



Wind Plant Performance Prediction Benchmark Phase 1 Technical Report

M. Jason Fields,¹ Mike Optis,¹ Jordan Perr-Sauer,¹
Austin Todd,¹ Joseph C. Y. Lee,² John Meissner,³
Eric Simley,¹ Nicola Bodini,¹ Lindy Williams,¹
Shuangwen Sheng,¹ and Rob Hammond¹

1 National Renewable Energy Laboratory

2 Pacific Northwest Laboratory

3 MXV Ventures

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5000-78715
October 2021



Wind Plant Performance Prediction Benchmark Phase 1 Technical Report

M. Jason Fields,¹ Mike Optis,¹ Jordan Perr-Sauer,¹
Austin Todd,¹ Joseph C. Y. Lee,² John Meissner,³
Eric Simley,¹ Nicola Bodini,¹ Lindy Williams,¹
Shuangwen Sheng,¹ and Rob Hammond¹

1 National Renewable Energy Laboratory

2 Pacific Northwest Laboratory

3 MXV Ventures

Suggested Citation

Fields, M. Jason, Mike Optis, Jordan Perr-Sauer, Austin Todd, Joseph C. Y. Lee, John Meissner, Eric Simley, Nicola Bodini, Lindy Williams, Shuangwen Sheng, and Rob Hammond. 2021. *Wind Plant Performance Prediction Benchmark Phase 1 Technical Report*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5000-78715. <https://www.nrel.gov/docs/fy22osti/78715.pdf>.

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5000-78715
October 2021

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

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 herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Acknowledgments

The authors thank Patrick Gilman, U.S. Department of Energy (DOE) Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office, for supporting this research. Thanks also to Richard Tusing, National Renewable Energy Laboratory (NREL), for providing strategic guidance in the development of this work. We thank our many external partners for providing support, in-kind labor contributions, as well as data for our analysis. We thank our sponsors at DOE, without whom this project and the important outcomes would not be possible. We thank our colleagues at NREL for giving advice on this project, including Eric Lantz, Rafael Mudafort, Mike Sprague, and the countless others who have served as sounding boards for this journey. Finally, we thank Erik Hale for his vision to create this collaboration; Cory Jog for keeping us technically sound; Patrick Pyle, Matthew Meyers, Dan Jaynes, Jared Kassebaum, Tadd Lisman, Matthew Breakey, Brian Healer, and Antonio Ruiz for keeping us grounded; Errol Halberg, John Bosche, Steve Hatlee, Lesley Dubois, Holly Burnett, Taurin Spalding, and all the consultants for their dedicated support over the years. Last, but not least, thank you to Michael Pram Nielsen, Oswaldo Rodriguez, and Paul Medina, who helped bring original equipment manufacturers to the table and provide critical data for the project.

List of Acronyms

AEP	annual energy production
DAP	Data Archive and Portal
DOE	U.S. Department of Energy
EYA	energy yield assessment
HVS	historical validation survey
IEC	International Electrotechnical Commission
IQR	interquartile range
NREL	National Renewable Energy Laboratory
OEM	original equipment manufacturer
OpenOA	Open Operational Assessment
PNNL	Pacific Northwest National Laboratory
PRUF	performance risk, uncertainty, and finance
ReEDS	Regional Energy Deployment System
RST	reStructuredText
SCADA	supervisory control and data acquisition
WP3	Wind Plant Performance Prediction

Executive Summary

To quantify and reduce the difference between the expected and the actual energy production of wind plants, the National Renewable Energy Laboratory (NREL) orchestrated an industry-wide data sharing initiative, the Wind Plant Performance Prediction (WP3) project. Through this multiyear effort, NREL gathered preconstruction and operational data from industry partners, developed the Open Operational Assessment (OpenOA) open-source operational analysis software, and studied historical trends of the energy prediction bias. NREL has recently completed Phase 1 of the benchmark exercise. This report summarizes the infrastructure, methodology, and tools NREL developed to facilitate this initiative as well as the results and lessons learned from the Phase 1 analysis.

The motivation for this work is that the financial risk associated with developing, owning, and operating wind power plants remains a barrier to reducing the levelized cost of energy for wind. Specifically, predicting the amount of energy a given wind plant will produce before it is constructed remains challenging, and industry has observed a bias toward overprediction (Lee and Fields 2021). Further, the uncertainty associated with these predictions remains high, resulting in a wide range of possible outcomes for wind plant owners and investors. Given the narrow operating margins of wind power plants, this overprediction of energy production—and therefore the overprediction of revenue—is damaging to project economics. A wind industry owner consortium (owners) approached NREL and the U.S. Department of Energy to develop an industry working group focused on reducing this source of financial risk through a first-of-its-kind data sharing initiative. The Wind Plant Performance Prediction (WP3) project was created from this request, and this report represents the successful completion of Phase 1 of the WP3 project to benchmark industry energy prediction performance.

In Phase 1, 6 wind plant owner participants agreed to share preconstruction and operational data for a total of 10 existing wind plants. The preconstruction data were shared with NREL from industry consultant participants, who typically conduct energy yield assessments (EYAs) before a wind project is financed and built. The results from these EYAs—conducted during 2018 and 2019—establish the expected energy production during the lifetime of the wind plant and form the basis for project economics and financial deal structures. The operational data were analyzed by researchers at NREL, who computed a benchmark, long-term corrected, annual energy production (AEP) value to which EYAs could be compared. NREL acted in the role of data aggregator and manager, ensuring the secure transfer of data among stakeholders and conducting analysis of the results. This role as an independent third-party data arbiter is a unique capability provided by government and federally funded research-and-development centers.

The WP3 team administered various foundational studies and data sharing exercises before the full inauguration of the Phase 1 analysis. NREL successfully conducted a historical validation survey (Lunacek, et al. 2018) demonstrating a 5.5% estimation bias, including curtailment across the industry; completed a pilot project to test the process for benchmarking consultant EYAs against operational data; and then rolled out Phase 1 of the WP3 benchmark to analyze 10 operating wind plants. NREL has developed and tested several internal capabilities to manage and share data, collect EYA results, and conduct an open-source operational assessment using the OpenOA platform.

NREL and the WP3 benchmark team have successfully completed Phase 1 of the benchmark. All 10 wind plants have undergone operational analysis by NREL, and 68 of a possible 80 EYA submissions have been received from 8 participating consultants. With this data set, NREL can compare EYA submissions with the operational data to establish a baseline for performance and can begin exploring the underlying reasons for error and uncertainty. A detailed explanation of the operational assessment methodology is provided in Section 2.3.4, and the resulting comparison of EYA data to operational data is discussed in Section 4.

Key findings of the Phase 1 analysis are as follows:

- We record mean and median underpredictions of AEP of -1.2% and -0.7 %, respectively, with a sample standard deviation of 4.8% which is four times greater than the mean underprediction.
- The energy prediction errors and the associated uncertainties demonstrate considerable variations among consultants. In particular, the average energy prediction biases for each consultant range from 4% underprediction to 2.6% overprediction. The associated standard deviations of the median consultant AEP estimate bias (P50) span from 3.7%–6.1%.
- The EYAs in our sample tend to have similar direction of overprediction or underprediction of bias for a given project but differing magnitudes of the predicted energy bias. Specifically, median energy prediction biases range from 6.3% underprediction to 4.1% overprediction at the project level.
- Gross energy comparisons reveal broad disagreement among consultants. Gross energy represents the total energy estimated to be available to the wind plant before physical losses are considered. For most projects, the gross energy estimates vary from 5%–10% among consultants. This is a significant source of spread in the resulting EYA process. The spread in gross energy suggests that efforts related to improving wind resource estimation, such as improved instrumentation and modeling, will help to reduce overall P50 variability.
- Total energy loss and categorical loss comparisons reveal the wide variation in the estimated energy lost from physical and operational processes. Individual consultant estimates of total loss range from 14%–26%, with a median value of 18.8% across all projects. These losses include but are not limited to electrical losses, wakes, and icing.
- In general, there is much broader disagreement among consultants in both International Electrotechnical Commission categorical uncertainties and total AEP uncertainties than in corresponding losses. The source of this variation is unclear, but it is likely driven by each participant’s internal, unique uncertainty estimation methods and training data sets. Uncertainty quantification appears to be the least mature area of the EYA process as demonstrated by the level of variation amongst participations.
- In a preliminary analysis of the disagreement in categorical loss and uncertainty estimates among consultants, we study the amount of disagreement by determining the interquartile range (IQR) of the difference between the estimates for each category and the corresponding project-level mean estimates. This analysis suggests that, for the data set studied, turbine performance, environmental, and wake loss estimates are the largest sources of disagreement among the loss categories. Specifically, disagreement for these three loss estimates (IQR values of 1.56% to 1.75%) is greater than the disagreement for

availability (IQR=1.12%) and electrical losses (IQR=0.71%). Further research is needed to draw industry-wide conclusions and to determine which loss categories have the greatest impact on total energy prediction variability.

- In a preliminary analysis of the disagreement in categorical energy prediction uncertainty estimates among consultants, project evaluation period uncertainty (IQR=2.61%) and plant performance uncertainty (IQR=2.26%) record the highest interquartile ranges. This implies that consultant estimates of 1) how well the wind resource over the evaluation period represents long-term average conditions, and 2) plant performance (which is directly related to energy losses) are among the largest sources of disagreement for uncertainty estimation. Further research is needed to draw industry-wide conclusions and to determine which uncertainty categories have the greatest impact on the variability of total energy prediction uncertainty.

Together, these results suggest that bias correction is generally effective at reducing average P50 bias; however, there remains broad disagreement on the methods to arrive at a consensus P50 value. Additionally, there remains fundamental disagreement on the steps in the EYA process, as demonstrated by commercially significant variations among consultants for estimates of gross energy, losses, and uncertainties. This demonstrates the need for consultant access to additional verification data as well as research to identify the best methods for gross energy quantification and uncertainty estimation.

The WP3 benchmarking initiative, while successful has generated many valuable lessons for future data sharing initiatives. There is a definitive need to improve data collection and data sharing practices to ensure scalability of future efforts. Acquiring access to high-quality data requires continuing industry support through adoption of common data standards and analysis definitions. Follow-on work should seek to scale to many projects with improved representation for geography, markets, and turbine technology. Future efforts in this space should investigate specific sources of bias between predictions and actual operations. This includes topics such as gross energy, wake modeling, and environmental effects.

Table of Contents

1	Introduction	1
1.1	The Challenge	1
1.2	The WP3 Solution.....	3
1.2.1	Project Participants.....	3
1.2.2	Project Structure and Operations.....	4
1.3	WP3 Project Goals	4
2	Capabilities and Methodology	6
2.1	Data Archive and Portal	6
2.2	Energy Yield Assessment Results Collection	6
2.2.1	EYA Results Collection Web Application.....	6
2.3	OpenOA (Operational Assessment Software).....	8
2.3.1	Software Architecture of OpenOA.....	9
2.3.2	Documentation	11
2.3.3	Open-Source Development	12
2.3.4	Operational Annual Energy Production Benchmark Methodology	12
3	Work Packages	19
3.1	Historical Validation Study.....	19
3.2	Benchmark Pilot.....	19
3.2.1	Consultant Response Collection.....	19
3.2.2	Pilot Results	20
3.2.3	Lessons Learned.....	24
3.3	WP3 Phase 1.....	25
3.3.1	Phase 1 Setup	25
3.3.2	Project Characteristics.....	26
4	WP3 Phase 1 Analysis	28
4.1	P50 Net Energy Estimate	28
4.2	Gross Energy	30
4.3	Energy Losses	31
4.4	Prediction Uncertainty.....	33
4.5	Energy Yield Assessment Performance Evaluation	35
5	Lessons Learned	38
5.1	Data Access	38
5.2	Data Management	39
5.3	Energy Yield Assessment Results Collection	39
5.4	Operational Data and Assessment.....	40
5.5	Metrics, Analysis, and Results	40
6	Conclusions and Next Steps	42
	References	43

List of Figures

Figure 1. Wind capacity additions from the “Wind Energy Technology Data Update: 2020 Edition” (Wiser et al. 2020).....	2
Figure 2. Prediction bias and uncertainty in wind energy estimates. The negative mean and median biases illustrate overprediction of annual energy production (AEP).	3
Figure 3. WP3 data sharing structure.....	4
Figure 4. Example of EYA application data collection fields	7
Figure 5. Home page and navigation for the EYA application.....	8
Figure 6. Schematic of dependencies among modules for an example computation of AEP.....	10
Figure 7. Unified Modeling Language diagram depicting the object structure that implements the OpenOA data model.....	11
Figure 8. Screenshots of OpenOA documentation: (left) application programming interface for the filters tool kit and (right) Monte Carlo P50 AEP Python Jupyter Notebook	12
Figure 9. A Monte Carlo approach to AEP estimation using operational data.....	14
Figure 10. Example of long-term monthly wind speeds (normalized) for different reanalysis products. An example period of record for a wind plant is shown in shaded blue.	15
Figure 11. Example of monthly gross energy during the plant period of record	15
Figure 12. Example of monthly availability and curtailment loss during the plant period of record	16
Figure 13. Scatterplots of gross energy and monthly averaged wind speeds for different reanalysis products. Outliers are identified as red crosses.	17
Figure 14. Long-term AEP distribution after 50,000 simulations	18
Figure 15. Screenshot from the Energy Yield Summary tab of the EYA results collection tool used for the pilot project. The consultants were asked to fill out the green cells.	20
Figure 16. P50 gross and net bias. Consultant estimates for z01 capacity factor. The blue and orange bars represent gross and net energy production, respectively. “AEP” represents the NREL-calculated benchmark capacity factor (normalized to 1), and the shaded bar represents the uncertainty.....	21
Figure 17. P50 bias comparison. Percentage difference between consultant P50 values and the NREL operational assessment. The zero horizontal line represents the NREL operational assessment value, and the grey band represents the uncertainty.....	21
Figure 18. Uncertainty comparison: consultant estimates for z01 uncertainty. The orange box plots represent net energy production and the uncertainty predicted by consultants. The blue shaded box plot represents the NREL-calculated benchmark capacity factor.	22
Figure 19. Consultant-estimated losses by category for z01. Each box plot distribution represents 10 different consultant estimates for z01 losses.....	23
Figure 20. Consultant estimates for z01 uncertainty by category. Each box plot distribution represents 10 different consultant estimates for z01 uncertainty.	23
Figure 21. Consultant estimates for per-turbine capacity factor for z01. Each box plot distribution represents 100 turbines, with standard deviation displayed at the top of the plot.....	24
Figure 22. z01 Wake deficit analysis of the predicted wakes from three consultants	24
Figure 23. z01 resource analysis of the turbine-level annual average free-stream winds from three consultants.....	24
Figure 24. Predicted near-term U.S. wind plant deployments used when selecting Phase 1 projects	26
Figure 25. P50 bias distribution for WP3 Phase 1. The positive values illustrate overprediction of AEP. 29	29
Figure 26. P50 bias distribution for WP3 Phase 1 color-coded by project and participant	29
Figure 27. P50 bias box plot distributions for WP3 Phase 1 by project and by consultant	30
Figure 28. P50 waterfall chart for WP3 Phase 1 by project and by consultant.....	31
Figure 29. Total loss box plot distributions for WP3 Phase 1 by project	32

Figure 30. IEC 61400-15 loss category box plot distributions for WP3 Phase 1 by project. Curtailment loss estimates are not included here because they were not considered as part of the WP3 analysis..... 33

Figure 31. Total energy prediction uncertainty (10-year standard deviation) box plot distributions for WP3 Phase 1 by project and by participant 34

Figure 32. IEC 61400-15 uncertainty category box plot distributions for WP3 Phase 1..... 35

Figure 33. WP3 Phase 1 comparison of predicted EYA to operational assessment by project..... 36

Figure 34. WP3 Phase 1 comparison of predicted EYA to operational assessment by participant 36

List of Tables

Table 1. Top Wind Installation States from 2016–2020 Modeled in ReEDS with Mid-Case Scenario..... 26

Table 2. P50 Bias Performance by Participant 37

1 Introduction

Financial risk resulting from the uncertainty associated with developing, owning, and operating wind power plants remains a barrier to reducing the levelized cost of energy. At the inception of the benchmark project, a National Renewable Energy Laboratory (NREL) analysis showed that modern wind power plants in the United States were underperforming in their expected annual energy output by 3.5%–4.5% after controlling for grid curtailment losses (Lunacek et al. 2018), with many underperforming by more than 10%. A more recent literature review conducted by NREL suggests that biases in wind power plant energy predictions are decreasing over time (Lee and Fields 2021). Nonetheless, to compensate for this uncertainty, investors require a larger return on investment and apply correction factors that mask much of the underlying sources of uncertainty; thus, wind energy projects have reduced access to low-cost capital. Further, operating wind plants often take a simple approach to estimating operation-and-maintenance costs (e.g., straight-line estimates based on similar plants), which can eat into profits. To overcome these issues, the wind industry must improve the models they use for estimating wind plant performance and operations.

A consortium of wind plant owners requested that NREL lead a U.S. Department of Energy (DOE) working group to benchmark the accuracy of wind power plant energy predictions against real operational data. The owners were also motivated by DOE and NREL’s potential to characterize systematic energy underperformance, identify sources of uncertainty, and explore root causes. The Wind Plant Performance Prediction (WP3) project was created, and this report represents the successful completion of Phase 1 of the WP3 project. Note that the WP3 benchmark project is not the first initiative to benchmark industry performance; it is, however, the first independent analysis of this depth and scale. You can find a summary of the various industry analyses in (Lee and Fields 2021).

During the project, wind plant owners provided both preconstruction and operational data to NREL. The preconstruction data were provided to wind resource assessment consultants so they could conduct energy yield assessments (EYAs). NREL took all the completed EYAs, along with the operational data, and conducted an operational assessment to benchmark the EYA results against actual operational data. Given the large amounts of sensitive data required for this effort, as well as the historical opposition to sharing data within industry, successful completion of Phase 1 represents an unprecedented milestone for industry data sharing.

1.1 The Challenge

The U.S. wind industry is facing the expiration of the production tax credit, which has been a driver of industry growth for more than a decade. As a result, many institutions are predicting a sharp decline in the development of wind power plants starting in 2021. Figure 1 shows this installation “cliff” as forecasted by several organizations.

