Forecasting of Wind Turbine Output Power Using Machine learning

Haroon Rashid
Sustainable Environment and Energy Systems
Middle East Technical University
Northern Cyprus Campus
Kalkanli, Guzelyurt, Mersin 10, Turkey
haroon.rashid@metu.edu.tr

Waqar Haider
Department of Computer Engineering
Middle East Technical University,
North Cyprus Campus
Kalkanlı, Guzelyurt, Mersin 10, Turkey
waqar.haider@metu.edu.tr

Canras Batunlu
Department of Electrical Engineering
Middle East Technical University
North Cyprus Campus
Kalkanlı, Guzelyurt, Mersin 10, Turkey
cbatunlu@metu.edu.tr

***Abstract*— Most of the countries around the world are facing huge environmental impact, and the most promising solution to mitigate these is the use of renewable energy, especially wind power. Though, the use of offshore wind energy is rapidly increasing to meet the elevating electricity demand. The researchers and policymakers have become aware of the importance of providing near accurate prediction of output power. Wind energy is tied to variabilities of weather patterns, especially wind speed, which are irregular in climates with erratic weather conditions. In this paper, we predicted the output power of the wind turbines using the random forest regressor algorithm. The SCADA data is collected for two years from a wind farm located in France. The model is trained using the data from 2017. The wind direction, wind speed and outdoor temperature are used as input parameters to predict output power. We test our model for two different capacity factors. The estimated mean absolute errors for the proposed model in this study were 3.6% and 7.3% for and 0.2 capacity factors, respectively. The proposed model in this study offers an efficient method to predict the output power of wind turbine with preferably low error.**

***Keywords*— power, prediction, machine learning, energy, wind speed**

I. Introduction

In recent years, the wind energy sector increase in strength. For example, the installed wind energy capacity in the European Union (EU) is 142 GW [1]. In the year 2019 651 GW wind energy was installed in world. Fig. 1 show the wind energy installed in world from the year 2001-2019.

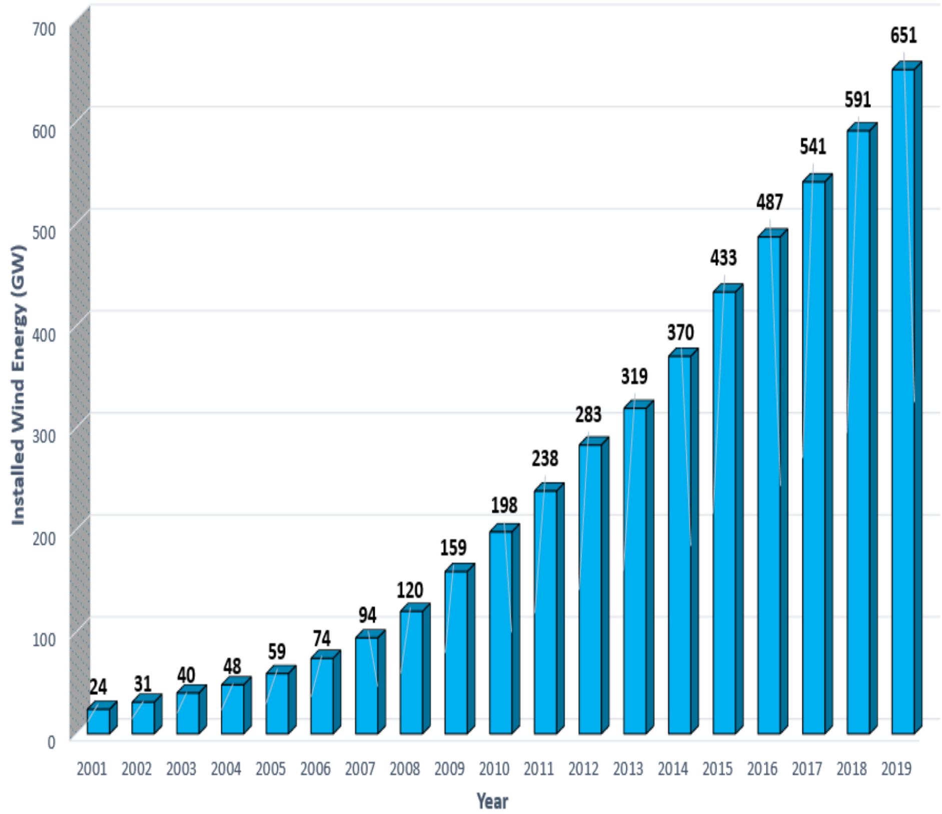

Fig. 1 Global wind energy installation during 2001-2019 [2]

