WE-Validate: An Open-Source Framework For Wind Power Validation

Malcolm Moncheur de Rieudotte, Allison M. Campbell, Larry K. Berg, Ye Liu,
Nader Samaan, Lindsay Sheridan, Heng Wang
Pacific Northwest National Laboratory, USA
Email: {malcolm.moncheurderieudotte, allison.m.campbell, larry.berg, ye.liu,
nader.samaan, lindsay.sheridan, heng.wang}@pnnl.gov

***Abstract*—Grid operators rely on historical weather time series at existing and planned wind power plants to make informed decisions when planning for a future power grid with a very high penetration of renewable power. While synthetic wind power time series have been developed using historical weather models, their validation with actual power production data remains complex because of variations in modeling practices and methodologies. This paper introduces the WE-Validate framework, originally designed for wind speed validation and now enhanced for wind power validation with a graphical user interface to support users with minimal programming experience. Validation of the wind power with WE-Validate is based on robust metrics: the root mean squared error (RMSE), centered RMSE, average bias, average percent bias, mean absolute error, mean absolute percent error, cross correlation, and calculations of the ramp magnitude, rate, and duration. This paper showcases WE-Validate with validation of the synthetically derived power for a wind plant in Washington state for July to December 2018. The synthetic power was validated by comparing two datasets with observations. This comparison shows that both series are strongly correlated with the observed data across monthly aggregations, despite a persistent negative bias. The metrics within WE-Validate facilitate immediate insight into the utility of the comparison datasets, enabling analysis across multiple dimensions. This user-friendly, open-source tool can be extended beyond wind power, making it a valuable resource for system planners and operators in different domains.**

I. Introduction

Grid operators require a long baseline of power time series for wind to make informed decisions about future infrastructure investments. To fill this need, synthetic wind power time series have been developed by researchers and industry; however, validation efforts can be complex and are often inconsistently applied. Variations in numerical modeling practices and methodologies drive the need for comprehensive, transparent, and user-friendly frameworks to validate synthetic wind power series with observational data.

Consistent validation processes help to build trust in synthetic power time series, providing critical data for project planning, grid integration, and overall energy system stability. By confirming the performance of synthetic models, validation helps investors, operators, and policymakers make informed decisions, reducing financial risks and optimizing energy production.

This paper leverages an existing framework called WE-Validate, which was originally developed to validate the wind speeds stored in the `netCDF` format [1]. WE-Validate has been expanded to allow for wind power validation with `CSV` files and to provide a graphical user interface (GUI) to facilitate consistent validation for users with minimal programming experience. The new capabilities of WE-Validate are demonstrated through the validation of actual and synthetic hourly wind power production datasets for October 2018 at a wind power plant in Washington state. Through this validation, we demonstrate WE-Validate's ability to quickly and easily produce visuals of time series, histograms, and correlations and tabulate seven built-in metrics—the root mean squared error (RMSE), centered RMSE (CRMSE), average bias, average percent bias, mean absolute error (MAE), mean absolute percentage error (MAPE), and cross correlation—at monthly timescales. WE-Validate improves access to consistent validation by providing users with a common and transparent framework for comparing two or more datasets.

The rest of the paper is organized as follows. Section II describes existing work and the gap that this work fills. Section III describes the WE-Validate GUI and metrics. In Section IV, results are presented for a use case and discussed in Section IV-D. Concluding thoughts are presented in Section V.

II. Background

Existing studies on the validation of synthetic wind power have predominantly focused on reporting metrics for the wind speed from numerical weather prediction models as opposed to the resulting power produced at a wind power plant. The previous version of WE-Validate provided a validation tool tailored to the wind speed and ramping events in the `netCDF` file format [1]. The National Renewable Energy Laboratory's Wind Integration National Dataset (WIND) Toolkit includes open-source meteorology validation code with their public wind meteorology data release [2]. Other efforts to extend the validation of wind meteorology to wind power have focused on presenting the results of bespoke validation frameworks. The authors in [3] presented extensive validation of plant-level power production at wind power plants in the Electric Reliability Council of Texas (ERCOT) electric grid with the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2), ECMWF Reanalysis v5 (ERA5), and High-Resolution Rapid Refresh (HRRR) reanalysis datasets, showing evidence for a strong diurnal sensitivity to the hourly resolution correlation [4], [5]. The authors proposed that the decrease in the hourly correlation during evening hours is primarily due to limitations in the boundary layer condition representations in the reanalysis datasets and suggested that turbine-specific wake effects contribute less to the error.

