Comparison of Models for Wind Speed Forecasting

J.C. Palomares-Salas, J.J. G. de la Rosa, J.G. Ramiro, J.Melgar, A. Agüera and A. Moreno
Research Unit PAIDI-TIC-168. University of Cádiz. Electronics Area. Escuela Politécnica Superior
Avda. Ramón Puyol, S/N. E-11202-Algeciras-Cádiz (*Spain*)
Email: juanjose.delarosa@uca.es

***Abstract*—In this paper an ARIMA model is used for time-series forecast involving wind speed measurements. Results are compared with the performance of a back propagation type NNT. Results show that ARIMA model is better than NNT for short time-intervals to forecast (10 minutes, 1 hour, 2 hours and 4 hours). Data was acquired from a unit located in Southern Andalusia (Peñaflor, Sevilla), with a soft orography (10 minutes between measurements). This feature is which makes performance of the ARIMA model and the NNT very similar, so a simple forecasting model could be used in order to administrate energy sources. The paper presents the process of model validation, along with a regression analysis, based in real-life data.**

Keywords-Short-term wind speed prediction; ARIMA; Neural networks; time-series; weather forecasting; wind speed.

I. INTRODUCTION

The use of wind energy has been developed significantly throughout the world, in order to get the ideal for a future with electricity without pollution. But the integration of wind farms in the power networks has become an important problem for the unity of commitment and control of power plants in electric power systems. Wind is considered one of the weather variables which more difficult to be predicted. Intermittent in nature, the electricity produced in a wind farm is difficult to be short-term forecasted. It is even difficult in the next few hours and, in general, any benefits obtained from the wind farms is not optimal, and may be necessary to increase the power plant spinning reserve.

Hence, the need to administer energy resources and the advent of alternative energy, particularly wind power, necessitate the use of advanced tools for short-term prediction of wind speed or what is the same thing, the wind production. End-users (independent power producers, electrical companies, system operator distribution, etc.) which recognize the contribution of wind forecast for a safe and economic operation of the network. Especially, in a liberalized electricity market, forecasting tools improve the position of wind energy compared with other available forms of generation.

II. EXPERIMENTAL PROCEDURE AND RESULTS

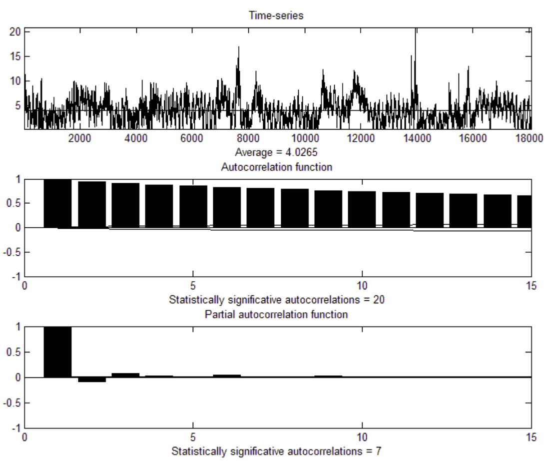
A. ARIMA model: identification, estimation of parameters, validation and forecasting

In this section we expose the procedure to obtain the possible models which better explain the time-series behavior.

The original measurement time-series comprises 18,090 data corresponding to wind speed in meters per second (m/s), acquired each 10 minutes.

The first step is to establish the ARIMA model which best fits the time-series behavior. With this goal the autocorrelation coefficient and the partial autocorrelation coefficient are evaluated and depicted in Fig. 1.

Fig.1. Original time-series, and autocorrelation functions for model identification.



The autocorrelation coefficient shown in Fig. 1 decays as the time-lag increases. This conveys the idea, as suggested by Box and Jenkins [1], that it may possible an autoregressive model to fit the time-series; the partial autocorrelation graph confirms that we are in presence of an autoregressive model, because the partial autocorrelation cuts beyond a certain time-lag, which by the way establishes the order of the model. More precisely, as the decay begins in the second lag, this is assumed as the order of the model.

Fig. 2 shows the model prospecting results from the first-derivative time-series. No model is concluded because both autocorrelation and partial autocorrelation coefficients decay to zero from the very beginning.