The intermittency of wind speed introduces challenges to the prediction of wind power operation during energy integration. This results in challenges associated with planning and regulation capabilities associated with sudden wind speed variations, which impacts on the reliability of power system predictions. Wind power generation and reliability planning rely on fast and robust wind speed prediction and response to system dynamics for better wind power prediction [3]. The global energy report shows that power generation from the wind rose to 54.6 gigawatts (GW) of installed capacity in 2018. China and the USA are leading with installed capacity.

Countries like Germany and India are showing a strong appetite for wind energy generation due to useful wind speed prediction capability [4].The role of wind power prediction is becoming increasingly crucial, while the wind penetration rate is continuously growing. Every power system has a reasonable capability to adapt demand changes as the demand estimation has never been entirely accurate. When the penetration rate is relatively low, power systems do not have to pay too much attention to the variance of wind power supply.

Machine learning is a field of computer science that focuses on improving the performance of the program by itself with experience. In this technique, the machine is not told how to solve the problem explicitly; instead, the experience is given to the device as different inputs, and the outputs are typically a model that can address future issues of the same kind. The complete procedure of machine learning includes several steps. First, past experience is usually gathered for the training in the later stage. At that point, the type of a unique target function is resolved, which portrays the relationship between inputs and outputs. From that point onward, an machine learning model is chosen to estimate the target function. At last, a fitting calculation is utilized to assemble the model from the training models [5]. The process of machine learning is summarized in Fig 2.

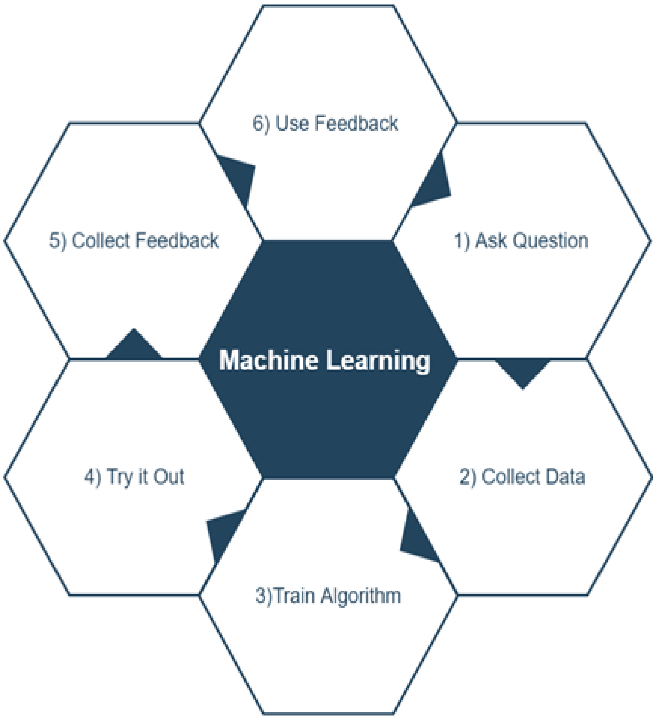

Fig. 2 Summary of machine learning

The elimination of variability of the wind is not possible, but by using machine learning, we can make wind power sufficiently more predictable and valuable. This will also help bring greater data rigor to wind farm operations, as these techniques can help wind farm operators make faster, more data-driven, and smarter assessments of how their power output can meet electricity demand [6]. The primary aim of this research work to predict wind turbine output power using machine learning techniques. Since there is a very limited study has been done on the wind turbine power prediction using machine learning it is greater interest to us how well it performs in this field.

The remaining sections of this study are structured as follows: In section II we will discuss the literature review related to power prediction in wind energy. Section III briefly describe the methodology of this study. Section IV presents the results and discussion of the study and finally, in section V this study will conclude with the future work direction.

II. LITERATURE REVIEW

This section will review the different studies done so far to predict wind turbine output power. The application machine learning on wind energy especially on output power prediction will briefly be described.