Besides wind power validation frameworks, the Solar Forecast Arbiter provides a platform for the streamlined validation of solar power forecasts [6]. While this validation platform is tailored to solar power, the developers are exploring the elements needed to support the validation of wind power forecasts. The forecasts and the results from synthetic power production validation have many validation metrics in common; however, the Solar Forecast Arbiter compares against a benchmark forecast, whereas synthetic power production must be compared against actual production. Absent consistent frameworks for computing validation metrics, researchers and industry professionals are left to implement validation analyses independently, often resulting in conflicting metrics for the same source data.

This paper builds upon previous work using open-source, Python-based code called WE-Validate to provide a wind power validation framework that is easily accessible to system planners and operators [1]. The previous version of WE-Validate allows two `netCDF` files containing wind speed data to be compared and applies *default* power curves to estimate power production. This enhanced version allows the comparison of two or more data sources of wind power in the `CSV` format with stock and user-defined metrics. A direct comparison of the wind power time series allows users to develop a sophisticated wind speed to power pipelines with plant-specific power curves. The metrics in WE-Validate are computed transparently and consistently, providing a common framework for comparison.

The extension of file formatting requirements from `netCDF` to include `CSV` means that WE-Validate is no longer limited to wind speed and power validation and can be applied to validate any two time series datasets, regardless of their specific domain. WE-Validate now also includes a GUI to empower users with disparate programming proficiency to easily leverage WE-Validate for their validation needs.

III. Open-Source Validation Framework

This section describes the WE-Validate framework and the enhancements to it. WE-Validate is a universal time series validation tool written in Python and can be found on the WE-Validate GitHub repository maintained by the Department of Energy (DOE) at [7]. It is designed to be user-friendly, open source, and easily accessible to the public. It is modularized and can be extended by users to suit their specific needs. The tool comes with built-in data quality control features to identify and address missing or duplicate data, align multiple time series according to user-defined evaluation period boundaries, and resample high-frequency data to match lower-frequency datasets as per the user's preference, either by averaging or selecting the first instance.

The new GUI allows users to quickly and easily view and download validation results plotted as time series, histograms of interval power, cross-correlation scatterplots, and time series of the ramp rate, magnitude, and duration of ramping events given a specified sensitivity to the period-to-period change. Plots are summarized monthly by default. The WE-Validate GUI supports zooming into specific time intervals to visually compare the baseline and comparison datasets.

In the following subsections, WE-Validate's seven built-in metrics (RMSE, CRMSE, average bias, average percent bias, MAE, MAPE, and cross correlation) are defined, and their uses for validation are described. Moreover, WE-Validate's ability to compare ramp detection events with signal compression is discussed.

A. WE-Validate Graphical User Interface

To make it easier for the user to setup WE-Validate and view the results, a GUI was built in Python with the Dash Plotly platform [8]. There are two parts to the GUI: an expandable configuration (Fig. 1) and the results (Fig. 2). Users can choose to drag and drop an existing YAML configuration file or complete a configuration form by providing the start and end times of the validation period, at least one validation metric (described in Section III-B) to be calculated and plotted from a drop-down menu, an absolute path for the output directory where `CSV` files and plots will be saved, a name for the active WE-Validate run, and the variable name and units for the time series. WE-Validate requires a baseline dataset and at least one comparison dataset. For each dataset, the user needs to provide the following information: the name of the dataset to be displayed in the results, the variable name, the data frequency in minutes, the flag value, and the full path to the dataset selected via the open file icon.

