Response to reviewer comments

Dear Editor and Reviewers,

We are very grateful for your time and valuable comments, which we found very helpful. We have addressed questions and comments raised by 4 reviewers in the revised manuscript with tracked changes. Please find our point-by-point response (in blue font) to the comments below. We hope our revisions have properly addressed your concerns.

Thanks again for your time.

Sincerely,
The authors

Reviewer 1

In this paper, the authors present interesting methods for wind speed corrections from the NWP model with multi-step methods. Below are a few minor suggestions for revision:

1. The main issue that I see in this paper is the short period for training and testing of the model, and the authors claim from this that the model is robust. Similar studies for wind speed correction from NWP models usually use several years for training and at least one year for testing. As I understood, this paper is trained only on data from February 2022, and the main conclusions are based on testing in December 2021 and January 2022, with some additional verification of stability over 10 months.

Response: Many thanks for pointing this out. While it is true that similar past studies for wind speed correction from NWP models usually use several years for training and at least one year for testing and our periods are shorter, the size of our data set is sufficient, if not greater than others'. For example, Sun et al. (2019) used a data set that contained 1827 days, from January 2012 to December 2016, using 143 grid points with a resolution of 0.5°*0.5° predicted by ECMWF, followed by 24 features for each sample, with a training set size of 1827*143*24 for each prediction time. Meanwhile, the size of our training set mentioned in lines 238-242 is about 2160*410*12. Therefore, even though it only took us a month to train, we actually trained millions of data; Second, the training data we used was obtained through daily operational runs of numerical weather forecasting, so we would have to run it for several years to get an equal amount of training data. The data we tested were mainly used to analyze the spatiotemporal changes after the model revision in December 2021 and January 2022. All the indicators of the proposed model (VMD-PCA-RF) are relatively robust for the other eight months. We will continue to add new training datasets going forward, however, it will be a challenge to train data over several million levels.

2. order of figures in the text: Fig. 1, Fig. 2, Fig. 3, Fig. 6, Fig. 4, Fig. 5, ... Fig. 11,

Fig. 14, Fig. 12.
Response: Many thanks for your suggestion. We have adjusted the order of the figures.

3. Sometimes authors refer to figures in the text as "Fig. NN" in other cases as "Figure NN", and even once as "figure NN". According to Journal rules, I think it should always be "Fig. NN." Fig. 6 and 9 are unreadable.
Response: Many thanks for your suggestion. We have corrected them all to "Fig. NN".

4. On lines 56–57, the authors state that "Currently,..." and cite a publication from 1999, but there are more recent publications for the HIRLAM model or consortium.
Response: Many thanks for your suggestion. We have updated to a more recent reference.

5. The authors claim in line 520 that "In general, VMD-PCA-RF is the best wind speed correction model for winter and even throughout the entire year in the five southern provinces," while on Fig. 14 for 2022-01, VMD-PCA-lightGBM is better.
Response: Thank you very much for pointing this out. As seen in Table 4, although VMD-PCA-lightGBM model has the best indicators for January 2022, compared to VMD-PCA-RF, the errors of the two models in various indicators are very small, and the error of MAE and RMSE is only 0.01 m/s. However, in Fig.14, the VMD-PCA-lightGBM model performed worse than VMD-PCA-RF in all of the other 9 months except January 2022.
To clarify this, we have added the following in the text: “In general, VMD-PCA-lightGBM is the superior wind speed correction model for the winter, and VMD-PCA-RF performs the best throughout the entire year in the five southern provinces.”

6. There should be more clarification about observational data. In line 132, the authors wrote "For the purposes of this paper, the 10-meter wind speed data is interpolated across 410 sites". Are those 410 sites the weather stations? Why did the authors use interpolation from this database instead of observations from stations?
Response: Thank you for drawing our attention to this.
The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html).
These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation, and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites.

To clarify this and we have rewritten the Data part into the following text:

"*The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set. According to the description of the documents on the official website (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html), the dataset is constructed through the integration of multiple sources, including ground and satellite data, and is refined using advanced techniques such as multi-grid variational assimilation, physical inversion, and terrain correction. This dataset exhibits superior quality in comparison to other products, offering higher spatial and temporal resolutions. The target observation data includes 2-m air temperature, 2-m specific humidity, 10-meter wind speed, surface pressure, and precipitation. These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites. We select the 10-meter wind speed data of 410 sites, as illustrated in Fig. 1.*"

Reviewer 2

Zhou et al. present a series of machine learning models including VMD-PCA-RF, a combination of Variational Mode Decomposition, Principal Component Analysis, and Random Forest, for correction of errors in WRF-predicted wind speeds. The manuscript presents various machine learning algorithms and uses two sets of experiments with different approaches to arrive at the model with better predictive capabilities. Accurate prediction of wind speed is important for the wind energy market for effective harvesting of wind energy, and this manuscript has potential in improving predictive capabilities for such uses. As a modeling paper it is fit for the scope of GMD, but the manuscript as presented has major shortcomings primarily in its presentation that require major revisions before further considation.

Major comments:

- Many of the figures in the text are unclear both in presentation and in purpose. Generally, the use of figures to illustrate points and their order in the text should be deliberate and help the flow of the reader to understand the text.

For example, figure 1 shows the elevation map of five southern provinces in China where observational data is used. Figure 2 shows the WRF simulation domain which appears to be a direct figure output from the WRF Pre-Processor (WPS). What is the purpose of these figures? It could be merged into one figure where the observation sites and provinces are marked. The purpose of the elevation maps in the analysis only shows up very late in the text in Section 4.2 about the RF feature importance and is not immediately clear to the reader.

Response: Many thanks for your suggestion. Fig. 1 shows the observation data evenly distributed in the five southern provinces, with the purpose of introducing specific locations of the observation data. The purpose of Fig. 2 is to illustrate the scope of WRF nesting regions. We took your advice and have combined the two figure. The combined figure is shown in Fig. 1* below and replaced in the manuscript.

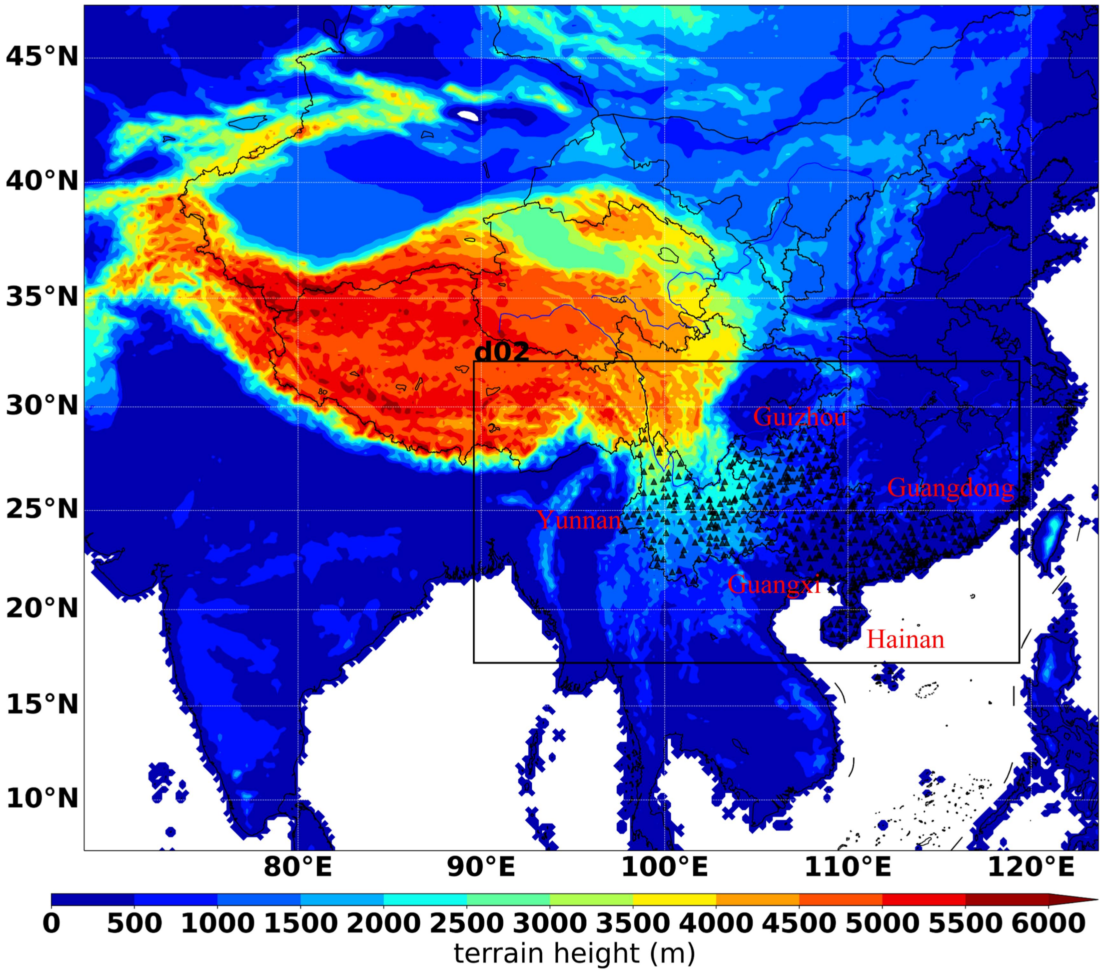

Figure 1*: WRF model simulation area elevation diagram. (d02 represents the nested area of the second layer of the WRF model, and the black triangles represent the meteorological sites).

2. Why was Lechang, Guangdong chosen for Figure 5, and where is this site, was it especially chosen? What is the purpose of the figure to the reader?

Response: Thank you very much for pointing this out. In Fig. 4 and Fig. 5, we randomly selected one (Lechang, Guangdong) of the 410 sites as a case study. Fig. 4 shows the meteorological elements of the station training and the division of the training set and the validation set. Fig. 5 shows the three-dimensional view of 12 wind speed components of the 10-meter forecast wind speed after VMD and PCA processing at the station in experiment 2.

3. In terms of presentation, Figure 3 bottom half is very unclear. The right section of Step 2 is completely unreadable. Step 3 - what do the colored boxes mean? Does their width represent some information? Define the error metrics (FA, ...) before presenting the figure;

Response: Thank you very much for pointing this out. We have changed the information at the bottom of Figure 3 and in step 2. The right half of Step 2 mainly introduces the brief architecture of DBN, MLP, RF, XGBoost and lightGBM models. Each color box represents the training set, the validating set, and the testing set. In this paper, we have introduced the evaluation indicators, such as FA before Fig. 2 below.

