https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

Deep mining of megawatt large wind turbine actual operating data: Exploration of accurate modeling & performance optimization

Weimin Wu[1], Xiongfei Liu[2], Yu Ren[2], Suocheng Zhang[3], Wanjun Yan[1], Wenqiang Du[2]

[1]School of Electronic and Information Engineering, Anshun University, Anshun, 561000, China

[2]Yinchuan University of Science and Technology, Yinchuan, 750001, China

[3]Inner Mongolia Three Gorges Mengneng Energy Co., Ltd., Hohhot, 010090, China

Correspondence to: Xiongfei Liu (xiao_fang_liu@yeah.net)

Abstract: The real-time operation data analysis and condition monitoring of large wind turbines are crucial for ensuring the efficient and safe operation of wind farms. In response to this, this paper proposes a precise prediction model architecture based on the multivariate linear regression algorithm to gain a deeper understanding of the actual operation of large wind turbines. By comparing different prediction variable combinations, we confirmed that the average wind direction and average wind speed play a core role in predicting active power, and found that their combined effect can capture more than 70% of power changes. Furthermore, this paper innovatively introduces Bayesian algorithm for parameter fusion, effectively improving the model's goodness of fit. However, the complexity of the data in actual applications poses a challenge to the effectiveness of the Bayesian fusion algorithm, suggesting that further optimization of the algorithm is needed to cope with the complex and variable real data environment. This study provides scientific evidence for the efficient operation, precise maintenance, and environmentally friendly design of wind turbines, promoting the continuous progress and development of wind power generation technology.

Keywords: Real-time data of wind turbine, multivariate linear regression, Bayesian parameter fusion, R-squared value, optimization of wind farm maintenance.

1. Introduction

Real-time analysis of large wind turbine operation data and close monitoring of its status are crucial for ensuring the efficient and safe operation of a wind farm. These data are like the "vital signs" of the wind farm, reflecting the health status, power generation efficiency, and potential faults of the wind turbine. By monitoring and analyzing in real time, potential problems can be detected and resolved in a timely manner, preventing downtime risks, optimizing maintenance strategies, and improving energy output. In addition, long-term data analysis can help drive continuous improvement and innovation in wind power technology, pushing the wind power industry towards a more intelligent and sustainable direction. The analysis of real-time operational data for large wind turbines is crucial for ensuring their efficient and safe operation. Luan and Moan (2021) highlighted the significance of considering startup and shutdown induced transient load processes on fatigue damage

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

in wind turbine towers. This emphasizes the need to account for dynamic operational conditions in fatigue analysis. Zhao et al. (2019) proposed a holistic monitoring system for wind turbines that provides real-time condition monitoring and data recording for post-event analysis, showcasing the importance of continuous monitoring for identifying potential issues. Uncertainties in key performance indicators (KPIs) for wind turbine operation were discussed by Pfaffel et al. (2020), underscoring the need for accurate data handling to support decision-making in the wind industry. Strbac et al. (2019) introduced a data-mining approach for fault detection in wind turbines using SCADA data analysis, demonstrating the effectiveness of machine learning techniques in maintenance activities. Trujillo-Franco et al. (2021) presented a method for identifying modal parameters of wind turbine blades in real-time, showcasing the potential for online operational assessment. Zhu et al. (2022) proposed a method for real-time operational state prediction of wind turbine gearboxes using deep learning and fuzzy synthesis, highlighting the importance of predictive maintenance to reduce costs and improve reliability. Rabie et al. (2022) discussed turbine curtailment strategies to reduce bat fatalities at wind energy facilities, emphasizing the need for operational adjustments to mitigate environmental impacts. Tsai and Wang (2022) introduced an acoustic-based method for identifying surface damage on wind turbine blades, showcasing the potential of convolutional neural networks in damage detection. Innovative approaches for anomaly detection in wind turbines were presented by Chen et al. (2024), introducing a new method based on norm-linear-ConvNeXt-TCN architecture for detecting abnormal operating conditions. These studies collectively highlight the importance of real-time data analysis and monitoring in ensuring the efficient and safe operation of large wind turbines.

Despite significant progress in the analysis and monitoring of real-time operation data for wind turbines, most of the previous studies focused on specific technical challenges or performance evaluation dimensions, and rarely covered all the critical links and subtle data changes in the complex operation environment of large wind turbines. For example, the real-time monitoring and analysis of the dynamic response of wind turbines under extreme weather conditions, the fluctuations of energy conversion efficiency under different wind speeds and directions, and the potential impact of material aging on structural safety in long-term operation are still insufficient. This study is aware of this gap and is committed to filling this knowledge gap by innovatively constructing a comprehensive data model, focusing on analyzing those often overlooked links and data in actual operation, such as the operating characteristics under non-standard conditions, early warning signals of minor faults, and the long-term correlation between environmental factors and turbine performance. The goal is to have a more comprehensive understanding of the actual operation of large wind turbines, providing scientific basis for the efficient operation, precise maintenance, and environmentally friendly design of wind farms, and promoting the continuous progress and development of wind power generation technology. This paper is organized as follows. Section 2 presents the algorithm theory and pros and cons of the multiple linear regression algorithm compared with other intelligent model algorithms. Section 3 a precise prediction model architecture is proposed based on the algorithm of multiple linear regression for wind turbine actual operation data. Section 4 conducts a precise model analysis on the actual large-scale wind turbine data in operation. The contribution closes with some conclusions and final remarks.