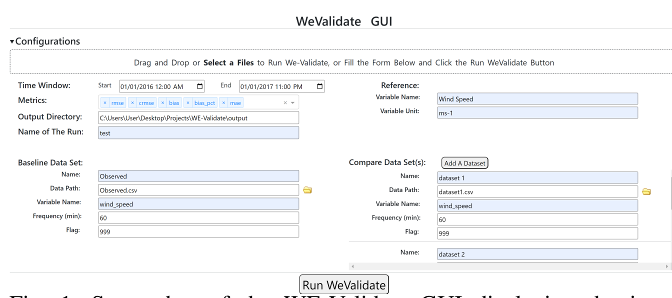

Fig. 1: Screenshot of the WE-Validate GUI displaying the input parameters.

The WE-Validate GUI provides the user with the following graphical outputs: time series (Fig. 2), histograms of the power, a scatterplot of the correlation between time series, and plots that display each user-defined metric on a monthly basis, illustrating the temporal dependence of the relationship between the baseline and the comparison data. WE-Validate also generates tabular data containing the seven user-defined metrics in the form of individual `CSV` files.

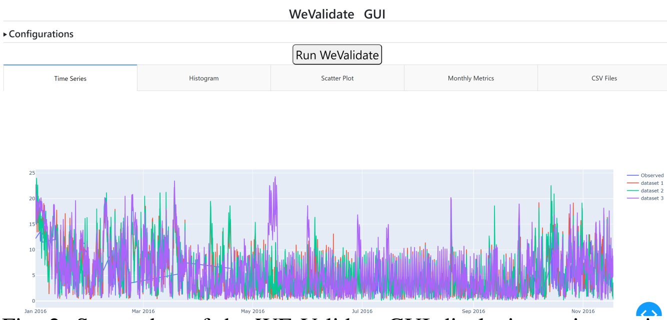

Fig. 2: Screenshot of the WE-Validate GUI displaying a time series output.

B. Validation Metrics

WE-Validate has seven built-in metrics and allows the user to incorporate additional metrics within a simple Python script. The relevant metrics for the validation of the synthetic wind power encompass a range of performance indicators crucial for assessing the accuracy and reliability of wind power estimates. While researchers and industry professionals often employ a single metric to summarize the performance of their synthetic time series, relying solely on one metric can result in the loss of valuable information when reducing multidimensional data. This underscores the importance of using multiple metrics that address different facets of the synthetic wind power time series, depending on the specific evaluation goal, to obtain a comprehensive understanding of the synthetic time series' veracity. Additionally, metrics must be computed consistently to provide a common framework for comparison. To support this, WE-Validate provides consistent calculations for the following metrics, where x is the observed time series, y is the synthetic series, n is a single time interval, and C is the installed capacity of the power plant.

Root Mean Squared Error: The RMSE shows the magnitude of the difference between time series in the units of the source series, such as megawatt hours (MWh) for wind power. This magnitude difference can indicate a systemic bias if the RMSE is large over all time intervals. The RMSE is calculated as

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^{N}(x_n - y_n)^2}{N}}.$$

Centered Root Mean Squared Error: The CRMSE can be used in place of the RMSE where there is suspected bias in the synthetic time series and the user wants to identify the difference in magnitude after correcting for bias. The CRMSE is calculated as

$$\text{CRMSE} = \sqrt{\text{RMSE}^2 - \left(\overline{x} - \overline{y}\right)^2}.$$

Average Bias: The average difference between the synthetic and observed data over a time series of length N conveys whether a model tends to overestimate (positive bias), underestimate (negative bias), or accurately represent (zero bias) the observed data. The average bias is calculated as

$$\text{Average Bias} = \frac{1}{N}\sum_{n=1}^{N} y_n - x_n.$$

Average Percent Bias: The average percent bias indicates the relative bias across the time series and is useful to highlight periods when the synthetic data is significantly larger or smaller than the reference series. The average percent bias is calculated as

$$\text{Average Percent Bias} = \frac{1}{N}\sum_{n=1}^{N}\frac{100 * \left(y_n - x_n\right)}{C}.$$

