models," 2001.

[26] A. Gholipour, B. Araabi, and C. Lucas, \Predicting Chaotic Time Series Using Neural and Neurofuzzy Models: A Comparative Study," Neural Processing Letters, vol. 24, pp. 217-239, 2006.

[27] Characterization of Strange Attractor. Grassberger, Peter and Procaccia, Itmar. s.l.: Physical Review Letters, 1983, Vols. Volume 50, Number 5.

[28] K. Ikeda. Opt . Commun. 30, 257 (1979) .

[29] Sprott, J. C. "Automatic Generation of Strange Attractors." Comput. & Graphics 17, 325-332, 1993. Reprinted in Chaos and Fractals, A Computer Graphical Journey: Ten Year Compilation of Advanced Research (Ed. C. A. Pickover). Amsterdam, Netherlands: Elsevier, pp. 53-60, 1998.

[30] D. J. C. MacKay, "A practical Bayesian framework for backpropagation networks," Neural Comput., vol. 4, no. 3, pp. 448–472, May 1992.

[31] Armstrong, J. S. (2001). Forecasting Principles. 2001.

[32] Mendoza, M. & de Alba, E. (2006). Forecasting an accumulated series based on partial accumulation II: A new Bayesian method for short series with seasonal patterns, International Journal of Forecasting, Issue 4, 781-798.