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

2. Multivariate Regression Algorithm and Related Comparisons

Multiple Linear Regression (MLR) is a statistical analysis method used to study the linear relationship between one dependent variable (response variable) and two or more independent variables (explanatory or predictor variables)(Alviso et al., 2020). In a multiple linear regression model, the dependent variable is considered to be a function of the linear combination of the independent variables, and may include a random error term to represent the variability outside the model.

2.1 Theoretical Foundation

Suppose we have a dataset $(x_1^{(i)}, x_2^{(i)}, \ldots, xn^{(i)}, y^{(i)})i = 1^m$, where $x_j^{(i)}$ is the value of the j-th independent variable for the i-th observation, $y^{(i)}$ is the corresponding value of the dependent variable, m is the total number of observations, and n is the number of independent variables (Egbueri and Agbasi, 2022). The multiple linear regression model can be expressed as:

$$y^{(i)} = \beta_0 + \beta_1 x_1^{(i)} + \beta_2 x_2^{(i)} + \cdots + \beta_n x_n^{(i)} + \epsilon^{(i)} \tag{1}$$

Here, β_0 is the intercept term, $\beta_1, \beta_2, \ldots, \beta_n$ are the regression coefficients, and $\epsilon^{(i)}$ is the random error term, typically assumed to be independent and identically distributed with a mean of 0 and some variance.

2.2 Parameter Estimation

The core of multiple linear regression lies in estimating the regression coefficients $\beta_0, \beta_1, \ldots, \beta_n$ (Roy, 2021). This is usually accomplished through the method of least squares, which involves finding the set of regression coefficients that minimizes the sum of the squared residuals (residual sum of squares, RSS).
The RSS is defined as:

$$RSS = \sum_{i=1}^{m}(y^{(i)} - (\beta_0 + \beta_1 x_1^{(i)} + \beta_2 x_2^{(i)} + \cdots + \beta_n x_n^{(i)}))^2 \tag{2}$$

By setting the partial derivatives of the RSS with respect to $\beta_0, \beta_1, \ldots, \beta_n$ equal to zero and solving the resulting equations, we obtain the estimated values of the regression coefficients, denoted as $\widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_n$.

2.3 R-squared Value

The R-squared value is a key metric used to assess the goodness of fit of a multiple linear regression model (Oh et al., 2020). It measures the proportion of the total variation in the dependent variable that is explained by the model's independent variables.
The R-squared value is calculated as:

$$R^2 = 1 - \frac{RSS}{TSS} \tag{3}$$

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

where TSS (Total Sum of Squares) is the total variation in the dependent variable around its mean, given by:

$$\mathrm{TSS} = \sum_{i=1}^{m} (y_i - \bar{y})^2 \tag{4}$$

with $\bar{y}$ being the mean of the observed values of Y.

2.4 Comparison of the Advantages and Disadvantages of Algorithms

In delving into the comparative analysis of Multiple Linear Regression (MLR) and other prominent algorithms like Decision Trees, Neural Networks, and Genetic Algorithms, it becomes evident that each approach boasts distinct strengths tailored to specific use cases. Multiple Linear Regression stands out for its simplicity, interpretability, and elegance within the confines of a linear relationship framework. By modeling the dependency of a continuous response variable on a set of explanatory variables, MLR offers straightforward coefficients that can be interpreted as the direct and quantifiable effect of each predictor. This feature renders MLR invaluable in disciplines where transparency and causality are paramount concerns, enabling decision-makers to grasp the nuances of predicted outcomes.

However, it is important to recognize that the complexity and nuances of real-world data often necessitate algorithms beyond the scope of MLR. Decision Trees (Sun et al., 2024), for instance, excel in handling non-linear relationships and categorical variables, offering a hierarchical, tree-like structure that can be intuitively understood by non-technical users. Their ability to automatically perform feature selection and handle missing data makes them versatile tools for classification and regression tasks. On the other hand, Neural Networks (Lavecchia, 2024), with their layered structure inspired by the human brain, are capable of learning and modeling extremely complex patterns in data, even those that defy mathematical formulations.

