https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Biases in preconstruction estimates of wind plant annul energy production

Rob Hammond and Eric Simley

National Wind Technology Center, National Renewable Energy Laboratory, Golden, CO 80401, USA

Correspondence: Rob Hammond (rob.hammond@nrel.gov)

Abstract. Estimating the energy yield of a wind plant during the preconstruction phase is an historically difficult task, even with industry improvements in these estimations. We build on prior research comparing the realized energy production of wind plants and their estimated annual energy production (AEP) P50 values (median energy production), using owner-provided energy production and losses. Similar to prior studies, we found a slightly increasing bias of overestimating median energy production (6.9 % to 6.5 % as opposed to 6.7 % to 5.5 %). In addition to assessing AEP P50 bias, we compared both the 1-year and long-term AEP P90 and uncertainty energy yield assessment (EYA) estimates to the observed long-term corrected energy production. We found that neither the EYA uncertainty nor the P90 are conservative enough compared to the observed distribution of prediction errors, suggesting significant room for improvement in the EYA process.

1 Introduction

The probability of exceeding varying thresholds of annual energy production (AEP) forms the basis of investment risk for wind power plants. AEP overestimation increases the risk of financial losses for both investors and owners (Clifton et al., 2016). Initial studies suggest that biases between wind power plant preconstruction energy yield assessment (EYA) estimates and energy production ranged between -6.7 % and -5.5 % (where the negative sign indicates overestimation), which is assumed to improve by 1–2 % each if accounting for unreported curtailment losses (Lunacek et al., 2018). More recent studies suggest that this bias is decreasing over time in the wind industry, with preconstruction estimates overpredicting energy yield by only 1–2 % (Lee and Fields, 2021). However, the uncertainty in EYA prediction accuracy remains high. To better understand EYA biases and characterize uncertainty in the EYA process we build upon the methods and findings of Lunacek et al. (2018).

In Lunacek et al. (2018), the authors relied on consultant-provided EYA data combined with publicly available net energy production from the U.S. Energy Information Agency (EIA). This rendered an analysis of 62 U.S.-based wind plants, primarily located in the southern and central United States, with commercial operation dates (COD) between 2008 and 2016 (23 of which

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

WIND ENERGY SCIENCE DISCUSSIONS

are post-2010). However, in this study, we compared the monthly gross energy production and the curtailment and availability losses obtained from wind plant owners and operators with the consultant-provided preconstruction EYA estimates for the long-term AEP. The AEP is estimated in terms of its PXX amount, such as P50 or P90, where XX represents the 100-XX percentile of the data, or the XX% chance of exceeding that level of energy production. For the P50 and P90, this corresponds to the 50th percentile, or median energy production, and the 10th percentile of energy production, respectively. The uncertainty of the AEP is also considered from a short-term (1-year variability) and long-term (variability in average AEP over 10+ or 20 years) perspective. In this study, we will consider the P50, short- and long-term AEP P90, and short- and long-term AEP uncertainty. Similar to (Lunacek et al., 2018), we will also focus on projects primarily in the United States and those with a COD later than 2010, when the the industry improved existing methods and data quality control for wind energy resource assessments used for EYA reporting (Dickinson et al., 2014). There are several contributions this work make to the literature. We improve on the data in Lunacek et al. (2018) by using owner-reported energy production, curtailment loss, and availability loss data, and updating the analysis for an additional seven years of data. Further, we improve on the methodology in in Lunacek et al. (2018) by accounting for the availability and curtailment losses in the long-term AEP estimation. Finally, we also benchmark the pre-construction AEP uncertainty estimates against the spread of AEP prediction errors.

This paper is organized as follows: The next section will cover the data collection and analysis methodology. The following section will provide a breakdown of the results, including comparing data from Lunacek et al. (2018) to this study and comparing operational AEP to EYA P50, P90, and uncertainty estimates. Finally, we will conclude with a discussion of the results and future directions for this research.

2 Methodology

The P50 and P90 estimates represent the AEP at which the project is expected to exceed 50 % or 90 % of the time, respectively, over the lifetime of the wind plant. In addition to the median and 10th percentiles of energy production, we also considered the uncertainty, which is modeled as the standard deviation divided by the mean energy production. For the P90 and uncertainty estimations we compared both the short-term and long-term values. The short-term estimates are considered to have a 1-year variability. The long-term estimates were provided as either a 10-year, 10 or more year, or 20-year variability, so our long-term comparisons will be a mixture of these values. For EYAs with multiple values, we chose the 20-year estimate to align with the analysis period. We also compared the net energy production from the monthly operating reports (MOR) to the results in Lunacek et al. (2018) using EIA monthly energy production data for wind plants in the United States.

2.1 Wind plants

The authors and their colleagues collected preconstruction EYAs or key data from them over the course of multiple years. To build on the work of Lunacek et al. (2018), we utilized with many of the same wind plant owners and operators who provided EYA estimates to collect monthly gross energy production, curtailment losses, and availability losses. Our data collection spanned MORs for 94 wind plants in 6 countries, covering 115 EYAs for 6 wind plant owners and 19 specified consultants. Of

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

those, 76 are in the United States (72 with corresponding EIA data), and 70 have a COD of 2011 or later – a reversed COD bias from Lunacek et al. (2018) where only 23 of the 56 plants had a COD of 2011 or greater. In total, we collected 707 complete wind plant years of data to conduct this study.

Figure 1 shows the distribution of key project characteristics provided in the EYAs for projects that also have MORs. Not all projects had metadata provided, and so there are fewer results for some variables. There are a handful of projects with CODs prior to 2010, but most plants came online between 2012 and 2017. Additionally, the provided project MORs tended to be 3.5 years or between 8 and 10 years of operational data as opposed to 2 years or less, providing sufficient data for modeling.

