Response to Reviewer #1's Comments

Wind energy is one of the most important renewable energy sources in the world. As the author said, China is currently confronted with an increasingly serious energy and climate situation. It is necessary to evaluate the wind energy resources on China. This study attempted to use the machine learning algorithms to evaluate the wind energy resource at eight coastal stations based on the wind speed profile and surface meteorological data. Three ML algorithms and the power law method were used and compared. The wind energy resource is then evaluated based on the wind speed from RF model. Overall, it is a nice and well-organized paper with a clear focus. However, before publishing, there are some problems need to be solved.

Response: We thank the anonymous reviewer for his/her comprehensive evaluation and thoughtful comments, which greatly improve the quality of our manuscript. We have made great efforts to adequately address the reviewers' concern one by one. For clarity purpose, here we have listed the reviewer' comments in plain font, followed by our response in bold italics.

1. My biggest concern is whether the reanalysis data (such as ERA5) can be used to evaluate wind energy? In this paper, the author actually uses ML algorithm to predict accurate high-level wind speed, and then conducts wind energy assessment. It means that the accurate high-level wind speed is the most important. As far as I know, ERA5 data can also provide hourly wind speed, so can this data be used? The author should explain this point in the text. Especially in the introduction, the application of reanalysis data is not mentioned at all.

Response: Good questions! The ERA5 data can provide hourly wind speed, and can also be used to evaluate wind energy. Compared to near-surface in-situ observations or ground-based remote sensing, it has better time continuity and spatial coverage, which can provide data support in the region with poor observational data. But the spatial resolution of the ERA5 data is 0.25 degree * 0.25 degree. This spatial resolution is much lower than the high-resolution model output such as WRF and the point-based observations. The hourly resolution of ERA5 reanalysis can be used to assess the wind energy in the absence of observational data, which has been reported in previous studies (Laurila et al., 2021; Jurasz et al., 2021; Gualtieri, 2021).

According to your suggestion, we added some descriptions in the introduction. "The reanalysis data, such as the fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis system (ERA5), can provide the hourly wind speed at a specific height (Hersbach et al., 2020; Liu et al., 2020). Compared to near-surface in-situ observations or ground-based remote sensing, it has better time continuity and spatial coverage, which can provide data support in the region with poor observational data. The hourly resolution of ERA5 reanalysis can be used to assess the wind energy in the absence of observational data (Laurila et al., 2021; Gualtieri, 2021). But the maximum spatial resolution of the ERA5 data is 0.25 degree * 0.25 degree. This spatial resolution is much lower than the high-resolution model output such as the weather research and forecasting (WRF) and the point-based observations."

References:

Laurila, T. K., Sinclair, V. A., & Gregow, H.: Climatology, variability, and trends in near-surface wind speeds over the North Atlantic and Europe during 1979–2018 based on ERA5. International Journal of Climatology, 41(4), 2253-2278, 2021.

Gualtieri, G.: Reliability of era5 reanalysis data for wind resource assessment: a comparison against tall towers. Energies, 14(14), 4169, 2021.

2. P5 Line 140: Why the wind shear coefficient was set to 0.15? Is this value for a site or all sites? What is the uncertainty of this parameter? Some recent publications should be discussed, for example, Retrieval of total and fine mode aerosol optical depth by an improved MODIS Dark Target algorithm; Accuracy, stability, and continuity of AVHRR, SeaWiFS, MODIS, and VIIRS deep blue long-term land aerosol retrieval in Asia; A High-Precision Aerosol Retrieval Algorithm (HiPARA) for Advanced Himawari Imager (AHI) data: Development and verification; Accuracy and error cause analysis, and recommendations for usage of Himawari-8 aerosol products over Asia and Oceania.

Response: Good questions! In engineering application of power law method, the value of wind shear coefficient is set to a fixed value according to the terrain type (Banuelos et al., 2010). The value of wind shear coefficient is ranged from 0.1 to 0.4 (Li et al., 2018). Here, the general value of α for coastal topography was set to 0.15

based on former studies (Patel et al., 2005; Banuelos et al., 2010). Such a setting will inevitably lead to certain errors, but we have not found clear uncertainties in the literature. Therefore, we propose to use machine learning to retrieve wind speed. We added a description to explain the problem. "In engineering application, the value of α is determined by the terrain type, and the variation range is from 0.1 to 0.4 (Li et al., 2018). Here, the general value of α for coastal topography was set to 0.15 based on former studies (Patel et al., 2005; Banuelos et al., 2010)." In addition, the relevant references were also discussed in the manuscript.

References:

Banuelos-Ruedas F, Angeles-Camacho C, Rios-Marcuello S. Analysis and validation of the methodology used in the extrapolation of wind speed data at different heights. Renew Sustain Energy Rev., 14(8):2383-91, https://doi.org/10.1016/j.rser.2010.05.001, 2010.

Li JL, Yu X. Onshore and offshore wind energy potential assessment near Lake Erie shoreline: A spatial and temporal analysis. Energy., 147: 1092-107, https://doi.org/10.1016/j.energy.2018.01.118, 2018.

Patel, M. R.: Wind and solar power systems: design, analysis, and operation. CRC press; 2005.

3. P5 Line 153: The author used a 5-fold crossover to train the model. What is the total number of samples? How to allocate samples for each training?

Response: Model training is conducted at Qingdao station. A total of 746 sample data are obtained after data matching. 5-fold cross validation means that 80% data is used for training and 20% data is used for testing each time. Repeat five times to traverse all data. We added an explanation to the manuscript.

4. Fig. 2: This figure is totally unnecessary. KNN, SVM and RF are very classic ML algorithms. It is not necessary to explain its principle. I suggest deleting it or changing it to a supplementary file.

Response: We change the Fig. 2 to the appendix. At the same time, the description of the ML models has also been modified.

5. Fig. 11: Similar to the first comment, the ERA5 data can also provide hourly wind speed, so can this data be used to evaluate wind energy? Here, the author used ERA5 data to evaluate wind energy, but did not discuss the application of ERA5 data. If the precision of ERA5 data is high enough, can the ERA5 data be used directly.

Response: The ERA5 data can provide hourly wind speed, and can also be used to evaluate wind energy. We explained and discussed the application of ERA5 data in the introduction.

"Compared to near-surface in-situ observations or ground-based remote sensing, it has better time continuity and spatial coverage, which can provide data support in the region with poor observational data. But the spatial resolution of the ERA5 data is 0.25 degree * 0.25 degree. This spatial resolution is much lower than the high-resolution model output such as WRF and the point-based observations. The hourly resolution of ERA5 reanalysis can be used to assess the wind energy in the absence of observational data, which has been reported in previous studies (Laurila et al., 2021; Jurasz et al., 2021; Gualtieri, 2021)."

6. There are many grammatical and spelling mistakes in the text. Please correct them carefully.

Response: We sought professional organizations to modify the language of the manuscript.