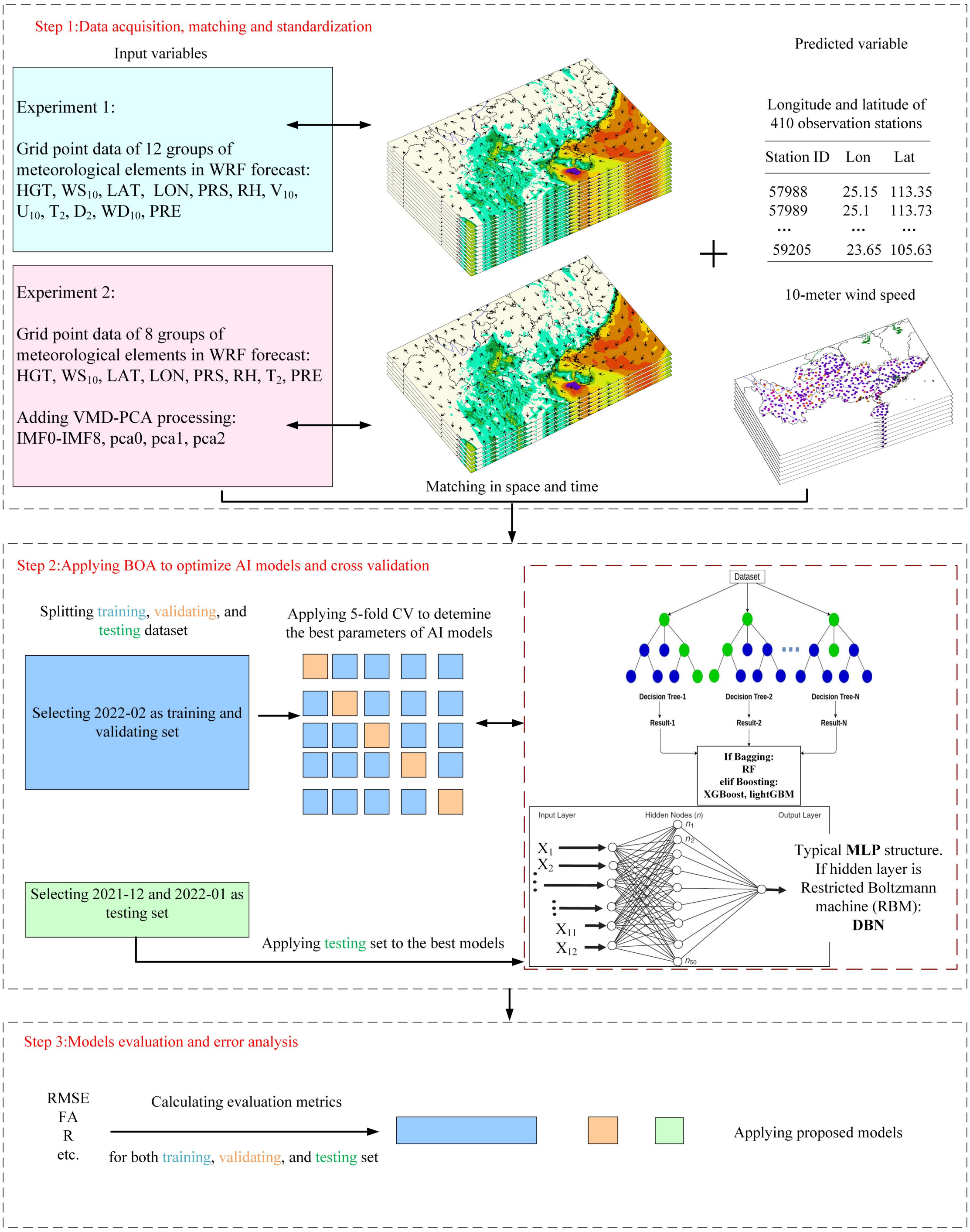

Figure 2: Flowchart of the AI model used to correct WRF-predicted wind speeds in the two main experimental pathways.

4. Figure 4 text is unreadable and the colors do not help discern the lines. Make the lines bolder. The backgrounds could just be white and grey to represent the training+validation & the test sets (label them with a legend).

Response: Many thanks for your suggestion. We have revised it according to your suggestion. The revised figure is shown in Fig. 3 below and in the manuscript.

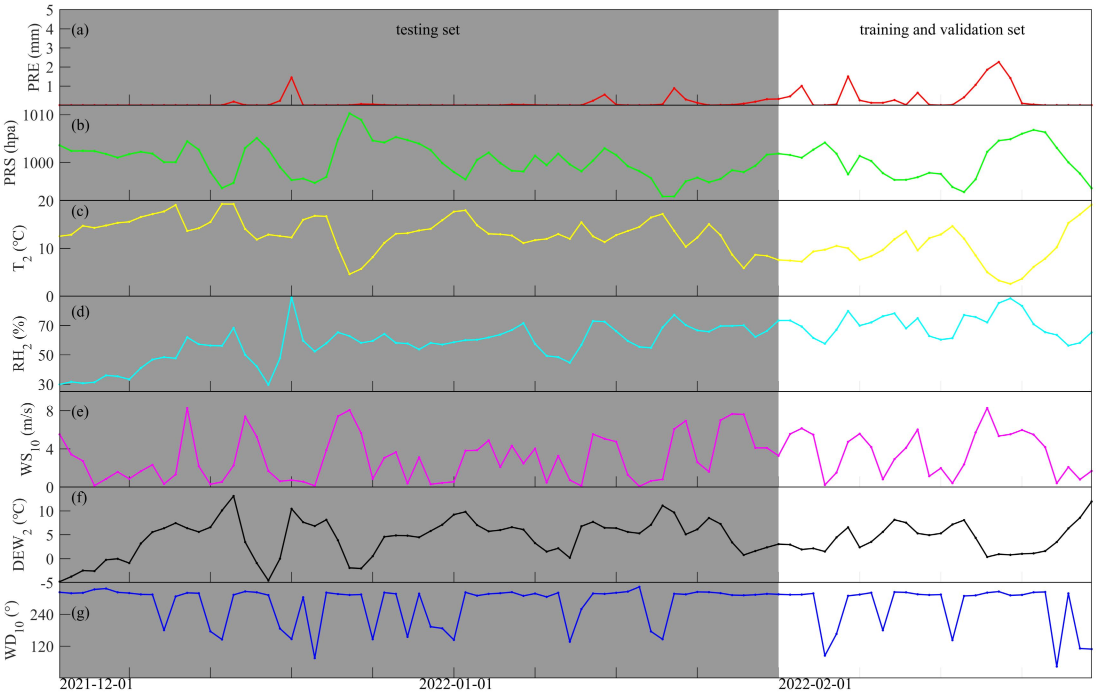

Figure 3: Daily average hourly rainfall (a), surface pressure (b), 2-meter temperature (c), 2-meter relative humidity (d), 10-meter wind speed (e), 2-meter dew point temperature (f), and 10-meter wind direction (g) which are located at Guangdong Lechang Station from December 1, 2021, to February 28, 2022. (February 2022 represents the training and verification sets, and December 2021 to January 2022 represents the testing set).

5. Figure 6 text on the right side is unreadable. Are the specific correlation coefficient text useful to the reader? The colorbar could be sufficient to illustrate the importance. The colorbars of the left and right panels could be the same size. Also, define the feature abbreviations in text as it is impossible to understand the figure and the corresponding feature names in the text if they're not clearly defined. Label the experiments 1 and 2 in Figure 6.

Response: Many thanks for your suggestion. Sure! We have added the explanation of the Pearson correlation coefficient (R) in the paper. We have added the following in the text: "*As illustrated in Fig. 5a, b, WS_{10} showed the strongest positive correlation with WS_{obs}, with the highest R of 0.51, which was consistent with the highest variable importance value of 31 % (23 %) in experiment 1 (experiment 2). In addition to WS_{10}, experiment 1 (experiment 2) also had another three dominant variables namely, LAT, HGT, and LON, with importance values of 16 % (14 %), 15 % (15 %), and 15 % (13 %), respectively. Meanwhile, in experiment 2, IMF0 and pca0 generated by VMD-PCA algorithm have a good importance value of 9 % and 4 %, and the R values of them with WS_{obs} are as high as 0.47 and 0.45.*"

We have also defined the feature abbreviations in the text. We have added the following in the text: "*In Experiment 1, as shown in Fig. 2, 12 sets of data are selected from the WRF forecast field, including altitude (HGT), 10-meter wind speed*

(WS_{10}), latitude (LAT), longitude (LON), surface pressure (PRS), relative humidity (RH), 10-meter meridional wind (V_{10}), 10-meter zonal wind (U_{10}), 2-meter temperature (T_2), 2-meter dew point temperature (D_2), 10-meter wind direction (WD_{10}), and hourly precipitation (PRE)."

We have added the following in the text: "*In this experiment, the wind speed is decomposed into 9 Intrinsic Mode Functions (IMFk, k=0, 1, 2, ..., 8) using VMD. Subsequently, a low-dimensional wind speed vector is extracted from the 9 IMF components via PCA dimensionality reduction (pca0, pca1, pca2), and all data are concatenated to construct the input factors for the model in Experiment 2.*"

We have labeled experiments 1 and 2 in Fig. 5.

