A concept for Wind Condition Estimation using TensorFlow and Random Forest models

Jeffrey Stegink, Jörn Froböse, János Zierath

W2E Wind to Energy GmbH
Grubenstraße 44, D-18055 Rostock,
jstegink@wind-to-energy.de

EXTENDED ABSTRACT

1 Introduction

The increasing importance of Artificial Intelligence (AI) is reflecting in the growing number of applications such as the use on Wind Turbines (WT) to forecast power yields based on wind speed and to predict necessary maintenance. Due to the energy transition and the associated development to build more and more powerful WTs, the requirements for structural integrity with respect to the maximum lifetime will be a challenging task. In order to meet these requirements, the load-reducing control strategies are being modernized and optimized. While AI is already being used for model optimization for multibody simulations, this approach is not yet applied to WTs [1].

2 Problem description

The anemometers used so far to measure the wind conditions at hub segmentat are disadvantageous with respect to their position behind the rotor. In this case, the wind conditions are influenced by the blade root, which can be several metres in diameter [3]. Therefore the aims were to use AI-based regression to analyze the system behaviour of wind turbines and identify the wind conditions in a data-driven manner. This includes the creation of a well-defined data set taken from SCADA and the development of automated AI-based (partial) system models.

3 Methods

As a first step, a multibody model, which was built up in alaska/Wind, already established by [2] was used for data acquisition. Different wind models were applied to simulate. For the different wind models, the parameters wind speed (4 m/s to 25 m/s, 0.5 m/s stepsize), turbulence intensity (8% to 18%, 1% stepsize) and the yaw error (-20° to 20°, 1° stepsize) were varied in four different seeds, whose purpose was the prediction by the AI model based on the different outputs of the simulation (label). Overall, more than 77.000 loadcases were simulated and evaluated. As an output the simulation generates time series of various sensors that reflect the system status. From these time series, the features are extracted in form of mean values, standard deviation, minima, maxima, median, first quartile as well as third quartile (see Figure 1).

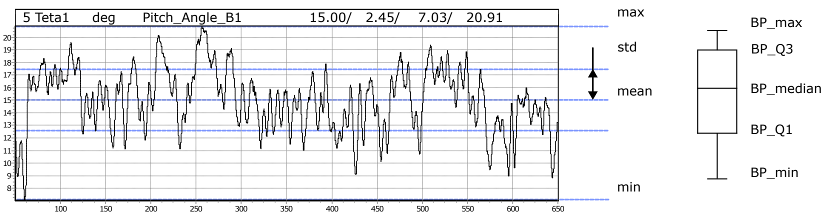

Figure 1: Representation of the outputs from the simulation and which features are taken from the time series.

Initially, 9 different sensors for which data can be measured or derived on a real system were evaluated. These include, among others: generator speed, rotor speed, generator torque, power.

To analyze the features and to select them with respect to their significance, the ensemble method Random Forest was used. The advantage of this method is that overfitting is avoided due to training partitioning of the data set. In addition to feature selection, the Random Forest model also predicts test features. For each label, the importance of the features is examined individually, by evaluating the feature that has been used most frequently on the path of the decision tree. After the most important features have been selected, they can be used for model training. First, the tree-based Random Forest model was trained. For the setup of this model the library of SciKit-Learn was used. Random Forest works like an ensemble of decision trees, which is combined into a stable model by a large variance, thus preventing overfitting of the model. As a second model, a Deep Neural Network was established by using the TensorFlow library. TensorFlow is based on a computational graph and attempts to establish a relationship between input and output variables during training. In order to build an efficient model, varying models with different numbers of hidden layers, numbers of nodes, and learning rates were tested.

4 Results

As described in the methods, the features were extracted from the simulation results during the postprocessing and entered into the data catalogue. The SciKit-Learn library was used to determine the meaningful features. Out of 90 features, the four most influental parameters for the prediction of wind speed had an combined informative value of 93,91%. Meanwhile, the four most influential parameters for turbulence intensity reach a combined informative value of 70.10% and for the yaw error 78.28%.

