IRJAEH

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 12 December 2024
Page No: 2734- 2739
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0378

Performance Comparison of Machine Learning Algorithms for Wind Energy Forecasting in the Coastal Region of Kerala

V.V. Muhammed Anees[1], K P. Abdul Nazar[2], Shijin Maniyath[3]
[1,2,3]Assistant Professor, Department of ME, Government College of Engineering Kannur, Kerala, India.
Emails: *anees@gcek.ac.in[1], nazar@gcek.ac.in[2], shijin@gcek.ac.in[3]*

Abstract

This paper presents the details predicting wind energy output with machine learning models. Accurate forecasts of wind power's future are essential to the feasibility of major renewable energy projects. Making precise forecasts of wind power generation requires accounting for changes in weather patterns over time. This is also essential for issuing early warnings and implementing risk-reduction measures. In this study, prediction models for wind energy are developed using wind data collected at coastal regions in Kerala. Accurate forecasting of wind power generation is necessary to balance supply and demand in the smart grid. In the present investigation, an extensive analysis of long-term wind power forecasting was undertaken utilizing daily wind speed data, employing five distinct machine learning algorithms like XG Boost, LASSO, Gradient Boosting, Random Forest, and Bayesian Ridge Regression.

Keywords: *XG Boost, LASSO, Gradient Boosting, Random Forest, Bayesian Ridge Regression*

1. Introduction

Wind energy constitutes a renewable resource that exploits the kinetic energy of the wind for the production of electricity. Its significance within the global energy paradigm has escalated considerably owing to its numerous benefits, such as sustainability, environmental advantages, and economic feasibility. The domain of wind power is garnering heightened attention globally, attributed to its ecological benefits and inherently renewable characteristics. In contemporary contexts, the rising demand for energy, coupled with the environmental challenges associated with non-renewable energy sources, has led to a substantial increase in global interest towards renewable alternatives. However, the intrinsic variability of wind necessitates precise forecasting of wind energy to optimize the utilization of renewable energy sources effectively. In practical scenarios, the development of accurate prediction systems may present considerable challenges. Real-world time series data commonly exhibit a combination of both nonlinear and linear patterns, which are interwoven in a manner that complicates their modeling. In this context, methodologies such as machine learning (ML) and artificial intelligence (AI) are employed to achieve more precise outcomes. The proficiency of machine learning algorithms in forecasting consumption and output has shown considerable promise. The fundamental rationale for the application of machine learning algorithms lies in their ability to construct models based on initial datasets rather than relying on a generalized framework, as well as their capacity to adapt to evolving trends within the data. In this research endeavor, the objective of forecasting wind speed and subsequently estimating wind energy production has been articulated as a specific problem. To tackle this challenge, various modeling approaches have been utilized. These diverse methodologies present unique advantages in capturing intricate patterns inherent in wind data. Through a systematic comparison and insightful information regarding the strengths and weaknesses of each approach, thereby facilitating the selection of the most efficacious model for accurate wind speed forecasting and precise wind energy estimation. Such comparative analyses of models are crucial for optimizing renewable energy generation and improving the reliability of wind energy forecasts.

1.1. Preliminary Analysis

This work is mainly simulation based where the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 12 December 2024
Page No: 2734- 2739
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0378

algorithm is developed in python to forecast the wind energy. In this work, we used past wind speed data to examine the use of various machine learning methods for long-term wind power prediction with a 1-year forecast horizon. The first step is the collection of data set. The data is obtained from the website of NASA. The data set consists of different environmental features like wind direction, wind speed, temperature and atmospheric pressure at different height etc. All these are generated from the website of NASA [1-3].

2. Methodology

Predicting the wind power in relation to previous data is the main goal. Additionally, understanding the characteristics of wind power output and enhancing the precision of wind power forecast can be achieved by the analysis of the effects of various elements on wind power. To achieve this, a comprehensive machine learning framework is created, involving data preprocessing, model selection, and extensive validation. The framework's efficiency is rigorously accessed through different evaluation metrics, like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), R2 Error and Adjusted R2 error. Another objective is to analyze whether the algorithms could manage various wind characteristics with areas having high wind speeds, pressure variations, temperature fluctuations and different precipitation rate etc. The data set consists of 22-year dataset of all the Environmental parameters. For this instance, the location coastal region Trivandrum was chosen. Also, it aims to know the best algorithms in this location. This holistic approach contributes to more efficient and sustainable wind energy utilization.