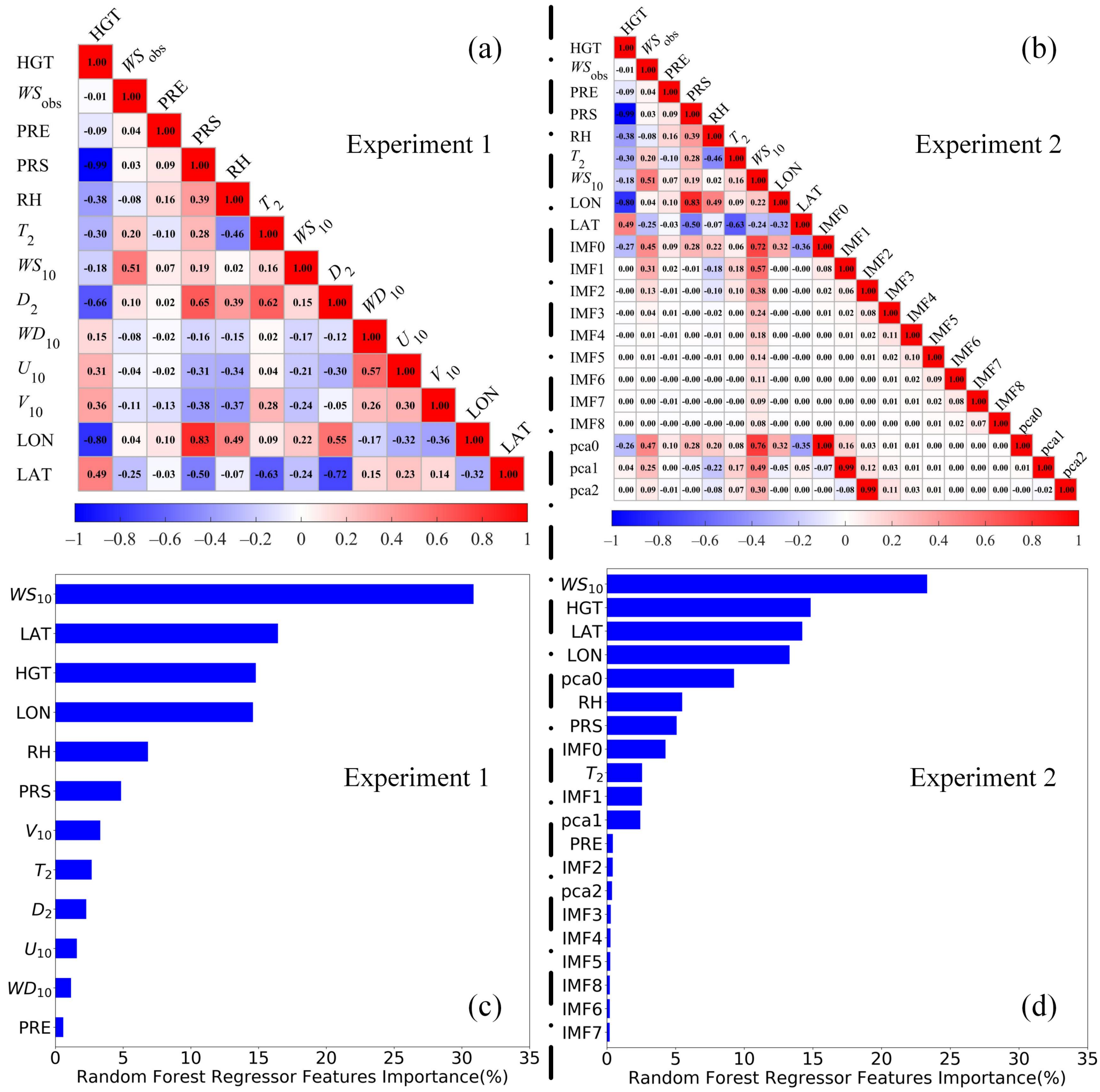

Figure 5: Schematic diagram of correlation and feature importance for two sets of experiments. (a) and (c) represent experiment 1, and (b) and (d) represent experiment 2.

6. Text in Figure 9, 14 is too small.

Response: Thank you very much for pointing this out. We have enlarged the font in Fig. 8 and Fig. 13.

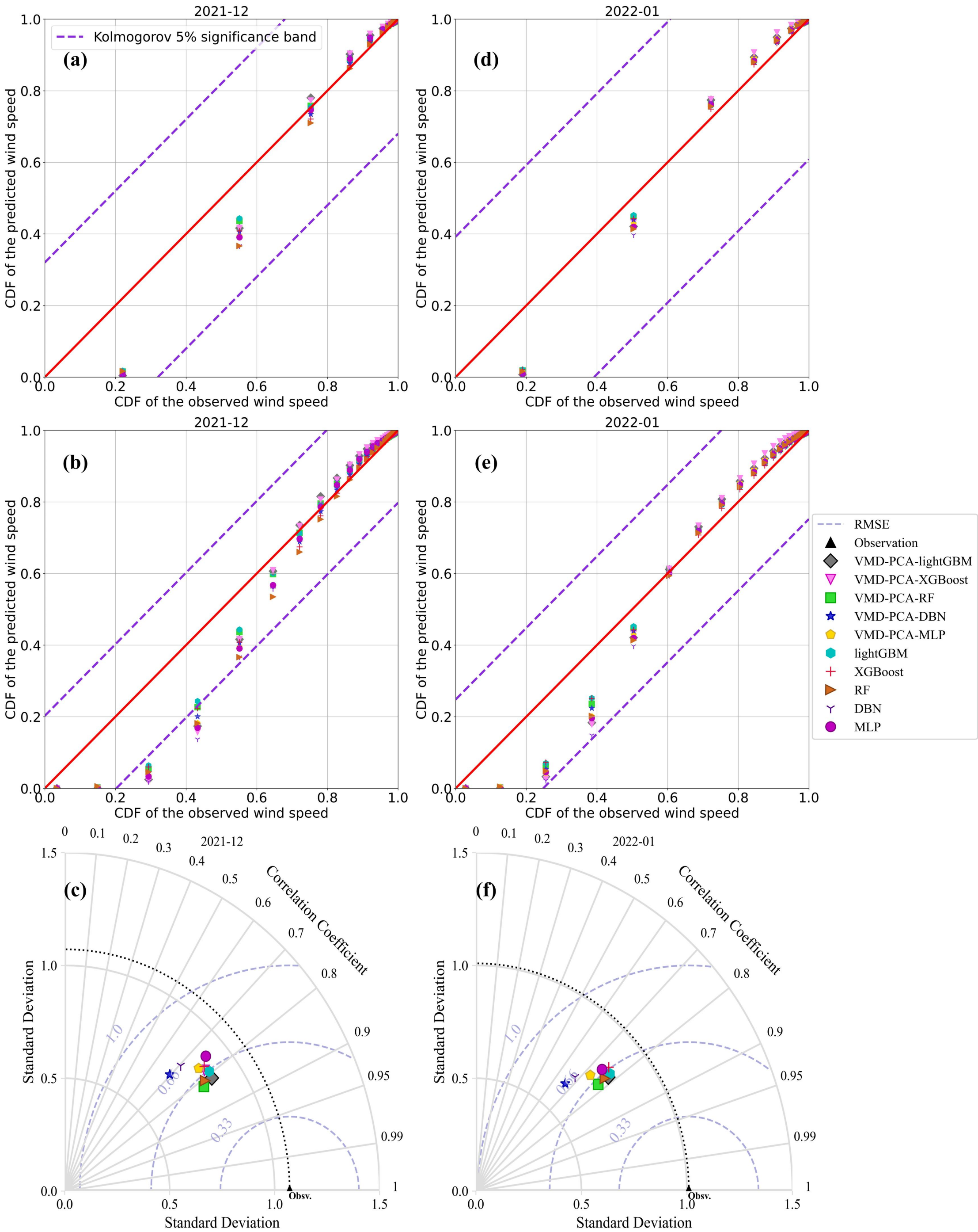

Figure 8: The cumulative distribution probability scatter plots of the actual wind speed and the predicted wind speed of 10 models in wind speed intervals of 0.5 m/s ((a) represents December 2021, (d) represents January 2022) and 0.2 m/s ((b) represents December 2021, (e) represents January 2022) respectively; Taylor distribution map ((c) represents December 2021, (f) represents January 2022).

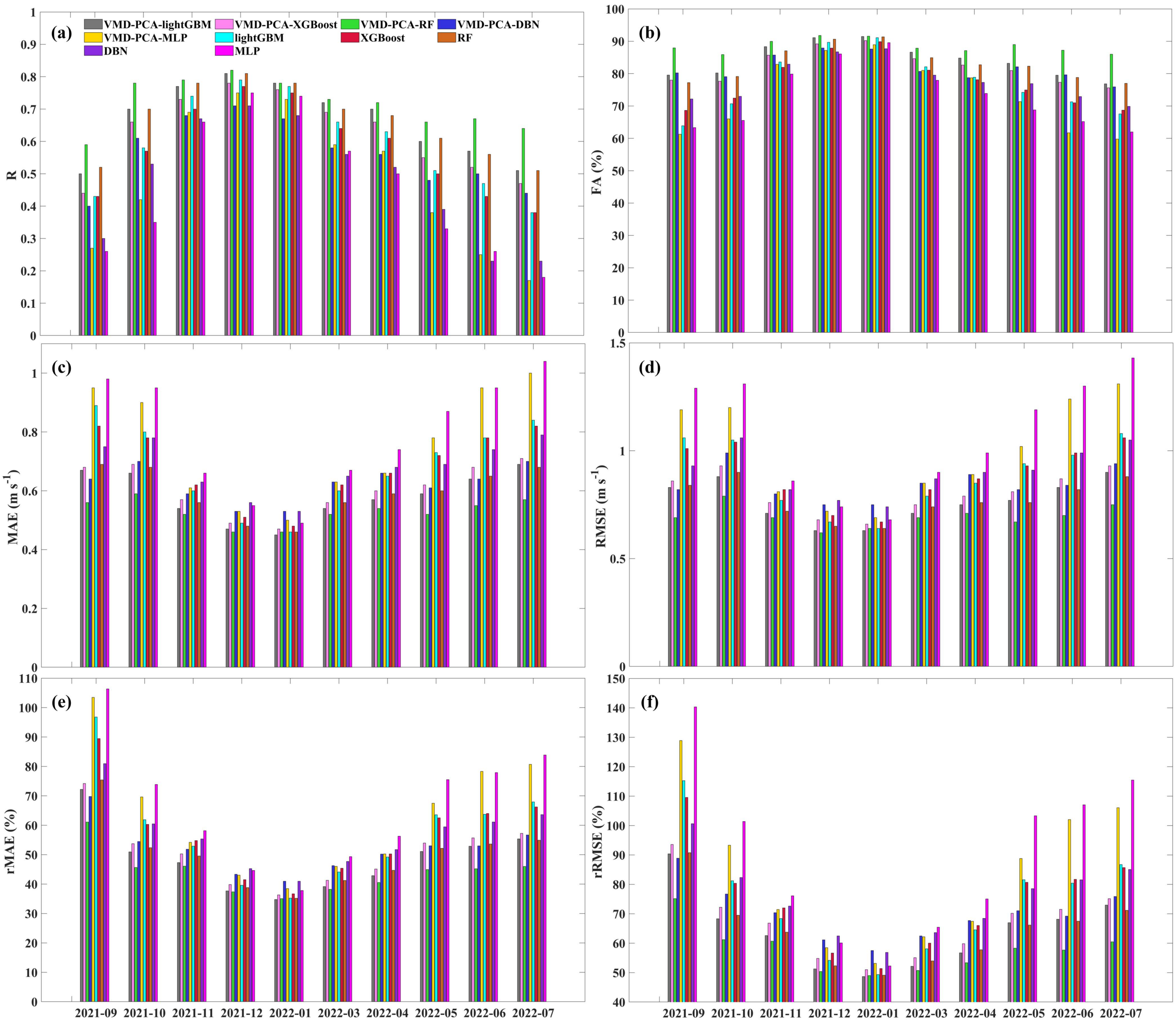

Figure 13: Evaluation histograms of 10-meter wind speed predicted by 10 models and actual wind speed in different months in Experiment 1 and Experiment 2 ((a), (b), (c), (d), (e), and (f) represent R, FA (%), MAE (m/s), RMSE (m/s), rMAE (%), and rRMSE (%) respectively).

7. - As voiced by Reviewer #1, the model was trained mostly based on winter data (DJF). Would the use of data from other seasons help the prediction?

