Article

Computing the real impact of wind turbine power curve upgrades: a SCADA-based multivariate linear method and a vortex generator test case

Davide Astolfi [1,‡]*, **Francesco Castellani** [1,‡], **Mario Luca Fravolini**[1,‡], **Silvia Cascianelli**[1,‡], **Ludovico Terzi** [2,‡]

1   University of Perugia - Department of Engineering, Via G. Duranti 93 - 06125 Perugia (Italy);
     davide.astolfi@unipg.it; francesco.castellani@unipg.it; mario.fravolini@unipg.it; silvia.cascianelli@unipg.it
2   Renvico srl, Via San Gregorio 34, Milano 20124, Italy; ludovico.terzi@renvico.it
*   Correspondence: davide.astolfi@unipg.it; Tel.: +39 075 585 3709
‡   These authors contributed equally to this work.

Abstract: Wind turbine upgrades have been spreading in the recent years in the wind energy industry, with the aim of optimizing the efficiency of wind kinetic energy conversion. This kind of interventions has material and labor costs and it is therefore fundamental to estimate realistically the production improvement. Further, the retrofitting of wind turbines sited in harsh environments might exacerbate the stressing conditions to which wind turbines are subjected and consequently might affect the residue lifetime. This work deals with a case of retrofitting: the testing ground is a multi-megawatt wind turbine from a wind farm sited in a very complex terrain. The blades have been optimized by installing vortex generators and passive flow control d evices. The complexity of this test case, dictated by the environment and by the features of the data set at disposal, inspires the formulation of a general method for estimating production upgrades, based on multivariate linear modeling of the power output of the upgraded wind turbine. The method is a distinctive part of the outcome of this work because it is generalizable to the study of whatever wind turbine upgrade and it is adaptable to the features of the data sets at disposal. In particular, applying this model to the test case of interest, it arises that the upgrade increases the annual energy production of the wind turbine of an amount of the order of the 2%. This quantity is of the same order of magnitude, albeit non-negligibly lower, than the estimate based on the assumption of ideal wind conditions. Therefore, it arises that complex wind conditions might affect the efficiency of wind turbine upgrades and it is therefore important to estimate their impact using data from wind turbines operating in the real environment of interest.

Keywords: wind energy; wind turbines; Supervisory Control And Data Acquisition; retrofitting; performance evaluation.

---

1. Introduction

The upgrade of operating wind turbines has been spreading in the wind energy industry, in order to improve the efficiency of the conversion of wind kinetic energy. This kind of interventions has material and labor cost and producible energy is lost during the installation. Therefore, the priority of wind farm owners dealing with wind turbine upgrades is to estimate realistically the profitability. Furthermore, the adoption of upgrades might exacerbate stressing mechanical conditions to which wind turbines are subjected. For example, a typical control system upgrade arising this kind of issues is the extension of the wind turbine operation above the cut-out wind speed [1,2]. For this reason, the analysis of wind turbine upgrades through operational data is attracting an increasing attention in the wind energy scientific literature.

Computing the energy improvement from a wind turbine upgrade is not straightforward because wind turbines operate under non-stationary conditions. Therefore, it makes little sense to compare the pre-upgrade production against the post-upgrade one. For this reason, the simplest consistent strategy

is inquiring how the relationship between wind speed and power output changes: the International Electrotechnical Commission (IEC) [3] has established widely accepted standards for analyzing the power curve. Basically, the power output of the wind turbine is averaged on wind speed intervals (commonly of 0.5 or 1 m/s); the dependency on ambient factors is addressed by normalizing the wind speed to standard air density conditions (at sea level and 15°C). Typically, data describing the wind turbine operating under the wake of a nearby one [4–8] are filtered away.

The power curve analysis is simple and intuitive, but has some drawbacks for the study of wind turbine upgrades. One drawback is that data describing the wind turbines operating under the wake of a nearby one are filtered away. The rationale for this choice is that the power curve is commonly used for detecting within what extent the performances under optimal conditions, *i.e.*, the absence of wakes, resemble the expected theoretical ones. However, in the case of wind turbine upgrades the focus should rather be in the actual performances (and therefore in the production improvement) in the real operating conditions, that include wakes, icing, wind shear induced by the terrain [9–12] and so on. This drawback can be circumvented by disregarding within a certain extent the IEC guidelines and, for example, deciding not to filter away data associated to wind turbines operating under wake. However, there is a further drawback that cannot be circumvented when it manifests: it might happen that the wind turbine upgrade affects the measuring chain of the Supervisory Control And Data Acquisition (SCADA) system. In such case, the power curve analysis cannot be performed because of the unavailability or the insufficient quality of the necessary measurements.

