

Answers to Reviewers

Juan Contreras and co-authors

REVIEWER 1

General comments

This study investigates how well ERA5 model-level wind speeds at different geometric heights represent 80-m mast observations in complex terrain and coastal/flat sites. The work addresses an important practical question for wind-energy applications: at which ERA5 heights are hub-height conditions best represented, and how does this depend on terrain complexity.

Overall, the paper poses a clear scientific question. The scientific approach is sound and the results are clearly presented. I therefore recommend minor revision, mainly to clarify some key statements, correct small inconsistencies in figures/tables, and explicitly summarize the main uncertainties and limitations in the Conclusions. The manuscript addresses an important and practical question for wind-energy applications and provides useful guidance on the representativeness of ERA5 model-level winds in different terrain types.

Response: We sincerely thank the reviewer for this positive and encouraging assessment of our manuscript. We are very pleased that the reviewer recognizes the importance and practical relevance of our study for wind-energy applications, as well as the clarity of the scientific question, the soundness of the approach, and the clear presentation of the results. We also greatly appreciate the constructive suggestions provided, which helped us improve the manuscript.

Minor comments

- 1. In Figure 2, I suggest adding the site elevation in parentheses next to each mast name in the legend (e.g., “M1 (3796 m)”, “M2 (3690 m)”), for easier cross-reference with the map.**

In addition, it would be very helpful to explicitly include the correlations between the mast observations and the ERA5 10-m and 100-m single-level winds for each site (e.g., additional markers or lines). This would make the comparison between single-level and model-level performance more transparent to the reader.

Response: Both suggestions were included in Figure 4 (originally Figure 2) in the revised version.

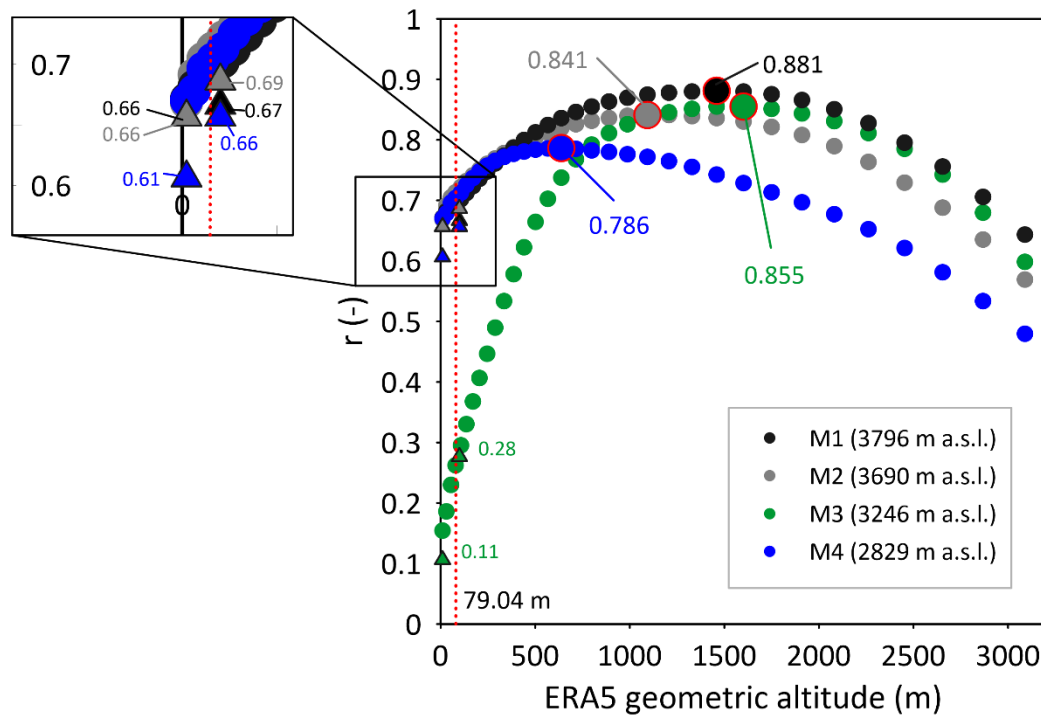


Figure 4. Correlation between observed wind speed at 80 m height at four sites and ERA5 data at different geometric altitudes for the period from January 2021 to December 2023. The 10 m and 100 m wind speeds from the ERA5 single level dataset are indicated by triangles, with their corresponding correlation magnitudes included. For each site, the highest correlation value is indicated by larger circles. The red dashed line indicates the model height closest to the observations. Site elevation is indicated in parenthesis in the legend.

2. In Figure A3, I only see three coastal sites, while the caption refers to “four Coast sites”. Please check and correct the caption and/or the figure.

Similar to Figure 2, I recommend also indicating the correlations between the mast observations and the ERA5 10-m and 100-m single-level winds for each coastal site, to facilitate comparison.

Response: There was a mistake in the caption. We considered only three coastal sites, as shown in the figure. The caption of Figure A2 was corrected in the revised version.

In addition, correlations between the mast observations and the ERA5 10-m and 100-m single-level winds for each coastal site were included in Figure A2 in the revised version to ensure consistency with Figure 2.

