DE GRUYTER OPEN

JAISCR, 2017, Vol. 7, No. 1, pp. 5 – 16
10.1515/jaiscr-2017-0001

ENERGY ASSOCIATED TUNING METHOD FOR SHORT-TERM SERIES FORECASTING BY COMPLETE AND INCOMPLETE DATASETS

Cristian Rodrìguez Rivero[1], Juliàn Pucheta[1], Sergio Laboret[1], Vìctor Sauchelli[1], Daniel Patiño[2]

[1]*Department of Electrical and Electronic Engineering, Universidad Nacional de Cordoba*
Velez Sarsfield Ave. 1611, Cordoba, Argentina

[2]*Advanced Intelligent Systems Laboratory, Institute of Automatic*
Universidad Nacional de San JuanSan Juan, Argentina

Abstract

This article presents short-term predictions using neural networks tuned by energy associated to series based-predictor filter for complete and incomplete datasets. A benchmark of high roughness time series from Mackay Glass (MG), Logistic (LOG), Henon (HEN) and some univariate series chosen from NN3 Forecasting Competition are used. An average smoothing technique is assumed to complete the data missing in the dataset. The Hurst parameter estimated through wavelets is used to estimate the roughness of the real and forecasted series. The validation and horizon of the time series is presented by the 15 values ahead. The performance of the proposed filter shows that even a short dataset is incomplete, besides a linear smoothing technique employed; the prediction is almost fair by means of SMAPE index. Although the major result shows that the predictor system based on energy associated to series has an optimal performance from several chaotic time series, in particular, this method among other provides a good estimation when the short-term series are taken from one point observations.

Keywords: short time series, forecasting, missing data, energy associated to series, complete and incomplete datasets

1 Introduction

Despite advances in missing data imputation techniques over the last three decades, the problem of missing data remains largely unsolved [1]. The problem of incomplete data poses a difficulty to time series analysis and decision making processes [2] which depend on this data, requiring methods of estimation more accurate and efficient for prediction systems [3]. Various techniques exist as a solution to this problem, ranging from data deletion [4] to methods employing statistical [5, 6] and artificial intelligence techniques [7, 8] to impute for missing variables.

In [9] the authors assessed the impact of missing data on general circulation statistics by systematically decreasing the amount of available training data. They determined that the ratio of the Root Mean Square Error (RMSE) in the monthly mean to the daily standard deviation was two to three times higher when the missing data was spaced randomly compared to space equally, and RMSE increased by up to a factor of two when the missing data occurred in one block.

In [10] found that highly correlated neighbor stations can be used to interpolate missing data in Canadian temperature datasets.

However, some methods, like average smoothing technique [11] can simplify the likelihood of producing biased estimates or make assumptions about the data that may not be true and can be used as a good estimator for the quality of decisions made based on this data. Replacing missing data with time series within the range of known data is crucial [12] for more accurate design proposals and performance evaluation.

In the statistical literature, missing data are usually imputed using maximum likelihood estimators corresponding to a specific underlying model. Very often, these estimators are not efficiently computed, which motivates the use of Expectation Maximization (EM) algorithms [13].

More recently, several data mining techniques for Computational Intelligence [41] have been proposed for short-term time series forecasting (ST-TSF). Examples applied to ST-TSF include: Artificial Neural Networks (ANN) [40, 43, 44, 45, 47], evolutionary computation [53], Support Vector Machines (SVM) [49], fuzzy techniques [51], or their combinations [42, 46, 50, 52].

The motivation of this work arises out of the forecasting problem with incomplete and missing information [14], which is applicable to a large class of learning algorithms [15, 16] including ANNs. One major advantage of the proposed solution is that the complexity does not increase with an increasing number of missing inputs. The solutions can be generalized to the problem of uncertain (noisy) inputs [17].

The estimation of incomplete data in vector elements in real time processing applications requires a system that possesses the knowledge of certain characteristics such as correlations between variables, which are inherent in the input space [18]. The benchmark to construct chaotic time series is chosen from Mackay-Glass (MG), Logistic (LOG) and Henon (HEN) equations, whose forecast is simulated by a Monte Carlo approach employing ANN.

The main contribution here is the forecast system based on energy associated to series (EAS) [19] for tuning the neural networks to predict short time series. The filter parameter is put in function of the roughness of the short time series, between its smoothness. A one-layered feed-forward neural network, trained by the Levenberg-Marquardt algorithm is implemented in order to give the next 15 values [20]. In order to show experimental results, a comparison with other techniques are proposed, such as neural network-based predictor modified (NNMod.) in the learning process [48] and ARMA [28]. The article is organized as follows: Section 2 describes the overall overview of energy associated tuning method and the assumption that the series behaviors are such a fractional Brownian motion path measured by the so-called Hurst parameter. The methodology and data used from benchmark short time series are derived to show the performance of the propose filter. In Section 3 numerical results of complete and incomplete series are presented, where EAS, NNMod. and ARMA models are used to forecast. Section 4 for give some interpretation of the results where is emphasized that the TSF methods proposed can build relationship among the nonlinear handling the irregularity and uncertainty of real data concerning missing or deletion of data in time series forecasting problems. Experimental results are concluded in Section 5. Lastly, some discussion are mentioned for drawing conclusions.