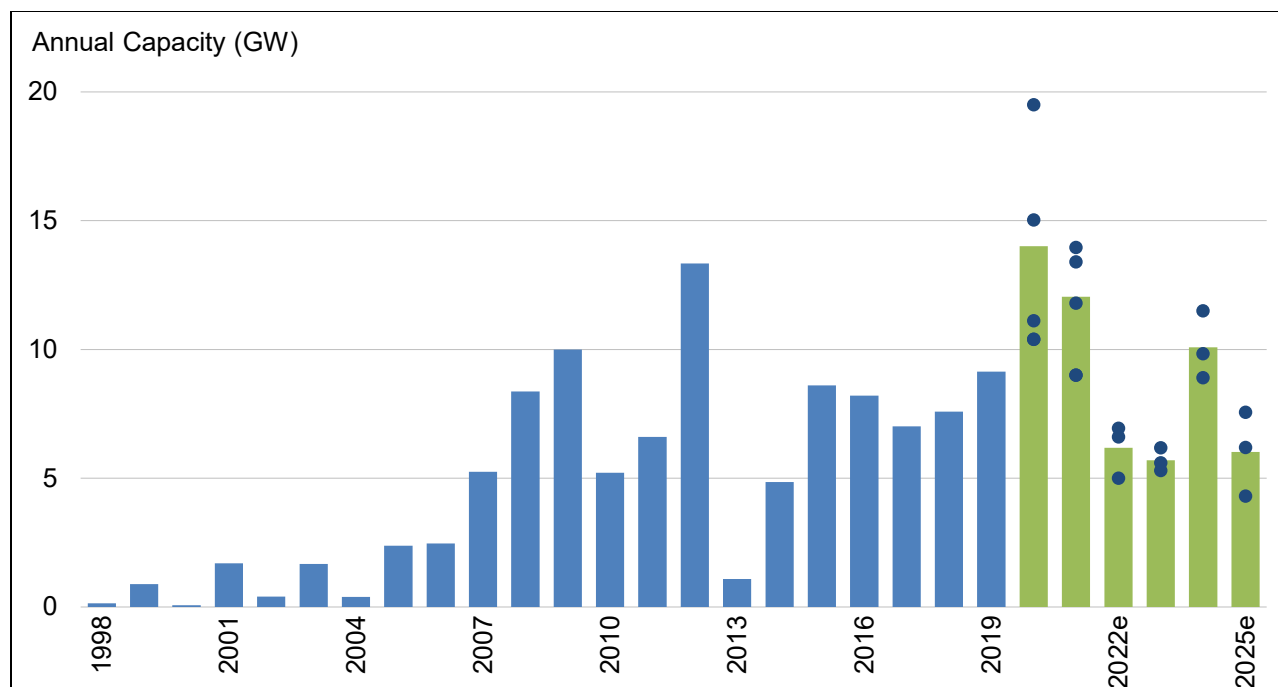


Figure 1. Wind capacity additions from the “Wind Energy Technology Data Update: 2020 Edition” (Wiser et al. 2020)

Wind plant owners can prepare for this potential decline in installations by reducing the cost of building, owning, and operating wind plants. One of the best, near-term ways to do this is by reducing the uncertainty and risk associated with the financing of new projects. The riskier it appears to own and operate a wind power plant, the more expensive it will be to acquire capital for development and construction. Further, large wind plant owners are concerned that their wind project portfolio might not meet long-term performance expectations, which dampens the future of wind development. One primary driver of this risk and uncertainty are preconstruction EYAs made by third-party consultants and in-house energy assessment teams at developers. In a validation study of historical data, NREL researchers found that EYAs overpredict energy production by 5.5% on average, including curtailment, and that the total error on those estimates ranges from 5%–12% (Lunacek et al. 2018). This study demonstrates the opportunity for improvement. Figure 2 demonstrates the bias and uncertainty found in the historical validation study (HVS).

The financial impacts of this performance gap will be project specific. Based on our conversations with wind plant owners, an overprediction of only 3% poses a significant threat to their profitability, which is a strong disincentive to continue investing equity capital. Further, a 1% change in energy yield is equal to approximately \$20/kW in the turbine price, or approximately \$4 million net present value. The typical energy yield assessment takes 100–200 hours for wind resource assessment consultants to complete, yet overprediction bias and uncertainty persist (Lee and Fields 2021). By addressing this performance gap, industry can reduce financial risk for investors and decrease the weighted cost of capital.

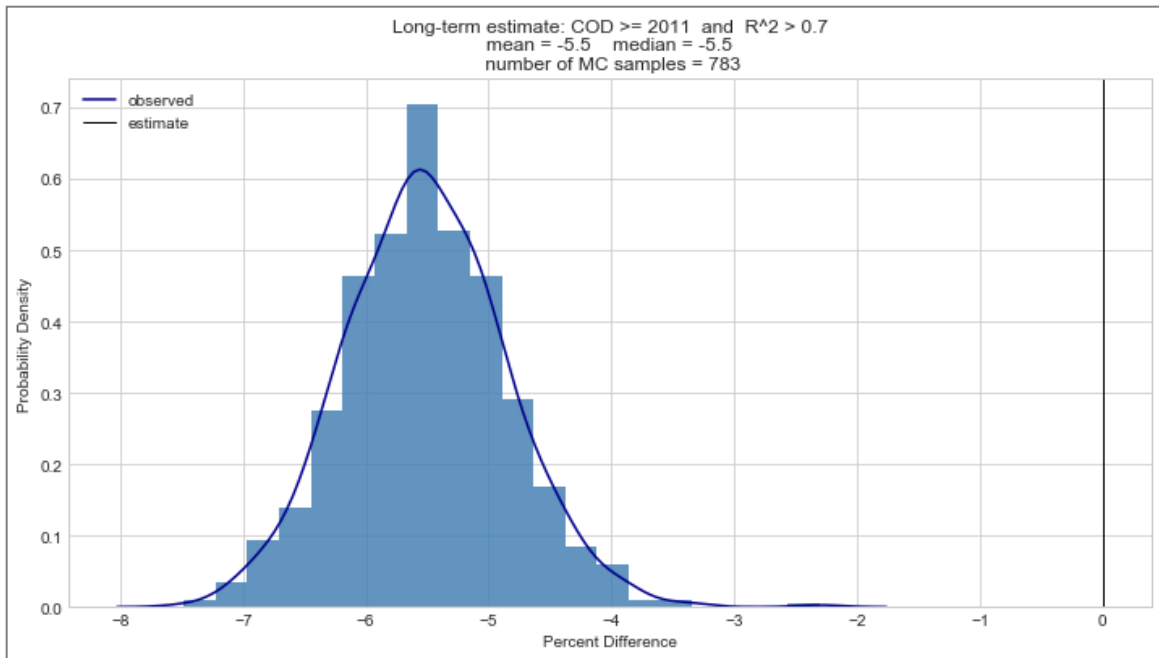


Figure 2. Prediction bias and uncertainty in wind energy estimates. The negative mean and median biases illustrate overprediction of annual energy production (AEP).

1.2 The WP3 Solution

To improve the accuracy and confidence of preconstruction EYAs, wind plant owners and investors need better, more certain energy yield predictions. The WP3 benchmark project is an industry-driven response to this reality. For the first time, industry has taken the important step of working together at scale, sharing valuable operational data with DOE and NREL to investigate the sources of bias and uncertainty in these energy estimates. The owners provide wind plant preconstruction and operational data to NREL in an organized and documented fashion, and NREL provides guidance and feedback as needed. The owners also provide introspection of the design of experiment, key metrics of success, data challenges, analysis best practices, and quality of results.

1.2.1 Project Participants

- **Project owners:** Project owners provide preconstruction and operational data to the project. They also worked closely with the NREL team to help develop internal capabilities for benchmarking preconstruction energy estimates against the operational data. For Phase 1 of the WP3 benchmark, the owner includes EDF Renewables Inc, EDP Renewables, RWE, Pattern Energy, Renewable Energy Systems, Enel Green Power, and Avangrid.
- **Resource assessment consultants:** The consultants use the preconstruction energy data to conduct EYAs. From there, NREL compares their EYA against operational data to benchmark the results. For Phase 1 of the WP3 benchmark, the consultants include WSP, ArcVera Renewables, UL (formerly AWS), EAPC-EMD, Luminate, Mott MacDonald, Natural Power, K2 Management, (formerly Prevailing Wind), and Wood Group.
- **Original equipment manufacturers (OEMs):** The equipment manufacturers offer insights into various aspects of the preconstruction data and operational data. For Phase 1

of the WP3 benchmark, the OEMs include Acciona-Nordex, General Electric, Vestas, Senvion, and Siemens Gamesa.

- **Government:** Government participants act as neutral third parties, holding and managing data as well as conducting objective analysis for the benchmark. For Phase 1 of the WP3 benchmark, the government participants include DOE, NREL, and the Pacific Northwest National Laboratory (PNNL).

1.2.2 Project Structure and Operations

NREL interfaces directly with wind plant owners and resource assessment consultants for data sharing, including the transmittal of project data and benchmark submissions. Nondisclosure agreements are in place to protect all stakeholder data from improper use. NREL serves as the centralized data collection and reporting hub for the initiative. Wind plant owners transfer project data to NREL for cataloging and distribution to the wind resource consultants. NREL then analyzes consultant responses for accuracy of preconstruction estimate methodologies compared to actual production. Upon submission of all project preconstruction EYAs, consultants are given access to operational data for the wind plants and asked to calibrate their EYA estimates. The operational data are valuable to the consultants and are seen as a benefit of participation in the WP3 process. DOE labs can also use the data for ongoing research per the terms of the nondisclosure agreement. Figure 3 displays a graphical representation of stakeholder responsibilities, roles, and data flows.

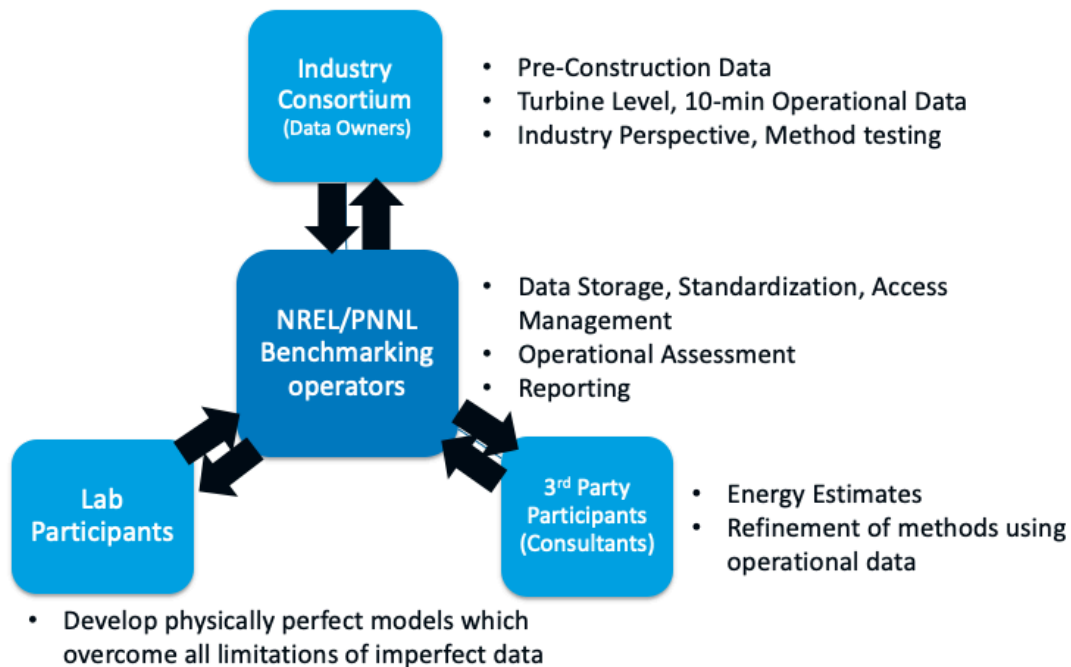


Figure 3. WP3 data sharing structure

1.3 WP3 Project Goals

Ultimately, the benchmark will verify the accuracy of the annual energy production (AEP) prediction process, including the 50% probability of exceedance of the AEP estimate (P50),

uncertainty estimation process values (P75, P90, P95, P99), and loss estimation processes. More broadly, WP3 has three high-level goals:

1. Create a platform for data sharing, data analysis, and model improvement.
2. Generate accurate, independent benchmarks of preconstruction energy assessment accuracy using operational performance data from industry.
3. Improve the accuracy and confidence of preconstruction energy estimates.

2 Capabilities and Methodology

NREL, PNNL, and DOE are actively developing capabilities that support the storage, aggregation, analysis, and reporting required to implement the WP3 benchmark initiative.

2.1 Data Archive and Portal

The DOE-sponsored WP3 benchmark research project leverages PNNL's Data Archive and Portal (DAP) to collect, store, and disseminate data among performance risk, uncertainty, and finance (PRUF) stakeholders and WP3 participants. DAP and PRUF are both part of DOE's ongoing wind research activities under the Atmosphere to Electrons program (a2e.energy.gov).

Wind plant owners provide both preconstruction and operational data to NREL via the DAP, which then goes through a quality-control step and iterative dialogue with owners to answer questions and to ensure the integrity of the data sets. Wind resource assessment consultants download the preconstruction data via DAP, conduct their resource assessment, and submit results to NREL via a collection form that is hosted on the DAP. NREL WP3 staff manage partner access to the DAP and can upload and download data as required to facilitate data sharing.

2.2 Energy Yield Assessment Results Collection

The WP3 benchmark process requires input data from consultants in the form of completed EYAs. There are, however, challenges with comparing the output from consultants' models, given the diversity in output formats, units, and other nomenclature. The WP3 team developed an online form, hosted by the DAP, to standardize submissions and to ensure a fair and accurate comparison. To perform this harmonization, NREL is relying on the draft requirements from the International Electrotechnical Commission (IEC) 61400-15 working group "Assessment of Wind Resource, Energy Yield and site suitability input conditions for wind power plants." This draft standard helps to bring together disparate results and descriptions from a number of consultants. The WP3 benchmarking initiative is one of the first times that the unified IEC framework has been applied in practice (Lee and Fields 2021; Fields and Sherwin 2018; Filippelli et al. 2018). As such, there are still some issues to be overcome, and NREL has developed several data collection approaches to remedy these issues during the WP3 benchmark pilot and Phase 1 portions of the work.

2.2.1 EYA Results Collection Web Application

The EYA application is a client-side web page that runs only in a web browser and facilitates the collection of robust preconstruction estimates. The application makes it possible for NREL to validate data before submission, which creates a higher likelihood that values will be easy to access and directly comparable once submitted. Prior to building the current EYA application, NREL experimented with other collection methods, such as Microsoft Excel spreadsheets. It was quickly found that entries often had incorrect units, and the data processing was difficult, if not impossible, to automate because of additional rows or columns being added or the variable descriptions and column headers being changed by our external participants. The EYA application addresses these data quality problems by providing a more rigid input form that provides real-time feedback to the users on what they are entering and how NREL will be interpreting those values.

Figure 4 shows an example of the collection form and the feedback we give users during the process. In the figure, one cell is highlighted in red because the entry is outside an expected range. This does not prevent users from submitting the value, but it highlights this potential error at the point of entry, providing the user with an opportunity to correct the data before they are submitted to NREL. Finally, when possible, the EYA application provides a summary of how entered values will be interpreted by NREL. Total losses, for example, are computed in real time by aggregating the reported loss categories. This provides users with one more cue to detect data entry errors. One important design decision was that all tabular data can be copied and pasted between the EYA web application and Microsoft Excel (or a similar spreadsheet platform) using the client's clipboard. This allows users to work with and reshape their data using a tool that is familiar to them before the values are transferred to the application in bulk. This reduces errors caused by entering multiple values by hand or from using unfamiliar software.

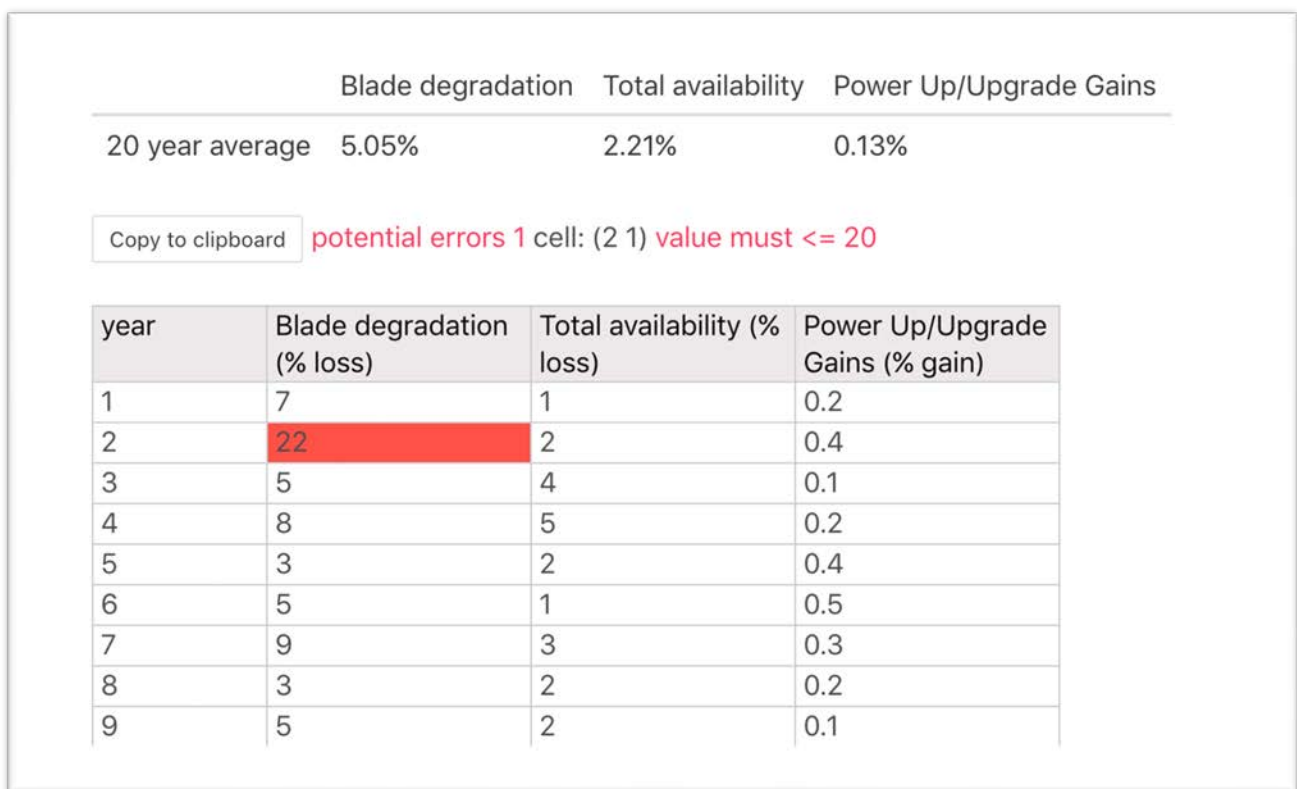


Figure 4. Example of EYA application data collection fields

There are certain advantages to using a client-side application. When a user submits results, the application posts a data packet to a server that simply forwards the packet to a collection email account. The application is statically served, and no user data are stored on the server. This architecture reduces maintenance requirements and eliminates the potential security vulnerabilities associated with running a back end. Because the application is served through a web browser, NREL can update the code and roll out to users instantly as well as track the different versions of the data model as changes occur. The application also includes an export feature that makes it possible download an EYA submission file to share with colleagues. Future versions of this application might add server-side support as more requirements are identified.

