Innovative methods for wind turbine power curve upgrade assessment

Ludovico Terzi\textsuperscript{1}, Andrea Lombardi\textsuperscript{1}, Francesco Castellani\textsuperscript{2}, Davide Astolfi\textsuperscript{2}

\textsuperscript{1} Renvico s.r.l., Via San Gregorio 34, 20124 Milano, Italy. www.renvicoenergy.com
\textsuperscript{2} University of Perugia - Department of Engineering, Via G. Duranti 93 - 06125 Perugia, Italy
E-mail: ludovico.terzi@renvico.it

Abstract. Wind turbine power curve upgrades have recently been attracting considerable investment in the operation of wind farms and noticeable attention in the wind energy literature. Due to the non-stationary conditions to which wind turbines are subjected, the most consistent strategy for quantifying the production improvement from the installation of an upgrade is comparing, after at least some months of operation, the post-upgrade production against a model of the pre-upgrade production under the same conditions. Formulating adequate models for the power of the upgraded wind turbines is in general non-trivial and it can be difficult to achieve the precision that wind turbine practitioners typically require for the production assessment.

In the present work, a multivariate linear method for selecting the most appropriate input for modeling a given output is presented and applied to a test case. The test case is a multi-megawatt wind turbine owned by Renvico, on whose blades vortex generators and passive flow control devices have been installed. Applying the proposed method, it is possible to compute with precision the production improvement in the first five months of post-upgrade operation (purely aerodynamic upgrade) and in the subsequent three months (aerodynamic and control upgrade). It is therefore possible to appreciate the different contributions to the production enhancement from the aerodynamic and control improvement. A non-upgraded wind turbine from the same wind farm is also studied and the precision of the results inspires the use of the proposed method for performance control and monitoring in general.

1. Introduction

Multi-megawatt wind turbine is nowadays a mature technology: this implies, on one side, that public subsides are gradually declining and, on the other side, that there is a continuous improvement in the efficiency of wind kinetic energy conversion and in the control and monitoring of operating wind farms.

Wind turbine power curve upgrades have therefore recently attracted a considerable attention from wind turbine practitioners. One subtle point about wind turbine power curve upgrades is that quantifying realistically their impact on the power production is tricky, because their efficiency might depend considerably on the features of the wind flow at the microscale level \cite{1}. It is evident that it doesn’t make much sense to compare the production of a wind turbine pre and post upgrade, because wind turbines operate under non-stationary conditions. The most appropriate approach is comparing the post-upgrade production, after at least some months of upgraded operation, against a simulation of how much the wind turbine would have produced if the upgrade hadn’t been installed.
Recently, this topic has been addressed in a certain number of studies in the wind energy literature. There are basically four studies. The first one is [2]: this manuscript contains an interesting discussion motivating why the binning method for the study of the power curve is insufficient for facing satisfactorily wind turbine power curve upgrades. Basically, as supported also in [3], precision modeling of the power of a wind turbine calls for multivariate dependencies. Therefore, on these grounds, in [2] a novel method for modeling wind turbine power output and quantifying power upgrades is proposed. It is a generalization of the conventional kernel regression method (which is, in essence, a weighted average of all the data points with the weighting coefficient associated to each input measurement decaying to zero when the euclidean distance between the evaluation point and the input measurement increases). Two issues are faced: dimensionality and bias. The first problem is approached by allowing hybrid kernels (additive-multiplicative). The second problem is tackled by selecting a calibration data set for the model that has similar features with respect to the target data set: this is done defining a distance conceptually similar to the Mahalanobis distance. In [2], this method is applied to the study of two test case: a real one (vortex generators installation) and an artificial one (pitch angle optimization: the data are created fictitiously basing on the logic of pitch angle control).

In [1], the problem of quantifying the production improvement from vortex generator installation is studied using SCADA data and time-resolved data having respectively ten minutes and seconds of sampling time. The point is that, using the kernel plus method and the SCADA data, the computation of the energy improvement results to be affected by considerable uncertainty. To enlarge the data set and therefore improve the statistical significance, the time-resolved data are employed: the authors of [1] show that the power-power approach is sufficient to provide reliable results, when this kind of data is used.

