Answers to Reviewer 1 comments

Maksims Pogumirskis and co-authors

October 2026

We thank Reviewer 1 for the valuable comments and suggestions, which helped us greatly in improving the quality of the manuscript.

The reviewers' comments are in **bold** while our answers are not. Line numbers refer to the revised version of the manuscript.

Introduction (last paragraph): The final paragraph of the introduction should be more informative. At present, it is excessively vague. For example, a sentence such as "Section 3.1 describes the results of the comparison between models and observations" should be reformulated. The purpose of this paragraph is not merely to announce the structure of the manuscript, but to briefly summarize the substantive content of each section. For example: "Section 3.1 shows that X and Y. . . "

We expanded the paragraph to be more informative. Lines 124-134.

Terminology ("winds"): The language is sometimes imprecise, whereas scientific writing requires precision. I noticed the term "winds" used without qualification (e.g., lines 132, 145, 146). On its own, this is ambiguous, as it may refer to wind speed, wind direction, or other wind-related quantities. If wind speed is intended, this should be stated explicitly.

Thank you for noting this. We intended to use "winds" as a term for both wind speed and wind direction at the same time. We agree that sometimes we used the term improperly referencing only the wind speed. We corrected the wording throughout the manuscript.

English language quality: The manuscript would benefit from thorough proofreading. The current level of writing often obscures the scientific content. Shorter sentences with straightforward grammatical structure (subject+ verb + complement) would substantially improve readability.

The manuscript was proofread, and the language was improved.

Figure 1: I am not convinced that "Scandinavian mountains" and "Central European plains" are the most appropriate regional descriptors, as these terms encompass more heterogeneous areas than those shown. Simpler geographic labels such as Northern Europe, Central Europe, and Southern Europe may be clearer.

We agree that our defined region boundaries are slightly different from the ones commonly understood by the naming we used. However, our results show that model biases depend on the terrain properties, which we would like to emphasise. Therefore, we think that using just "Northern Europe" and "Central Europe" would not be informative. Based on the suggestion we decided to adjust the naming to "Scandinavian mountainous region", "central European lowland region", and "Southern Europe" throughout the text.

Figure 7: A latitude/longitude scale may improve clarity.

A latitude/longitude scale was added to the figure.

Lines 221–228: This paragraph is difficult to read and should be reformulated in clearer English. Also, using SD for standard deviation feels somewhat unconventional in this context. I would

suggest using notation such as sigma z since sigma is commonly used for standard deviation. If an abbreviation is retained, it should be introduced more clearly.

We reformulated the paragraph to improve readability. We changed $SD(z)$ to σ_z throughout the manuscript.

Figure readability: The font size in the figures should be approximately consistent with the font size used in the captions. In particular, the text in Figures 3, 5, and 7 is too small to be comfortably readable.

Thank you for noting this. We increased the font size for figures 3, 5, 6, and 7.

Lines 256-259: The justification for using hourly observations refers to turbulence characteristics, but turbulence is not resolved at mesoscale resolution, so this argument is not convincing as written. A clearer justification would simply be that hourly resolution provides a practical compromise: using 10-minute data may introduce excessive aliasing relative to the model resolution, whereas averaging to 3-hour intervals would unnecessarily reduce the number of available data points.

Thank you for the comment. The paragraph was corrected (lines 267-270).

Section 4: Quantitative support in the discussion: Discussing the paper scope and impact qualitatively is useful, but the discussion would be stronger if it included more quantitative elements. Long discussions based mainly on qualitative statements can become difficult to follow and, at times, confusing. Please support the main interpretations with numerical values, ranges, differences, percentages, or other quantitative indicators where possible.

We expanded the discussion section with quantitative details, where possible.

Across the manuscript: I suggest using "elevation bias" instead of "elevation difference between the model and the real world" for brevity and clarity.

We defined the term "elevation bias" in the introduction and adjusted it throughout the manuscript.

Line 807-808: Microscale modelling and terrain representation: The statement that microscale modelling can "significantly reduce the elevation discrepancy" is too broad. Higher model resolution does not automatically reduce elevation bias if the underlying terrain data or terrain preprocessing is inadequate. Please reformulate this sentence to make clear that improvements from microscale modelling depend on the quality and accuracy of the terrain representation.

Thank you for the suggestion. We reformulated the sentence.

Answers to Reviewer 2 comments

Maksims Pogumirskis and co-authors

October 2026

We thank Reviewer 2 for the valuable comments and suggestions, which helped us greatly in improving the quality of the manuscript.