Response: Thank you very much for your question. The purpose of this paper is to compare 5 various machine learning methods, try to introduce additional wind velocity volume, and finally get a hybrid machine learning method with highest robustness and highest wind velocity correction accuracy. Our model was trained in February 2022. the size of our training set mentioned in lines 238-242 is about 2160*410*12. Therefore, even though it only took us a month to train, we actually trained millions of data. It is unclear whether using data from other seasons instead of winter would help with the prediction. But one thing is certain, in general, for machine learning models, the introduction of more training data will improve the prediction effect to a certain extent.

8.The observational dataset presented in Section 2.1 is unclear. Where does this observational dataset come from? Did the authors create this blended data set, and if so what is the source data and the relevant citations?

Response: Thank you very much for pointing this out. The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html).

These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation, and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites.

To clarify this and we have rewritten the Data part into the following text:

"*The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set. According to the description of the documents on the official website (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html), the dataset is constructed through the integration of multiple sources, including ground and satellite data, and is refined using advanced techniques such as multi-grid variational assimilation, physical inversion, and terrain correction. This dataset exhibits superior quality in comparison to other products, offering higher spatial and temporal resolutions. The target observation data includes 2-m air temperature, 2-m specific humidity, 10-meter wind speed, surface pressure, and precipitation. These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites. We select the 10-meter wind speed data of 410 sites, as illustrated in Fig. 1.*"

8. Specific comments:

- Figure 2 has a contour map but it is not labeled (I assume this is topography), only the unit (m) is specified.

Response: Thank you very much for pointing this out. We have corrected the label of contour map and incorporated Fig. 2 (original) into Fig.1* below (updated).

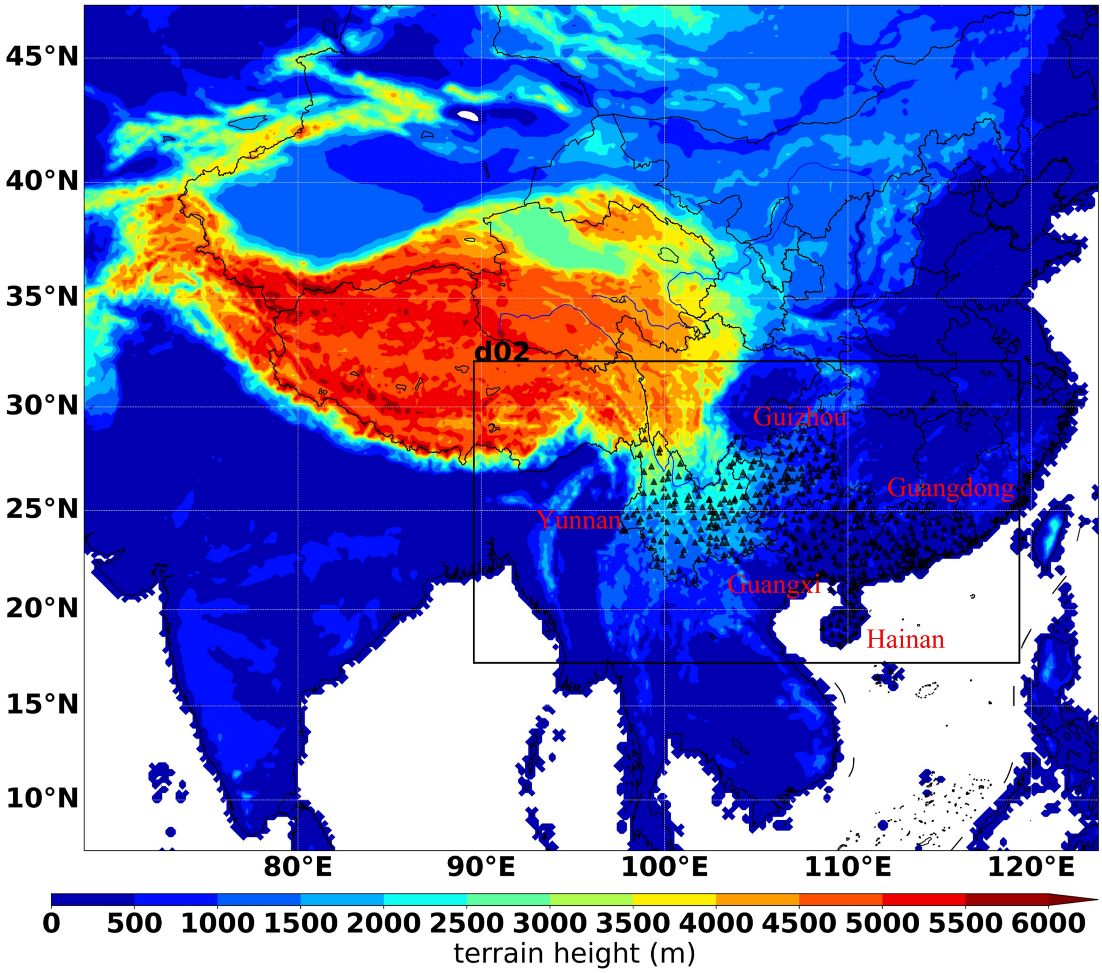

Figure 1*: WRF model simulation area elevation diagram. (d02 represents the nested area of the second layer of the WRF model, and the black triangles represent the meteorological sites).

9.- Line 16: "safe"? Elaborate on the purpose of wind speed prediction for use of wind speed resources.

Response: Thank you very much for pointing this out. In lines 45-48, we summarize the following conclusions based on the literature (Guo et al., 2021; Xiong et al., 2022; Tang et al., 2021): *"Therefore, accurate and stable wind speed prediction (WSP) is very important for the safe and stable operation of the power grid system and improving the utilization rate of wind energy and economic development (Guo et al., 2021; Xiong et al., 2022; Tang et al., 2021)."* Of course, the purpose of accurate prediction of wind speed is also efficient use of wind speed. Therefore, we have updated line 16 to the following text: *"Accurate wind speed prediction is crucial for the safe and efficient utilization of wind resources."*

10.- Line 26: Define "BOA" here.

Response: Thank you very much for pointing this out. "BOA" is defined in line 23 as "Bayesian Optimization Algorithm (BOA)."

11. - Line 26: Why "debug"? Is there a bug in the models? I suggest "analyze".

Response: Thank you very much for pointing this out. We originally used "debug" to mean parameter selection and optimization. There's not a bug in the models. we have corrected line 25-27 to the following text: "*We then perform two sets of experiments with different input factors and apply BOA optimization to tune the four artificial intelligence models, ultimately building the final models.*"

12.- Line 33 shows many metrics of the presented model compared to observations. How much better is this against WRF-predicted values before correction?

Response: Thank you very much for your question. We have already expressed the R, FA, and RMSE of WRF-predicted values before correction from September 2021 to June 2022 in FIG. S10 of the supplementary materials and Fig. 7c and Fig. 7e. As seen in the figures above, WRF evaluation indices for 10 months remain relatively poor: correlation coefficient R is below 0.59, accuracy rate FA is below 52%, RMSE is above 1.77m/s. Therefore, The VMD-PCA-RF model proposed can effectively correct the wind speed predicted by WRF and greatly improve the accuracy of wind speed correction.

13. - Line 43-45 talks about the decline of wind markets. Could authors elaborate on the relationship of this to wind speed prediction? It could be more useful for the reader to understand how better wind speed prediction serves the wind energy markets.

Response: Thank you very much for your question. Accurate wind speed prediction is of great significance for the operation and grid connection of wind farms (Huang et al., 2019).

We have added the following in the text: "*The instability and unpredictability of wind power generation can lead to instability in the power system. In addition, the decline of the wind energy market also makes it more challenging to improve the accuracy of wind speed forecasts. An accurate wind speed prediction method is needed to reduce the instability risk of power system and the economic loss of wind power enterprises (Huang et al., 2019).*"

Huang, Y., Yang, L., Liu, S., and Wang, G.: Multi-Step Wind Speed Forecasting Based On Ensemble Empirical Mode Decomposition, Long Short Term Memory Network and Error Correction Strategy, Energies, 12, 1822, https://doi.org/10.3390/en12101822, 2019.

14.- Line 59: Cite the original WRF whitepapers as well (Skamarock et al.) instead of just the wind speed prediction part.

Response: Thank you very much for pointing this out. we have updated line 59 to the following text: "*… and the Weather Research and Forecasting Model (WRF) (Skamarock et al., 2021) are extensively utilized for wind speed prediction.*"

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Model Version 4, 2021.

15. - Line 116: Define DBN here.
Response: Thank you. "DBN" was defined in line 22 as "Deep Belief Network (DBN)."

16. - Line 140-141: WRF is not just developed by NCEP. The WRF website states it is a "collaborative partnership of the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (represented by the National Centers for Environmental Prediction (NCEP) and the Earth System Research Laboratory), the U.S. Air Force, the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA)."
Response: Thank you very much for pointing this out.
We've corrected line 140-141 to the following text: "*The WRF 4.2 model (Skamarock et al., 2021), developed by the National Center for Atmospheric Research (NCAR), represents a new generation of mesoscale numerical models with numerous applications in research forecasting.*"

17. - Line 145-146: WRF can use other input fields other than GFS. I suggest just stating that your run of WRF uses GFS as initial and lateral boundary conditions.
Response: Thank you very much for pointing this out. We have already mentioned in lines 156-157: "*The regular Global Forecast System (GFS) forecast field data serve as the initial field and lateral boundary conditions for the WRF model.*"
Therefore, we have updated line 145-146 to the following text: "*When forecasting meteorological elements, the WRF model normally uses the GFS data developed by NCEP.*"

18. - Overall, section 2.2.1 could be improved to be more relevant and shortened. The background of WRF is well stated in literature and the manuscript should focus on parts relevant to wind speed prediction. "Boilerplate" text about WRF (e.g., L166-167 about "WRFOUT") is not exactly relevant and could be shortened (authors already state previously in text that output frequency is 1-hour to line up with observational data).
Response: Many thanks for your suggestion.
We've updated section 2.2.1 to the following text:
"*The WRF 4.2 model (Skamarock et al., 2021), developed by the National Center for Atmospheric Research (NCAR), represents a new generation of mesoscale numerical models with numerous applications in research forecasting. When forecasting meteorological elements, the WRF model normally uses the GFS data developed by the National Centers for Environmental Prediction (NCEP). Using the WRF model in combination with daily GFS data resolution of 0.25° × 0.25°, the GFS data updates at 06:00 UTC and generates forecasting every 3 hours for a total duration of 90 hours.*

We selected the 24-h forecasting data from the WRF-resulted file after a spin-up time of 18 hours. The GFS data as the initial field and lateral boundary conditions for the WRF model. Surface static data, such as terrain, soil data, and vegetation coverage, are derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite with a resolution of 15 seconds (approximately 500 meters). Incorporating a two-layer grid nesting configuration, the forecast area is illustrated in Fig. 1. The WRF configuration process is detailed in Table 1. Given that the time scale of the meteorological station data in the study area is 1 hour, the forecast data time interval of the WRF model is also set to 1 hour. As a widely used numerical weather forecast model, the WRF model is suitable for weather studies from a few meters to several thousand kilometers. Therefore, this paper uses the WRF model to predict 10-meter wind speed as the input factor for the error correction model."