The power-power approach is an intuitive method and it is based on quantifying how the power difference between the target upgraded wind turbine and a reference non-upgraded wind turbine changes after the upgrade of the wind turbine of interest. Possibly, as in [1], data can be pre-treated by removing measurements characterized by wind turbine operation under the wake of nearby ones. Doing this, operation conditions are selected that resemble as much as possible the ideal ones in real environment. The main result of [1] is that vortex generators can have an efficiency that considerably depends on the flow conditions at the microscale level.

This method has been generalized in [4]: the test case (the same as in the present work) is a multi-megawatt wind turbine sited in a very complex terrain. Vortex generators and passive flow control devices have been installed on this test wind turbine. The idea of [4] is modeling the power of the upgraded wind turbine as a function of the powers of a certain number of nearby wind turbines (possibly and realistically more than one, while the power-power approach employs only one reference wind turbine). Since the relation between the input and the output is unknown, a data-driven approach has been employed: a feed-forward Artificial Neural Network is the selected model. In [4], it is shown that a bootstrap technique on the training and validation data sets is useful for improving the precision of the results. Other wind turbine power curve upgrades (including a real test case of pitch angle optimization) are studied, according to the same philosophy, in [5].

This study moves from [4] and [5]. It is a collaboration between the University of Perugia and the Renvico company, owning and managing 335 MW of wind turbines in Italy and France (www.renvicoenergy.com). At least three considerable steps forward with respect to [4] are addressed.

The first step forward regards the fact that the model of [4], despite more general than the power-power approach, still contained a certain degree of arbitrariness in the selection of the input for modeling the power of the upgraded wind turbine. One of the objectives of this work is overcoming this limitation: an appropriate multivariate linear model has therefore been identified by means of stepwise regression. Since the stability and the quality of the results were
required features, particular care has been devoted to the validation of the model. Therefore, the outcome of this work is not only the computation of the production improvement, but also the procedure for selecting the appropriate inputs and validating the model. The method is generalizable to the study of whatever kind of wind turbine upgrades and the precision obtained in this very complex test case is impressive. As regards the test case of interest, it is important to notice that three input variables (from three nearby wind turbines) are selected for modeling the power of the upgraded wind turbine: this implies that using only one, as in the power-power approach, might not be robust enough.

The second improvement with respect to [4] regards the fact that the upgrade has been installed in two steps: first, the purely aerodynamic part has been installed and the wind turbine has been operating for five months. Subsequently, the control of the revolutions per minute has been optimized and the wind turbine has been operating with the entire upgrade for three months up to when this study has been performed. Therefore, with this work it is possible to separate the effects of the two sub-upgrades and quantify them.

The third important point of this work is that the adopted method provides sufficient precision in the computation of the energy production in order to employ it for performance monitoring in general. In this work, this is supported by modeling the power output of a non-upgraded wind turbine and assessing the precision of the model.

Summarizing, the structure of the manuscript is therefore the following: first, the test case and the data sets are briefly described in Section 2; subsequently the model is formulated in Section 3; the results are summarized in Section 4; conclusions and further directions are sketched in Section 5.

2. The wind farm and the data set
The layout of the wind farm is reported in Figure 1. From the contour lines, it is possible to appreciate the complexity of the terrain. T7 (in red in Figure 1) is the wind turbine that has been upgraded with vortex generators and passive flow control devices. The wind turbines are multi-megawatt and the hub height is 80 meters above ground level.
The vortex generators installed on the blades are injection moulded thermoplastic components made out of a compound polymeric material and they are mounted both inboard at the root section of the blade as well as outboard at the tip of the blade. The adopted passive flow control devices are flaps that are glued onto the trailing edge of blade root and of the blade tip. All these devices produce an increase of the lift on the basis of different principles: the vortex generators delay the airflow separating from the blade surface, the blade tip flow control devices extend the blade chord. For detailed discussions on the physics of the different types of flow control devices and on their use as smart rotor control for loads reduction and therefore wind turbine life-cycle increase, see for example [6, 7, 8, 9].

In Figure 2, a photo is reported of a blade of T7 equipped with the flow control devices. It has been reported for a qualitative comprehension of the kind of devices. For a detailed discussion of the technology, see [10].