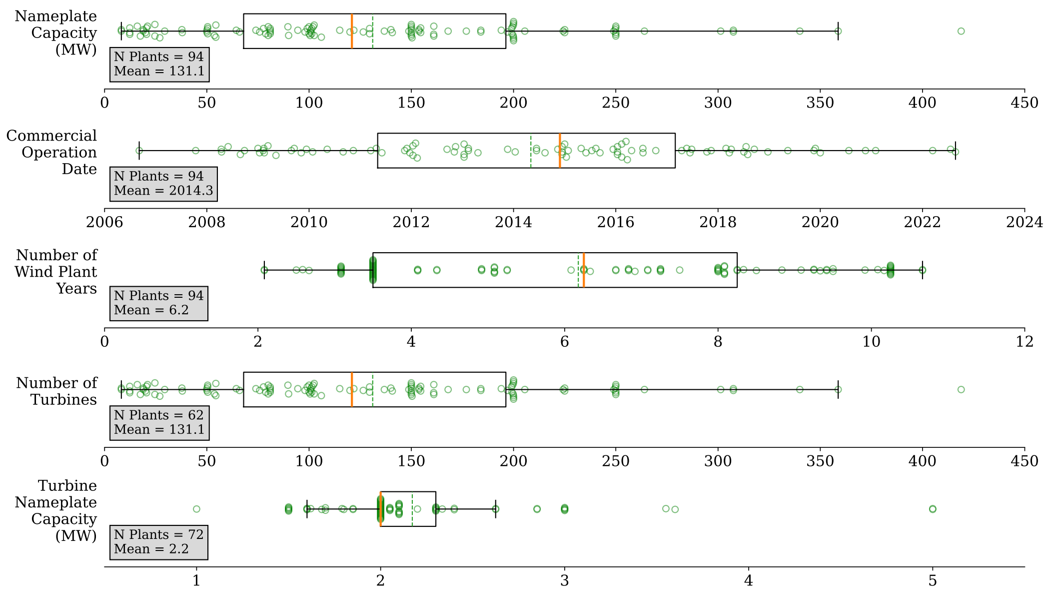

Figure 1. Summary of the key characteristics of the wind plants that have both an EYA and monthly operating data. From top to bottom, the boxplots show the distribution of the project's nameplate capacity, COD, number of turbines, and turbine capacity, where that information is available in the EYA data.

For each wind plant, we also collected the ERA-5 (Hersbach et al., 2018) and MERRA-2 (Global Modeling and Assimilation Office (GMAO), 2015) reanalysis datasets for the long-term wind resource corrections. The MERRA-2 data are made available at 0.125° resolution, and the ERA-5 data are available at 0.25° resolution, so we identified the closest latitude and longitude to each plant's owner-provided centroid and used the atmospheric data corresponding to that point. For each point, we downloaded the hourly wind speed, wind direction, temperature, humidity, and surface air pressure and aggregated these readings to the monthly mean reading for each variable. Figure 2 demonstrates the high degree of correlation between gross energy production (discussed further in Sect. 2.2) and each of the MERRA-2 and ERA-5 wind speed data as measured by the Pearson correlation coefficient. While ERA-5 demonstrates a higher average Pearson correlation coefficient (0.81) than the MERRA-2 data (0.79), both exhibit a median value of 0.89. The close alignment of reanalysis wind speed and gross energy suggest that for most

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

projects, the reanalysis wind conditions will be highly suitable for analysis. We removed the five projects with a correlation coefficient of less than 0.4 from the operational analysis.

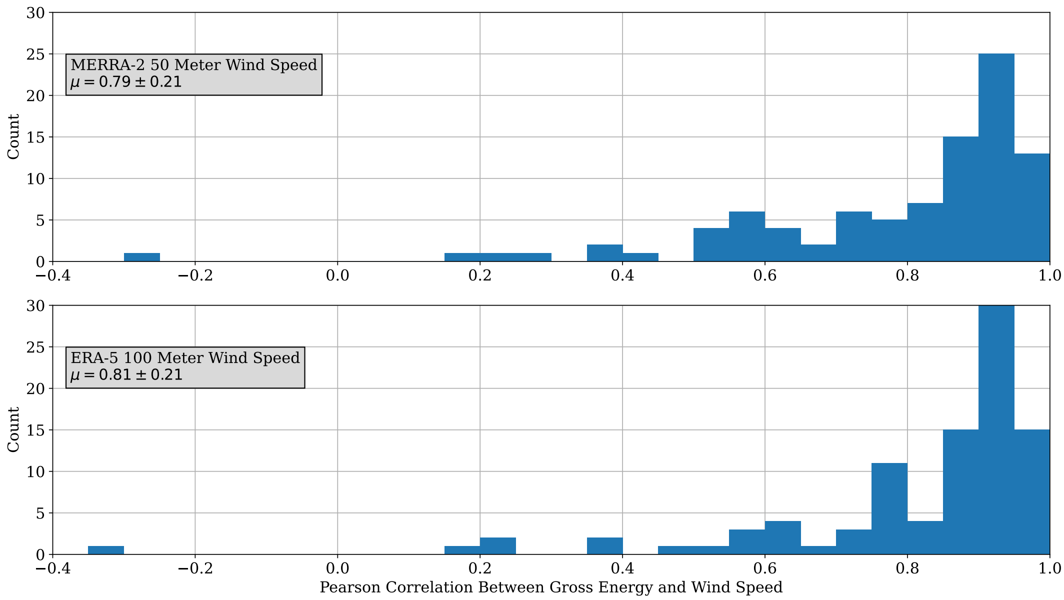

Figure 2. Histogram of the correlation between the monthly gross energy production and each of the monthly MERRA-2 50 m wind speed (top) and ERA-5 100 m wind speed (bottom) for each wind plant for the nearest available coordinate.

2.2 Operational analysis

The first set of analyses assessed the similarities between the P50 biases using the annual net energy provided through the EIA and MORs. To annualize the energy production, we summed each consecutive 12-month period starting from the COD, excluding any partial years at the end of the time series. We then calculated the percentage difference between each annualized reading and the EYA estimated P50.

To compute the long-term corrected P50, P90, and uncertainty estimates, we utilized the Monte Carlo AEP analysis method in NREL's OpenOA software (Perr-Sauer et al., 2021) as described in Bodini and Optis (2020), outlined below, and shown as a flowchart in Fig. 3. The method includes the following steps:

1. Compute the density-corrected wind speeds for each reanalysis product, and average the hourly wind speed, wind direction, and temperature to a monthly level to align with the operating data.

2. Flag months where the losses were unreported and remove months where no energy data were available.

3. Calculate gross energy by adding the revenue meter energy, availability loss, and curtailment loss, and normalize each month to a 30 day period.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

4. Run a linear regression between the gross energy production and density-corrected wind speeds, the sine and cosine components of the wind direction, and temperature.

5. Apply the linear regression to the long-term, density-corrected wind speed, wind direction, and temperature data.

6. Correct the long-term gross energy estimates, which are based on 30-day months for the correct number of days in each month.

7. Subtract the monthly average availability losses to estimate monthly net energy

8. For analyses comparing the 1-year uncertainty, include the interannual variability (IAV) uncertainty factor, and for analyses compared to the long-term uncertainty, exclude the IAV. The IAV uncertainty component is sampled from a normally distributed random variable, centered on the 20-year standard deviation of the energy production estimates.

