

EEM 2017 Forecast Competition: Wind power generation prediction using autoregressive models

Ilias Dimoukas
School of Electrical Engineering
Electric Power and Energy Systems
KTH Royal Institute of Technology
Stockholm, Sweden
Email: iliasd@kth.se

Peyman Mazidi
Comillas Pontifical University
Madrid, Spain and
KTH Royal Institute of Technology
Stockholm, Sweden
Email: peyman@comillas.edu

Lars Herre
School of Electrical Engineering
Electric Power and Energy Systems
KTH Royal Institute of Technology
Stockholm, Sweden
Email: lherre@kth.se

Abstract—Energy forecasting provides essential contribution to integrate renewable energy sources into power systems. Today, renewable energy from wind power is one of the fastest growing means of power generation. As wind power forecast accuracy gains growing significance, the number of models used for forecasting is increasing as well. In this paper, we propose an autoregressive (AR) model that can be used as a benchmark model to validate and rank different forecasting models and their accuracy. The presented paper and research was developed within the scope of the European energy market (EEM) 2017 wind power forecasting competition.

I. INTRODUCTION

Wind power is one of the fastest growing means of power generation. In a time of paradigm change in energy policy and intensified competition of renewable energy sources, wind power offers several environmental benefits. However, due to its intermittent nature, the integration of wind power poses challenges on power system operation [1], [2], [3]. In order to minimize power imbalance, appropriate means of forecasting power production are essential. Since forecasting is inherently erroneous, the target is to minimize the forecast error. A vast number of forecast models have been developed in research. The state-of-the-art in wind power forecasting is summarized in [1] and [4]. In the review of [5], five basic types of forecasting models are identified.

Autoregressive (AR) and autoregressive moving average models (ARMA) have been widely used to predict the wind speed or the wind power generation directly. Reference [6] for example, provides a methodology to generate statistically dependent wind speed scenarios based on ARMA modeling. It is concluded that this methodology is accurate in reproducing wind speed historical series. In [7], an ARMA model is developed to predict wind speeds up to a forecast horizon of 10 hours. A different model is developed for each calendar month. Forecasts are proven to perform significantly better in short terms.

A multi-variate ARMA model is presented in [8] for wind power generation forecasting and for simulation of realistic wind speed predictions with adjustable accuracy. In [9], the performance of an autoregressive model and a neural network (NN) model are compared based on the root mean square error. In their analysis, neural networks with a varying number of

hidden layers perform inferior to the autoregressive model. However, this investigation is limited to forecast horizons of 0.5 to 3 hours.

The k-nearest neighbors algorithm (k-NN) has also been used in wind power forecasting [10], [11]. The models developed in the previous references were used with success in the global energy forecasting competition 2014 (GEFCOM'14) [12]. The competition included a wind power forecasting track where participants required to forecast the probabilistic distribution of the wind power generation of 10 wind farms. They were provided with historical wind power generation data and weather forecasts.

This paper presents the autoregressive model (AR) model of the team **4C** which finished second in the European energy market (EEM) 2017 forecast competition [13]. It should be noted that the setting of the competition was special in the way that there was no weather forecast data available. Therefore, in this paper, the forecasting is done based only on the past production data. Two more different approaches were tested for possible use in the competition, based on the k-nearest neighbors algorithm and neural networks. k-NN model is also based only on past wind generation data, and NN model is based on past weather and generation data.

The rest of this paper is structured as follows. In Section II, the competition setting and data are introduced. We describe the procedure of the competition and the data that were provided. The methodology for the AR model is detailed in Section III. Section IV shows an analysis of the competition results. We further conduct an ex-post analysis comparing the results of various AR models and the k-NN and NN models. The conclusions of this paper are given in Sections V.