Mean Absolute Error: The MAE is a measure of the difference in magnitude similar to the RMSE; however, it is less sensitive to outliers, as the absolute value of the errors are calculated. It is most useful with datasets that have significant noise. The units of the error are those of the source dataset. The MAE is calculated as

$$\text{MAE} = \frac{1}{N}\sum_{n=1}^{N}|y_n - x_n|.$$

Mean Absolute Percentage Error (MAPE): The MAPE provides a measure of the relative accuracy of the comparison dataset. As the units are scaled to those of the source dataset, the interpretation of this error metric highlights significant deviations. The MAPE should not be used with time series that have values close to zero, as interval normalization to this value will inflate the error measure. The MAPE is expressed as

$$\text{MAPE} = \frac{1}{N}\sum_{n=1}^{N}\frac{100 * |y_n - x_n|}{C}.$$

Cross Correlation: The cross correlation (also called the Pearson correlation coefficient, denoted by r) is a measure of the relation between two time series. The two series are deemed co-related if, when one is regressed against the other, the slope of the regression is close to one. The cross correlation can indicate whether two series follow a similar trend; however, a high correlation can exist with series that suffer from a large bias. The cross correlation is calculated as

$$r = \frac{\sum_{n=1}^{N}(x_n - \bar{x})(y_n - \bar{y})}{\sqrt{\sum_{n=1}^{N}(x_n - \bar{x})^2 \sum_{n=1}^{N}(y_n - \bar{y})^2}}.$$

The metrics available to the user can be applied to the power time series or to the comparison of the ramps detected in the series. Ramp detection and the application of these metrics to the power time series are discussed in the following section.

IV. Use Case

This section outlines the results provided to the user from WE-Validate for a use case of a wind power plant in Washington state. The validation results are shown for one month to illustrate the default characteristics of WE-Validate. Results are provided in a tabular form for the metrics described in Section III-B on a monthly basis. The plots shown in Section IV-B are produced with monthly binning by default.

A. Data

Validation of the wind power plant data in Washington State is shown here with six months of data in 2018. The example wind power plant is in the Avista balancing authority, has 58 turbines, and has a total rated capacity of 104 MW [9]. The actual hourly production at the wind plant is reported by the Hourly Electric Grid Monitor (EIA-930) provided by the U.S. Energy Information Administration (EIA) [10]. This database contains hourly generation data for each balancing authority in the U.S. starting in July 2018. Between July 2018 and December 2018, the example wind plant was the only one actively producing power in its balancing authority; the hourly generation during this time period is representative of the plant-level production. The synthetic datasets come from the National Wind Power Database (NWPDB) fed with meteorology data from HRRR (NWPDB HRRR) and PLUSWIND modeled with HRRR (PLUSWIND HRRR). The NWPDB is a national database of hourly plant-level power production at 1,390 plants created from the HRRR meteorology in 2018 and the exact power plant configurations specified in the EIA-860 database and the U.S. Wind Turbine Database [11], [9], [12]. PLUSWIND is a national plant-level database of wind power at 1,175 plants using HRRR, ERA5, and MERRA2 meteorology for the years 2018–2021, modeled with the turbine power curves from the U.S. Wind Turbine Database [13]. The NWPDB and PLUSWIND datasets can be found on the DOE's Wind Data Hub [14].

B. Wind Power Validation Plots

WE-Validate provides monthly time series plots of baseline observed time series and all comparison time series. This is shown in Fig. 3 for the example wind plant for October 2018. This figure highlights the variability seen in wind power generation with variations between the synthetic and actual data.

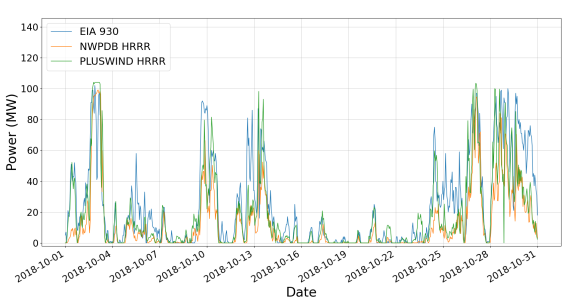

Fig. 3: Hourly power generated by the example wind plant in October 2018.

Although WE-Validate produces monthly plots by default, the user can zoom to a specific interval such as the first week in October 2018, as shown in Fig. 4. With this zoomed-in perspective, the user can immediately identify the differences between synthetic wind power series. For example, neither of the candidate synthetic series produced from the HRRR meteorological datasets fully reproduce the observed decrease in power on October 3 or the spike in power on October 6.

