Answers to Reviewer 2 comments

Juan Contreras and co-authors

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.

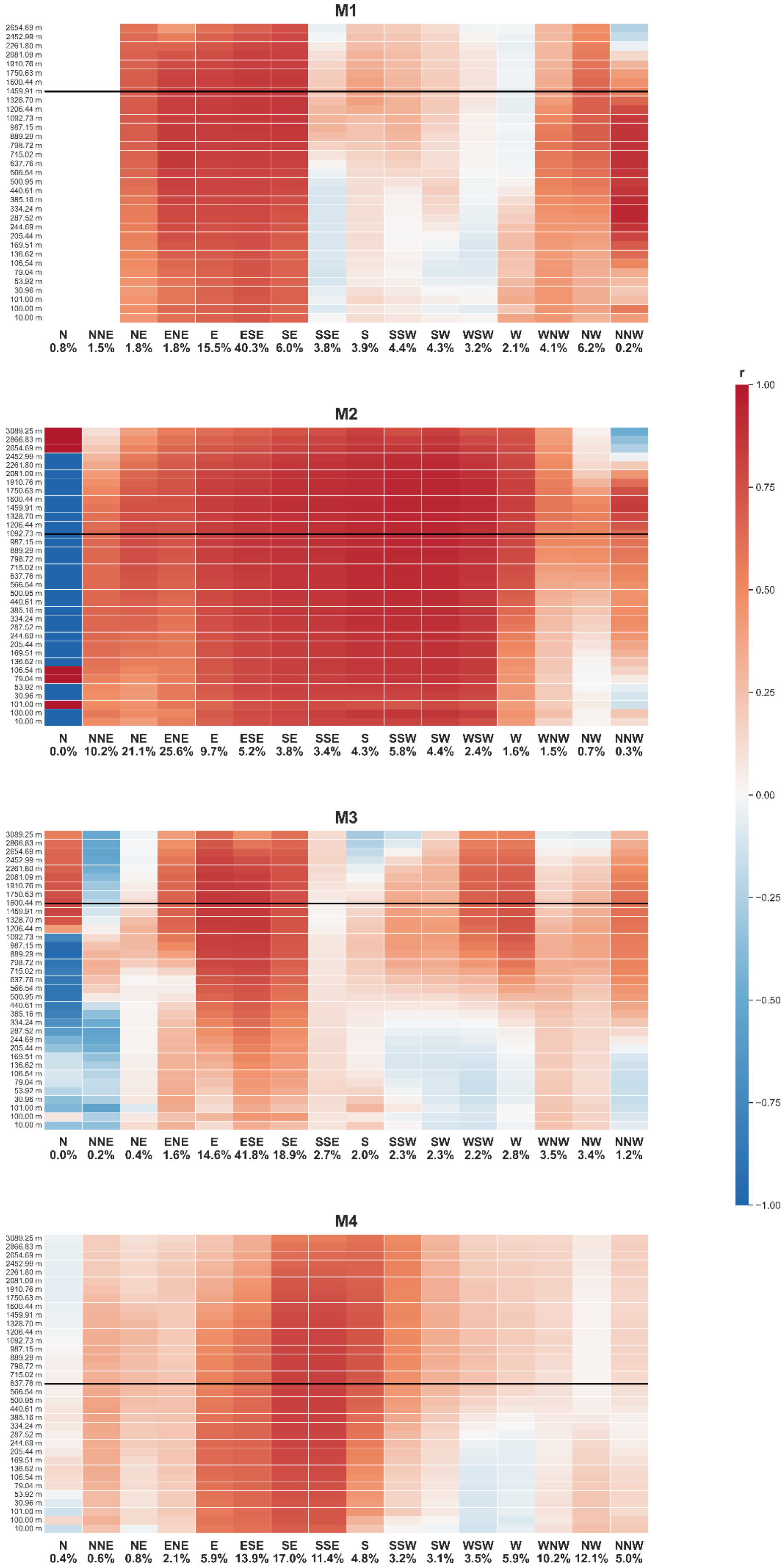
First, I would be interested to see typical model errors expressed for different hub-height wind speeds and directions. Do patterns emerge? For instance, does the error change dramatically when the wind direction is aligned with the ridge as opposed to cases where it is perpendicular to the mountain 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."*.