II. COMPETITION SETTING AND DATA

In cooperation with several technical sponsors, EEM organizers provided historical weather and production data of a wind power plant portfolio. The teams were asked to predict the quarter-hourly wind power generation of a wind power plant portfolio for the next 3-40 hours (152 quarter-hour values) on a daily rolling basis for the competition period of 14 days. The competition started on April 3, 2017 and ended on April 16, 2017. The data were released every day at 08:00

TABLE I
LIST OF VARIABLES

Notation	Explanation
u	zonal component of wind speed
v	meridional component of wind speed
ρ	temperature
H	global radiation
h	height above ground (20,50,75,100)
L	location (1,2,3,4,5,6,7,8,9,10)
y	generated wind power
N	number of observations

and the teams were required to submit the forecasts by 10:00. This means that the forecast period was from the hour 10:00 on the submission day (D) until the hour 24:00 of its next day (D+1). A total of 44 teams joined the competition in the first round where 26 teams continued until the last (14th) round. The timeline of the competition can be seen in Fig. 1.

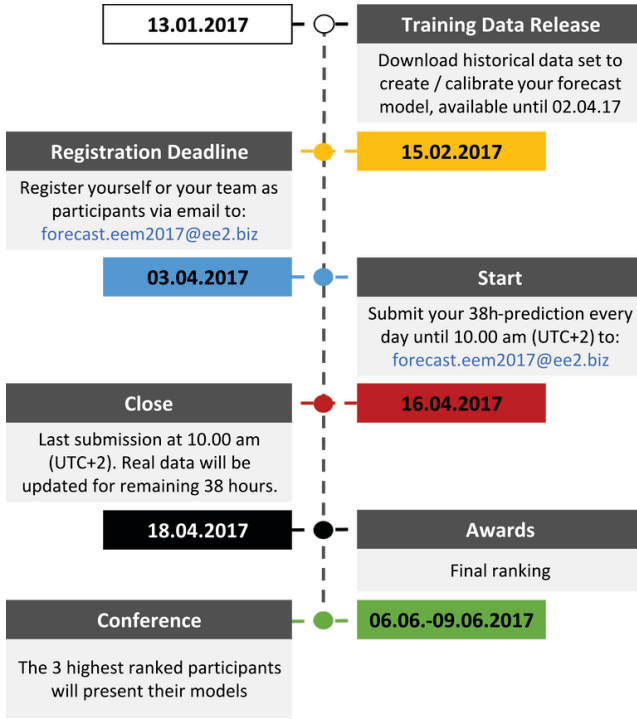


Fig. 1. Timeline of the competition

Two sets of data were provided. The first set included four meteorological variables in a three-hours-resolution. These variables are: zonal component of wind speed vector, meridional component of wind speed vector, temperature and global radiation. Furthermore, two additional inputs were provided: height above ground and location.

There were variable measurements from four height levels (25, 50, 75 and 100 m.) and ten different locations (L_1 , L_2 etc.). Therefore, for each meteorological variable, there were 40 values at each time stamp. The teams were allowed to select as many number of variables as they required.

The second data set included generation of the wind power plant portfolio in MW in a fifteen-minutes-resolution (Fig. 2).

In order to train the model, one year of historical data from both sets were provided.

Accuracy of the forecasts was assessed by two references based on mean absolute error (MAE) measure. MAE score is given by Eq. (1).

$$MAE = \frac{1}{N} \sum_{t=1}^N |X_t - \hat{X}_t| \quad (1)$$

where X_t is the actual value at period t , \hat{X}_t is the predicted value at the same period and N is the number of point forecasts.

The first reference calculates the accuracy of the forecasts for the day D and the second reference calculates the accuracy of the forecasts for the day D+1. Both of these two references have equal weights. The final score is the average of all the MAE scores at each round minus the two worst performances. The team with the lowest cumulative weighted MAE wins the competition.

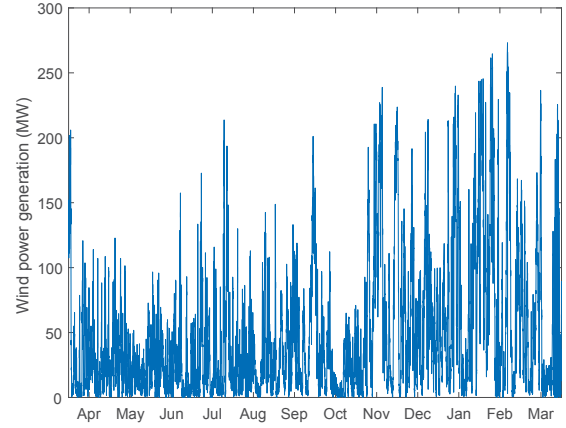


Fig. 2. Competition's wind power generation data in 15-min resolution covering a full year

III. METHODOLOGY

Forecasting procedure consists of three steps: a) Data cleaning, b) Model building and data fitting and c) Monte Carlo simulation of the fitted model. These steps are detailed below.