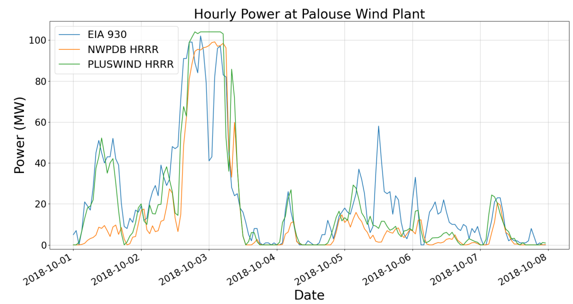

Fig. 4: Hourly power generated by the example wind plant for the 1st week of October 2018.

WE-Validate provides histograms of the base and comparison time series binned monthly. This is shown in Fig. 5 for the example wind plant. The histogram outputs allow users to understand the shape and spread of the data and to identify outliers. A visual comparison of these series shows that both NWPDB HRRR and PLUSWIND HRRR underestimate the frequency of large power outputs in October 2018.

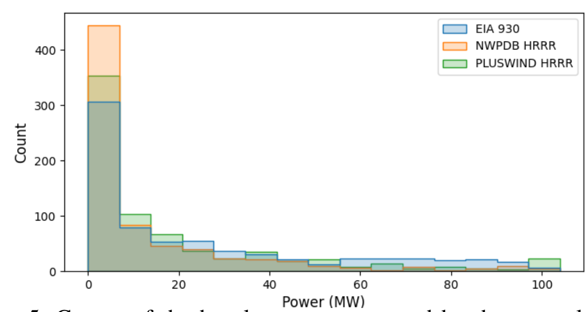

Fig. 5: Counts of the hourly power generated by the example wind plant in October 2018.

Cross-correlation scatterplots for the example wind plant are shown in Fig. 6. The baseline EIA-930 data are regressed against PLUSWIND HRRR (left) and NWPDB HRRR (right) for October 2018. These cross-correlation plots show the systematic underestimation of actual power production by both synthetic power time series. In October 2018, the correlation coefficients between the observed data in EIA-930 and the two comparison sets (PLUSWIND, NWPDB) are $r = 0.83, 0.82$, respectively. PLUSWIND HRRR and NWPDB HRRR are based on the same meteorological dataset; the similarity in the cross correlation is expected.

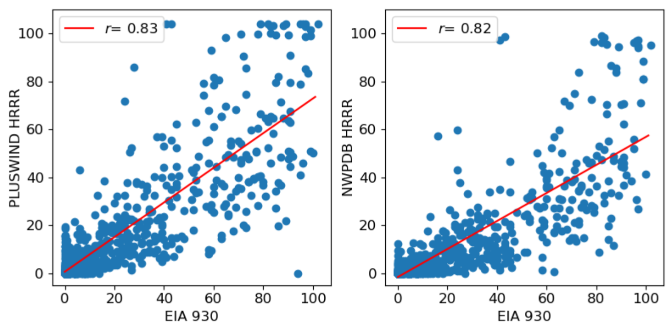

Fig. 6: Cross-correlation scatterplots of the hourly power production between EIA-930 and PLUSWIND HRRR (left) and EIA-930 and NWPDB HRRR (right) for the example wind plant in October 2018.

WE-Validate provides users with the ability to compare ramp detection using a signal compression algorithm called the swinging door algorithm. This algorithm compresses time series into a magnitude, rate, and duration depending on a sensitivity threshold. A threshold of zero means that the value in each consecutive interval is interpreted as a ramp; increasing the threshold increases the tolerance for small interval-to-interval variance. The swinging door algorithm has been widely applied in power systems to convert automatic generation control signals into reserve requirements [15] and in the detection of wind power ramps [16]. In the latter, the sensitivity threshold has been updated to a dynamic threshold. In WE-Validate, the threshold is static and set by the user. The three series produced by the swinging door algorithm can be further compressed into user-selected metrics such as the cross correlation between ramp magnitudes.

Comparisons of ramp detection with the observed EIA-930 hourly power for the example plant are shown in Fig. 7 for NWPDB HRRR and in Fig. 8 for PLUSWIND HRRR. These plots show the magnitude of the ramp (top), the ramp rate (middle), and the duration of the ramp (bottom). Most of the ramp events are captured fairly well by both NWPDB and PLUSWIND; however, neither captures the ramp of roughly 60MW in the first half of the month (blue line, top vignette).