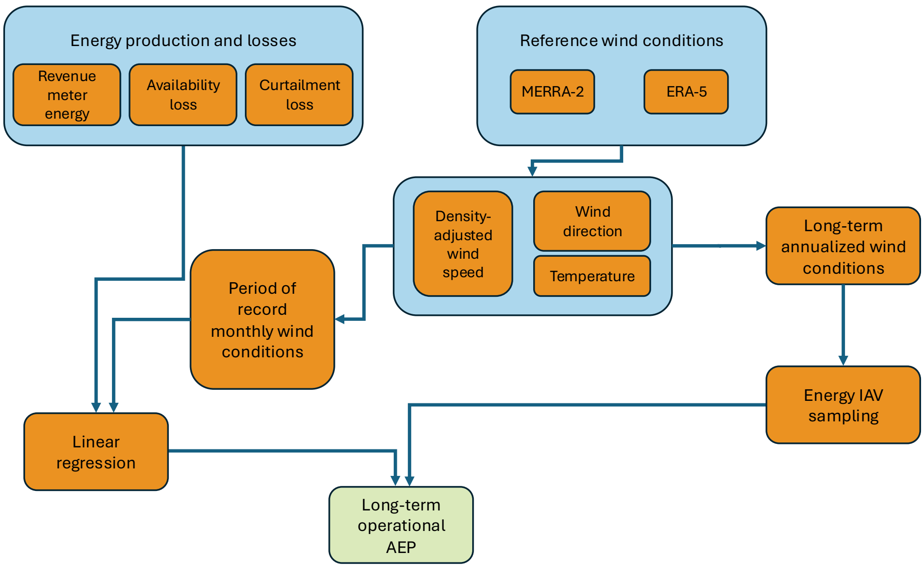

Figure 3. Flowchart of the OpenOA Monte Carlo AEP methodology.

The long-term correction is similar to the approach taken in Lunacek et al. (2018), with a few essential differences in the data used, and how it is filtered. Both the reanalysis data were of a different vintage, and only the wind speed data were used in Lunacek et al. (2018). No curtailment or availability data are provided through the EIA, so there was no gross energy calculation, nor was a 30-day normalization applied in Lunacek et al. (2018). The outlier filtering in Lunacek et al. (2018) was based on the standard deviation from the mean, rather than the more robust approach taken in this study.

For each iteration in the Monte Carlo analysis, we randomized the usage of the following variables: the reanalysis product used for the long-term correction, uncertainty percentage applied to the net energy (normal distribution centered on 0.5 %), uncertainty percentage applied to the long-term losses (normal distribution centered on 5 %), number of years to use for the windiness correction (uniformly distributed between 10 and 20 years), maximum combined availability and curtailment losses as a percentage of energy production to identify and remove months that have higher than expected energy losses (defaults to uniformly distributed between 10 % and 20 %), and the tuning constant for the Huber loss function for a robust linear regression to identify and remove outliers. We used the Huber loss method, as it is more robust to outlier data, such as high curtailment or low availability months, that are not representative of regular plant operations. For each combination of model configurations, we ran 5,000 Monte Carlo iterations, including bootstrapping the monthly operational data and randomly sampling different analysis parameters to understand the uncertainty of the operational assessments for each plant.

We ran the following suite of analysis combinations to identify the analysis settings that yield the most accurate and reliable estimates:

- Use of temperature data
- Use of wind direction data
- Use of the Huber loss algorithm to identify and filter out any outliers
- The maximum combined curtailment and availability loss threshold for a month to be included in the analysis
  - Model defaults: sampled between 10 % and 20 %
  - Between the 60th percentile and the 80th, 90th, or 100th percentile of actual losses
  - Between the 80th percentile and the 90th or 100th percentile of the actual losses.

To select the representative model for each of the combinations, we chose the lowest total rank by comparing the R^2 of each bootstrapped model and the total uncertainty (σ_{AEP}/μ_{AEP}). Using each of these metrics, we summed the ranking of the lowest total uncertainty and maximum median R^2. We selected the lowest total rank model as the representative model for each plant as it would have minimized uncertainty and maximized how well the model fits the test data. Using the 5,000 iterations, we computed the P50, P90, and uncertainty of the AEP estimations.

When computing the results, we also gradually filtered out time periods and projects to arrive at a more accurate assessment of each wind plant and the industry more broadly. First, we removed the first year of operational data to account for non-representative issues that arise in the first year of operations. Then, we removed projects with a COD earlier than 2011 to account for changes in EYA methodology. Finally, we considered the combination of both filters.

3 Results

The results of our analyses indicate that a significant P50 bias still exists for wind plant EYA methods despite industry improvements for estimating energy production. Additionally, we note that both the short- and long-term EYA uncertainties are underestimated relative to the observed spread in AEP P50 prediction errors and, similarly, that the P90 estimates are not conservative enough.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

3.1 Validation of the original approach

As a first step, we confirmed the validity of the approach to rely on monthly EIA net energy data to determine AEP from the original study (Lunacek et al., 2018). To do this, we compared the EIA net energy output to the MOR net energy where the EIA identifier was provided with the EYA data. As seen in Fig. 4, the correlation between the MOR energy production received from plant owners and operators and the energy reported by the EIA was strong ($\rho = 0.98$). The line of best fit for the EIA and MOR data yields $y = 0.99. \times \ x + 0.26$, indicating that the EIA data are highly consistent with those reported by wind plant owners. This result indicates that the methods used by Lunacek et al. (2018) to set a baseline for P50 bias remain valid overall, though significant errors exist for individual months.

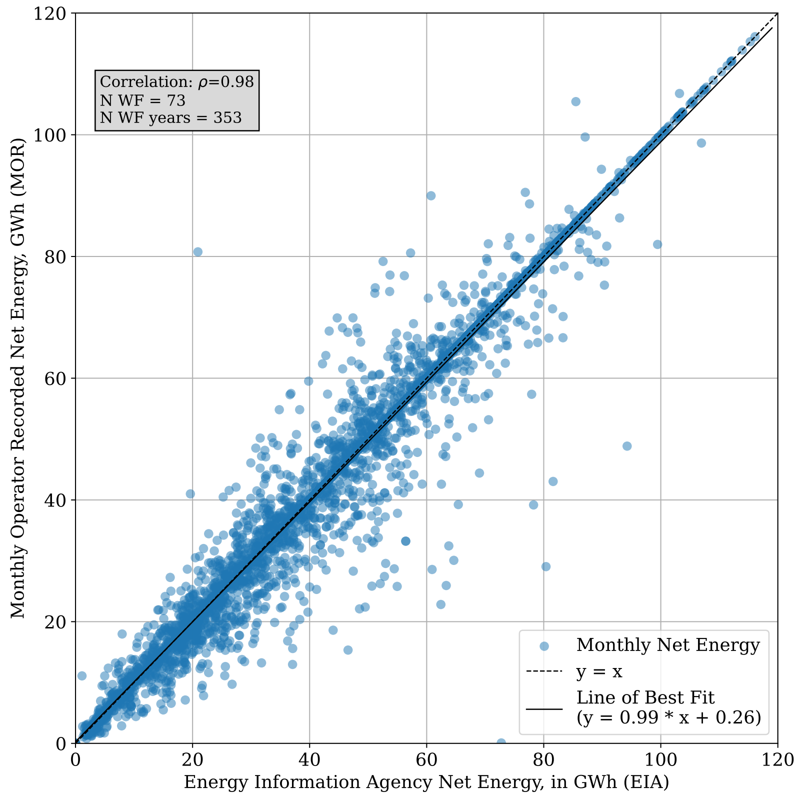

Figure 4. Relationship between EIA reported monthly net energy and MOR net energy. The line of best fit in this demonstrates that the EIA net energy closely tracks that of the MORs.