This flexibility comes at the cost of interpretability, as Neural Networks operate as "black boxes," but their predictive power can be unparalleled in certain domains. Lastly, Genetic Algorithms (Alhijawi and Awajan, 2024), inspired by natural selection, offer a unique approach to optimization and search problems, evolving solutions over generations of iterations. They are adept at finding solutions in vast, unconstrained search spaces and can be applied to both continuous and discrete problems. By highlighting MLR's strengths while acknowledging the distinct characteristics of Decision Trees, Neural Networks, and Genetic Algorithms, we gain a nuanced understanding of the diverse algorithmic landscape and their respective roles in data analysis and predictive modeling.

3. A Precise Prediction Model Architecture

3.1 Data Preprocessing

Data preprocessing for multiple linear regression also requires a series of steps to ensure the accuracy of the data and the effectiveness of the model. Here are the specific steps taken for data preprocessing in this study. The specific illustration is shown in Figure 1.

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

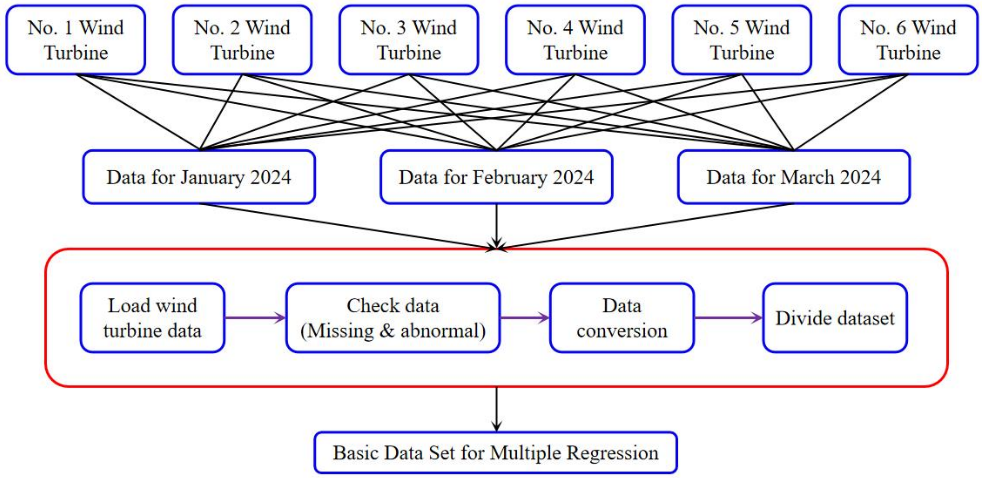


Figure 1. Preprocessing of wind turbine operation monitoring data

The data preprocessing flow for the multivariate linear regression algorithm mainly includes the following steps: ① Load the data and conduct initial exploratory analysis to ensure the completeness and quality of the data and prepare for the subsequent modeling process. Among them, missing value detection and outlier detection are very necessary, which may have an adverse effect on the regression model. According to the nature of the outliers and the characteristics of the data distribution, the values can be deleted or adjusted. Here, the real-time monitoring data from 1 to 6 wind turbines for three months is targeted. ② After completing the data loading and initial exploration, the next stage is the data preparation and transformation stage, whose task is to convert the data into a form suitable for analysis by the multivariate linear regression model. The function value variable is selected as the average active power, and the independent variable is selected as the average blade angle, average wind direction, average blade angle of the 1st blade, and average blade angle of the 2nd blade. Since the average blade angle of the 3rd blade is seriously missing, it is ignored here. ③ The data set is divided according to the 80/20 principle, and it is divided into a training set (80%) and a test set (20%). The training set is used to train the model, and the test set is used to evaluate the model's performance. When dividing the data set, attention should be paid to maintaining the consistency of data distribution between different sets to avoid introducing bias.

3.2 Build and Train the Model

In this section, the basic data of No.1 wind turbine is used to train a multivariate linear regression model and conduct detection. Then, the basic data of No.2 wind turbine is used to train a similar multivariate linear regression model, and the model parameters of No.1 Wind Turbine are used for fusion and correction. Following this logic, the regression models for No.3 wind turbine and No.4 wind turbine are obtained by correcting the models based on their respective basic data. The logical flow framework for the specific correction is shown in Figure 2.