Therefore, the general strategy for quantifying the effects of an upgrade is comparing the post-upgrade production against a reliable simulation of the pre-upgrade production in the same conditions. The model for simulating the power output of the upgraded wind turbine can be the power curve when possible, or it can be a more complicated model if necessary, as in the test case of the present work. The study of this kind of problems and the formulation of adequate data-driven models addressing the occurring criticality is quite recent and there are substantially three relevant studies on this topic: [13], [14] and [15]. In [13], a kernel plus method is proposed for computing wind turbine performance upgrades. In general, kernel regression is a non-parametric method: the available measurements are employed for simulating an output after being weighted with a kernel function (typically Gaussian). In [13], a modified version of the kernel is proposed (hence, named kernel plus) for addressing dataset dimensionality and bias issues: it has a hybrid structure that includes multiplicative kernel functions in an additive model. This method is employed for studying pitch angle control optimization and aerodynamic retrofitting (vortex generators installation on the blades [16]). In [14], an academia-industry joint study is presented and the test cases are two onshore wind farms that have been upgraded through the installation of vortex generators. The approach is twofold: on one hand, SCADA data with 10 minutes of sampling time are employed; on the other hand, high-frequency power data are employed. The estimates of production improvement are shown to be similar. The adopted methods are the kernel plus (as in [13]) and a power-power approach, based on observing how the difference between the power of the upgraded wind turbine and the power of a reference wind turbine changes after the upgrade.

This study, as well as [15], also is a partnership between academia and industry: the involved company is Renvico[1], owning and managing 335 MW of wind turbines in Italy and France. In [15], three cases of upgrades on the wind turbines owned by Renvico are studied: (1) the pitch angle optimization near the cut-in, (2) the installation of vortex generators and passive flow control devices, (3) the extension of the power curve for high wind intensity through a soft cut-out strategy. It is argued that a data-driven study of the impact of wind turbine power curve upgrades in complex environments can lead to estimates that are non-negligibly different from what can be obtained basing on the hypothesis of ideal operating conditions as flat terrain, absence of wake interactions and so

---

[1]   www.renvicoenergy.com

on. The complexity of one of the test cases in [15] has inspired a systematic investigation of wind turbine power curve upgrades: this is one of the motivations of the present work and of its philosophy, focused on the formulation of the method as well as the computation of the upgrade impact.

The test case of the present work is a multi-megawatt wind turbine from a wind farm sited in a very complex terrain [11,17,18] that has been upgraded through the installation of vortex generators [16,19,20] and passive flow control devices [21], increasing the lift and therefore the energy production of the turbine. This has been a pilot test and the decision of adopting the retrofitting on the whole wind farm has been based on the quantification of the benefits by means of this study. One key point of this study is that the measurements from the nacelle anemometer of the retrofitted wind turbine are unreliable. This implies that the power curve analysis is impossible and, moreover, the wind speed of the upgraded wind turbine cannot be used to model its power output. For this reason, the wind turbines nearby the upgraded one must be used as reference for the wind conditions. Furthermore, the complexity of the terrain makes it difficult to reliably use the power-power approach of [14], because the power difference between two wind turbines fluctuates severely. Therefore, a more complex and robust procedure must be formulated.

On these grounds, an appropriate multivariate linear model has 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 (2% of the annual energy production is the estimated order of magnitude), 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. Moreover, having tested it for the study of an upgrade in complex terrain on a data set having the severe issue of the unreliability of nacelle anemometry, it is demonstrated its robustness and versatility.

The structure of the paper is the following. In Section 2, the test case and the data sets are described. Section 3 is devoted to the description of the method. The results are collected in Section 4. In Section 5 conclusions are drawn and further directions are indicated.

2. The test case and the data set

The wind farm under investigation has recently attracted a certain attention in the scientific literature, because it is an interesting test case for the study of wind flow, wake-terrain interactions and wind turbine performances in complex terrain [11,17,18]. In Figure 1, the layout of the wind farm is sketched. From the contour lines, it is possible to appreciate the complexity of the terrain. The upgraded wind turbine object of this study is denoted as T7 (colored in red in Figure 1). The wind turbines in the farm are multi-megawatt, the rotor diameter with blades is 93 meter and the hub height is 80 meters above ground level. The cut-in is 4 m/s and the cut-out is 25 m/s and the nominal wind speed is 13-14 m/s. The rotor rotational speed goes from 6 to 16 rpm. The gearbox is 3-stage, the generator is asynchronous having synchronous speed of 1500 rpm and the main brake is aerodynamic through pitch angle adjustment.
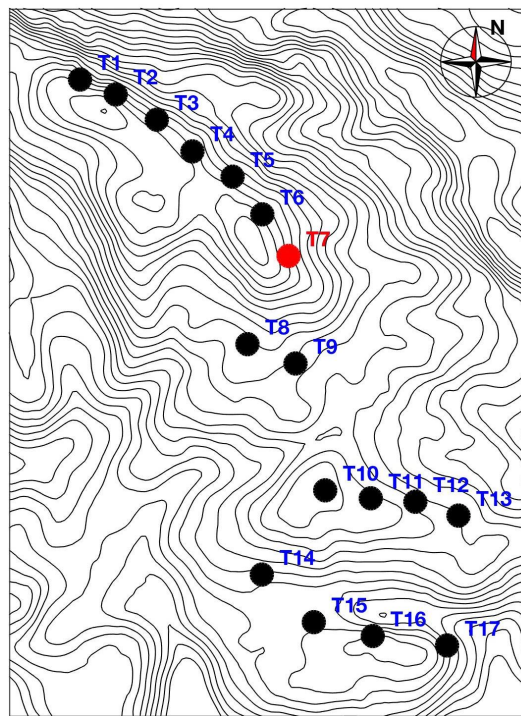
Figure 1. The layout of the wind farm.

