Are Uncertainty Categories in a Wind Farm Annual Energy Production Estimate Actually Uncorrelated?

Nicola Bodini [1] and Mike Optis [1]

[1]National Renewable Energy Laboratory, Golden, Colorado, USA

Correspondence: Mike Optis (mike.optis@nrel.gov)

Abstract. Calculations of annual energy production (AEP) from a wind farm—whether based on preconstruction or operational data—are critical for wind farm financial transactions. The uncertainty in the AEP calculation is especially important in quantifying risk and is a key factor in determining financing terms. Standard industry practice assumes that different uncertainty categories within an AEP calculation are uncorrelated and can therefore be combined through a sum of squares approach. In this analysis, we assess the rigor of this assumption by performing operational AEP estimates for over 470 wind farms in the United States. We contrast the standard uncertainty assumption with a Monte Carlo approach to uncertainty quantification in which no assumptions of correlation between uncertainty categories are made. Results show that several uncertainty categories do, in fact, show weak to moderate correlations, namely: wind resource interannual variability and the windiness correction (positive correlation), wind resource interannual variability and regression (negative), and wind speed measurement uncertainty and regression (positive). The sources of these correlations are described and illustrated in detail in this paper, and the effect on the total AEP uncertainty calculation is investigated. Based on these results, we conclude that a Monte Carlo approach to AEP uncertainty quantification is more robust and accurate than the industry standard approach.

Copyright statement. This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

1 Introduction

Calculations of wind farm annual energy production (AEP)—whether based on preconstruction data before a wind power plant is built or on operational data after a wind farm has started its operations—are vital for wind farm financial transactions. Preconstruction estimates of AEP are needed to secure and set the terms for new project financing, whereas operational estimates of AEP are required for important wind farm transactions, such as refinancing, purchasing/selling, and mergers/acquisitions. The need for AEP analyses of wind farms is increasing, as global wind capacity increased to 539 GW in 2017, representing

25 11% and 91% increases over 1-year and 5-year periods, respectively; and capacity is expected to increase by another 56% to 841 GW by 2022 (Global Wind Energy Council, 2018).

This rapid growth of the wind energy industry is putting an increased spotlight on the accuracy and consistency of AEP calculations. For preconstruction AEP estimates, there has been considerable movement towards standardization. The International Energy Commission (IEC) is currently developing a standard (IEC 61400-15:draft), and there have long been guidance

30 and best practices available (Brower, 2012). By contrast, operational AEP estimates do not have such extensive guidance or standards. Only limited standards covering some operational analyses exist: IEC 61400-12-1:2017 addresses turbine power curve testing, and IEC 61400-26-3:2016 addresses the derivation and categorization of availability loss metrics. There are, however, no standards and very limited published guidance on calculating AEP from operational data. Rather, documentation seems to be limited to a consultant report (Lindvall et al., 2016), an academic thesis (Khatab, 2017), and limited conference

35 proceedings (Cameron, 2012; Lunacek et al., 2018).

Documentation and standards for preconstruction AEP methods are of limited use for operational-based AEP methods, given the many differences between the two approaches. In general, operational AEP calculations are much simpler than preconstruction estimates because actual measurements of wind farm power production at the revenue meter replace the complicated preconstruction estimate process (e.g., meteorological measurements, wind and wake-flow modeling, turbine performance,

40 estimates of wind farm losses). However, the two methods do share several similarities, including regression relationships between on-site measurements and a long-term wind speed reference, the associated windiness correction, and estimates of uncertainty in the resulting AEP calculation. The uncertainty categories for operational AEP calculations are simplified relative to those in a preconstruction estimate (IEC 61400-15:draft); shared categories between the two methods are listed in Table 1.

Category	Description
On-site measurements	Accuracy in measured met mast wind speeds (preconstruction) or power at the revenue meter (operational)
Long-term reference measurements	Accuracy in long-term reference measured or modeled wind speed data
Losses	Accuracy in estimated or reported availability and curtailment losses
Regression	Confidence in the regression relationship between on-site measurements and long-term reference wind speeds
Windiness correction	Confidence in the long-term correction applied to the on-site measurements
Interannual variability of resource	Uncertainty in future energy production because of resource variability

Table 1. Main Sources of Uncertainty in an Operational AEP Estimate

The uncertainty values from each category listed in Table 1 must be combined to produce a total estimate of AEP uncertainty.

45 We found no guidance in the literature for combining uncertainty categories in an operational AEP estimate. However, con-

siderable guidance exists for combining preconstruction uncertainties (Lackner et al.; Brower, 2012; Vaisala, 2014; Kalkan; Clifton et al., 2016). In every case, recommended best practices assume that all uncertainties, σ_i, are uncorrelated and can therefore be combined using a sum of squares approach to give the total AEP uncertainty, $\sigma_{\text{tot,uncorr}}$:

$$\sigma_{\text{tot,uncorr}} = \sqrt{\sum_i \sigma_{\text{i}}^2} \tag{1}$$

50   To better understand how uncertainties are combined in operational AEP calculations, we reached out to several wind energy consultants who regularly perform these analyses. These conversations revealed that uncertainties in an operational AEP calculation are also assumed uncorrelated and combined using Equation 1.

1.1   Goal of Study

The purpose of this study is to examine the extent to which the assumption of uncorrelated uncertainties—and therefore
55   the combination of those uncertainties through a sum of squares approach—is accurate and appropriate for operational AEP calculations. Specifically, this study aims to identify potential correlations between AEP uncertainty categories and propose a Monte Carlo approach to capture such correlations when combining individual uncertainty categories. The focus here is on operational AEP uncertainty, given publicly available wind farm operational data and the more simple AEP calculation relative to the preconstruction method. However, results from this analysis—namely the potential identification of correlated
60   uncertainty categories—are equally relevant for informing and improving preconstruction AEP methods.

In Section 2, we first describe the data sources used in this analysis—namely wind farm operational data and reanalysis products—as well as the Monte Carlo approach to calculate AEP. Section 3 presents the main results of our analysis, in terms of uncertainty contributions and correlation among the different categories. We conclude and suggest future work in Section 4.

2   Data and Methods

2.1   Wind Farm Operational Data and Reanalysis Products

Wind farm energy production data for this analysis were obtained from the publicly available Energy Information Administration (EIA) 923 database (EIA, 2018). This database provides reporting of monthly net energy production from all power plants in the United States, including wind farms. A total of over 670 unique wind farms were available from this data set.