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

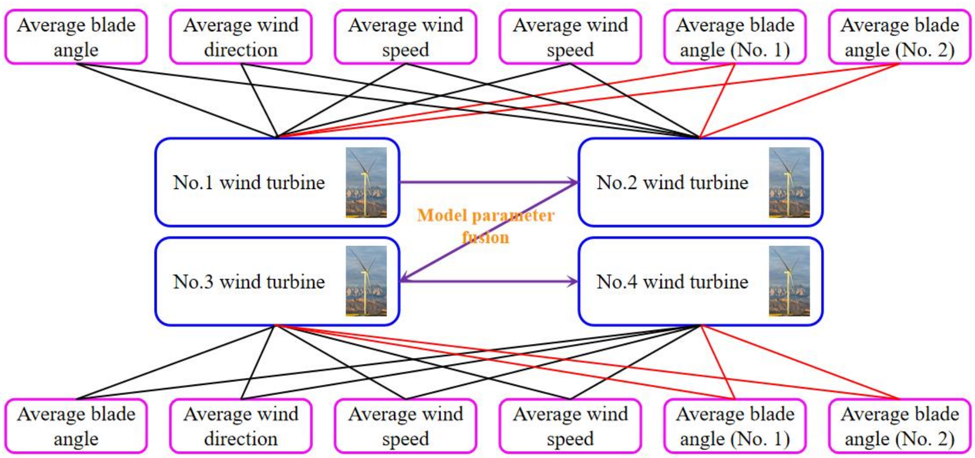


Figure 2. The domain division for integral wind turbine

The fusion of parameters from multiple linear regression models at this juncture involves employing the Bayesian Model Averaging (BMA) framework, specifically leveraging the posterior probabilities derived from the Akaike Information Criterion (AIC) as weights for the averaging process (Berkhout et al., 2024). AIC is a widely used metric in model selection that balances the goodness of fit of a model with its complexity, penalizing models with more parameters. In the context of BMA, AIC can be utilized to inform the relative importance or contribution of each individual model to the overall predictive ensemble.

To expand on this process and provide the corresponding computational formulas, let's outline the steps: Fit Multiple Models: First, multiple linear regression models are fitted to different datasets or subsets of the same dataset, resulting in a collection of models $M_1, M_2, \ldots, M_k$, where k is the total number of models. Calculate AIC for Each Model: For each model M_i, the AIC is computed using the formula:

$$\mathrm{AIC_i} = 2p_i - 2\log(L_i) \tag{5}$$

where p_i is the number of parameters (including the intercept) in model M_i, and $\log(L_i)$ is the log-likelihood of the model given the data. Note that the actual log-likelihood function used depends on the underlying assumptions of the model, but for linear regression with Gaussian errors, it can be approximated based on the residual sum of squares.

Compute Relative Weights: The relative weights w_i for each model M_i are then computed based on the AIC values. A common approach is to use the exponential of the negative half of the difference between each model's AIC and the minimum AIC among all models: where $\Delta\,\mathrm{AIC_i} = \mathrm{AIC_i} - \min(\mathrm{AIC_1}, \mathrm{AIC_2}, \ldots, \mathrm{AIC_k})$. Perform Bayesian Model Averaging: Finally, the predictions or estimates from the individual models are averaged using the computed weights. For a given prediction or estimate $\hat{y}_{i,new}$ from model M_i for a new data point, the BMA prediction is:

$$\hat{y}_{BMA,\,new} = \sum_{i=1}^{k} w_i \hat{y}_{i,new} \tag{6}$$

By following these steps, the parameters from multiple linear regression models are fused in a principled manner, accounting for both the goodness of fit and complexity of each model through the use of AIC-based posterior probabilities (or rather, weights informed by AIC). This approach helps mitigate the risk of overfitting while still leveraging the information contained in each individual model.

3.3 Evaluation and Validation of Fusion Model

After the fusion of the multiple linear regression models through Bayesian Model Averaging (BMA), the combined model will undergo further validation and evaluation using the foundational data from No.5 Wind Turbine and No.6 Wind Turbine. This validation and assessment process is crucial to ensure the robustness and applicability of the fused model to the specific context of these two wind turbines. The specific illustration is shown in Figure 3.

