Answers to Reviewer 3 comments

Juan Contreras and co-authors

Summary:

The paper examines whether ERA5 wind speeds from higher model levels can better represent wind conditions in mountainous terrain than surface level windspeeds. The main finding is that wind speeds from ERA5 at higher model levels are better suited to characterize wind conditions at exposed location in mountainous terrain, as they show significantly higher correlation to wind speed measurements (10 m and 100 m). The results provide interesting new insights, which are of high relevance for the application of reanalysis data for wind energy. They significantly contribute to the discussion on how to best apply reanalysis data in mountainous terrain. The paper is well written and generally well-structured and provides novelty. The experimental data and the model data choices seem sound. However, major revisions especially in the methodology are needed before it can be published.

Response: We sincerely thank Reviewer 3 for the careful evaluation of our manuscript and for the positive and encouraging comments. We are pleased that the reviewer finds the study relevant, novel, and of practical importance for improving the use of reanalysis data in mountainous terrain for wind energy applications. We also appreciate the recognition that the paper is generally well written, well structured, and based on sound experimental and model data choices.

We have carefully considered all of the reviewer's comments and have revised the methodological section to improve its clarity, rigor, and transparency. In the revised manuscript, we provide additional explanations, clarifications, and supporting analyses to address the concerns raised. We believe these revisions have strengthened the manuscript and improved the presentation of our approach and results.

Main concern:

The strongest scientific point of the paper is the identification of the higher model levels as being more representative for the wind climate on ridges and exposed locations. However, the current methodology weakens this finding by combining the analysis of the higher levels with a machine learning model for long term correlation (LTC)/as a measure correlate predict (MCP) method. The use of a Random Forest (RF) model makes the analysis less transparent and introduces disadvantages. While using a RF model is not inherently a bad choice for LTC/MCP it has several disadvantages for the study:

1. **A RF seems overly complex for a model with only one input feature. It essentially is a "step-wise" function fitted to the data. A linear model would be a lot more transparent here. Even if low wind speeds might be mapped below 0, low wind speeds are in general not relevant for wind energy applications. Alternatively, higher order polynomials could be used.**

   Response: We included a linear regression model with residuals in our analysis to gain insight into the improvements obtained by using optimal height levels and different methods (RF and linear regression). Details of this analysis are reported bellow in "Recommendations".

In addition, we better justified the use of RF in our study in Lines 227-230: "*In addition, a recent study found that RF most effectively mitigates the influence of interannual wind variations in long-term referencing compared with classical linear regression, and that RF outperformed conventional models, especially in complex mountainous terrain (Borowski et al., 2026).*".

Finally, we made explicit the use of RF with a single input feature and provided insights into the inclusion of additional suitable variables in future studies. This was added in Lines 517-521: *"It should be noted that as our main objective was to highlight the suitability of an optimal model level height to estimate wind speed, no other heights were included in the RF model. Although the application of RF with one input variable (i.e., ERA5 optimal level height) may be considered unconventional, further studies could consider additional closest level heights to the optimal level as input features to improve the wind speed estimates."*.

2. **RF models are not distribution-conserving. In general, the variance will be reduced by using a RF. As most machine learning models, RF will miss extremes and tend to bias predictions towards the mean. The reduced dispersion in the distribution is therefore likely to be at least in part due to the application of the RF model to the ERA5 data.**

   Response: We thank the reviewer for raising this important point. We agree that RF models are not distribution-conserving and that, in many applications, they may reduce variance and smooth extreme values by biasing predictions toward the mean. To assess whether this behaviour affected our results, we calculated the coefficient of variation (CV) of the RF and LR wind speed estimates and the raw ERA5 wind speeds. We included the results in Table 5, together with the following text in Lines 412–416: *"These biases suggested yet challenges in capturing the full variability of wind speeds, particularly at the distribution extremes, which is a well-known problem with the RF algorithm and other machine learning methods. In this regard, the LR method seems to better fit extreme values, as shown in Fig. 6. We confirmed a greater reduction in the dispersion of the distribution when using RF compared with LR (see Table 5)."*.

   Table 5. Coefficient of variation (CV) calculated for wind speed observations at 80 m, ERA5 wind speed at the optimal height, and wind speed estimates obtained using linear regression with residuals (LR) and random forest (RF) with ERA5 optimal height as input.