19. - In Section 2.2.4, first summarize the major difference (and purpose) of experiments 1 & 2. It is hard for the reader to see the importance of the two experiments when it is mixed together with the analysis.

Response: Thank you very much for pointing this out. The correlation between the WRF-predicted 10-m wind speed and the observed wind speed is the highest. The purpose of the second experimental path is using VMD-PCA algorithm to dig out the hidden wind speed characteristics of the 10-meter forecast wind speed, reduce the input of other meteorological factors such as WD_{10} and D_2, and further prove that the VMD-PCA algorithm is effective before correcting the WRF-predicted wind speed.

We have updated the first paragraph of section 2.2.4 to the following text:

"*This study used five machine learning algorithms to conduct ten experiments following two main paths. The first path involves increasing the meteorological variables possibly related to wind speed in the forecast field. The correlation between the WRF-predicted 10-m wind speed and the observed wind speed is the highest. The purpose of the second experimental path is using VMD-PCA algorithm to dig out the hidden wind speed characteristics of the 10-meter forecast wind speed, reduce the input of other meteorological factors such as WD_{10} and D_2, and further prove that the VMD-PCA algorithm is effective before correcting the WRF-predicted wind speed. The overarching goal is to achieve accurate correction of the forecast field wind speed. The flowchart of the artificial intelligence models used to correct the WRF predicted wind speed for the two main experimental paths is illustrated in Fig. 2 and comprises the following three steps:*"

20. - Line 234: "Missing and outlier values are removed from the dataset" - isn't this WRF model outputs, why would there be missing values?

Response: Thank you very much for your question. As mentioned in question 8, we cooperate with China Meteorological Public Service Center through the project (the second batch of service public bidding projects for EHV transmission companies in 2022 (2022-FW-2-ZB)).

The 3km observation data transmitted by China Meteorological Public Service Center was interrupted or incomplete sometimes. Therefore, we need to eliminate the corresponding time point of the observation data when matching the WRF-predicted and observation data.

21.- Section 3.1: Better to describe the RMSE, R, error metric values of different model configurations in a table for clarity. A table only shows in Section 3.3 in the form of Table 3 & 4 and it is unclear of the relationship of these and experiments 1 & 2. The flow could be much improved here.

Response: Many thanks for your suggestion. The analysis in Section 3.1 is mainly carried out from Fig. S1-5 in supplementary materials. The purposes for section 3.1 and section 3.3 are different. Section 3.1 mainly studies the comparison of the training set and validation set of 5 artificial intelligence models in experiment 1 in February 2022 to further determine whether there is overfitting. Section 3.3 mainly analyzes the comparison of various error indicators in the test sets of December 2021 and January 2022 of the two experiments. Of course, for clarity, we have added Table 3 to show the error indicators of the training set and validation set of the 10 AI models in two sets of experiments in February 2022.

Table 3. Table of evaluation indices of wind speed error trained and verified by 10 models in February 2022

Model	training set			validation set		
	R	RMSE（m/s）	FA	R	RMSE（m/s）	FA
VMD-PCA-lightGBM	0.96	0.33	0.99	0.88	0.53	0.94
VMD-PCA-XGBoost	0.96	0.31	1.00	0.87	0.54	0.94
VMD-PCA-RF	0.89	0.52	0.94	0.86	0.57	0.93
VMD-PCA-DBN	0.74	0.75	0.87	0.74	0.75	0.87
VMD-PCA-MLP	0.84	0.60	0.91	0.81	0.66	0.90
lightGBM	0.93	0.41	0.98	0.88	0.54	0.94
XGBoost	0.96	0.31	0.99	0.87	0.56	0.93
RF	0.89	0.52	0.94	0.86	0.57	0.93
DBN	0.76	0.73	0.88	0.76	0.73	0.88
MLP	0.85	0.59	0.92	0.83	0.62	0.91

21. - A lot of the feature labels could be better explained instead of being just listed in the text (e.g., in conclusion Line 542). What does pca0, IMF0 represent physically?

Response: Thank you very much for pointing this out.

Xu et al., 2021 states: VMD is adopted to obtain unknown but meaningful features hidden in the wind speed series predicted using the WRF model. The set of stationary sub-series contains more valid information than the previous non-stationary wind speed series when they are used as the inputs of the error correction model. PCA is a dimensional-reduction method that recombines the original variables into a new set of several independent variables and comprehensively reflects the information of the original variables.

In this study, the original variables are a set of sub-series of the wind speed that contain valid features and noise. PCA method is adopted to extract the pcax (x=0, 1, 2) and remove the illusive components.

As shown in Fig. 4, IMF0 physically represents the wind speed stationary series with a specific lowest center frequency after the original wind speed series has been processed by VMD.

pca0 physically represents the lowest frequency wind speed series after PCA treatment of all IMFk (k=0, 1, 2, …, 8) sub-series with reduced dimension.

We have updated the following in the text: "*Feature importance analysis revealed that the top eight contributing factors for correcting WRF forecasted wind speed include WRF forecast 10-meter wind speed (WS_{10}), latitude, longitude, altitude, pca0 (pca0 physically represents the lowest frequency wind speed series after PCA treatment of all IMFk (k=0, 1, 2, …, 8) sub-series with reduced dimension), humidity, pressure, IMF0 (IMF0 physically represents the wind speed stationary series with a specific lowest center frequency after the original wind speed series has been processed by VMD).*"

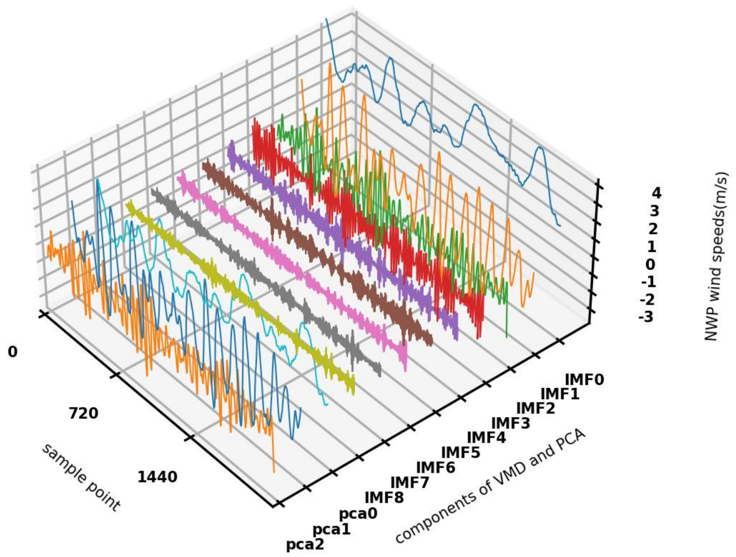

Figure 4: Three-dimensional view of 12 wind speed components after VMD and PCA processing of the 10-meter forecast wind speed at Lechang Station in Guangdong from December 1, 2021, to February 28, 2022.

Reviewer 3

General comments:

This paper is a description of several candidate post-processing approaches for producing point wind speed forecasts from numerical weather prediction simulations for sites in southern China. While the overall methods appear reasonable, and the conclusions appear valid, the paper needs some work to clarify the approach in some regards.

Response: Thank you very much for your recognition and encouragement of our work, we will further modify it according to your comments.

Specific comments:

1. I think more detail is needed on the gridded meteorological dataset that you describe creating in section 2.1 "Data". In particular, what is the source of the meteorological in situ observations? Are these wind towers all at a consistent height? How do you combine the surface observations with satellite data? Can you show some proof that your dataset "exhibits superior quality compared to other products", or at least provide some references that evaluate the dataset?

Response: Thank you very much for pointing this out. The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html).

These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation, and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites.

Yes! These actual observed wind speed data are obtained from the meteorological station location at a height of 10 m.

"Combining the surface observations with satellite data" and "exhibits superior quality compared to other products" are CLDAS-V2.0 official website documentation description results.

To clarify this and we have rewritten the Data part into the following text:

"*The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set. According to the description of the documents on the official website (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html), the dataset is constructed through the integration of multiple sources, including ground and satellite data, and is refined using advanced techniques such as multi-grid variational assimilation, physical inversion, and terrain correction. This dataset exhibits superior quality in comparison to other products, offering higher spatial and* *temporal resolutions. The target observation data includes 2-m air temperature, 2-m specific humidity, 10-meter wind speed, surface pressure, and precipitation. These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites. We select the 10-meter wind speed data of 410 sites, as illustrated in Fig. 1.*"

2. In the step 1 description (first part of Fig. 3, and text description in lines 222-242), it seems like a lot is being changed between Exp. 1 and Exp. 2. How can you control for this? It makes it somewhat hard to interpret the results.

Response: Thank you very much for pointing this out. In terms of controlled experiments, the control of variables in our two experiments is not very strict. However, on the basis of experiment 1, experiment 2 eliminated four meteorological variables whose feature importance was less than 4 %, which were called U_{10}, V_{10}, D_2, and WD_{10}. Most importantly, when we introduced the VMD-PCA algorithm, the feature importance of pca0 and IMF0 both exceeded 5 %. In other words, we retained the most important meteorological variable for correcting forecast wind speed in experiment 1, and introduced pca0 and IMF0 wind speed sub-series processed by VMD-PCA algorithm.

3. I was confused about why no bias statistics were shown in the verification section. Showing only mean absolute error and root mean square error type verification is only part of the story; can you say anything about the mean biases of the different approaches explored in this study? I think that is an important part of the analysis that is not shown yet.

Response: Thank you very much for your question. I have read some literature (Xiong et al., 2022; Zhang et al., 2019; Xu et al., 2021) about correcting wind speed forecasting, and I hardly saw the mean biases as a statistical indicator. According to our understanding, if at some time point the predicted wind speed is higher than the actual wind speed, then its bias is a positive value such as +x. If at other time point the predicted wind speed is lower than the actual wind speed, then its bias is negative such as -x. In this case, the calculated mean biases may be 0, which is not very suitable for evaluating the forecast of wind speed.

Xiong, X., Guo, X., Zeng, P., Zou, R., and Wang, X.: A Short-Term Wind Power Forecast Method via XGBoost Hyper-Parameters Optimization, Front. Energy Res., 10, 905155, https://doi.org/10.3389/fenrg.2022.905155, 2022.