2 Methodology and Data

Analysis of data sets with missing values is a pervasive problem for which standard methods are of limited value. The nature of data can determine what forecasting method can be used. For instance, it is impossible to use ARIMA forecasting techniques [21] if sufficient sample data are unavailable; it is also unnecessary to use a complicated nonlinear technique to forecast a simple linear time series [22]. For that reason, this research points out that simply models can outperform [23], for particular nonlinear time series, predictions with little information and incomplete datasets [24].

2.1 Overview on fractional Brownian motion

The so-called Hurst's parameter is used by this research in the learning process to modify on-line the number of patterns, the number of iterations, and the number of filter's inputs and is defined its stochastic representation [25] as follows

$$B_H(t) = \frac{1}{\Gamma(H+\frac{1}{2})}\begin{pmatrix} \int_0^\infty \left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right) \\ + \int_0^\infty (t-s)^{H-\frac{1}{2}}\, dB(s) \end{pmatrix} dB(s), \quad (1)$$

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

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad (2)$$

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 (1). 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, an *fBm* [26] is a time continuous 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.

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

$$Var\left(B_H(t) - B_H(s)\right) = \nu\, |t-s|^{2H}, \quad (3)$$

where υ is a positive constant.

2.2 Overview on energy associated tuning method

During the learning process of the proposed approach, the primitives of the time series are calculated as a new entrance to the ANN [19], in which the prediction attempts to even the area of the forecasted area to the primitive real area predicted. After each pass stage, the number of inputs of the nonlinear filter is tuned —that is the length of input-delay line, according to the following heuristic criterion. The hypothesis follows that both sequences, the real and the forecasted one, should have the same H parameter. The error between the smoothness of the time series data and the forecasted data (energy associated of series) modifies the number of the filter parameters.

The startup of the algorithm [27] achieves the long term stochastic dependence of the Hurst parameter in order to make more precisely the prediction. The forecasted time series area is set as a new entrance to the NN and serve to be compared with the real area of the time series.

2.3 Benchmark chaotic time series

Despite the series being short (ranging from 51 to 69 observations), the last 15 observations of each time series was separated for performance assessment (the test set) aiming at preserving statistical dependence from the parameter estimation process [28].

Simulations are performed on three common benchmarks. The first one is Primitives of time series from sampling the Mackay-Glass (MG) [29] equations defined by

$$\dot{x}(t) = \frac{ax(t-\tau)}{1-x(t-\tau)^C} - bx(t), \quad (4)$$

with *a,b,c*, τ setting parameters shown in Table 1. The second one is the logistic series (LOG) [30] defined by

$$x(t+1) = ax(t)\,[1-x(t)]. \quad (5)$$

When a=4, the iterates of Eq. 2 form a chaotic time series. The third one is the Henon equation [31] which has a simple format described by

$$x(t+1) = b - ax^2(t), \quad (6)$$

which generates chaotic time series, where the constants are taken to be *a* = 1.3, *b* = 0.22, *x(0)* = 0 and x(1) = 0.

The construction and selection of the benchmark series parameters are shown in Table 1, Table 2 and Table 3, respectively. 50 samples are for the selected short series, the first 35 values are used for training and the remaining 15 values are kept for validation and test data.

The short-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.

Table 1. Parameters to generate short MG time series

Series No.	β	α	c	τ	H	
					Complete	Incomplete
MG1.6	1.6	30	10	100	0.26	0.095
MG1.8	1.8	30	10	100	0.54	0.15
MG17	0.2	0.1	10	17	0.98	0.94
MG30	0.2	0.1	10	30	1	0.98

Table 2. Parameters to generate short HEN time series

Series No.	a	b	H	
			Complete	Incomplete
HEN01	1.4	0.3	0.39	0.21
HEN02	1.3	0.22	0.83	0.7

Table 3. Parameters to generate short LOG time series

Series No.	N	a	X_0	H	
				Complete	Incomplete
LOG402	50	4	0.2	0.14	019
LOG404	50	4	0.4	0.3	0.12

This H serves to have an idea of roughness of a signal [32] and the time series are considered as a trace of an fBm depending on the so-called Hurst parameter 0<H<1 [33].

We select some time series data from the NN3 competition [34]. The complete dataset of 111 time series of the NN3 dataset was chosen containing between 68 and 144 observations. The dataset consists of a representative set of long and short, monthly time series drawn from a homogeneous population of empirical business time series.

Two out of the 111 series were selected since they represent the two most common types of behavior found: annual repetition and financial behavior. These correspond to the 17th and 33rd series from the 111-series. These time series contain both seasonal and non-seasonal patterns, with only minor trends and different time series lengths.

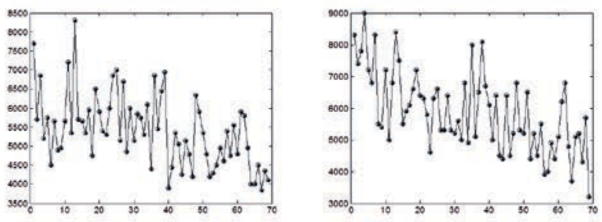


Figure 1. NN3 Competition reduced series chosen; a) NN3_007, b) NN3_008.