A. Data cleaning

Data cleaning or data cleansing is the process of detecting incorrect or missing data in the time series. It also includes the detection of outliers and the data smoothing. The Hampel filter included in Matlab [14] is used to remove the outliers and smooth wind power generation data. The effect of this filter can be seen in Fig. 3. The provided data show some steep and rapid changes at several instances. These may be true changes in the wind power generation for some reason or they can be measurement errors. In this work, these step changes were smoothed to increase the fitness of the regressed model.

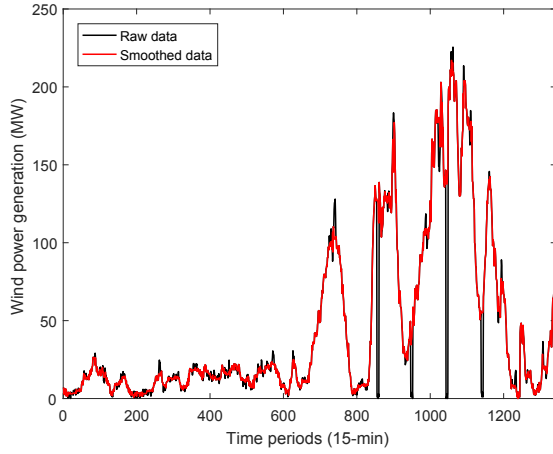


Fig. 3. Data smoothing using the Hampel filter

B. AR/ARMA modeling

AR and ARMA models can be used to model stochastic processes. In an AR model, Eq. (2), the output variable depends linearly on its own previous values and a stochastic term (error term). In an ARMA model, the output variable depends linearly on its own previous values, the error term and past values of the error term as well. An AR model is represented as

$$y_t = c + \sum_{i=1}^p \varphi_i \cdot y_{t-i} + \varepsilon_t \quad (2)$$

and an ARMA model can be written as

$$y_t = c + \sum_{i=1}^p \varphi_i \cdot y_{t-i} + \sum_{i=1}^q \theta_i \cdot \varepsilon_{t-i} + \varepsilon_t \quad (3)$$

where y_t is the output variable (wind power generation) at period t , c is a constant, $\varphi_1, \dots, \varphi_p$ are the autoregressive parameters (order p), $\theta_1, \dots, \theta_q$ are the moving average parameters (order q) and ε_t is the error term, which is a white noise process with zero mean and constant variance σ^2 .

Box-Jenkins methodology is used for building a qualitative forecasting model [15]. This is a trial and error process consisting of four steps:

- 1) Model identification: After applying the logarithmic transformation and integrating the time-series as many times needed to become stationary if needed, the autocorrelation (ACF) and partial autocorrelation functions (PACF) are used to select the orders of p , q .
- 2) Model estimation: Given the orders of the model from the previous step, a least squares or maximum likelihood method is used to estimate the parameters of the model.
- 3) Diagnostic checking: In this step, the error is checked to verify it follows a white noise process. The fitness of the model can also be checked with some tests such as the Bayesian information criterion (BIC) and the Akaike information criterion (AIC). If the tests fail, then the we go back to the first step and modify the initial model.

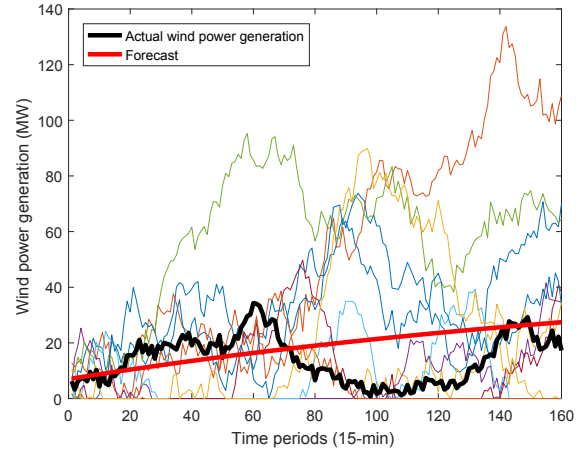


Fig. 4. Wind power forecast and scenario generation using AR model and Monte Carlo simulation

- 4) Forecasting: After the model has been fitted to the historical data, it can be used to forecast future values of the time-series.