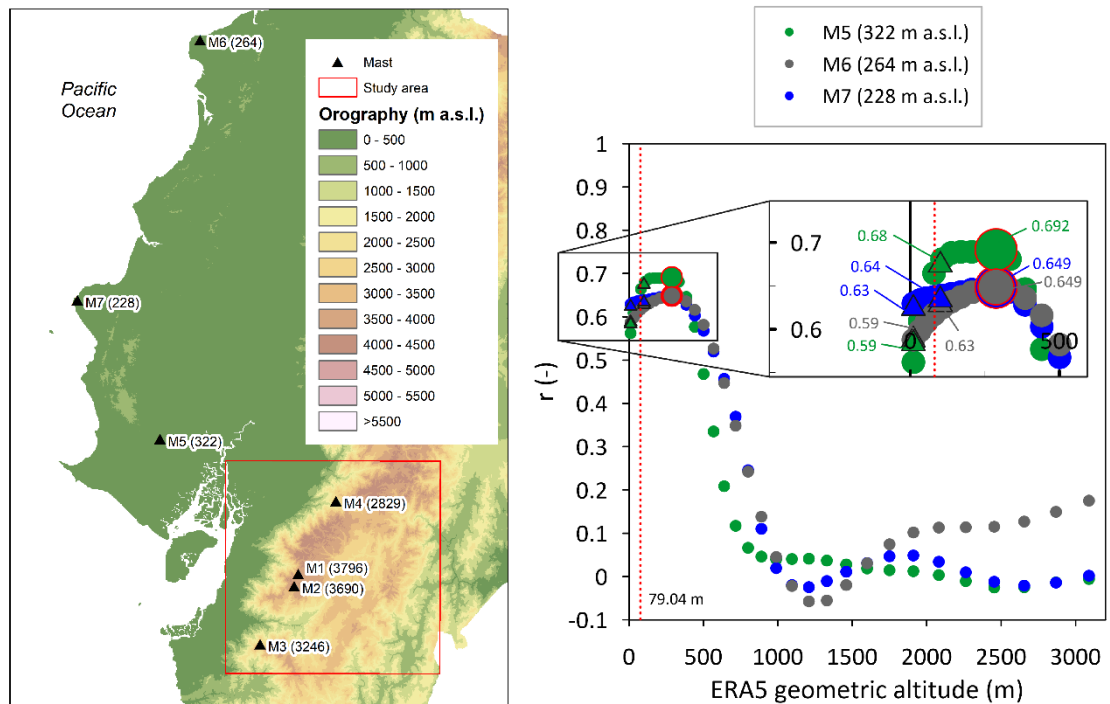


Figure A2: Correlation between observed wind speed at 80 m height at three Coast sites and ERA5 data at different geometric altitudes for the period from January 2021 to December 2022 (different analysis period compared with the mountain sites due to data availability constraints). On the left, the map shows the location of the sites, with the actual elevation of each site indicated in parentheses (m a.s.l.). On the right, the highest correlation value for each site is indicated by larger circles. The 10 m and 100 m wind speeds from the ERA5 single-level dataset are indicated by triangles, with their corresponding correlation magnitudes. The red dashed line indicates the model height closest to the observations. Site elevation is indicated in parenthesis in the legend.

3. Lines 303–305: I suggest explicitly emphasizing in the Conclusions that, in complex terrain, better correlations are obtained at well-exposed locations (e.g., ridge-top sites) than at sites strongly affected by local flow behaviour (e.g., sheltered hillslopes or valleys).

In addition, the important findings stated in Lines 444–445 (“higher underestimations would be expected for sites with high frequencies of low wind speed”) and in Lines 289–292 (“wind speeds over flat terrain are representative of hub-height conditions at coastal and inner flat sites but are not representative in mountainous areas”) could also be briefly restated in the Conclusions section so that the main messages are more clearly visible to the reader.

Response: All these findings were highlighted in the Conclusions section in the revised version in Lines 542–546: “This finding indicates that while near surface and hub-height wind speeds over flat terrain are representative of observed hub-height conditions at coastal and inner flat sites, they are not representative in mountainous areas. Particularly, better correlations between ERA5 and observed winds were obtained at well-exposed locations (e.g., ridge-top sites) than at sites strongly affected by local flow behaviour (e.g., sheltered hillslopes or valleys).”; and in Lines 551–554: “In the latter site, higher underestimations would be expected due high presence of low wind speeds and the challenges of the RF models to predict the frequency of extreme low values (0–1 m s⁻¹); however, these values showed no

major impact in AEP estimates as wind turbines considered here operate at higher wind speeds ($>2.4\text{--}4\text{ m s}^{-1}$). ”.

4. **In the caption of Table 4 it is stated that “The best-performing models ... are highlighted in bold”, but no boldface appears in the current version of the table. Please either apply the bold formatting or adjust the caption accordingly.**

Response: The caption was modified in the revised version by deleting the phrase “... are highlighted in bold” to avoid confusion.