Guo et al. [7] performed the multi-step forecasting of wind speed using an ensemble or combination of two models EMD and FFNN. For each of these techniques, the nonlinear wind speed is decomposed into small chunks. The continuing the counterpart EMD enhanced insights on the data structures involving monthly mean wind speed data over three years. The performances of these models is evaluate using mean absolute error, mean square error and some trials independently.

The authors in [8] introduce Bayesian forecasting based on a truncated model approach, which can join area information about wind information. The model is applied for ultra- transient wind speed forecast. The direct attributes of the exhibited structural break strategy limit the capacity of this model to address all the more testing forecast issues with longer estimating time horizons.

Kanna et al. [9] proposed an ”adaptive wavelet neural network for mapping the NWP’s wind speed and wind direction forecasts to wind power forecasts. Wind direction inherently being a circular variable, for better training and function approximation, a transformed version of wind direction variables are used as inputs. Further, the closest set of patterns based on Euclidean distance are chosen for training patterns and block- wise training and forecast strategy is employed for carrying wind power forecast. The results show that the significant improvement over the persistence method is achieved.

Lionel et al. [10] compare various algorithms for wind power forecasting and show that random forests with and without random input selection yield a prediction performance similar to SVR, but recommend to prefer a linear model when the computation time grows too large.

III. METHODOLOGY

This section will explain the methodology used in this research work. Firstly we will discuss the wind turbine data collection. The data obtain from SCADA is preprocessing in the second step before the data analysis. Thirdly the useful features are selected from the data. In fourth step machine learning model is train using the data. In last step the wind turbine power is predicted. The complete flow of methodology is show in Fig 3.

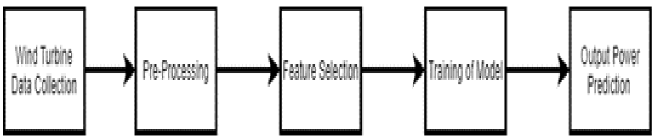

Fig. 3 Methodology of the study

A. *Wind Turbine Data Collection*

In every wind turbine farm the Supervisory control and data acquisition (SCADA) is already installed which collect the data from different sensors. The dataset for this research work is obtain from the wind farm in France. The wind farm consist of 4 wind turbines. The data contain data from 1st January 2017 to 29th January 2018 for each turbine. The collected data is recorded after every 10 minutes. All installed wind turbines have same specifications which is show in Table I.

TABLE I WIND TURBINE SPECIFICATIONS USED IN THIS STUDY

Rated Power	2050kW
Cut-in Wind Speed	3.5 m/s
Cut-out Wind Speed	25 m/s
Nominal Wind Speed	14.5 m/s
Diameter	82 m
Rotor area	5281 m2
Length of rotor blades	40 m

B. *Data pre-processing*

The model accuracy of wind turbine is highly effected by inaccurate SCADA data caused by null entries and irregular operation. Therefore it is really important to pre-process this data which will remove these confusing entries before further analyses. The SCADA data collected have mismatch in date and time which is fixed using Python. The data is further cleaned by applying the cut in and cut out wind speed limit. The data contain some negative output power which is unphysical value which is removed before further analysis. The pitch angle of wind turbine also consist of some irregular value which normally is very low or slightly negative so before further analysis the data is excluded. The wind speed and Power of wind turbine after preprocessing is show in Fig 4 and Fig 5 respectively.