	CV - M1	CV - M2	CV - M3	CV - M4
Observed	0.552	0.598	0.653	0.914
ERA5 optimal height	0.599	0.616	0.598	0.616
LR (ERA5 optimal height)	0.551	0.577	0.613	0.799
RF (ERA5 optimal height)	0.520	0.535	0.586	0.827

3. **Except for the correlation the results are only shown for the combination of ERA5 data with a machine learning approach. This makes it difficult to judge, what the contribution of the ML-model to the quality of the results was.**

   Response: In the revised version, we added a comparison between the observations and the ERA5 wind speeds at the selected optimal model level used as input to the RF and LR models for each site. This comparison allows us to better assess the specific contribution of the RF and LR models to the final results. The figure and modified text are included below in "Recommendations".

4. **The discussion compares the results with several other studies using machine learning. However, some of these are different applications and, in some cases, the direct comparison of performance indicators does not make sense. E.g. [1] developed a model to estimate wind speeds at locations with no prior measurement. This is a significantly different use case from the LTC/MCP investigated here, for which a measurement on site is needed. Also, the last paragraph of the discussion does not seem to be well rooted in the presented results, as it discusses a completely new topic (climate change).**

   Response: We agree that some of the studies cited in the Discussion address different applications of machine learning and are not directly comparable to the LTC/MCP approach used in this study. Therefore, we revised this section to better frame the comparison with previous work and to avoid direct comparisons of performance indicators when the objectives, data availability, or application contexts differ. The revised discussion text is shown below in “Recommendations”.

Recommendation:

1. **Include a direct comparison of the distributions found in the ERA5 data at different model levels with the measurement data. This way wind speed distributions of ERA5 can be better evaluated without the possibly distorting effect of the RF.**

   Response: This comparison was included in Figure 7 in the revised version. The following text was included in Lines 394-399: *“To quantify the contribution of the RF model to wind speed predictions, a direct comparison between ERA5 wind speeds at the optimal level and the observations was included in Fig. 7. This comparison shows that ERA5 optimal height data generally exhibit wind speed distributions skewed toward lower values, with a marked underestimation of frequencies above 10 m s^{-1} at all sites in comparison to the estimates obtained using LR and RF.”*.
   In addition, due to the changes done in Figure 6 (before Figure 4), the comparison with ERA5 single level data as inputs is now reported in Table 4.

Table 4. Performance of wind speed estimates using ERA5 single level dataset (10 m and 100 m) and optimal ERA5 heights as inputs in LR and RF models.

Site	Input	Model	R^2 (-)	MAE (m s^{-1})	RMSE (m s^{-1})	PSS (-)
M1	ERA5 10 m + 100 m	LR	0.10	4.50	5.61	0.85
		RF	0.52	3.27	4.10	0.74
	ERA5 optimal height	LR	0.61	2.96	3.70	0.91
		RF	0.73	2.41	3.05	0.90
M2	ERA5 10 m + 100 m	LR	0.19	3.61	4.52	0.87
		RF	0.52	2.76	3.47	0.75
	ERA5 optimal height	LR	0.53	2.74	3.45	0.89
		RF	0.67	2.26	2.88	0.88
M3	ERA5 10 m + 100 m	LR	-0.28	5.18	6.52	0.79
		RF	0.24	3.94	5.04	0.62
	ERA5 optimal height	LR	0.57	2.97	3.76	0.85
		RF	0.68	2.51	3.25	0.89
M4	ERA5 10 m + 100 m	LR	0.11	3.67	4.82	0.64

	RF	0.47	2.57	3.71	0.73
ERA5 optimal height	LR	0.41	3.04	3.92	0.65
	RF	0.61	2.27	3.20	0.86