5. **Lines 369–371: M1 shows the smallest discrepancies, whereas M2 exhibits the largest. It would be useful if the authors could briefly discuss the possible reasons for this difference (e.g., differences in exposure, local flow features, representativeness of the ERA5 grid point, or data quality), as this may help readers understand how site characteristics affect model performance.**

Response: Thanks to the reviewer for this helpful observation. We clarified this point by adding topographic contour lines to Figure 1c, which helps readers better interpret the results in relation to the site characteristics. The revised figure shows that M1 is located on a high plain, with no higher surrounding mountains, making it a well-exposed site to the prevailing easterly winds. In contrast, M2 is located on a mountain foothill and is closely surrounded by higher terrain ($\geq 4000\text{ m a.s.l.}$), which likely produces a more sheltered environment and promotes local flow effects.

The following text was included in Lines 441–447 of the revised version: *“This higher underestimation is mainly related to the underestimation of the frequency of wind speeds above 16 m s^{-1} compared with the other sites (Fig. 7). It should be noted that the M1 and M2 sites are located within the same ERA5 pixel. Thus, the larger error at M2 may be attributed to the limited representativeness of the ERA5 pixel in capturing the local flow conditions at this site. As shown in Fig. 1c, M1 appears to be more representative of the dominant topographic characteristics within the ERA5 pixel, whereas M2 is surrounded by higher mountains, which likely influence the local flow conditions.”.*

6. **In the Conclusions section, I recommend adding a short paragraph summarizing the main sources of uncertainty and limitations of the study. For example, this could include:**

- **the limited number of masts and regions considered;**
- **the representativeness of individual masts for complex terrain;**
- **the dependence on the chosen ERA5 grid point and the computation of geometric heights from model levels;**
- **and the relatively short analysis period (2021–2024).**

Such a paragraph would help frame how far the results can be generalized to other complex-terrain regions.

Response: Thanks for this valuable recommendation. The following paragraph describing the uncertainty and limitations of the study was included in the revised version in Lines 557–570: *“Despite these promising results, some limitations are acknowledged for interpreting*

and generalizing the findings. The analysis was based on a limited number of masts located in a specific region of the tropical Andes, and most sites correspond to well-exposed locations with relatively high wind potential. Therefore, additional observations from other complex-terrain regions, including more sheltered hillslope or valley sites, would be needed to assess the broader applicability of the results. Moreover, the representativeness of individual masts may vary according to their local exposure and surrounding terrain characteristics. Such results may also depend on the selected ERA5 grid point, particularly where the reanalysis topography differs from the real terrain. Alternatively, future studies could employ the weighted-pixel extraction approach proposed by Gualtieri (2021), which considers neighbouring pixels to improve the spatial representativeness of the sites. Finally, although the sensitivity analysis showed no major interannual changes in the selected optimal height levels, the relatively short analysis period considered here (2021–2024) may not fully capture longer-term climate variability or periods with anomalously high or low wind speeds. Extending the analysis to longer periods and additional sites would help to better evaluate the robustness and transferability of the proposed approach.”

Gualtieri, G. (2021). Reliability of ERA5 reanalysis data for wind resource assessment: A comparison against tall towers. *Energies*, 14(14), 4169.

REVIEWER 2

The manuscript sets out to answer a clear scientific question, can better hub-height wind speed predictions be made in complex terrain by using other ERA5 model heights than the standard 10 m and 100 m heights, and answers it in a clear and straightforward fashion. Given the simplicity of the question, though, I am not sure of the scientific significance of the work. It seems rather obvious that you can improve wind speed estimates by searching for and selecting the ERA5 height that correlates most with the quantity you want to predict. I would recommend considering the following additions to enhance the impact of the work.

Response: We thank the reviewer for this thoughtful and constructive general assessment. We appreciate the recognition that the manuscript addresses a clear scientific question and answers it in a clear and straightforward manner. We also appreciate the reviewer's suggestion to strengthen the impact of the work.

Regarding the scientific significance of the study, we agree that wind speed estimates are expected to improve when using predictors that are more strongly correlated with observations. However, the key point is that this relationship must be first demonstrated for ERA5 model-level winds in complex terrain. In particular, before our study, it was not known whether higher ERA5 model levels could provide a better representation of hub-height wind speeds than the commonly used 10 m and 100 m single-level data, nor which model level is most representative across sites with different terrain characteristics. To our knowledge, such an analysis has not been systematically carried out before using hub-height mast observations. In addition, applying the approach in the tropical Andes offers an interesting and extreme scenario (altitudes above 3000 m a.s.l.) for complex terrain. Therefore, the main contribution of our study is to provide empirical evidence of this relationship and to demonstrate the added value of ERA5 model-level data for wind resource assessment in complex mountainous terrain. This motivation was included in Lines 108-111 and 114-116 in the revised manuscript.

First, I would be interested to see typical model errors expressed for different hub-height wind speeds and directions. Do patterns emerge? For instance, does the error change dramatically when the wind direction is aligned with the ridge as opposed to cases where it is perpendicular to the mountain ridges? This would help to establish the topography as the main reason why upper elevation ERA5 wind speeds correlate so well with the wind speed measurements (even better than any ERA5 heights for the coastal cases).

Response: Thank you for this valuable and interesting recommendation. Following the reviewer's suggestion, we further analysed the relationship between observed hub-height wind speed and ERA5 wind speeds at different model-level heights as a function of wind direction. Although the analysis is based on correlation rather than model error, it provides complementary evidence on the directional dependence of ERA5 representativeness in complex terrain.