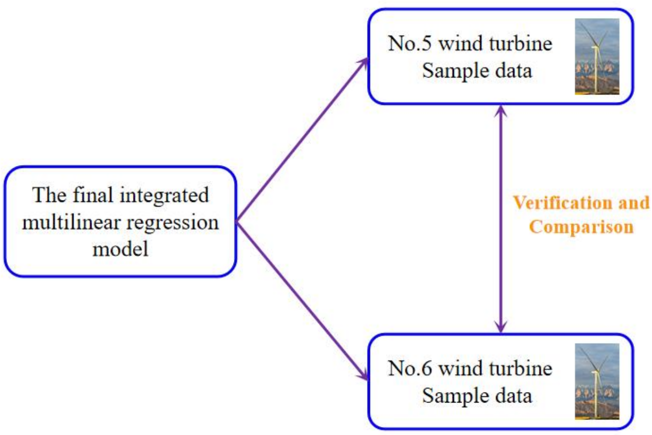


Figure 3. Validation and Evaluation of the Fusion Model

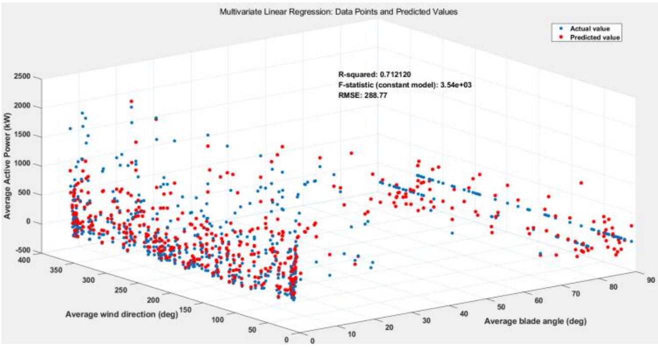
The fused BMA model will be applied to the prepared data from Wind No.5 and No.6 Turbines. This involves using the model's coefficients and intercept, weighted according to the BMA weights, to make predictions for the response variable of interest. The predictions made by the fused BMA model will be compared to the actual observed values for Wind Turbines No.5 and No.6. This comparison can be done both qualitatively, through visual inspection of the predictions versus actuals, and quantitatively, using various statistical metrics.

4. Analysis of Results on Real Operational Data

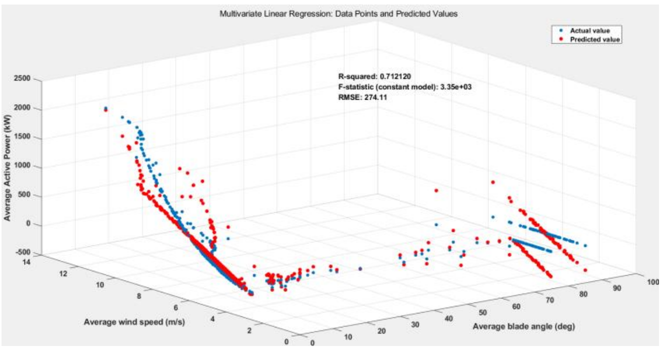
The relevant computational results of the multivariate linear regression framework present a series of detailed data analysis results, including not only the specific numerical values of the regression coefficients such as the coefficient of determination (R^2) and the adjusted coefficient of determination, but also other evaluation information. In order to present these results in a more intuitive way and assist in understanding the model's performance, the evaluation information will be

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

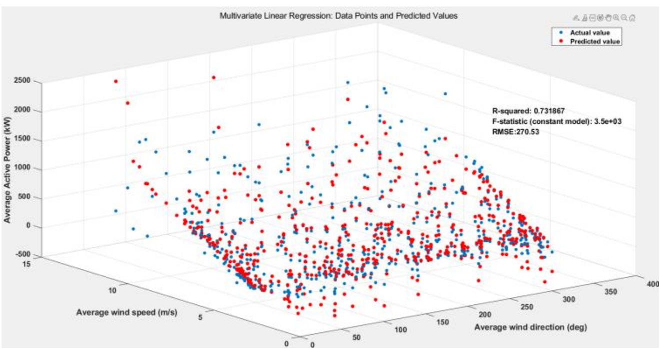
carefully added to the main fitting results displayed in the Figure 4.



(a) Prediction of Data for Wind Turbine No. 1



(b) Prediction of Data for Wind Turbine No. 1



(c) Prediction of Data for Wind Turbine No. 1

Figure 4. Validation and Evaluation of the Fusion Model

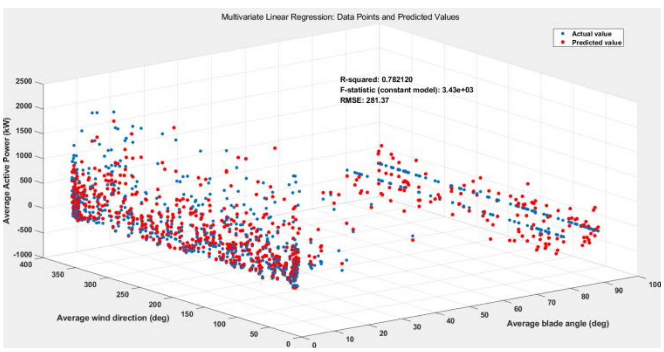
The above predictions and comprehensive validation results not only reveal the great potential of the multivariate linear regression model in explaining the active power output of large wind turbines, but also highlight the outstanding performance of wind direction average and wind speed average as core predictor variables. The high degree of determination coefficient (R-squared), a key indicator for measuring the degree of model fitting, provides a direct visualization of the model's ability to capture more than 70% of the active power variation when these two major factors work together. This

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

achievement is significantly better than that of single variables or other variable combinations, further confirming their core position in wind power generation efficiency prediction. It is worth noting that while the root mean square error (RMSE) and F-statistic value show relative consistency in comparing different predictor variable combinations, a deeper analysis reveals that the average pitch angle is the most sensitive predictor variable, with minor adjustments having a significant impact on output power. This finding not only aligns with the practical experience of frontline technicians, but also provides scientific evidence for optimizing wind turbine operation and maintenance strategies.