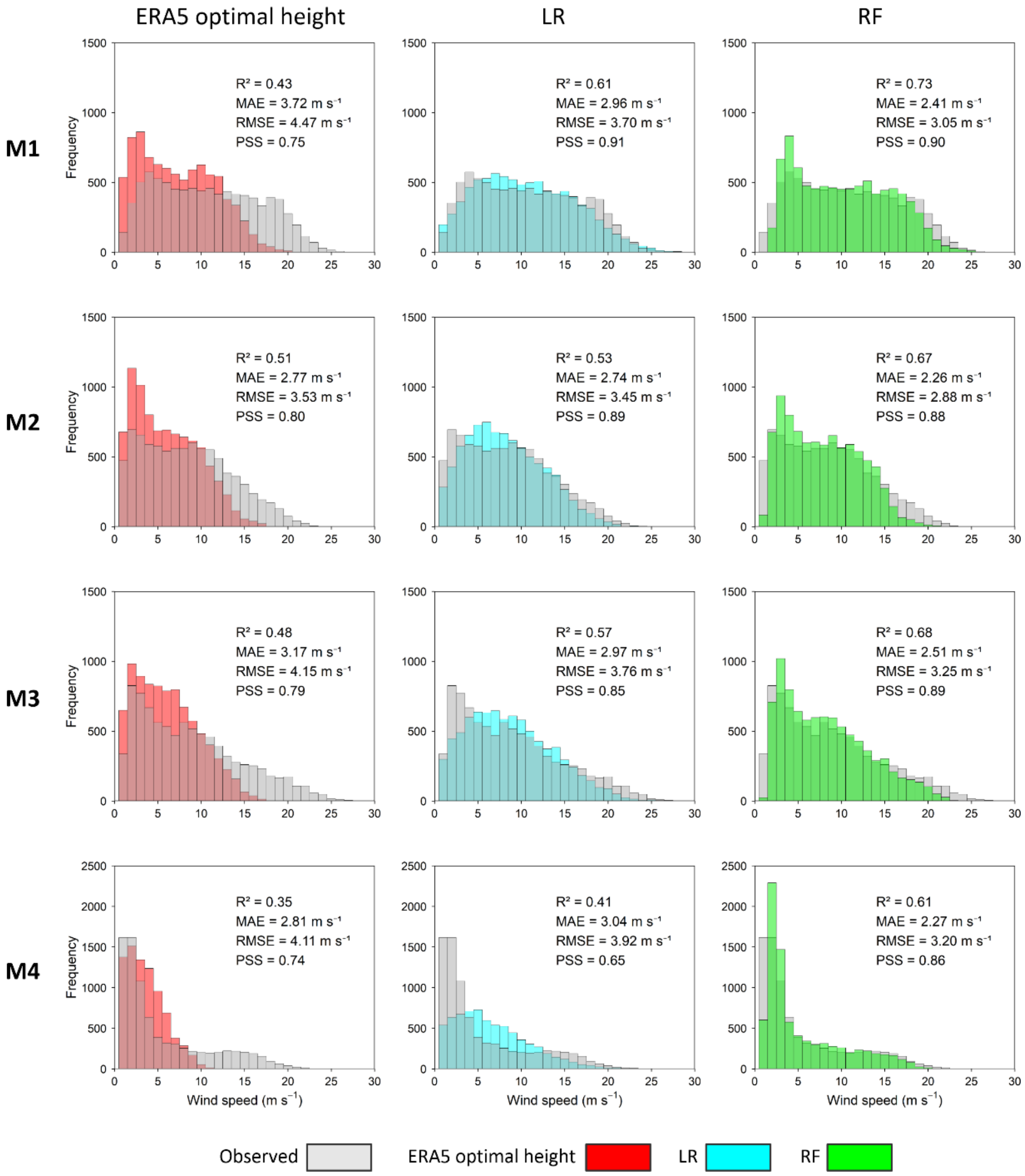

Figure 7: Comparison of wind speed frequency distributions between observed wind speeds, ERA5 wind speeds at the optimal levels, and wind speed estimates obtained using LR and RF with ERA5 optimal-level data as input across the four study sites for 2024. For each site, the figure also displays the performance metrics.

2. **The application of a simple linear model (in addition to the RF) would increase the transparency of the analysis and remove some of the black box effect the RF has.**

Response: This recommendation was addressed in previous comment.

3. **The application of widely used MCP methods like linear regression (with residuals) or variance ratio [2] would improve the comparability to other studies evaluating MCP approaches. This would also help to judge, if the performance of the MCP in high mountain ranges reaches similar levels as in flat terrain if the appropriate height level is used.**

   Response: This recommendation was addressed in the previous comment and is further discussed in the comment below.