Long-term wind speed data (needed to perform the "windiness correction" in an AEP estimate) are used from three reanalysis
70   products over the period of January 1997 through December 2017:

- The Modern-Era Retrospective analysis for Research and Applications v2 (MERRA-2) (Gelaro et al., 2017). We specifically use the M2T1NXSLV data product, which provides diagnostic wind speed at 50 m above ground level (AGL), interpolated from the lowest model level output (on average about 32 m AGL), using Monin Obukhov similarity theory. Data are provided at an hourly time resolution.
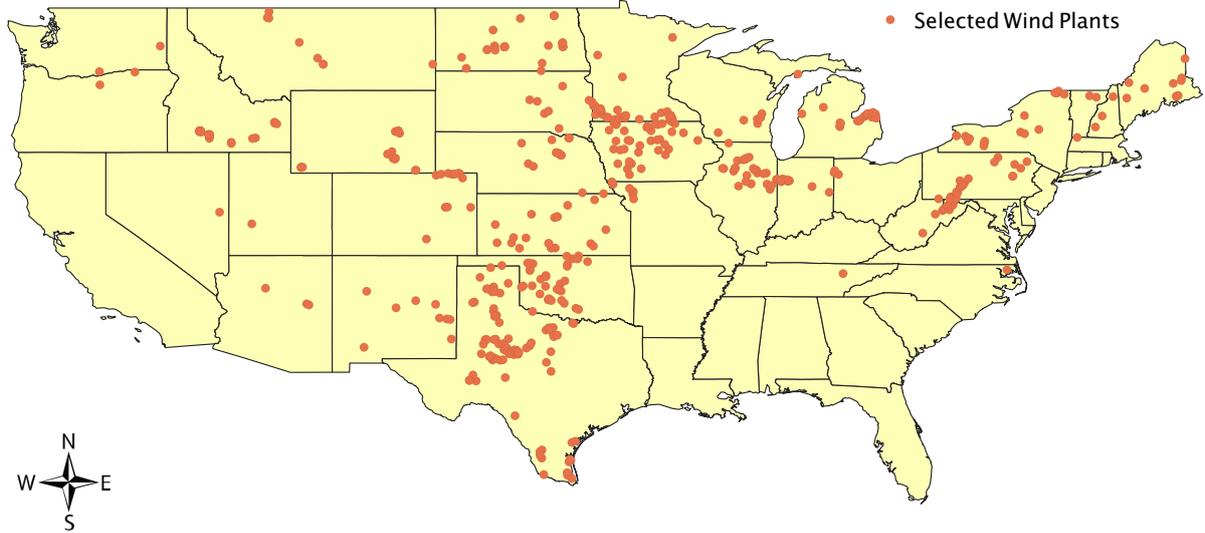
Figure 1. Map of the 472 wind farms that were considered in this study

75   – The European Reanalysis Interim (ERA-interim) (Dee et al., 2011). We specifically use output at the 58th model level, which on average corresponds to a height of about 72 m AGL. Data are provided at 6-hourly time resolution.

     – The National Centers for Environmental Prediction v2 (NCEP-2) (Saha et al., 2014). We specifically use diagnostic wind speed data at 10 m AGL. Data are provided at a 6-hourly time resolution.

The wind speed data are density-corrected at their native time resolutions to correlate more strongly with wind farm power
80   production (i.e., higher density air in winter produces more power than lower density air in summer, wind speed being the same):

$$U_{\mathrm{dens,corr}} = U \left(\frac{\rho}{\rho_{\mathrm{mean}}} \right)^{1/3} \tag{2}$$

where $U_{\mathrm{dens,corr}}$ is the density-corrected wind speed, U is the wind speed, ρ is air density (calculated at the same height as wind speed), and ρ_{mean} is the mean density over the entire period of record of the reanalysis product.

85   To calculate air density at the same height as wind speed, we first extrapolate the reported surface pressure to the wind speed measurement height, assuming hydrostatic equilibrium:

$$p = p_{\mathrm{surf}} \exp \left[\frac{gz}{RT_{\mathrm{avg}}} \right] \tag{3}$$

where p is the pressure at the wind speed measurement height, p_{surf} is the surface pressure, g is the acceleration caused by gravity, z is the wind speed measurement height, R is the gas constant, and T_{avg} is the average temperature between the
90   reported value at 2 m AGL and at the wind speed measurement height. To compute air density at the wind speed measurement height, the ideal gas law assumption is used.

To lessen the impact of limited and/or poor quality data on the results of our analysis, we filter for wind farms with at least 8 months of data and with a moderate-to-strong correlation with all three reanalysis products ($R^2 > 0.6$). A threshold of 8 months is selected in order to investigate uncertainty as it relates to a low number of data points but not so low as to make the use of a regression relationship questionable. A total of 472 wind farms were kept for the final analysis, and their locations are shown in Figure 1. Because obtaining an accurate representation of wind data in complex terrain by reanalysis products is challenging, most of the selected wind plants are located in the Midwest and Southern Plains. Notably, no wind farms in California pass the filtering criteria, because they are predominately located in areas with thermally driven wind regimes such as Tehachapi Pass, where coarse-resolution reanalysis products are poor predictors of wind energy production.

2.2 Operational AEP Methodology

Given the current lack of existing guidelines that offer a standard approach for operational AEP calculations, we instead base our methodology from conversations with several wind energy consultants. These conversations overwhelmingly revealed the following characteristics of an industry standard and bankable[1] operational AEP analysis:

1. Analysis is performed on a monthly timescale (i.e., monthly energy production data, monthly average availability and curtailment losses, and monthly average wind speeds from a long-term wind resource product).

2. Linear regression between monthly energy production and average wind speeds is used to perform the windiness correction.

3. Monthly revenue meter data are corrected for monthly availability and curtailment (i.e., to calculate gross energy) to improve the linear regression relationship[2].

4. Monthly energy production is normalized to 30-day months to improve the accuracy of the regression relationship.

5. Slope and intercept values from the regression relationship are then applied to 10–20 years of long-term wind resource data to perform the windiness correction. Long-term monthly gross energy production (i.e., average January wind speed, average February wind speed, and so forth) is then calculated.

6. The resulting long-term monthly gross energy estimates are then "denormalized" to the normal number of days per month, and long-term estimates of availability and curtailment losses are finally applied to arrive at an long-term calculation of operational AEP.

A diagram outlining this general process is shown in Figure 2.

---

[1]Results are accepted by banks, investors, and so on for use in financing, buying/selling, and acquiring wind farms.