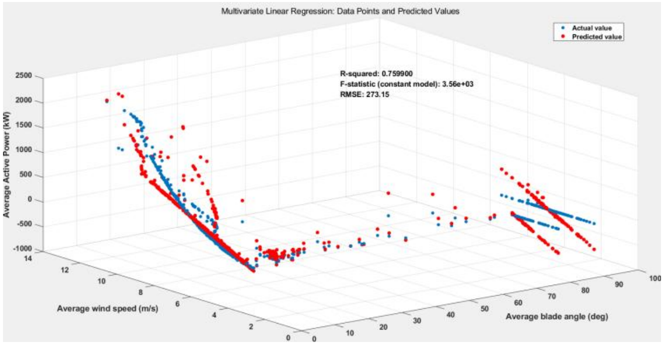
Moreover, all three main predictor variables (including wind direction, wind speed, and pitch angle) in the model have demonstrated excellent performance with a coefficient of determination above 0.71, which is not only a strong proof of the model's effectiveness but also indicates that further refinement of the model and parameter tuning may lead to higher prediction accuracy (Zhang and Wang, 2014). Given that the multivariate regression model has shown good initial fitting effects, introducing the Bayesian Model Averaging (BMA) method will be a forward-looking strategy. BMA method effectively avoids the overfitting or underfitting problems of a single model by considering all possible model combinations and assigning them corresponding weights based on posterior probabilities. This method not only enhances the robustness and accuracy of the model's predictions, but also helps us better understand the relative importance of different predictor variables in the prediction process and their interactions.

After applying the Bayesian Information Criterion (AIC) algorithm for the first time to fuse and optimize the parameters of the multivariate linear regression model, we obtained a series of encouraging prediction results. At the same time, we used AIC, an efficient model selection standard, to automatically balance the complexity of the model and the degree of fitting, thereby achieving more precise and robust predictions. Specifically, as shown in Figure 5.

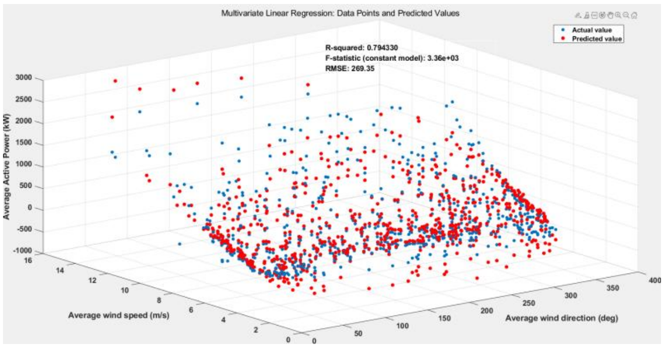


(a) The first Bayesian fusion result

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024



(b) The first Bayesian fusion result



(c) The first Bayesian fusion result

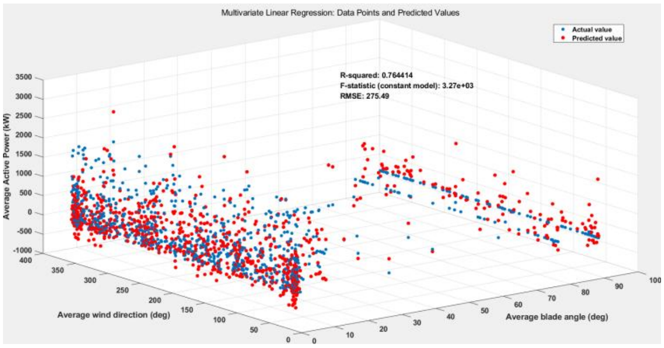
Figure 5. Fusion model integrating data from No.2 wind turbine

Through in-depth exploration and application of Bayesian algorithms for parameter fusion, we observed a significant improvement in the R-squared value, which not only validated the rationality and innovativeness of our research approach at the theoretical level, but also confirmed its effectiveness in improving the accuracy of model predictions at the practical level. The increase in R-squared value directly reflects the improvement in the model's goodness of fit, indicating that the fused parameters are better able to accurately capture the complex relationships between data, thereby reducing the deviation between predicted values and actual values. However, we also recognize that although the Bayesian parameter fusion method brings significant improvements, its effectiveness is constrained by various factors in actual applications. The complexity of data, including non-linear relationships, outliers, missing values, and potential interactions between variables, may pose challenges to the method's performance. These factors limit the optimal prediction accuracy that can be achieved solely by relying on the Bayesian fusion algorithm.