The reviewers' comments are in **bold** while our answers are not. Line numbers refer to the revised version of the manuscript.

1 Major comments

My strongest recommended change is to soften the language on causality (e.g. the wind biases can be explained by differences in terrain height" (line 11). Based on the evidence, such language is too strong, and I suggest that "terrain is a statistically important explanatory factor". This is especially true, given that the authors acknowledge several possible causal factors (lines 569 – 574), and state that identifying the exact causes of wind direction biases requires manual inspection at each site.

Thank you for noting this. We carefully read through the manuscript and corrected the language on causality.

2 Minor comments

Line 31: Consider defining a range for typical hub height "... which are less useful for estimating the wind at turbine heights (typically x to x m).

Thank you for the suggestion. We added the height range of interest. In addition, we changed the wording throughout the paper to avoid ambiguity. We are interested in winds at heights spanned by turbine rotors rather than just the hub heights.

Line 45: "match observations" – numerical models will never align perfectly with observations. I suggest "generally agree with" can replace "match".

The wording was corrected.

Line 50: "atmospheric", not "atmosphere".

The wording was corrected.

Line 54: "Smaller scale processes than the effective resolution..." suggest "Processes that are at a scale smaller than the effective resolution..."

The wording was corrected.

Line 69: please check (jaf, 2011). There is a referencing formatting error here. In the reference list, there is an entry where the author field is missing from the bibliography entry: Terrain Effects on Wind Flow: Simulations With an Immersed Boundary Method, vol. Volume 1: Aircraft Engine; Ceramics; Coal, Biomass 3 and Alternative Fuels; Wind Turbine Technology of Turbo Expo, https://doi.org/10.1115/GT2011-46240, 2011. This may be that reference.

The reference was corrected.

Line 73: "covering the entire Europe." suggest changing to "covering the entire European continent."

The wording was corrected.

Line 100: remove "models" and change "overestimate" to "overestimates", so that the sentence reads correctly.

The wording was corrected.

Lines 107-108: "a representativeness bias is present in models, since the nearest model grid-cell is not always the most representative of the actual observation site." This is true, although, for example, WRF allows for output of U and V on the staggered and unstaggered grid, which allays some of those challenges. I suggest noting this in the text.

The sentence "a representativeness bias is present in models, since the nearest model grid-cell is not always the most representative of the actual observation site." is based on the previous reference to Jimenez 2012. Jimenez 2012 does not mention anything related to grid staggering. We adjusted the sentence to clarify that it is based on Jimenez 2012. We also added a sentence about potentially improving the representativeness bias using grid staggering/interpolation. Lines 110-112.

Line 139: Suggest adding the reference to support ERA5s use as the gold standard for wind energy applications/modelling (Olauson, 2018).

Thank you for the suggestion. We added the reference. Lines 144-145.

Line 140: "than in the..." change in to "is".

The wording was corrected.

Lines 146-150: I was going to suggest that, due to its significantly higher resolution (9 km vs. 28 km), the authors examine ERA5 Land as part of their study (or at least run a quick check against the other datasets), but this is not advisable considering recent developments (Bass, 2026). While not strictly related to the paragraph at hand, I think that the authors can pre-empt any criticism about not considering the higher resolution ERA5 Land dataset in their work by citing Bass (2026) (and address a wider-issue at the same time). I suggest the authors read this work and draft a short paragraph to this effect.

Thank you for the suggestion. However, there is a simpler reason why ERA5-Land was not considered in our work. It provides winds only at 10 m height, which is not sufficient to extrapolate winds to observation mast heights with the methodology we used.

Our work focusses on extracting the model bias that is correlated with the elevation discrepancy. Therefore, we do not think that the poor performance of ERA5-Land on its own as is mentioned in Baas (2026) is a reason to exclude it from our analysis.

ERA5-Land is not directly produced using CFD numerical simulations. Winds in ERA5-Land are interpolated from ERA5 and processed using a land surface model. In contrast to temperature and pressure, there is no accounting for terrain elevation discrepancy between resolutions for wind speed in ERA5-Land. Therefore, it would definitely be interesting to investigate how the wind biases related to the elevation discrepancy propagate from ERA5 to ERA5-Land. However, we think that this would be outside the scope of our work, which focuses on mesoscale/synoptic scale CFD models.

We expanded the paragraph mentioning the existence of ERA5-Land and justifying the reason it was not considered in our work. Lines 152-155.