The new figure shows that the ERA5–observation relationship is strongly dependent on wind direction and height. In general, the highest correlation values occur for the dominant wind sectors, which also correspond to the directions associated with higher wind speeds at the mast sites. These high-correlation sectors are mainly concentrated around easterly to southeasterly directions at several sites, whereas weaker or even negative correlations are observed for some less frequent sectors. This directional behaviour suggests that the improved performance of upper ERA5 model levels is related to the interaction between the large-scale flow and the local topography, particularly the exposure of each mast to the prevailing wind regimes and ridge orientation.

We have added Fig. 5 to illustrate these direction-dependent correlation patterns. The black horizontal line indicates the optimal ERA5 model level selected from the overall correlation

analysis in Fig. 4. The results show that this optimal level generally intersects the height range with stronger correlations for the dominant wind directions, supporting the interpretation that upper ERA5 model levels better represent hub-height winds in the complex terrain of the tropical Andes.

We introduce the following text in Lines 353-362: *“To further explore whether the improved agreement between ERA5 upper model levels and hub-height observations is related to terrain-induced directional effects, we analysed the correlation patterns as a function of wind direction and model-level height. Figure 5 shows that the relationship between observed wind speed at 80 m and ERA5 model-level winds is strongly dependent on wind direction. In general, the highest correlations occur within the dominant wind sectors at each mast site, particularly for easterly to southeasterly flows. The selected optimal ERA5 model level generally intersects these high-correlation bands, indicating that upper ERA5 model levels better represent hub-height winds when the large-scale flow is aligned with the prevailing local wind regimes. Conversely, weaker or negative correlations in less frequent sectors suggest a stronger influence of local terrain effects and direction-dependent flow modification.”*