Table 4. H Parameter measured from NN3 reduced time series.

Series No.	H	
	Complete	Incomplete
NN3_007	0.19	0.086
NN3_008	0.10	0.099

2.4 Missing Data Mechanisms

In [1] Little and Rubin distinguish between three missing data mechanisms: missing at random, missing completely at random and missing not at random. In this research Missing not at random implies that the missing data mechanism is related to the missing values. It is also referred to as the non-ignorable case [35, 36] as the missing observation is dependent on the outcome of interest. Thus, the methodology of this research follows the missing not a random mechanism and contribute with an ANN technique for missing data imputation.

2.5 Average smoothing technique

The main issue when forecasting a time series is how to retrieve the maximum of information from the available data. In order to predict one step ahead, an average smoothing approach is assumed. It is proposed to fill these empty values by dividing the dataset into subsets of 12. Then a matrix is formed by 12 columns and the number of the rows will depend on the dataset size. To complete the missing information, the prior and posterior data is used for the average smoothing technique as follows: four dataset are built, the first one is an incomplete dataset with $x$labeled in red color. The second one is completed with zeros, the third one is using the same ensemble of the row above and below in order for averaging the prior and posterior row as shown in Figure 2.

The same analogy is used to construct benchmark series.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	x_1	x_2	x	x	x	x_6	x_7	x_8	x_6	x_6	x_6	x_{12}
2	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x	x	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}
3	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}
4	x_{37}	x_{38}	x_{39}	x	x	x	x_{43}	x_{44}	x_{45}	x_{46}	x_{47}	x_{48}
5	x_{49}	x_{50}	..	..	..	..	..	..	..	..	..	..

Figure 2. Average smoothing technique: missing data marked in red.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	x_1	x_2	$x_{15}/2$	$x_{16}/2$	$x_{17}/2$	x_6	x_7	x_8	x_9	x_6	x_6	x_{12}
2	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	$x_{30}+x_6/2$	$x_{31}+x_7/2$	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}
3	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}
4	x_{37}	x_{38}	x_{39}	$x_{28}/2$	$x_{29}/2$	$x_{30}/2$	x_{43}	x_{44}	x_{45}	x_{46}	x_{47}	x_{48}
5	x_{49}	x_{50}	..	..	..	..	..	..	..	..	..	..

Figure 3. Average smoothing technique: infilled with zeros in red.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	x_1	x_2	0	0	0	x_6	x_7	x_8	x_6	x_6	x_6	x_{12}
2	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	0	0	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}
3	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}
4	x_{37}	x_{38}	x_{39}	0	0	0	x_{43}	x_{44}	x_{45}	x_{46}	x_{47}	x_{48}
5	x_{49}	x_{50}	..	..	..	..	..	..	..	..	..	..

Figure 4. Average smoothing technique: completed with the prior and posterior row.

3 Experimental results

3.1 Error Metrics

For each time series, the last 15 observations are used as out-of-sample to compare and evaluate the accuracy of forecasting model. For each out-of-sample observation, its previous data are used as training samples to set the forecasting model for making one-step-ahead forecast.

In order to test the proposed design procedure of the ANN-based nonlinear predictor, the performance given for predicting the chaotic time series is evaluated using the Symmetric Mean Absolute Percent Error (SMAPE) proposed in the most of metric evaluation [37], defined by

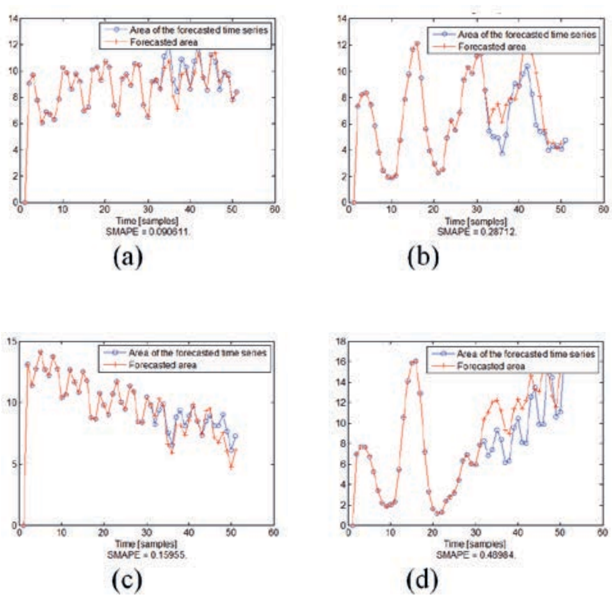
$$SMAPE_S = \frac{1}{n}\sum_{t=1}^{n} \frac{|X_t - F_t|}{(|X|_t + |F_t|)/2} \cdot 100, \qquad (7)$$

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 of each series s calculates the symmetric absolute 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 [38]. SMAPE index self-limits to an error rate of 200%. It has both a lower bound and an upper bound that provides a result between 0% and 200%.

3.2 Numerical results on chaotic time series

In this subsection, the objective of the experiment is the comparison to evaluate the forecasting performances of energy associated tuning method for time series forecasting. Numerical experiments have thus been conducted based on each type of data series, respectively. For producing the horizon forecasts, the same methodology used in section 2.5 was applied to all series with the difference that the 15 last data observations reserved as test set before were implemented in MG, LOG and HEN series included in the modeling data as validation set. However, the last 18 observations were used in NN3 reduced series. 10 time series are used in total for performance comparisons of univariate forecasting energy associated tuning technique, which consist of 50 samples.