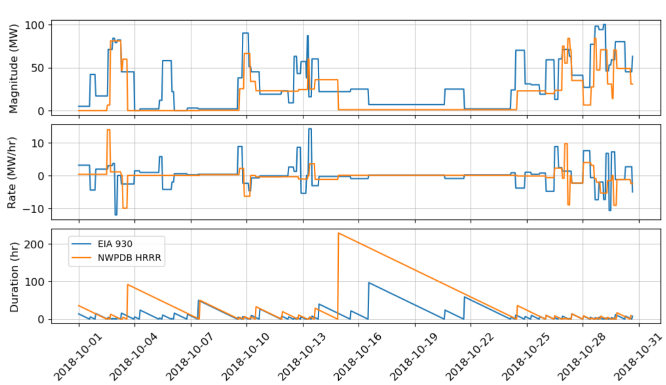

Fig. 7: Ramp magnitude (top), rate (middle), and duration (bottom) for the EIA-930 (blue) and NWPDB HRRR (orange) datasets for the example wind plant in October 2018.

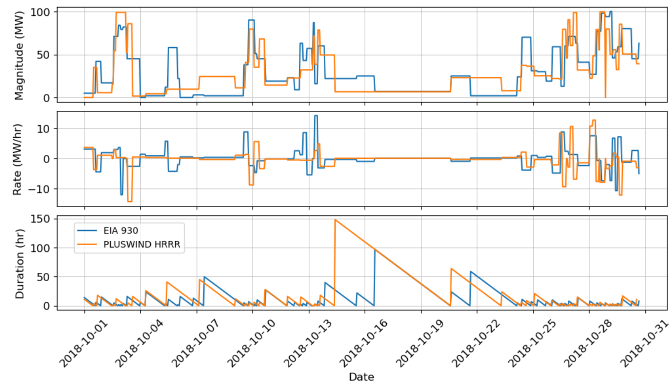

Fig. 8: Ramp magnitude (top), rate (middle), and duration (bottom) for the EIA-930 (blue) and PLUSWIND HRRR (orange) datasets for the example wind plant in October 2018.

C. *Wind Power Validation Metrics*

WE-Validate generates tabular data containing the seven user-defined metrics in the form of individual `CSV` files, calculated on a daily, weekly, monthly, and annual basis. These files are saved in a user-specified output folder.

Table I lists the seven default WE-Validate metrics for EIA-930 and NWPDB HRRR for the months of July through December 2018. The same metrics for the validation of PLUSWIND HRRR with respect to EIA-930 are listed in Table II.

Table III summarizes the monthly average metrics for the swinging door ramp magnitude for July–December 2018.

D. *Use Case Discussion*

The results of the validation with WE-Validate for the use case of a wind plant in Washington state highlight the importance of consistent and transparent frameworks to evaluate wind power time series, especially at the plant level. Although PLUSWIND HRRR and NWPDB HRRR were built from the same HRRR meteorology and EIA-860 plant configuration, the process of producing synthetic wind power time series is fraught with inconsistencies such as the application of weather to power conversion, making it difficult to produce plant-level wind power data.

For example, during the first week of October 2018, neither NWPDB HRRR nor PLUSWIND HRRR captures the decrease in the power production on October 3rd or the spike in power production on October 5th, as shown in Fig. 4. This is reflected in the metrics as a consistent low bias for both synthetic power series. Fig. 6 shows that while the values of r for both datasets are similar, the spread of the data is different. With regards to ramping events, Fig. 7 shows that NWPDB HRRR fails to capture a ramp between October 19th and 22nd, whereas Fig. 8 shows that PLUSWIND HRRR captures the magnitude but not the duration of that same ramp. These results emphasize the importance of using multiple validation metrics to provide a comprehensive understanding of the variability of wind power production and emphasize the significance of validation frameworks that ensure consistent calculations of these metrics.