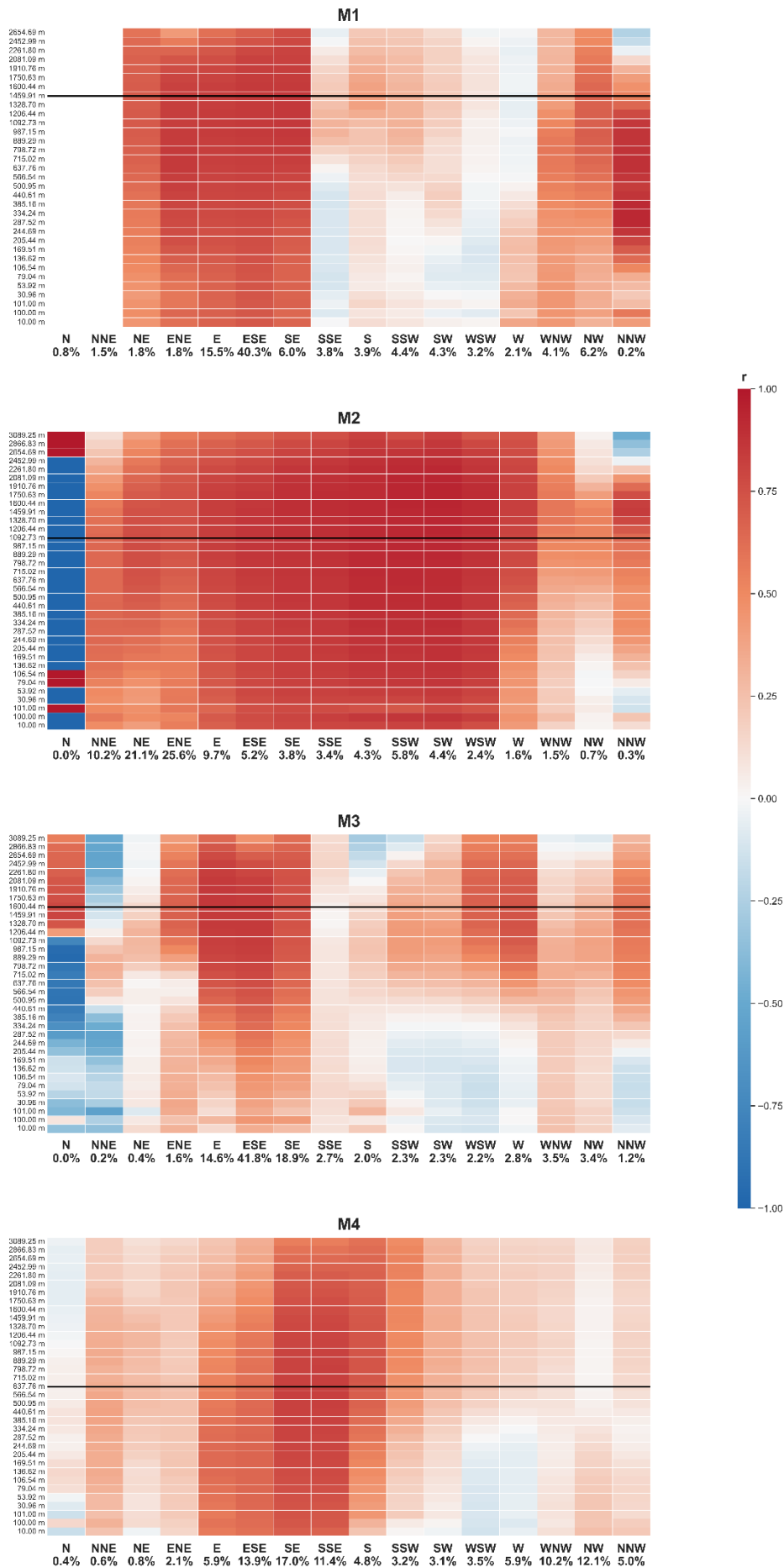


Figure 5: Correlation magnitude between observed wind speed at 80 m and ERA5 wind speeds at different model-level heights, stratified by wind direction for each mast site. The colour scale represents the Pearson correlation coefficient (r). The percentage of recorded data for each direction is shown below the direction labels. The black horizontal line indicates the optimal ERA5 model level selected for each site according to Fig. 4.

Second, I would consider methods to remove the need to select the optimal ERA5 height prior to model training. For instance, instead of a single, site-specific optimal height, multiple standard heights identical across all sites could be used. This helps show that the ERA5 data contains useful physical information that the model can utilize to improve wind speed estimates. As it currently stands, the optimal height selection actually includes the testing data as well as the training data. This results in a data leakage that could artificially increase model performance. If the authors decide not to use a multi-height standard model, then the optimal height selection should at least be performed on the training data only. Alternatively, a simple sensitivity analysis could be done to show if the optimal height at each site varies significantly from year to year.

Response: Thanks for this valuable comment. Regarding the first part of the observation, we followed the second suggestion. We chose it to strength our motivation about specifically demonstrating that higher ERA5 model levels can provide a better representation and prediction of hub-height wind speeds than the commonly used 10 m and 100 m single level data.

Regarding the second part of the observation, in the revised version of the manuscript, to select the optimal model height at each site, we excluded the testing period (2024 year) from the analysis. This change resulted in the selection of the L133 height instead of L114, which was selected in the first version of the manuscript, at M3. This change led to minor variations in the results but did not change the main findings or conclusions. We updated the text and figures with the new analysis for M3 in the revised manuscript. We also conducted a sensitivity analysis to assess whether the selected height remained the same, or nearly the same, across different years and periods. The results indicate that the identified heights are broadly consistent over time, with only minor year-to-year variations, supporting the robustness of the selection. This analysis has been included in Fig. A3 and is described in Lines 347–352 of the revised manuscript: “*The model height selection was done based on the analysis of 3 years (2021 to 2023). This is 2024, which is later used as testing dataset, was excluded from this analysis in order to avoid artificially increase model performance. In addition, a sensitivity analysis was performed to assess whether the selected height remained the same, across different years. The results indicate that the identified heights are broadly consistent over time, with only minor year-to-year variations (± 1 height level from the optimum identified in Fig. 4), thereby supporting the robustness of the selection (see Fig. A3).*”.

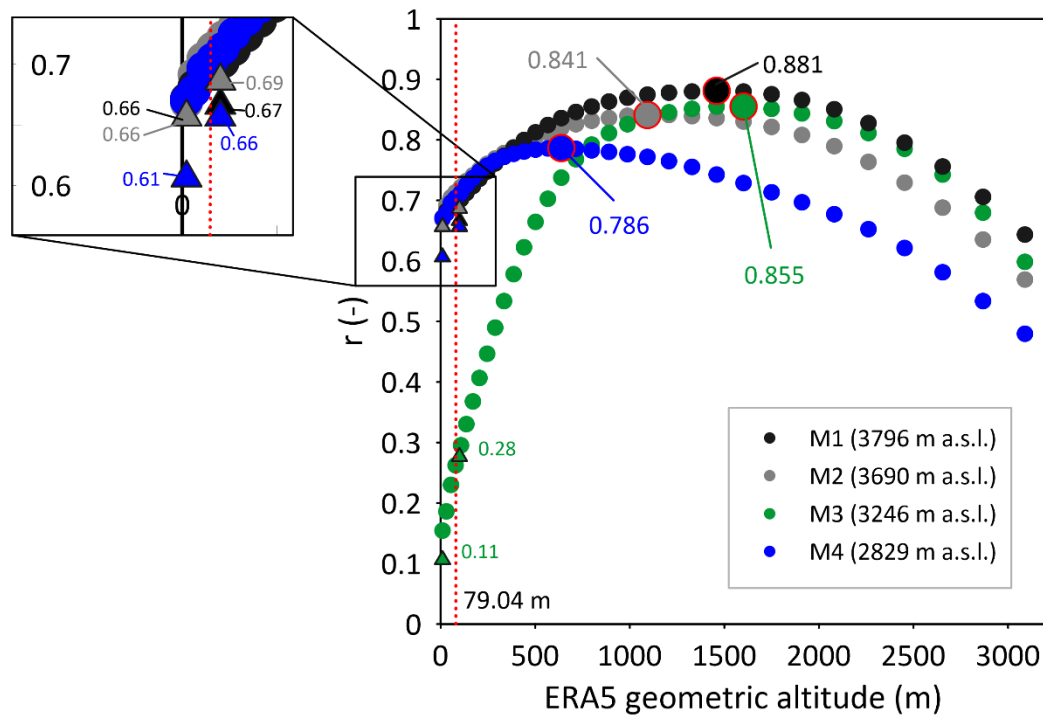


Figure 4. Correlation between observed wind speed at 80 m height at four sites and ERA5 data at different geometric altitudes for the period from January 2021 to December 2023. The 10 m and 100 m wind speeds from the ERA5 single level dataset are indicated by triangles, with their corresponding correlation magnitudes included. For each site, the highest correlation value is indicated by larger circles. The red dashed line indicates the model height closest to the observations. Site elevation is indicated in parenthesis in the legend.