For this study, the SCADA collected data have been exploited and organized in two datasets as follows:

- The first dataset is denoted as D_{bef} and contains the data collected from 01/01/2016 to 01/07/2017. It is a period during which the standard blade configuration was adopted.
- The second dataset is denoted as D_{aft} and contains the data collected from 01/09/2017 to 01/05/2018. It is a period during which T7 has been operating with the improved blade configuration.

The SCADA data are recorded with 10 minutes of sampling time through a microprocessor controller and are WPS-transmitted via modem.

3. The method

Before discussing the method, it is interesting to motivate why a power curve analysis cannot be performed for a qualitative assessment of the production improvement. In Figure 2, the power coefficient C_p as estimated from the SCADA data during the D_{aft} data set is reported as a function of the wind speed for the upgraded turbine T7 and for another sample wind turbine from the wind farm, T2. Note that T2 has been selected as reference, but whatever wind turbine (except, of course, T7) could have been chosen for a qualitative comparison. The power coefficient is defined as:

$$C_p = \frac{P}{\frac{1}{2}\rho A v_\infty^3} \tag{1}$$

where P is the measured power output, ρ is the air density on site, A is the blade swept area and v_∞ is the undisturbed wind speed as reconstructed from the nacelle wind speed through the nacelle transfer function. From Figure 2, it can be noticed that the measurements for T7 are implausible, because well above the Betz limit. The possible explanation for this fact is that the nacelle transfer function has not been updated by the manufacturer after the upgrade of T7.
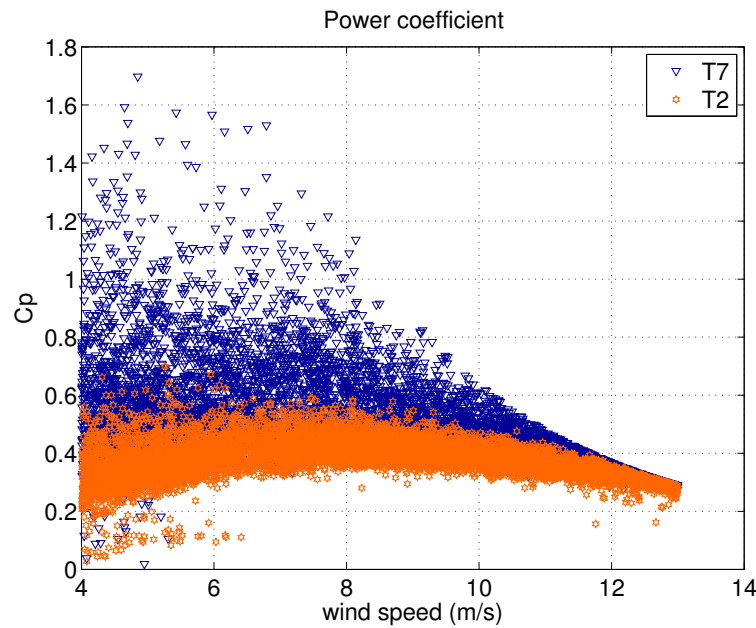
Figure 2. The power coefficient C_p, as computed from the SCADA data, vs. nacelle wind speed: T7 and a sample wind turbine (T2), D_{aft} data set.

The same conclusion can be drawn also by another point of view: assume that the wind speed measurements at T7 during D_{aft} are plausible. Therefore, an order of magnitude estimation of the production improvement can be simply computed by comparing the produced power against the theoretical power, according to the pre-upgrade power curve provided by the wind turbine manufacturer. Doing this, it is obtained that the improvement is of the order of 10% of the annual energy production. This is of course implausible, because vortex generators and passive flow control devices are known to provide an improvement of the order of few percents. The conclusion is that the wind speed measurements at T7 are unreliable after the installation of the upgrade.

On the grounds of the above discussion, it is necessary to adopt the wind turbines nearby T7 as references for assessing the energy improvement. In the following, an argument is reported why a power-power approach (similar to the one in [14]) can't be considered robust enough for the present test case sited in a very complex terrain.