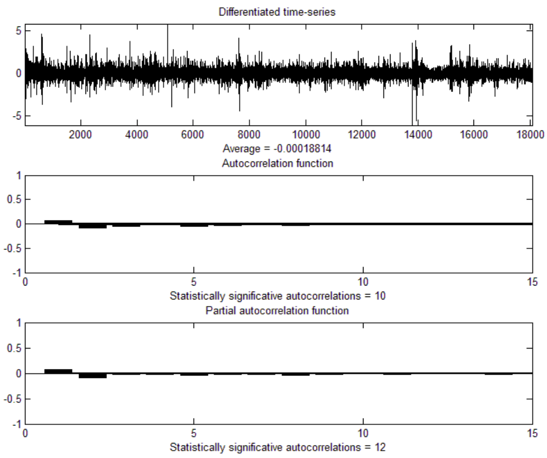


Fig.2. Differentiated original time-series, and autocorrelation functions for model identification.

Fig. 3 shows the model prospecting results from the second-derivative time-series. A possible model is an autoregressive-integrated-moving average (2,2,0); because the decay in the partial autocorrelation graph from the second lag. A second possible model is (0,2,1), because the autocorrelation function decay from the first lag. The second index "2" is associated to the order of the integral.

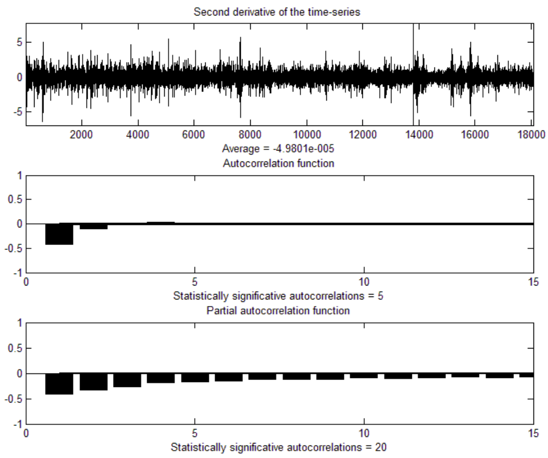


Fig.3. The original time-series, two times differentiated, and autocorrelation functions for model identification.

In the estimation process, mistakes of the model computation have also been calculated and assessed, along with the values of the parameters. The values for the autoregressive coefficients are 0.9 and 0.1 (2,0,0). This is the easiest model to implement and the model with least computation-time, in the forecasting procedure. These are the reasons why we have selected this model.

To assess the model the following 3 parameters have been selected. The Pearson correlation coefficient associated with the original N-point time-series (x) and the forecasted series ($\hat{x}$):

$$r_{x,\hat{x}} = \frac{\sigma_{x\hat{x}}}{\sigma_x \sigma_{\hat{x}}} ; \tag{1}$$

the Index Of Agreement (IOA), Willmot [2]:

$$IOA_{x,\hat{x}} = 1 - \frac{\sum_{i=1}^{N} (\hat{x}_i - x_i)^2}{\sum_{i=1}^{N} \left(|\hat{x}_i - x_i| + |\hat{x}_i + x_i| \right)^2} , \tag{2}$$

and the RMSE:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2} . \tag{3}$$

Eqs. (1), (2) and (3) are the quality indexes or indicators used to assess models.

Tables 1, 2 and 3 summarize the assessment results for the three selected models. Two situations are studied: considering null initial conditions (conditional) and conditioned to concrete initial values (non conditioned). These tables have been obtained from forecasting results and for a few horizons of prediction of 10 minutes, 1, 2 and 4 hours.

TABLE I. MODEL ASSESSMENT ARIMA (2,0,0).

	Conditional				**Non conditional**			
Forecasting horizon	10 min	1 h	2 h	4 h	10 min	1 h	2 h	4 h
$r_{x,\hat{x}}$	0.97	0.88	0.80	0.67	0.97	0.88	0.79	0.65
$IOA_{x,\hat{x}}$	0.98	0.93	0.88	0.79	0.98	0.93	0.89	0.80
$RMSE$	0.57	1.07	1.31	1.55	0.58	1.12	1.41	1.66

TABLE II. MODEL ASSESSMENT ARIMA (2,2,0).