[2]These loss data are not available in the EIA-923 database and therefore are not considered in this analysis.
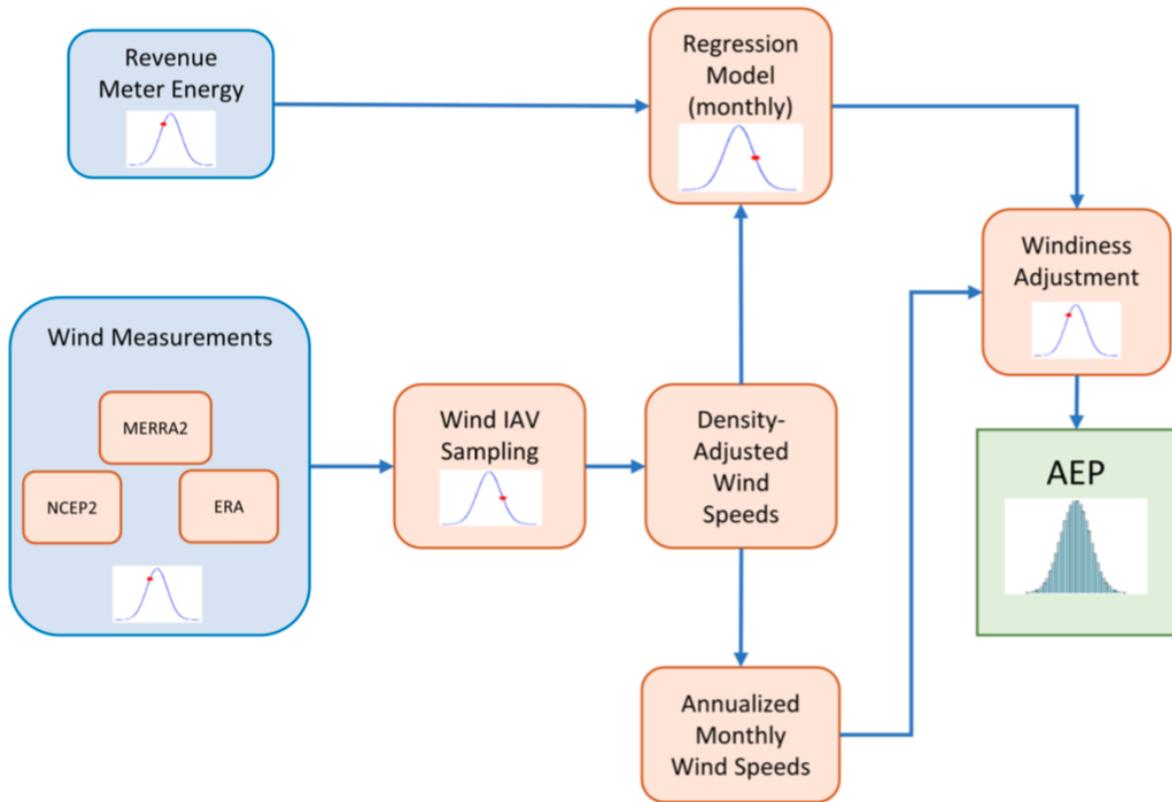
Figure 2. Annual energy production estimation process using operational data under a Monte Carlo approach; source of uncertainty and points of Monte Carlo sampling are denoted by probability distribution images. Note: IAV for inter-annual variability.

2.3 Monte Carlo Analysis

To quantify uncertainty from the AEP calculation described in the previous section, we implement a Monte Carlo approach.

120   In general, a Monte Carlo approach involves the randomized sampling of inputs to or calculations within a method which, when repeated many times, results in a distribution of possible outcomes from which uncertainty can be deduced (usually calculated as the standard deviation of the distribution). We apply this approach to the operational AEP calculation to quantify the key sources of uncertainty. The procedure is repeated 10,000 times under random sampling of the key uncertainty sources to produce a distribution of AEP values from which total uncertainty can be quantified. In this process, we consider five

125   uncertainty categories and ways to incorporate them in the Monte Carlo approach, as listed in Table 2. Note that uncertainty categories related to availability and curtailment losses are not considered because the EIA 923 database does not include measurements of these losses.

Category	Incorporation in Monte Carlo approach
Revenue meter accuracy	Sampling monthly revenue meter data from a distribution with an imposed 0.5% uncertainty.
Wind measurement accuracy	Randomly selecting one of the three reanalysis products for each Monte Carlo iteration.
Wind interannual variability (IAV)	Sampling the long-term average calendar monthly wind speeds (i.e., average January, average February) based on corresponding long-term uncertainties for each calendar month (calculated from 20-year long reanalysis data).
Regression model	Sampling slope and intercept values from the distribution derived from their standard errors.
Windiness adjustment	Sampling the number of years to use in the windiness correction (between 10 and 20).

Table 2. Sources of Uncertainty and their incorporation in the Monte Carlo approach for Operational AEP Estimate
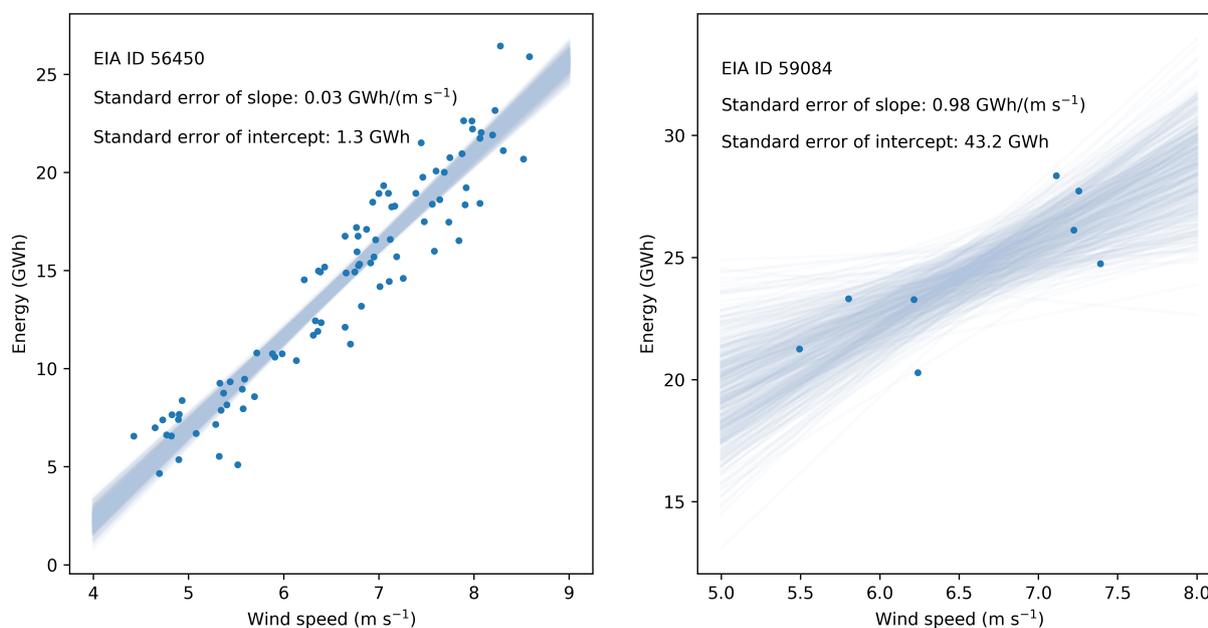
Figure 3. Sampling set of regression lines corresponding to the slope and intercept values derived from their standard errors in the Monte Carlo approach, for two stations in the EIA dataset.

