Improving site-dependent power curve prediction accuracy using regression trees

S Barber and H Nordborg
University of Applied Sciences Rapperswil, Oberseestrasse 10, 8640 Rapperswil, Switzerland
E-mail: sarah.barber@hsr.ch

Abstract. The accurate prediction of the power production of a wind turbine at a particular site is important in both the planning and operation phases; however, the standard power curve binning method is not specific to the atmospheric conditions at the site. In this work, the application of machine learning for improving the accuracy of site-specific power predictions by taking into account turbulence intensity and shear is investigated, by creating a set of 8,000 ten-minute long aero-servo-elastic simulations of the NREL 5MW reference wind turbine at a random combination of hub-height wind speeds, turbulence intensities and shear factors using cloud computing with the software ASHES. A regression tree with maximum depth of eight and optimised with Adaptive Boosting is trained using a random selection of half of the data. For a set of 50 random test cases, the Root Mean Square Error of the predicted power compared to the simulated power is found to be three times smaller than for the standard power curve method of binning. OEMs could use this method to train a model for the dependency of power production on wind speed, turbulence intensity and shear in the certification phase of a given wind turbine type, which could then be applied by the wind farm planner or operator at any given site for which measurements of the atmospheric conditions are available. Similar success is observed for out-of-plane tip deflections. This result is highly dependent on the quality and distribution of the input data, and therefore this work is being used as a basis for the analysis of real measurement data, as well as for comparisons to other methods such as Artificial Neural Networks.

Keywords: Power curve, machine learning, atmospheric conditions, regression trees, tip deflection

1. Introduction

1.1. Background

The accurate prediction of the expected power production of a wind turbine at a particular site is important in both the wind farm planning and operation phases. For wind farm planning, the power curve provided by the manufacturer (OEM) is used in combination with the expected long-term wind speed frequency distribution to calculate
the expected Annual Energy Production (AEP), and therefore the expected Levelized Cost Of Electricity (LCOE) and Return On Investment (ROI). For wind farm operation, the OEM’s power curve and historical measurements of power and wind speed are used to create reference power curves for each wind turbine in a given wind farm, which are then used for various applications including examining anomalous behaviour of single wind turbines via simple comparisons, calculating compensation for forced curtailment as well as making short-term power predictions for optimising the revenues by buying and selling electricity at optimal times.

However, the power curve provided by the OEM is not specific to the atmospheric conditions at the site, and therefore only applies if the conditions are the same as those at the test site. Atmospheric conditions such as turbulence intensity and shear differing from those at the test site may affect the power production significantly. Previous studies have shown that atmospheric conditions can affect the turbine output by 10% or more at the same wind speed, e.g. [1, 2, 3, 4, 5]. In the planning phase, this can lead to a significant error in AEP and therefore in revenues and ROI. For example, a loss of just 1% in rated power of a 1.5 MW wind turbine can lead to a lost income of $1,000 per year per turbine [6]. In the operation phase, this can lead to an error in comparisons between measured and expected power curves, leading to operators potentially coming to false conclusions when attempting to optimise operation or revenues.

It is clear that it is not possible for OEMs to provide reference power curves for every site in the world, because the power production of a wind turbine has a complex dependency on the atmospheric conditions, and it is not trivial to establish these relationships. The standard method of calculating the power curve based on measurement data at a specific site is described in IEC 61400-12-1 [7], and involves the method of binning. This consists of dividing the measured data into separate wind speed bins based on wind speed measurements at hub-height, and calculating the mean of all the measured powers in each bin. This method has a limited accuracy, especially for the prediction of single ten-minute average power values, as can be deduced by comparing the bin-averaged power to a given ten-minute value of any measured power curve.

1.2. Machine learning

Machine learning has the potential to improve the accuracy of power predictions at a given site, as well as to allow the prediction of power of a given wind turbine model at other sites at which different atmospheric conditions occur, by taking the effects of turbulence intensity and shear into account. Provided sufficient measurement data is available for training a model, transfer functions can be developed between several input conditions, such as the wind speed, turbulence intensity and shear factor, and the power output.