	Conditional				**Non conditional**			
Forecasting horizon	10 min	1 h	2 h	4 h	10 min	1 h	2 h	4 h
$r_{x,\hat{x}}$	0.96	0.86	0.79	0.65	0.96	0.86	0.79	0.64
$IOA_{x,\hat{x}}$	0.98	0.92	0.88	0.79	0.98	0.92	0.88	0.78
$RMSE$	0.67	1.30	1.63	2.12	0.67	1.30	1.63	2.12

TABLE III. MODEL ASSESSMENT ARIMA (0,2,1).

	Conditional				Non conditional			
Forecasting horizon	10 min	1 h	2 h	4 h	10 min	1 h	2 h	4 h
$r_{x,\hat{x}}$	0.93	0.78	0.68	0.57	0.74	0.48	0.45	0.23
$IOA_{x,\hat{x}}$	0.96	0.87	0.81	0.72	0.83	0.60	0.56	0.40
RMSE	0.90	1.73	2.1	2.56	2.15	4.09	4.39	6.28

As a general remark, we can say that as the forecasting horizon increases, all quality indexes degenerate. This indicates that models are not valid for long term forecasting. On the other hand the model (2,0,0) has the best quality index.

Fig. 4 shows the time-series forecasting results using the model ARIMA (2,0,0). Fig. 5 shows the regression of the time series in Fig. 4.

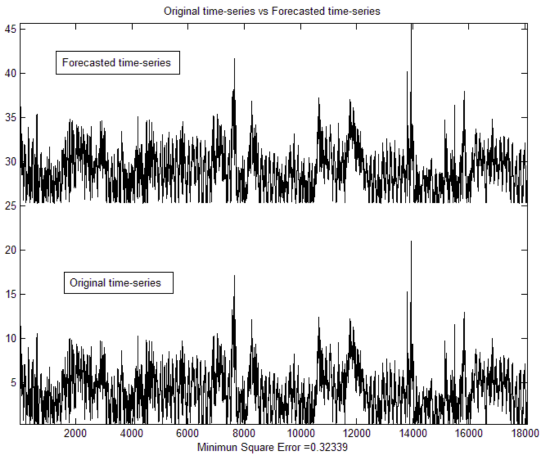


Fig. 4. Original time-series and forecasted data.

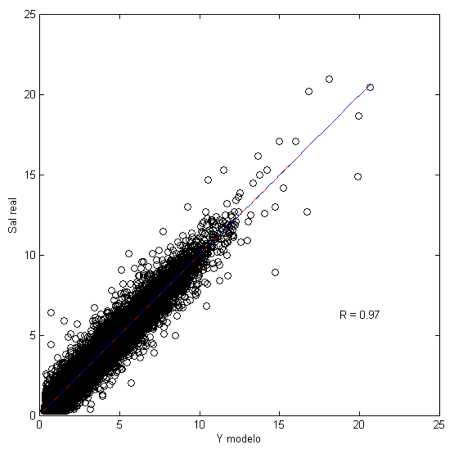


Fig. 5. Regression of the original time-series and forecasted data.

B. Neural networks

We used a layer of backpropagation networks in which employs three sampling techniques for conducting training and validating them. These techniques have been: Bootstrap, double cross-validation and 10 cross-validations.

The bootstrap analyses repeatedly subsamples data. Each set of training, which has the same size as the original data set is obtained through a selection process with replenishment on all available data. The data not selected for the joint training became part of the overall validation, which makes the size of the latter is not uniform.

The double cross-validation [3] is to choose two sets of data available, one for training and another for test.

The 10 cross-validations [4] in which breaks down the data set in 10 subsets disjunct. All data have been used at some moment so as to train to validate what we get a more reliable estimate.

The algorithm backpropagation used for the learning has been the method Levenberg-Marquardt [5] for being a numerical technical optimization most powerful and very fast but has the disadvantage that requires a lot of memory to run. The number of employed has been hidden layers of 30 units and we have made management of data for forecasting in the form of parent autoregressive deep *N* samples. This means that for the prediction of the process in a time *t* used *n* previous values of the process. Also normalize the data so that they are in the interval [-1.1] so that training is faster.

The models Backpropagation neuronal network have been, on one hand: the Bootstrap, the double cross-validation and the 10 cross-validations, all with an autoregressive management of $n = 6$; and on the other hand: the same ones but with a management of $n = 12$.