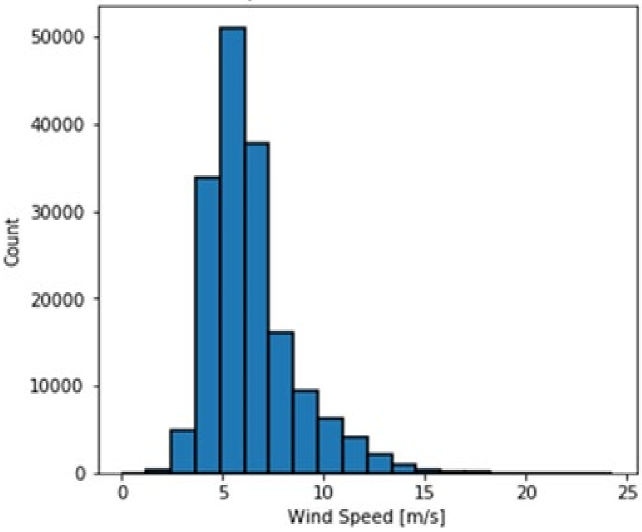

Fig. 4 Wind speed after pre-processing

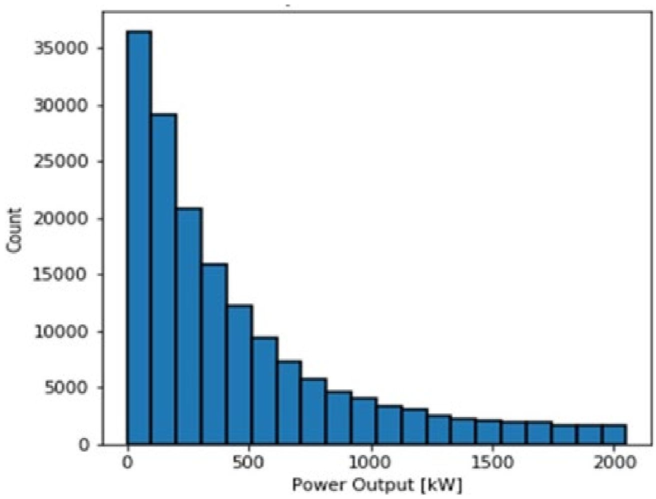

Fig. 5. Output power of wind turbine after preprocessing

C. *Feature Selection*

The features selected for this model are wind turbine generator ambient (external) conditions. The parameters selected for this research are absolute wind direction (Wa), wind speed (Ws) and outdoor temperature (Ot). The average, minimum, maximum and standard deviation values of all four features are consider while analysis.

D. *Training of Model*

The dataset after pre-processing and features selection divided into test and training set. The training model is shown in Fig 6 . The total data set consist of 215612 samples. The data of wind turbine for year 2017 is selected as training set while data for year 2018 for the month January is used for the prediction of output power of one of the wind turbine. The summary of training and test set is show in Table II.

TABLE II SUMMARY OF MODEL

	Training set	***Test set***
Year of record	2017	2018
Number of data points	208917	6695

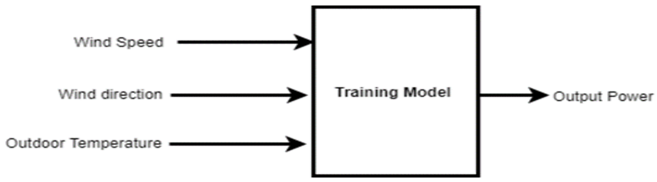

Fig. 6. Inputs and output of the proposed model

IV. RESULTS AND DISCUSSION

The separation of test and train set is done in continuous ways in order to avoid over-fitting risk. Random Forest Regressor is used to predict the output power of wind turbine. This model have lot of advantages over other existing models in literature including low over-fitting tendency, simple and fast to train. The capacity factor is a design parameter which is used to evaluate the performance of wind farm. The capacity factor F is the ratio between real and rated values which in our case is given as

$$F = \frac{P(real)}{P\ (rated\ )}$$

The capacity factor for the wind turbines used for this study is between 20-40 %. The monthly average capacity factor for the wind turbine is show in Fig. 7.

The Mean Squared Error (MSE) and Mean Absolute Error (MAE) are calculated to evaluate the performance of model. The results obtain in this research show the MAE for our model is 30 which is satisfactory. The summary of model performance is show in Table III.

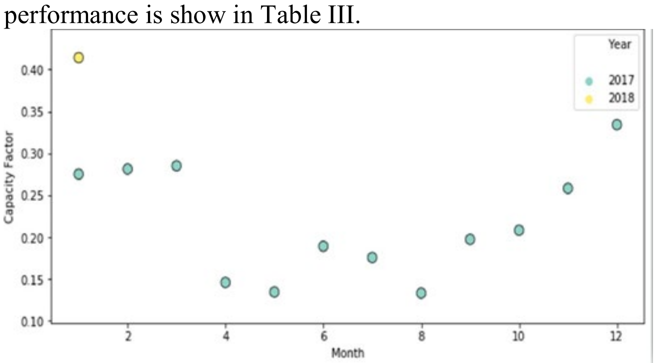

Fig. 7. Monthly average capacity factor of wind turbine

TABLE III MODEL PERFORMANCE EVALUATION

Performance metrics	***Value***
MAE	30.012
MSE	8712.0145

The nominal power of wind turbine is 2050kW as from Table I. MAE gave us the idea of average error in our model if we compare it with the average capacity factor of wind turbine through the year .To calculate the percentage error we use following equation.