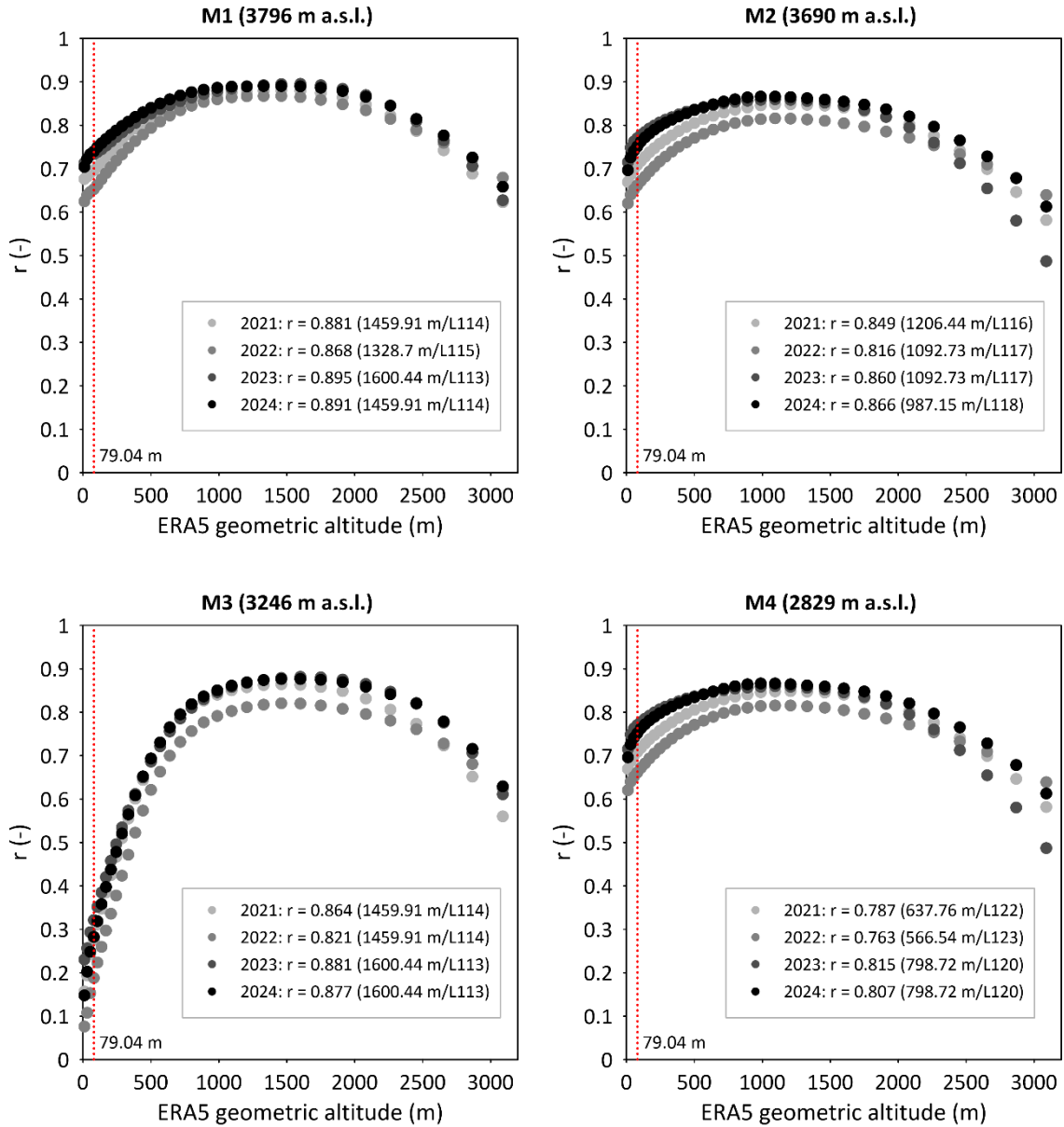


Figure A3. Correlation between observed wind speed at four sites at 80 m height and ERA5 data at different geometric altitudes for each year of the study period. The red dashed line indicates the model height closest to the observations. The highest correlation value and their respective altitude and model level (in parenthesis) are indicated for each year and site.

Below are minor technical comments:

1. For figure 4, I would like to also see the PDFs of the training wind speeds. This helps confirm that out-of-sample predictions are not taking place and that there is a different reason for the poor performance at the PDF extremes.