C. Monte Carlo simulation and forecasting

Multi-step ahead forecasting can be done using the AR/ARMA model in an iterative way: assuming that y_{t-1} is the last known value, the next period forecast y_t is calculated using Eq. (2) or Eq. (3). Then, the forecast y_t is used for the next prediction y_{t+1} and so on. In each round, the error term is assumed to be zero. This gives a prediction series that starts from a value close to the last known value y_{t-1} and moves towards the mean value of the model. For an AR(1) model, this means that the value is given by Eq. (4). Such a prediction is depicted with the red line in Fig. 4.

$$\mu = \frac{c}{1 - \varphi_1} \quad (4)$$

Monte Carlo simulation is another approach to make a forecast by simulating various scenarios of wind power generation. The procedure is similar as before with the difference that the error term takes some random value picked from a normal distribution with zero mean and σ standard deviation. The average value of all scenarios per period (average scenario) is the forecast. If adequate numbers of scenarios are used, the average scenario approximates the forecast of the previous way. Fig. 4 shows 10 scenarios created with this method. It should be noted that some scenarios for some periods may be negative. Since wind power generation is a physical quantity and cannot take negative values, the negative values of the scenarios are set to zero.

IV. MODEL TESTING AND RESULTS

This section consists of two parts. The first part presents the testing phase which includes the analyses carried out to find a suitable model. It resulted in choosing the AR(1) model during the competition phase. The second part presents the competition results.

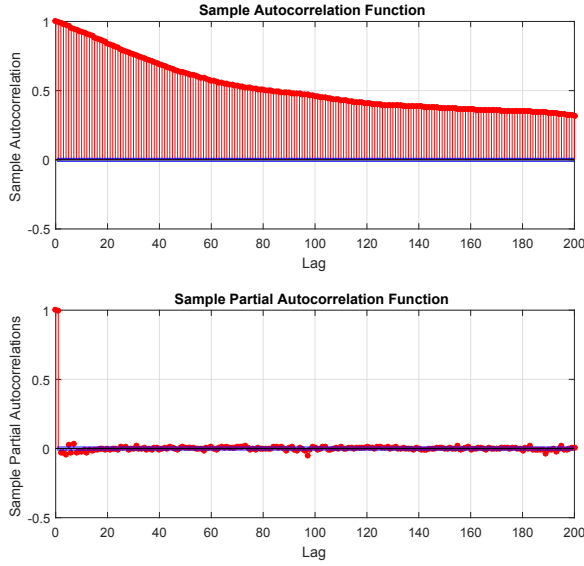


Fig. 5. Autocorrelation and partial autocorrelation functions of wind power generation time series

A. Testing phase

Initially, the training data set was split into two subsets. The first one was used as a training-fitting set for the candidate models and was contained the first 359 days (1-359). The second one was used as an out-of-sample test set to measure the performance of the candidate models. This set contained the last 7 days of the initial training data set (360-366). The Hampel filter was applied to the wind power generation data in the training data set.

To decide about the orders of p and q , the sample ACF and PACF are plotted in Fig. 5. ACF decreases gradually while PACF cuts off after the lag 1. This is a sign of an AR(1) stochastic process. However, there is some minor dependence in the lags 2-5 and 97. The big lag corresponds to 24 hours (24 hours by 4 $\frac{\text{quarters}}{\text{hour}}$). For this reason, apart from the AR(1) model, various other AR and ARMA models were tested. Table II contains the average MAE scores achieved by models AR(1), AR(2), AR(97) and ARMA(1,1) in the training step. After the models were fitted to the data, they generated predictions similar to the competition setting for the whole period of the training set. Table III contains the MAE scores achieved by the aforementioned models in the independent test set. The setting was exactly the same as in the competition. Six rolling predictions were done in days 359-365 covering the period 8:00 of the day 359 till 24:00 of the day 366. The table provides the average MAE scores for each round and the average MAE score of all six rounds.

