1. There are many speed forecasting method available in literatures? How this method is specifically different from other works in terms of performance in prediction? Whether the authors did any comparison with existing works available in literatures? This shall be included in the results section.

Response:

Thank you for your insightful comments regarding our work. We appreciate your attention to detail and your questions concerning the distinction of our speed forecasting method. Considering that the complex hybrid model prediction methods mentioned in the literatures are tailored to specific prediction durations and application scenarios, these hybrid models may not necessarily be applicable to the target wind field of the experiments in this research paper. Therefore, in the subsection 'Comparison experiment' of the results section, we use the popular machine learning methods mentioned in the literature to build hybrid models to predict the wind speed and visualize the prediction performance of this paper's method through the comparison of various error indicators.

2. How the historical data from April to July 2023 is alone sufficient to predict data for August 2023? Whether this work considers any uncertainties?

Response:

Thank you for your review and in-depth comments on our work. Regarding the limitations of the historical data we used in our study, we would be happy to provide a more detailed explanation of the data sources and usage. Our data came from designated wind farms that provided limited historical data to April through July 2023, which places some limitations on our ability to analyze and forecast. Hourly or daily forecasts in short-term forecasting models can be made simply by using a few months of historical data.

We used the methods mentioned in the article for wind speed forecasting, which typically capture seasonal variations and short-term trends. And we used the following countermeasures to cope with uncertainty: first we used flexible predictive models (time series analysis and machine learning algorithms) to extract as much as possible the patterns of the data over the time period and make predictions. In addition, we performed cross-validation to assess the consistency and stability of the models on different slices of data to improve the confidence of the forecasts. Finally, in the robustness experiments, we also conducted experiments using data of several months' length from another wind farm to ensure the stable type and validity of the BCRBR model. We realize that more data will help improve the accuracy of the model, so in subsequent studies, we plan to work with additional data providers to obtain a wider range of historical data.

3. Is there any other data considered such as moisture, or air density for wind speed prediction?

Response:

Thanks for your careful suggestion. Among the features used in the wind speed prediction model as shown in the article, we considered and used the 2-meter humidity as one of the input features. In more complex models and application scenarios, air density is a factor of interest. However, in the WRF data we use, there is no air quality output, and the data are difficult to access. And we may prioritize the historical data of wind speed and other variables directly related to wind speed in the wind speed prediction process, rather than introducing additional complex features such as air density. During the experiment, the data in the test set may not be coherent after the label prediction step in different modules. We pay more attention to the accuracy of the data used and the correlation of wind speed to ensure the validity of the input features and avoid overfitting situations during training.

4. In Figure 6, which characterizes the results of the BCMMC model, the colors used to distinguish the individual clusters may not be easily distinguishable. It is recommended to use an alternative color scheme for better recognition.

Response:

Thank you for your valuable comments on the figure. We have adjusted the color scheme of Figure 6 to improve the recognition effect.