Response: To provide a clearer comparison between both datasets, the PDFs of the training and testing wind speeds have been included as Figure 3 of the revised manuscript. The following description was introduced in Lines 249-252: “Figure 3 compares the probability density functions (PDFs) of observed wind speed during the training and testing periods for

the four mast sites. Overall, the PDFs show that the main wind speed ranges observed during the testing period are also present in the training period, suggesting that the training data broadly capture the wind regimes later used for validation.”.

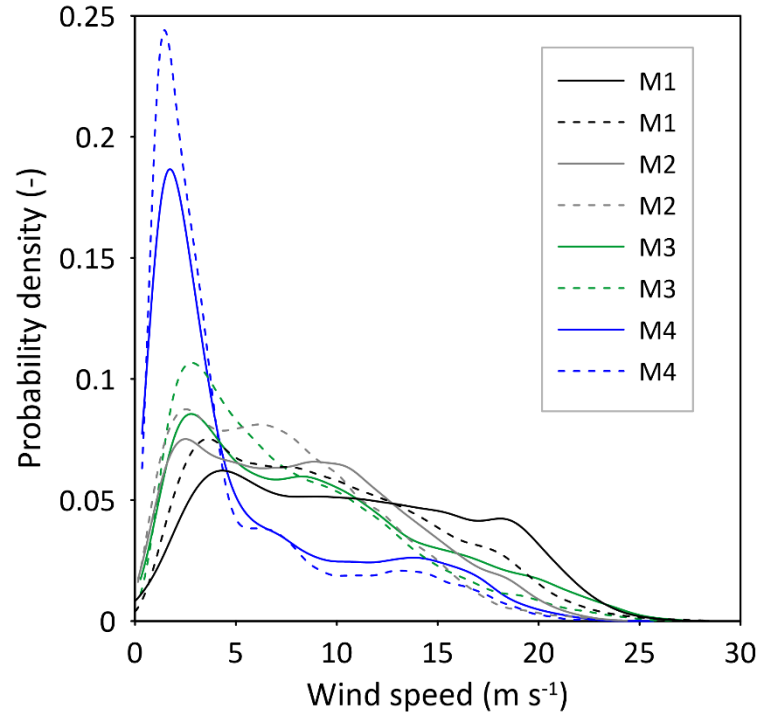


Figure 3. Probability density functions (PDFs) of observed wind speed data at 80 m for the training period (solid lines) and testing period (dashed lines) across all mast sites.

2. Having noted the Cavaiola et al. 2023 study in the introduction (line 69), it would be helpful to give more explanation as to how this study differs from the 2023 study. For instance, more details could be added in the discussion on lines 411-423.

More details about differences between our work and Cavaiola et al. (2023) is discussed in Lines 495-500 of the revised version: “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.”

REVIEWER 3

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 below 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 suggest 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. 7. Based on the coefficient of variation we confirmed a greater reduction in the dispersion of the distribution when using RF compared with LR (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 393-398: *“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 7 (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