Little work has been published on the application of machine learning methods to power prediction. The most in-depth publication involves using simulation data of a 1.5 MW wind turbine to train and test regression trees for a range of atmospheric conditions.
conditions [6], showing that the method can be three times more accurate than the power curve method. More work has been done on using machine learning to predict site-specific loads, including multivariate regression models [8], expansions using orthogonal polynomial basis [9], polynomial chaos expansion (PCE) [10] and Artificial Neural Networks [11]. This last study involved creating training data using simulations of the DTU 10 MW reference wind turbine using the HAWC2 tool, and showed that a feedforward neural network with two hidden layers and 11 neurons per layer, trained with the Levenberg Marquardt back propagation algorithm, is able to estimate blade root flapwise damage-equivalent loads more accurately and faster than a PCE trained on the same data set.

Due to the comparable ease of tuning and modelling with categorical variables when dealing with multi-dimensional complex data typical for wind turbines, regression trees have been chosen for further investigated in the present work. An optimised regression tree is developed and applied to 8,000 simulation results of the NREL 5 MW reference wind turbine using the commercial aero-servo-hydro-elastic simulation software ASHES. This ensures a high quality and full range of wind speed, turbulence intensity and shear input data, which is required to train and test the model. In this paper, the creation of the training and test data is described in section 2, the development and application of the regression tree model is described in section 3, and a further investigation of its application to rotor blade tip deflection is shown in section 4.

2. Training and test data

In this work, a data set for training and testing the regression tree models was generated for the NREL 5 MW onshore reference wind turbine using the simulation programme ASHES. ASHES is a commercial aero-hydro-servo-elastic wind turbine design tool developed by the company Simis in Norway that uses the Blade Element Momentum Method (BEM) to calculate the loads on the blades, and the Finite Element Method (FEM) in conjunction with a co-rotational formulation of beam elements to determine the dynamic response of the structure. The simulation data was generated by firstly randomly generating 8,000 separate combinations of wind speeds in the range 3-25 m/s, turbulence intensities (TI) in the range 5-45% and shear factors (alpha) in the range -0.5 to 0.5, and then creating 20 zero turbulence cases with an alpha of 0.2 as a reference. This was done by creating a batch file containing these 8,020 input conditions, and carrying out the simulations using the cloud computing option. For each simulation, ASHES used the TurbSim software from NREL to create ten-minute long turbulent fields in a 20 by 20 grid covering the rotor area based on the IEC Kaimal model [12]. The simulations took just under 24 hours on 18 nodes in the Cloud. It is clear that these conditions do not fully represent real conditions entirely correctly; however, the approach has been used for a first test of the method. Future work would have to look at more realistic combinations of turbulent structures and shear profiles.

Figure 1 shows the resulting power curve - the ten-minute averages for each of
the 8,000 turbulent simulations in part (a) and in a box plot in part (b), which shows the median power production as a black horizontal line, the interquartile range as the box, whiskers extended to the 5th and 95th percentiles, and individual outliers marked as dots. The black line represents the zero-turbulence case. The largest variation in power can be seen just below rated speed, and could result in a large variation between predicted and observed power output. Additionally, it can be seen that the effect of non-zero TI is to increase the power production at wind speeds below rated, and to decrease the power production at or slightly above rated. This agrees with previous work, e.g. [3], supporting the expectation that power predictions at sites where the atmospheric conditions are different to the test site at which the OEM’s power curve was measured, such as in mountainous regions, may be inaccurate.

Figure 1. Power curves of the 8,000 simulations: (a) Ten-minute averages; (b) Box plot with zero-TI case marked with black line.

Figure 2 shows the dependency of the power production on TI and alpha for three different wind speed bins of 7-8 m/s (below rated wind speed), 11-12 m/s (at rated wind speed) and 17-18 m/s (above rated wind speed) from the simulation data. The individual coloured squares represent the bin-averaged normalised power production in each of 20 equally-spaced TI and 20 equally-spaced alpha bins. The power has been normalised with the zero turbulence power production in the relevant wind speed bin. It can be seen that there are many gaps in the data - this is due to the limited number of ten-minute simulation periods in each wind speed bin. Despite these gaps, it is possible to see that the dependency of the power on TI and alpha is significantly higher below rated wind speed that at or above rated wind speed. This is because the wind turbine is not being regulated to limit the power below rated wind speed. At wind speeds below rated, the power increases with increasing TI. This is to be expected, due to the cubic relationship between the turbulence fluctuations of the wind and the power. Additionally, the power appears to increase with increasing alpha, both positively and negatively. This is because increasing the shear leads to a larger difference between the wind speed at the bottom of the rotor and the top of the rotor, increasing the power averaged over the entire rotor. As well as having a much smaller dependency at wind speeds at or above rated, the power production is reduced with increasing TI and alpha, as opposed to the increase observed below rated wind speed. This is because the wind
turbine cannot convert the extra power from short-term wind speed increases above rated power, but does lose power when the wind speed reduces. This observation agrees with previous work, e.g. [5] and [6].