Given the approach to calculating regression uncertainty described in Table 2, we describe it in more detail here. For a regression model between an independent variable, x, and a dependent variable, y, we can define the standard error of the

130 regression:

$$e_y = \sqrt{\frac{\sum \left(y_i - \hat{y_i}\right)^2}{n-2}}, \tag{4}$$

where $\hat{y_i}$ is the regression-predicted value for y_i, and n is the number of data points used in the regression. We can then introduce the standard error of the regression slope:

$$e_a = \frac{e_y}{\sum \left(x_i - \overline{x_i}\right)^2}, \tag{5}$$

135 and the standard error of the intercept:

$$e_b = e_y \, e_a \sqrt{\frac{\sum x_i^2}{n}}. \tag{6}$$

Slope and intercept values are strongly negatively correlated, which is captured by the covariance result when performing linear regression. Therefore, to avoid sampling unrealistic combinations, we constrain the random sampling of slope and intercept values based on this covariance. An example of this sampling is shown in Figure 3 for two projects of different regression

140 strengths. We sample 500 slope and intercept values from a normal distribution centered around the best-fit parameters, and with standard deviation equal to the standard errors of slope and intercept. As shown in the Figure, the low standard errors found for the leftmost regression relationship constrain the possible slope and intercept values that can be sampled while the high standard errors in the rightmost regression relationship allow for a much wider sampling.

Each of the previously mentioned sources of uncertainty, which corresponds to a Monte Carlo sampling, is highlighted

145 by a probability distribution in the flowchart in Figure 2. For each wind farm, we estimate both the total AEP uncertainty and its single components from each uncertainty category considered. Each uncertainty contribution is quantified from the coefficient of variation of the AEP distribution obtained by running the Monte Carlo simulation with a single category of sampling (revenue meter, wind measurement, IAV, regression, or windiness) only. Finally, the total uncertainty is determined by running the Monte Carlo simulation with all five samplings performed simultaneously.

150 Calculations were performed on Eagle, NREL's high-performance computing cluster. Specifically, each wind farm was assigned a different processor and run in parallel. Given the general simplicity of the AEP method used here, computational requirements were moderate despite the 60,000 simulations (10,000 x 6 uncertainty setups) required for each wind farm.

Code used to perform the AEP calculations is published and documented in NREL's open-source operational assessment software, OpenOA.[3]

155 ## 3 Results

3.1 Uncertainty Contributions

The application of the Monte Carlo approach first allows for an assessment of the distributions of the different components of the AEP uncertainty (Figure 4). Uncertainty connected to wind resource IAV is found to contribute the most (average 4.1%

---

[3]https://github.com/NREL/OpenOA
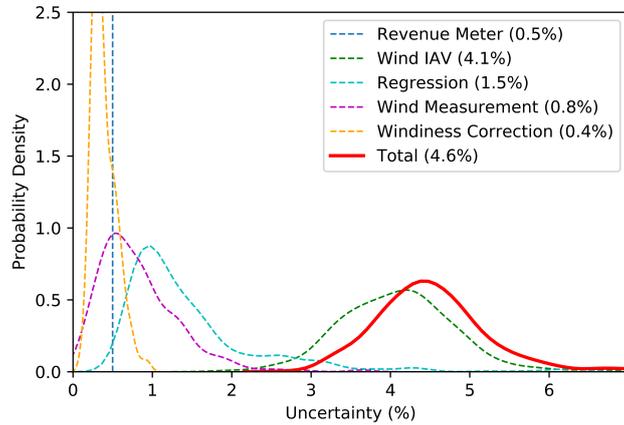
Figure 4. Uncertainty distributions across projects for the different uncertainty categories; mean values across projects are shown in the legend. Note that the sum of squares of the average values of the single components does not add up to the average of the total uncertainty.

across all wind farms). The uncertainty in the regression model has the second largest contribution (1.5%), followed by the
160 uncertainty of the wind measurements (0.8%; here, of the reanalysis products), and revenue meter data (here, imposed at 0.5%, see Table 2). The windiness correction has the smallest uncertainty component (0.4%). Therefore, the number of years used for the long-term windiness correction does not have a large impact on the overall uncertainty in operational AEP, at least for the sampled range of 10–20 years. Using as few as 10 years seems sufficient to give stability to the AEP estimate, and adding additional years does not provide a significant reduction in uncertainty.

165 The total AEP uncertainty calculated with the Monte Carlo approach ($\sigma_{\text{Monte Carlo}}$) can be compared with the uncertainty calculated with the current usual industry standard, which assumes uncorrelated components and calculates the total uncertainty ($\sigma_{\text{uncorrelated}}$) with a sum of squares approach. For the sum of squares approach, each uncertainty contribution is quantified from the coefficient of variation of the AEP distribution obtained by running the Monte Carlo simulation with a single category of sampling. Figure 5 shows the results of this comparison from the 472 considered wind farms, both in terms of a scatterplot and
170 a histogram of the percentage difference, Δ_{σ}, between the two versions of the total uncertainty, calculated as

$$\Delta_{\sigma} = \frac{\sigma_{\text{Monte Carlo}} - \sigma_{\text{uncorrelated}}}{0.5 \cdot (\sigma_{\text{Monte Carlo}} + \sigma_{\text{uncorrelated}})} \cdot 100 \qquad (7)$$

A weak bias can be observed with a median value of -2% in uncertainty percentage difference (which corresponds to a -0.25% median difference in the actual total uncertainty value). In other words, if the correlations between the different uncertainty components are taken into account, the whole AEP uncertainty is then slightly reduced. This difference can be explained by
175 considering that the two biggest sources of uncertainty (regression and IAV) are slightly negatively correlated (as will be shown in detail in the next section), thus making the Monte-Carlo-based uncertainty lower, on average, than the one derived with the uncorrelated assumption. Moreover, ignoring the existing correlation between the uncertainty components can introduce significant errors in the assessment of the AEP uncertainty for the single projects, with about 47% (16%) of the considered wind