We then compared the AEP P50 bias between the EYA P50 estimates and both the net EIA energy and the net MOR energy without a long-term correction, and the bias resulting from Lunacek et al. (2018), as seen in Fig. 5. Here, we took the best-case scenario where we removed the first operational year of the wind plant and removed plants with a COD prior to 2011. This approach is similar to the approach taken in Lunacek et al. (2018), which reported a 5.5 % overprediction when using a long-

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

term correction. Here, the average P50 biases for the uncorrected EIA and MOR data are -14.9 % and -11.5 %, respectively. This finding indicates that there is still an overprediction of AEP P50 in the EYAs, given it is at least double the long-term corrected value reported in (Lunacek et al., 2018).

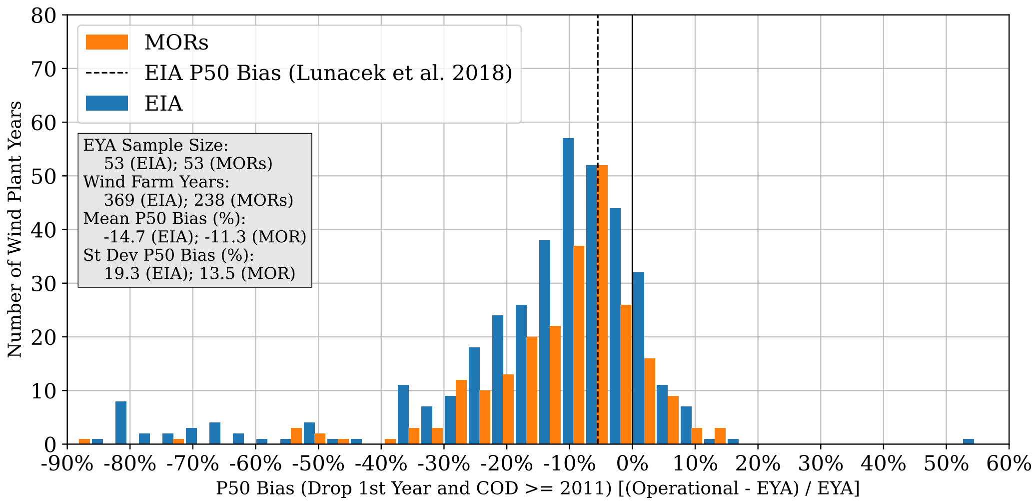

Figure 5. Comparison of the P50 bias between the EIA net energy and MOR net energy.

3.2 Long-term corrected results

The next phase of the analysis was to estimate the long-term corrected AEP and uncertainty. As described in Sect. 2.2, we found each project's ideal analysis settings; we report on the long-term corrected results in this section. We found that 84 % of the best combinations included the use of the Huber loss algorithm outlier detection and filtering, 74 % used the reanalysis temperature data, and 33 % used the reanalysis wind direction. For the uncertainty loss thresholds, we found that 21 % of the best combinations used the OpenOA default value, and another 46 % used thresholds with a minimum of the 80th percentile. Combined, 47 % of the best model combinations also used the outlier detection, temperature data, and no wind direction data.

In Fig. 6 we demonstrate the overall long-term corrected P50 bias by COD where there is modest yet inconsistent improvement in the P50 bias between 2011 and 2019 relative to projects prior to 2011. However, in 2019 and later, the P50 bias worsens, though it is worth noting that there are fewer projects and therefore less operational data in our dataset in this COD grouping, so the results may not be representative. Even with later COD P50 bias increases, the post-2010 COD wind plants exhibit an improving P50 bias, -6.9 % vs. -6.6 %, as seen in the text box of Fig. 6, which is in-line with the baseline results trends from (Lunacek et al., 2018).

We also demonstrate the P50 bias by consultant in Fig. 7 to determine if there were any significant variations across the 19 consultants. Most deviation from the average P50 bias occurred in consultants with a small sample size or where the spread was large. In particular, Consultants 1, 3, 6, and 8 had a large spread in their individual estimates, with only Consultant 5

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

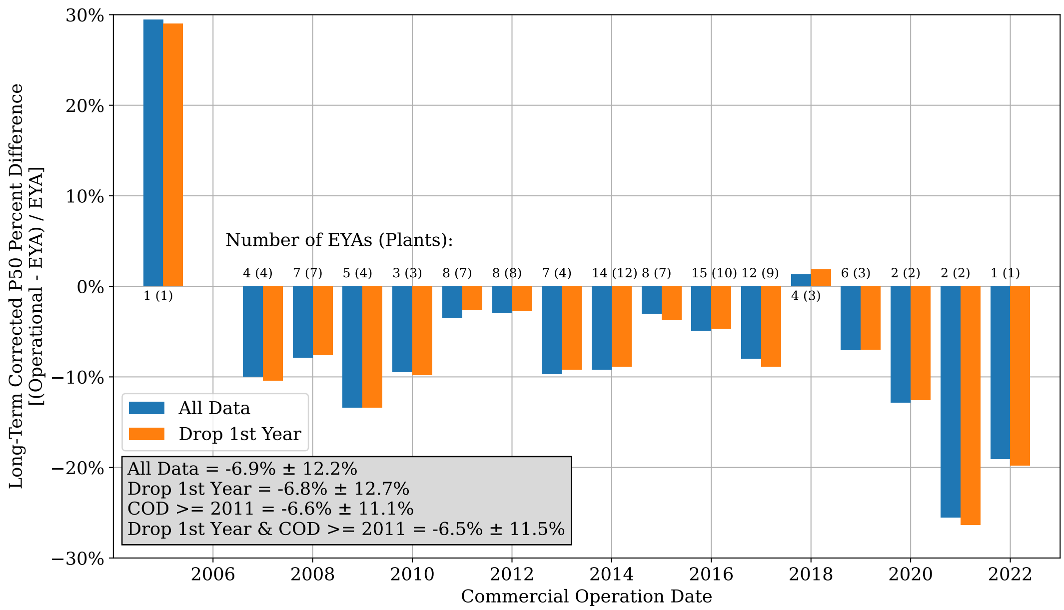

Figure 6. Histogram of the P50 bias by COD for all projects, and all projects when removing their first year of operations.

demonstrating a high level of precision in their estimates. The unspecified grouping, corresponding to no consultant provided, however, yielded the largest sample size with some of the smallest spread in EYA P50 bias.