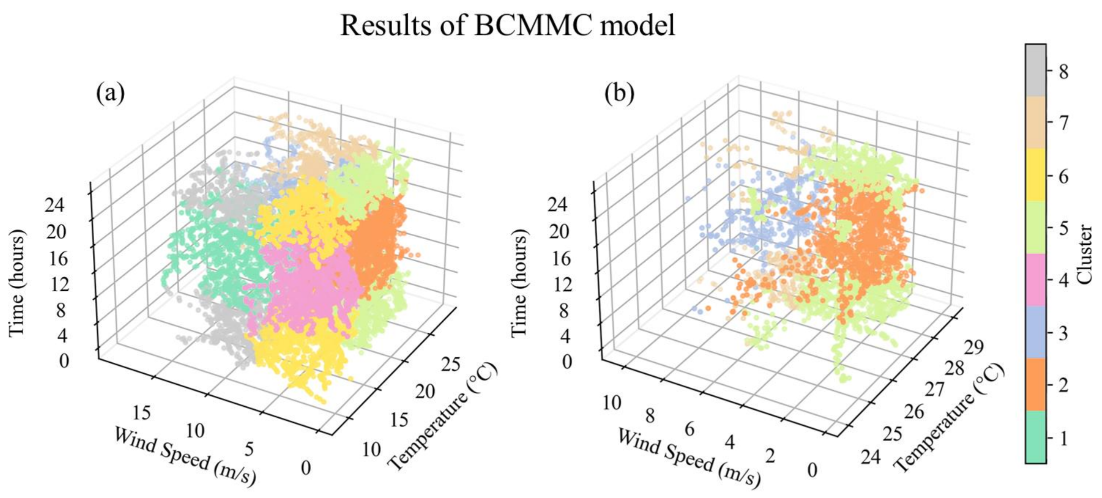

Figure 6: Results of the BCMMC model. (a) Bi-clustering classification, and (b) label prediction.

5. Why don't the authors try normalized data for training instead of actual data?

Response:

Thank you very much for your valuable suggestion. As we explained in the data preprocessing in Step 1 of Section 2.2.1, we standardized all the actual data and used standardized data to ensure that the training process of the model is more stable and effective. The reason we used standardization instead of normalization is because wind speed data are volatile with extreme maximum and minimum values, so normalization is not recommended for processing. Standardization is more suitable for algorithms that require data to have a uniform distribution, especially since many machine learning models are sensitive to the distribution of features.

6. Although the paper summarizes the advantages and potential applications of the BCRBR model in the conclusion section, the discussion section lacks an in-depth exploration of the model's potential limitations and directions for future improvements. It is recommended that the authors further discuss the model's limitations in the discussion section, such as its dependence on specific wind field conditions and computational resource requirements, and propose possible directions for future research improvements.

Response:

Thank you for reviewing our paper and for your valuable comments. We agree that the

Conclusion section needs further expansion and in-depth description. We have rewritten the Conclusion section based on your comments:

4 Conclusions

This study presented a multimodal short- and medium-term wind speed prediction correction approach based on bi-clustering and machine learning, aiming to address the crucial scientific problems of incomplete consideration of influencing factors and limited feature extraction in traditional wind speed prediction. In this paper, the BCRBR model was innovatively constructed, with offshore wind farms as the research object, and the complex influence mechanisms of sea-land breeze and atmospheric circulation on wind speed were thoroughly considered from a meteorological perspective. By introducing bi-clustering strategies and innovative features such as atmospheric circulation indices, precise correction of wind speed predictions was achieved in this study. Compared to the traditional WRF model, the errors of the corrected wind speed predictions were significantly reduced, specifically the following. RMSE, rRMSE, MAE and MAPE all decreased by more than 60%, the wind speed forecast accuracy rate increased from 30.2% to 78.4%, and the correlation coefficient R increased from 0.77 to 0.9.

This research not only offers new thoughts for wind speed prediction in methodology, but also provides data source guarantees to enhance the accuracy of wind power prediction. The research findings have significant theoretical value and practical significance in promoting the large-scale development of wind power and the sustainable development of power systems. The results of comparative experiments demonstrated that the BCRBR model outperforms the prevalent single-model prediction methods in the current literature. The results of the robustness experiments indicated that this method also performed better than the traditional WRF model in the prediction of wind speed in mountain wind farms in terms of various performance indicators, but there was a slight disparity compared to the results in offshore wind farms, which provided an important direction for subsequent improvement. Future research will focus on the following directions to improve the accuracy of wind speed prediction in mountainous areas:

(1) Integrating terrain factors and surface parameters—such as elevation, slope, and roughness—as input features for wind speed prediction models.

(2) Enhancing the quality of observational data, especially in complex terrain regions, by implementing data quality control protocols to supply more reliable samples for model training and validation.

(3) Refining model parameters and extracting informative features to improve prediction performance under complex terrain conditions, thereby broadening the applicability of the method and supporting efforts to mitigate instability in renewable energy generation.