The idea is studying how the difference between the power of the upgraded wind turbine (T7) and a reference wind turbine (T6 is selected because it is the nearest) changes after the upgrade of T7. Therefore, the difference of the powers of T7 and T6 is computed on the data sets D_{bef} and D_{aft}, once they are opportunely filtered. The requests for the filter are that the T7 and T6 are producing power and that T7 catches free wind: therefore, the nacelle position corresponding to the sectors free from wakes (on the grounds of IEC guidelines [3]) are selected. The results are collected in Figure 3: the x-axis is the power of the reference wind turbine (T6) and the y-axis is the power difference between T7 and T6. Data are averaged in T6 power intervals having amplitude of the 10% of the rated power. In Figure 3, for the post upgrade data set, also the standard deviation of the data inside each interval is reported. It clearly arise that the average behavior of the power difference is compatible with the conclusion that there has been a power improvement for T7, but the data sets before and after the upgrade are perfectly compatible between one standard deviation of the data inside each T6 power interval.
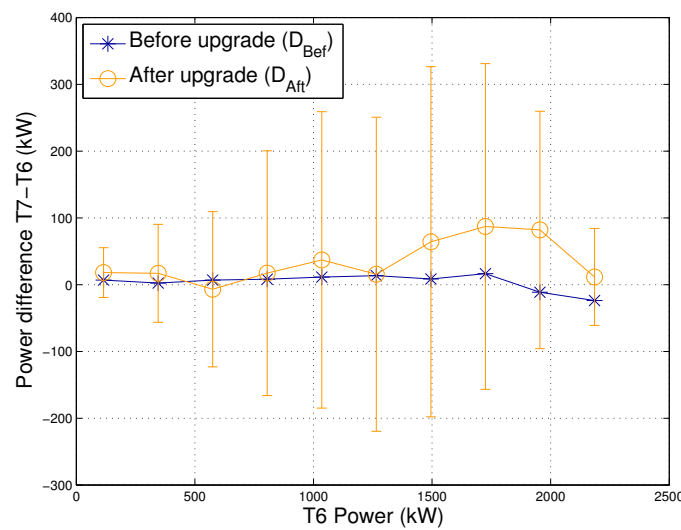
Figure 3. The power difference between T7 and T6, before and after the upgrade.

On these grounds, the power-power method has been considered not solid enough for this test case and this is definitely reasonable, because the terrain is very complex. Therefore, it has been decided to formulate a more robust model, at the possible cost of increasing its complexity. This has been done in a completely general way.

The output of the model, denoted as y, is the power output of T7. The possible input data at disposal, from each of the remainder 16 wind turbines in the farm, are:

- nacelle wind speed,
- power output,
- individual blade pitch angles,
- rotor revolutions per minute,
- high speed rotor temperature.

The data are filtered on the condition of power output production from all the wind turbines in the farm (using the appropriate counter of grid production, available in the SCADA data set) and of power production of T7 below the rated, because at rated power the upgrade has no visible effect.

In the following, the procedure is reported for selecting the inputs appropriately. Note that, the measurements of the output y are assembled in a column vector $\boldsymbol{y}$ and the values of the inputs x are assembled in a matrix $\underline{x}$. A stepwise regression is performed: it is a systematic method for adding and removing terms from a multilinear model, based on their statistical significance in a regression task. The algorithm begins with an initial model and then compares the explanatory power of incrementally larger and smaller models, where 'larger' and 'smaller' refers to the number of used regression terms. The candidate models are obtained via the stepwise regression algorithm based on the value of the p_{remove} parameter. This is an exit tolerance for the probability of rejecting the hypothesis of a zero coefficient for a given input in the model.

In this work, to obtain a robust input selection strategy and consequently a reliable estimate of the power output of the turbine under investigation, the value of the p_{remove} parameter has been chosen after an extensive study. The tested values of p_{remove} are $10^{-\gamma}$ with $\gamma = 1, \ldots, 15$. For each value of p_{remove}, the input selection and estimation pipeline is subjected J times to k-fold cross-validation [22]. In particular, the data set D_{bef} is divided J times randomly in two subsets: (k-1)/k of the data are used for training the model and the remaining 1/k are used for for validation. k has been set to 10 for this study, because the objective is validating the model for very short folds, in order to test its robustness. J has been set to 300 to increase the significance of the study.