Figure 2. Power for TI and alpha for three wind speed bins, normalised with zero turbulence power in each bin.

3. Regression trees for site-specific power curve prediction

Decision trees are predictive models that use a set of binary rules to calculate a target value. Each individual tree is a fairly simple model that has branches, nodes and leaves. A decision tree reaches an estimate by asking a series of questions to the data, each question narrowing our possible values until the model get confident enough to make a single prediction. The order of the questions as well as their content is determined by the model. There are two types of decision trees - classification trees and regression trees. Classification trees are based on categorical targets, whereas regression trees are applied when the targets are real numbers. Regression trees are therefore appropriate for this work.

Regression trees work by firstly defining a root node representing the entire data, and then splitting this node into sub-nodes, which contain subsets of the original data. These sub-nodes are then further split into decision nodes. Node splitting takes place by grouping the input variables into homogeneous regions, chosen automatically by the regression tree itself by assessing the Mean Square Error of the splitting possibilities. The number of splits can be limited by setting the maximum depth of the tree. As decision trees have no guarantee of returning the globally optimal result, they have to be optimised by training multiple trees. Two common methods for doing this are the Random Forest Algorithm (e.g. [13]) and the Adaptive Boosting Method [14]. The Random Forest Algorithm involves averaging the predictions from a multiple number of trees created with random samples of the training data. Adaptive Boosting involves using the prediction error to adjust input weightings, and converts many so-called weak learners into a single strong learner. Weak learners are poor when it comes to learning the relationships between inputs and target, and are identified by making some distributions and forming small decision trees from them.

For this study, regression trees were built by randomly choosing half of
the data (4,000 simulations) for training and half for testing and applying the `DecisionTreeRegressor` from Python's `sklearn.tree`. The maximum depth of the tree was increased between three and 20 in order to find the best accuracy for the lowest computational power. No significant improvement was found above a maximum depth of eight; therefore this value was chosen. For optimising the method, the Random Forest Algorithm was compared to the Adaptive Boosting Method by comparing the Root Mean Square Error (RMSE) of the predicted power curve to the actual power curve of the test data. The Adaptive Boosting Method was found to offer a 20% improvement in accuracy compared to the Random Forest Algorithm and was therefore chosen for this study and applied using the Python `AdaBoostRegressor`. An example of one section of the optimised regression tree is shown in Figure 4.

Using this optimised regression tree, the power was then predicted for every combination of the 20 equally-spaced TI and 20 equally-spaced alpha bins for the three wind speed bins examined above (7-8 m/s, 11-12 m/s and 17-18 m/s). This allowed the data gaps visible in Figure 2 to be filled in, as shown in Figure 3. The power has been normalised with the zero turbulence power production in the relevant wind speed bin. Compared to Figure 2, this shows much clearer gradients in power output as a function of TI and alpha. It can now be seen that the increase in power with TI below rated wind speed is less significant for negative values of alpha than for positive values. On the other hand, the power is hardly dependent on alpha for positive shear, but becomes highly dependent upon it for negative shear. This is due to the cubic relationship between the turbulence fluctuations of the wind and the power, causing changes in TI to dominate much more strongly at positive shear than at negative shear. It can also now be seen that there is a significant difference in power dependency at rated wind speed and above rated wind speed. At rated wind speed, a small dependency of power on alpha for negative shear can still be seen, whereas above rated wind speed, this dependency has completely disappeared, for the reasons described in the previous section.