We can see that the AR/ARMA models in the tables achieve similar results both in the in-sample and out-of-sample sets. We can also see that the score in the in-sample set is lower than the score in the out-of-sample set. This is not so much due to the imposed bias as explained before, but due to the fact

that the training set contains a period with low wind power variability (April-September) that all models achieve very low MAE scores reducing its average value.

Model parameter values are given in Table IV. Since the models achieve similar results, and φ_2 and θ_1 parameters are close to zero in the models AR(2) and ARMA(1,1), the model AR(1) was chosen as the best one for the wind power generation data.

TABLE II
IN-SAMPLE TRAINING SET MAE SCORES

	AR(1)	AR(2)	AR(97)	ARMA(1,1)
MAE	29.40	29.40	29.43	29.40

TABLE III
OUT-OF-SAMPLE TEST SET MAE SCORES

Round	1	2	3	4	5	6	Mean
AR(1)	40.8	55.2	61.5	76.2	54.2	31.5	55.3
AR(2)	41.6	61.5	61.2	73.9	61.6	31.9	55.3
AR(97)	42.0	62.0	61.7	74.0	60.4	30.9	55.1
ARMA(1,1)	41.6	61.5	61.2	73.9	61.6	31.9	55.3
NN	39.5	47.5	76.4	66.3	68.9	39.9	56.4
KNN	36.2	53.0	53.8	120.4	43.2	29.0	55.9

TABLE IV
AR/ARMA MODEL PARAMETERS

Round	c	φ_1	φ_2	φ_{96}	θ_1	σ^2
AR(1)	0.209	0.995	-	-	-	24.23
AR(2)	0.209	0.995	$2.17e^{-5}$	-	-	24.23
AR(97)	0.221	0.995	-	-0.0005	-	24.22
ARMA(1,1)	0.209	0.995	-	-	$-1.11e^{-5}$	24.23

Apart from the AR/ARMA models, models based on the k-nearest neighbors algorithm (k-NN) and the neural networks (NN) were also tested. In short, the forecasting procedure for these models was as follows:

- 1) k-NN: The algorithm was used to find the most similar periods in terms of wind power generation as the last day before the forecasting period. For this reason, only wind power generation data were used. Assuming that a prediction of day D+1 is needed, the algorithm finds the days in the past that have similar wind power generation pattern as day D. Assuming that these days are D-100 and D-12, then the prediction for the day D+1 was the average wind power generation of the days D-99 and D-11.
- 2) NN: The NN was trained to fit past weather data to future wind power generation. For this reason, both weather and wind power generation data were used. Assuming that a prediction of day D+1 is needed, weather variables of the day D are used to make the prediction.

Results of the two methods are given in Table III. We can see that on average, both methods achieve a little worse results compared to the AR/ARMA models. k-NN seems to be more

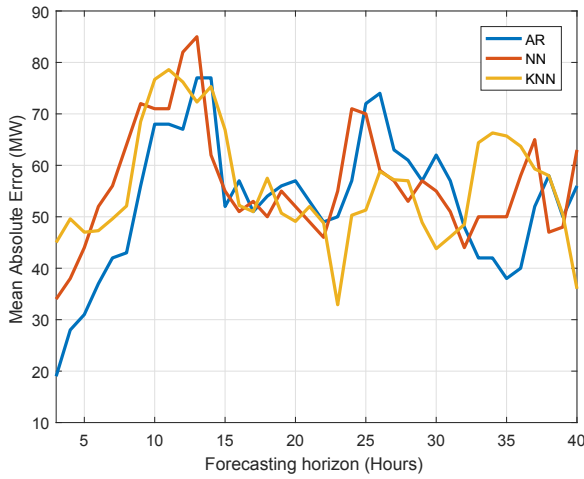


Fig. 6. Average MAE score of AR, NN and k-NN models for each hour of the forecasting horizon

promising as it achieved better results in each round apart from one compared to the other models. Furthermore, Fig. 6 shows the average MAE scores of the six testing rounds for each hour of the forecasting horizon. This is done for models AR(1), NN and k-NN. From the figure, it is obvious that the AR(1) outperforms the other two models in short-term forecasting (up to 6-10 hours ahead). For the rest of the period of the forecasting horizon (up to 40 hours ahead) all models achieve similar scores.