For each of the $j = 1, \ldots, J$ runs of the cross-validation, for a given value of p_{remove}, the most significant inputs are selected and the estimation model is obtained as follows. The training input

matrix $\underline{x}_{train_{j,\gamma}}$, constructed with the inputs using the measurements selected for the training, is normalized and its pseudo-inverse is used to compute the weight matrix as

$$\underline{W}_{j,\gamma} = \underline{x}^{-1}_{train_{j,\gamma}} \cdot \boldsymbol{y}_{train_{j,\gamma}} \tag{2}$$

and the estimated output on the validation data set is computed as

$$\hat{\boldsymbol{y}}_{valid_{j,\gamma}} = \boldsymbol{x}_{valid_{j,\gamma}} \cdot \underline{W}_{j,\gamma} \tag{3}$$

For each of the J runs, the mean absolute error is computed as follows:

$$\bar{\delta}_{j,\gamma} = \overline{|\hat{\boldsymbol{y}}_{valid_{j,\gamma}} - \boldsymbol{y}_{valid_{j,\gamma}}|} \tag{4}$$

Subsequently, it is possible to average the $\bar{\delta}_{j,\gamma}$ over j and obtain the average mean absolute error for a given γ:

$$\bar{\delta}_p = \sum_{j=1}^{J} \frac{\bar{\delta}_{j,\gamma}}{J}$$

Table 1. Results of the model validation

γ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\bar{\delta}_\gamma$ (kW)	97	97	98	98	98	99	99	100	100	100	100	100	100	100	100
N. of regressors	13	11	6	6	6	8	6	5	4	3	3	3	3	3	3

It arises that $\bar{\delta}_\gamma$ is of the same order of magnitude ($\simeq 100$ kW) for each value of p_{remove}: the improvement in choosing $p_{remove} = 10^{-1}$ instead of $p_{remove} = 10^{-15}$ is averagely only 3 kW in $\bar{\delta}_\gamma$, at the cost of employing many more regressors (13 against 3) and having a less stable model (the number of possible regressor configurations is much higher).

For whatever choice of $p_{remove} \leq 10^{-10}$, instead, one obtains that the selected inputs are the same (except for a small number of outliers explaining the values in the third row of Table 1) 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.

In the following Section 4, it is shown how to use this model to obtain an estimate of the energy improvement since the adoption of the T7 upgrade.

4. The results

The data sets at disposal are employed as follows:

- D_{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 $\underline{W}$, D1 is used for validating the model.

213   • D_{aft} (also named D2 for simplifying the notation in the following) is used in its entirety to
214   quantify the performance improvement.

The approach is based on the analysis, for the data sets D1 and D2, of the residuals R between
the measurement y and the estimation $\hat{y}$ as a function of the inputs x (power of T6, power of T9, rotor
revolutions per minute of T8, as indicated in Section 3). In other words, the interest is in how the
residuals vary after the upgrade with respect to before. Therefore, consider Equation 5 with $i = 1, 2$.

$$R(x_i) = y(x_i) - \hat{y}(x_i). \tag{5}$$

For $i = 1, 2$, one computes

$$\Delta_i = 100 * \frac{\sum_{x \in Di} (y(x) - \hat{y}(x))}{\sum_{x \in Di} y(x)} \tag{6}$$

Since Δ_i is constructed with the relative discrepancies of power data each having the same sampling
time (10 minutes), the quantity

$$\Delta = \Delta_2 - \Delta_1 \tag{7}$$

215   provides a percentage estimate also of the energy improvement.
216   The above procedure has been repeated several times, with several random choices of D0 and D1
217   for training and validating the model, until the standard deviation of the estimates of Δ has reached a
218   plateau. In Table 2, statistics are reported for the discrepancy between estimation and measurement for
219   the different choices of the D0 and D1 data sets: average residual between measurement and estimation
220   (δ_{ave}), average mean absolute error ($\bar{\delta}_{ave}$), average standard deviation of the residuals (σ_δ).

Table 2. Statistical behavior of the residuals between measurement and estimation, for the different
random choices of the D0 and D1 data set.

Residual	δ_{ave} (kW)	$\bar{\delta}_{ave}$ (kW)	σ_δ (kW)
$R(x_1)$	-0.5	100	146
$R(x_2)$	+48	120	160

221   From Tables 2, it arises that the average discrepancy between estimation and measurement is
222   negligible on the D1 data set after having trained the model with the data sets D0; the measurements
223   instead are averagely 50 kW higher in the post upgrade data set D2 with respect to the estimate
224   provided by the model trained with pre-upgrade data. This is a clear indication of the fact that the
225   wind turbine upgrade produces a non-negligible power improvement.

Further, a Student's t-test can be performed to detect the difference in the residuals $R(x_1)$ and
$R(x_2)$. The t statistic is computed as

$$t = \frac{\bar{R}_2 - \bar{R}_1}{\sigma_R \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}. \tag{8}$$

In Equation 8, N_1 and N_2 are the number of measurements respectively in D1 and D2, $\bar{R}_2$ and $\bar{R}_1$ are
the average residuals between measurement and model respectively in D1 and D2 and σ_R is given by

$$\sigma_R = \sqrt{\frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}}, \tag{9}$$

226   where S_1 and S_2 are the standard deviations of the residuals in data sets D1 and D2. For each of the
227   model runs, the value of the t statistic is computed to be less than 10^{-14}: this means that the probability
228   that the data sets are compatible with the hypothesis that there has not been an improvement in the
229   production of T7 is practically zero.