In a real application, measured wind speed, shear, turbulence intensity and power data would be used for a limited period of time in order to build the regression tree, which would then be applied to further measurements of wind speed, shear and turbulence intensity data in order to predict the power for these previously unseen conditions.
This process was simulated for 50 test cases, created by randomly choosing half of the simulation data 50 times. The quality of the predictions was assessed by calculating the RMSE of the predicted power compared to the actual power for each test case. These RMSE results were then compared to the RMSE for the standard power curve method of binning for each case. The results are shown in Figure 5. It can be seen that the regression tree method is much more accurate than the standard power curve binning method, with an average RMSE of 62 kW, compared to 182 kW from the power curve method. The difference between these two methods is visualised in Figure 6, which compares the average power predictions for one test case for (a) the regression tree method and (b) the power curve method. It can be seen that the power curve method is less accurate because it only predicts one constant power value per wind speed bin.

Figure 4. Example section of optimised regression tree.

Figure 5. Root Mean Square Error (RMSE) comparison.

Figure 6. Power predictions compared to actual power for one test case: (a) Regression tree method; (b) Power curve method.

It can therefore be concluded that the regression tree method that takes shear and turbulence intensity into account works three times more accurately than the standard power curve method of binning for this application. OEMs could use this method to train a model for the dependency of power production on wind speed, turbulence intensity and shear in the certification phase of a given wind turbine type, which could then be applied by the wind farm planner or operator at any given site for which measurements of the atmospheric conditions are available.
As a limited amount of data may be available, the effect of data availability on the results has been investigated, by altering the proportion of training data from the original value of 50% (4,000 ten-minute periods, corresponding to about 28 days of measurements). The effect of reducing the number of training data points on the average RSME is shown in Figure 7. It can be seen that increasing the amount of training data reduces the RMSE, as expected. However, even reducing the amount of training data to 400 ten-minute points (2.8 days) does not increase the RMSE to above the power curve method for 4,000 data points.

It is clear that this result is highly dependent on the quality and distribution of the input data. Furthermore, the simulated wind conditions do not fully represent real conditions entirely correctly. The next step of this work is to test the method on real data sets, which contain more realistic combinations of turbulent structures and shear profiles and are likely not to be as equally spaced or well-distributed as the simulation data, as well as comparisons to other methods such as Artificial Neural Networks. This work is on-going.

4. Regression trees for site-specific tip deflection prediction

As the regression tree method worked successfully for power curve predictions, its application to other physical quantities has been investigated. It is relatively simple to extend this method to other quantities using the existing simulation results. In this study, the out-of-plane tip deflection was chosen as a suitable quantity, although it could be equally well applied to load measurements. Tip deflection was chosen because it is becoming more and more important to measure this quantity in the field in order to validate design models, as wind turbine blades are getting longer and more flexible. The simulated out-of-plane tip deflection is shown in Figure 8, showing the expected increase up to a maximum near rated speed, followed by a decrease above rated speed. This decrease occurs because of the pitch regulation of the wind turbine reducing the flapwise blade loading.

![Figure 7](image1.png)  
**Figure 7.** Effect of changing the training data size on average Root Mean Square Error (RMSE).

![Figure 8](image2.png)  
**Figure 8.** Ten-minute averages of tip deflection for 8,000 simulations.
The same method for optimising the regression tree was then applied to the tip deflection data. The optimised regression tree was used to predict the out-of-plane tip deflection for every combination of the 20 equally-spaced TI and 20 equally-spaced alpha bins for the three wind speed bins examined above (7-8 m/s, 11-12 m/s and 17-18 m/s), as shown in Figure 9. At wind speeds below rated, the tip deflection increases for increasing negative alpha, and is hardly dependent on TI. The largest changes are seen for varying TI at rated wind speed. Above rated wind speed, increasing alpha both positively and negatively decreases the tip deflection, and the tip deflection is only dependent on TI around zero shear.

![Filled Contours Plot](image)

**Figure 9.** Tip deflection dependency on TI and alpha for three wind speed bins, average prediction for 100 regression trees. Colour scale shows deflection in m.

As for the power predictions described above, tip deflection predictions for previously unseen atmospheric conditions were then made with the optimised regression tree for 50 test cases, created by randomly choosing half of the simulation data 50 times. Again, the regression tree method showed a similar improvement over the standard binning method, with an average RMSE of 0.08 m, compared to 0.21 m with the binning method.