2.1. Description of the Dataset

In the current work, machine learning techniques were utilized to forecast the daily total wind power based on the daily mean wind speed. The original dataset had wind speed hourly values. The hourly wind speed values were transformed into the daily mean wind speed. Due to the variable and unpredictable nature of wind speed, we choose to employ daily wind power predictions. The dataset contains certain environmental features like temperature, pressure at 10m and 50m height, relative humidity, precipitation, wind speed at 10m. and 50m height and wind direction All these are collected from the website of NASA [22]. Figure 1, The selected location is coastal region Trivandrum (Latitude: 8.31268, Longitude: 76.56118) [4-7].

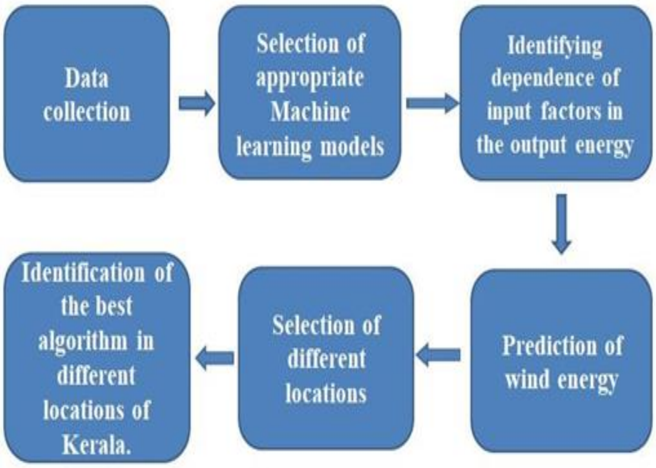

Figure 1 Methodology

2.2. Correlation between Weather Parameters and Power

The dataset contains different weather parameters such as Date, temperature, wind direction, relative humidity, precipitation, pressure and different wind speeds at 10m and 50m.Fig.2 provides correlation between weather parameters and wind power. The map illustrates that wind speed has a high correlation with wind power. Wind direction and relative humidity are other parameters which has correlation with wind power. Amount of rainfall intensity has very less dependence on wind power. The other parameters like date, pressure and temperature has very small correlation with wind speed and it is negative value.

The power from the wind turbine P as,

$$P \quad = ½ \rho A V^3$$

P_w = the wind power (in watts or kilowatts),
ρ = the air density (kg/m^3),
A = Swept area rotor (m^2),
V = speed of wind (m/s)

Theoretical wind power is represented by this equation. It is clear from this equation that wind speed has a significant role in producing wind energy. As a result of the turbine blades rotating more quickly

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 12 December 2024
Page No: 2734- 2739
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0378

due to increasing wind speed, the generator connected to the turbine will produce more mechanical power and electricity. But there are many other factors which effects the amount of power obtained from the wind turbine. That is there will be always a loss when we convert one form of energy into another. Approximately 59% of the total energy produced by wind may be converted into electrical power. This is not reliant on the turbine model and is referred to as the Betz limit. Wind turbines operating at utility scale can actually achieve 75–80% of the Betz limit, shown in figure 2.

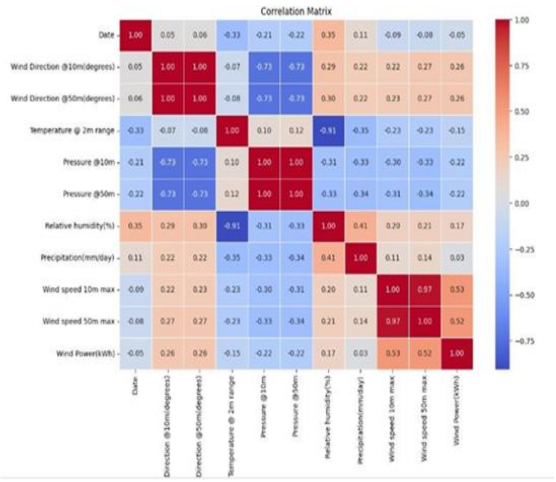

Figure 2 Correlation of Weather Parameters

2.3. Machine Learning Algorithms Used for The Comparison

The Machine learning algorithms used for the comparison for wind energy forecasting in the coastal region of Kerala are XG Boost, LASSO, Gradient Boosting, Random Forest, and Bayesian Ridge Regression.