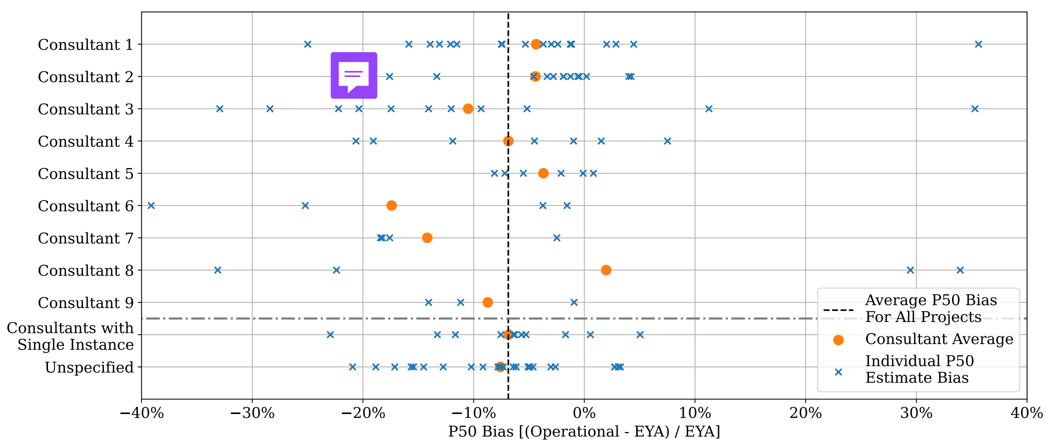

Figure 7. Scatter plot of each consultant's individual P50 estimate biases, including for estimates without a specified consultant.

Using the 5,000 simulations, we then computed the P50 and P90 as their respective 50th and 10th percentiles and the uncertainty as σ_{AEP}/μ_{AEP}. The boxplots in Figs. 8, 9, and 10 demonstrate the distribution of the per-project spread of the P50, long-term P90, and long-term uncertainty bias, respectively, for each of our four down-selection criteria. Compared to the

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

(Lunacek et al., 2018) results, we can see a slightly worsening P50 bias. We observe that the short-term P90 bias is positive, indicating that the P90 of the estimated operational AEP exceeds the EYA AEP P90 value on average. However, this result does not necessarily suggest that the EYAs use overly conservative means for estimating high-probability energy production levels; to reach this conclusion, we would need to demonstrate that the estimated operational AEP exceeds the EYA P90 predictions 90 % of the time.

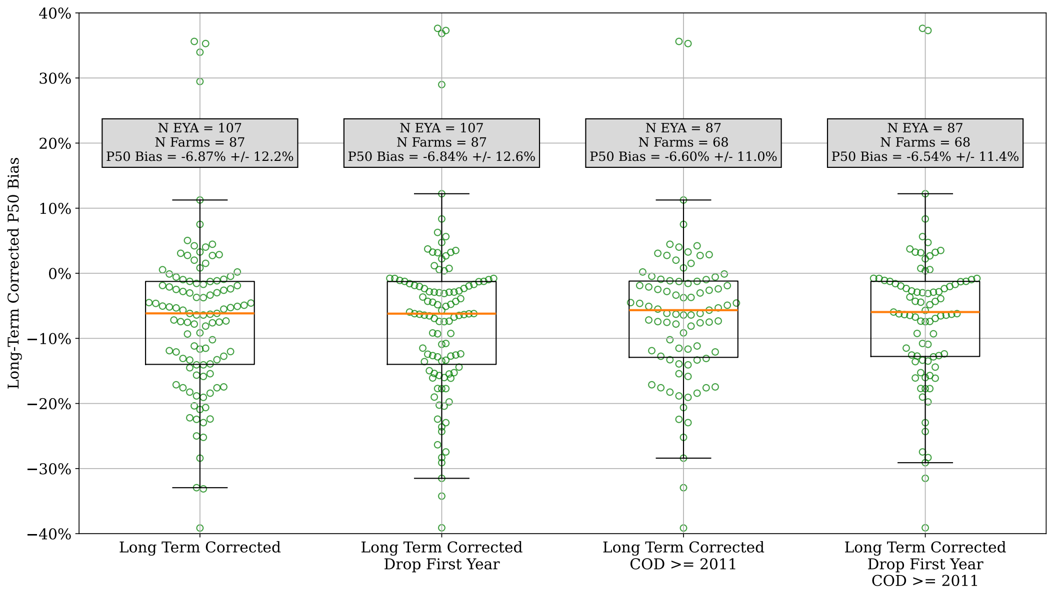

Figure 8. Boxplot showing the project-level P50 bias for each of the project down-selection criteria.

As indicated, the uncertainty bias for the long-term corrected energy estimates relative to the 1-year EYA uncertainties appears to suggest that the EYAs are routinely conservative estimates of uncertainty with the average bias in the -52 % to -48 % range. It should be noted, though, that the uncertainties for the operational AEP estimates and the EYA AEP predictions are not defined equivalently. The operational AEP uncertainties represent the uncertainty in the long-term AEP based on realized energy production during an initial period of operation. EYA estimates, on the other hand, represent the uncertainty in the predicted AEP before the plant is built. The EYA estimates are expected to have a higher degree of uncertainty because there are a significant number of individual model uncertainties and potential sources of loss to consider when predicting energy production prior to building a wind plant.