The comparison results for each type of series are described as follows: each time series is composed by samples of MG, HEN, LOG and NN3 competition series. The average smoothing technique construction is depicted by means of Figure 10 up to Figure 13 for NN3_008 series. Three classes of datasets are used. The first one is the original time series used by the algorithm to train the predictor filter, which comprises 64 values. The second one is the primitive obtained by integrating the original time series data. The last one is used to compare whether the forecast is acceptable or not, in which the last 15 observations can be used to validate the performance of the prediction system.



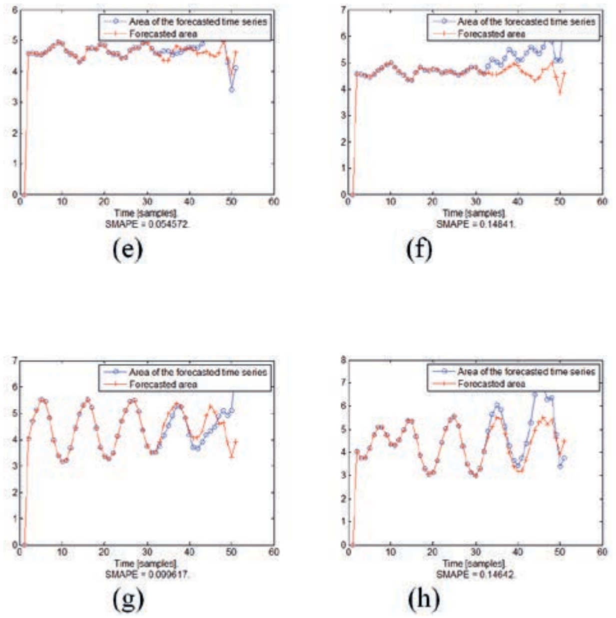


Figure 5. MG series forecasts; a) MG1.6 with complete dataset; b) MG1.6 with incomplete dataset; c) MG1.8 with complete dataset; d) MG1.8 with incomeplete dataset; e) MG17 with complete dataset; f) MG17 with incomplete dataset; g) MG30 with complete dataset; h) MG30 with incomplete dataset.

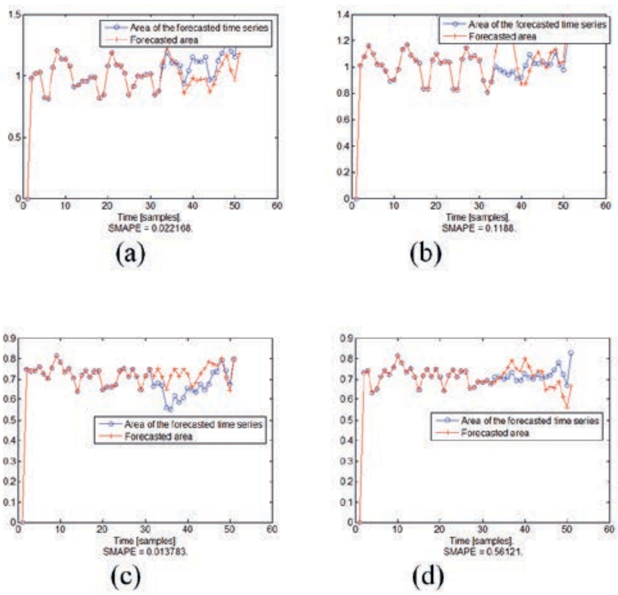


Figure 6. NN3_008 series forecasts; a) NN3_008 with complete dataset; b) NN3_008 with incomplete dataset; a) NN3_007 with complete dataset; b) NN3_007 with incomplete dataset.

The Monte Carlo method was used to forecast the horizon of the benchmark series. Such outcomes are shown from Figure 5 to Figure 8. The validation set was used for measuring the minimum mean squared error as training stopping criterion again. The validation set was also used for improving performance measured by the SMAPE metrics.

Based on the SMAPE metrics measured on the NN3 validation set, 18 steps ahead, only the single best performing combination of architecture and training algorithm was selected for each series' training. Each forecasting ensemble was then used for producing the 15 forecasts for its series. The median forecast of each ensemble was then compared to the test data set of their respective series.

Table 5. Comparison of the proposed approach by SMAPE Index

Series No.	SMAPE EAS	
	Complete Dataset	Incomplete Dataset
MG1.6	0.090	0.287
MG1.8	0.159	0.489
MG17	0.054	0.148
MG30	0.059	0.146
HEN01	0.022	0.118
HEN02	0.013	0.561
LOG402	0.202	0.430
LOG404	0.548	0.764
NN3_007	20.94	21.23
NN3_008	19.93	23.51

Table 6. Comparison using NNmod predictor filter by SMAPE Index

Series No.	SMAPE NNmod	
	Complete Dataset	Incomplete Dataset
MG1.6	0.265	0.361
MG1.8	0.390	0.632
MG17	0.090	0.140
MG30	0.033	0.148
HEN01	0.058	0.072
HEN02	0.037	0.047
LOG402	0.962	1.02
LOG404	0.905	1.11
NN3_007	22.17	24.63
NN3_008	16.48	26.36