2.4. Performance Measures

In this final stage, the actual or original values are compared with the predicted/forecasted values to determine certain performance measures. The performance measure which I have used in this work are Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R2 Value. Root Mean Square Error (RMSE) is a number that quantifies the average divergence between the predicted values of a model and the real values within a dataset. It serves as a measure to assess how effectively a model matches with the dataset. A model's ability to fit a dataset improves with a decreased root mean square error (RMSE). The RMSE can be evaluated using the following formula.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}(Ai - Fi)^2}$$

The MAE is a statistical measure that calculates the average absolute difference between individual data points and the average of the dataset. It is used to assess the dispersion or variability of data points in a dataset. MAE can be computed using the following formula.

$$MAE = \frac{1}{n}\sum_{i=0}^{n-1}|(Ai - Fi)|$$

R^2 (or the coefficient of determination) measures the percentage that the independent variable can account for in explaining the variance in the dependent variable or R^2, indicates how well the data fit the model. The value can be calculated using the following formula.

$$R^2 = 1 - \frac{\sum_{i=0}^{n}(Ai - Fi)^2}{\sum_{i=0}^{n}(Ai - \bar{F}i)^2}$$

A_i = The Actual data
F_i = The Forecasted data
n = Number of Data

3. Results and Discussion

The specified location is located in the southern section of Kerala. That region has a lot of seaside regions. The study found that the seaward region is most suitable for the development of wind energy since it has the highest effective wind speed frequency. High production areas are primarily found in outer sea areas, while developed areas are found close to shore, and tidal and sea-land interface areas are potential development zones. A coastal region serves as the study area. Thus, there is a relatively high amount of wind speed available. Figure. 3 illustrates the wind speed characteristics for a year.It is evident from the figure that there is seasonal variation. However, it is also generating a wide range of wind speeds. This makes the location as a best place for a Wind farm, shown in Figure 5 & Figure 6.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 12 December 2024
Page No: 2734- 2739
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0378