WIND
ENERGY
SCIENCE
DISCUSSIONS
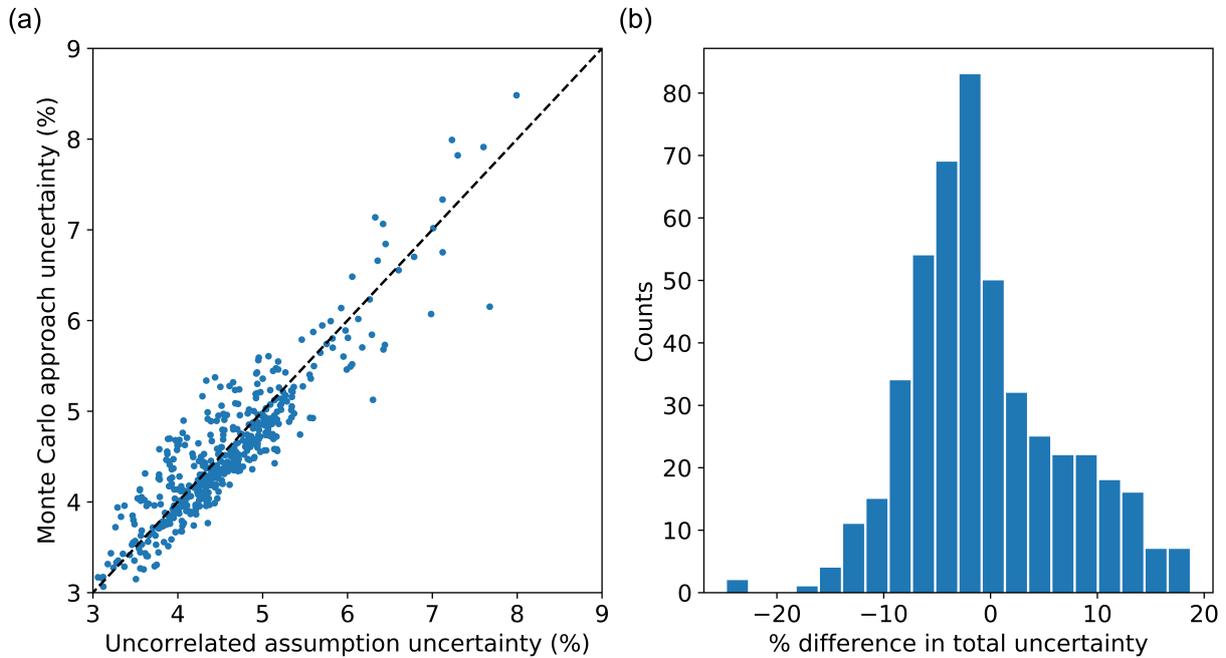
Figure 5. (a) Scatterplot of total AEP uncertainties calculated with the proposed Monte Carlo approach and assuming uncorrelated uncertainty components for the 472 considered wind farms. (b) Histogram of percentage differences between the uncertainties calculated using the two different approaches.

farms showing a ±5% (10%) uncertainty difference compared to the values from the Monte-Carlo-based approach. The mean
180   absolute error of the distribution of uncertainty percentage differences is approximately 6% (Figure shown in the Supplement).

3.2   Uncertainty Correlations

Because AEP uncertainty calculated by ignoring the correlation among its different components can greatly differ from the uncertainty values obtained when considering these correlations, it is worth exploring which contributions are responsible for this difference. By calculating the Pearson correlation coefficients between the different uncertainty components from the 472
185   wind farms, we derived the average correlation matrix in Figure 6. Out of the 10 possible assessments of correlation between uncertainty categories, three pairs are correlated with a $p-$value less than 10^{-5} and therefore of strong statistical significance:

- The wind IAV and the windiness correction uncertainties are moderately correlated ($R = 0.49$, $p = 1.9 \cdot 10^{-29}$).

- The regression and wind measurement uncertainties are weakly correlated ($R = 0.35$, $p = 2.5 \cdot 10^{-15}$).

- The wind IAV and the regression uncertainties appear weakly negatively correlated ($R = -0.21$, $p = 2.6 \cdot 10^{-6}$).

190   The first correlation noted earlier (resource IAV and windiness) is explained simply by the fact that both uncertainties are driven by wind resource variability. At a site with large wind variability, IAV will be large by definition, and so will the uncer-
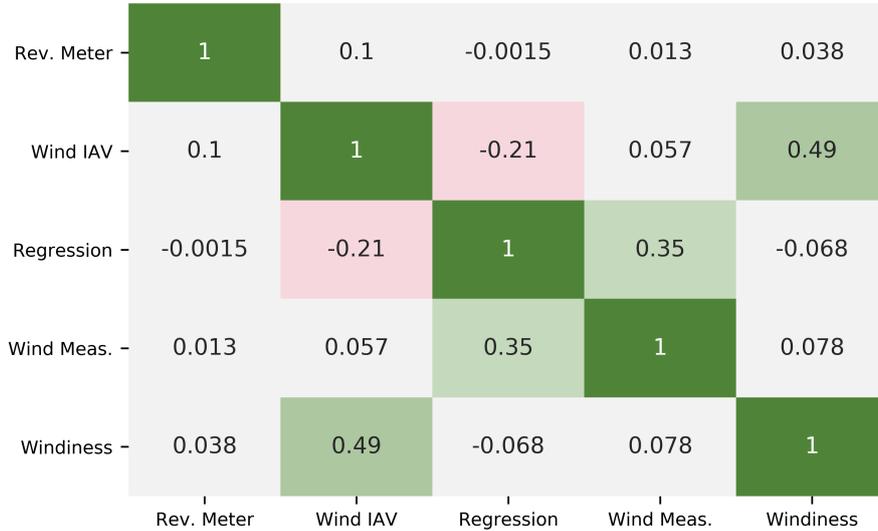
Figure 6. Correlation coefficient heat map between uncertainty categories. Note: "Rev." denotes "Revenue."

tainty introduced by different lengths of time series used for the AEP calculation.

The correlation between regression and measurement uncertainties can be justified, given the dependence of both these uncertainty components on the number of data points (Figure 7). Both the slope and intercept error (Equations 5 and 6), from which the regression uncertainty depends, are inversely proportional to the number of data points, so that when a regression is performed on few data points, its uncertainty increases. This relationship is exemplified in Figure 3, where we compare the sampling sets of regression lines for two stations in the EIA data set: for this case, the standard errors of regression slope and intercept for the station with 8 data points are 30-50 times larger than what is found for the station with 90 data points.