4. **Revise the discussion and focus more on studies looking at LTC/MCP. It would be interesting to know how good estimates in mountainous terrain can get in comparison to simpler terrain given the appropriate height level is chosen. Remove the last paragraph from the discussion.**

   Response: The discussion section was revised and rewritten in Lines 488-506: "*A strictly fair comparison between our findings and previous studies is challenging because most studies used additional reanalysis-derived covariates, included surface-level wind speed observations as input variables, or reported different performance metrics. Borowski et al. (2026) assed the performance of different MCP methods using ERA5 reanalysis data in Europe and United States. They found better estimates with RF than liner regression model, especially in complex mountain terrain. Although these results may be influenced by the inclusion of additional atmospheric variables rather than only 100 m wind speed used in linear regression, our results confirm the better performance of RF than LR when considering the same input variables. Cavaiola et al. (2023) used ERA5 10 m wind speed, along with other ERA5 atmospheric variables, and Quantile RF (a method similar to RF) to calibrate wind energy estimates at mountainous sites in Italy. Their results showed biases of approximately −20% to 40% in the 20-year accumulated wind energy estimates across the study sites. Although our results are based on annual energy production rather than accumulated long-term production, we manage to provide accurate estimates using only the optimal height, with errors below 7% at all sites.*
   Compared with studies conducted in simple terrain, our wind speed estimates using RF and LR with ERA5 data at the optimal height still showed lower performance than those reported for flat terrain using ERA5 wind speeds at 100 m. For instance, Schwegmann et al. (2023) reported an R^2 of approximately 0.81 and 0.75 in north-eastern France when using RF and LR, respectively. These values are significantly higher than the R^2 values found in this study, which ranged from 0.61 to 0.73 for RF and from 0.41 to 0.61 for LR. Nevertheless, we emphasize that our results are obtained for a highly complex-terrain.".
   The last paragraph of the Discussion was removed from the revised manuscript.

Borowski, J., Schwegmann, S., Avila, K., & Dörenkämper, M. (2026). Evaluating the impact of inter-annual variability on long-term wind speed predictions. Wind Energy Science, 11(2), 661-677.

Cavaiola, M., Tuju, P. E., Ferrari, F., Casciaro, G., & Mazzino, A. (2023). Ensemble machine learning greatly improves ERA5 skills for wind energy applications. Energy and AI, 13, 100269.

Schwegmann, S., Faulhaber, J., Pfaffel, S., Yu, Z., Dörenkämper, M., Kersting, K., & Gottschall, J. (2023). Enabling Virtual Met Masts for wind energy applications through machine learning-methods. Energy and AI, 11, 100209.

Specific comments:

- **L16: Insert "the": "the ERA5 model level data set"**

  Response: The correction was included in the revised version.

- **L19: Write "to" instead of "for".**

  Response: The correction was included in the revised version.

- **L51: There are also regional reanalysis. Please rephrase.**

  Response: We rephrased the text in Lines 47–49 of the revised version to acknowledge the existence of regional reanalysis.

- **L53: ... (ERA5) is the 54 preferred reanalysis in the wind power meteorology community (Olauson, 2018)." This is somewhat misleading. The study evaluates the dataset but does not investigate its acceptance in the wind energy community. Please rephrase.**

  Response: The sentence was corrected in the revised version.

- **Section 2.3: I think it is not justified to dedicate a section to the fact that Pearson's coefficient of correlation was used to identify the height of the best correction. Recommendation: merge with Section 2.4.**

  Response: Both sections were merged in the revised version.

- **A RF with just one input feature is an unconventional choice (see comments above). please clearly state and justify this here. Also, the application of more transparent approaches is recommended here (see comment above).**

  Response: This was clarified in previous comments (see comments above).

- **Section 2.1: It would be great to have wind roses of the sites here. The main ridge of the mountain range of the Andes has a clear orientation. Therefore, this would help to understand the influencing topography better.**

  Response: Thanks to the reviewer for this suggestion. The wind roses for each site were included in Figure 2 in the revised version of the manuscript. A description of the main prevailing wind directions and their link with topography is provided in Lines 155-165: *"M1 and M3 are located on ridge tops, with no nearby obstacles (Fig. 1c), and show a clearly dominant wind direction from the ESE (Fig. 2). These sites also present the highest wind speed classes, particularly for winds from the ESE–SE sector, indicating well-exposed conditions to the prevailing flow. In contrast, M2 is located at the foothill of a ridge and is surrounded by higher hills (Fig. 1c), which likely influence the local wind direction and result in prevailing winds mainly from the ENE–NE sector (Fig. 2).*

Compared with M1 and M3, M2 shows a broader directional distribution and lower frequencies of the highest wind speed classes, suggesting a stronger influence of local topographic effects. Finally, M4 is located in the western foothills of the Andes, at a lower elevation than the other sites, and is situated on a hilltop surrounded by higher mountain ridges, resulting in a more sheltered site (Fig 1c). The wind rose for M4 shows the most heterogeneous wind distribution, with prevailing directions from ESE–SSE and WNW–NW, and a lower frequency of high wind speeds compared with the ridge-top sites (Fig. 2).".