To better understand the uncertainty and prediction errors, we also computed the prediction error of the EYA P50 and P90 estimates for each plant relative to each Monte Carlo iteration's AEP estimation for the corresponding plant (percentage difference between the EYA reading and each iteration's AEP estimate), as shown in Figs. 11–13. In Fig. 11, which shows the distribution of P50 prediction errors including IAV, we smooth the error histogram to highlight the centering around the mean values. Importantly, we found that the standard deviation of the sample errors ranged from 12 % to 13.6 %, which closely

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

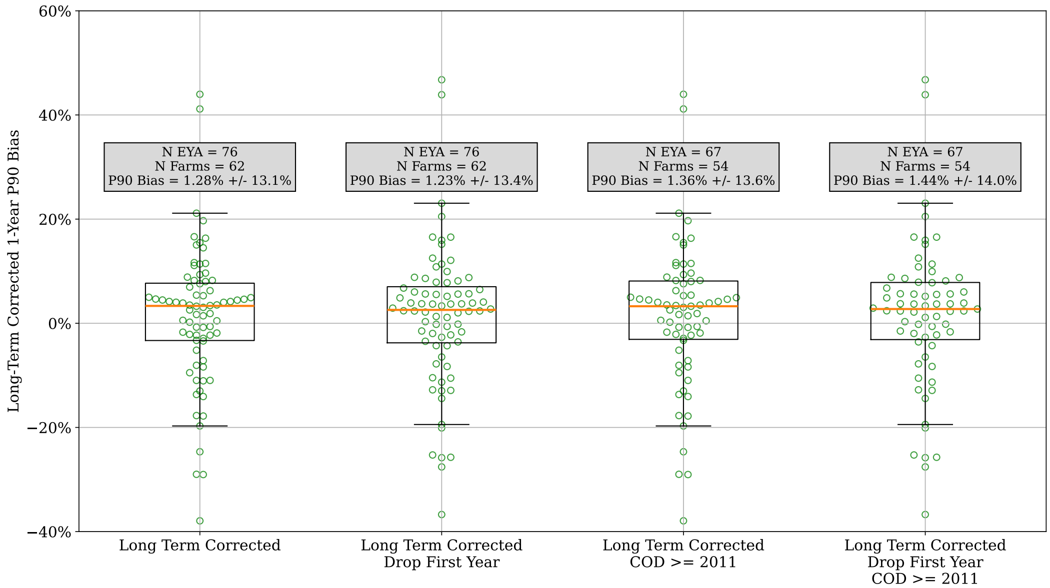

Figure 9. Boxplot showing the project-level, long-term P90 bias for each of the project down-selection criteria.

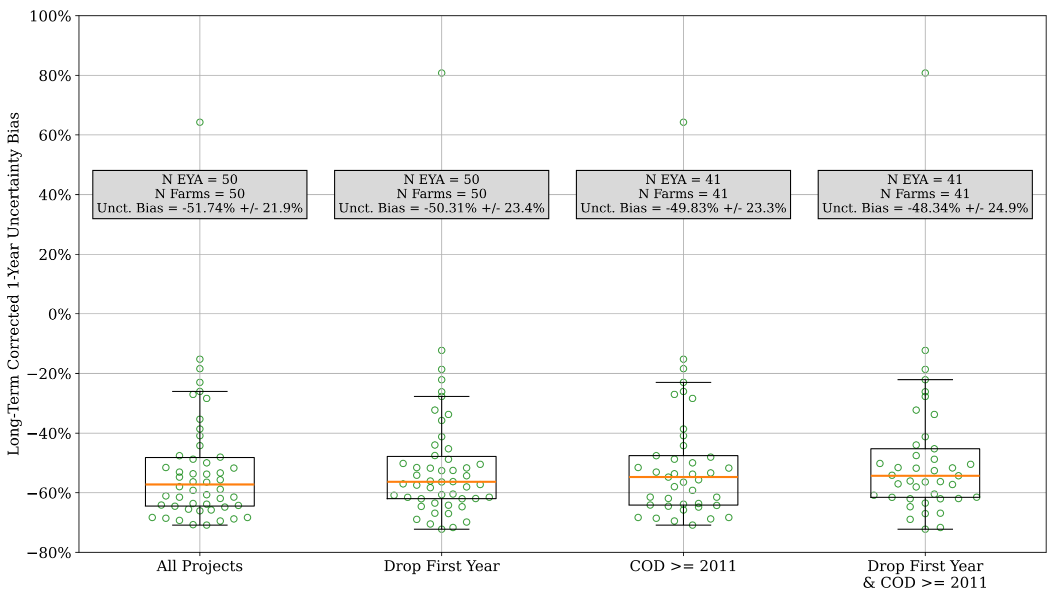

Figure 10. Boxplot showing the project-level, long-term corrected uncertainty biases relative to their 1-year EYA uncertainty estimate for each of the project down-selection criteria.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

aligns with the average 1-year uncertainty in the corresponding EYAs of $10.5\ \% \pm 2.6\ \%$, shown on the right side of Fig. 11. However, when comparing the EYA long-term uncertainty to the P50 prediction errors that exclude IAV, as shown in Fig. 12, our standard deviation ranges from 11.4 % to 12.9 %, which is much larger than the corresponding EYA long-term uncertainties of $7.4\ \% \pm 2.1\ \%$. As opposed to the direct comparison in Fig. 10, these results indicate that the EYA uncertainty, especially the long-term uncertainty, is being underestimated.

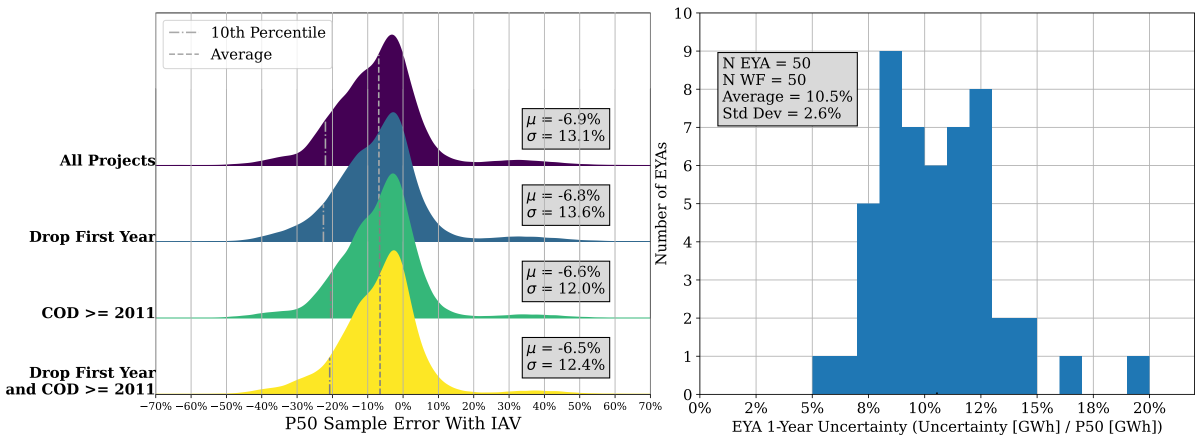

Figure 11. (left) Ridge plot showing the percentage error of each sample's AEP value calculated with IAV relative to EYA P50. (right) Histogram of the EYA 1-year uncertainty values.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