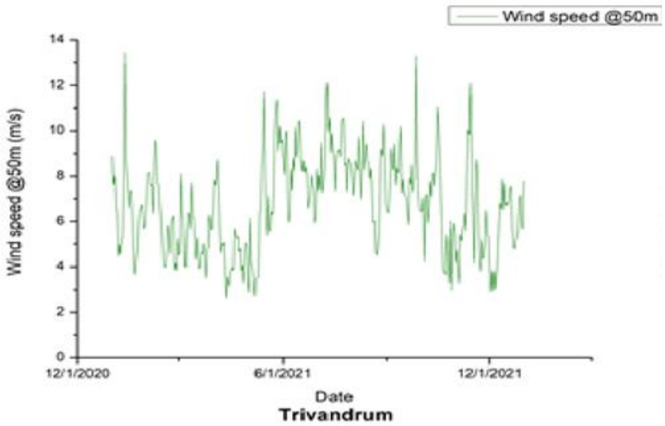

Figure 3 Wind speed 50m Hub Height at Trivandrum

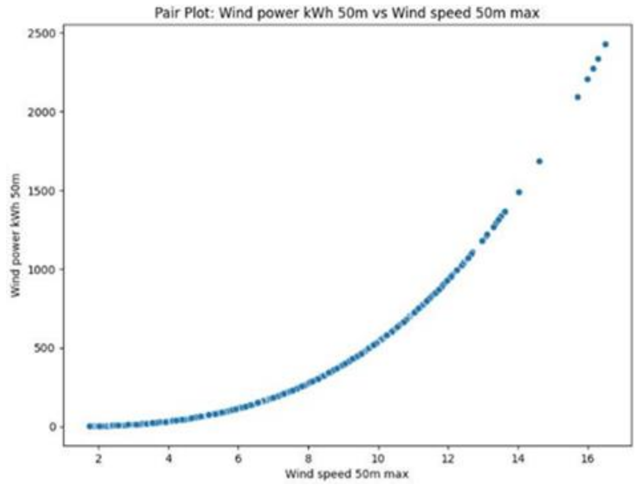

Figure 4 Correlation of Wind Power with Wind Speed at 50m Height Trivandrum

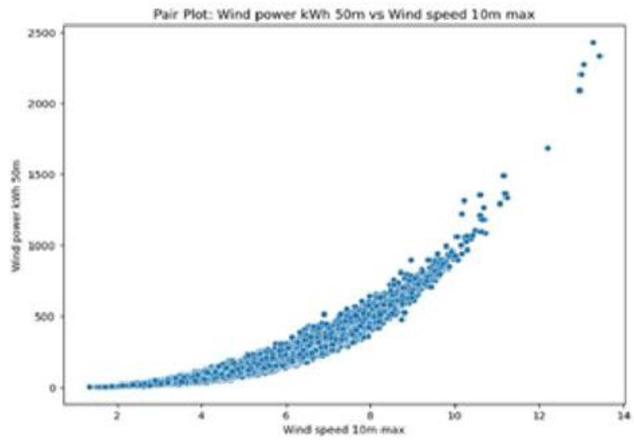

Figure 5 Correlation of Wind Power with Wind Speed at 10m Height

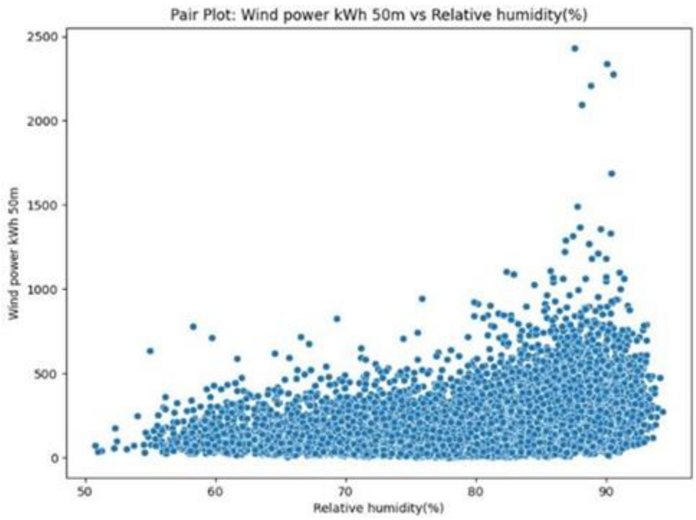

Figure 6 Correlation of Wind Power with Relative Humidity at 50m height

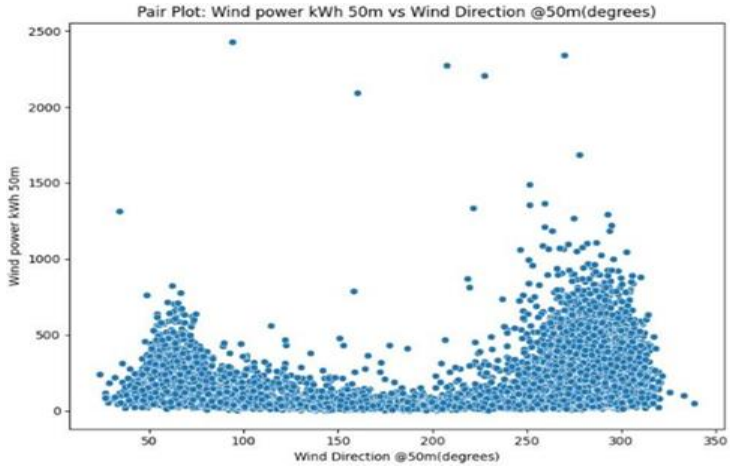

Figure 7 Correlation of Wind power with Wind Speed at 50m Height Trivand

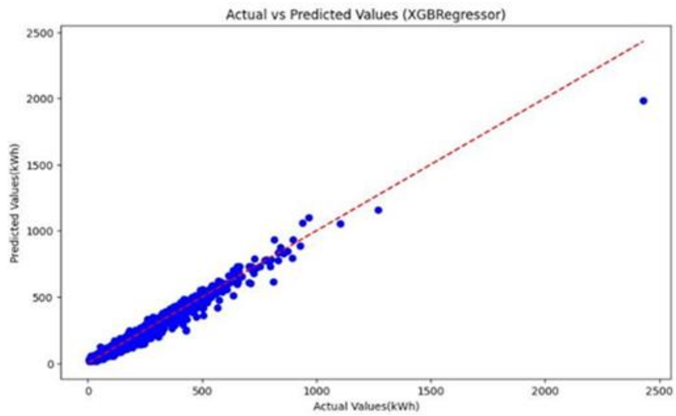

Figure 8 Wind Power Prediction by XGBOOST

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 12 December 2024
Page No: 2734- 2739
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0378