Table 7. Comparison using linear ARMA predictor filter by SMAPE Index

Series No.	SMAPE ARMA	
	Complete Dataset	Incomplete Dataset
MG1.6	0.288	0.450
MG1.8	0.394	0.812
MG17	0.112	0.263
MG30	0.042	0.195
HEN01	0.025	0.118
HEN02	0.451	0.092
LOG402	0.072	0.067
LOG404	0.23	1.82
NN3_007	36.72	38.70
NN3_008	36.58	40.23

4 Interpretation of the results

The results presented show the forecast error measures selected from the known data separated into validation and test sets, as described in the previous Section. The measurable results on these series are presented in Table 4, Table 5 and Table 6 which show the performance of the system according to SMAPE metrics [37, 38] applied to each series validation and test sets, averaged across short and long forecast horizons, for time series categorized as long and short [34] using different forecasting methods, such as EAS, NNMod. and ARMA.

It can be observed that on short time series, the EAS predictor filter has the smallest SMAPE for medium to long horizons, and over forecast lead time, 1-18. The performance of method reflects an advantage of both sorts of series, chaotic and financial. The availability of sufficient data for training is particularly important where the time series is short.

For short series, the size of the training and validation set from using complete and incomplete datasets, results in better training of the proposed approach network and as the results suggest, improved forecast accuracy. When sufficient data is available for training and validation, the increase in the SMAPE index is shown particularly from complete datasets. As expected, the financial series has much worse performance than the chaotic time series.

The forecast of series NN3_008 is much better than that of series NN3_007. Despite not revealed on a series basis, the competition results measured a 16.87% SMAPE for this approach.

5 Discussion

The evaluation of the results across the ten analyzed cases was performed by the same initial parameters for each predictor filters. The parameters and the structure of the filters are tuned by considering their stochastic dependency.

It can be noted that in Figure 9 the SMAPE index is computed between the complete time series horizon (it includes the series validation and test horizon) and the incomplete series horizon, as indicates the Eq. (7) for each series, performed by the three filters. Note that there is no improvement of the forecast for any given time series, which results from the use of a stochastic characteristic to generate a deterministic result, such as using complete and incomplete datasets for short-term prediction.

In addition, a comparison was performed by other predictor filters proposed such as NNmod. [28] and ARMA [48] in term of error metrics. The result highlights that the energy associated predictor filter (EAS) supplied to short time series has an optimal performance from several chaotic time series, in particular to time series whose H parameter has a high roughness of signal, which is evaluated by H parameter, respectively. This fact encourages us to apply the proposed approach to meteorological and financial time series when the observations are taken from a linear statistical point.

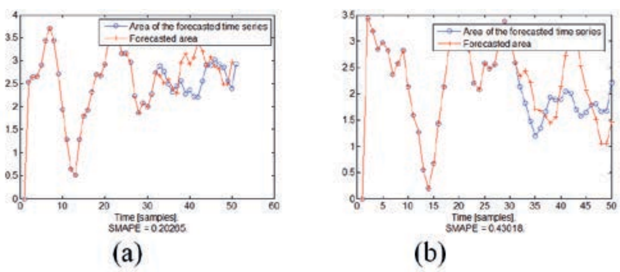


Figure 7. The SMAPE index applied over the 10 time series.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	8300	7400	7800	9000	7200	6800	8300	5500	5400	7200	5000	6800
2	8400	7500	5500	5900	6100	6600	7200	6400	6300	5800	4600	6300
3	6600	5300	5300	6400	5300	5200	5600	5000	6800	4900	8000	5100
4	6500	8100	6700	6100	5000	6400	4500	4400	6400	4500	5200	6800
5	5300	5200	6500									

Figure 8. Original NN3_008 time series, data marked in red color are selected to construct the proposed average technique.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	8300	7400	0	0	0	6800	8300	5500	5400	7200	5000	6800
2	8400	7500	5500	5900	6100	0	0	0	6300	5800	4600	6300
3	6600	5300	5300	6400	5300	5200	5600	5000	6800	4900	8000	5100
4	6500	8100	6700	0	0	0	4500	4400	6400	4500	5200	6800
5	5300	5200	6500									

Figure 9. NN3_008 infilled with zeros marked in red color.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	8300	7400	2750	2950	3050	6800	8300	5500	5400	7200	5000	6800
2	8400	7500	5500	5900	6100	6000	6950	5250	6300	5800	4600	6300
3	6600	5300	5300	6400	5300	5200	5600	5000	6800	4900	8000	5100
4	6500	8100	6700	3200	2650	2600	4500	4400	6400	4500	5200	6800
5	5300	5200	6500									

Figure 10. NN3_008 completed with average between prior and posterior row data marked in red color.

x_n	1	2	3	4	5	6	7	8	9	10	11	12
1	4400	5200	4500	5500	3900	4000	4900	4400	5100	6200	6800	4800
2	3700	5100	5200	4300	5700	3200						