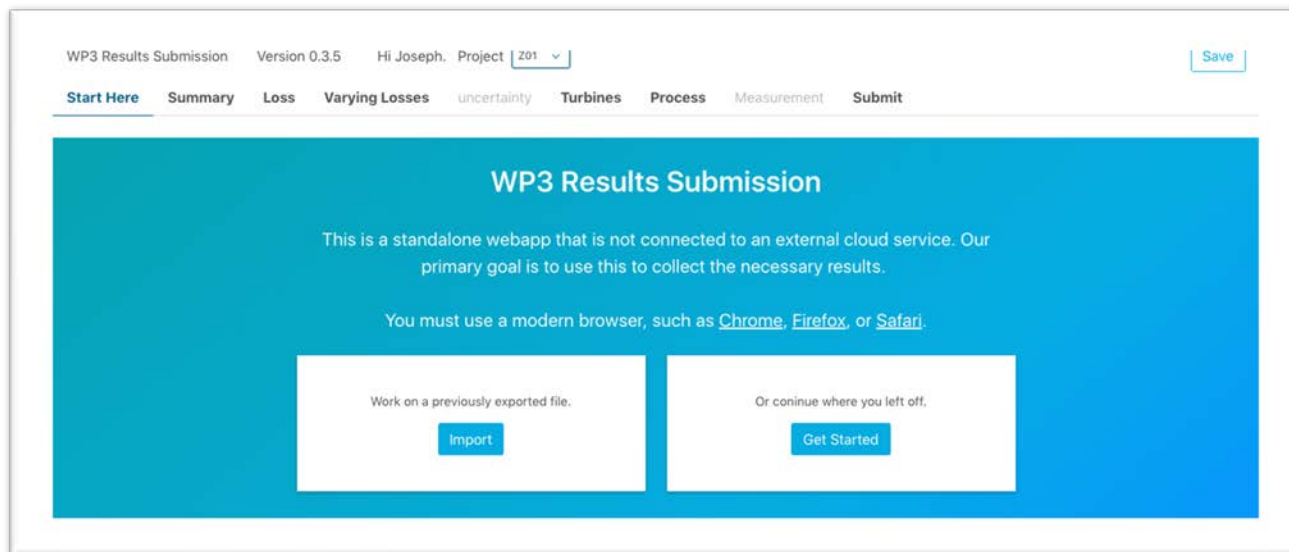


Figure 5. Home page and navigation for the EYA application

Figure 5 shows the home page of the application, which is organized into tabs that represent the various aspects of the EYA data. At the beginning of Phase 1, it was found that the amount of data requested from consultants was too great and that full compliance with the data collection application was simply not feasible. NREL then defined a subset of all the fields (the bold tabs in Figure 5) to kick-start the data collection phase and reduce the data entry burden on consultants. The reduced set of data allows NREL to collect enough information to facilitate the stated goals of the Phase 1 benchmark. Note that the analysis is therefore limited by the data we could reasonably collect from our partners. Additional data might be requested by the WP3 team to provide the remaining portion of the request (the non-bold tabs in Figure 5) to take the analysis to an even more granular level. So far, the results received through the collection application have been much easier to analyze in an automatic fashion, and the feedback from consultants who use the app to enter their data has been widely positive. NREL expects to implement feature enhancements and to begin requesting more detailed data requests via the non-bold tabs in the form (e.g., the “uncertainty” tab) for future phases of the project.

2.3 OpenOA (Operational Assessment Software)

On the operational assessment side, the WP3 team developed an open-source operational assessment software library, OpenOA (Perr-Sauer et al. 2021), hosted on GitHub (<https://github.com/NREL/OpenOA>), to produce consistent industry standard operational analysis metrics. These metrics can be compared with the energy predictions in the EYA data. These metrics, computed from operational data, are a critical part of the benchmark analysis because they represent the “ground truth” energy production values. As such a critical piece of the analysis, it became apparent that our software should be made open source so that it could be shared with and reviewed by WP3 stakeholders. At the time of the benchmark analysis, OpenOA software contained two high-level analysis methods.

The plant-level analysis summarizes monthly energy generation, availability, curtailment, and other high-level losses to generate a plant-wide, long-term corrected AEP value. It is important to understand which losses are included in this value and which are not so that an apples-to-

apples comparison can be made with the consultant EYA estimates. Drawing from the discussions with industry partners, NREL developed a set of industry standard programmatic tools and algorithms to compute this key performance metric. A detailed description of this analysis method is provided in Section 2.3.4.

The turbine-level analysis includes high-resolution data (10-minute averages or finer data) for each turbine, which allows a granular look at energy production, loss categories, and sources of uncertainty. These data allow NREL researchers to estimate the individual loss categories from operational data.

OpenOA provides the first open-source tool kit for this purpose. Outside of its immediate use for the PRUF project, the design of this software was also supported by the NREL Research Data Initiative and provides NREL researchers and the broader wind industry with a capability to perform consistent, reproducible operational analysis on wind plant data. More importantly, the operational analysis conducted through OpenOA reflects industry standards and future capabilities that do not exist consistently across industry.

2.3.1 Software Architecture of OpenOA

OpenOA is a package for the Python programming language, which is compatible with Python 3.6 or greater. The package contains three modules: (1) types, (2) tool kits, and (3) methods. Figure 6 shows the relationship between each component and provides an example of a computed distribution for AEP. In this example, we use data from the supervisory control and data acquisition (SCADA) system, the plant-level energy meter, and meteorological reanalysis data. The statistical flagging, gap-filling, and time-series tool kits are then used to process these raw data. Finally, a distribution of AEP metrics is computed using the operational AEP benchmark methodology. This example provides a high-level overview of an analysis using OpenOA and is not meant to represent any particular project.

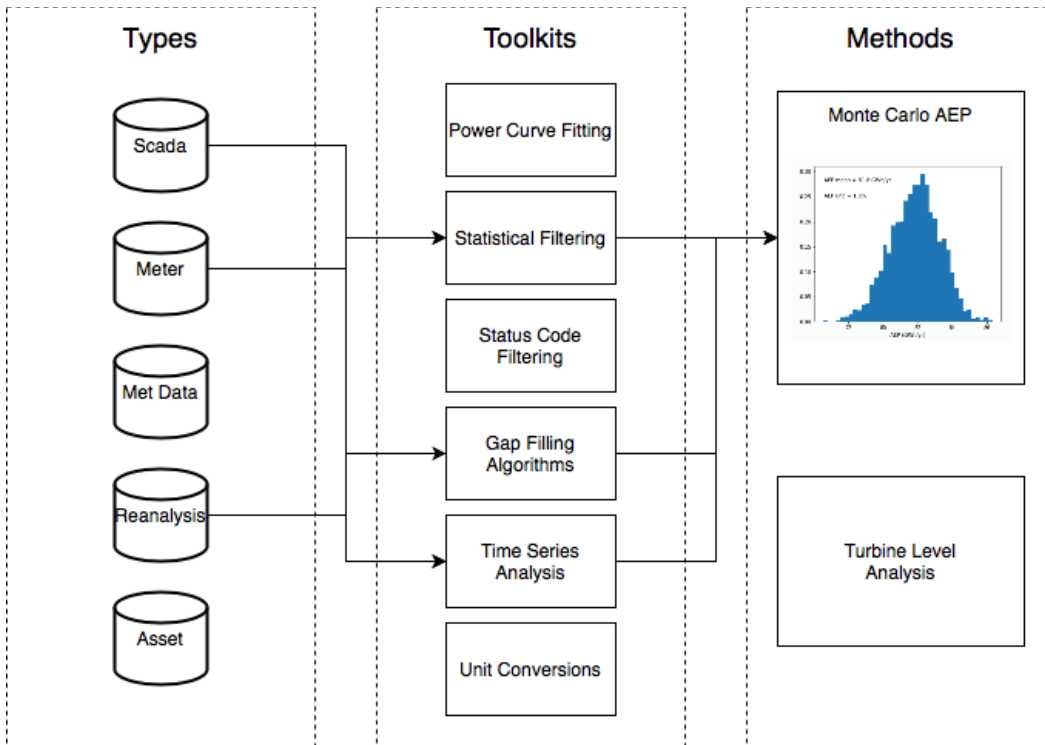


Figure 6. Schematic of dependencies among modules for an example computation of AEP

Here, we enumerate the three modules of OpenOA and explain their functionalities. The tool kits module is a container for general helper functions that represent common tasks in the analysis of wind plant operational data. These functions are written to operate on pandas data frames and NumPy series, making these functions useful to all Python developers and users, even if they do not leverage the other components of OpenOA. The tool kits module is organized by purpose, with submodules such as unit conversion, time-series imputing, and power curve fitting, to name a few.

The types module implements the model for wind plant operational data and associated data. The model is built around a class called PlantData, which acts as a high-level container for constituent data. This class is central to the processes of data management, cleaning, and transformation. The goal of this module is to help standardize operational data sets in the wind industry. We propose a data model with a schema that is based on the IEC 61400-25 standard as shown in Figure 7. The schema is provided in a JSON format and can be leveraged and extended by software developers.

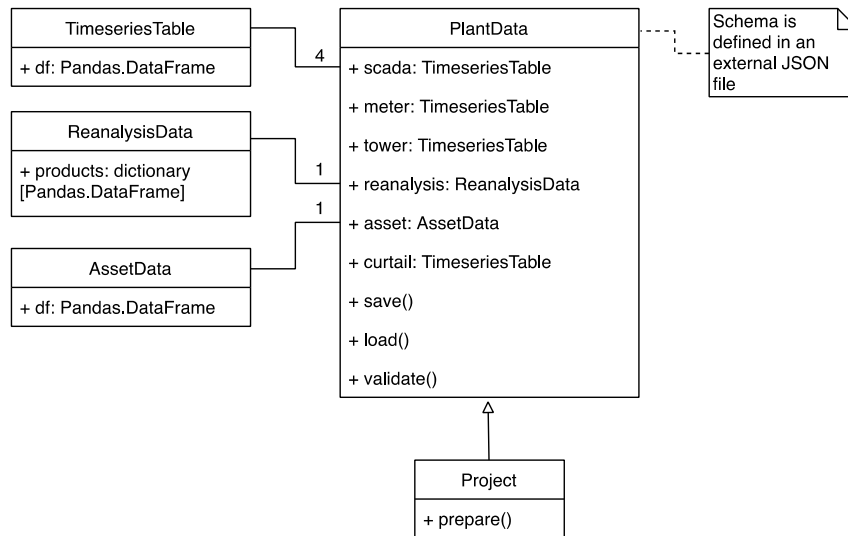


Figure 7. Unified Modeling Language diagram depicting the object structure that implements the OpenOA data model

The methods module contains Python classes that use the PlantData model and low-level tool kit functions to perform an end-to-end analysis. Two methods—the plant-level and turbine-level analyses—are available in this module at the time of writing. Both methods are implemented as Python classes and can be used in an interactive manner through the Python interpreter. These built-in analysis methods can be extended by creating subclasses and overriding their methods.

2.3.2 Documentation

Documentation is a critical part of any open-source codebase. It is the key that unlocks the code for unfamiliar users. The OpenOA documentation is comprehensive and includes the full application programming interface (API) of each tool kit module, installation instructions, and a set of Python Jupyter Notebooks with example outputs and figures. The documentation is drawn from internal docstrings and reStructuredText (RST) files throughout the codebase. It is automatically compiled after each merged pull request on GitHub. Documentation is then automatically pushed to a publicly hosted website on Read the Docs (<https://openoa.readthedocs.io>) as shown in Figure 8.

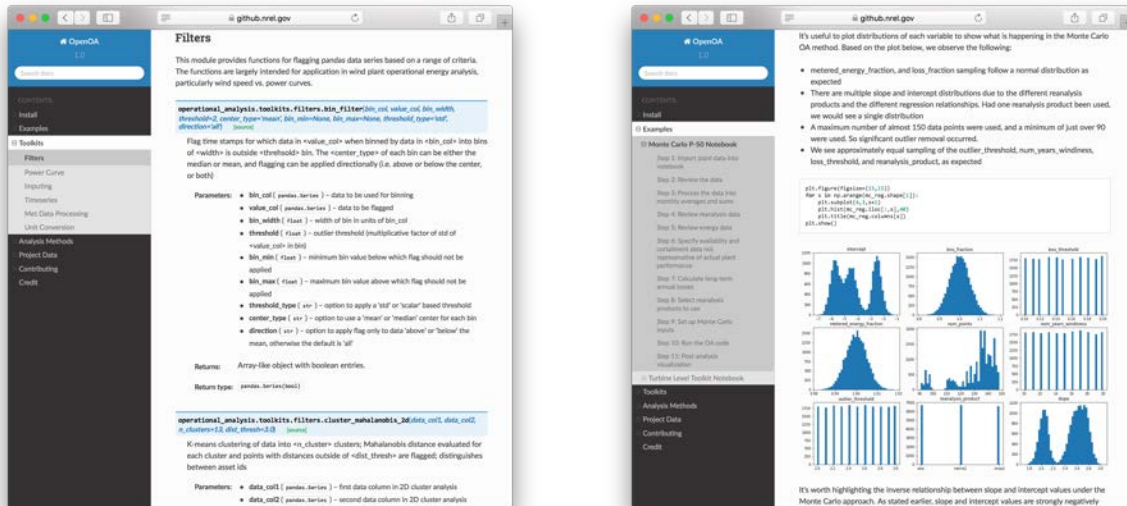


Figure 8. Screenshots of OpenOA documentation: (left) application programming interface for the filters tool kit and (right) Monte Carlo P50 AEP Python Jupyter Notebook

2.3.3 Open-Source Development

OpenOA is a multiyear effort, and the number of developers has been expanding over the years. After multiple internal beta releases throughout 2018, Version 1 of OpenOA was released to the public on September 10, 2018. The software is now developed in a public repository where the development team can collaborate with and accept contributions from the broader community.

2.3.4 Operational Annual Energy Production Benchmark Methodology

The WP3 method for generating an AEP estimate based on operational data follows an industry standard approach. This approach establishes a relationship between reported, loss-corrected monthly energy production and the concurrent, monthly average wind speed from a long-term data set. This relationship is then applied to a full long-term wind speed database to estimate the long-term AEP.

Uncertainty in the AEP estimate is calculated through a Monte Carlo approach (Montgomery and Runger 2014; Wilks 2011). Specifically, inputs into the OpenOA code, as well as intermediate calculations, are randomly sampled based on their specified or calculated uncertainties. By performing the operational assessment tens of thousands of times under different combinations of the random sampling, a distribution of AEP values results from which uncertainty can be deduced.

An overview of the Monte Carlo-based AEP calculation is shown in Figure 9. Uncertainty is assessed in the following input data and AEP calculation processes:

1. Reported energy data:
 - A. Uncertainty in reported revenue meter data is assumed to be 0.5%.
2. Modeled wind resource data:
 - A. Uncertainty is quantified by the use of different reanalysis products.

- B. ERA-Interim (Dee et al. 2011), the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) (Bosilovich et al. 2016; Gelaro et al. 2017), and Global Forecast System reanalysis data are currently considered.
3. Reported availability and curtailment losses:
 - A. Uncertainty in the reported losses is assumed to be 5%.
 4. Large combined losses:
 - A. Accounts for the fact that large, combined availability and curtailment losses for a given month result in a highly uncertain gross energy calculation.
 - B. It is industry standard to remove the data points when their combined losses exceed some threshold (e.g., 15%).
 - C. We account for uncertainty in the analyst decision here by varying the threshold from 10%–20% (Craig et al. 2018).
 5. Identification of regression outliers:
 - A. Accounts for analyst choice in assessing what is and is not an outlier in the linear regression relationship between the monthly energy production and the monthly average resource
 - B. Implemented through robust linear regression and random samples based on the Huber’s t weight function to select outliers between values from 2–3 (Huber and Ronchetti 2009; Seabold and Perktold 2010).
 6. Regression relationship:
 - A. Accounts for the imperfect relationship between monthly energy production and monthly average wind speed
 - B. Implemented by calculating the error in the slope and intercept values, as well as the covariance between the two, to resample the slope and intercept values that are then used for the remainder of the AEP calculation.
 7. Windiness correction:
 - A. Accounts for the variable historical wind resource and our confidence in the long-term correction applied to the AEP
 - B. Implemented by varying the period of record for the long-term correction from the last 10 years to the last 20 years.

Note that interannual variability, which captures the uncertainty of the wind resource in any given future year (Lee et al. 2018a, 2018b), is not considered for this operational AEP benchmark. Rather, the total uncertainty in the operational AEP reflects NREL’s confidence in the long-term AEP based on the historical data and time period available to NREL. For this reason, any uncertainty associated with future annual variations in the wind speed is beyond the scope of this total uncertainty calculation.