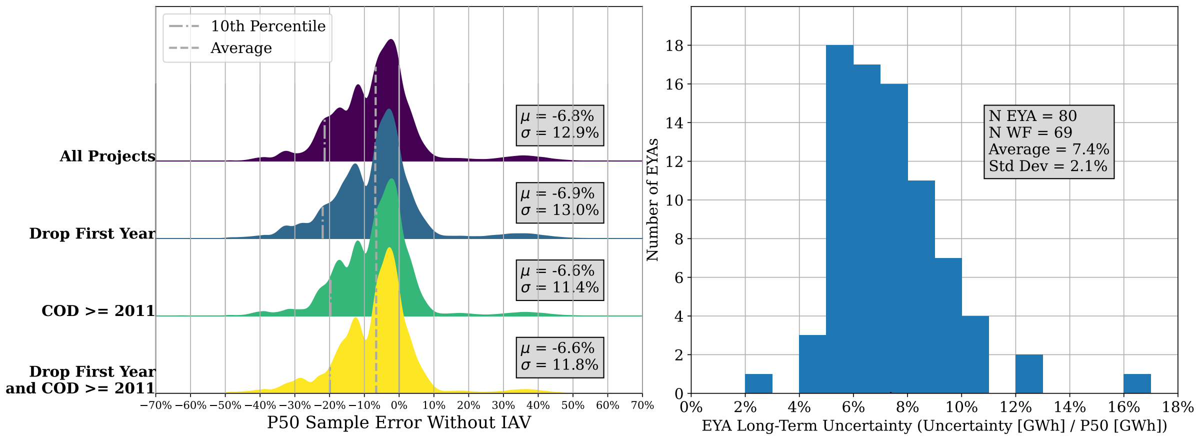

Figure 12. (left) Ridge plot showing the percentage error of each sample's AEP value calculated without IAV relative to EYA P50. (right) Histogram of the EYA long-term uncertainty values.

In line with our project-level P90 estimate comparisons (though more positive), the individual P90 prediction errors for each Monte Carlo iteration center around positive values (Fig. 13); however, the 10th percentile of the prediction errors is roughly -10 % for all cases, demonstrating that EYA P90 estimates are not conservative enough for both the 1-year and long-term P90 predictions. Ideally, the 10th percentile should be 0 %, meaning the realized operational AEP exceeds the EYA P90 predictions at least 90 % of the time. Note that we can still see the impact of outlier data in the spread of the distribution and the presence of additional peaks.

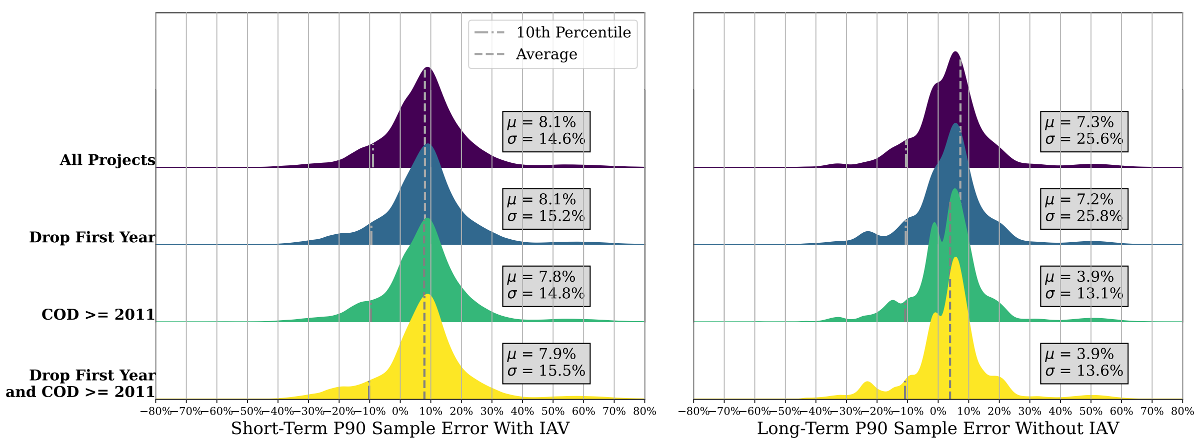

Figure 13. Ridge plots showing the percentage error of each sample's AEP value relative to the EYA P90. (left) The comparison is made between the analyses run with IAV and the EYA short-term P90. (right) The comparison is made between the analyses run without IAV and the EYA long-term P90.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

WIND ENERGY SCIENCE DISCUSSIONS

Finally, because EYA availability and curtailment loss estimates were provided for 36 projects, we also computed the P50 value for each of the long-term estimated operational availability and curtailment losses in the Monte Carlo simulations by project, as shown in Fig. 14. We have excluded a detailed comparison between the EYA estimated curtailment losses and operational curtailment losses, as 31 of the EYAs provided an EYA-estimated 0 % curtailment loss, and only two EYAs estimated 0 % availability losses, which were removed from the comparison. As such, the results of the availability comparison are shown in Fig. 14. We found that the average difference between the operational P50 values and EYA estimates for availability losses was 0.8 percentage points, indicating a slight underestimation of availability loss; however, the median value is 0.07 percentage points, and the standard deviation is 3.6 percentage points. Taken together, this suggests an inconsistent understanding of availability across the board, even if it contributes a minimal amount to the AEP P50 biases. Further, it should be noted that availability data were provided in mixed formats – energy lost and time-based availability – so the comparisons are not wholly on the same basis.

The curtailment P50 bias only consists of seven comparisons, so the underestimation of losses by 13 ± 23 percentage points may not be representative of reality. In fact, many of the EYAs with curtailment loss information state outright that situations where the energy grid may force the power plant to stop sending power are not considered, so the comparison of realized curtailment and EYA estimated curtailment likely provides little insight without further details or delineations.

4 Conclusions

In this paper, we built on the methodology of the work performed in (Lunacek et al., 2018) by gathering operational data from wind plant owners and operators that include the curtailment and availability losses that could not be accurately accounted for in the original work. We observed many of the same relationships as the original study, such as an overprediction of energy yield from consultants and a shrinking P50 bias following the change in methodology in 2011. However, we found that P50 bias has remained in line with, but slightly worse than, (Lunacek et al., 2018) now that we include more recent projects and are able to account for curtailment and availability losses. Additionally, we observed that the EYA uncertainty and P90 estimates are not conservative enough when compared to the observed distributions of AEP prediction errors relative to the operational AEP estimates, although the EYA 1-year AEP uncertainties are a reasonable approximation of the observed uncertainty.