Figure 11. NN3_008 test data

6 Conclusion

In this work a feedforward neural networks-based on nonlinear autoregression (NAR) filter by means of energy associated to series (EAS) tuning approach for forecasting time series has been presented. The difficulties in forecasting short-term series are attributable to the missing information in the dataset relative to the limited number of samples. The smoothing technique proposed is adopted to complete the data. In contrast, we emphasized the importance of inferences based on one-step ahead forecasting performances in the practically more relevant context of missing data [38]. We illustrate the proposed approach by deriving closed-form solutions for a selection of benchmark time series used in the literature. Then we compare performances with NNMod. and ARMA predictor filters one-step ahead based on a selection of simulated as well as practical time series by the SMAPE index. Our empirical findings confirm that the smallest forecast errors for a given forecast arise from the corresponding EAS method shown in Figure. 9 for complete and incomplete datasets. Even though a linear rule is used to complete the missing data, the finally proposed model to short term forecast with minimum least squared errors in addressing whether missing data necessarily needs to be imputed using complicated techniques, this work found that a imputations using linear smoothing approach are fairly acceptable.

The learning rule proposed to adjust the neural net weights is based on the Levenberg-Marquardt method. The parameters estimated in the modeling stage were then in function of the long and short term stochastic dependence of the time series, evaluated by the Hurst parameter H, to update the neural net topology, number of input taps, and the number of patterns and iterations at each time stage. Future work may focus on applications of the model in some relevant fields and real-life problems. We are interested in addressing more complex forecasting problems such as computation of concurrent trend or seasonal-adjustment filters in univariate and multivariate frameworks, of particular interest to time-series data with concept drift or with high levels of noise.

Acknowledgment

This work was supported by Universidad Nacional de Crdoba (UNC), FONCYT-PDFT PRH N°3 (UNC Program RRHH03), SECYT UNC, Universidad Nacional de San Juan – Institute of Automatics (INAUT), National Agency for Scientific and Technological Promotion (ANPCyT) and Departments of Electrotechnics – Electrical and Electronic Engineering - Universidad Nacional of Cordoba.

References

[1] Little, R.J.A. and D.B. Rubin, Statistical Analysis with Missing Data. John Wiley Publishers Company, 2002.

[2] Tsikriktsis, N., A review of techniques for treating missing data in OM survey research, Journal of Operations Management 24, 2005, pp. 53–62.

[3] Maravall A, Pea D., Missing observations and additive outliers in time series models. In: Mariano RS (ed) Advances in statistical analysis and statistical computing. JAI Press, Stanford, 1986.

[4] Mitat Uysal, Reconstruction of Time Series Data with Missing Values, Journal of Applied Sciences, 7 (6): ISSN 1812-5654, 2007, pp. 922–925 .

[5] Kornelsen, K., & Coulibaly, P., Comparison of Interpolation, Statistical, and Data-Driven Methods for Imputation of Missing Values in a Distributed

Soil Moisture Dataset. Journal of Hydrologic Engineering, 19(1), 2012, pp. 26–43.

[6] Kharin, Yu.S. and Huryn, A.S., Statistical analysis and forecasting of autoregressive time series under missing values. Bulletin of the International Statistic Institute, 1, 2003, pp. 612–613.

[7] Hong, B. and CH. Chen, Radial basis function neural network-based nonparametric estimation approach for missing data reconstruction of nonstationary series. IEEE Int. Conf. Neural Networks and Signal Processing Nanjing, China, December, 14-17, 2003, pp. 75–78.

[8] Coulibaly P. Comparison of neural network methods for infilling missing daily weather records. Journal of Hydrology. v. 341, 2007, pp. 27–41.

[9] Kidson, J.W., and K.E. Trenberth, Effects of missing data on estimates of monthly mean general circulation statistics, J. Climate, 1, 1988, pp. 1261–1275.

[10] Vincent, L. A., and D. W. Gullet., Canadian historical and homogeneous temperature datasets for climate change analyses. International Journal of Climatology. 19, 1999, 1375–1388.

[11] Rodrguez Rivero, Cristian; Patio, Hector Daniel; Pucheta, Julian Antonio, Short-term rainfall time series prediction with incomplete data, in Neural Networks (IJCNN), 2015 International Joint Conference on, 2015, pp. 1–6, 10.1109/IJCNN.2015.7280315.

[12] Wasito, I.: 2003, Least Squares Algorithms with Nearest Neighbour Techniques for Imputing Missing Data Values, PhD thesis, University of London.

[13] Nelwamondo, F. V., Mohamed, S. and Marwala, T.: n.d., Missing data: A comparison of neural networks and expectation maximization techniques, Current Science 93(11), 2007.

[14] Cristian Rodrguez Rivero, Julin Pucheta, Sergio Laboret, Daniel Patio, Vctor Sauchelli. Forecasting short time series with missing data by means of energy associated of series. Applied Mathematics, 6, 2015, pp. 1611–1619, http://dx.doi.org/10.4236/am.2015.69143.

[15] Richman, M. B.; Trafalis, T. B.; Adrianto I.; Missing data imputation through machine learning algorithms, in Artificial Intelligence Methods in the Environmental Sciences. Ed. by H. Sue Ellen, P. Antonello, M. Caren. Springer Netherlands Press, 2009, pp.153–169, doi:10.1007/978-1-4020-9119-3_7.