TABLE I: Monthly Average Metrics for EIA-930 Compared With NWPDB HRRR (July–December 2018)

Month	RMSE	CRMSE	Average Bias	Average Percent Bias	MAE	MAPE	Cross Correlation
July	20.8	18.6	−9.3	−8.9%	12.1	11.7%	0.77
August	23.5	19.5	−13.1	−12.6%	15.2	14.6%	0.78
September	26.0	23.0	−12.2	−11.7%	17.2	16.5%	0.68
October	20.7	16.9	−12.0	−11.5%	13.2	12.7%	0.82
November	29.1	23.3	−17.4	−16.7%	20.2	19.4%	0.78
December	21.6	19.2	−9.8	−9.4%	14.0	13.5%	0.89

TABLE II: Monthly Average Metrics for EIA-930 Compared With PLUSWIND HRRR (July–December 2018)

Month	RMSE	CRMSE	Average Bias	Average Percent Bias	MAE	MAPE	Cross Correlation
July	18.6	18.0	−4.6	−4.4%	10.7	10.3%	0.81
August	21.2	20.2	−6.5	−6.2%	13.6	13.0%	0.77
September	22.2	21.6	−4.8	−4.6%	14.4	13.9%	0.74
October	17.3	16.1	−6.5	−6.2%	10.8	10.3%	0.84
November	23.1	21.5	−8.4	−8.1%	15.3	14.7%	0.83
December	19.8	19.4	−3.9	−3.8%	12.2	11.8%	0.89

TABLE III: Monthly Average Swinging Door Ramp Magnitude Metrics Compared With EIA-930 (July–December 2018)

Month	Synthetic Dataset	RMSE	CRMSE	Average Bias	Average Percent Bias	MAE	MAPE	Cross Correlation
July	NWPDB HRRR	20.2	17.4	−10.3	−9.75%	15.2	14.5%	0.73
July	PLUSWIND HRRR	22.1	21.5	−5.1	−4.73%	17.2	16.5%	0.71
August	NWPDB HRRR	24.0	21.6	−10.5	−10.1%	18.5	17.8%	0.61
August	PLUSWIND HRRR	23.5	23.1	−4.2	−3.98%	17.1	16.0%	0.64
September	NWPDB HRRR	27.8	26.2	−9.3	−9.21%	21.4	20.4%	0.45
September	PLUSWIND HRRR	25.5	25.2	−4.2	−4.0%	18.1	17.5%	0.59
October	NWPDB HRRR	21.2	17.7	−11.8	−11.3%	14.7	14.1%	0.76
October	PLUSWIND HRRR	19.7	19.6	−2.0	−1.98%	14.5	14.1%	0.72
November	NWPDB HRRR	27.4	25.0	−11.2	−10.7%	22.0	21.2%	0.63
November	PLUSWIND HRRR	26.2	25.6	−5.3	−5.29%	19.3	18.9%	0.66
December	NWPDB HRRR	22.6	20.9	−8.5	−8.16%	17.7	17.0%	0.80
December	PLUSWIND HRRR	24.0	23.4	−5.0	−4.8%	18.7	18.1%	0.78

V. Conclusion

This paper demonstrates the new capabilities of WE-Validate by presenting results for the validation of synthetic hourly wind power production datasets for July–December 2018 for a wind power plant in Washington state. The results from WE-Validate show that both datasets have strong correlations with the observed dataset across weekly and monthly aggregations, albeit with a persistent negative bias.

Given the need for consistent frameworks that allow cross-study comparisons, WE-Validate contains seven built-in metrics with the ability to edit existing metrics or include new ones. WE-Validate outputs these metrics in a `CSV` tabular format for daily, weekly, monthly, and annual aggregations. These results allow users to quickly and easily compare metrics of interest over various time scales. WE-Validate generates plots of time series, histograms, and scatterplots and provides ramp detection to facilitate data visualization on the fly.

WE-Validate is a user-friendly, open-source, and easily accessible Python-based code base that can be used to validate two or more time series of time-stamped `CSV` data. Originally designed for wind speed validation, the tool has been extended to include wind power validation and was improved with the addition of a GUI to enhance accessibility, making it usable even for individuals with minimal coding experience. As the tool only requires time-stamped data to compute validation metrics, it can be applied to validate any time series datasets, regardless of their specific domain.