Based on the above analysis of the testing results, the AR(1) model was chosen for the competition phase.

B. Competition results

The MAE results for each round are given in Table V. The final cumulative average MAE score is 20.5070. Fig. 7 shows the actual wind power generation during the 14 rounds of the competition and the forecasts that were submitted. It is clear that the forecast series for each round start with a value close to the last known value and converges towards the model's mean value which is approximately 41.8. Table VI presents the cumulative average MAE scores of the 5 best teams in the competition.

TABLE V
COMPETITION RESULTS - ROUND MAE SCORES

Round	1	2	3	4	5
MAE	15.7667	26.2059	13.8404	25.4674	40.9869
Round	6	7	8	9	10
MAE	29.8597	32.6275	9.7445	16.6536	30.3870
Round	11	12	13	14	
MAE	11.2611	17.2053	17.0645	34.1513	

V. CONCLUSIONS

This paper presents the methodology used by team 4C to produce wind power generation predictions in the EEM17 forecasting competition. The forecasting method is based on

TABLE VI
COMPETITION RESULTS - CUMULATIVE MAE SCORE OF FIRST 5 TEAMS

Team rank	1	2	3	4	5
MAE	19.5906	20.5070	20.5590	21.6042	22.1197

an autoregressive model, AR(1). The advantage of using such model is its simplicity, the minimalized requirements for only past wind power generation data and the possibility to create wind power generation scenarios using Monte Carlo simulation. The model is quite good for short-term forecasting while it gives average results for longer forecasts. For this reason, this model can be used as a benchmark tool for more sophisticated models. On the other hand, the forecasts produced by this model cannot capture the wind power variations, though it is possible to create various scenarios.

ACKNOWLEDGMENT

Ilias Dimoulkas and Peyman Mazidi would like to thank SweGRIDS [16], the Swedish Centre for Smart Grids and Energy Storage and UPGRID project [17] respectively for funding this work. Lars Herre would like to thank the Swedish Energy Agency for funding this work.

REFERENCES

- [1] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The state-of-the-art in short-term prediction of wind power: A literature overview," ANEMOS. plus, Tech. Rep., 2011. [Online]. Available: http://orbit.dtu.dk/fedora/objects/orbit:83397/datastreams/file_5277161/content
- [2] R. Rajagopal, E. Bitar, P. Varaiya, and F. Wu, "Risk-limiting dispatch for integrating renewable power," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 615–628, Jan. 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0142061512004085>
- [3] W. Zhou, H. Sun, and Y. Peng, "Risk Reserve Constrained Economic Dispatch Model with Wind Power Penetration," *Energies*, vol. 3, no. 12, pp. 1880–1894, Dec. 2010. [Online]. Available: <http://www.mdpi.com/1996-1073/3/12/1880/>
- [4] C. Monteiro, H. Keko, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, and INESC Porto, "A quick guide to wind power forecasting : state-of-the-art 2009." Tech. Rep. ANL/DIS-10-2, 968213, Nov. 2009, doi: 10.2172/968213. [Online]. Available: <http://www.osti.gov/servlets/purl/968213-QPwKb9/>
- [5] M. Lei, L. Shiyang, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, May 2009. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S1364032108000282>
- [6] J. M. Morales, R. Míguez, and A. J. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, vol. 87, no. 3, pp. 843–855, Mar. 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261909004140>
- [7] J. L. Torres, A. García, M. De Blas, and A. De Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)," *Solar Energy*, vol. 79, no. 1, pp. 65–77, Jul. 2005. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0038092X04002877>
- [8] A. Boone, "Simulation of short-term wind speed forecast errors using a multi-variate ARMA (1, 1) time-series model," Ph.D. dissertation, 2005. [Online]. Available: <http://www.diva-portal.org/smash/record.jsf?pid=diva2:604579>
- [9] A. H. Nielsen and H. Madsen, "Wind Power Prediction Using ARX Models and Neural Networks," 1996. [Online]. Available: <http://www.xenfor.dk/pub/bib/paprep/public/arx-nn-iasted-1996.pdf>