Xu, W., Liu, P., Cheng, L., Zhou, Y., Xia, Q., Gong, Y., and Liu, Y.: Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy, Renewable Energy, 163, 772–782, https://doi.org/10.1016/j.renene.2020.09.032, 2021.

Zhang, Y., Chen, B., Pan, G., and Zhao, Y.: A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting, Energy Conversion and Management, 195, 180–197, https://doi.org/10.1016/j.enconman.2019.05.005, 2019.

4. I think somewhere (maybe in Fig. 1 and/or Fig. 2) you need to label the provinces, as readers from outside of China may not know which is which.

Response: Many thanks for your suggestion.

We have labeled the provinces in Fig. 1.

5. Minor comments:

Page 1, line 24: GFS stands for Global Forecast System.

Response: Thank you very much for pointing this out. We have corrected line 24 to the following text: "*We first construct WRF-predicted wind speeds using the Global Forecast System (GFS) model output based on prediction results*."

6. Lines 29: indexes > indices.

Response: Thank you very much for pointing this out. We have corrected line 29 to the following text: "*We find that the VMD-PCA-RF evaluation indices exhibit relative stability over nearly a year*:"

7. Page 3, line 83: training the > training on the

Response: Thank you very much for pointing this out. We have corrected line 83 to the following text: "*The error correction model improves the accuracy of the NWP model by training on the relationship between the NWP predictor variables and the observed correlation variables.*"

8. Page 5, line 118: Can you specify that these provinces are in China?

Response: Thank you very much for pointing this out. We have corrected line 118 to the following text: "*We analyze six distinct wind speed error indicators to compare and identify the most suitable wind speed error correction schemes for five southern provinces (Yunnan, Guizhou, Guangxi, Guangdong, Hainan) in winter and throughout most of the year.*"

8. Page 6, line 141: NCEP does not develop WRF, but rather NCAR (National Center for Atmospheric Research). While there are contributors to WRF from NCEP, there are also contributors from universities and many other organisations.

Response: Thank you very much for pointing this out. We've corrected line 140-141 to the following text: "*The WRF 4.2 model, developed by the National Center for Atmospheric Research (NCAR), ...*"

9. Page 6, lines 144-147: This section about GFS is confusing. Are you saying WRF uses GFS initial and lateral boundary conditions? It has the capability, but is not

required to use GFS data. Also, NCAR did not have a role in developing GFS to my knowledge.

Response: Thank you very much for pointing this out.

We've updated section 2.2.1 to the following text:

"*The WRF 4.2 model (Skamarock et al., 2021), developed by the National Center for Atmospheric Research (NCAR), represents a new generation of mesoscale numerical models with numerous applications in research forecasting. When forecasting meteorological elements, the WRF model normally uses the GFS data developed by the National Centers for Environmental Prediction (NCEP). Using the WRF model in combination with daily GFS data resolution of 0.25° × 0.25°, the GFS data updates at 06:00 UTC and generates forecasting every 3 hours for a total duration of 90 hours. We selected the 24-h forecasting data from the WRF-resulted file after a spin-up time of 18 hours. The GFS data as the initial field and lateral boundary conditions for the WRF model. Surface static data, such as terrain, soil data, and vegetation coverage, are derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite with a resolution of 15 seconds (approximately 500 meters). Incorporating a two-layer grid nesting configuration, the forecast area is illustrated in Fig. 1. The WRF configuration process is detailed in Table 1. Given that the time scale of the meteorological station data in the study area is 1 hour, the forecast data time interval of the WRF model is also set to 1 hour. As a widely used numerical weather forecast model, the WRF model is suitable for weather studies from a few meters to several thousand kilometers. Therefore, this paper uses the WRF model to predict 10-meter wind speed as the input factor for the error correction model.*"

10. Page 7, line 166: surface process plan > land surface model.

Response: Thank you very much for pointing this out. We have deleted this content according to Reviewer 2's opinion.

11. Page 9, line 206: Can you define and capitalize your acronym "pcs".

Response: Thank you very much for pointing this out. It means principal components (PCs). We've corrected line 206 to the following text: "*When principal components (PCs) are used as the input of the error prediction algorithm, the PCs fully reflect the characteristics of the subsequence and reduce the model complexity.*"

12. Fig. 3: validing > validating.

Response: Thank you very much for pointing this out. The revised figure is shown in Fig. 2.

13. Page 10, line 228: selected WRF field forecast data, including > selected WRF field forecast data to include only…

Response: Thank you very much for pointing this out. We've corrected line 228 to the following text: "*Experiment 2, derives 8 sets of data by reducing the selected WRF field forecast data to include only altitude, 10-meter wind speed, latitude, longitude,* *surface pressure, relative humidity, 2-meter temperature, and hourly precipitation.*"

14. Page 10, line 235: 8+9+3 does not equal 12. Are you counting the 9 IMF components as one set of meteorological elements? Please clarify your wording here.

Response: Thank you very much for pointing this out.

We have corrected line 235 to the following text: "*Experiment 1 (Experiment 2) standardize 12 sets of meteorological elements (8 sets of meteorological elements in Fig. 3, 9 IMF components, and three PCA vectors in Fig. 4) and wind speed observation data, respectively.*"

15. Page 13, line 276: Where does the "FA" acronym come from? I normally interpret that as false alarm, but it seems you have a different definition.

Response: Thank you very much for pointing this out. It means Forecasting Accuracy (FA), which has been used in past literature (Sun et al., 2019).

Sun, Q., Jiao, R., Xia, J., Yan, Z., Li, H., Sun, J., Wang, L., and Liang, Z.: Adjusting Wind Speed Prediction of Numerical Weather Forecast Model Based on Machine Learning Methods. Meteorological Monthly, 45(3): 426-436. https://doi.org/10.7519/j.issn.1000-0526.2019.03.012, 2019.

16. Line 278: index > indices.

Response: Thank you very much for pointing this out. We have corrected line 278 to the following text: "*The formula for calculating the error indices is as follows:*"

17. Lines 303-308: Can you put these verification results in a Table? That would make it much easier to read, and to compare the different approaches. The same goes for further lists of results in other sections.

Response: Thank you very much for pointing this out.

Sure! The testing set results in lines 303-308 are shown in Tables 3 and 4 of the original paper. As for the results of training and verification, as mentioned by Reviewer 2, we put the results in Table 3.

Of course, for clarity, we have added Table 3 to show the error indices of the training set and validation set of the 10 AI models in two sets of experiments in February 2022.

18. Line 309: Indexes > indices.

Response: Many thanks for your suggestion. We've corrected Line 309 to the following text: "*Considering different evaluation indices,*"

19. Lines 309-311: These sentences don't make much sense. It would be better to say "in" instead of "is that". For example FA in January 2022 is generally higher

than in December 2021.

Response: Thank you very much for pointing this out.

We've corrected Line 309-311 to the following text: "*Considering different evaluation indices, the revision effects of the five models in two months demonstrate that RMSE in January 2022 is generally lower than in December 2021; FA in January 2022 is generally higher than in December 2021; R in January 2022 is generally lower than in December 2021.*"

20. Lines 356-359: Can you point out which figure this text refers to? I see some panels in Figs. 7 and 8 have blue and red scatter plots. Which model(s) are you specifically referring to about the day vs. night issues?

Response: Thank you very much for your question. It refers to *Figs. S6-8d, f and Figs. 6-7d, f.*

We've corrected Line 309 to the following text: "

As is shown in Figs. S6-8d, f and Figs. 6-7d, f, the red scatters represent the nighttime wind speed, which is more concentrated on the 1:1 line. In contrast, the blue scatters represent the afternoon wind speed, which is slightly away from the 1:1 line. This suggests that the correction effect of the five models (VMD-PCA-lightGBM, VMD-PCA-XGBoost, VMD-PCA-RF, VMD-PCA-DBN, and VMD-PCA-MLP) exhibits a noticeable diurnal variation."

21. Figs. 7 and 8: Please clarify in the caption that the scatter plots are by hour. Is there some pattern to the models and months that are being shown in each panel? If so, it is above my head. Also, what is the difference between Fig. 7 and Fig. 8?

Response: Thank you very much for pointing this out.

For clarity, we have refined the headings of Fig. 7 and Fig. 8. The difference between Fig. 7 and Fig. 8 is the result of two different models, VMD-PCA-RF, VMD-PCA-lightGBM, respectively.

For example, we have corrected the title of Fig. 7 to the following text: "*Figure 7: The scatter density map compared with the actual 10-meter wind speed: (a) 10-fold cross-validation training set of VMD-PCA-RF model in February 2022, (b) 10-fold cross-validation validation set of VMD-PCA-RF model in February 2022.*

The 24-hour scatter map compared with the actual 10-meter wind speed: (c) WRF forecasts in December 2021, (d) VMD-PCA-RF model forecasts in December 2021, (e) WRF forecasts in January 2022, and (f) VMD-PCA-RF model forecasts in January 2022."

22. Fig. 11: Can you clarify in the caption which panels show FA and which show RMSE? It is not clear.

Response: Thank you very much for pointing this out.

For clarity, we have corrected the title of Fig. 11 to the following text: "*Figure 11: FA ((a), (b), (c), and (d)) and RMSE ((e), (f), (g), and (h)) distribution maps of VMD-PCA-RF, VMD-PCA-lightGBM and WRF models on 410 sites in five southern* *provinces ((a), (c), (e), and (g) represent December 2021; (b), (d), (f), and (h) represent January 2022).*"

23. Line 477: I think it would be clearer to say "elevation above sea level" rather than "height". When I read "height" in this sort of study, it makes me think of anemometer height above ground level.

Response: Thank you very much for pointing this out.

For clarity, we have corrected Line 477 to the following text: "*To further understand the feature importance ranking of the RF models, we divided the model prediction results and actual wind speeds of the 410 stations into 20 equal parts according to terrain height above sea level (Fig. 11).*"

23. Line 494-495: This sentence is poorly worded and doesn't make sense.

Response: Thank you very much for pointing this out.

For clarity, we have corrected Line 494-495 to the following text: "*With 1 km as the center, the measured 10-meter wind speed is more variable in areas where the station terrain height increases or decreases. However, the pink box of the 10-meter wind speed predicted by WRF becomes wider as the station terrain height decreases (Fig. 11). The distance between the gray box and the pink box is greater as the station terrain height decreases. It shows that the 10-meter wind speed predicted by WRF has less accuracy with the station terrain height decreases.*"

24. Lines 493-498: The use of the word "unstable" or "instability" in this section is confusing. I might say something more like "variability".

Response: Many thanks for your suggestion.