[16] Haviluddina, Ahmad Jawahir, Comparing of ARIMA and RBFNN for short-term forecasting, International Journal of Advances in Intelligent Informatics, Vol. 1, No 1, 2015, pp. 15–22.

[17] Zhang G. P., Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 2003, pp. 159–175.

[18] Schneider, T., Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. J. Climate, 14, 2001, pp. 853–871.

[19] C. Rodrguez Rivero, M. Herrera, J. Pucheta, J. Baumgartner, D. Patio and V. Sauchelli. High roughness time series forecasting based on energy associated of series, Journal of Communication and Computer, USA, David Publishing Company, Vol. 9 No. 5, 2012, pp. 576–586, ISSN 1548-7709.

[20] C. Rodriguez Rivero, J. Pucheta, H. Patio, J. Baumgartner, S. Laboret and V. Sauchelli. Analysis of a Gaussian Process and Feed-Forward Neural Networks based Filter for Forecasting Short Rainfall Time Series, 2013, International Joint Conference on Neural Networks. doi: 10.1109/IJCNN.2013.6706741.

[21] Kohn, R., & Ansley, C. F., Estimation, Prediction, and Interpolation for ARIMA Models with Missing Data, Journal of the American Statistical Association, 81, 1986, pp. 751–761.

[22] Jones R. H., Maximum Likelihood Fitting of ARMA Models to Time Series with Missing Observations, Technometrics, 22(3), 1980, pp. 389–395.

[23] Shumway R., and D. Stoffer, An approach to time series smoothing and forecasting using the EM algorithm, Journal of Time Series Analysis, 3, 1982, pp. 253–264.

[24] Tresp V., & Hofmann R., Nonlinear Time-Series Prediction with Missing and Noisy Data. Neural Computation, 10, 1998, pp. 731–747.

[25] Mandelbrot B. B., The Fractal Geometry of Nature, Freeman, San Francisco, CA, 1983.

[26] Dieker, T. Simulation of fractional Brownian motion. MSc theses, University of Twente, Amsterdam, The Netherlands, 2004.

[27] C. Rodrguez Rivero, J. Pucheta, J. Baumgartner, M. Herrera, D. Patio y B. Kuchen. A NN-based model for time series forecasting in function of energy associated of series, Proc. of the International Conference on Applied, Numerical and Computational Mathematics (ICANCM11), Barcelona, Spain, 2011, pp. 80–86, ISBN 978-1-61804-030-5.

[28] Pucheta J., Patio 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) 2008, Beijing, China.

[29] Glass L. and M. C. Mackey. From Clocks to Chaos, The Rhythms of Life. Princeton University Press, Princeton, NJ, 1988.

[30] Robert M. May, Simple mathematical models with very complicated dynamics, Nature, vol.261, 1976, pp. 459–467.

[31] Henon M., A two-dimensional mapping with a strange attractor. Communications in Mathematical Physics. Vol. 50, 1976, pp. 69–77.

[32] 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, Birkhuser, 2003, pp. 527–556.

[33] 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, Birkhuser, 2003, pp. 557–577.

[34] S. F. Crone, M. Hibon, and K. Nikolopoulos, Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction, International Journal of Forecasting, vol. 27, 2011, pp. 635–660.

[35] Rubin D. B., Multiple Imputation for Nonresponse in Surveys, New York:Wiley, 1987.

[36] Schafer J., Analysis of Incomplete Multivariate Data, Chapman & Hall, 1997.

[37] Armstrong J.S. (Ed.) Principles offorecasting: Handbook for researchers and practitioners. Kluwer, 2001.

[38] Makridakis, S. & M. Hibon, The M3-Competition: Results, conclusions and implications, International Journal of Forecasting, 16, 2000, pp. 451–476.

[39] Shang Zhaowei, Zhang Lingfeng, Ma Shangjun, Fang Bin, Zhang Taiping. Incomplete Time Series Prediction Using Max-Margin Classification of Data with Absent Features Hindawi Publishing Corporation. Mathematical Problems in Engineering, Volume 2010, Article ID 513810, doi:10.1155/2010/513810.

[40] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, C. Cobelli, A new neural network approach for short-term glucose prediction using continuous glucose monitoring time-series and meal information, 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011, pp. 5653–5656.

[41] L.P. Wang and X.J. Fu, Data Mining with Computational Intelligence, Springer, Berlin, 2005.

[42] Y. Ren, P. N. Suganthan, N. Srikanth, G. Amaratunga, Random Vector Functional Link Network for Short-term Electricity Load Demand Forecasting, Information Sciences, 2016.

[43] L. P. Wang and Shekhar Gupta, Neural networks and wavelet de-noising for stock trading and prediction, Time Series Analysis, Modeling and Applications, Witold Pedrycz and Shyi Ming Chen (eds.), Springer, 2013, pp. 229–247.

[44] L. P. Wang, K.K. Teo, and Z.P. Lin, Predicting time series with wavelet packet neural networks, 2001 IEEE International Joint Conference on Neural Networks (IJCNN 2001), 2001, pp. 1593–1597.

[45] Dong-Chul Park, "A Time Series Data Prediction Scheme Using Bilinear Recurrent Neural Network," 2010 International Conference on Information Science and Applications (ICISA), 2010, pp. 1–7.