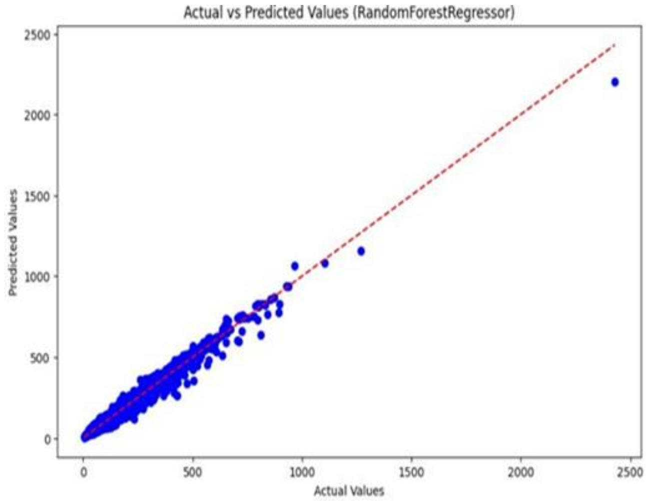

Figure 9 Wind Power Prediction by Random Forest

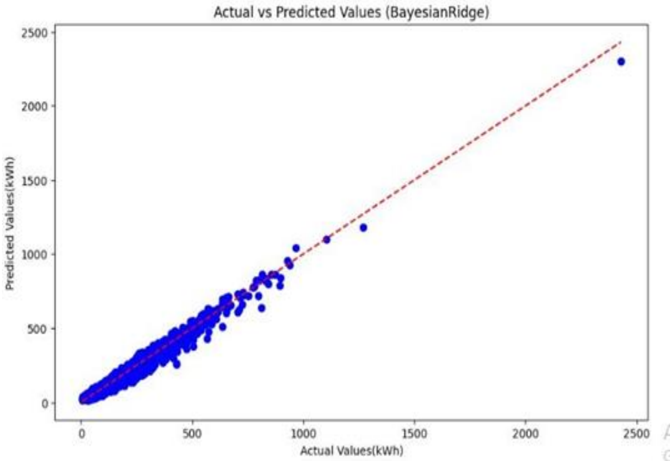

Figure 10 Wind Power Prediction by Bayesian Ridge

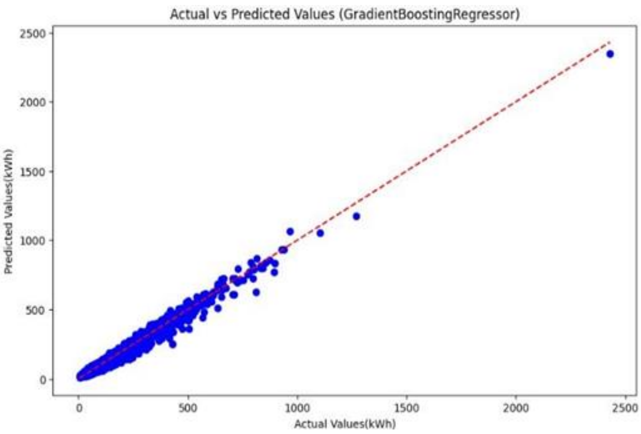

Figure 11 Wind Power Prediction by Gradient Boosting

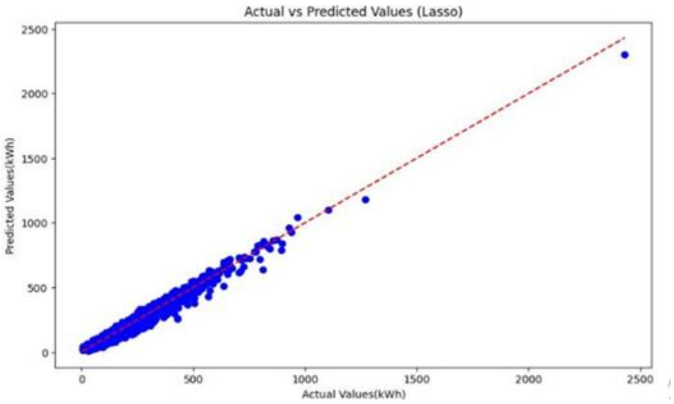

Figure 12 Wind Power Prediction by LASSO Regression

The relationship between the wind power generated at Trivandrum and other characteristics is shown in Figures 4 through Figure 7. According to this, wind power and wind speed are most correlated. In comparison, wind power has little relationship with other parameters. The relationship between actual and expected power for various algorithms is displayed in Figures 8 through Figure 12. It indicates if the predicted and actual values fit in a linear plane and displays the linearity of the power values for every algorithm. Table 1 shows that the XGBOOST algorithm is the worst of the five algorithms since it has a lower R2 and modified R2 value, as well as a high RMSE and MAE value, shown in Figure 9, Figure 10, Figure 11. Good R2 and Adjusted R2 values are provided by Bayesian Ridge, Random Forest, Gradient Boosting, and LASSO regression at the RMSE and MAE values are better for Gradient Boosting.ie;