Figure 2 depicts the distribution of the predictions for the wind conditions, showing the deviation for the TensorFlow model in orange and that for the Random Forest model in blue. For the TensorFlow model, 90% of the estimates for wind speed were between -0.17 m/s and 0.34 m/s, for turbulence intensity between -0.46% and 0.37% and for yaw error between -2.00° and 2.58°. Concerning the Random Forest model, 90% of the estimates for wind speed were between -0.16 m/s and 0.14 m/s, for turbulence intensity between -0.36% and 0.37% and for yaw error between -1.6° and 1.7°.

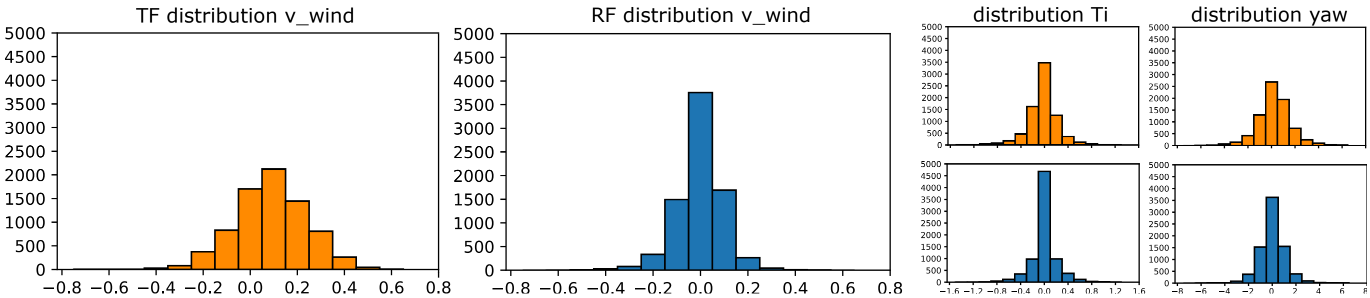

Figure 2: Plot of the distribution of the predictions of the Random Forest model (blue) and the TensorFlow model (orange) for wind speed, turbulence intensity and yaw error

5 Discussion

With the use of AI-based models, the current wind conditions can be accurately estimated. With this, the problems described in 2 with the use of an anemometer can be well circumvented. Since the machine learning models work with the system parameters and are trained on them, it could be used as a support or even a replacement for the anemometer. Despite measuring the yaw error with expensive lidar measurement systems, the standard deviation is still around 15° yaw error [4, 5]. This can be reduced to 1.7° yaw error through the use with the provided AI-based model, which greatly increases the power output of a wind turbine and reduces the loads.

6 Conclusion

This demonstrates that the AI-based models can be a profitable factor on the wind turbine, on the one hand to save sensor technology which minimizes the erection costs and on the other hand to increase the power yield through accurate wind condition estimation.

Acknowledgments

The authors would like to thank the German Federal Ministry of Education and Research.

References

[1] Poliakov, Vladimir: The Artificial Intelligence and Design of Multibody Systems with Predicted Dynamic Behavior. 14 (2020), S. 972–977

[2] Zierath, J. ; Rachholz, R. ; Woernle, C.: Field test validation of Flex5, MSC.Adams, alaska/Wind and SIMPACK for load calculations on wind turbines. In: *Wind Energy* 19 (2016), S. 1201–1222

[3] Deiss, Olga ; Lackmann, Frank ; Hilling, Christoph ; Kameier, Frank: Turbulenzeinflüsse bei der Messung der Windgeschwindigkeit. (2001), 01

[4] Zhang, Le ; Yang, Qiang: A Method for Yaw Error Alignment of Wind Turbine Based on LiDAR. In: *IEEE Access* PP (2020), 01, S. 1–1

[5] Zhao, Han ; Zhou, Lawu ; Liang, Yu ; Zhang, Shuowang ; Ma, Mianzhuo: A 2.5MW Wind Turbine TL-EMPC Yaw Strategy Based on Ideal Wind Measurement by LiDAR. In: *IEEE Access* PP (2021), 06, S. 1–1