For clarity, we have corrected Lines 493-498 to the following text: "*With 1 km as the center, the measured 10-meter wind speed is more variable in areas where the station terrain height increases or decreases. However, the pink box of the 10-meter wind speed predicted by WRF becomes wider as the station terrain height decreases (Fig. 11). The distance between the gray box and the pink box is greater as the station terrain height decreases. It shows that the 10-meter wind speed predicted by WRF has less accuracy with the station terrain height decreases. The VMD-PCA-RF and VMD-PCA-lightGBM models significantly reduce the variability of the 10-meter wind speed predicted by WRF. When the height of the station increases or decreases at 1 km, the correction intensity tends to increase gradually. This further explains the higher importance of the height factor in the RF model training.*"

25. Fig. 14: The text claims this figure shows the actual wind speed in each month, but I cannot find that.

Response: Thank you very much for pointing this out.

For clarity, we have corrected the title of Fig. 14 to the following text: "*Figure 13: Evaluation histograms of 10-meter wind speed predicted by 10 models in different months in Experiment 1 and Experiment 2 ((a), (b), (c), (d), (e), and (f) represent R, FA (%), MAE (m/s), RMSE (m/s), rMAE (%), and rRMSE (%) respectively).*"

26. Line 513: Indexes > indices

Response: Many thanks for your suggestion.

We've corrected Line 513 to the following text: "*As shown in the Fig. 13, the evaluation indices of the model trained in Experiment 2*"

Reviewer 4

General comments

The manuscript “A robust error correction method for numerical weather prediction wind speed based on Bayesian optimization, Variational Mode Decomposition, Principal Component Analysis, and Random Forest: VMD-PCA-RF (version 1.0.0)” by Zhou et al. introduces a hybrid method for correcting 10-meter wind speed predicted by WRF. The authors compare the performance of two sets of experiments with different predictors and report the best model for wind speed correction during December 2021 to January 2022. In general, this manuscript fits the scope of the Geoscientific Model Development. However, after reading the manuscript, I find it still has a few major flaws. Firstly, the descriptions for the observation data and methods are unclear and ambiguous, and some citations should be implemented in the main text. Secondly, the information in the main text, figures, and tables is repeated. For example, the authors just simply report many statistics for model validation and comparison in Section 3, which are also showed in the tables. I would suggest the authors to summarize the key points and analyze the potential reasons for the differences in the main text rather than listing the statistics, which can be better for readers’ understanding. Finally, the writing and figures should be improved. Some figures should be combined, e.g., Figure 1 and 2. The captions for some figures are very simple, e.g., Figure 5, Figure6, and Figure 10. The labels and legends might be enlarged for a better readability. This reviewer requests major revisions listed below.
Response: Thank you very much for your recognition and encouragement of our work, we will further modify it according to your comments.

Specific comments

1. P5, Section 2.1: The description of the observation data is unclear. I would suggest the authors to give more details on this dataset. What are the data sources for the ground and satellite data? How do the authors process the data? How do the authors interpolate the data across 410 sites? Please cite the data sources and related techniques.
Response: Thank you very much for pointing this out.
The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html).
These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation, and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites.

"Combining the surface observations with satellite data" and "exhibits superior quality compared to other products" are CLDAS-V2.0 official website documentation description results.

To clarify this and we have rewritten the Data part into the following text:

"*The observed data comes from the China Meteorological Administration land data assimilation system (CLDAS-V2.0) real-time product data set. According to the description of the documents on the official website (https://data.cma.cn/data/cdcdetail/dataCode/NAFP_CLDAS2.0_RT.html), the dataset is constructed through the integration of multiple sources, including ground and satellite data, and is refined using advanced techniques such as multi-grid variational assimilation, physical inversion, and terrain correction. This dataset exhibits superior quality in comparison to other products, offering higher spatial and temporal resolutions. The target observation data includes 2-m air temperature, 2-m specific humidity, 10-meter wind speed, surface pressure, and precipitation. These data are processed by the China Meteorological Public Service Center to equivalent latitude and longitude grid scale, covering a geographical range of 15-32.97°N and 94-120.97°E. The spatial resolution of the grid is 0.03° × 0.03° (3km by 3km) and the temporal resolution is 1 hour. China Meteorological Public Service Center applied the nearest neighbor interpolation for precipitation and bilinear interpolation for the other four meteorological elements with downscaling from 3km to 410 sites. We select the 10-meter wind speed data of 410 sites, as illustrated in Fig. 1.*"

2. P6, Line 155: Do the authors consider the spin-up time for WRF simulations?

Response: Thank you very much for pointing this out.

Using the WRF model in combination with daily GFS data resolution of 0.25° × 0.25°, the GFS data updates at 06:00 UTC and generates forecasting every 3 hours for a total duration of 90 hours. We selected the 24-h forecasting data from the WRF-resulted file after a spin-up time of 18 hours.

We've updated section 2.2.1 to the following text:

"*The WRF 4.2 model (Skamarock et al., 2021), developed by the National Center for Atmospheric Research (NCAR), represents a new generation of mesoscale numerical models with numerous applications in research forecasting. When forecasting meteorological elements, the WRF model normally uses the GFS data developed by the National Centers for Environmental Prediction (NCEP). Using the WRF model in combination with daily GFS data resolution of 0.25° × 0.25°, the GFS data updates at 06:00 UTC and generates forecasting every 3 hours for a total duration of 90 hours. We selected the 24-h forecasting data from the WRF-resulted file after a spin-up time of 18 hours. The GFS data as the initial field and lateral boundary conditions for the WRF model. Surface static data, such as terrain, soil data, and vegetation coverage, are derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite with a resolution of 15 seconds (approximately 500 meters). Incorporating a two-layer grid nesting configuration, the forecast area is illustrated in Fig. 1. The WRF configuration process is detailed in Table 1. Given that the time scale of the* *meteorological station data in the study area is 1 hour, the forecast data time interval of the WRF model is also set to 1 hour. As a widely used numerical weather forecast model, the WRF model is suitable for weather studies from a few meters to several thousand kilometers. Therefore, this paper uses the WRF model to predict 10-meter wind speed as the input factor for the error correction model.*"

3. P6, Line 162-166: Please add the citations for these WRF parameterizations and schemes.

Response: Thank you very much for pointing this out. We have deleted Line 162-166 in section 2.2.1, and we have added the citations in Table 1.

Table 1: WRF configuration scheme

Model (Version)	WRF (V4.2)	
Domains	D1	D2
Horizontal grid points	600*500	967*535
Δx (km)	9	3
Vertical layers	58	
Longwave radiation	RRTMG (Iacono et al., 2008)	
Shortwave radiation	RRTMG (Iacono et al., 2008)	
Land surface	Noah LSM (Chen et al., 1997)	
Surface layer	MYJ (Janjić, 1994)	
Microphysics	Thompson (Thompson et al., 2008)	
Boundary layer	MYJ (Janjić, 1994)	
Cumulus	Tiedtke (Tiedtke, 1989; Zhang et al., 2011)	

Chen, F., Janjić, Z., and Mitchell, K.: Impact of Atmospheric Surface-layer Parameterizations in the new Land-surface Scheme of the NCEP Mesoscale Eta Model, Boundary-Layer Meteorology, 85, 391 – 421, https://doi.org/10.1023/A:1000531001463, 1997.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res., 113, D13103, https://doi.org/10.1029/2008JD009944, 2008.

Janjić, Z. I.: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes, Monthly Weather Review, 122, 927 – 945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2, 1994.

Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, Monthly Weather Review, 136, 5095–5115, https://doi.org/10.1175/2008MWR2387.1, 2008.

Tiedtke, M.: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in

Large-Scale Models, Monthly Weather Review, 117, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2, 1989.
Zhang, C., Wang, Y., and Hamilton, K.: Improved Representation of Boundary Layer Clouds over the Southeast Pacific in ARW-WRF Using a Modified Tiedtke Cumulus Parameterization Scheme*, Monthly Weather Review, 139, 3489–3513, https://doi.org/10.1175/MWR-D-10-05091.1, 2011.

4. P5 and P7: Figure 1 and Figure 2 both show the terrain heights in the study region. What's the difference between the two figures? I would suggest the authors to combine the two figures.
Response: Many thanks for your suggestion. We have combined the two figures. The revised figure is shown in Fig. 1.

5. P10, Line 234: What's the criteria for the outliers?
Response: Thank you very much for your question. We cooperate with China Meteorological Public Service Center through the project (the second batch of service public bidding projects for EHV transmission companies in 2022 (2022-FW-2-ZB)).
The 3-km and 410-sites observation data transmitted by China Meteorological Public Service Center was interrupted or incomplete sometimes. Therefore, we need to eliminate the corresponding time point of the observation data when matching the WRF-predicted and observation data.
We have corrected line 234 to the following text: "*The time points in the dataset where missing values are located are eliminated.*"

6. P10, Line 235: There are only 7 meteorological elements in Figure 4. Please add the missing one in Figure 7.
Response: Thank you very much for pointing this out.
The eight meteorological elements are altitude (HGT), 10-meter wind speed (WS_{10}), latitude (LAT), longitude (LON), surface pressure (PRS), relative humidity (RH), 2-meter temperature (T_2), and hourly precipitation (PRE). But we have not drawn the longitude, latitude and altitude, because the longitude, latitude and altitude are already shown in Fig. 1.
We have corrected line 235 to the following text: "*Experiment 1 (Experiment 2) standardize 12 sets of meteorological elements (8 sets of meteorological elements in Fig. 3, 9 IMF components, and three PCA vectors in Fig. 4) and wind speed observation data, respectively.*"

7. P10, Line 245: Why do the authors only use the data in February as training and validation dataset? I think there may be some seasonal variability for meteorological fields in the three months. I'm wondering if these machine learning models can successfully capture the relationship between the predictors and target variables in other two months.
Response: Thank you very much for your question.

We chose one of those months at random as the training set and the validation set. We did not only test in December and January, we tested over nearly a year and received a robust result (Fig. 13). We find that the VMD-PCA-RF evaluation indexes exhibit relative stability over nearly a year: correlation coefficient (R) is above 0.6, accuracy rate (FA) is above 85 %, mean absolute error (MAE) is below 0.6 m s^{-1}, root mean square error (RMSE) is below 0.8 m s^{-1}, relative mean absolute error (rMAE) is below 60 %, and relative root mean square error (rRMSE) is below 75 %.
In general, the other two months can also capture the relationship between the predictor and the target variable.

8. P16, Figure 6: How does the feature importance calculate? What's the correlation coefficients represent? Please add more details in the main text or in the caption.
Response: Thank you very much for your question.
Feature importance is calculated by "RandomForestRegressor.Feature_importances_" function of scikit-learn python package. Feature importance has been calculated in the paper (Duan et al., 2021). Correlation coefficients represent the WS_{10} and input variables in two sets of experiments.
We have corrected title of Fig. 5 to the following text: "*Schematic diagram of correlation coefficients (represented the WS_{10} and input variables) and feature importance (calculated by the scikit-learn python package) for two sets of experiments. (a) and (c) represent experiment 1, and (b) and (d) represent experiment 2.*"

Duan, Z., Yang, Y., Zhou, S., Gao, Z., Zong, L., Fan, S., and Yin, J.: Estimating Gross Primary Productivity (GPP) over Rice–Wheat-Rotation Croplands by Using the Random Forest Model and Eddy Covariance Measurements: Upscaling and Comparison with the MODIS Product, 21, 2021.