The data sets at disposal are the following:

- $D_{bef}$ goes from 01/01/2016 to 01/07/2017. It is a period during which the standard blade configuration was adopted.
• $D_{\text{ft}}^1$ goes from 01/09/2017 to 01/03/2018. It is a period during which T7 has been operating with the improved blade configuration.
• $D_{\text{ft}}^2$ goes from 01/03/2018 to 01/06/2018. It is a period during which T7 has been operating with the improved blade configuration and the optimized revolutions per minute control.

3. The method
The proposed approach is based on linear regression. The objective is to estimate the power output of T7, denoted as $y$ in the following. This has to be estimated based on the data from the other 16 wind turbines in the farm because, as explained in [4], the nacelle wind speed measurements from T7 are biased (in [4], it is hypothesized that the nacelle transfer function hasn’t been updated after the upgrade). In particular, for each wind turbine the data at disposal are:
• the nacelle wind speed,
• the power output,
• the individual blade pitch angles,
• the rotor revolutions per minute,
• the high speed rotor temperature.

The conditions for data filtering are that all the wind turbines are producing output and that the power of T7 is below rated: this second condition is requested because the power doesn’t upgrade when it is already rated. The filter on the request of power production has been based on the appropriate counter of grid production, available in the SCADA data set.

The procedure for selecting the input to the model is based on the stepwise regression algorithm [11], which is a data-driven iterative algorithm for automatically selecting the most appropriate inputs for a multilinear regression task.

Starting from an initial model, at each step a potential regression term is added to or removed from the model. This is done by calculating the $p-value$ of an $F$-statistics and testing the performance of the model with and without that potential term. Only the regressors whose probability of not having coefficients in the model is less then a specified threshold on the $p-value$ are included and maintained. The threshold on the $p-value$ is denoted as $p_{\text{remove}}$.

Due to its crucial role in the input selection process, the value for the $p_{\text{remove}}$ parameter must be chosen carefully. To this end, an extensive study has been carried one on the best value for $p_{\text{remove}}$. The tested values are $10^{-\gamma}$ with $\gamma = 1, \ldots, 15$. For each value of $p_{\text{remove}}$, the input selection and estimation pipeline is subjected $J$ times to $K$-fold cross-validation [12]. In particular, the data set $D_{\text{bef}}$ is divided $J$ times randomly in two subsets: $(K-1)/K$ of the data are used for training and the remaining $1/K$ are used for validation. $K$ has been selected to be 10 for this study, because the objective is having a robust model and therefore short validation folds are selected. $J$ has been set to 300 to increase the statistical significance of the study.

Given a value $\gamma$ for $p_{\text{remove}}$, for each of the $j = 1, \ldots, J$ runs of the cross-validation, the most significant inputs are selected and the estimated power of T7 $\hat{y}_{\text{valid},j,\gamma}$ is modeled via linear regression. The performance of the linear model for the given value $\gamma$ for $p_{\text{remove}}$ is estimated in terms of the mean absolute error, averaged over the $j = 1, \ldots, J$ folds:

$$\bar{\delta}_{\gamma} = \frac{1}{J} \sum_{j=1}^{J} \frac{|\hat{y}_{\text{valid},j,\gamma} - y_{\text{valid},j,\gamma}|}{J}$$

where $y_{\text{valid},j,\gamma}$ is the real power of T7 for the validation subset of that run.

From the feature selection process, it arises that $\bar{\delta}_{\gamma}$ is of the same order of magnitude ($\approx 100$ kW) for each value of $p_{\text{remove}}$; the improvement in choosing $p_{\text{remove}} = 10^{-1}$ instead
of $p_{\text{remove}} = 10^{-15}$ is averagely only 3 kW in $\bar{\delta}$, at the cost of employing many more regressors (13 against 3) and having a less robust model since, furthermore, the number of possible regressor configurations is much higher.

For whatever choice of $p_{\text{remove}} \leq 10^{-10}$, instead, one obtains that the selected inputs are the same (except for a small number of outliers) and they are:

- the power output of T6,
- the power output of T9,
- the rotor revolutions per minute of T8.