This method could be applied for any other physical quantity, such as the blade loads. As for the power predictions, OEMs could use this method to train a model for the dependency of tip deflection on wind speed, turbulence intensity and shear in the certification phase of a given wind turbine type, which could then be applied by the wind farm planner or operator at any given site for which measurements of the atmospheric conditions are available. This would probably be most applicable to the loads, as these are measured by the OEM as part of the certification process anyway. Additionally, it could be useful in order to extend short-term measurement campaigns to a range of unseen atmospheric conditions, saving significant amounts of time and costs.

5. Conclusions

In this work, the application of regression trees for improving the accuracy of site-specific power predictions by taking into account turbulence intensity and shear has been investigated using the Python *DecisionTreeRegressor*. This was done by creating a set of training and test data by carrying out 8,000 ten-minute long aero-servo-elastic
simulations of the NREL 5 MW reference wind turbine at a random combination of hub-height wind speeds from 3 m/s to 25 m/s, turbulence intensities from 5% to 45%, and shear factors from -0.5 to 0.5 using cloud computing with the software ASHES. Next, half of the wind and power data (4,000 simulations) was randomly chosen for training and the other half for testing. The optimal maximum tree depth was found to be eight, and the best optimisation method was found to be the Adaptive Boosting Method. Power predictions using the optimised regression tree allowed a detailed analysis of the effect of turbulence intensity and shear for different wind speed regimes to be carried out. For a set of 50 random test cases, the Root Mean Square Error of the predicted power compared to the simulated power was found to be three times smaller than for the standard power curve method of binning. OEMs could use this method to train a model for the dependency of power production on wind speed, turbulence intensity and shear in the certification phase of a given wind turbine type, which could then be applied by the wind farm planner or operator at any given site for which measurements of the atmospheric conditions are available. Finally, the same method was repeated for simulated out-of-plane tip deflections in order to demonstrate the ability of the technique to predict other physical quantities. A similar improvement in accuracy over the standard binning method was observed. This could be useful for extending load measurement campaigns to unseen conditions, saving significant amounts of time and costs. This result is highly dependent on the quality and distribution of the input data, and therefore this work is being used as a basis for the analysis of real measurement data, as well as for comparisons to other methods such as Artificial Neural Networks.

6. References

[1] I Antoniou, S M Pedersen, and P B Enevoldsen. Wind shear and uncertainties in power curve measurement and wind resources. Wind Engineering, (33):449–468, 2009.
[2] E Hedevang. Wind turbine power curves incorporating turbulence intensity. Wind Energy, (17(2)):173–195, 2014.
[3] B J Vanderwende and J K Lundquist. The modification of wind turbine performance by statistically distinct atmospheric regimes. Environmental Research Letters, (7):034035, 2012.
[4] R Wagner, M Courtney, J Gottschall, and P Lindelow-Marsden. Accounting for the speed shear in wind turbine power performance measurement. Wind Energy, (14):993–1004, 2011.
[5] S Wharton and J K Lundquist. Atmospheric stability affects wind turbine power collection. Environmental Research Letters, (7):014005, 2012.
[6] A. Clifton, L. Kilcher, J. K. Lundquist, and P. Fleming. Using machine learning to predict wind turbine power output. Environmental Research Letters, 8(2):024009, 2013.
[7] IEC International Electrotechnical Commission. Wind energy generation systems - part 12-1: Power performance measurements of electricity producing wind turbines, 03.03.2017.
[8] H S Toft, L Svenningsen, W Moser, J D Sorensen, and M L Thogersen. Assessment of wind turbine structural integrity using response surface methodology. Eng. Struct., (106):471–483, 2016.
[9] JPMurcia, PER ´ethor´e, NDimitrov, ANatarajan, JDSorensen, PGraf, and TKim. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. Renewable Energy, (119):910–922, 2018.
[10] N Dimitrov, M Kelly, A Vignaroli, and J Berg. From wind to loads: wind turbine site-specic load
estimation using databases with high-fidelity load simulations. *Wind Energy Science*, (3):767–790, 2018.

[11] Laura Schröder, Nikolay Krasimirov Dimitrov, David Robert Verelst, and John Aasted Sørensen. Wind turbine site-specific load estimation using artificial neural networks calibrated by means of high-fidelity load simulations. *Journal of Physics: Conference Series*, 1037:062027, 2018.

[12] IEC International Electrotechnical Commission. Wind energy generation systems - part 1: Design requirements, 08.02.2019.

[13] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[14] Y Freund and R E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, (55):119–139, 1997.