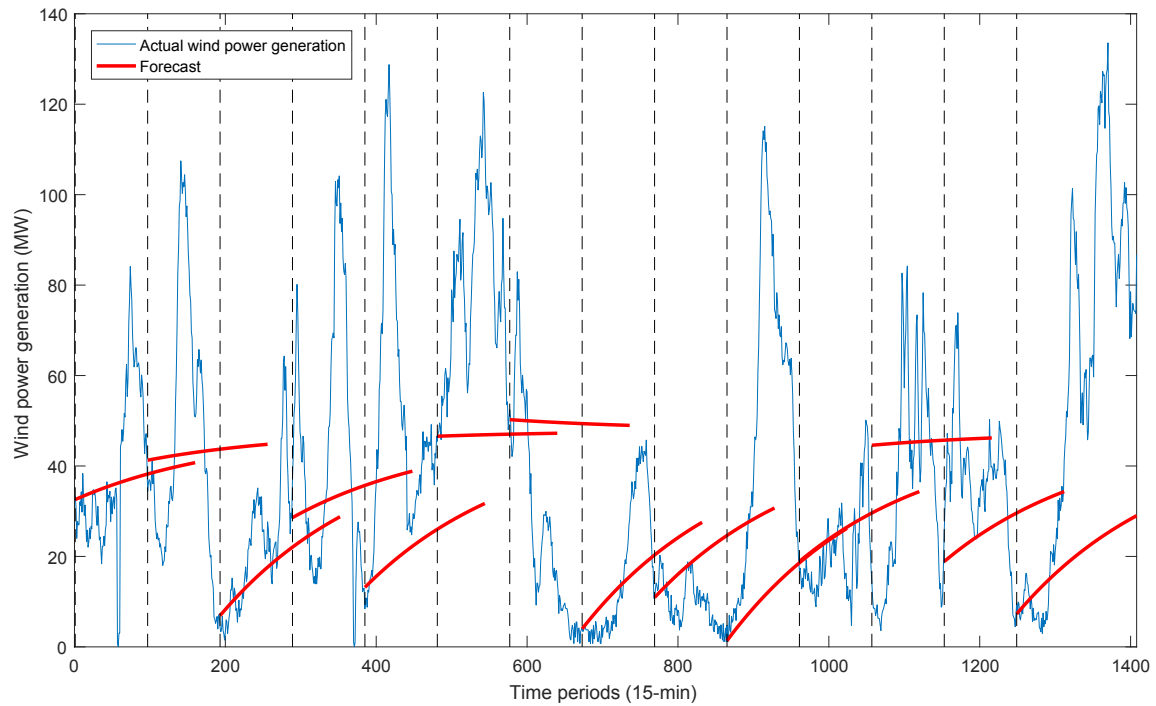


Fig. 7. Actual wind power generation and forecasts during the 14 rounds of the competition

- [10] E. Mangalova and O. Shesterneva, "K-nearest neighbors for GEFCom2014 probabilistic wind power forecasting," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1067–1073, Jul. 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207015001429>
- [11] Y. Zhang and J. Wang, "K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1074–1080, Jul. 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207015001417>
- [12] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, Jul. 2016. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0169207016000133>
- [13] "EEM17 Forecast Competition," Dec. 2016. [Online]. Available: <http://eem2017.com/program/forecast-competition>
- [14] "Outlier removal using Hampel identifier - MATLAB hampel." [Online]. Available: https://www.mathworks.com/help/signal/ref/hampel.html?s_tid=gn_loc_drop
- [15] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: forecasting and control*. Hoboken, N.J.: John Wiley, 2008.
- [16] "KTH | SweGRIDS." [Online]. Available: <https://www.kth.se/en/ees/omskolan/organisation/centra/swegrids>
- [17] "Real proven solutions to enable active demand and distributed generation flexible integration, through a fully controllable low voltage and medium voltage distribution grid." [Online]. Available: <http://upgrid.eu/>