For measurement uncertainty, short periods of wind plant operation record can lead to different interpretations from the reanalysis products as to whether that period of record was above, equal to, or below the long-term average resource. Over a longer period of record, these potential discrepancies between reanalysis products tend to average out, therefore leading to a reduced measurement uncertainty. We illustrate this phenomenon by exploring the long-term trend of the reanalysis products for the wind farm with one of the highest reported measurement uncertainties (EIA ID 60502, reported 3.7% wind speed measurement uncertainty). Figure 8 shows the result. The period of record for wind farm operation (shown as shaded blue in Figure 8) was only 12 months. As shown in the figure, the various reanalysis products have very different interpretations of the period of record wind resource relative to the long-term (ERA-i: 4% above average, MERRA-2: 1% below average; NCEP-2: 1% above average). Consequently, each reanalysis product will make different magnitude (both positive and negative) windiness corrections, leading to high uncertainty in the resulting AEP calculation.
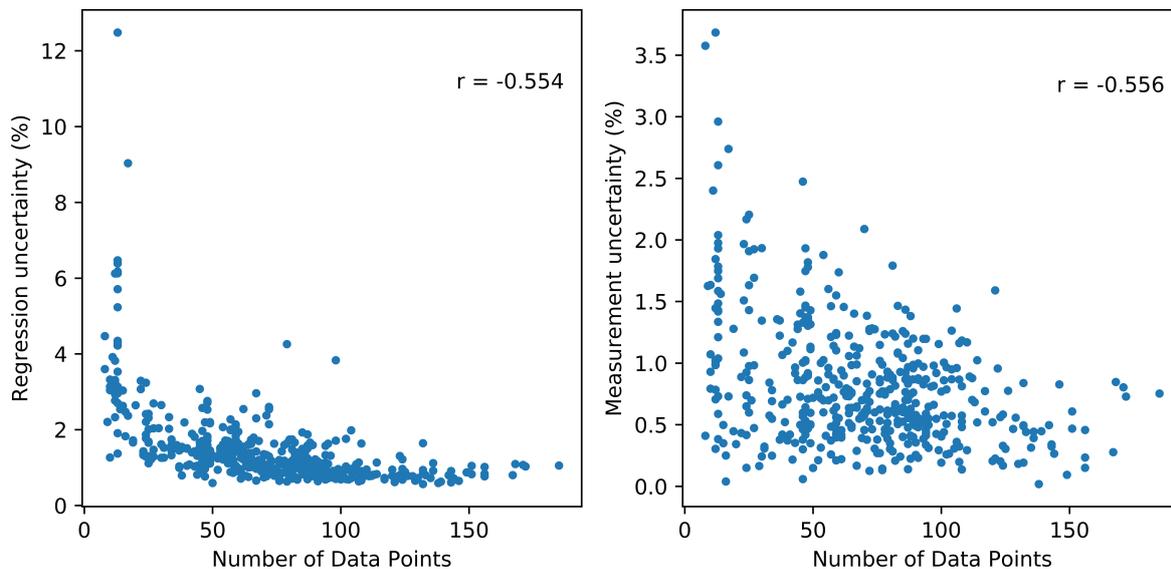
Figure 7. Dependence of regression uncertainty and measurement uncertainty on the number of data points
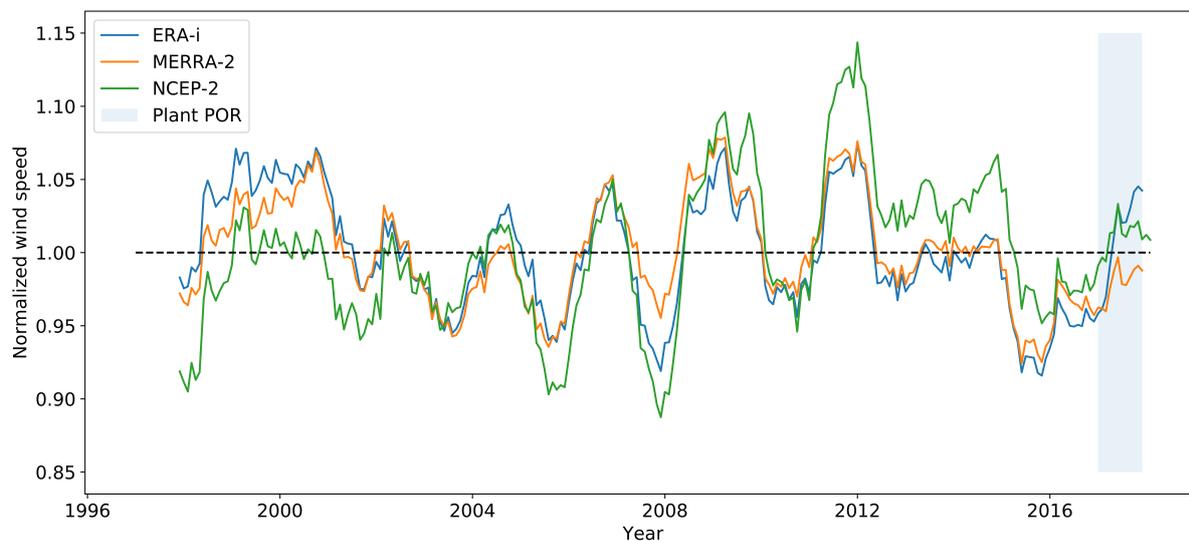
Figure 8. Long-term time series of normalized wind speed for EIA station ID 60502 from the three reanalysis products used in the study. The period of record for the wind farm is highlighted in light blue.

210   By increasing the period of record (i.e., increasing the number of data points), such discrepancies tend to average out. This is illustrated in Figure 9, where we show how the period of record to long-term wind speed ratio varies as we extend the period
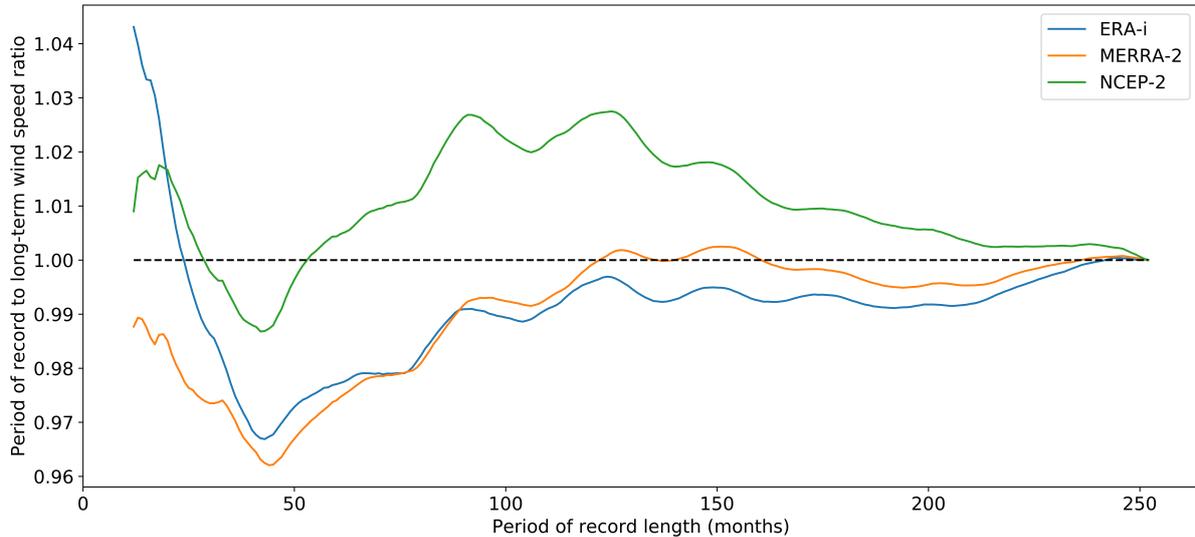
Figure 9. Ratio of wind speed to the long-term, 20-year average for periods of record of different lengths (all ending in December 2017), for EIA station ID 60502 using data from the three reanalysis products in the study

of record by increasing the number of months while keeping December 2017 as the constant ending time. For short periods of record, there is considerable deviation of this ratio between the different reanalysis products (i.e., high wind speed measurement uncertainty). As the length of the period of record increases, this ratio tends to converge to 1.0, and the spread between

215 the three reanalysis products decreases (i.e., low wind speed measurement uncertainty).

