

Response to Reviewer #2's Comments

This study introduces a Phy-RF method to extend wind profiles beyond the surface layer, overcoming limitations of the traditional model based on the Monin–Obukhov similarity theory. By combining the power law model (PLM) with the random forest (RF) algorithm, the Phy-RF method addresses errors in the PLM above the surface layer attributed to the α setting. Comparing performance over China, the Phy-RF model outperforms PLM and RF, demonstrating better accuracy and stability. Temporally, it is not significantly affected by seasonal variations but shows limitations during specific time periods. Spatially, the model performs worse in highland areas due to the absence of consideration for terrain factors. After some minor revisions, I am in favor, that this paper gets published in ACP.

Response: We greatly appreciated the reviewer's comments on our manuscript, which greatly improve the quality of our manuscript. We have made 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. The text in general should be carefully checked during the English language copy-editing process.

Response: Thanks for pointing these issues out. We tried our best to correct spelling and grammatical errors in the revised manuscript.

2. The acronym Phy-RF confuses me a bit, because you state in line 11-12: "...we propose a novel method that combines the power law method (PLM) with the random forest (RF) algorithm to extend wind profiles beyond the surface layer, called the Phy-RF method." Why you do not use PLM-RF as acronym if it is based on PLM. RF-PHY acronym (Radio Frequency Physical Layer) is also used in the wireless communication sector, and you may want to separate the name of your new method better.

Response: Good point! To avoid misunderstandings, the proposed method of "Phy-RF" has been revised to "PLM-RF" throughout this whole revised manuscript, from the main text to figures.

3. In the introduction, I find the absence of a concise overview and reference of the Prandtl layer, which encompasses the initial tens of meters within the atmospheric boundary layer.

Response: Agreed! It is well known that Dr. Prandtl proposed the concept of Prandtl layer in 1904. In the thin layer near the solid wall, the influence of viscous force cannot be ignored, and this thin layer is called the Prandtl boundary layer. Therefore, we added a concise overview and reference of the Prandtl layer in the introduction, which is shown as follows:

“The Prandtl layer encompasses the initial tens of meters within the atmospheric boundary layer (Anderson, 2005).”

Anderson, J. D.: Ludwig Prandtl’s boundary layer. Physics today, 58(12), 42-48, <https://doi.org/10.1063/1.2169443>, 2005.

4. The Root Mean Squared Error (RMSE) quantifies the accuracy of a regression model in predicting the response variable's value in absolute terms, whereas R-Squared measures how effectively the predictor variables account for the variability in the response variable. I encourage the authors to also have a look and include R-Squared or Adjusted R-Squared metrics in their model evaluation. If the authors stick with RMSE, I think they must better justify their decision.

Response: According to your suggestion, the statistical parameters such as coefficient of determination (R^2), mean absolute error (MAE) and root mean squared error (RMSE) are used in the model evaluation. The modifications can be seen in Fig.7, Fig.8, Fig.13 and Fig. S6-S8.

5. Just a curiosity, now you focused on different land-cover types, can you make a statement about the performance of the Phy-RF model above water surfaces yet? The emphasis was on comparing the performance over China. Do you plan to investigate a more global performance estimate of the models in the future?

Response: Because the radiosonde stations are mainly located over land and the drifting route of sounding balloon varies sharply over time and space, we are unable to analyze the performance of the PLM-RF model on the water surface. In the future,

we plan to use global RS observation data to train and test the PLM-RF model, and evaluate its performance on a global scale.

According to your suggestion, we make a statement about the performance of the PLM-RF model above water surfaces in the last paragraph of the conclusions. The modifications in the revised manuscript are as follows:

“However, due to the limitations in data size and terrain factors, the performance of the PLM-RF model above water surfaces is uncertain. In the future, the global RS observation data will be used to train and test the PLM-RF model, and evaluate its performance on a global scale.”

6. Lines 25-26: “These findings have great implications for the weather, climate and renewable energy.” The noun of the sentence is missing. Maybe add in the end of the sentence the word “sector” or “research.”

Response: Good suggestion! We added the “sector” in the end of the sentence.

7. Lines 37-38: “..., in which can be assimilate into atmospheric models to produce global wind profile products.” This sounds a bit wrong. I suggest writing: “Satellite observations, such as those from Aeolus, can provide horizontal line-of-sight wind profile data that can be assimilated into atmospheric models to generate global wind profile products.”

Response: Amended as suggested.

8. Line 41: I suggest to better formulate the following part: “...ground-based observations like wind tower, wind profile radar, and wind profile lidar...” into “...ground-based wind measurements from towers, radar or lidar-based profilers...”.

Response: Amended as suggested.

9. Line 90: I recommend not to describe the color bar in the text here. Just refer to the Figure 1 and maybe specify the land cover types in the caption of this Figure, if needed.

Response: Good suggestion! We deleted the description about color bar in here.

10. Line 157: Here you talk about previous studies, but no references are given. Either add the references again or reformulate.

Response: We added the references.

11. Delete full stop in title of Section 4.1.

Response: Amended as suggested.

12. Wording in title of Section 4.2 wrong. I guess should be “Wind speed evaluation of the PhyRF model”.

Response: The title of Section 4.2 modified to “Wind speed evaluation of the PLM-RF model”.

13. Paragraph 346-349: First sentence misses something g in the end like e.g. “sector”: “...implications for the weather, climate and renewable energy sector.”.

Response: We added the “sector” in the end of the sentence.

14. Please also reformulate the second part. What is meant by “limitations of data time”? This is not clear to me.

Response: Amended as “However, due to the limitations in data size and terrain factors, the performance of the PLM-RF model above water surfaces is uncertain. In the future, the global RS observation data will be used to train and test the PLM-RF model, and evaluate its performance on a global scale.”

15. Figure 3: I suggest adding the units in the y-axis’s captions. The scatter points, for my impression overlap to strong here and this could lead to misinterpretations by the reader.

Response: Amended as suggested.

16. Figure 4: The figure does look too pixelated, please increase the resolution of this figure. In 4(c) correct “Wooldland” to “Woodland”. The typo (“wooldland”) is also in the caption of the figure.

Response: Amended as suggested.

17. Figure 5: Here please write in the caption the meaning of all variable acronyms.

Response: Amended as suggested.