Therefore, the decision is modeling the power output of T7 as a function of the three above inputs. Besides the statistical significance of the method adopted for identifying the appropriate inputs for modeling the power of T7, support to the decision comes from the fact that it is definitely plausible. In fact, the wind turbines selected for the inputs are the nearest to T7 and one input per nearby wind turbine is selected, therefore there is no redundant information. Furthermore, it is reasonable that in such a complex terrain the power output or the rotor revolutions are more stable than the wind speed (because the wind turbine acts like a filter) and are preferable to model the power of T7.

The power of T7 is modeled via linear regression as function of the inputs selected as described above. The input measurements are normalized and organized in a matrix $x_{\text{train}}$. The pseudo-inverse of $x_{\text{train}}$ is used to compute the weight matrix of the linear regression model as:

$$W = x_{\text{train}}^{-1} \cdot y_{\text{train}}$$  \hspace{1cm} (2)

Finally, the estimated power output $\hat{y}_{\text{valid}}$ on the test data set is computed as a linear combination of the input at test time, weighted using the model weight matrix $W$:

$$\hat{y}_{\text{valid}} = x_{\text{valid}} \cdot W$$  \hspace{1cm} (3)

3.1. Modeling the power of a non-upgraded wind turbine

Consider applying the same method for a wind turbine that has not been retrofitted. Wind turbine T4 has been selected for this test, but any other non-upgraded turbine in the farm could have been selected. The inputs for modeling the power production of T4 have been selected with the same procedure described above and are:

- the power of T2,
- the power of T3,
- the power of T5,
- the rpm of T5.

4. Results

In the following, the procedure is reported for estimating the energy improvement. The data sets at disposal are employed as follows:

- $D_{\text{bef}}$ is randomly divided in two subsets: D0 (1 year of data) and D1 (6 months of data). D0 is used for training the model and constructing the weight matrix $W$, D1 is used for validating the model.
- $D_{\text{aft}}^i$ with $i = 1, 2$ (also named $D2^i$ for simplifying the notation in the following) is used in its entirety to quantify the performance improvement.
The union of $D_{\text{aft}}^1$ and $D_{\text{aft}}^2$ is named as $D_{\text{aft}}^{\text{tot}}$.

The residuals between the measurement $y$ and the simulation $\hat{y}$, for the data sets $D_1$ and $D_2$, are studied. In particular, the interest is in how the residuals vary after the upgrade with respect to before. Therefore, consider Equation 4 with $i = 1, 2$.

$$R(x_i) = y(x_i) - \hat{y}(x_i).$$

For $i = 1, 2$, one computes

$$\Delta_i = 100 \times \frac{\sum_{x \in \text{Data}_i} (y(x) - \hat{y}(x))}{\sum_{x \in \text{Data}_i} y(x)}.$$  \hspace{1cm} (5)

Since $\Delta_i$ is constructed with the relative discrepancies of power data each having the same sampling time (10 minutes), the quantity $\Delta = \Delta_2 - \Delta_1$ provides a percentage estimate also of the energy improvement.

The above procedure has been repeated several times with several random choices of $D_0$ and $D_1$ for training and validating the model. Doing this, it is possible to obtain an average estimate of the energy improvement, as well as the standard deviation (and therefore reasonable lower and upper limits for the energy improvement).

4.1. T7: aerodynamic upgrade

In Figure 3, for a sample selection of $D_0$ and $D_1$, the sets $R(x_1)$ and $R(x_2)$ are plotted after being averaged within intervals having amplitude of the 10% of the rated power.

![Figure 3](image)

**Figure 3.** The average difference between measurements and simulation (Equation 1), for data sets $D_1$ and $D_{\text{aft}}^1$ and for a sample run of the model.

The average energy improvement is computed as $\Delta = 3.6\%$. In other words, the estimate is that T7 has produced, during data set $D_{\text{aft}}^1$ and below rated power, the 3.6% more than it would without aerodynamic improvement. The standard deviation is computed as $\sigma_\Delta = 0.4\%$; therefore, reasonable upper and lower limits of the energy improvement are $\Delta_+ = 4.0\%$ and $\Delta_- = 3.2\%$. This corresponds to an improvement in the whole production during $D_{\text{aft}}^1$ given by $\Delta_E = 1.8\% \pm 0.2\%$. 