Finally, the negative correlation between regression and IAV uncertainties is linked to the fact they respond differently to the R^2 coefficient between the reanalysis wind speed and the energy production data (Figure 10). Predictably, the regression uncertainty is inversely proportional to the coefficient of determination because a stronger correlation between winds and

220 energy production will lead to a reduced uncertainty of the regression between the two variables. On the other hand, IAV uncertainty shows a direct correlation with R^2. We hypothesize that higher IAV leads to large ranges of wind speed in the regression relationship, which acts to "stabilize" regression and increase the regression strength. This phenomenon is illustrated in Figure 11(a). Here, the data set in blue has an equal spread in the regression relationship than the data set in orange but over a large range of wind speeds. As shown in the figure, this longer range (quantified by the coefficient of variation of the wind

225 speeds) leads to a higher R^2 in the regression. We test this hypothesis in Figure 11(b) where this coefficient of variation in a period of record wind speeds is calculated for each wind farm and compared to the regression correlation coefficient. As expected, a moderate correlation is observed. Therefore, we conclude that sites that experience a more variable wind resource tend to have a broader distribution of monthly wind speeds over their period of record. This broadness augments the range of the linear regression, which stabilizes the regression itself, and lowers its uncertainty.
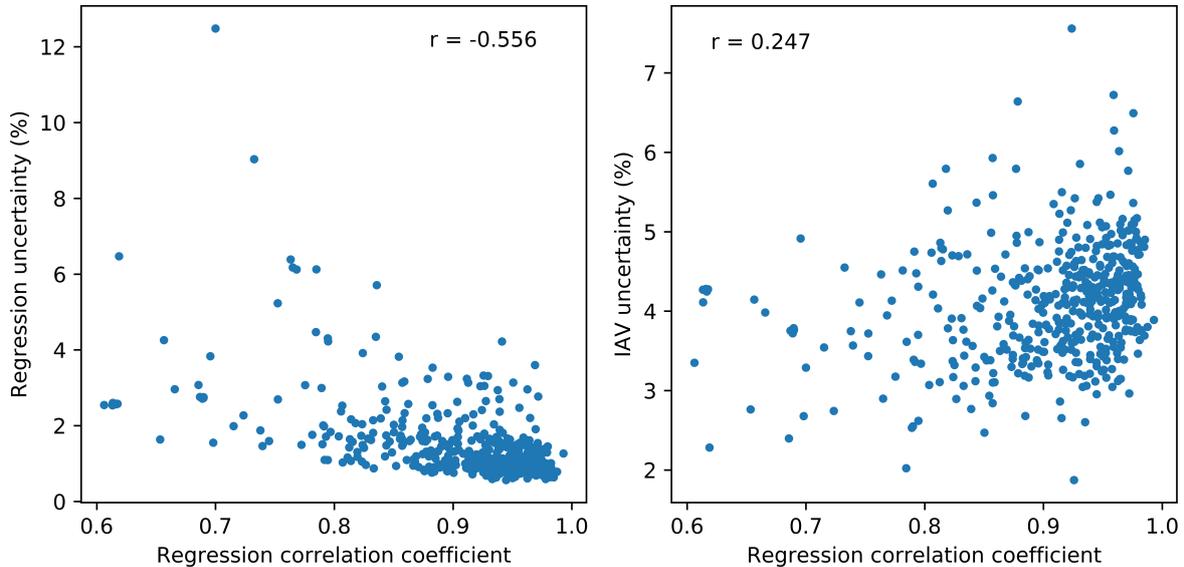
Figure 10. Dependence of regression uncertainty and IAV uncertainty on the R^2 of the regression between reanalysis wind speed and energy production data

4 Conclusions

Financial operations related to wind farms require accurate calculations of the annual energy production (AEP) and its uncertainty, both prior to the construction of the plant and in the context of its operational analysis. As the wind energy penetration keeps increasing globally, the need for accurate techniques to assess AEP uncertainty is a priority for the wind energy industry. Typically, the current industry standard assumes that the uncertainty components in AEP estimates are uncorrelated, and it combines them with a sum of squares approach. However, we have shown that this assumption is not valid on the EIA data set.

In this study, we investigated the assumption of uncorrelated uncertainty components by proposing a Monte Carlo approach to assess annual energy production. Our technique not only directly accounts for correlations between uncertainty categories, but also provides quantitative insight into aspects of the AEP process that drive this uncertainty. We have applied this approach using operational data from 472 wind farms across the United States in the EIA-923 database.