230 The improvement is appreciable also from Figure 4: it is a plot of $R(x_1)$ and $R(x_2)$ on a sample
231 model run. The data are averaged in power production intervals, whose amplitude is 10% of the rated.
232 This plot allows to appreciate how the residuals between measurements and estimations vary after the
233 upgrade: the difference between measurement and estimation increases in the post-upgrade period,
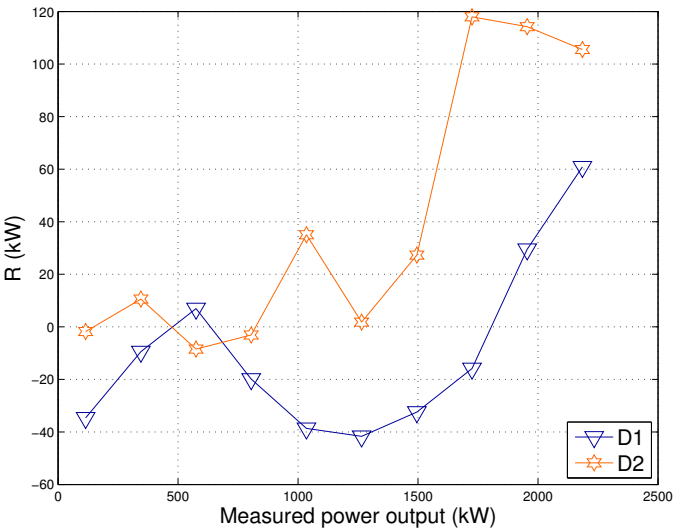234 especially approaching rated power.

Figure 4. The average difference R between measurements and simulation (Equation 5), for data sets
D1 and D2 and for a sample run of the model.

235 The described procedure, being based on repeated random choices of D0 and D1, allows to have
236 not only an average estimation of the energy improvement, but also a standard deviation (and therefore
237 reasonable lower and upper limits for the energy improvement). The average energy improvement is
238 $\Delta = 3.9\%$. In other words, since the upgrade has been installed, T7 has produced below rated power
239 the 3.9% more than it would have done without aerodynamic improvement. The standard deviation is
240 $\sigma_\Delta = 0.2\%$: therefore, reasonable upper and lower limits of the energy improvement are $\Delta_+ = 4.1\%$
241 and $\Delta_- = 3.7\%$. This corresponds to an estimation of the improvement of annual energy production
242 (AEP) of the order of $\Delta_{AEP} = 2.0 \pm 0.1\%$.

5. Conclusions and further directions

244 In this work, a test case of wind turbine upgrade (installation of vortex generators and passive flow
245 control devices) has been studied using SCADA data. This work has been organized as a collaboration
246 between academia and industry and it is hopeful that the outcomes stimulate further cooperation. The
247 objective has been a realistic computation of the production improvement since the adoption of the
248 upgrade.
249 This kind of problems is non-trivial in general, because wind turbines are subjected to
250 non-stationary operation conditions and therefore the most appropriate approach is comparing
251 the post-upgrade production against a simulation of the pre-upgrade production under the same
252 conditions.
253 In particular, the selected test case was challenging for two reasons: the first is that the data set at
254 disposal has a severe limitation. In particular, the nacelle anemometer of the upgraded wind turbine is
255 unreliable in the post-upgrade period. Therefore, the wind turbines nearby the upgraded one must
256 be employed as reference of the external conditions. The second peculiarity of the test case is that
257 the wind farm is sited in a very complex terrain and therefore it is difficult to select the appropriate
258 inputs for modeling the power of the upgraded wind turbine. This has stimulated the formulation of a
259 general method for studying wind turbine upgrades, based on stepwise regression for selecting the

most appropriate inputs for modeling a given output. In particular, it has been observed that in this complex-terrain test case the difficulty is mainly in selecting a robust model.

The particular result of this work is that the value of the upgrade has been estimated of the order of the 2% of the annual energy production. It is important to notice that it is of the same order of magnitude as the estimate provided by the wind turbine manufacturer, but it is non-negligibly lower. This can be explained by the fact that commonly wind turbine upgrades are estimated through simulation or field tests in conditions that are quite different from the ones to which wind turbines operating in real environments are subjected. Therefore, the results of the present work point to the conclusion that the complex flow conditions have an impact on the efficiency of passive flow control devices. This observation motivates a valuable further direction of this study: it is important to elaborate on the wake-terrain interaction in this environment and to study how the mechanical conditions of the upgraded wind turbine have changed and could possibly affect the residue lifetime. An important lesson of this work is that it is very important to estimate wind turbine upgrades on real environments through a judicious use of operational data.