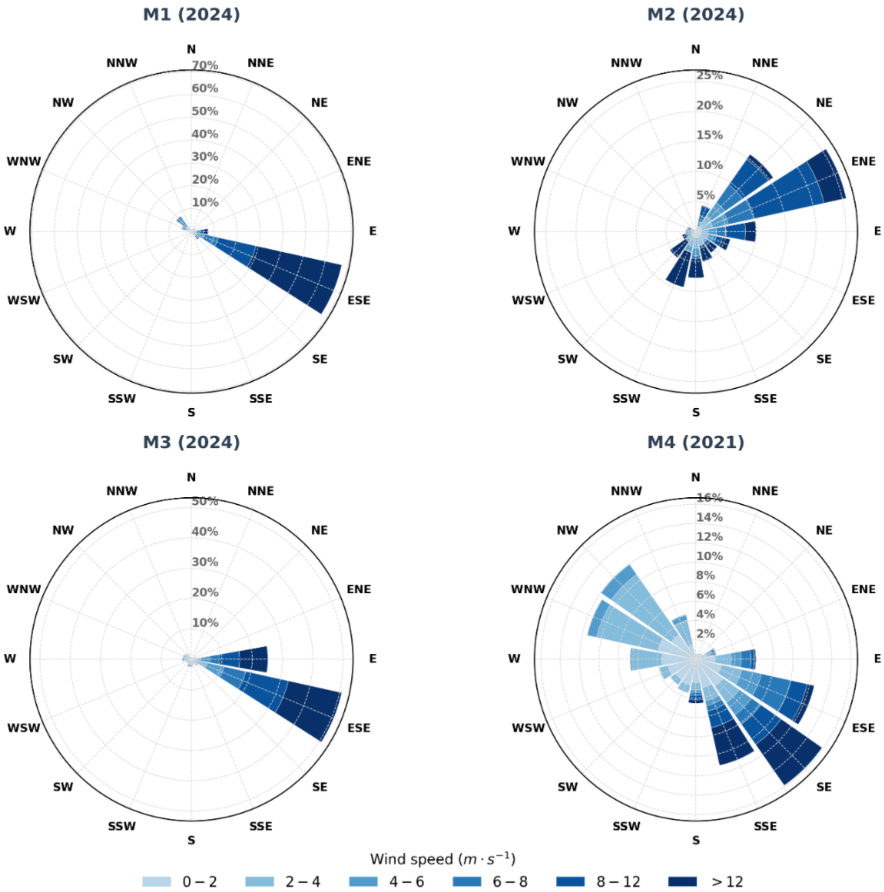

Figure 2: Wind roses showing the observed wind speed and wind direction distributions at the four mast sites. The plots correspond to 2024 for M1, M2, and M3, and to 2021 for M4 (selected based on wind direction data availability). Bar length represents the frequency of wind occurrence from each direction, while colours indicate wind speed classes in m s^{-1}. Note that the frequency range differs among plots to improve visualization.

- **97: Typo: "...has been stablished..."**

  Response: The correction was included in the revised version.

- **157: Grammar needs to be corrected, e.g. "This was done following MEASNET".**

  Response: The correction was included in the revised version.

- **188: Can you provide a justification why no interpolation was performed?**

  Response: A justification is provided in the revised version in Lines 211-212.

- **"For each site, both models were trained using observed hourly wind speed data from the first three years of the monitoring campaign (i.e., January 2021 to December 2023)." Usually, one year of data is used for LTC. This would also make results more comparable to other studies. Why did you deviate from this? Please change or justify.**

  Response: We agree that one year of on-site measurements is commonly used in LTC/MCP applications, particularly due to practical constraints in wind resource assessment. In this regard, we performed a sensitivity analysis of the specific years used to select the optimal model height, confirming that there were no substantial differences between using the 3-year period (January 2021–December 2023) and using a single year. Nevertheless, we choose to use *three years of hourly observations in order to reduce interannual variability and calibrate the models under a broader range of wind conditions.*

  This choice was particularly relevant for our study sites, which are located in complex mountainous terrain where local wind regimes can vary substantially from year to year. Using a longer training period allowed the transfer function to be calibrated with a broader range of wind speed conditions, including seasonal variability and less frequent high-wind events. In addition, this approach enabled us to reserve 2024 as a fully independent year for testing the models, providing a robust out-of-sample evaluation.