M1
N 0.8% NNE 1.5% NE 1.8% ENE 1.8% E 15.5% ESE 40.3% SE 6.0% SSE 3.8% S 3.9% SSW 4.4% SW 4.3% WSW 3.2% W 2.1% WNW 4.1% NW 6.2% NNW 0.2%
M2
N 0.0% NNE 10.2% NE 21.1% ENE 25.6% E 9.7% ESE 5.2% SE 3.8% SSE 3.4% S 4.3% SSW 5.8% SW 4.4% WSW 2.4% W 1.6% WNW 1.5% NW 0.7% NNW 0.3%
M3
N 0.0% NNE 0.2% NE 0.4% ENE 1.6% E 14.6% ESE 41.8% SE 18.9% SSE 2.7% S 2.0% SSW 2.3% SW 2.3% WSW 2.2% W 2.8% WNW 3.5% NW 3.4% NNW 1.2%
M4
N 0.4% NNE 0.6% NE 0.8% ENE 2.1% E 5.9% ESE 13.9% SE 17.0% SSE 11.4% S 4.8% SSW 3.2% SW 3.1% WSW 3.5% W 5.9% WNW 10.2% NW 12.1% NNW 5.0%
r
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

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).".*

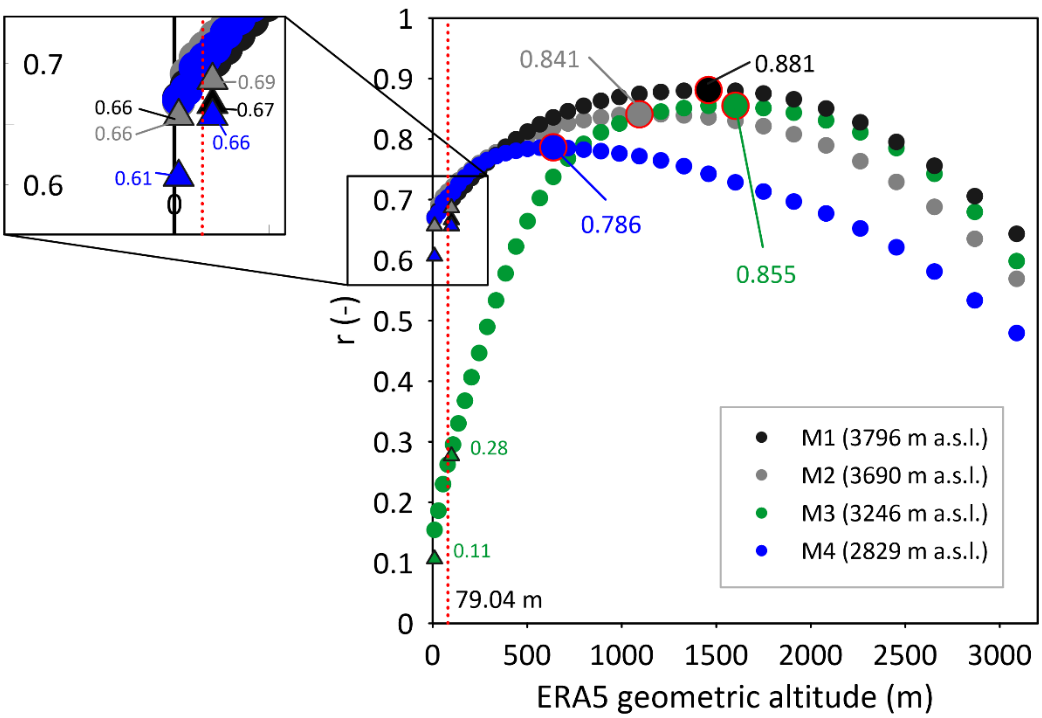


Figure 4. Correlation between observed wind speed at 80 m height at four sites and ERA5 data at different geometric altitudes for the period from January 2021 to December 2023. The 10 m and 100 m wind speeds from the ERA5 single level dataset are indicated by triangles, with their corresponding correlation 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.