Table1 The Performance Measures of Wind Power Forecast in Trivandrum

MODELS	R^2	Adjusted R^2	RMSE	MAE
LASSO	0.972	0.97	31.36	23.289
XG Boost	0.967	0.96	36.91	23.289
Gradient Boosting	0.973	0.97	30.36	22.111
Bayesianridge	0.972	0.97	31.32	23.291
Random Forest	0.972	0.97	31.18	22.395

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 12 December 2024
Page No: 2734- 2739
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0378

As accuracy and performance improve, the error rate falls. Other algorithms are good in comparison. Additionally, since all algorithms provide higher R^2 values, they are all acceptable, shown in Table 1.

Conclusion

Wind energy prediction using machine learning is an important and evolving field with significant potential for improving the efficiency and reliability of wind power generation. In case of a seashore area Gradient Boosting algorithm gives a better result. Because they provide a good R2, Adjusted R2 value, RMSE and MAE value. Conclusion should contain the confirmation of the problem that has been analyzed in result and discussion section. Gradient boosting, LASSO and Bayesian Ridge regression shows a better performance with large datasets. It shows comparatively high R2 value and low RMSE and MAE values [8-9].

References

[1]. Najeebullah, A., Zameer, A., Khan, A., & Javed, S. G. (2015). Machine Learning based short-term wind power prediction using a hybrid learning model. Computers and Electrical Engineering, 45. http:// dx.doi.org/10.1016/j.compeleceng.2014.07.009.

[2]. Demolli, H., Dokuz, A. S., Ecemis, A., & Gokcek, M. (2019). Wind power forecasting based on daily wind speed data using machine learning algorithms. Energy Conversion and Management, 198,111823. https:// doi.org/10.1016/j.enconman.2019.111823

[3]. Aierken Dilimulati, Ted Stathopoulos, & Marius Paraschivoiu. (2018). Wind turbine designs for urban applications: A case study of shrouded diffuser casing for turbines. Journal of Wind Engineering & Industrial Aerodynamics, 175, 179–192. doi: https://doi.org/10.1016/j.jweia.2018.01.003

[4]. Reto Spielhofer, Jonas Schwaab, & Adrienne Gret-Regamey. (2023). How spatial policies can leverage energy transitions Finding Pareto-optimal solutions for wind turbine locations with evolutionary multi-objective optimization. Environmental Science and Policy, 142, 220–232. doi:https:// doi.org/10.1016/j.envsci.2023.02.016

[5]. Katja Franke, Frank Sensfu, Gerda Deac, Christoph Kleinschmitt, & Mario Ragwitz. (2021). Factors affecting the calculation of wind power potentials: A case study of China. Renewable and Sustainable Energy Reviews, 149,111351.doi:https://doi.org/10.1016/j.rser.2021.111351

[6]. Potter, C., & Negnevitsky, M. (2006). Very Short-Term wind Forecasting for Tasmanian power generation. IEEE Transactions on Power Systems, 21(2), 965–972. https://doi.org/10.1109/tpwrs.2006.873421.

[7]. Wan, C., Xu, Z., Pinson, P., Dong, Z. Y., & Wong, K. P. (2013). Probabilistic Forecasting of wind power generation using extreme Learning Machine. IEEE Transactions on Power Systems, 29(3), 1033–1044. https://doi.org/10.1109/tpwrs.2013.2287871

[8]. Du, P. (2018). Ensemble Machine Learning-Based Wind Forecasting to Combine NWP Output with Data from Weather Station. IEEE Transactions on Sustainable Energy, 10(4), 2133–2141. https:// doi.org/10.1109/ tste.2018.2880615

[9]. Zhang, Z., Zhang, R., Fang, D., & Wang, J. (2015). Prediction study and application of wind power development based on filtering error threshold. Environmental Progress & Sustainable Energy, 34(5), 1536–1546. https://doi.org/10.1002/ep.12138