8. P21, Line 382: I think there is no significant differences in statistics for most models except for the DBN and VMD-PCA-DBN based on the Taylor chart in Figure 9. The Taylor chart and Table 3 provide the same information. I would suggest the authors to remove the Taylor chart in Figure 9.
Response: Many thanks for your suggestion.
In addition to the correlation coefficient and root mean square error, there is also an index (standard deviation) of its own data in the Taylor chart. If the standard deviation of the 10-meter wind speed corrected by the model is closer to the standard deviation of the actual 10-meter observation data, the effect of the revision is more consistent with the actual distribution. Therefore, from the Taylor chart, we can see that although the correlation coefficients and mean square deviations of other models are relatively close except DBN and VMD-PCA-DBN models, their standard deviations are still different.

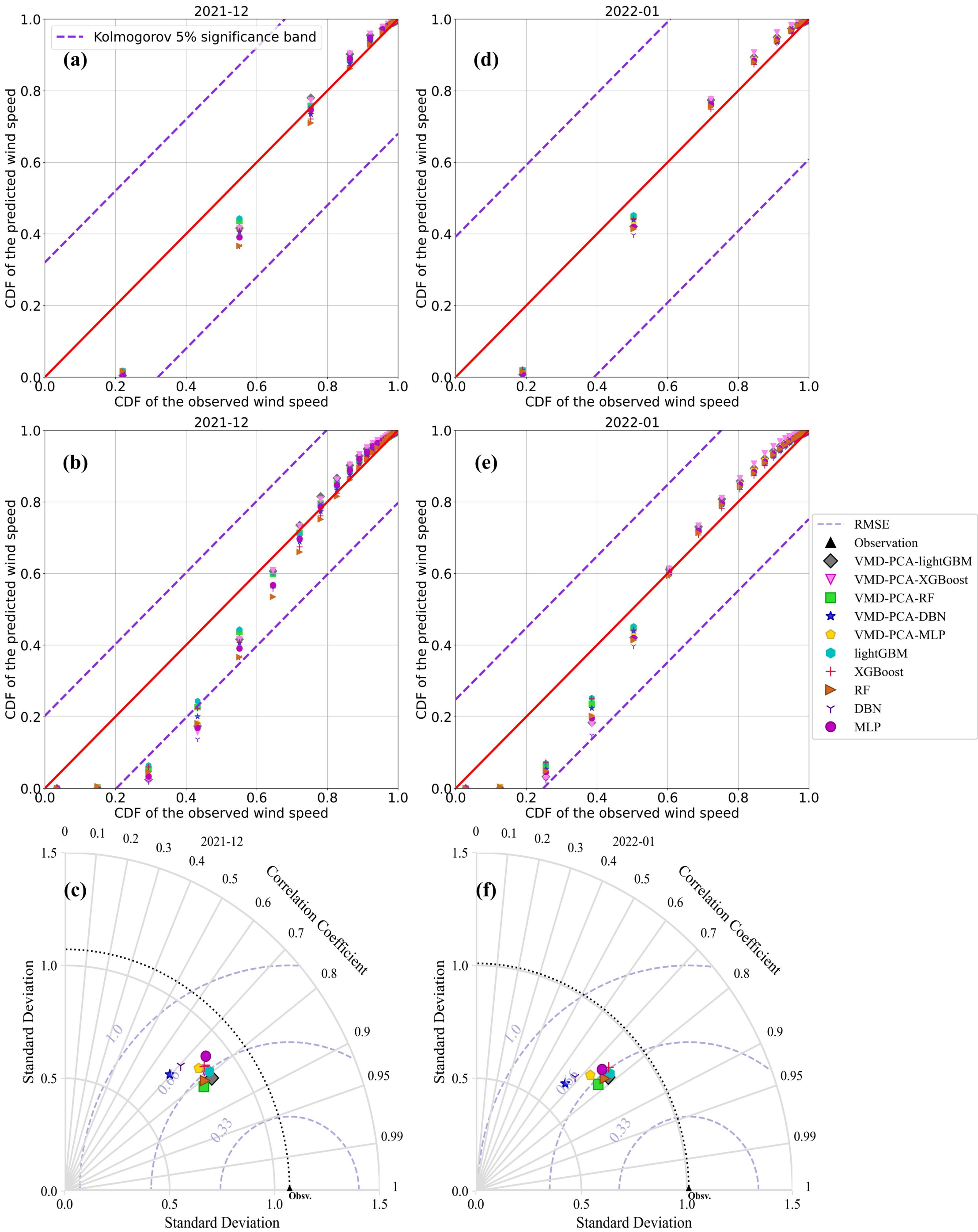

Figure 8: The cumulative distribution probability scatter plots of the actual wind speed and the predicted wind speed of 10 models in wind speed intervals of 0.5 m s^{-1} ((a) represents December 2021, (d) represents January 2022) and 0.2 m s^{-1} ((b) represents December 2021, (e) represents January 2022) respectively; Taylor distribution map ((c) represents December 2021, (f) represents January 2022).

9. P24, Figure 10: What do the shading areas and colored curves in Figure10c and 10d represent? Please clarify in the caption.

Response: Thank you very much for pointing this out. Colored curves are shown in the legend.

For clarity, we have corrected title of Fig. 9 to the following text: “Figure 9: VMD-PCA-lightGBM,VMD-PCA-RF and WRF daily variation of predicted and actual wind speeds in December 2021 and January 2022. (The shading areas represent an interval of 1 standard deviation, which is a 68% confidence interval.)”

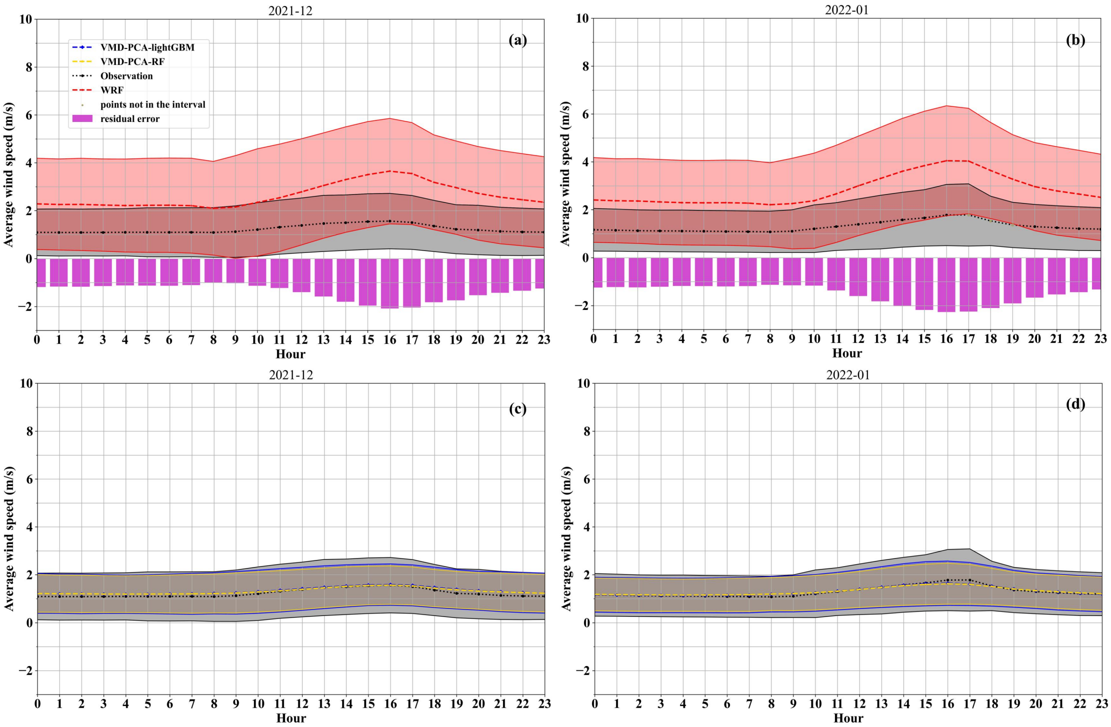

Figure 9: VMD-PCA-lightGBM,VMD-PCA-RF and WRF daily variation of predicted and actual wind speeds in December 2021 and January 2022. (The shading areas represent an interval of 1 standard deviation, which is a 68% confidence interval.)

10. Technical corrections

P6, Line 140: Please add citations for the WRF v4.2 model.

Response: Thank you very much for pointing this out. we have updated line 140 to the following text: *“The WRF 4.2 model (Skamarock et al., 2021), developed by the National Center for Atmospheric Research (NCAR), …”*

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Model Version 4, 2021.

11. P6, Line 140: It should be “National Centers for Environmental Prediction (NCEP)”.

Response: Thank you very much for pointing this out. We have updated line 140 to the following text: *“The WRF 4.2 model (Skamarock et al., 2021), developed by the*

National Center for Atmospheric Research (NCAR), represents a new generation of mesoscale numerical models with numerous applications in research forecasting. When forecasting meteorological elements, the WRF model normally uses the GFS data developed by the National Centers for Environmental Prediction (NCEP)."

12. P10, Line 224: Please rearrange the order of the figures and tables, which should be numbered in the order of their appearance in the main text.

Response: Thank you very much for pointing this out.

We have rearranged the order of the figures, and added the following in the text: *"In Experiment 1, as shown in Fig. 2, 12 sets of data are selected from the WRF forecast field, including altitude (HGT), 10-meter wind speed (WS_{10}), latitude (LAT), longitude (LON), surface pressure (PRS), relative humidity (RH), 10-meter meridional wind (V_{10}), 10-meter zonal wind (U_{10}), 2-meter temperature (T_2), 2-meter dew point temperature (D_2), 10-meter wind direction (WD_{10}), and hourly precipitation (PRE)."*

12. P13, Line 276: Please spell out the acronyms "FA".

Response: Thank you very much for pointing this out.

It means Forecasting Accuracy (FA), which has been used in past literature (Sun et al., 2019).

Sun, Q., Jiao, R., Xia, J., Yan, Z., Li, H., Sun, J., Wang, L., and Liang, Z.: Adjusting Wind Speed Prediction of Numerical Weather Forecast Model Based on Machine Learning Methods. Meteorological Monthly, 45(3): 426-436. https://doi.org/10.7519/j.issn.1000-0526.2019.03.012, 2019.

13. P14: Please change the "m/s" to "m s-1".

Response: Thank you very much for pointing this out.

We have changed the "m/s" to "m s^{-1}".