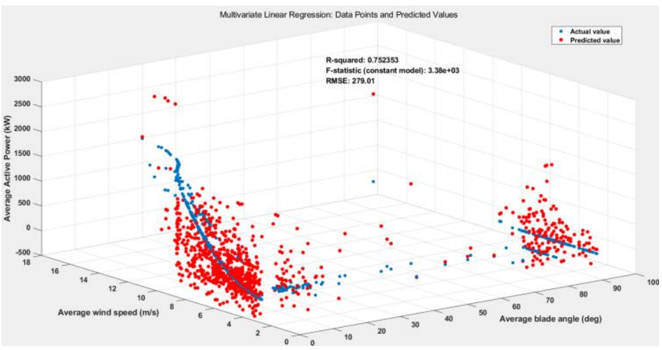
To further tap the potential of the model and achieve more precise predictions, we should consider combining the Bayesian parameter fusion method with other advanced algorithms and technologies to form a comprehensive predictive framework with complementary advantages. For example, we can introduce machine learning's ensemble learning methods, such as random forests and gradient boosting trees, to capture non-linear features in the data; or utilize deep learning technology to learn and express high-order dependencies between data by constructing more complex network structures. In addition,

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

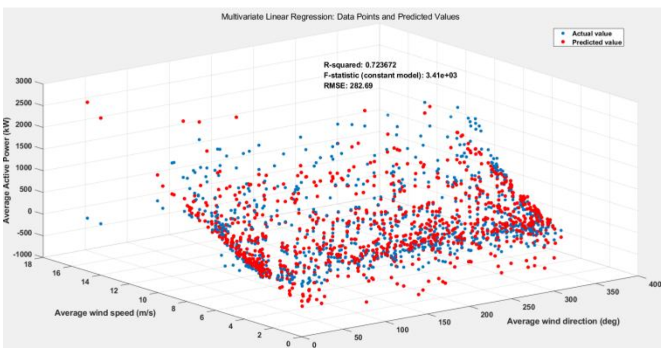
optimizing the data preprocessing stage, such as feature selection, outlier handling, and data normalization, is also a key step in improving the overall performance of the model. For the final verification on the data of No.5 and No.6, the Figure 6 is as follows.



(a) Final Bayesian fusion result



(b) Final Bayesian fusion result



(c) Final Bayesian fusion result

Figure 6. Fusion model integrating data from wind turbines of No.2 to No 4

From the current analysis results, the R-squared index has decreased, and this change is not accidental, but contains profound implications from multiple aspects. First of all, although the wind turbines selected for comparison and optimization have shown a high degree of similarity in their working environment and conditions, which is an ideal basis for model comparison

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

and optimization, the complexity and diversity of the real world often exceed our theoretical assumptions. Especially when considering the various unforeseen factors that may arise in actual operation, these external disturbances are likely to have a significant impact on the basic data of the wind turbines. Specifically in this analysis, there may be higher levels of uncertainty factors hidden in the basic data of wind turbines 5 and 6. These factors may originate from various sources, including but not limited to, performance deviations due to equipment aging, incomplete maintenance records, minor differences in operating environment (such as subtle changes in wind direction and speed), and inconsistencies in human operation, etc. The cumulative effect of these uncertainty factors is reflected in the deviation between the model prediction results and actual observations, ultimately leading to a decrease in R-squared.

To address this challenge and enhance the accuracy and practicality of the model, we plan to adopt a more refined data collection and processing strategy. Specifically, we will further refine the scope of data extraction, paying attention not only to the overall operating data of the wind turbine, but also to the monitoring data of each wind turbine's individual blades. The advantage of doing so is that it can capture the subtle changes of the wind turbine in operation more comprehensively, thus providing more rich and realistic feature information for the training of the model.

5. Conclusion

This study conducts in-depth analysis of real-time operating data for large wind turbines and successfully builds an accurate prediction framework based on a multivariate linear regression model. By comparing the performance of different prediction variable combinations, we confirm the crucial role of wind direction average and wind speed average as core predictor variables, which together capture more than 70% of active power changes. This significantly outperforms single variables or other variable combinations. Furthermore, by introducing Bayesian algorithms for parameter fusion, we further enhance the model's goodness of fit (R-squared value), validating the effectiveness and rationality of this innovative method in improving prediction accuracy. However, the complexity of real-world data (such as nonlinear relationships, outliers, missing values, and potential interactions between variables) limits the optimal prediction accuracy of the Bayesian fusion algorithm. Therefore, future work needs to further optimize the algorithm to cope with the complex and variable real-world data environment, thereby continuously improving the efficiency of wind turbine operation and prediction accuracy.

Acknowledgement

This work was funded by the Anshun University Doctoral Special Fund Project (No.anxybsjj202411), the Ningxia Higher Education & Research Project of China (No. NGY2018-259), and the Ningxia Natural Science Foundation of China (No.2023AAC03389).

Code/Data availability

Not applicable.

Author contribution

Weimin Wu and Xiongfei Liu: Designed and performed the experiments, analyzed the data and prepared the paper.
Yu Ren and Suocheng Zhang: Designed the experiments and revised the manuscript.
Wanjun Yan and Wenqiang Du: Participated to collect the materials related to the experiment.

Competing interests

The authors declare that they have no conflict of interest.