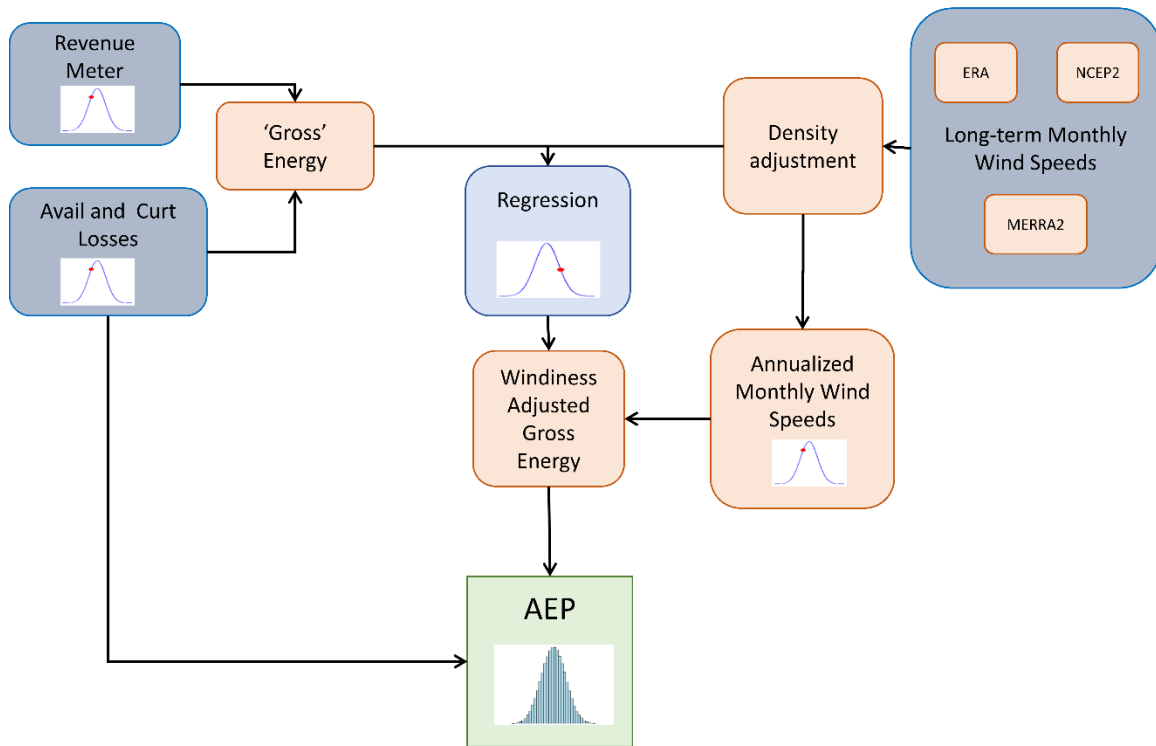


Figure 9. A Monte Carlo approach to AEP estimation using operational data

The steps in implementing this Monte Carlo-based AEP approach to calculate an AEP and uncertainty metric are summarized in the remainder of this section.

- **Step 1: Process and review reported monthly energy data.**
In this step, wind plant data (availability, curtailment, and revenue meter energy) provided by owner/operators are processed into monthly values (if not already done so) and under quality-control checks. Next, monthly gross energy is calculated by adding back the reported monthly availability and curtailment losses to the reported monthly revenue meter energy.
- **Step 2: Review reanalysis data and select products to use in AEP calculation.**
Reanalysis data are used to correct the operational energy during the plant period of operation to the long term. The reanalysis data are plotted (example provided in Figure 10) and checked for reasonableness, i.e., the reanalysis data showing reasonable trends over time with no noticeable discontinuities would pass the quality control test. If unreasonable trends or discontinuities are identified for a reanalysis product, it is not considered in the AEP calculation process.

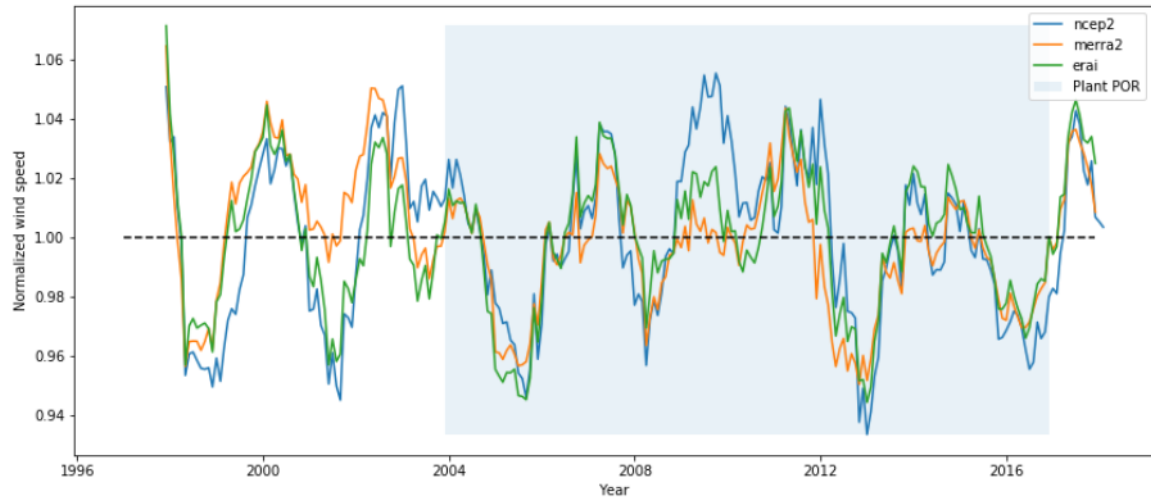


Figure 10. Example of long-term monthly wind speeds (normalized) for different reanalysis products. An example period of record for a wind plant is shown in shaded blue.

- Step 3: Review wind plant data and calculate long-term losses. Monthly time series of gross energy and availability and curtailment losses are plotted and manually checked for reasonableness (examples provided in Figure 11 and Figure 12). Large, anomalous availability and curtailment losses are flagged. Flagged data require follow-up with either the data provided or the owner/operator to determine if these losses are considered representative of long-term plant performance. Losses that are deemed not representative of long-term performance are excluded when calculating the long-term availability and curtailment losses for the wind plant.

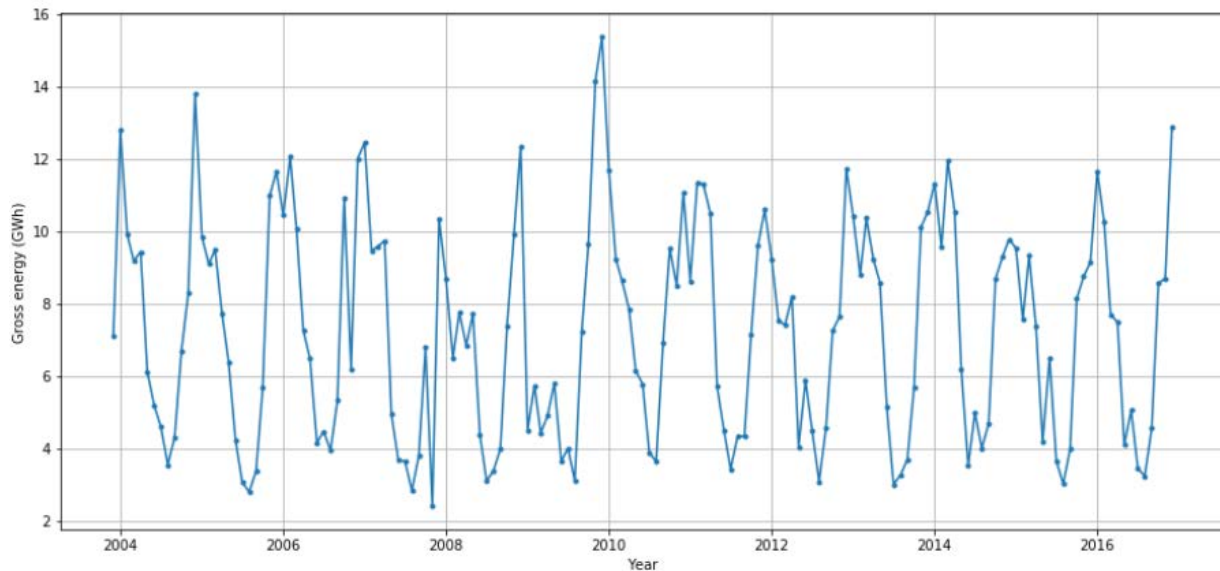


Figure 11. Example of monthly gross energy during the plant period of record

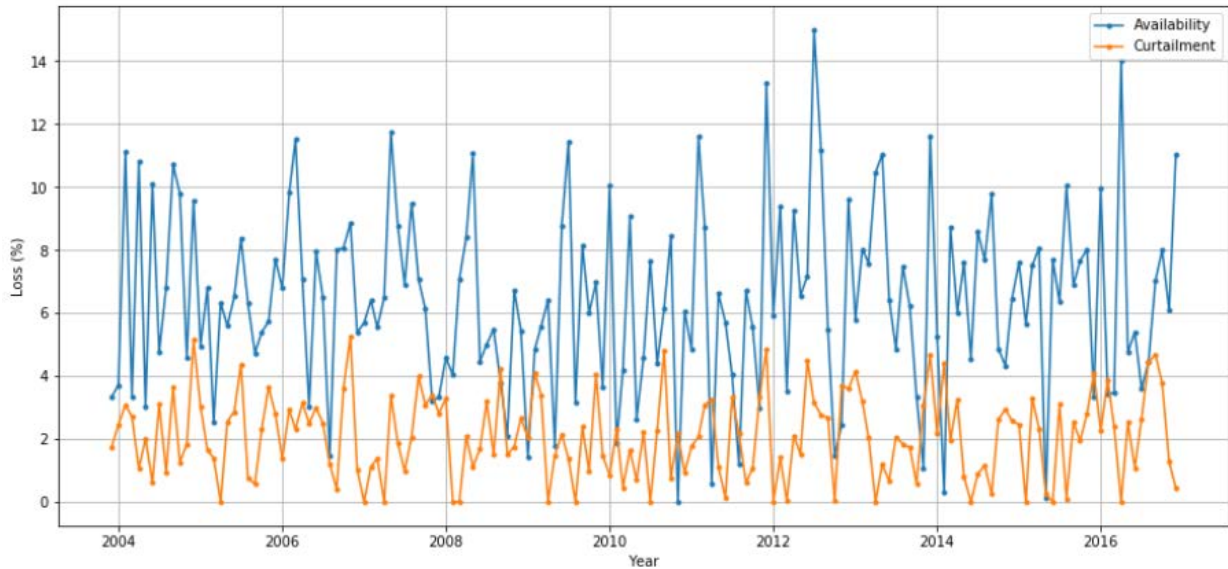


Figure 12. Example of monthly availability and curtailment loss during the plant period of record

- Step 4: Review regression relationship.
Next, linear regression plots of the monthly reported gross energy and monthly average wind speeds for each reanalysis product are plotted to check for reasonableness and to identify outliers (example shown in Figure 13). Before plotting, monthly gross energy values are normalized to 30 days to provide a more accurate regression relationship.

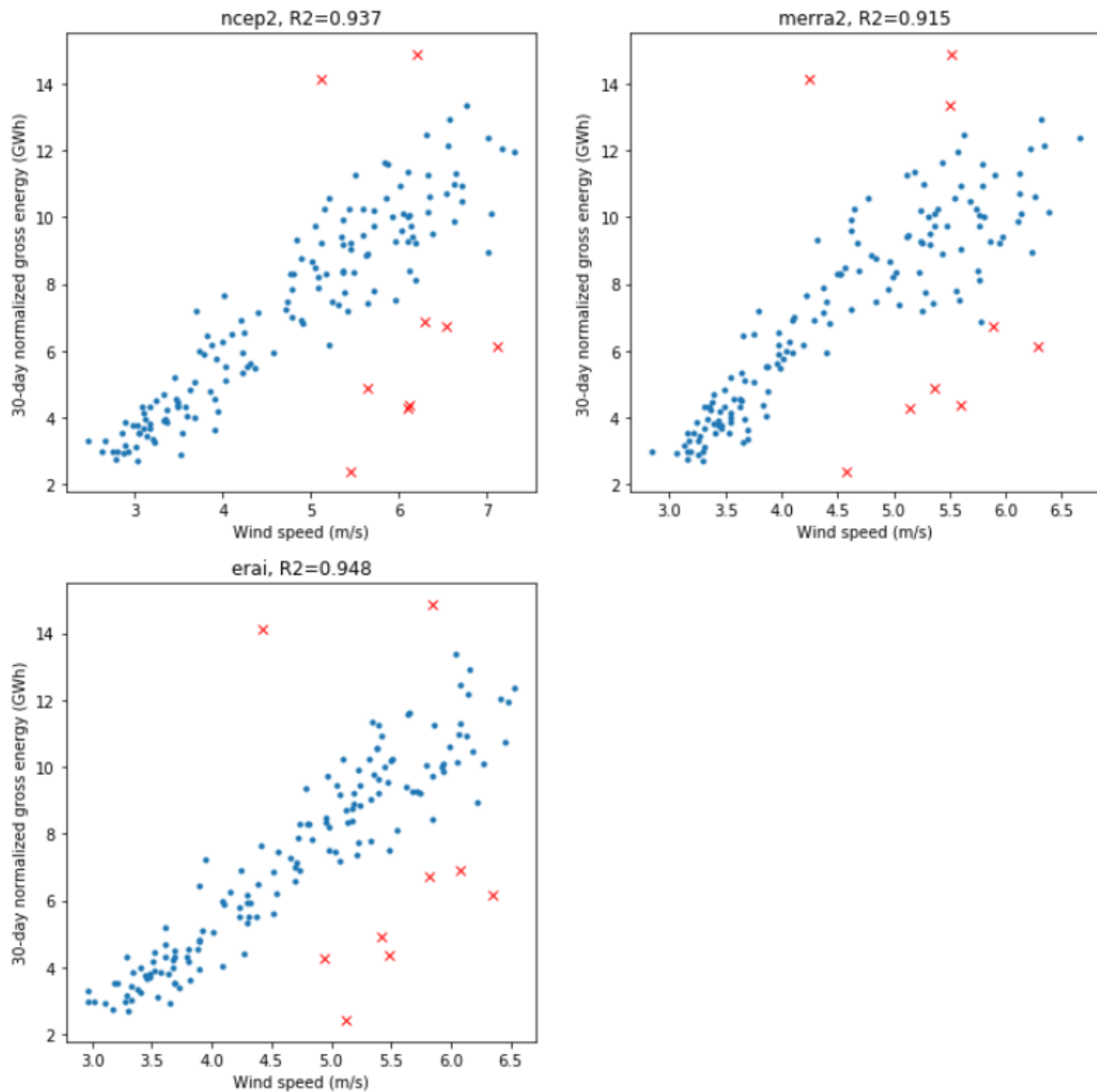


Figure 13. Scatterplots of gross energy and monthly averaged wind speeds for different reanalysis products. Outliers are identified as red crosses.

- Step 5: Run the Monte Carlo AEP calculation. Once data have been adequately reviewed and decisions on the input data have been made (i.e., choice of reanalysis products, removal of anomalous loss months from consideration of long-term loss calculation), the Monte Carlo AEP calculation is executed. A total of 50,000 iterations are considered. For each iteration, the following steps are performed:
 - First, randomly select the following:
 - Reanalysis product to be used
 - The combined availability and curtailment loss threshold
 - The outlier threshold in the Huber's t weight function within the robust linear regression (or the Huber's t criteria) (Huber and Ronchetti 2009)

- The number of years to use in the long-term correction.
- Normalize the gross energy data to 30 days.
- Perform the linear regression with the applied Huber's t criteria, and record the slope and intercept values, their standard errors, and the covariance between the slope and intercept.
- Based on the calculations in the previous step, create a joint distribution of possible slope and intercept values, and then randomly sample a pair of slope and intercept values from that distribution.
- Use the sampled slope and intercept values to calculate 30-day normalized gross energy during the long-term period of record (defined by the number of years selected in the first step).
- Calculate the monthly average 30-day normalized gross energy for each calendar month, convert it back to the regular number of days per month, and then sum to calculate the long-term gross energy.
- Remove the long-term availability losses to obtain the long-term AEP (noting that for this benchmark, long-term curtailment losses are not removed).

After completing the 50,000 simulations, a distribution of AEP values will result from which uncertainty can be reduced (example provided in Figure 14).

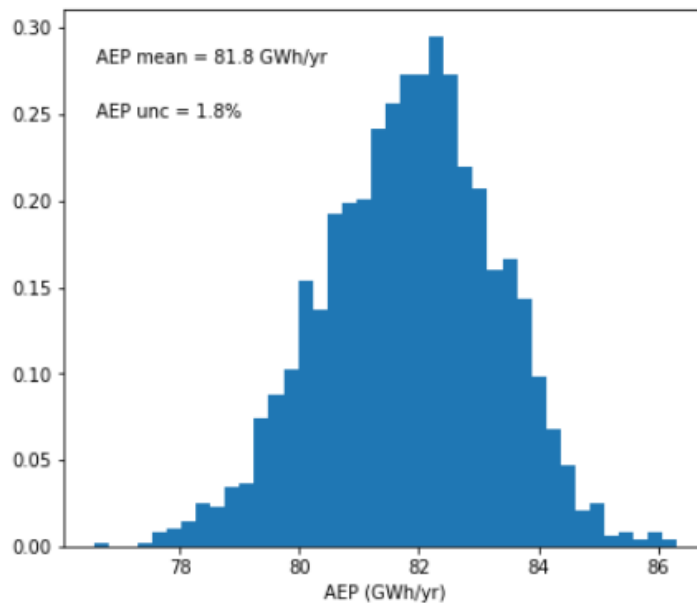


Figure 14. Long-term AEP distribution after 50,000 simulations

3 Work Packages

The WP3 project currently has three work packages defining the products being produced. Each work package used feedback from and interactions with NREL staff, PNNL DAP management, wind plant owners, OEMs, and wind resource consultants.

3.1 Historical Validation Study

Before embarking on the larger data collection effort required for the full benchmark, the NREL team conducted a baselining exercise to try to validate some of the industry-wide resource assessment bias estimates being provided by industry. The HVS compared monthly production data, obtained from the U.S. Energy Information Administration, with preconstruction P50 energy estimates collected from the owners of 56 projects in the United States. HVS results indicate significant bias between predicted and measured energy, even when controlling for factors such as commission date and grid curtailment. Results were highly dependent on the exogenous variables used during the analysis, which were accounted for with a sensitivity analysis over a range of reasonable values. For projects starting after 2011, which makes the analysis more representative of modern wind modeling capabilities, the calculated bias was $-5.5\% \pm 1.28\%$, with 1%–2% of that being attributable to curtailment (Lunacek et al. 2018). Figure 2 in the introduction summarizes the results from the HVS.

3.2 Benchmark Pilot

To kick-start the benchmarking effort and test the process for collecting and analyzing EYA submissions against operational data, NREL chose a pilot wind power plant project from one owner in the stakeholder group. This pilot project, coded as z01 for anonymity, was representative of Phase 1 projects and provided valuable insights into potential issues that could come up during the rollout of the full 10-project Phase 1 effort. NREL acquired preconstruction data from the owner-operator of the project and released the data to the consultants via the DAP. NREL also developed and released the initial attempt at the EYA collection form in Microsoft Excel, which was the starting point for and precursor to the EYA results collection web application described in Section 2.2.1. NREL used EYA submissions from a total of 10 consultants: ArcVera, DNV GL, EAPC, K2 Management, Luminate, Mott MacDonald, NRG Systems, WindGuard North America, Wood Group, and WSP (note that submissions from additional consultants were received after the analysis presented in this section was completed). Through a series of virtual and in-person meetings, NREL also received feedback from the consultants on potential process improvements.