  We added the following text in the revised manuscript in Lines 246-249: *"Although one year of data is commonly used in MCP applications, three years of hourly observations were used to reduce interannual variability and calibrate the models under a broader range of wind conditions. The year 2024 was then reserved for independent model evaluation.".*

- **247: Grammar: "to quantify"**

  Response: The correction was included in the revised version.

- **L207: "In this study, the RF model was preferred over a conventional regression model because it provides more 207 consistent predictions (e.g., non-negative values), particularly at low wind speeds." Usually low wind speeds (so low that the linear regression would provide negative values) are not of interest for wind energy applications. I think it would still be good to include a widely used regression method like linear regression with residuals or variance ratio in the analysis as a reference or to make a better argument against their usage.**

  Response: We included an analysis using linear regression with residuals as a reference method (see comments above).

- **290: Check grammar.**
  L300: "This pronounced difference, compared to the other sites, may be attributed to the larger discrepancy between the actual site elevation and the ERA5 model elevation." Please state differences here.

Response: Grammar issues were corrected in the revised version in Lines 328-330: *"These results using indicate that ERA5 model level data indicates that wind speeds over flat terrain is are representative of hub-height conditions at coastal and inner flat sites, but are not representative in mountainous areas."*. The elevation differences were included in the revised version in Lines 340-342: *"This pronounced difference, compared to the other sites, may be attributed to the larger discrepancy between the actual site elevation and the ERA5 model elevation (e.g., ΔM1 = -755 m a.s.l., ΔM2 = -649 m a.s.l., ΔM3 = -1125 m a.s.l., ΔM4 = -132 m a.s.l.) see Fig. 1."*.

- **331: It is stated that M4 is located on a hilltop and should be suitable for wind energy purposes. Can these very low wind speeds for such a site be explained in more detail from the specific terrain characteristics. This is mentioned in line 305, but could be discussed more.**

  Response: This comment was clarified in previous comment related to wind direction and topographic characteristics of the site (see comments above).

- **L386: "The results presented above confirm that ERA5 consistently underestimates wind speed variability in the tropical Andes in line with other studies in complex terrain." The paper does not analyse the distribution of ERA5 directly but rather the output of the RF model. From the current analysis this conclusion cannot be drawn. Also, it would be nice to differentiate stronger between the characteristics of the different height levels.**

  Response: To improve the transparency of the method and allow a direct comparison with the ERA5 distribution, we redesigned Figure 7 by adding the original ERA5 distribution. We refer this part of the text with Figure 7 in the revised version.

- **386: Check grammar: "confirms"- "confirm"**

  Response: The correction was included in the revised version.

- **398: Spelling: "intricated"- indicated**

  Response: The correction was included in the revised version.

- **L408: "The lower performance in M4 is caused by the lower performance of the RF models in estimating particularly lower wind speeds (i.e., 0-2 m s-1) where high frequency values within this range are common in this site." I don't understand this sentence.**

  Response: The sentence was corrected to improve clarity in the revised version in Lines 483-487: *"The lower performance at M4 is mainly due to the difficulty of the RF model in estimating low wind speeds, particularly within the 0–3 m s^{-1} range. The models show a marked underestimation of frequencies between 0–1 m s^{-1} and an overestimation of wind speed values between 2–3 m s^{-1}. Since this site has a higher frequency of wind speeds within these ranges, the overall model performance is compromised in this case."*.

- **425: Typo "transformers"**

Response: The correction was included in the revised version.

- **436: Typo, correct to “not commonly assessed”**

  Response: The correction was included in the revised version.

References

[1] Wenxuan Hu, Yvonne Scholz, Madhura Yeligeti, Lueder Von Bremen, and Ying Deng. Downscaling ERA5 wind speed data: a machine learning approach considering topographic influences. Environmental Research Letters, 18(9):094007, September 2023.

[2] Anthony L. Rogers, John W. Rogers, and James F. Manwell. Comparison of the performance of four measure–correlate–predict algorithms. Journal of Wind Engineering and Industrial Aerodynamics, 93(3):243–264, 2005.