The general outcome of this work is the robust and generalizable method. In fact, the study of wind turbine upgrades is conceptually and practically quite different from power curve modeling (a subject about which there is plenty of scientific literature: see for example [23–25]). Therefore, another interesting further direction of this work is extending to other test cases: different wind turbine upgrades, different environments, different peculiarities of the data sets. Furthermore, it is possible to explore non-linear approaches to the problem and investigate what kind of similar problems actually require non-linearity to be successfully studied.

Appendix   A crosscheck of the results

Consider applying the same method for a wind turbine that has not been retrofitted. While the results in Section 4 for wind turbine T7 clearly point to the detection of a production upgrade, it is expected that, selecting a non-upgraded wind turbine, the proposed method indicates that no considerable performance upgrade has occurred. 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 in in Section 2) and are:

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

Adopting the same procedure as in Section 4, one obtains the results reported in Table A1. It arises that the statistical features of the residuals are extremely similar in the sets $R(x_1)$ and $R(x_2)$, differently from what happens for wind turbine T7 (Table 2). In particular, the mean difference between measurement and estimation is zero within 1 kW of tolerance for the sets $R(x_1)$ and $R(x_2)$.

Table A1. Statistical behavior of the residuals between measurement and simulations, for the different random choices of the D0 and D1 data set.

Simulation	δ_{ave} (kW)	$\bar{\delta}_{ave}$ (kW)	σ_δ (kW)
$R(x_1)$	-0.9	129	187
$R(x_2)$	+0.3	140	202

The average AEP variation (basing on Equations 7 and 6) is estimated to be $\Delta_{AEP} = 0.05\%$: practically zero.
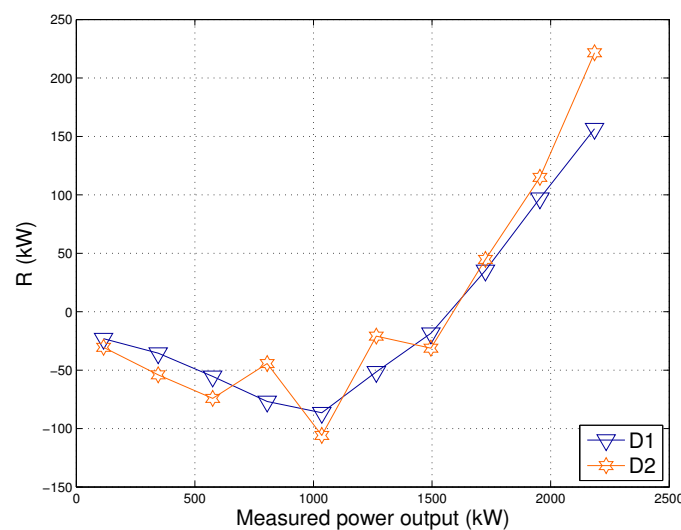
Figure A1. The average difference R between measurements and simulation (Equation 5), for data sets D1 and D2 and for a sample run of the model. T4 test case.

Figure A1 is, similarly to Figure 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 Figure 4 is evident: in Figure A1, 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.

This test demonstrates the versatility and the good estimation capabilities of the proposed approach, which make it suitable to be deployed in operative contexts.

References

1. Bossanyi, E.; King, J. Improving wind farm output predictability by means of a soft cut-out strategy. European Wind Energy Conference and Exhibition EWEA, 2012, Vol. 2012.

2. Petrović, V.; Bottasso, C.L. Wind turbine optimal control during storms. Journal of Physics: Conference Series. IOP Publishing, 2014, Vol. 524, p. 012052.

3. IEC. Power performance measurements of electricity producing wind turbines. Technical Report 61400–12, International Electrotechnical Commission, Geneva, Switzerland, 2005.

4. Barthelmie, R.J.; Hansen, K.; Frandsen, S.T.; Rathmann, O.; Schepers, J.; Schlez, W.; Phillips, J.; Rados, K.; Zervos, A.; Politis, E.; others. Modelling and measuring flow and wind turbine wakes in large wind farms offshore. *Wind Energy* **2009**, *12*, 431–444.

5. Barthelmie, R.J.; Pryor, S.C.; Frandsen, S.T.; Hansen, K.S.; Schepers, J.; Rados, K.; Schlez, W.; Neubert, A.; Jensen, L.; Neckelmann, S. Quantifying the impact of wind turbine wakes on power output at offshore wind farms. *Journal of Atmospheric and Oceanic Technology* **2010**, *27*, 1302–1317.

6. Hansen, K.S.; Barthelmie, R.J.; Jensen, L.E.; Sommer, A. The impact of turbulence intensity and atmospheric stability on power deficits due to wind turbine wakes at Horns Rev wind farm. *Wind Energy* **2012**, *15*, 183–196.