4.2. **T7: aerodynamic and control upgrade**

In Figure 4, for a sample selection of D0 and D1, the sets $R(x_1)$ and $R(x_2)$ are plotted after being averaged within intervals having amplitude of the 10% of the rated power.

![Graph](image)

**Figure 4.** The average difference between measurements and simulation (Equation 1), for data sets D1 and D2 and for a sample run of the model.

The average energy improvement is computed as $\Delta = 5.2\%$. The standard deviation is computed as $\sigma_\Delta = 0.2\%$: this corresponds to an improvement in the whole production during $D_{aft}^2$ given by $\Delta_E = 2.6\% \pm 0.1\%$.

This result allows to appreciate, comparing against the estimate in subsection 4.1, that the control optimization provides a production improvement of the order of $\Delta_E^{control} = 0.8\%$. This is compatible with the estimates in [5] for other test cases: the typical order of magnitude of energy improvement from control optimization is the percent.

Furthermore, it is interesting to study the characteristic curves before and after the control system upgrade. T7 is analyzed in comparison with a sample wind turbine from the rest of the wind farm (T1). The curves after the aerodynamic upgrade and before the control upgrade (data set $D_{aft}^1$) are reported in Figure 5 and it arises that for T1 and T7 they are basically indistinguishable.
Figure 5. The power - rotor revolutions per minute curve, for data set D1\textsubscript{aft}. Turbines T7 and T1.

The curves after the aerodynamic upgrade and the control upgrade (data set D2\textsubscript{aft}) are reported in Figure 6 and it arises that the curves for T1 and T7 are distinguishable. For a given value of the number of rotor revolutions per minute, T7 produces more power than T1.

Figure 6. The power - rotor revolutions per minute curve, for data set D2\textsubscript{aft}. Turbines T7 and T1.

4.3. T4
Figure 7 is, similarly to Figures 3 and 4, a plot of $R(x_1)$ and $R(x_2)$ on a sample model run, once the data are averaged in power intervals of 10% of the rated. The difference with respect to Figures 3 and 4 is evident: in Figure 7, the average residuals $R(x_1)$ and $R(x_2)$ as a function of the power are almost identical the ones with respect to the others.
Figure 7. The average difference between measurements and simulation (Equation 1), for turbine T4. Data sets D1 and $D_{tot}^{-2}$ and sample run of the model.

The average production variation (basing on Equation 5) is estimated to be $\Delta E = 0.05\%$: practically zero. This result supports that, even for a wind farm sited in a very complex terrain as the selected test case, it has been possible to model the production of a wind turbine with remarkable precision. This opens the perspective of employing the proposed method not only for upgrade assessment, but also for ordinary control and monitoring of the performance of operating wind turbines.

5. Conclusions and further directions
The Renvico company and the University of Perugia have been cooperating for years and recently they have devoted particularly to the study of wind turbine power curve upgrades. Actually, this topic has become relevant in the operation and management of wind farms and requires devoted techniques that are commonly in the toolbox of the academia rather than industry. For this reason, the literature about the data-based study of wind turbine power upgrades after at least some months of operation is somehow at its early stages, but it is boosting rapidly [2, 1, 4, 5].

In this work, a test case has been selected because it has been considered the most instructive between those analyzed jointly by Renvico and UniPG and the study has been performed in depth. This has allowed a methodological approach, opening the perspective of generalizing the use of the proposed techniques, and has provided interesting insight about how a wind turbine power upgrade works.

The test case has been a multi mega-watt wind turbine from a wind farm sited in a very complex terrain. On the blades, vortex generators and passive flow control devices have been installed. As discussed in [4] and [5], a qualitative assessment of the upgrade through the study of the power curve is impossible because the nacelle wind speed measurements after the upgrade are presumably biased. This limitation of the data set at disposal has been turn into a power, because it has motivated the formulation of a general method for modeling a given output (in this case, the power of the upgraded wind turbine) as a function of the most appropriate input.

An automatic feature selection algorithm based on stepwise multivariate regression has been adopted and it has allowed to select the most meaningful input variables for a linear model whose output is the power of the upgraded wind turbine. Furthermore, a bootstrap technique on the pre-upgrade data set at disposal has allowed to artificially produce the repeatability of the tests.
and this has considerably improved, with respect to the state of the art in the literature, the statistical significance and the precision of the results.

