Short time ahead wind power production forecast.

A.Sapronova* (1), C.Meissner (2), M.Mana (2)

(1) UniResearch AS, (2)WindSim AS

Abstract.

An accurate prediction of wind power output is crucial for coordinating the cooperative production of different energy sources more efficiently. Accurate prediction for wind power output long-term ahead (from 6 to 24 hours) for onshore wind parks can be achieved by employing a model coupling mesoscale weather prediction data and computational fluid dynamics as presented in [1]. When a shorter time ahead (less than 1 hour) forecast is anticipated, accuracy of predictive model that uses hourly mesoscale weather data decreases, as the higher frequency fluctuations of the wind speed are lost when data averaged over an hour. It is observed in [2] that wind speeds can vary up to 50% in magnitude over a period of 5 min and the analysis of wind speed fluctuations over periods from 1 min to hours performed in [3] has shown that higher frequency variations of wind speed and direction have to be taken into account for accurate short-term ahead energy yield forecast.

In this work a new model for accurate wind power output forecast from 5 to 30 minutes ahead is presented. The model uses historical park production time series, rather than mesoscale weather data analysis; and employs machine learning techniques and categorization approach as proposed in [4]. The model was tested and validated on real wind park production data.

Approach.

Proposed model for short-time ahead energy output prediction based on time series analysis of historical park production. Since the quality of forecast depends on ability of the model to predict the wind flow near the ground in the complex terrain, where roughness and complexity affect the flow at microscale, the model employs machine learning techniques as they are proven to be efficient for nonlinear multivariable functions approximation when explicit physical based models have limited application or not available. In this work a supervised learning methods combined with kernel methods to increase model's accuracy. The model uses time series of historical power production to issue 5 to 30 minutes ahead forecast and was trained, tested and validated on data from several on-shore wind parks from Sweden. The data preprocessing for machine learning included data cleaning, one- or 5-minutes time intervals averaging, and normalization.

Main body of abstract.

The model comprise of artificial neural network that uses historical park power production timeseries of various length to issue a prediction of power output 5 or 30 minutes ahead.

At first, two feed forward artificial neural networks (ANN) have been constructed: one has used inputs time series from multiple environmental variables, and second has used inputs time series from a single variable - wind power production.

Each ANN has been trained with backpropagation method and Encog Workbench open source software package [6] has been used for model training and validation.

Both ANNs architecture and performance are shown in Table 1.

	Several-variables time series	Single-variable time series
Validation root mean square percentage error (RMSPE)	14.9	9.7

Table 1. ANN architectures and performances for short time ahead wind power forecast.

ANN that utilizes time series from a single variable has been selected for further modification to increase the forecast accuracy.

A double-layer model has been suggested to improve forecast accuracy: the original ANN has been used to issue a “coarse” prediction that was submitted as one of the inputs of the “corrected” model. With double-layer approach RMSPE has been lowered 5.9.

Also the model that uses supervised machine learning methods has been compared to model that uses kernel methods.

The performance of the double layer model was also compared with model uses data categorization approach. 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 [5] 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 ANN as input in addition to time series variables. The model with categorization predicts power output with RMSPE 3.3 for on the validation data sets. The RMSPE on validation data set for all the above mentioned models summarized in Table 2.

Single layer model, single variable	Single layer model, multiple variables	Double-layer, supervised learning model	Double-layer, kernel methods model	Data categorization model
14.9	9.7	5.9	3.3	3.3

Table 2. The error (RMSPE) for short time ahead power output forecasting of different models.

Conclusions.

- In this work a new model for accurate 5-30 minutes ahead wind power output forecast is presented.
- The model uses historical production time series and employs machine learning methods and categorization approach.
- The best performance is shown for combination of supervised learning method and kernel method for double layer model.
- If categorization of an input variable is used, the same accuracy of the model can be achieved with smaller number of hidden neurons. It is discussed that ANN based model's generalization arises from the model's ability to find similarity in the training data that usually consists of continuous numeric data. Since numbers are rarely exactly the same from one example to the next, the model can fail in selecting the margins for identical properties. In this case, the generalization can be improved by classification.

It is shown that choice of methods for categorization is irrelevant to the generalization improvement. Methods, quite different by nature, give similar and reasonable results and all lead to improved model's generalization.

Learning objectives.

The model allow accurate wind power output prediction short time ahead. As model uses only park production data, no weather forecast data are required. The model also robust to uncertainty in the data sets and therefore can be used in a real time application where short periods of cutoffs can be expected.

References.

1. A.Sapronova, C.Meissner, M.Mana, Mesoscale-microscale coupling model. 2014

2. A. Sapronova, NORCOWE report: NORCOWE-RR-C-14-WP2-018
3. C.L.Vincent, P.Pinson, and G.Giebela, Wind fluctuations over the North Sea, Int. J. Climatol. 31: 1584–1595 (2011)
4. A.Sapronova, C.Meissner, M.Mana, Categorization approach. 2015
5. L.Li, A.Pratap, Hsuan-Tien Lin, and Y.S.Abu-Mostafa, Improving Generalization by Data Categorization., Lecture Notes in Computer Science Vol. 3721, 2005, pp 157-168.
6. http://www.heatonresearch.com/wiki/Encog_Workbench

Acknowledgment.

This work is sponsored by the Norwegian Research Council, project ENERGIX, 2015-2016.