[46] M. Zhu and L. P. Wang, Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms, in The 2010 International Joint Conference on Neural Networks (IJCNN), 2010, pp. 1–5.

[47] K.K. Teo, L.P. Wang, Z.P. Lin, Wavelet packet multi-layer perceptron for chaotic time series prediction: effects of weight initialization, Computational Science – ICCS 2001, Proceedings Pt 2, Volume: 2074, 2001, pp. 310–317.

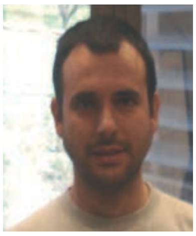
[48] J. Pucheta, D. Patio and B. Kuchen. Neural Networks-Based Time Series Prediction Using Long and Short Term Dependence in the Learning Process". In proc. of the 2007 International Symposium on Forecasting, New York 2007.

[49] C.-N. Ko, C.-M. Lee, Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended kalman filter, Energy 49, 2013, pp. 413–422.

[50] R.-A. Hooshmand, H. Amooshahi, M. Parastegari, A hybrid intelligent algorithm based short-term load forecasting approach, International Journal of Electrical Power & Energy Systems 45, 2013, pp. 313–324.

[51] Kamal S. Selim and Gihan A. Elanany, A New Method for Short Multivariate Fuzzy Time Series Based on Genetic Algorithm and Fuzzy Clustering, Advances in Fuzzy Systems, vol. 2013, Article ID 494239, 2013, doi:10.1155/2013/494239.

Cristian Rodriguez Rivero received Electrical Engineering, and his Ph.D. degree in electrical engineering from Faculty of Exact, Physical and Natural Sciences at Universidad Nacional de Córdoba (UNC), in the topic of computational intelligence applied for intelligence agriculture. His research interest includes topics of automatic control, stochastic and optimal control for intelligence agriculture, time series forecasting, machine learning: neural networks, support vector regression and genetic algorithms. Board of Directors of the LA-CIS and Steering Committee of the LA-CCI (IEEE Latin America Conference on Computational Intelligence - la-cci.org). IEEE CIS Argentina Chapter chair since 2013. He is a member of: INNS (International Neural Network Society), SIAM (Society for Industrial and Applied Mathematics) Control and Optimization, and AADECA (Argentine Association of Automation and Control).



J. A. Pucheta received the Electrical Engineering degree from National Technological University - Córdoba Regional Faculty, Argentina, the M.S. and Ph.D. degrees from Universidad Nacional de San Juan, Argentina, in 1999, 2002 and 2006, respectively. He is currently Professor at Mathematics Research Laboratory applied to Control (LIMAC), Universidad Nacional de Córdoba, Argentina. His research interests include stochastic and optimal control, time series forecast and machine learning. IEEE member.

H. Daniel Patiño was born in San Juan, Argentina. He received the electronic engineering degree, and the doctorate degree in engineering from the Universidad Nacional de San Juan (UNSJ), Argentina, in 1986 and 1995 respectively. He is an Associate Professor in the Institute of Automatics and the Electronics and Automatics Department, Faculty of Engineering, UNSJ. From 2001 to 2003, he was also Secretary of Research in this Faculty. From 1987 to 1993, he was fellowship holder of the National Council for Scientific and Technical Research (CONICET) of Argentina. From 1997 to 1998, he was a post-doctoral research associate and taught in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. From 2003, he is in charge of the Coordination of the Strategic Program of Information Technology and Communications (ITC) of the Secretary of Science, Technology and Production Innovation of Argentina. His current research areas cover computational intelligence, operations research, robotics, systems and control theory, data mining, computational finance. Specific research and teaching topics include artificial neural networks, hybrid intelligent systems, adaptive critic designs, neurodynamic programming and evolutionary computation for optimization and control, intelligent agents, and autonomous machines and vehicles. In 1994, as a Member of the Research Staff of the Institute of Automatics received the National Award, Dr. José A. Balseiro, for the Best Research Group in the technological area and the technology transfer to industry. Dr. Patiño is a member of the IEEE and he has collaborated and served as a technical reviewer in the IEEE Transactions on Systems, Man, and Cybernetics, Circuits and Systems, and in numerous IEEE Conferences.

Sergio O. Laboret received the Electrical Engineering degree from Universidad Nacional de Córdoba and M.Sc. degrees in University Education from National Technological University, Córdoba Regional Faculty (UTN, FRC,) Argentina, in 1991 and 2012, respectively. He is currently Professor of Control Systems II and Signal Processing in electrical and electronic engineering. His interests include fractional calculus, robust control, neural networks and signal processing.

V. H. Sauchelli received the Electrical Engineering degree and Ph.D. degrees from Universidad Nacional de Cordoba (UNC) a, Major in University Education from National Technological University, Córdoba Regional Faculty (UTN, FRC,) Argentina, in 1973 and 1997, respectively. He is currently Principal at Telecommunication Postgraduate Program, National University of Cordoba, and Professor of Control Systems and Signal Processing in electrical and electronic engineering. He is also Head of Mathematics Research Laboratory applied to Control (LIMAC) and Research Projects at Secretary of Sciences and Technology (SECyT), Universidad Nacional de Cordoba. His interests include automatic control, neural networks, fractional-order controllers and signal processing.