Tables 4, 5 and 6 summarize the assessment results for the three selected models, so much with $n = 6$ as with $n = 12$. These tables have been obtained f*rom forecasting* results and for a few horizons of prediction of 10 minutes, 1, 2 and 4 hours.

TABLE IV. MODEL ASSESSMENT BOOTSTRAP.

	***n* = 6**				***n* = 12**			
Forecasting horizon	10 min	1 h	2 h	4 h	10 min	1 h	2 h	4 h
$r_{x,\hat{x}}$	0.96	0.77	0.59	0.43	0.95	0.72	0.55	0.37
$IOA_{x,\hat{x}}$	0.98	0.87	0.75	0.61	0.97	0.83	0.72	0.56
RMSE	0.67	1.54	2.17	3.03	0.76	1.82	2.5	3.57

TABLE V. MODEL ASSESSMENT DOUBLE CROSS-VALIDATION.

	***n* = 6**				***n* = 12**			
Forecasting horizon	10 min	1 h	2 h	4 h	10 min	1 h	2 h	4 h
$r_{x,\hat{x}}$	0.95	0.65	0.47	0.28	0.93	0.68	0.46	0.33
$IOA_{x,\hat{x}}$	0.98	0.76	0.66	0.51	0.96	0.81	0.62	0.52
RMSE	0.69	2.34	2.79	3.61	0.89	2.03	3.3	3.88

TABLE VI. MODEL ASSESSMENT 10 CROSS-VALIDATIONS.

	***n* = 6**				***n* = 12**			
Forecasting horizon	10 min	1 h	2 h	4 h	10 min	1 h	2 h	4 h
$r_{x,\hat{x}}$	0.96	0.82	0.59	0.4	0.95	0.78	0.56	0.39
$IOA_{x,\hat{x}}$	0.98	0.9	0.76	0.62	0.97	0.88	0.7	0.59
RMSE	0.66	1.32	2.05	2.76	0.73	1.53	2.74	3.06

Fig. 6 presents a window of 100 records of length belonging to the results obtained by the models for a horizon of prediction of 10 minutes. Also the coefficient of correlation is indicated in the graph for every model with regard to the original series:

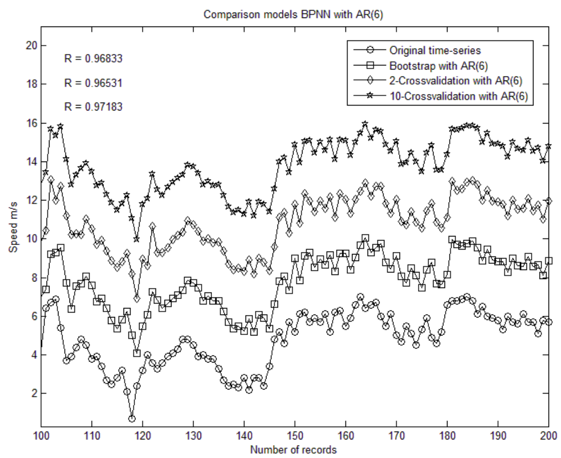


Fig. 6. Original time-series and forecasted data with BPNN with n = 6.

We have also realized the same graph using the same methods of re-sampling but now using an autoregressive management of n = 12, giving us the results showed in the figure 7:

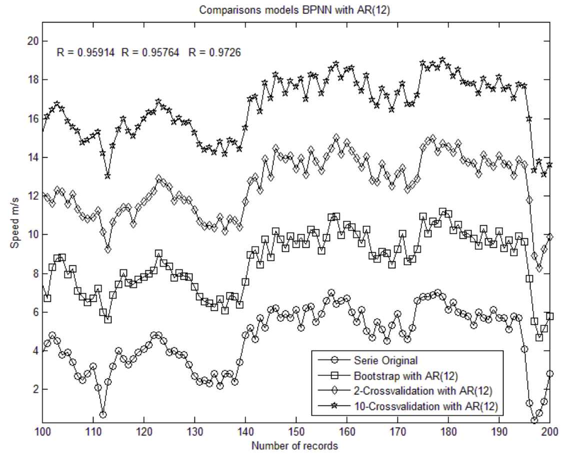


Fig. 7. Original time-series and forecasted data with BPNN with n = 12.