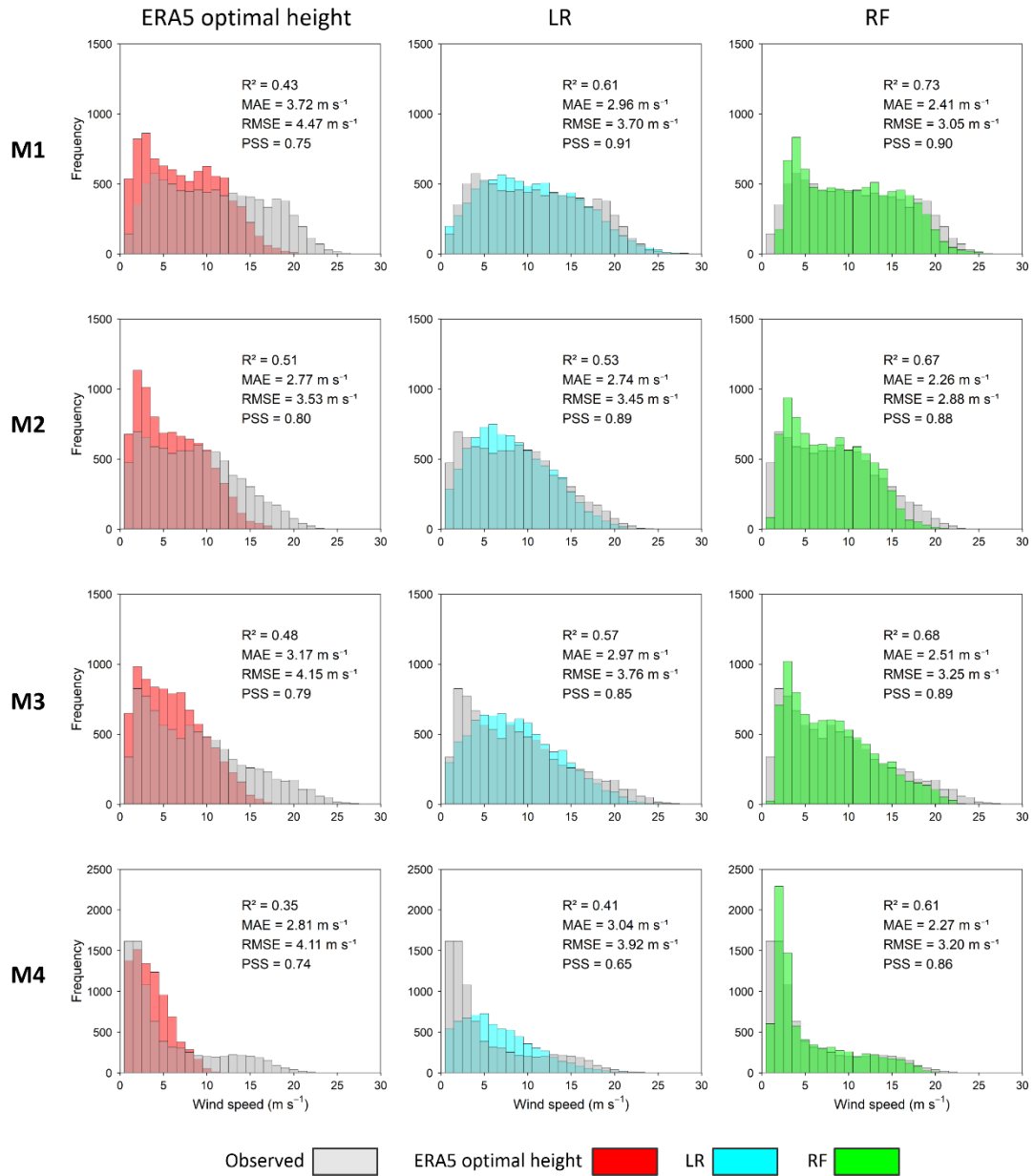


Figure 7: Comparison of wind speed frequency distributions between observed wind speeds, ERA5 wind speeds at the optimal levels, and wind speed estimates obtained using LR and RF with ERA5 optimal-level data as input across the four study sites for 2024. For each site, the figure also displays the performance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

- 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.

- 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).”.

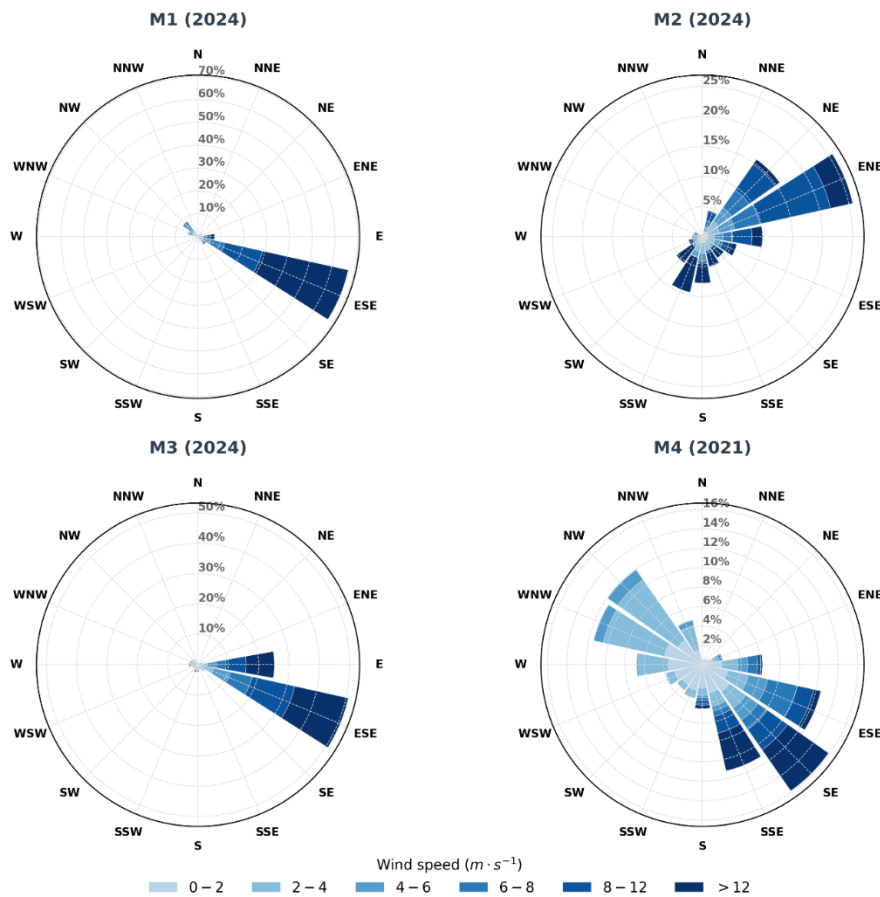


Figure 2: Wind roses showing the observed wind speed and wind direction distributions at the four mast sites. The plots correspond to 2024 for M1, M2, and M3, and to 2021 for M4 (selected based on wind direction data availability). Bar length represents the frequency of wind occurrence from each direction, while colours indicate wind speed classes in $m \cdot s^{-1}$. 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., $\Delta M1 = -755$ m a.s.l., $\Delta M2 = -649$ m a.s.l., $\Delta M3 = -1125$ m a.s.l., $\Delta M4 = -132$ 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 482-487: *“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.