Deep learning for wind power production forecast

Alla Sapronova[1,*], Klaus Johannsen[1], Eirik Thorsnes[1], Catherine Meissner[2], Matteo Mana[2]

[1] Center for Big Data Analysis, Uni Research AS,Bergen, Norway
[2] WindSim AS, Tonsberg, Norway

alla.sapronova@uni.no

Abstract. An accurate prediction of wind power output is crucial for efficient coordination of cooperative energy production from different sources. Long-term ahead (from 6 to 24 hours) prediction of wind power for onshore wind parks can be achieved by using a coupled model that would bridge mesoscale weather prediction data and computational fluid dynamics. When a forecast for shorter time horizon (less than 1 hour) is anticipated, an accuracy of predictive model that utilizes hourly mesoscale weather data is decreasing as the higher frequency fluctuations of the wind speed are lost when data is averaged over an hour. The analysis of wind speed fluctuations over periods from minutes to hours shows that higher frequency variations of wind speed and direction have to be taken into account for accurate short-term ahead energy production forecast. In this work a deep learning approach was used for accurate wind power output forecast for 5 to 30 minutes ahead. The model uses historical park production time series and hourly numerical weather forecasts to predict the total power production of the wind park.

1. Introduction

Currently, wind energy plays a significant role in total electricity production and it is vital to predict the wind energy output timely and accurately. The variability of the wind is a main challenge for obtaining an accurate forecast for minutes to hours ahead. Many studies have been devoted to the improvements of wind forecasting techniques with a number of models developed and launched. These models are based on physical, statistical or hybrid approaches and often using Numerical Weather Pre-

diction (NWP) data at tens of kilometers resolution as a main component. But NWP based models often are not providing satisfactory accuracy for site-specific and short horizon forecasts. For site-specific forecasts the mesoscale-microscale coupling model was proposed in [1]: there artificial intelligence methods were used to issue 1 to 3 hours ahead wind forecast from NWP data and CFD module was used to calculate the flow at finer scales [5]. The accuracy of the coupling model [1] was found to be superior comparing to that based on polynomial fittings as well as ARMA models, although accuracy for shorter than one hour forecast remains unsatisfactory. In this work, the model proposed in [1] is adapted to forecast a wind power output for less than one hour ahead by employing deep learning. The time series of historical power production, NWP data and categorization approach [4] are submitted as inputs to the model.

2. Model and Results

The model for short-time ahead prediction of the total wind power production is based on time series analysis of historical park production and NWP data. Since the quality of forecast depends on ability of the model to predict the wind flow near the ground in the complex terrain, where the roughness and complexity affect the flow at microscale, the model employs machine learning techniques as those are proven to be efficient for nonlinear multivariable functions approximation when explicit physical based models have limited application or not available.

In this work, deep learning has been used to forecast the total wind park power production 5 to 30 minutes ahead. The modules are trained, tested and validated on 6 month historical data from several on-shore wind parks from Sweden and Norway. The historical time series include nacelle wind speed, wind direction, and power production data at sub-seconds frequency as well as hourly NWP data. Two different sets of input variables were used. For one set, the model used the historical power production values as inputs and obtained as output the total power production minutes ahead. And for another set, the model used the time-series of multiple environmental variables as inputs, like wind speed and direction time-series and hourly NWP data, in addition to historical power production time-series.

The length of a time window was two times longer than the prediction horizon. For example, for 5 minutes ahead prediction the past 10 minutes of historical production were used as inputs.

The pre-processing included: data cleaning, normalization, and time intervals averaging. The data was averaged by 1 minute for 5 to 10 minutes ahead prediction, and 5

minutes averages were used for 15 to 30 minutes ahead prediction. The data set was split in the ratio 70:20:10 for training, testing and validation accordingly.

To measure the differences between values predicted by a model and the values actually observed, the normalized root mean squared error (root mean squared percentage error, RMSPE) has been used. As the negative and positive errors don't cancel out each other and the higher weight is given to larger errors in RMSPE, it makes an excellent general purpose error metric for numerical predictions: the smaller the RMSPE, the better fit of the model.

The performance of the new model has been compared to the data categorization approach based model. This approach is described in [4] and suggests that the entire data-set is grouped into several discrete categories which allow identical category values to be treated in the same manner for non-equal continues numerical data.

Previously it was shown [6] that the selection of methods for categorization is not critical, so in this work the wind speed was categorized. The information obtained from categorization of wind speed variable was supplied to neural network as input in addition to time series variables.

New model with deep learning approach shows higher accuracy of power prediction for 5 and 10 minutes ahead, comparing to previously reported results when using only a feed forward artificial neural network (ANN) [4] or linear regression approaches, as shown in Table 1.

The new model performance is not affected by lack of NWP data for very short (up to 20 minutes) or longer (30 minutes) prediction horizons as.

One of the model's modification also employs categorized wind speed data: a better performance is shown for such models with no regards whether NWP data was present or not.

Table 1. Comparison of the model's accuracy for 5 and 10 minutes ahead forecasting for test and validation runs.

	Deep learning		Feed forward ANN		Linear regression	
	5 min ahead	10 min ahead	5 min ahead	10 min ahead	5 min ahead	10 min ahead
Test run						
Mean absolute per-	3.64	3.68	4.77	4.78	5.24	5.27

centage error						
Root mean square error, percent	5.3	5.32	6.02	6.02	7.36	7.37
Validation run						
Mean absolute percentage error	4.35	4.35	8.58	8.61	9.54	9.57
Root mean square error, percent	6.21	6.21	9.1	9.17	12.32	12.8

3. Conclusions

Deep learning approach has been used to develop a new model that forecasts wind park energy yield 5 to 30 minutes ahead. The model shows higher accuracy comparing to models based on feed forward neural network or linear regression methods. The deep learning model can also take into account the categorization approach proposed earlier by authors. Indeed, using deep learning allows data of various frequency and nature to be used and is robust for lack of data from one of the input channels (temporary or constant), e.g. NWP data or categorization data.

The new model shows that use of NWP data does not significantly improve the accuracy for both very short time ahead forecast (5-10 minutes ahead) and 30 (and longer) minutes ahead forecasts.

It is observed that when categorization of one of the input variables is used (wind speed, for instance), the better accuracy of the forecast can be achieved with shorter time window for historical time series variables (like power production).

4. Acknowledgment

5. References

A. Sapronova, C. Meissner, M. Mana, "Mesoscale-microscale coupled model based on artificial neural network techniques for wind power forecast," EWEA Offshore 2014, PO ID 385 (2014)

A. Sapronova, NORCOWE report, NORCOWE-RR-C-14-WP2-018 (2014)

C.L. Vincent, P. Pinson, G. Giebela, "Wind fluctuations over the North Sea," Int. J. Climatol. 31: 1584–1595 (2011)

A. Sapronova, C. Meissner, M. Mana, "Improving an Accuracy of ANN-Based Mesoscale-Microscale Coupling Model by Data Categorization: With Application to Wind Forecast for Offshore and Complex Terrain Onshore Wind Farms," Lecture Notes in Computer Science, 2014, Volume 8817, pp.61-66

M. Mana, "Short-term Forecasting of wind energy production using CFD simulations," EWEA 2013 (2013)

L. Li, A. Pratap, H-T. Lin, Y.S. Abu-Mostafa, "Improving generalization by data categorization," In: A.M. Jorge, L. Torgo, P.B. Brazdil, R. Camacho, J. Gama,(eds.) PKDD 2005. LNCS (LNAI), vol. 3721, pp. 157–168. Springer, Heidelberg (2005).