Furthermore, the installation of the upgrade has taken place in two steps: for five months, the wind turbine has run with the aerodynamic upgrade. Subsequently, for three months up to when this study has been conducted, the wind turbine has been operating with the aerodynamic upgrade and a control upgrade managing the revolutions per minute. Therefore, studying separately the two data sets, it has been possible to quantify the production improvement from the complete upgrade and from each of the two parts of the upgrade. The estimate is that the purely aerodynamic upgrade has provided a production improvement that is slightly less than the 2%; the aerodynamic and control upgrade has provided a production improvement of the order of the 2.6%. Therefore, the impact of the control upgrade can be estimated as the difference between these two and it amounts to 0.8%. This is consistent with other test cases of control system optimization studied in [5]: the typical production upgrade from control system optimization is of the order of the percent.

The proposed method has been applied also to a non-upgraded wind turbine as a consistency check. The result are remarkably precise: once one has trained the model on a data set, the residuals between model and measurements in two different data sets behave impressively similar. The obtained precision opens the perspective of employing this kind of method not only for the study of wind turbine power upgrades, but also for the control and monitoring of the performances in normal operation.

Two are the main further directions of this work: the first one is employing the proposed method for the study of other test cases. The second one is testing other kinds of method (for example, non-linear) on this test case and analyzing how the precision of the results changes.

Bibliography

[1] Hoon Hwangbo, Yu Ding, Oliver Eisele, Guido Weinzierl, Ulrich Lang, and Georgios Pechlivanoglou. Quantifying the effect of vortex generator installation on wind power production: An academia-industry case study. Renewable Energy, 113:1589–1597, 2017.
[2] Giwhyun Lee, Yu Ding, Le Xie, and Marc G Genton. A kernel plus method for quantifying wind turbine performance upgrades. Wind Energy, 18(7):1207–1219, 2015.
[3] Bruce Stephen, Stuart J Galloway, David McMillan, David C Hill, and David G Infield. A copula model of wind turbine performance. IEEE Transactions on Power Systems, 26(2):965–966, 2011.
[4] Davide Astolfi, Francesco Castellani, and Ludovico Terzi. A scada data mining method for precision assessment of performance enhancement from aerodynamic optimization of wind turbine blades. In Journal of Physics: Conference Series, volume 1037, page 032001. IOP Publishing, 2018.
[5] Davide Astolfi, Francesco Castellani, and Ludovico Terzi. Wind turbine power curve upgrades. Energies, 11(5):1300, 2018.
[6] Thanasis K Barlas and GAM Van Kuik. Review of state of the art in smart rotor control research for wind turbines. Progress in Aerospace Sciences, 46(1):1–27, 2010.
[7] Kuo-Chang Tsai, Cheng-Tang Pan, Aubryn M Cooperman, Scott J Johnson, and CP Van Dam. An innovative design of a microtab deployment mechanism for active aerodynamic load control. Energies, 8(6):5885–5897, 2015.
[8] Unai Fernández-Gáiniz, Clara Marika Veite, Pierre-Elouan Réthoré, Niels N Sørensen, and Eduard Egusquiza. Testing of self-similarity and helical symmetry in vortex generator flow simulations. Wind Energy, 19(6):1043–1052, 2016.
[9] Inigo Aramendia, Unai Fernandez-Gamiz, Jose Antonio Ramos-Hernanz, Javier Sancho, Jose Manuel Lopez-Guede, and Elaizzt Zulueta. Flow control devices for wind turbines. In Energy Harvesting and Energy Efficiency, pages 629–655. Springer, 2017.
[10] J. Laursen, P. Fuglsang, and P.B. Enevoldsen. Dinosaurs, racecars and blades: Discover the connection. Technical report, Wind Power World, 2012.
[11] PT Pope and JT Webster. The use of an F-statistic in stepwise regression procedures. Technometrics, 14(2):327–340, 1972.
[12] Payam Refaeizadeh, Lei Tang, and Huan Liu. Cross-validation. In Encyclopedia of database systems, pages 532–538. Springer, 2009.