3.2.1 Consultant Response Collection

For the pilot, the data request was rigorous, with the following specific requests captured in the spreadsheet response form (Figure 15):

- **Project summary:** a summary of the 20-year P50 total project losses and uncertainty
- **Energy losses:** project energy losses categorized into wake effect, availability, electrical, turbine performance, environmental, and curtailment/operational strategies
- **Uncertainty:** project uncertainty values associated with measurements, horizontal extrapolation, vertical extrapolation, plant performance, historical wind resource, and project evaluation period variability

- **Process description:** survey of each consultant’s in-house EYA process, including the choice of numerical models, analyst experience, etc.
- **Measurement details:** metadata associated with how each consultant processes and uses meteorological masts data
- **Measurement statistics:** temporal wind speed data used during the assessment process
- **Energy yield per turbine:** predicted average wind speed, wake loss, net energy yield per year, and turbulence experienced by each turbine of the project for 20 years.

Purpose: Categorizing and quantifying losses in details fosters the understanding of the WRA processes and the differences between consultancies				
Energy Yield Summary				
Wind Farm Name Plate			MW	
Gross Output			GWhr/yr	
#	Category	Loss %	Distribution type	Comments/Description/Inclusions/Exclusions
1	Wake effect			
1a	Internal wake effect			Wake effects internal to the wind plant
1b	External wake effect			Wake effects generated externally to the wind plant
1c	Future wake effect			Wake effects that will impact future energy projections based upon either confirmed or predicted new project development or decommissioning
1e	Other (please describe)			
2	Availability			
2a	Turbine availability			Turbine availability (energy-based), considering: warranted availability, non-contractual availability, restart after grid outage, site access, downtime (or speed) to energy ratio, first year/plant start-up availability
2b	Balance of plant availability			Availability of substation and collection system, other non-turbine availability, warranted availability, site access, first year/plant start-up availability
2c	Grid availability			Grid being outside grid connection agreement operational parameters, actual grid downtime, delays in restart after grid outages
2d	Other (please describe)			
3	Electrical			

	Parameters			
Distribution type	1	2	3	4
Normal	Mean	Standard deviation		
Normal Truncated	Mean	Standard deviation	Lower bound	Upper bound
Weibull	k	A	Alternatively: k and mean to find A	
Perturbed	Minimum	most likely	Maximum	
Triangle	Minimum	most likely	Maximum	

Figure 15. Screenshot from the Energy Yield Summary tab of the EYA results collection tool used for the pilot project. The consultants were asked to fill out the green cells.

3.2.2 Pilot Results

The pilot project was the first iteration of NREL’s benchmark methodology. NREL collected 10 consultant responses and compared them to the benchmark AEP as calculated from owner-provided operational data. As shown in Figure 16, 5 of the 10 consultants estimated net energy production higher than the AEP derived by NREL using operational data via OpenOA. Moreover, as shown in Figure 21, annual average per-turbine production varied widely across different consultants, which implies that despite some agreement at the plant-level AEP, there are still significant differences in the underlying methods. It appears that each consultant has “tuned” their various methods to ultimately output similar plant-level results, but the underlying calculations are inconsistent.

Figure 16, Figure 17, and Figure 18 show Level 1 results from the pilot, which represent the primary metrics of interest for the benchmark (e.g., the P50 bias and the total project uncertainty). Specifically, this includes plots for P50 gross and net bias, P50 bias comparison, and uncertainty comparison.

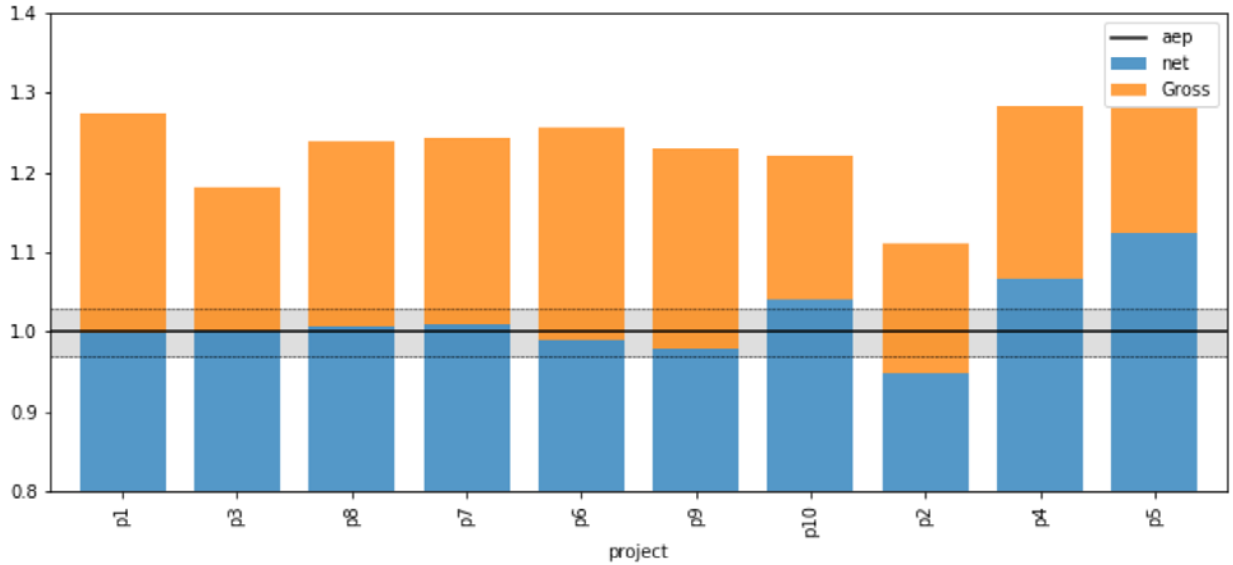


Figure 16. P50 gross and net bias. Consultant estimates for z01 capacity factor. The blue and orange bars represent gross and net energy production, respectively. “AEP” represents the NREL-calculated benchmark capacity factor (normalized to 1), and the shaded bar represents the uncertainty.

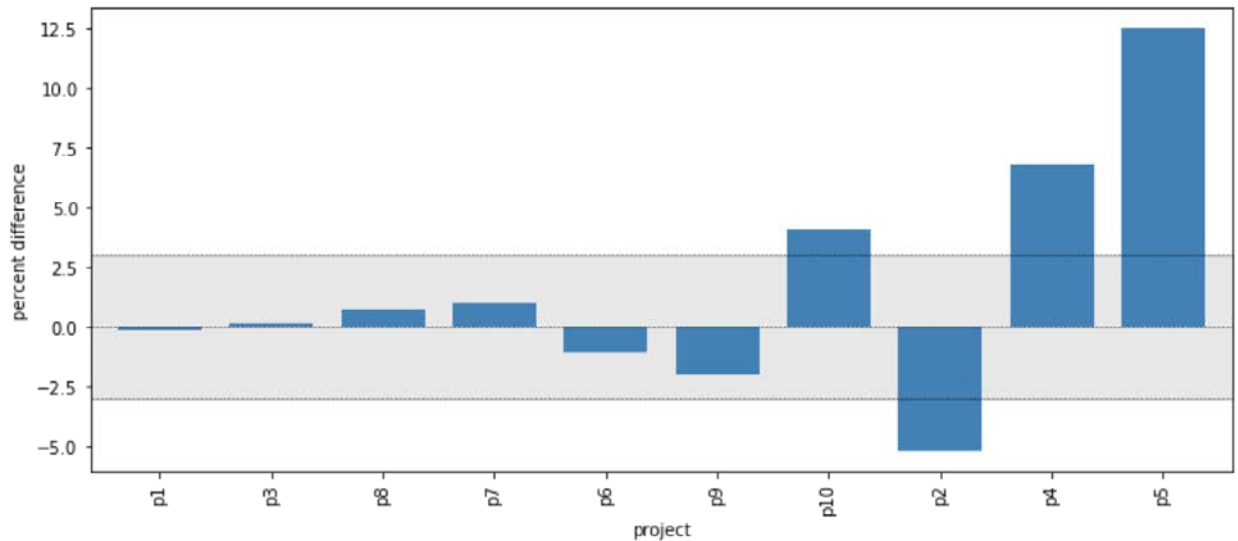


Figure 17. P50 bias comparison. Percentage difference between consultant P50 values and the NREL operational assessment. The zero horizontal line represents the NREL operational assessment value, and the grey band represents the uncertainty.

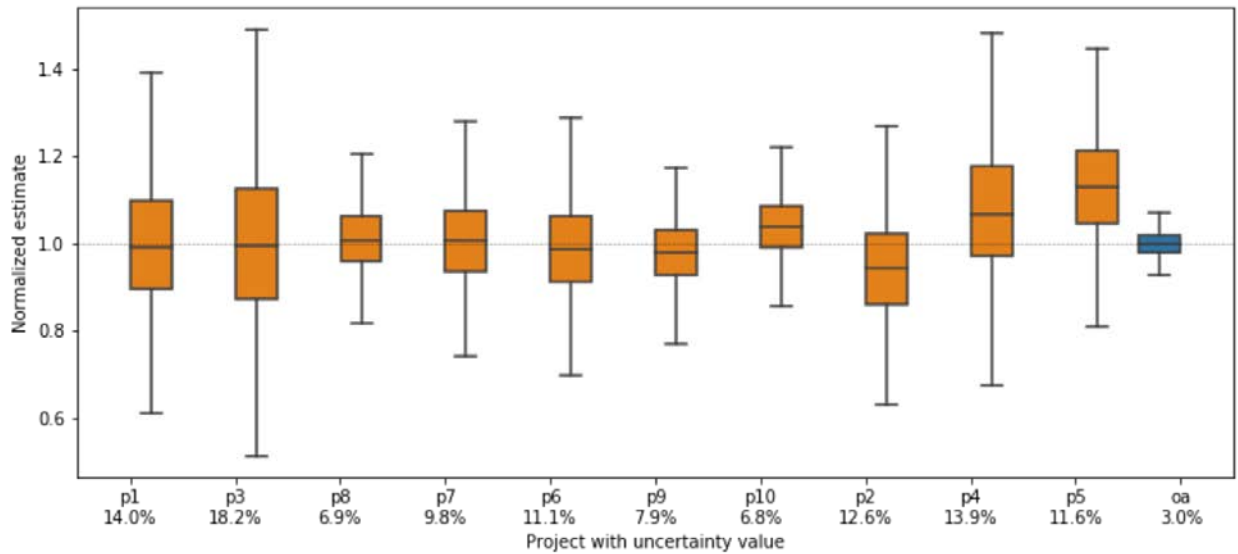


Figure 18. Uncertainty comparison: consultant estimates for z01 uncertainty. The orange box plots represent net energy production and the uncertainty predicted by consultants. The blue shaded box plot represents the NREL-calculated benchmark capacity factor.

To take a closer look at the underlying causes behind the Level 1 pilot results, NREL also conducted a more granular Level 2 analysis, which represents secondary metrics of interest for the benchmark (e.g., categorical energy losses and uncertainties). These are metrics that will aid in understanding why the Level 1 metrics might disagree from each other and the calculated NREL operational assessment values. Specifically, this includes plots for categorical loss comparison (Figure 19), categorical uncertainty comparison (Figure 20), turbine-level capacity factor comparison (Figure 21), wake deficit comparison (Figure 22), and wind resource comparison (Figure 22).

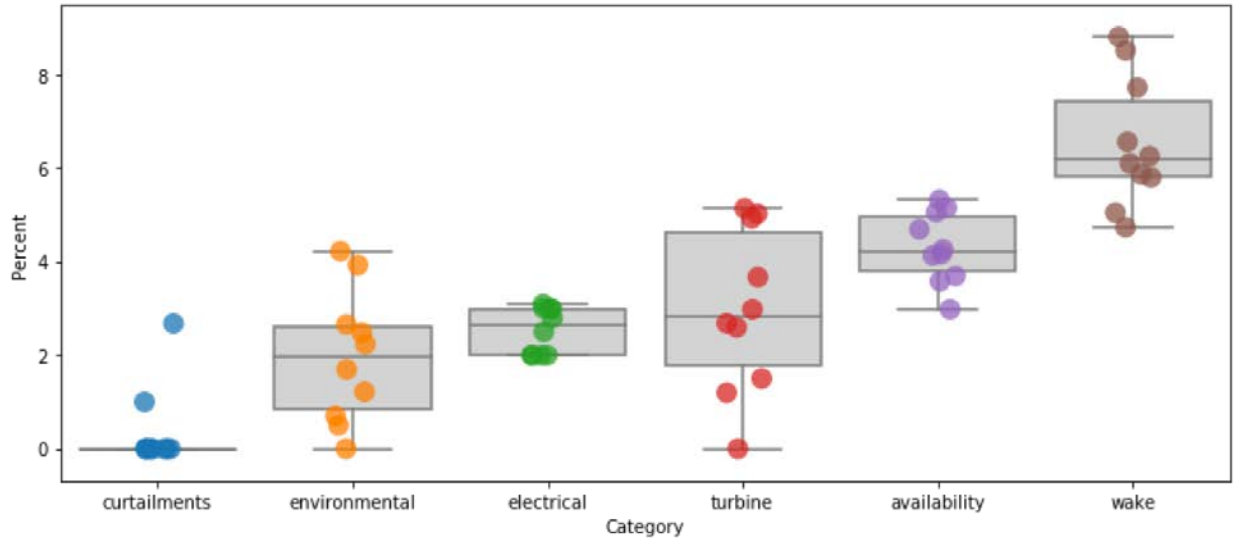


Figure 19. Consultant-estimated losses by category for z01. Each box plot distribution represents 10 different consultant estimates for z01 losses.

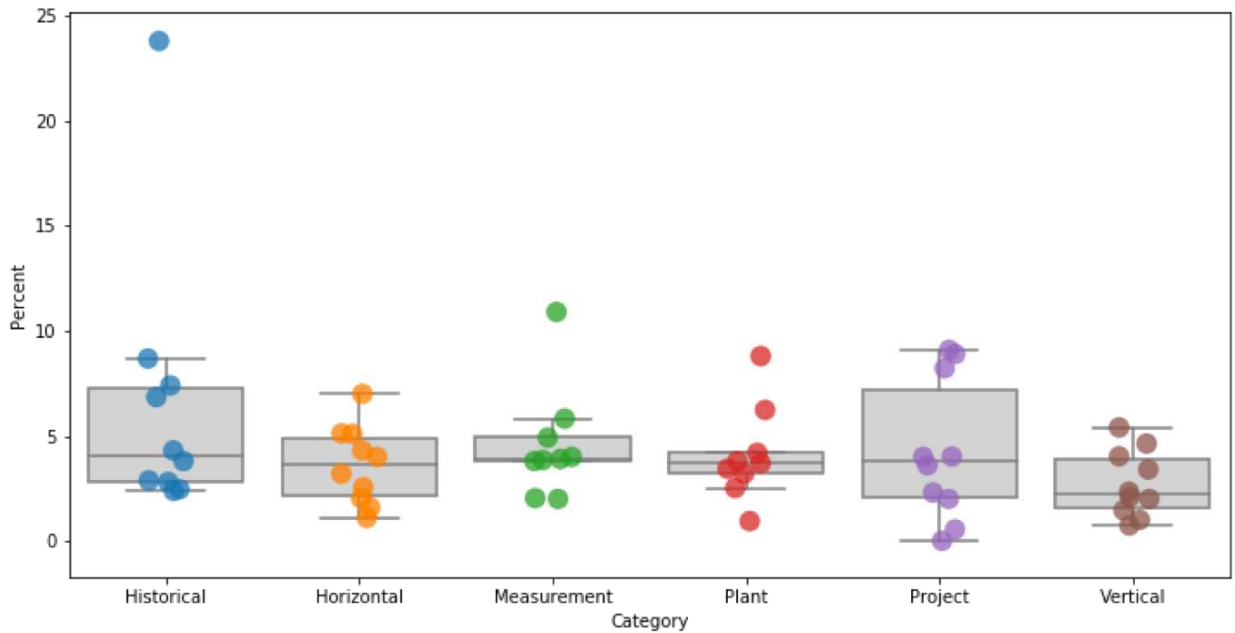


Figure 20. Consultant estimates for z01 uncertainty by category. Each box plot distribution represents 10 different consultant estimates for z01 uncertainty.

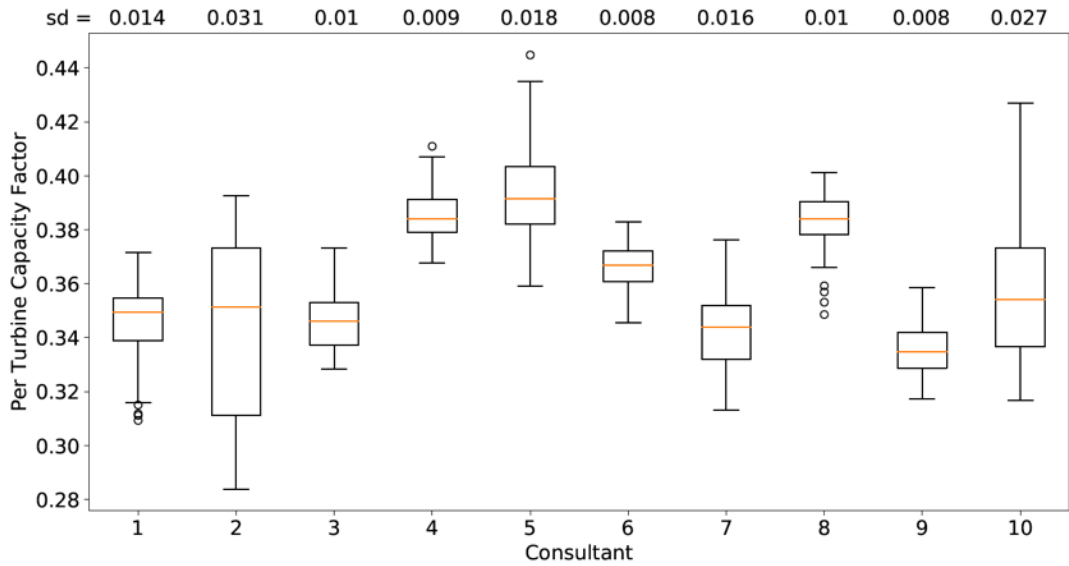


Figure 21. Consultant estimates for per-turbine capacity factor for z01. Each box plot distribution represents 100 turbines, with standard deviation displayed at the top of the plot.