$$Percentage\ error = \frac{MAE}{F * P(rated)}$$

By using above equation we calculate the percentage error in predicted power for our model. The capacity factor analysis for our model is below

For 20% capacity factor

$$\frac{30}{0.2 * 2050} = 7.3\%$$

For 40% capacity factor

$$\frac{30}{0.4 * 2050} = 3.6\%$$

The capacity factor analysis show that the results are solid for wind turbine power prediction because wind is very unpredictable energy source and based of the external parameter is mostly uncertain. The comparison between predicted and actual value for one day in January 2018 is show in Fig 8.

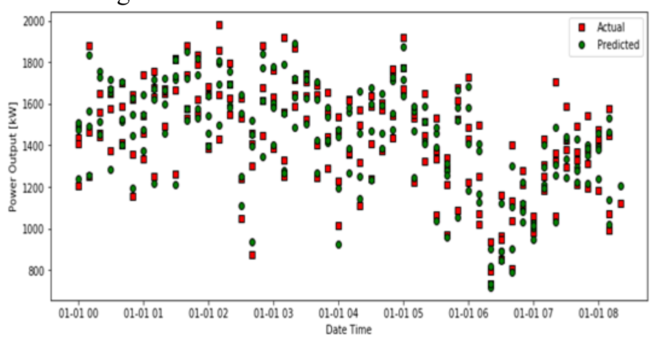

Fig. 8. Comparison between actual and predicted output power

The results are highly correlated with each other.

V. CONCLUSIONS AND FUTURE WORK

The result of this paper is a powerful method for precise and efficient wind power prediction using machine learning. The SCADA data is collected from 2 MW wind turbine in the wind farm in France from 1st January 2017 to 29th January 2018 were used to forecast the output power of the wind turbines. The data is pre-processed before the analysis for better performance. Wind turbine output power for the year 2018 is forecast using the 2017 measurements. The use of appropriate features helps to improve the prediction. We used the absolute wind direction, wind speed and outdoor temperature to predict output power . We estimate the wind turbine performance by the capacity factor for real power output and annual effected power output. Random forest regressor machine learning model to predict the output power. The estimated mean absolute error for our proposed model for the capacity factor 0.4 and 0.2 is 3.6 % and 7.3 % respectively. These results are promising in uncertain and unpredictable wind forecast. The method offers an efficient and comfortable balancing of a preferably low prediction error.

In future the problem of missing data can be solved using different imputation methods. The prediction model can be further extended to predict other parameter like fault in the wind turbine.

REFERENCES

[1] The European Wind Energy Association. Wind in power – 2015 European statistics. https: // windeurope. org/ wp-content/ uploads/ files/ about-wind/ statistics/ EWEA-Annual-Statistics-2015. pdf . 2016.

[2] Council, Global Wind Energy. "Global wind statistics 2019." Global Wind Report (2020).

[3] Iessa, Athraa and Wahab, Noor Izzri Abdul and Mariun, Norman and Hizam, Hashim, "Method of estimating the maximum penetration level of wind power using transient frequency deviation index based on COI frequency," 016 IEEE International Conference on Power and Energy (PECon), pp. 274–279, 2016.

[4] CNBC, "China and US lead way with wind power installations, says global energy report.," 05 May, 2017 13 Feb 2017

[5] Mitchell, T. M. :Machine Learning. Singapore: McGraw-Hill Companies, Inc" 1997.

[6] Carl Elkin, Sims Witherspoon . "Machine learning can boost the value of wind energy". "deep mind field study" Feb 2019

[7] Z. Guo, W. Zhao, H. Lu, and J. Wang, Multi-step forecasting for wind speed using a modified EMDbased artificial neural network model. 2012, pp. 241–249.

[8] S. Baran, "Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components," Computational Statistics Data Analysis, vol. 75, pp. 227-238, 2014/07/01/ 2014.

[9] B. Kanna and S. N. Singh, "Long term wind power forecast using adaptive wavelet neural network," in 2016 IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering (UPCON), 2016, pp. 671-676.

[10] L. Fugon, J. Juban, G. Kariniotakis, et al. Data mining for wind power forecasting. In Proceedings European Wind Energy Conference Exhibition EWEC 2008, 200