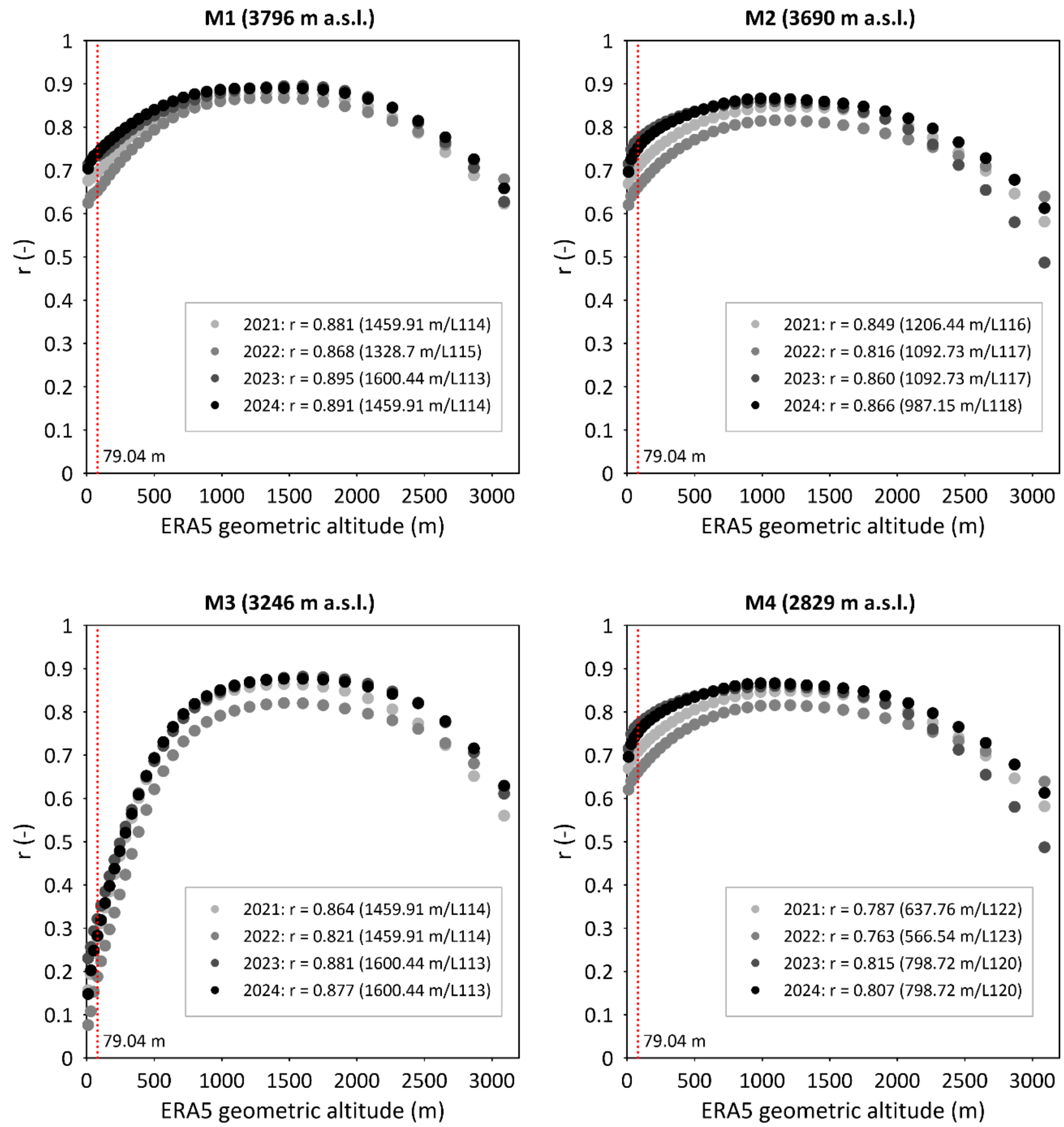


Figure A3. Correlation between observed wind speed at four sites at 80 m height and ERA5 data at different geometric altitudes for each year of the study period. The red dashed line indicates the model height closest to the observations. The highest correlation value and their respective altitude and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

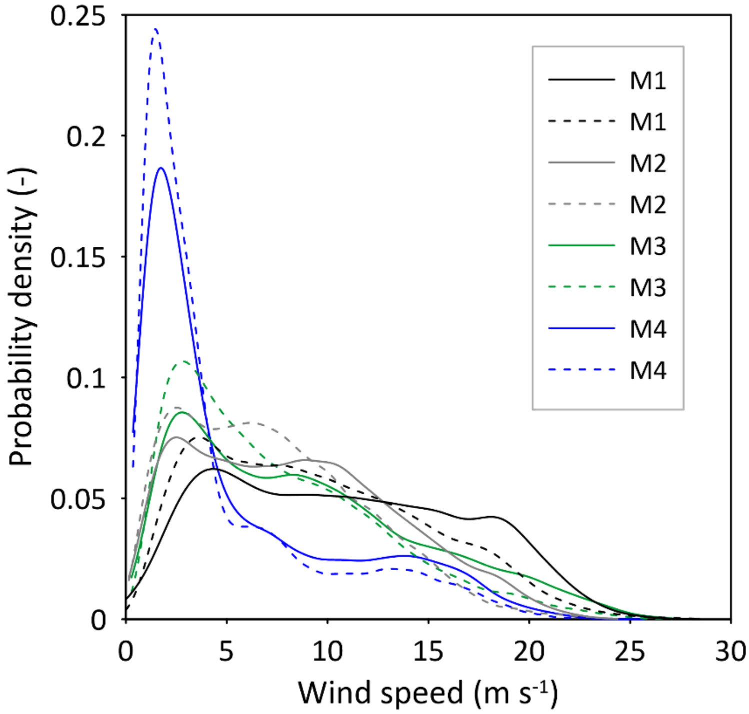


Figure 3. Probability density functions (PDFs) of observed wind speed data at 80 m for the training period (solid lines) and testing period (dashed lines) across all mast sites.

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."*