Line 184: Suggest also stating the approx. km resolution for the JRA-3Q dataset.

Was added.

Lines 207-208/Figure 2: While not necessary, I think that to aid interpretation of the study, it is useful to include a figure alongside the observation density figure, that shows the density of wind farms in each country. I appreciate that the focus of the study is terrain bias – it

is helpful to appreciate the existing wind farm density across Europe relative to observation location density and SD(z). It's true that measurements are carried out at the potential locations of wind farms, and one can also assume that the historical EMD mast record covers wind farms that went ahead post-measurement. By having an appreciation for existing density, your work gains even more application, as many studies have examined real-world impact of wind turbines using parameterizations (e.g. Fitch, Volker etc.) that might be impacted by terrain factors, and modelled bias from wind resource assessment.

Thank you for the suggestion. We included a map showing installed wind power capacity by country and discussed how it relates to the significance of our work. Figure 2, lines 221-224.

Line 251-255: I know that a previous reviewer criticised the lack of commentary on quality control. I appreciate the challenges with NDAs etc. Can the authors perhaps provide rough statistics rather than specifics? The revised paragraph is fine, and it's just helpful to include some high-level statistics, if possible, to allay any reviewer criticism.

Thank you for the comment. We included that handling of raw measurement data follows the procedure described by "MEASNET: Evaluation of Site-Specific Wind Conditions".

Line 268: "360°was added. . . ." Need a space between the degree sign and was.

We corrected this issue throughout the manuscript.

Line 383: Please reference the previous studies showing stronger wind speed biases in complex terrain.

We included the references.

Line 414: It's important to cite such studies, although this is not relevant to the study area. I suggest the statement is adjusted to reflect that elevation source datasets aren't necessarily materially wrong across Europe e.g. "Model biases can arise not only from smoothing and scale mismatches between model and real-world terrain, but also, in some regions, from inaccuracies in the underlying terrain datasets themselves." Alternatively, there is stronger, more applicable evidence across Europe that will allow for a firmer statement e.g. "Previous studies have shown that mismatches between model and observed terrain elevation are associated with substantial climatological biases in complex terrain." Matiu et al. (2024) for elevation-dependent climate-model biases across the Alpine region. This is suitable evidence to show that terrain representation matters in Europe, rather than evidence of erroneous DEM source data, although the authors do discuss this (lines 424 – 429).

Based on your suggestion, we expanded the paragraph with the provided reference. Lines 421-422.

Line 421: The authors are quite correct, although it is possible to do some detective work to trace lineage (this is easier for MERRA-2 than JRA-3Q, which is more circumstantial). I suggest changing "To the authors knowledge, the terrain elevation source data for MERRA2 and JRA-3Q are not publicly disclosed", to: "The model terrain fields used by MERRA-2 (GEOS-5) originates from global topographic datasets, principally USGS GTOPO30, and is subsequently interpolated and smoothed to the native atmospheric model grids." "JRA-3Q employs a fixed model orography inherited from the JMA global NWP system (December 2018 configuration), represented through invariant surface geopotential fields. The public JRA-3Q documentation reviewed here does not explicitly identify the original DEM from which the model orography was derived."

Thank you for the suggestion. The paragraph was adjusted.

Line 491: The figure ordering is incorrect – Figure 10 is first mentioned here, whereas Figure 8/9 are mentioned further down in the text. Figure 10 should be moved to Figure 8, and so on, so that the figures are presented in the correct order that they are first mentioned in the text.

The ordering of figures was corrected.

Line 498: "describe the modes how the temporal patterns..." this sentence needs to be corrected – it appears that 'the modes' was inserted in the previous revision, but the sentence now doesn't make sense grammatically.

The wording was adjusted.

Line 506: "Further in the work PC1 of the wind speed" – insert comma so that the sentence reads "Further in the work, PC1 of the wind speed".

The issue was corrected.

Line 520: "Wind speed bias is more positive during the day and more negative during the night." I understand what the authors mean – it is confusing. I suggest: "The magnitude of negative wind speed bias is greater at night, and less negative during the day" (or language to that effect).

Thank you for the comment; we implemented your suggestion.

Figure 10 caption: Please add to the caption for (a) that this is for mountainous regions, and for (c) that this is for the Central European Plain. For (b), I am assuming that this is for the entire domain – if so, add that to the caption for (b). Also, the definition for (a) in the caption should come before (b) and (c). I suggest checking all relevant figure captions for clarity and adding specifics where necessary.