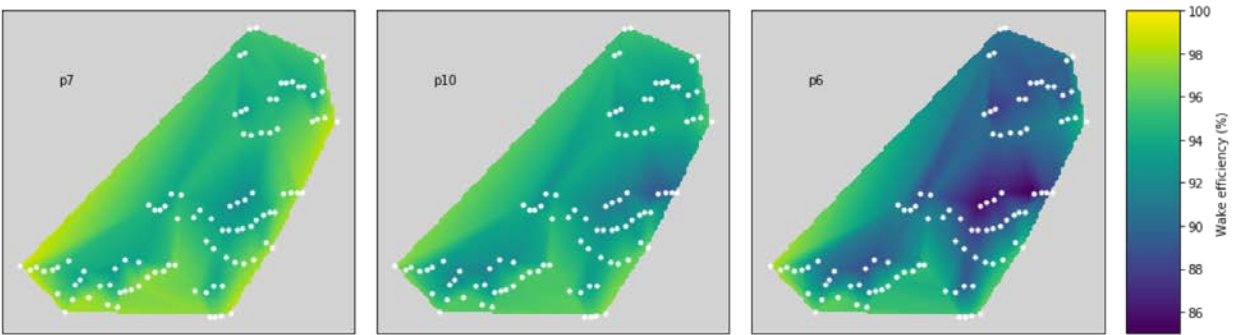


Figure 22. z01 Wake deficit analysis of the predicted wakes from three consultants

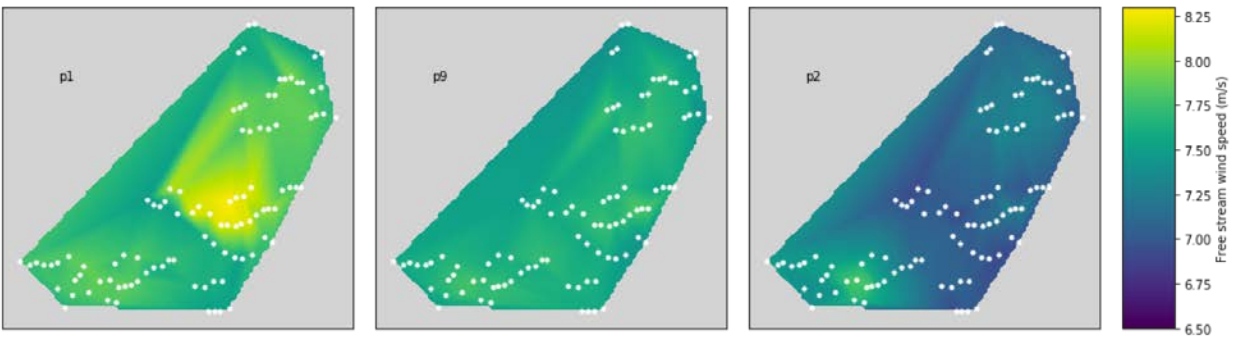


Figure 23. z01 resource analysis of the turbine-level annual average free-stream winds from three consultants

3.2.3 Lessons Learned

The pilot project provided important lessons learned for benchmarking Phase 1 and offered an excellent opportunity to revise the EYA results-submission process. The most important takeaway was the need for more effective and engaged communication with the consultants and

owners to reduce obstacles to results submissions and to increase overall participation and timeliness.

The spreadsheet collection method ran into several challenges, and the pilot project made it clear that Phase 1 would need a better solution for collecting consultant responses. Filling out the spreadsheet was time-consuming for the consultants, the Excel files after consultants' editing were difficult to parse, data validation was difficult to enforce, and version control of the responses was challenging. Particularly, using an Excel-based form resulted in consultants modifying tables, data structures, and data fields in the provided template. These modifications were extremely common and led to problems in data post-processing and analysis. In some cases, NREL could not derive meaningful conclusions with statistical significance. To address these data collection challenges, the WP3 team developed the EYA collection application described in Section 2.2.1 to improve uniformity, ease of use, validation, and version control.

The pilot project also revealed a need to improve the metadata documentation for preconstruction and operational data. As a result, for Phase 1, NREL prepared project overview documents consisting of metadata such as wind plant nameplate capacity, turbine layout, and meteorological tower records. Each project overview document was included with the preconstruction energy data provided to the consultants in Phase 1.

3.3 WP3 Phase 1

Phase 1 represents the formal launch of the WP3 project and includes the data from 10 operational wind power plants. Each consultant conducts one EYA per project, resulting in 10 submissions from each consultant. Upon submitting all 10 EYAs to NREL and completing Phase 1 of the benchmark exercise, consultants receive operational data for all 10 projects to revise their methods and to perform their own research activities.

3.3.1 Phase 1 Setup

For Phase 1 of WP3, NREL selected projects that mirror expected near-term future deployments in the U.S. wind belt. In this region, NREL selected projects at random after filtering for specific characteristics (Section 3.3.2). With a small set of 10 projects for Phase 1, defining a project group with similar features will allow for greater statistical significance; therefore, as a key theme of this initial phase of WP3, in Phase 1 NREL chose wind projects with relatively simple terrain. Tackling this straightforward group of projects sets up the WP3 collaborative to pursue more complex “edge” cases in subsequent project phases.

NREL used the Regional Energy Deployment System (ReEDS) Standard Scenarios Results Viewer (Cole et al. 2019) to examine modeled deployments from 2016–2020. We ranked them in descending order of installed capacity, then looked at states installing at least 1 GW of wind in this time frame. As expected, most deployment is expected in the middle of the United States, in the so-called “wind belt” states (Table 1). Figure 24 shows these states on a U.S. map, where the modeled near-term deployment is indeed concentrated within the central United States.

Table 1. Top Wind Installation States from 2016–2020 Modeled in ReEDS with Mid-Case Scenario

State	Installed Capacity (GW)	% of Total Installations	Cumulative % Total Installations
Texas	27.29	42%	42%
Wyoming	6.50	10%	52%
Oklahoma	5.35	8%	60%
Iowa	4.14	6%	66%
North Dakota	3.42	5%	71%
Montana	3.31	5%	76%
South Dakota	2.29	3%	80%
Minnesota	1.85	3%	83%
Indiana	1.61	2%	85%
Ohio	1.57	2%	87%
Kansas	1.27	2%	89%
Remaining States	6.93	11%	100%
TOTAL	65.5		

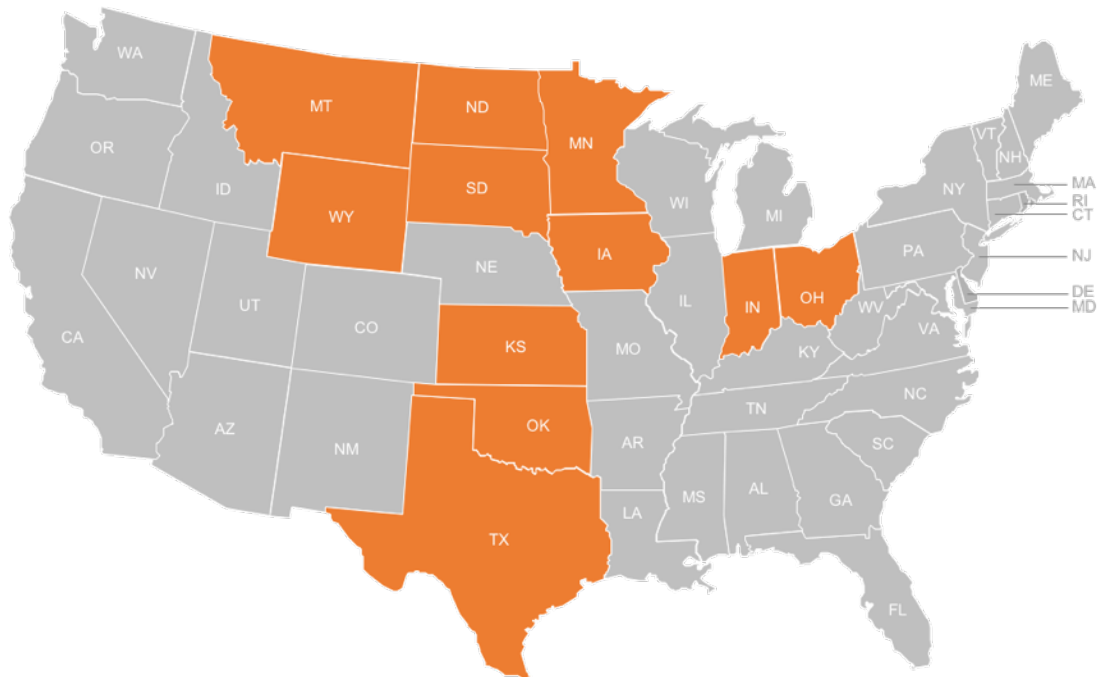


Figure 24. Predicted near-term U.S. wind plant deployments used when selecting Phase 1 projects

3.3.2 Project Characteristics

The following characteristics and associated criteria represent the next level of filtering for the Phase 1 project list.

- **Wind plant characteristic filtering:** to represent a diversity of wind plant configurations, including plant size, array design, and a representative number of measured production years.
 - Size: 100 MW or greater
 - Hub height: At least 80 meters aboveground
 - Turbine size: 1.5 MW or greater
 - Operation: minimum 3 years of operation
 - Geography: Central United States or simple terrain
 - Curtailment: 5% or less (alternatively remove data periods that are above the threshold)
 - Availability: 95% or greater (alternatively remove data periods that are below the threshold).
- **Owner/operator diversity:** to represent several operation-and-maintenances approaches and to avoid overrepresenting specific operators/owners.
 - Operations: documented operation-and-maintenance strategy including any upgrades or modifications
 - Curtailment: minimum monthly curtailment values; preferred is hourly or finer temporal resolution of curtailment signals from utilities.
- **Turbine manufacturer diversity:** to represent a diversity of turbine technologies and avoid overrepresenting specific turbine manufacturers.
 - Turbine OEM: active OEMs capable of supporting warranties, operations, and maintenance. This eliminates defunct OEMs from consideration.

4 WP3 Phase 1 Analysis

NREL and the WP3 benchmark team have successfully completed Phase 1 of the benchmark. All 10 Phase 1 wind plants have undergone operational analysis by NREL, and 70 of a possible 90 EYA submissions have been received from 9 participating consultants. Eight consultants have completed at least half of the project submissions, constituting 68 submissions, and have been included in the Phase 1 analysis discussed in this report. With this data set, NREL can compare EYA submissions with the operational data to establish a benchmark for energy-estimation performance and can begin exploring the underlying reasons for error and uncertainty. A detailed explanation of the operational assessment methodology is provided in Section 2.3.4, and the resulting comparison of EYA data to operational data is included in the following sections. Note that the consultant EYA submissions were received during 2018 and 2019; some consultants' EYA methods may have changed since then. Similarly, the preconstruction wind measurements used for wind resource assessment in this study do not necessarily represent current best practices in the wind industry.

The Phase 1 analysis includes the investigation and comparison of the P50 net energy estimate (Section 4.1), gross energy (Section 4.2), energy losses (Section 4.3), and prediction uncertainty (Section 4.4), with more detailed analysis of spatial variability, modeling approaches, project parameters, and examination of other metadata planned for future work.

4.1 P50 Net Energy Estimate

P50 represents the central estimate of the predicted energy distribution. This can be thought of as the long-term median energy production or the most likely outcome of the consultant predictions. Next, we compare the consultant EYA to the measured AEP as created by the OpenOA analysis. The total distribution of the differences between the estimated P50 and the measured AEP is shown in Figure 25 for the WP3 Phase 1 sample. This plot represents the 10 projects in the Phase 1 sample and might not be representative of total industry prediction capability. Industry-wide analyses, such as the NREL HVS discussed in Section 3.1, are intended to represent actual industry performance. The average bias of our sample indicates underprediction but is still relatively small, at -1.2%; however, the standard deviation of the P50 bias remains high, at 4.8%.

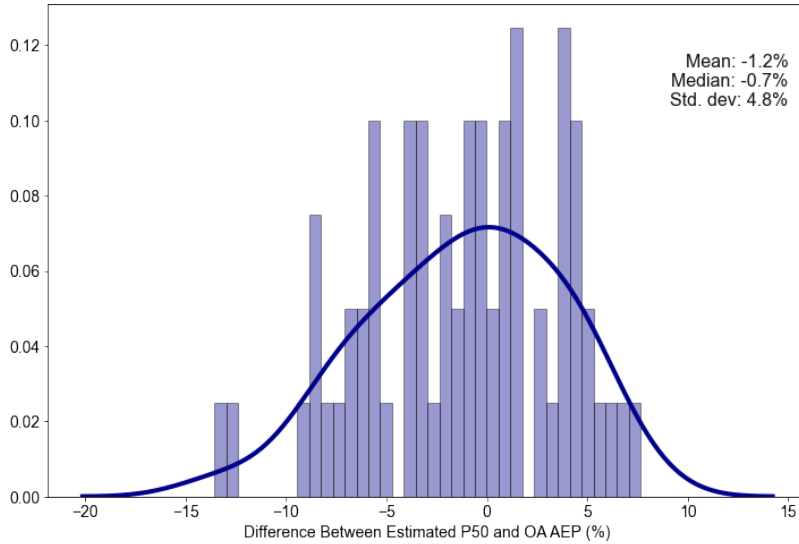


Figure 25. P50 bias distribution for WP3 Phase 1. The positive values illustrate overprediction of AEP.

For the 10 Phase 1 projects, on average, consultants were largely successful at estimating energy performance. But given the near 5% margin of error, it is still likely that any given consultant estimate for any given project could have been drastically different from the measured operational data. It is clear that a deep level of analysis is justified to begin to understand the underlying differences between consultant responses and predictions. In particular, an in-depth analysis approach is required to truly understand P50 performance and the implications of estimation methodologies across different industry players and project types. The breakdown of P50 bias by consultant and by project (Figure 26) reveals that P50 bias estimates tend to vary more between projects than between individual consultants, although certain consultants tend to consistently overpredict or underpredict.

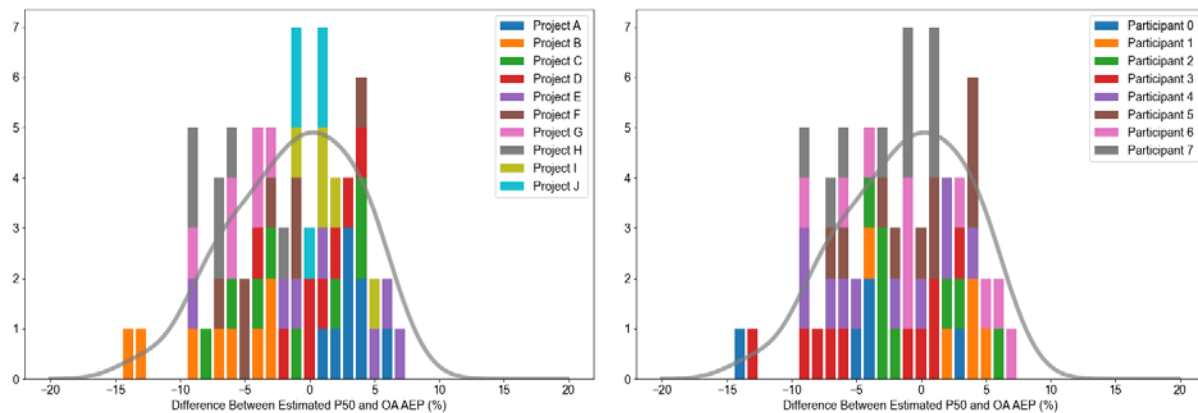


Figure 26. P50 bias distribution for WP3 Phase 1 color-coded by project and participant

Figure 27 shows how the P50 bias differs between each individual project and by participant. This demonstrates a systematic tendency toward over- or underprediction relative to the operational assessment values on the project level. For example, Project A is unanimously overpredicted, whereas Project B is unanimously underpredicted. Moreover, the ranges in the prediction errors among consultants for each project are broad. Project B, for example, displays

that the consultant’s range of underprediction errors is from 3%–14%, whereas the range in prediction errors in Project E spans from approximately 8% underprediction to 8% overprediction.

One key finding of this work is that the distribution of the P50 bias varies more by project than by consultant. This is clearly evident when comparing the left and right panels in Figure 27, where the ranges in P50 bias distributions are very different across the individual projects (left panel) and across participants (right panel). That is, P50 bias estimates tend to vary more between projects than between individual consultants, although certain consultants tend to consistently over- or underpredict the AEP.

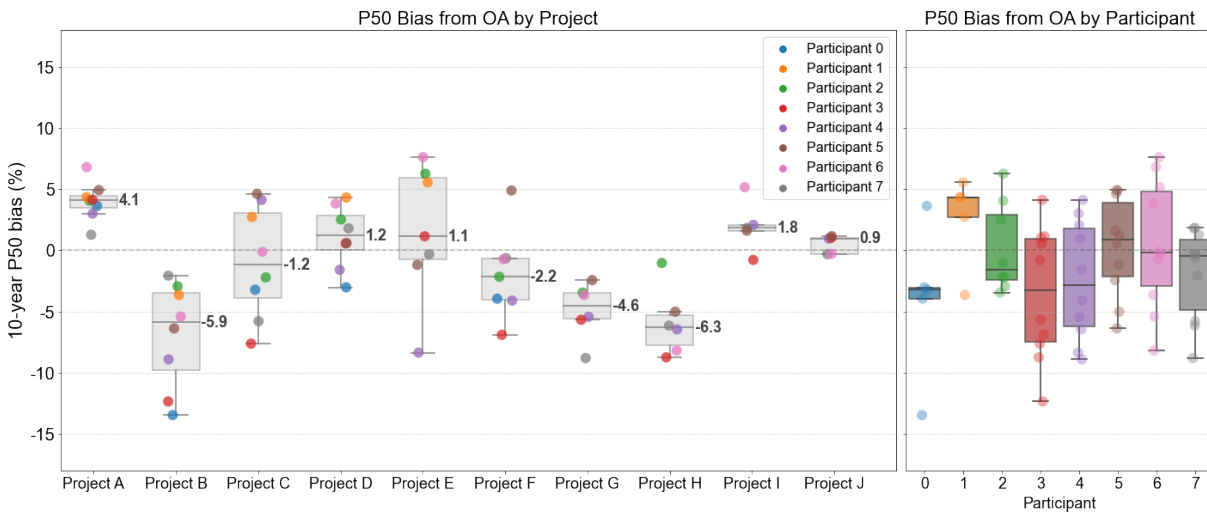


Figure 27. P50 bias box plot distributions for WP3 Phase 1 by project and by consultant

4.2 Gross Energy

Gross energy is defined as the total energy available to the wind plant from the wind resource alone. This can also be thought of as an ideal energy before losses are subtracted. The WP3 benchmark definition of gross energy excludes all losses, such as wakes, although gross energy does include the natural variation of the wind resource at the site. A valid comparison to measurements is not yet available for gross energy estimates, so we simply compare the gross energy estimated by consultants within projects. This comparison reveals that the total energy estimated to be available to the wind plant before physical losses are considered varies widely from 5%–10% between consultants; however, some of the variability could be caused by slight discrepancies in how consultants define gross energy (i.e., whether any losses are included). Figure 28 shows how these variations in gross energy among consultants for each project, combined with estimates of different types of losses, lead to a consultant’s P50 prediction, where the units of energy are normalized by the operational assessment of P50.