To facilitate easy access alongside existing synthetic wind power datasets on the DOE's Wind Data Hub, a link to the WE-Validate GitHub repository and user manual will be added to the Wind Data Hub.

Future work should address seasonal and diurnal visualizations within the GUI, moving beyond the default monthly visualization. Users are encouraged to incorporate additional metrics.

VI. Acknowledgment

This study is funded by the U.S. Department of Energy (DOE) Wind Energy Technologies Office (WETO). The project team wants to especially thank Dr. Jian Fu and Dr. Cynthia Bothwell from the Grid Integration Subprogram at DOE's WETO for their continuing support, help, and guidance.

References

[1] J. C. Y. Lee, C. Draxl, and L. K. Berg, "Evaluating wind speed and power forecasts for wind energy applications using an open-source and systematic validation framework," *Renewable Energy*, vol. 200, pp. 457–475, Nov. 2022. [Online]. Available: https://doi.org/10.1016/j.renene.2022.09.111

[2] W. Lieberman-Cribbin, C. Draxl, and A. Clifton, "A guide to using the WIND Toolkit validation code," National Renewable Energy Laboratory, Tech. Rep., 2014.

[3] M. R. Davidson and D. Millstein, "Limitations of reanalysis data for wind power applications," *Wind Energy*, vol. 25, no. 9, p. 1646–1653, Jun. 2022. [Online]. Available: http://dx.doi.org/10.1002/we.2759

[4] R. Gelaro, W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A. M. da Silva, W. Gu, G.-K. Kim, R. Koster, R. Lucchesi, D. Merkova, J. E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S. D. Schubert, M. Sienkiewicz, and B. Zhao, "The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2)," *Journal of Climate*, vol. 30, no. 14, p. 5419–5454, Jul. 2017. [Online]. Available: http://dx.doi.org/10.1175/JCLI-D-16-0758.1

[5] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J. Thépaut, "The ERA5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, p. 1999–2049, Jun. 2020. [Online]. Available: http://dx.doi.org/10.1002/qj.3803

[6] C. W. Hansen, W. F. Holmgren, A. Tuohy, J. Sharp, A. T. Lorenzo, L. J. Boeman, and A. Golnas, "The solar forecast arbiter: An open source evaluation framework for solar forecasting," in *2019 IEEE 46th Photovoltaic Specialists Conference (PVSC)*, 2019, pp. 2452–2457.

[7] "DOE Wind Data Hub GitHub repository." [Online]. Available: https://github.com/a2edap/WE-Validate/tree/main

[8] "Dash Plotly." [Online]. Available: https://dash.plotly.com/

[9] B. Hoen, J. E. Diffendorfer, J. Rand, L. A. Kramer, C. P. Garrity, A. D. Roper, and H. Hunt, "United States wind turbine database," 2018. [Online]. Available: https://www.sciencebase.gov/catalog/item/57bdfd8fe4b03fd6b7df5ff9

[10] E. I. A. U.S. Department of Energy, "Hourly electric grid monitor, EIA-930." [Online]. Available: https://www.eia.gov/electricity/gridmonitor/about

[11] ——, "Form EIA-860 detailed data with previous form data (EIA-860A/860B)." [Online]. Available: https://www.eia.gov/electricity/data/eia860/

[12] "National Wind Power Database." [Online]. Available: https://a2e.energy.gov/project/nwpdb

[13] "The Plant-Level US multi-model WIND and generation (PLUSWIND) data." [Online]. Available: https://a2e.energy.gov/project/pluswind

[14] C. Sivaraman and M. MacDuff, "The Wind Data Hub, also known as The Data Archive and Portal." 2023. [Online]. Available: https://www.osti.gov/biblio/1910052

[15] Y. V. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational impacts of wind generation on california power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1039–1050, 2009.

[16] Y. Cui, Y. He, X. Xiong, Z. Chen, F. Li, T. Xu, and F. Zhang, "Algorithm for identifying wind power ramp events via novel improved dynamic swinging door," *Renewable Energy*, vol. 171, p. 542–556, Jun. 2021. [Online]. Available: http://dx.doi.org/10.1016/j.renene.2021.02.123