This work motivates additional research to understand how the EYA process can be improved to reduce the AEP P50 overprediction bias and better characterize the uncertainty in the predictions, especially because P50 bias appears to be increasing compared to the findings of HVS1. Our results ultimately raise further questions that were out of scope for this study. One example is to measure the impacts of external wakes from neighboring farms that came online after a plant's EYA was, and the extent to which they contribute to P50 overprediction bias. Many projects represented in this dataset are older and are likely to have had neighboring projects built since their inception, creating unaccounted losses in the EYAs that would be present in the MORs.

A more fundamental question would be whether working with higher-resolution operational data, turbine operating status, and consistently delineated loss data combined with more advanced long-term operational AEP modeling could provide a more

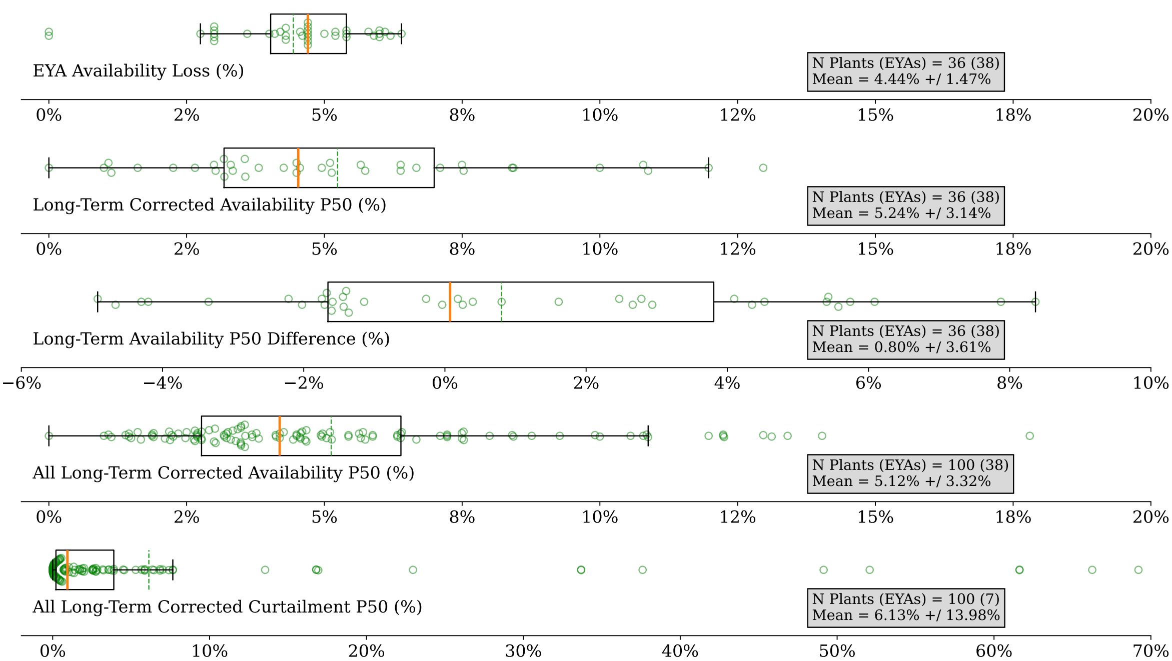

Figure 14. Summary of the the EYA estimated availability losses (mixture of time- and energy-based), long-term corrected P50 of operational availability and curtailment losses, and the percentage point difference between the operational availability P50 and EYA availability losses. From top to bottom, the boxplots show the distribution of EYA availability loss, the long-term corrected operational availability loss P50, the difference between the long-term corrected operational P50 and EYA estimated availability losses, the long-term corrected operational availability loss P50 for all projects, and the long-term corrected operational curtailment loss P50 for all projects.

accurate or less uncertain long-term correction. Similar to (Lunacek et al., 2018), it remains an open question of how to curate a broader selection of wind plant MORs, EYAs, and plant characteristics to understand the effects that plant characteristics such as geography or turbine technology can have on energy production.

Code availability. The OpenOA software is publicly available on GitHub at https://github.com/NREL/OpenOA. However, the analysis code is tied to NDA-protected data references and is not made public.

Author contributions. RH and ES met with wind plant owners to obtain EYA and operational data. RH performed the analyses and drafted the manuscript. ES advised on the work, and reviewed and edited the manuscript.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government.

Acknowledgements. The authors would like to thank the anonymous wind farm operators who provided the wind farm data that was used in this analysis. Hersbach, H. et al., (2018) was downloaded from the Copernicus Climate Change Service (2023).

Financial support. This work was authored by the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under contract no. DE-AC36-08GO28308. Funding was provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office.

https://doi.org/10.5194/wes-2025-127
Preprint. Discussion started: 8 September 2025

References

Bodini, N. and Optis, M.: Operational-based annual energy production uncertainty: are its components actually uncorrelated?, Wind Energy Science, 5, 1435–1448, 2020.

Clifton, A., Smith, A., and Fields, M.: Wind Plant Preconstruction Energy Estimates. Current Practice and Opportunities, https://doi.org/10.2172/1248798, 2016.

Dickinson, K., Monache, L. D., McCormack, K., and Magontier, P.: Wind Energy Resource Assessment : Information Production, Uses, and Value - Survey Report, https://sciencepolicy.colorado.edu/admin/publication_files/2014.60.pdf, 2014.

Global Modeling and Assimilation Office (GMAO): MERRA-2 tavg1_2d_slv_Nx: 2d,1-Hourly,Time-Averaged,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4, https://doi.org/10.5067/VJAFPLI1CSIV, 2015.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, https://doi.org/10.24381/cds.adbb2d47, 2018.

Lee, J. C. Y. and Fields, M. J.: An overview of wind-energy-production prediction bias, losses, and uncertainties, Wind Energy Science, 6, 311–365, https://doi.org/10.5194/wes-6-311-2021, 2021.

Lunacek, M., Jason Fields, M., Craig, A., Lee, J., Meissner, J., Philips, C., Sheng, S., and King, R.: Understanding Biases in Pre-Construction Estimates, Journal of Physics: Conference Series, 1037, https://doi.org/10.1088/1742-6596/1037/6/062009, publisher Copyright: © Published under licence by IOP Publishing Ltd.; 7th Science of Making Torque from Wind, TORQUE 2018 ; Conference date: 20-06-2018 Through 22-06-2018, 2018.

Perr-Sauer, J., Optis, M., Fields, J. M., Bodini, N., Lee, J. C., Todd, A., Simley, E., Hammond, R., Phillips, C., Lunacek, M., Kemper, T., Williams, L., Craig, A., Agarwal, N., Sheng, S., and Meissner, J.: OpenOA: An Open-Source Codebase For Operational Analysis of Wind Farms, Journal of Open Source Software, 6, 2171, https://doi.org/10.21105/joss.02171, 2021.