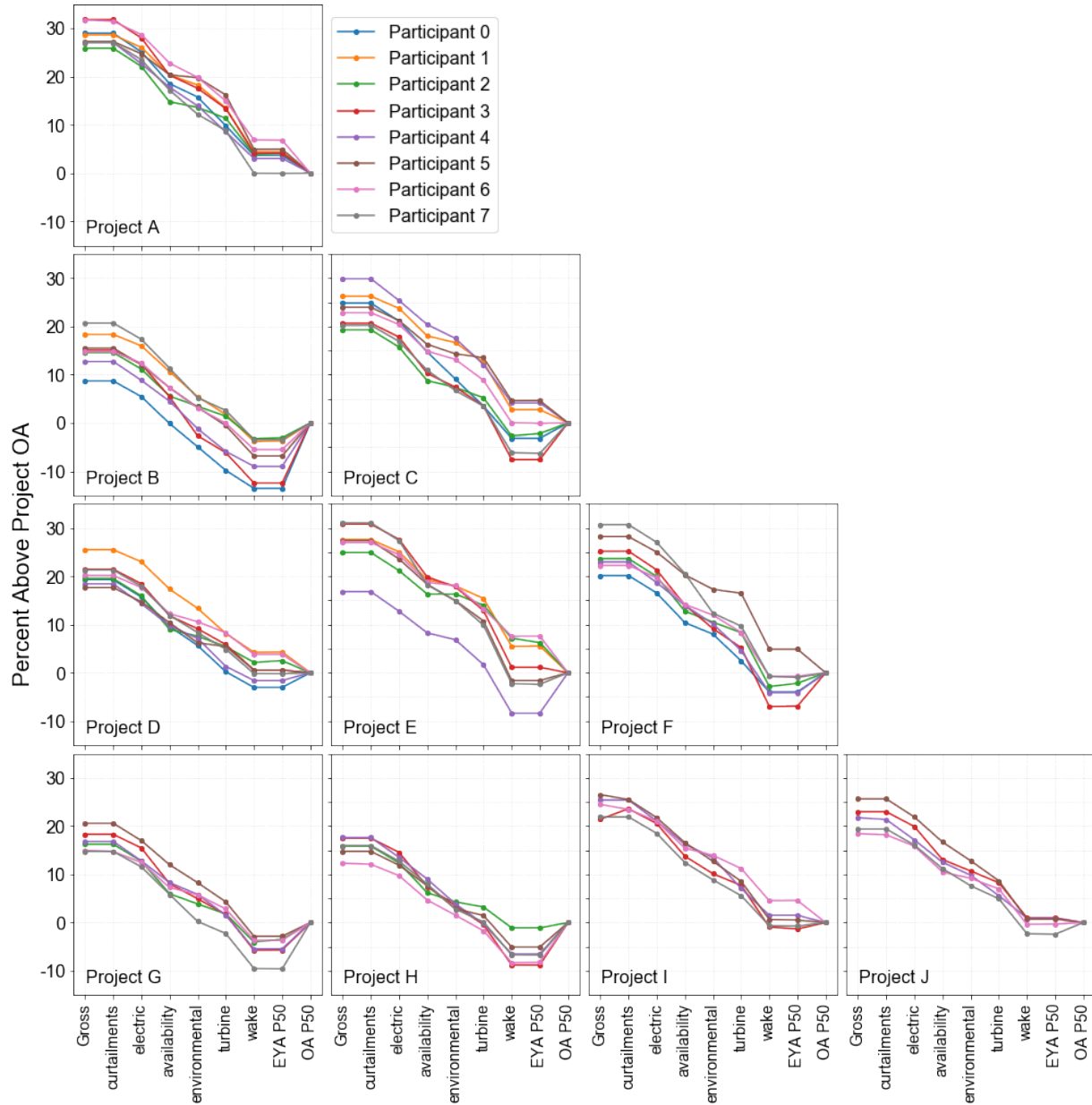


Figure 28. P50 waterfall chart for WP3 Phase 1 by project and by consultant

4.3 Energy Losses

Energy losses are the physical energy sinks that subtract from the gross energy to arrive at the P50 net energy estimate. Figure 28 demonstrates how aggregating the estimates of individual loss categories leads to a consultant’s EYA P50. This work examines variability in the total combined loss (the difference from gross energy to net energy) as well as the variability of each individual loss category level. The WP3 benchmark process used the proposed IEC 61400-15 energy loss framework to define the standard loss categories (Lee and Fields 2021; Fields and Sherwin 2018; Filippelli et al. 2018). These include the following major categories and corresponding subcategories (note that we do not investigate loss estimates at the subcategory level in this report):

- Wake effect (internal wake effects, external wake effects, future wake effects)
- Availability (turbine availability, balance-of-plant availability, grid availability)
- Electrical (electrical efficiency, facility parasitic consumption)
- Environmental (icing, performance degradation, environmental loss, exposure)
- Turbine performance (suboptimal performance, generic power curve adjustment, site-specific power curve adjustment, high-wind hysteresis)
- Curtailment/operational strategies (load curtailment, grid curtailment, environmental/permit curtailment, operational strategies).

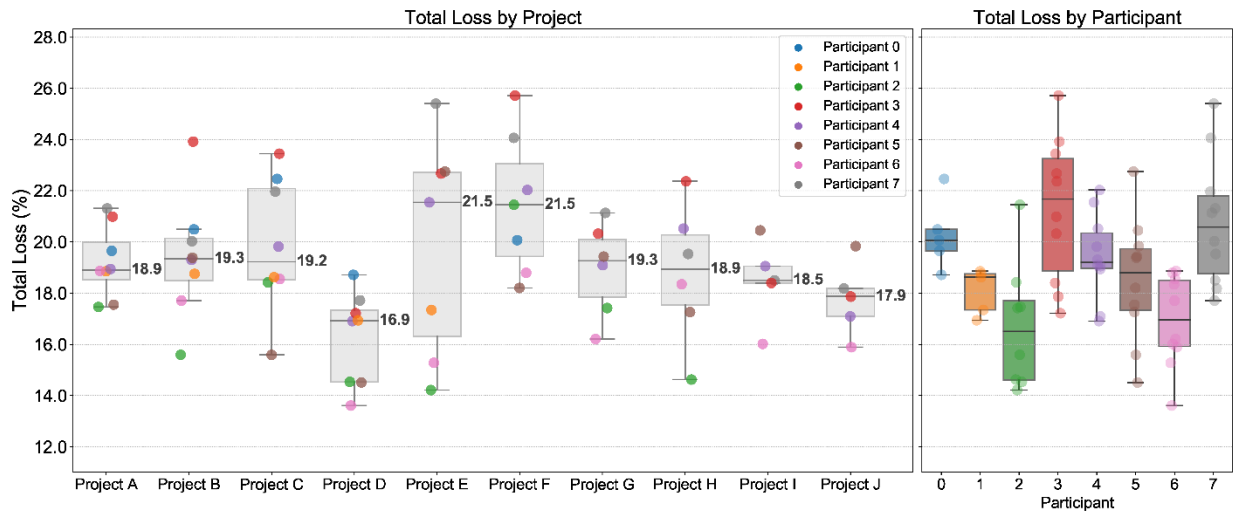


Figure 29. Total loss box plot distributions for WP3 Phase 1 by project

Figure 29 shows the variation in total losses by project. The variability within projects is considerably large, with some projects varying by greater than 10% from lowest to highest consultant estimate. Despite the large variability within projects, the median loss estimates are similar between projects, with a tendency for the interquartile ranges of each project’s loss estimates to overlap. There is also considerable variability in the total loss estimates for each consultant, where a consultant’s loss estimates might vary anywhere from 3%–11% across the range of projects. This range in variability could be caused by the degree to which consultants vary their estimates of individual loss categories from one project to the next.

The categorical losses also exhibit considerable variability within each specific loss category (Figure 30). Figure 30 represents all loss estimates from all projects and consultants in the WP3 Phase 1 sample, expressed as the difference between the individual estimates and the corresponding average categorical loss values for each project. By subtracting the project-level average losses from the individual estimates, we attempt to remove the impact of interproject variability from the assessment of consultant agreement (i.e., although categorical losses could be significantly higher for some projects than others, we are interested in the degree of consultant agreement within each project). We consider the level of consultant agreement for each loss category by examining the interquartile range (IQR) of the estimates for each category. In particular, we see that turbine performance (IQR=1.75%), environmental (IQR=1.62%), and wake losses (IQR=1.56%) have the largest interquartile ranges—suggesting that these loss categories are among the largest sources of consultant disagreement—followed by availability (IQR=1.12%) and electrical losses (IQR=0.71%). Note that the order of the loss categories

arranged by IQR does not match the ranking when sorted by the median loss estimate across all projects (see Figure 30). For example, the median wake loss estimate (6.75%) is roughly twice as large as the median turbine performance loss estimate (3.35%), but the IQR values suggest that turbine performance losses are a comparable source of disagreement. Additionally, although not investigated here, the level of consultant agreement may vary across projects. Further research is needed to understand the impact of the disagreement for each loss category on the overall energy prediction variability.

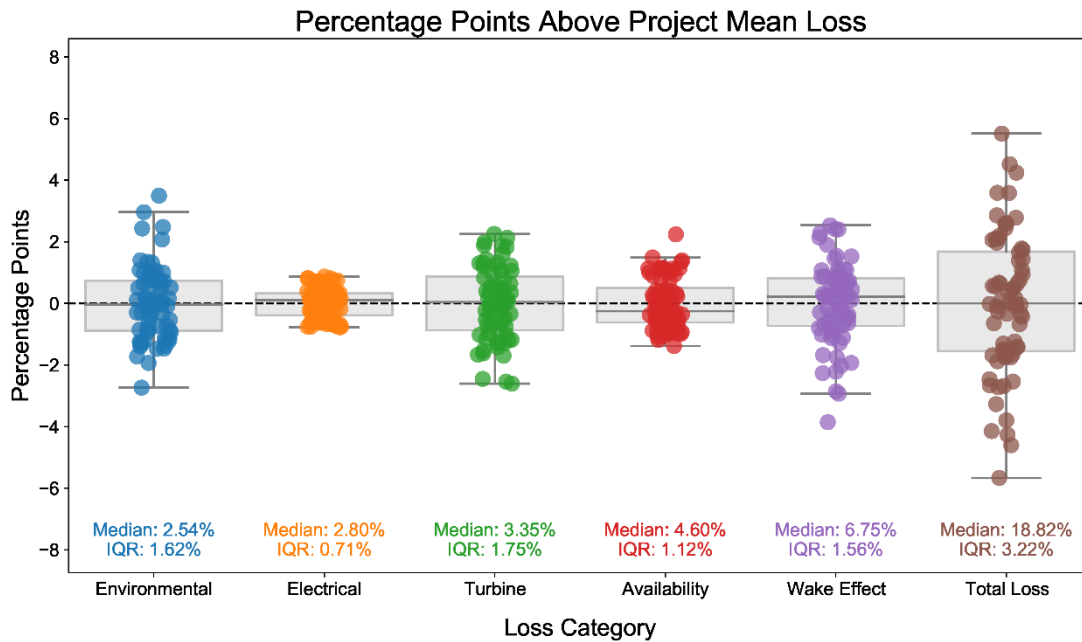


Figure 30. IEC 61400-15 loss category box plot distributions for WP3 Phase 1 by project. Curtailment loss estimates are not included here because they were not considered as part of the WP3 analysis.

4.4 Prediction Uncertainty

The energy prediction uncertainty for an EYA is a critical indicator of potential risk for investors. This analysis examines both the total combined uncertainty on the P50 net energy and the variability at each uncertainty category level. The WP3 benchmark process used the proposed IEC 61400-15 energy prediction uncertainty framework to define standard uncertainty categories (Lee and Fields 2021; Fields and Sherwin 2018; Filippelli et al. 2018). These include the following major categories:

- Historical wind resource
- Project evaluation period variability
- Measurement uncertainty
- Vertical extrapolation
- Horizontal extrapolation
- Plant performance.

Figure 31 compares the 10-year standard deviation of the energy predictions for the WP3 Phase 1 sample. There is significant spread in the total estimated uncertainties for most projects in the Phase 1 sample, with ranges in uncertainty predictions up to 10%.

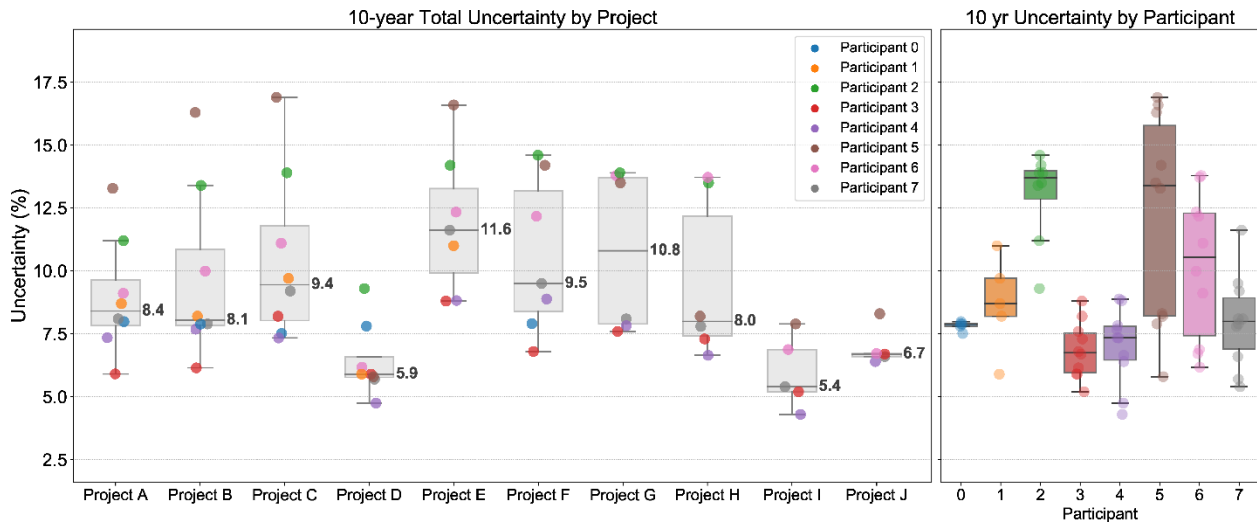


Figure 31. Total energy prediction uncertainty (10-year standard deviation) box plot distributions for WP3 Phase 1 by project and by participant

But there exists a much lower spread in each consultant’s individual estimates of total project uncertainty (with the exception of Participant 5). This is likely driven by the consultants’ unique, internal uncertainty estimation methods or training data sets. The large spread in uncertainty values for Participant 5 could be caused by multiple analysts within the consultant group contributing EYA submissions. Further examination of the categorical uncertainties as shown in Figure 32 also demonstrates substantial variability for specific uncertainty categories. Figure 32 represents all uncertainty estimates from all projects and consultants in the WP3 Phase 1 sample. All uncertainties plotted in Figure 32 are presented as the difference between the individual uncertainty estimates and the average categorical uncertainty values for the corresponding project. Similar to Section 4.3, we use interquartile ranges to indicate the level of consultant agreement for individual uncertainty categories. In general, broader disagreement in uncertainties is observed compared to the loss categories (Figure 30), with interquartile ranges up to 2.61%. We also see less clear signals about which uncertainties are the greatest sources of disagreement; IQR values range from 1.57% for horizontal extrapolation to 2.61% for project evaluation period uncertainty, with the other uncertainty categories roughly evenly spaced throughout this range. Note that further research is needed to draw conclusions about the relative impact of each uncertainty category on the variability of the total energy prediction uncertainty estimates.

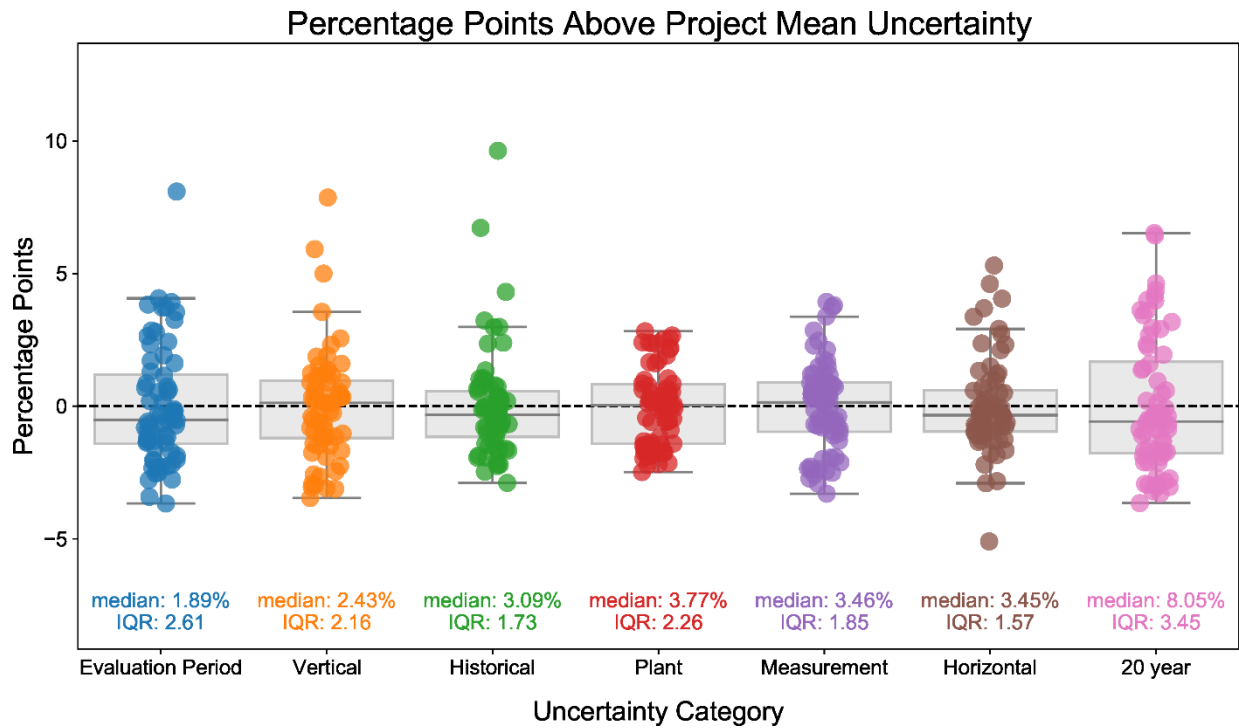


Figure 32. IEC 61400-15 uncertainty category box plot distributions for WP3 Phase 1

4.5 Energy Yield Assessment Performance Evaluation

The total accuracy or performance of an EYA is still an ongoing area of research and discussion, and we discussed several metrics in the previous sections that encapsulate the interests of a variety of stakeholders, including project owners and investors. This section provides an overview comparison of the predicted energy distributions and the total bias. Figure 33 and Figure 34 demonstrate both the predicted P50 net energy and the uncertainty relative to the OpenOA operational energy and uncertainty. These analyses visually demonstrate the large variability in both P50 energy levels and uncertainties. In particular, the average P50 prediction biases for each consultant range from -4%–+2.6%, and the associated uncertainties of the P50 error span from 3.7%–6.1% (Table 2).