C. Models Comparison

As soon as we have selected the possible models who correspond with our information, so much the ARIMA as the BPNN, we realize a comparison of all of them to choose which or which are better. For it we realize the following procedure: first we apply a test ANOVA [6] which allows us to prove the equality of several averages, being based on an analysis of variability (to deduce if the void hypothesis or the alternative is fulfilled). In case the averages are different, we apply BONFERRONI's adjustment [7-8], which we are determined what model is better in the average.

We realize 10 experiments for selected model obtaining 10 information for every index of quality and model. We obtain a matrix for every index of 12 rows (models) for 10 columns (values of indexes).

Finally, models to comparing are:

1) Bootstrap con n = 6.
2) Bootstrap con n = 12.
3) 2-fold Crossvalidation con n = 6.
4) 2-fold Crossvalidation con n = 12.
5) 10-fold Crossvalidation con n = 6.
6) 10-fold Crossvalidation con n = 12.
7) ARIMA (2,0,0) with conditional estimation.
8) ARIMA (2,2,0) with conditional estimation.
9) ARIMA (0,2,1) with conditional estimation.
10) ARIMA (2,0,0) with not conditional estimation.
11) ARIMA (2,2,0) with not conditional estimation.
12) ARIMA (0,2,1) with not conditional estimation.

We apply to every matrix the test ANOVA and obtain that for all indexes, averages are different.

We apply Bonferroni´s joust in different horizons of prediction which compares models two to two and we obtain which have the best average for every index. Tables 7, 8, 9 and 10 show obtained results.

TABLE VII. HORIZON OF PREDICTION OF 10 MINUTES.

Index	Average	Models
$r_{x,\hat{x}}$	0,97	5 7 10 11 8
$IOA_{x,\hat{x}}$	0,98	5 7 10 11 8 1
RMSE	0,57	5 7 10

TABLE VIII. HORIZON OF PREDICTION OF 1 HOUR.

Index	Average	Models
$r_{x,\hat{x}}$	0,88	7 10 11 8
$IOA_{x,\hat{x}}$	0,93	7 10 11 8
RMSE	1,07	7 10

TABLE IX. HORIZON OF PREDICTION OF 2 HOURS.

Index	Average	Models
$r_{x,\hat{x}}$	0,80	7 10 11 8
$IOA_{x,\hat{x}}$	0,89	7 10 11 8
RMSE	1,32	7 10

TABLE X. HORIZON OF PREDICTION OF 4 HOURS.

Index	Average	Models
$r_{x,\hat{x}}$	0,67	7 10 11 8
$IOA_{x,\hat{x}}$	0,80	7 10 11 8
RMSE	1,55	7 10

We verify which are models who repeat themselves in all the indexes and horizons. In this case only there are two models: ARIMA (2,0,0) with conditional estimation and ARIMA (2,0,0) with not conditional estimation.

For the reasoning of Occam's razor we choose the model ARIMA (2,0,0) with conditional estimation for being the easiest to implement and the one that minor time of calculation needs.

III. CONCLUSIONS

In this paper we have identified three ARIMA models which match the short term behavior of wind speed time-series. The model (2,0,0) exhibits the best performance. The validation has been done using three common quality indexes, based in correlation procedures.

The models ARIMA compared with the neuronal networks have given a few very similar results but with a few very low times of calculation.

REFERENCES

[1] Box G.E.P., Jenkins G.M., Time series analysis: forecasting and control, Ed. Prentice-Hall. New Jersey, 1976.

[2] C. Willmott, On the Validation of models, *Physical Geography* **2** (1981), pp. 183–194.

[3] Bishop C.M. Neural Networks for Pattern Recognition, Oxford: Oxford University Press, 1995.

[4] Stone, M. Math. Operationsforsch. Statist. Ser. Statistics 9 (1978) 127.

[5] Hagan M. T, Menhaj M. Training feedforward networks with the Marquardt algorithm, IEEE Transactions on Neural Networks, 1994; 5(6):989-993.

[6] Scheff H. Analysis of variance. New York: Wiley; 1959.

[7] Hochberg Y, Tambane AC. Multiple comparison procedures. New York: Wiley; 1987.

[8] Jobson J.D. Applied Multivariate Data Analysis. Springer Texts in Statistics, Springer-Verlag New-York, 1991;(1).