Our results show that assuming uncorrelated uncertainties determines a mean absolute percentage difference of 6% compared to the uncertainty calculated with the Monte-Carlo-based approach, with larger deviations (up to 20%) for specific sites. Moreover, three pairs of uncertainty components reveal a statistically significant correlation, which is neglected in the current industry standard: wind IAV and windiness (positive correlation), wind IAV and regression (negative), and wind measurement and regression (positive). Wind IAV and windiness uncertainties are correlated because they both depend on wind resource variability. Wind IAV uncertainty is correlated with regression uncertainty because they are both inversely proportional to the number of data points in the period of record. Finally, measurement (reanalaysis) uncertainty and regression uncertainty show
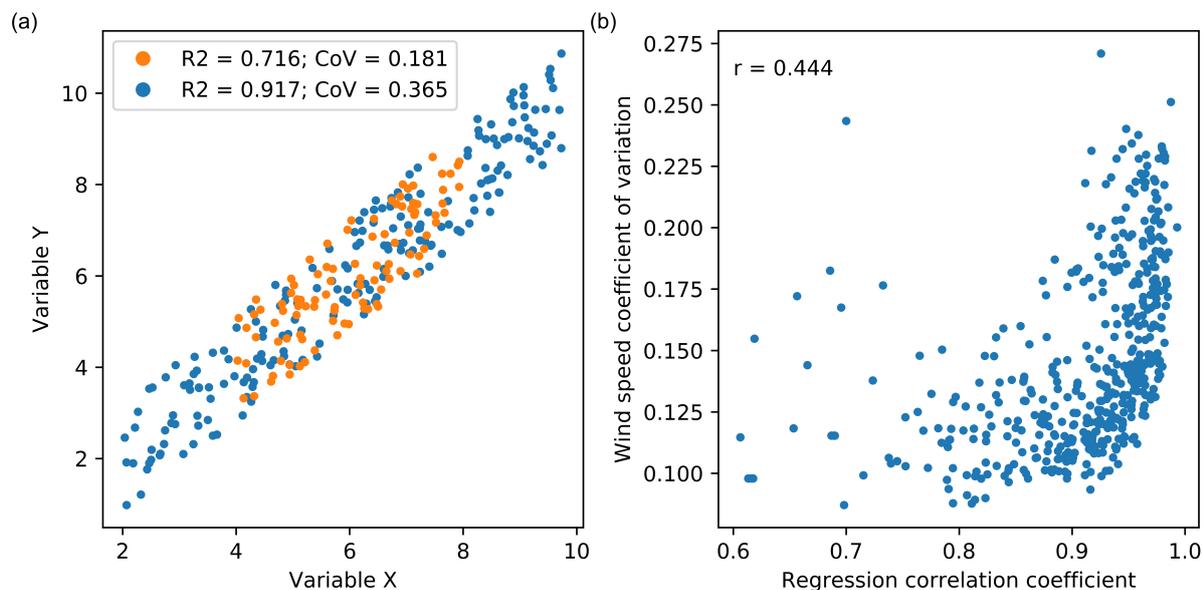
Figure 11. (a) Scatterplot of two ideal variables with equal spread, but different data ranges, and impact on the correspondent R^2 and coefficient of variation. (b) Dependence of the coefficient of variation of MERRA-2 wind speeds on the R^2 of the regression between reanalysis wind speed and energy production data.

a negative correlation because they respond oppositely to the R^2 coefficient between the reanalysis wind speed and energy production data. Therefore, our results suggest that a Monte Carlo approach should be preferred to take into account these correlations between uncertainty components to lead to more accurate results, compared to the current industry standard approach.

250 To facilitate the transition towards this new industry standard, NREL's open-source OpenOA software [4] already supports the recommended Monte Carlo approach to assess AEP. In addition, the benefit of this technique will be further described in a guideline document in preparation for publication by the AWEA TR-1 working group.

Additional categories of uncertainty in an operational AEP were not considered in our study because of limited reporting in the EIA-923 database. These categories include reported availability, curtailment uncertainty, and various uncertainties

255 introduced through analyst decision-making (e.g., filtering high-loss months from analysis and regression outlier detection). Future studies could include the impact of these additional sources of uncertainty on the operational AEP assessment. Finally, this study focused on correlations between operational AEP uncertainty categories. Future work could explore correlations between preconstruction AEP categories. Given the numerous categories (e.g., wake loss, wind speed extrapolation, wind flow model) and their intercomplexities, a Monte Carlo approach could reveal correlations that are at present not considered.

[4]https://github.com/NREL/OpenOA

Code and data availability. EIA data used in this study are accessible from https://www.eia.gov/electricity/data/eia923/. Geographical data of the EIA wind farms are available at https://www.eia.gov/maps/layer_info-m.php. Software used to assess operational AEP is available from https://github.com/NREL/OpenOA.

Author contributions. NB and MO are equal contributors to this work. MO performed the AEP estimates on the wind farms considered in the study. NB and MO analyzed the processed data. NB wrote the manuscript, with significant contributions by MO.

Competing interests. The authors declare that they have no conflict of interest.

References

Brower, M.: Wind resource assessment: a practical guide to developing a wind project, John Wiley & Sons, Hoboken, New Jersey, https://doi.org/10.1002/9781118249864, 2012.

Cameron, J.: Post-construction Yield Analysis, http://www.ewea.org/events/workshops/wp-content/uploads/proceedings/Analysis_of_

270   Operating_Wind_farms/EWEA%20Workshop%20Lyon%20-%202-3%20Jessica%20Cameron%20Natural%20Power.pdf, European Wind Energy Association Technical Workshop, 2012.

Clifton, A., Smith, A., and Field, M.: Wind Plant Preconstruction Energy Estimates: Current Practice and Opportunities, Tech. rep., https://www.nrel.gov/docs/fy16osti/64735.pdf, 2016.

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, d. P., et al.: The

275   ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the royal meteorological society, 137, 553–597, 2011.

EIA: A Guide to EIA Electric Power Data, Standard, Energy Information Administration, https://www.eia.gov/electricity/data/guide/pdf/guide.pdf, 2018.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R.,

280   et al.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, 2017.

Global Wind Energy Council: Global Wind Report - Annual Market Update 2017, Tech. rep., Global Wind Energy Council, 2018.

IEC 61400-12-1:2017: Wind energy generation systems - Part 12-1: Power performance measurements of electricity producing wind turbines, Standard, International Electrotechnical Commission, 2017.

285   IEC 61400-15:draft: Assessment of site specific wind conditions for wind power stations, Standard, International Electrotechnical Commission, in draft.

IEC 61400-26-3:2016: Wind energy generation systems - Part 26-3: Availability for wind power stations, Standard, International Electrotechnical Commission, 2016.

Kalkan, A.: Uncertainty in Wind Energy Assessment, http://www.windsim.com/documentation/UM2015/1506_WindSim_UM_Inores_

290   Akgun_Kalkan.pdf.

Khatab, A. M.: Performance Analysis of Operating Wind Farms, Master's thesis, Uppsala University, Department of Earth Sciences, Campus Gotland, 2017.

Lackner, M., Rogers, A., and Manwell, J.: Uncertainty Analysis in Wind Resource Assessment and Wind Energy Production Estimation, https://doi.org/10.2514/6.2007-1222, https://arc.aiaa.org/doi/abs/10.2514/6.2007-1222.

295   Lindvall, J., Hansson, J., Undheim, O., and Vindteknikk, J.: Post-construction production assessment of wind farms, Tech. Rep. 2016:297, Energyforsk, 2016.

Lunacek, M., Fields, M. J., Craig, A., Lee, J. C. Y., Meissner, J., Philips, C., Sheng, S., and King, R.: Understanding Biases in Pre-Construction Estimates, Journal of Physics: Conference Series, 1037, 062 009, http://stacks.iop.org/1742-6596/1037/i=6/a=062009, 2018.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., et al.: The NCEP climate

300   forecast system version 2, Journal of Climate, 27, 2185–2208, 2014.

Vaisala: Reducing Uncertainty in Wind Project Energy Estimates, Tech. rep., https://www.vaisala.com/sites/default/files/documents/Triton-DNV-White-Paper.pdf, 2014.