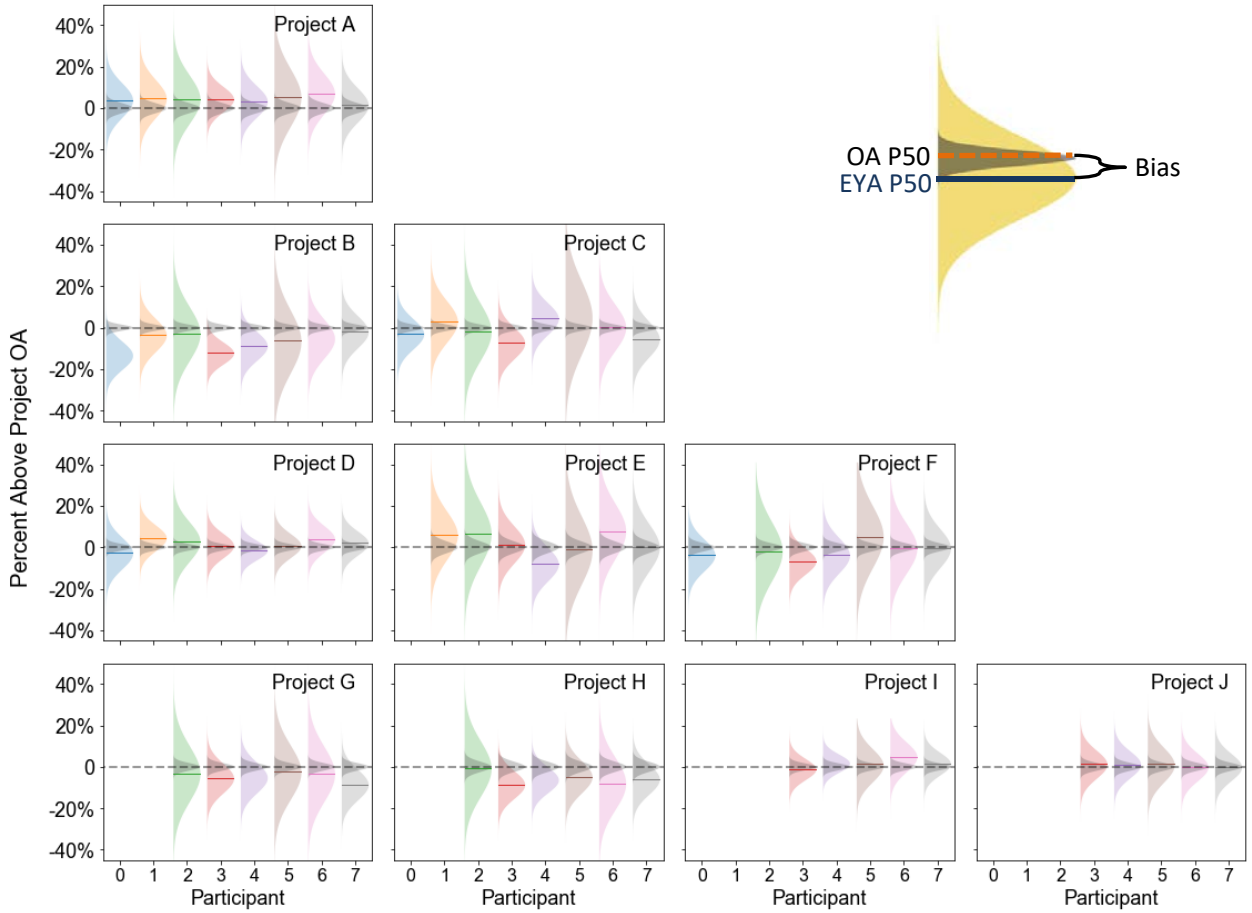


Figure 33. WP3 Phase 1 comparison of predicted EYA to operational assessment by project

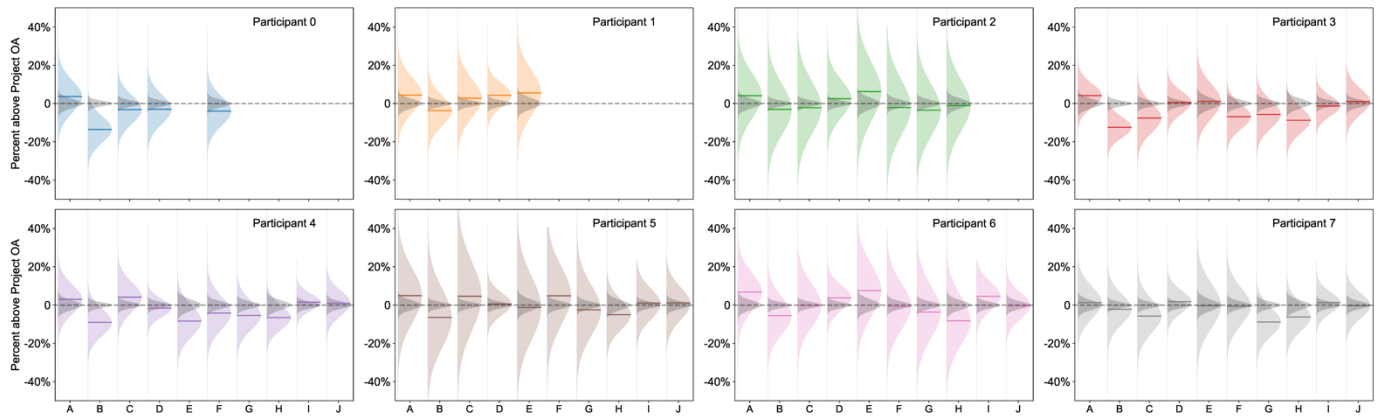


Figure 34. WP3 Phase 1 comparison of predicted EYA to operational assessment by participant

Table 2. P50 Bias Performance by Participant

Participant	P50 Net/OA Bias Mean (%)	P50 Net/OA Bias Std. (%)	Mean Absolute Error (%)	Projects Completed
0	-4.02	6.11	5.46	5
1	2.64	3.66	4.11	5
2	0.11	3.64	3.08	8
3	-3.53	5.41	4.89	10
4	-2.48	4.83	4.50	10
5	0.26	4.03	3.27	10
6	0.48	5.30	4.18	10
7	-1.93	3.73	2.90	10

5 Lessons Learned

Throughout Phase 1 of the WP3 benchmark, several lessons were learned that will positively impact future initiatives. First and foremost, we learned that massive data sharing initiatives are possible within the wind industry (Lee et al. 2020). This was not completely certain at the outset of this project, considering that the benchmark required asking all stakeholders involved to share their proprietary and extremely sensitive data. The initiative was made possible by leveraging DOE laboratories as a neutral third party. The project has ultimately been a success and can be used as a model for other data sharing efforts. There were also many specific lessons learned throughout Phase 1—enough so that it is possible to divide them into categories.

5.1 Data Access

“Data access” refers to the ability to define, collect, and aggregate the various types of preconstruction and operational data required to develop a credible benchmark. Preconstruction data were collected from owners and distributed to consultants; operational data were collected from owners and used by NREL to develop operational assessment benchmarks; metadata and process data were collected from stakeholders. Specific data access lessons from Phase 1 include:

- Wind plant preconstruction and operational data are nonuniform and not standardized across different wind plants and owners, making them difficult to aggregate. For example, the variable names and units differ among project owners. This, in turn, made it challenging to develop uniform methods for summarization and analysis. Careful thought should be given to the data collection procedure, such as the creation of a data collection application.
- Defining and making data requests as early as possible can help overcome data collection challenges because it helps owners assess their internal capabilities and resources. Owners might not have ready or easy access to all required data, and they might not know this is the case until they try to collect them. Different wind plants might have different methods for storing data, even within an individual owner’s project portfolio. Sharing data of jointly owned projects that involve multiple investors and owners can also be a challenge.
- Wind plant selection criteria should include data accessibility and data quality as a critical component to overcome the challenges of collecting useable data. In the future, wind plants where there is high confidence in data quality and accessibility should be prioritized for selection in benchmark efforts.
- Similarly, the high variance among owners or projects for operational data is a challenge that will require more diligence and planning in the future regarding the data quality, data availability, and data temporal resolution.
- Application programming interface (API) access to operational data would be highly beneficial. This would allow an analysis team to directly access owner databases and make specific data calls on demand, removing the need for owners to relay and manage every single data request.

5.2 Data Management

“Data management” refers to the ability for DOE, NREL, and PNNL to store confidential data internally and to manage access to that data across a diverse set of stakeholders. There are two competing needs to meet this requirement: (1) day-to-day storage requiring frequent updates and (2) long-term storage of large, sensitive data sets. Specific data management lessons from Phase 1 include:

- The current system for collecting and sharing data needs to be streamlined going forward, with a reassessment of how to best handle the (1) day-to-day storage and management and (2) long-term storage. User experience and speed are critical to (1), as well enable project staff to have more direct control over how the solution is managed. Metadata tracking and the visibility of data sets are critical to serving the need of (2). Security is a deeply important concern across both types of data management.
- It is challenging to develop and deploy data management systems in parallel to executing a data-intensive project such as WP3. On several occasions during Phase 1, there were delays in making data accessible to project stakeholders because of unforeseen challenges, such as deploying new data management features or through system bugs that needed to be addressed before data could be shared. These types of unforeseen delays are common in software development efforts, but they are compounded when deployed at the same time as a new and untested business process.

5.3 Energy Yield Assessment Results Collection

After each consultant conducted an EYA for a project, they submitted the results to NREL through a custom-built data-upload portal attached to the DAP. This was ultimately successful, and the method proved effective and secure. At the same time, there were several lessons learned about process and technology improvements that can improve EYA collection in the future:

- Consultants have unique in-house processes that do not map directly to each other, let alone to any single collection system that NREL can design. As a result, we can never fully eliminate user error or ambiguities in the data. We can and should still strive to reduce error via improved validation, by better and more documentation, and by adding flexibility to how consultants interact with the EYA collection application.
- The types and granularity of data requests from consultants should be optimized and agreed upon before a new project phase. A balance must be achieved between the amount of data requested from consultants and the burden of actually entering that data into the collection application.
- Updates to user management, data storage, user interface design, and the process for resubmitting results will be essential for a streamlined Phase 2 of the WP3 benchmark.
- Turbine-level results, met campaign data, and granular statistical information are more time consuming for consultants to port to the EYA collection application, and as a result, some finer resolution data might not be practical to collect. For example, initially NREL asked consultants to describe things such as loss distribution shapes and detailed measurement calculations, but these proved impractical and were eventually removed from the data ask.

- Submission timelines and deadlines should align with commercial timelines of consulting firms. The industry has a clear “busy time” during the last quarter of the calendar year, during which it is unlikely they will be able to make many EYA submissions.
- Going forward, NREL needs to improve the way time-varying losses are collected from consultants (e.g., blade degradation assumptions, refinishing schedules, short-term and long-term degradation). These data were proved critical to the benchmark but were not recognized as such until well into the project. Time-varying losses are needed to correct long-term energy production estimates to the operational assessment period of record.

5.4 Operational Data and Assessment

NREL collected operational data for each project in Phase 1, including availability data, turbine SCADA data, metadata, and related information. There were several data challenges, process challenges, and analysis challenges, each of which were overcome throughout Phase 1. That said, many lessons were learned for future operational assessment efforts:

- The OpenOA project proved to be an important open-source analysis tool that helps support the transparency and accuracy of the benchmark findings. It is also increasingly being used by other researchers and industry stakeholders for their own operational analysis needs. Further development of OpenOA will be critical for future AEP benchmark projects as well as any other DOE projects involving operational data from power plants.
- The operational data model specified in IEC 61400-25 proved cumbersome to implement with ambiguous naming conventions. An industry-wide effort to build a reference implementation based on this data model, such as the ENTR Alliance (<https://www.entralliance.com/>), or OSDU (<https://osduforum.org>) which could be used across different software packages, would be valuable.
- Performance upgrades (e.g., vortex generators or control system upgrades) are not always listed in preconstruction data packages, which means they might not be obvious when making operational data requests. Because these upgrades impact plant performance, not including them in EYAs or in the operational assessment can result in discrepancies in results.
- Better coordination with owners on the operational period of record, as well as on project-specific details such as the construction of neighboring wind plants, will lead to more complete preconstruction data packages being sent to consultants. This will avoid having to provide changes to EYA guidance mid-project, which can be time-consuming for consultants and NREL analysts.

5.5 Metrics, Analysis, and Results

Once NREL collected EYA results from the consultants and operational data from owners, the benchmark team began conducting comparative analysis to develop the benchmark and to develop a set of Phase 1 results. In doing so, the NREL team learned how to best approach this analysis:

- It is critical to be on the same page with all stakeholders about important metrics of success. Each stakeholder group (DOE, NREL, owners, consultants, turbine OEMs) might have a unique set of priorities that will result in different desired types of analysis.

Project results should reflect priorities of the stakeholders and should consider the overall goals. This will help to appropriately tailor the message and define the best way to convey the value of the project.

- Developing early results and socializing them with all stakeholders is critical to ensuring clarity and value of results. NREL was able to iterate quickly based on feedback from the stakeholder group, ultimately resulting in a higher level of confidence in the presented results.
- Clarity is key, and effective communication with industry collaborators ensures continuing success of this effort. Open analysis and unambiguous comparative metrics can alleviate this.

6 Conclusions and Next Steps

Across multiple years, the WP3 benchmark initiative successfully conducted an HVS of energy estimation performance (Lunacek et al. 2018), completed a pilot project benchmarking consultant EYAs against operational data for a single operating wind plant, and with this report completed Phase 1 of the WP3 benchmark. To support these efforts, NREL developed and tested several internal capabilities to manage and share data, collect EYA results, and conduct an operational assessment using the OpenOA platform. To date, the WP3 benchmark analysis has shown that prediction bias and energy production uncertainty have been significant challenges for the industry and that the underlying causes of both bias and uncertainty remain largely unknown. Prediction bias has improved over time; however, detailed analysis of pilot data and Phase 1 results indicate a lack of agreement on underlying methods. Specifically, we see broad disagreement in the quantification of wind resource levels, losses, and uncertainties. Additionally, there is broad disagreement about turbine-level production when given identical source information for the predictions.

The original purpose and goals of WP3—particularly improving the accuracy and confidence of preconstruction EYAs—remain intact and highly relevant given the support we have received from industry to continue our efforts and in light of the increasing market pressures for low-cost, high-performing assets. The WP3 benchmark represents a near-term pathway to improve the financial performance of wind power plants, resulting in a reduced cost of energy for wind.

In parallel to the completion of the Phase 1 benchmark, NREL is performing a more detailed analysis looking at sources of P50 energy production and uncertainty prediction error. These results will be published in additional peer-reviewed journal articles.

References

- Bosilovich, M. G., R. Lucchesi, and M. Suarez, 2016: *MERRA-2: File Specification*. GMAO Office Note No. 9 (Version 1.1). 73 pp. http://gmao.gsfc.nasa.gov/pubs/office_notes (Accessed August 1, 2017).
- Cole, W., N. Gates, T. Mai, D. Greer, and P. Das, 2019: *2019 Standard Scenarios Report: A U.S. Electricity Sector Outlook*. 69 pp.
- Craig, A., M. Optis, M. J. Fields, and P. Moriarty, 2018: Uncertainty quantification in the analyses of operational wind power plant performance. *J. Phys. Conf. Ser.*, **1037**, 052021, <https://doi.org/10.1088/1742-6596/1037/5/052021>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Fields, M. J., and R. Sherwin, 2018: *IEC 61400-15 Working Group Update 2*. 17 pp.
- Filippelli, M., B. Sherwin, and J. Fields, 2018: IEC 61400-15 Working Group Update. *AWEA Wind Resource and Project Energy Assessment Workshop 2018*, Austin, TX, AWEA.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Huber, P. J., and E. M. Ronchetti, 2009: *Robust Statistics*. 2nd ed. Wiley,.
- Lee, J. C. Y., and M. Jason Fields, 2021: An Overview of Wind Energy Production Prediction Bias, Losses, and Uncertainties. *Wind Energy Sci.*, **6**, 311–365, <https://doi.org/10.5194/wes-6-311-2021>, 2021.
- Lee, J. C.-Y., M. J. Fields, J. K. Lundquist, and M. Lunacek, 2018a: Determining variabilities of non-Gaussian wind-speed distributions using different metrics and timescales. *J. Phys. Conf. Ser.*, **1037**, 072038, <https://doi.org/10.1088/1742-6596/1037/7/072038>.
- , ———, and J. K. Lundquist, 2018b: Assessing variability of wind speed: comparison and validation of 27 methodologies. *Wind Energy Sci.*, **3**, 845–868, <https://doi.org/10.5194/wes-3-845-2018>.
- , and Coauthors, 2020: The Power Curve Working Group’s assessment of wind turbine power performance prediction methods. *Wind Energy Sci.*, **5**, 199–223, <https://doi.org/10.5194/wes-5-199-2020>.
- Lunacek, M., M. J. Fields, A. Craig, J. C. Y. Lee, J. Meissner, C. Philips, S. Sheng, and R. King, 2018: Understanding Biases in Pre-Construction Estimates. *J. Phys. Conf. Ser.*, **1037**, 062009, <https://doi.org/10.1088/1742-6596/1037/6/062009>.

Montgomery, D. C., and G. C. Runger, 2014: *Applied statistics and probability for engineers*. 6th ed. Wiley, 811 pp.

Perr-Sauer, J., Mike Optis, Jason M. Fields, Nicola Bodini, Joseph C.Y. Lee, Austin Todd, Eric Simley et al. 2021. “OpenOA: An Open-Source Codebase For Operational Analysis of Wind Farms.” *Journal of Open Source Software*, 6(58), 2171. <https://doi.org/10.21105/joss.02171>.

Seabold, S., and J. Perktold, 2010: Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference* <http://statsmodels.sourceforge.net/> (Accessed January 11, 2021).

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*. Academic Press, 676 pp.

Wiser, R., Mark Bolinger, Ben Hoen, Dev Millstein, Joe Rand, Galen Barbose, Naim Darghouth et al. 2020. *Wind Energy Technology Data Update: 2020 Edition*. https://emp.lbl.gov/sites/default/files/2020_wind_energy_technology_data_update.pdf.