References

Alhijawi, B., & Awajan, A.: Genetic algorithms: Theory, genetic operators, solutions, and applications. Evolutionary Intelligence, 17(3), 1245-1256, doi: 10.1007/s12065-023-00822-6, 2024

Alviso, D., Artana, G., & Duriez, T.: Prediction of biodiesel physico-chemical properties from its fatty acid composition using genetic programming. Fuel, 264, 116844, doi: 10.1016/j.fuel.2019.116844, 2020

Berkhout, S. W., Haaf, J. M., Gronau, Q. F., Heck, D. W., & Wagenmakers, E. J.: A tutorial on Bayesian model-averaged meta-analysis in JASP. Behavior Research Methods, 56(3), 1260-1282, doi: 10.3758/s13428-023-02093-6, 2024

Chen, N., Shao, C., Wang, G., Wang, Q., Zhao, Z., & Liu, X.: Anomaly detection of wind turbine based on norm-linear-ConvNeXt-TCN. Measurement Science and Technology, 35(7), 076107, doi: 10.1088/1361-6501/ad366a, 2024

Egbueri, J. C., & Agbasi, J. C.: Data-driven soft computing modeling of groundwater quality parameters in southeast Nigeria: comparing the performances of different algorithms. Environmental Science and Pollution Research, 29(25), 38346-38373, doi: 10.1007/s11356-022-18520-8, 2022

Lavecchia, A.: Advancing drug discovery with deep attention neural networks. Drug Discovery Today, 29(8), 104067, doi: 10.1016/j.drudis.2024.104067, 2024

Luan, C., & Moan, T.: On short-term fatigue analysis for wind turbine tower of two semi-submersible wind turbines including effect of startup and shutdown processes. Journal of Offshore Mechanics and Arctic Engineering, 143(1), 012003, doi: 10.1115/1.4047542, 2021

Oh, Y.R., Lee, G., Jun, K.S., Sunwoo, W., Baek, S.W., Chung, G.H.: A Study on the Prediction of Daily Snowmelt Depth using Multiple Linear Regression. Journal of the Korean Society of Hazard Mitigation, 20(6), 311-321, doi: 10.9798/KOSHAM.2020.20.6.311, 2020

https://doi.org/10.5194/wes-2024-163
Preprint. Discussion started: 27 November 2024

Pfaffel, S., Faulstich, S., & Rohrig, K.: Considering uncertainties of key performance indicators in wind turbine operation. Applied Sciences, 10(3), 898, doi: 10.3390/app10030898, 2020

Rabie, P. A., Welch-Acosta, B., Nasman, K., Schumacher, S., Schueller, S., & Gruver, J.: Efficacy and cost of acoustic-informed and wind speed-only turbine curtailment to reduce bat fatalities at a wind energy facility in Wisconsin. PLoS ONE, 17(4), e0266500, doi: 10.1371/journal.pone.0266500, 2022

Roy, A.: Atmospheric pollution retrieval using path radiance derived from remote sensing data. Journal of Geovisualization and Spatial Analysis, 5(2), 26, doi: 0.1007/s41651-021-00093-8, 2021

Strbac, A., Martini, T., Greiwe, D., Hoffmann, F., & Jones, M.: Analysis of Rotorcraft Wind Turbine Wake Encounters using Piloted Simulation. 45th ERF. Warsaw, Poland, 2019.

Sun, Z., Wang, G., Li, P., Wang, H., Zhang, M., & Liang, X.: An improved random forest based on the classification accuracy and correlation measurement of decision trees. Expert Systems with Applications, 237(Part B), 121549, doi: 10.1016/j.eswa.2023.121549, 2024

Trujillo-Franco, L. G., Abundis-Fong, H. F., Campos-Amezcua, R., Gomez-Martinez, R., Martinez-Perez, A. I., & Campos-Amezcua, A.: Single output and algebraic modal parameters identification of a wind turbine blade: Experimental results. Applied Sciences, 11(7), 3016, doi: 10.3390/app11073016, 2021

Tsai, T. C., & Wang, C. N.: Acoustic-based method for identifying surface damage to wind turbine blades by using a convolutional neural network. Measurement Science and Technology, 33(8), 085601, doi: 10.1088/1361-6501/ac68d0, 2022

Zhang, Z. Y., & Wang, K. S.: Wind turbine fault detection based on SCADA data analysis using ANN. Advances in Manufacturing, 2, 70-78, doi: 10.1007/s40436-014-0061-6, 2014

Zhao, L., Zhou, Y., Matsuo, I. B., Korkua, S. K., & Lee, W. J.: The design of a remote online holistic monitoring system for a wind turbine. IEEE Transactions on Industry Applications, 56(1), 14-21, doi: 10.1109/TIA.2019.2951088, 2019

Zhu, Y., Zhu, C., Tan, J., Wang, Y., & Tao, J.: Operational state assessment of wind turbine gearbox based on long short-term memory networks and fuzzy synthesis. Renewable Energy, 181, 1167-1176, doi: 10.1016/j.renene.2021.09.070, 2022