We adjusted the caption for this and similar figures throughout the manuscript.

Line 527: Again, the figure numbering is out of order – Figure 12 is mentioned before Figure 11, so Figure 12 should be relabelled to the order it is mentioned in text.

The ordering of figures was corrected.

Line 555: I appreciate NDAs etc., but I am not certain it is wise to add findings if the authors aren't able to show the data to support those findings. I suggest removing statements such as this where the data can't be shown to support.

Because of the NDA we cannot show exact locations of observations, and we can only show spatially aggregated results. While spatial aggregation works well when a clear spatial pattern is present, it cannot be reliably used to demonstrate the absence of clear spatial patterns. We decided to use the correlations between values of PC1 of wind direction bias for each pair of models to demonstrate that spatial patterns in all models are similar. We adjusted the paragraph to support our claims without strongly relying on the fact that there is a lack of evident spatial patterns. Lines 568-574.

Line 569 - 574: See my general comment re: line 11 about overstating results given that there are several plausible causal factors that the authors identify here.

See the reply for the general comment.

Line 607: Show, not "shows".

The wording was corrected.

Line 617: Suggest adding between March – October to this sentence. Do you consider a difference of 0.2-0.3 m/s to be sharp?

We think that the effect for November-February is also noticeable in the plot. The magnitude of the effect is just smaller. We agree that "sharp" is not the best word to use to describe our observation. We adjusted the wording to better describe the nature of the effect. Lines 635-640.

Line 626: Initialize [the] model....

The wording was corrected.

Line 646: This is a particularly important point and should be emphasised in the conclusions. Strong elevation gradients are particularly problematic, and I have personally experienced

Courant–Friedrichs–Lewy (CFL) violations when running WRF with >60 vertical levels and an insufficiently low time step on high resolution domains. It often becomes impossible to achieve convergence without a time step that is sufficiently low, which is impossible due to computational resource constraints. Thus, the authors suggestion is entirely valid and essential to state, as I expect enhanced terrain would cause convergence problems in WRF.

Thank you for the comment. We emphasised the importance of balancing the terrain smoothness to represent terrain features, while ensuring the numerical stability, in the conclusions. Lines 901-906.

Comments:

- **Line 664: WRF model simulations. Not models.**
- **Line 667: As above, "Similarly to observations, the principal components of the differences between two models show correlations with the elevation difference between these two models." It should be simulations, not models. Please correct all instances where the WRF simulations are referred to as "models", as there are several instances where this is the case.**
- **Line 675: "...when two models are compared". This should be simulations.**
- **Line 675: "...as when models are compared with observations". Only one model is used in this study, thus this statement needs to be changed to "...when the WRF model simulations are compared with observations."**
- **Line 675-678: Again, only 1 model is used in this study (WRF). So, it is not possible to make generic conclusions about all models – I suspect this is a continuation of the use of "models" instead of "simulations" to describe the 2 WRF simulations as models instead of simulations. Please correct these sentences.**
- **Line 681-700: As above – please correct all instances where the 2 WRF simulations are referred to as "models". In reference to the above comments – what you have are 2 sensitivity experiments/simulations using WRF – not two separate models. Changing the terrain dataset does not constitute a new model – it is a sensitivity experiment using a model (WRF).**

We carefully read through the manuscript and corrected the usage of words: model, model simulation, model run, etc.

Line 764: "...near the very surface". Should read "near the surface."

The wording was corrected.

The discussion can be bolstered by including additional references in their discussion of unresolved terrain and atmospheric stability:

- **One representative Helbig et al. paper on unresolved terrain/subgrid topography. I suggest Helbig et al. (2017) as they demonstrated that unresolved sub-grid terrain characteristics exert a measurable influence on near-surface winds and can be parameterised using terrain metrics beyond mean elevation.**
- **For stability. The authors could state that: "Atmospheric stability influences vertical wind profiles and wind-resource estimates (Kelly & Troen, 2016), providing a plausible mechanism by which model biases may vary with season and time of day."**

Thank you for the suggestion. We incorporated both references into the discussion section.

Appendix A: Please correct the Table A1 caption to reflect WRF simulations, not "models".

The wording was corrected.

Appendix B: Please expand figure captions in line with my previous comment regarding checking the in-text figure captions for clarity and adding specifics where necessary.

We clarified that the spatially averaged biases are for the entire domain of study.