7. Grassi, S.; Junghans, S.; Raubal, M. Assessment of the wake effect on the energy production of onshore wind farms using GIS. *Applied Energy* **2014**.

8. Moriarty, P.; Rodrigo, J.S.; Gancarski, P.; Chuchfield, M.; Naughton, J.W.; Hansen, K.S.; Machefaux, E.; Maguire, E.; Castellani, F.; Terzi, L.; others. Iea-task 31 wakebench: Towards a protocol for wind farm flow model evaluation. part 2: Wind farm wake models. Journal of Physics: Conference Series. IOP Publishing, 2014, Vol. 524, p. 012185.

9. Politis, E.S.; Prospathopoulos, J.; Cabezon, D.; Hansen, K.S.; Chaviaropoulos, P.; Barthelmie, R.J. Modeling wake effects in large wind farms in complex terrain: the problem, the methods and the issues. *Wind Energy* **2012**, *15*, 161–182.

10. Rodrigo, J.S.; Gancarski, P.; Arroyo, R.C.; Moriarty, P.; Chuchfield, M.; Naughton, J.W.; Hansen, K.S.; Machefaux, E.; Koblitz, T.; Maguire, E.; others. Iea-task 31 wakebench: Towards a protocol for wind farm flow model evaluation. part 1: Flow-over-terrain models. Journal of Physics: Conference Series. IOP Publishing, 2014, Vol. 524, p. 012105.

11. Castellani, F.; Astolfi, D.; Mana, M.; Piccioni, E.; Becchetti, M.; Terzi, L. Investigation of terrain and wake effects on the performance of wind farms in complex terrain using numerical and experimental data. *Wind Energy* **2017**, *20*, 1277–1289.

12. Hyvärinen, A.; Segalini, A. Effects from complex terrain on wind-turbine performance. *Journal of Energy Resources Technology* **2017**, *139*, 051205.

13. Lee, G.; Ding, Y.; Xie, L.; Genton, M.G. A kernel plus method for quantifying wind turbine performance upgrades. *Wind Energy* **2015**, *18*, 1207–1219.

14. Hwangbo, H.; Ding, Y.; Eisele, O.; Weinzierl, G.; Lang, U.; Pechlivanoglou, G. Quantifying the effect of vortex generator installation on wind power production: An academia-industry case study. *Renewable Energy* **2017**, *113*, 1589–1597.

15. Astolfi, D.; Castellani, F.; Terzi, L. Wind Turbine Power Curve Upgrades. *Energies* **2018**, *11*, 1300.

16. Øye, S. The effect of vortex generators on the performance of the ELKRAFT 1000 kw turbine. 9th IEA Symposium on Aerodynamics of Wind Turbines, Stockholm, ISSN, 1995, pp. 0590–8809.

17. Castellani, F.; Astolfi, D.; Burlando, M.; Terzi, L. Numerical modelling for wind farm operational assessment in complex terrain. *Journal of Wind Engineering and Industrial Aerodynamics* **2015**, *147*, 320–329.

18. Astolfi, D.; Castellani, F.; Terzi, L. A study of wind turbine wakes in complex terrain through RANS simulation and SCADA data. *Journal of Solar Energy Engineering* **2018**, *140*, 031001.

19. Mueller-Vahl, H.; Pechlivanoglou, G.; Nayeri, C.; Paschereit, C. Vortex generators for wind turbine blades: A combined wind tunnel and wind turbine parametric study. ASME Turbo Expo 2012: Turbine Technical Conference and Exposition. American Society of Mechanical Engineers, 2012, pp. 899–914.

20. Gao, L.; Zhang, H.; Liu, Y.; Han, S. Effects of vortex generators on a blunt trailing-edge airfoil for wind turbines. *Renewable Energy* **2015**, *76*, 303–311.

21. Fernandez-Gamiz, U.; Zulueta, E.; Boyano, A.; Ansoategui, I.; Uriarte, I. Five megawatt wind turbine power output improvements by passive flow control devices. *Energies* **2017**, *10*, 742.

22. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-validation. In *Encyclopedia of database systems*; Springer, 2009; pp. 532–538.

23. Lydia, M.; Kumar, S.S.; Selvakumar, A.I.; Kumar, G.E.P. A comprehensive review on wind turbine power curve modeling techniques. *Renewable and Sustainable Energy Reviews* **2014**, *30*, 452–460.

24. Shokrzadeh, S.; Jozani, M.J.; Bibeau, E. Wind turbine power curve modeling using advanced parametric and nonparametric methods. *IEEE Transactions on Sustainable Energy* **2014**, *5*, 1262–1269.

25. Ouyang, T.; Kusiak, A.; He, Y. Modeling wind-turbine power curve: A data partitioning and mining approach. *Renewable Energy* **2017**, *102*, 1–8.