Review of *Evaluating effects of the terrain on modelled winds in multiple atmospheric model datasets*. Maksims Pogumirskis et al. (2026). Submitted to *Wind Energy Science*.

Summary

This paper is not a routine mast vs. reanalysis dataset validation assessment. The authors are attempting a more ambitious task – to test a hypothesis that terrain elevation discrepancies between atmospheric datasets and reality is the primary driver of wind speed and direction biases across multiple atmospheric datasets, with the implication that modelling the wind resource at potential wind farm locations, is significantly impacted.

Using principal component analysis to identify the physical mechanism driving the biases, seven atmospheric reanalysis datasets are assessed against a large proprietary mast dataset from observational campaigns. A model sensitivity experiment, where the terrain resolution is changed, complements the mast and reanalysis dataset assessment.

The primary result of this work is a worthwhile contribution to the literature: Terrain elevation mismatch between a model and reality ***is a statistically important factor*** associated with systematic wind speed and some wind direction biases across multiple datasets.

General comments

Thank you for the opportunity to review. This paper is an excellent addition to the literature and suitable for publication in WES. The work is ambitious and I like the fact that the authors are attempting to identify a common mechanism underlying wind speed biases across multiple datasets. A task that is profoundly difficult to quantify. This work is not only applicable to wind resource assessment, but topics such as European windstorm, where the advancement of modelling can help serve the insurance industry in pricing risk. I think that the wind energy community is going to benefit greatly from this work and I am pleased that it will be published.

I note that this paper has been reviewed extensively by two previous reviewers. The authors have responded faithfully to those revisions, such that the paper is in good shape compared to the original submission.

My view is that ***the paper is ready for publication pending minor revisions***. I have read the previous reviews, and the authors have addressed the identified major methodological issues. My view is that the methodological approach (particularly the analytical methods and modelling sensitivity study) is how I too, would approach a study of this nature for submission to WES.

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.

Moreover, the evidence demonstrates a robust statistical relationship repeated across multiple atmospheric reanalysis datasets and is partially reproduced in the controlled WRF terrain experiment, but it does not conclusively demonstrate that terrain is the dominant physical cause. Indeed, the strongest correlation for wind speed is 0.71 (ERA5) and considerably lower for other datasets (a correlation of 0.37-0.46 for NEWA and EMD-EUR+ is less definitive).

Overall, the revisions have substantially improved the interpretation. Causal wording, however, remains inconsistent, particularly in the abstract, introduction and conclusions (e.g. line 843). The observational analysis establishes a robust statistical association, while the WRF experiment shows that changing the terrain input produces comparable wind-response patterns. Together these provide strong support for terrain representation as an important contributor, but they do not isolate terrain-elevation mismatch as the sole or complete cause of the observed biases. Please use consistently qualified language such as "is associated with", "provides explanatory power for", or "is likely to contribute to" instead of the definitive statements that are still currently used after the first revision.

My other comments are minor and should be easy to address. I recommend them to improve the quality of the final manuscript.

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).

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

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

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

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*

and Alternative Fuels; Wind Turbine Technology of Turbo Expo, https://doi.org/10.1115/GT2011-46240, 2011. This may be that reference.

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

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

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.

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

Line 140: “than in the…” change in to “is”.

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.

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

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.

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.

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

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

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).

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."

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.

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.

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".

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).

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.

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.

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.

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.

Line 607: Show, not "shows".

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?

Line 626: Initialize [the] model….

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.

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).

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

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."*

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

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.

References

Baas, C.W., 2026, Systematic under-representation of ERA5 10 m wind speeds in the ERA5-land product undermines wind energy studies. Applied Energy, vol 412, https://doi.org/10.1016/j.apenergy.2026.127634

Helbig, N., Mott, R., van Herwijnen, A., Winstral, A., & Jonas, T. (2017). *Parameterizing surface wind speed over complex topography.* Journal of Geophysical Research: Atmospheres, 122, 651–667. https://doi.org/10.1002/2016JD025593

Kelly, M. C., & Troen, I. (2016). Probabilistic stability and "tall" wind profiles: theory and method for use in wind resource assessment. Wind Energy, 19(2), 227–241. https://doi.org/10.1002/we.1829

Matiu, M., Napoli, A., Kotlarski, S. *et al.* Elevation-dependent biases of raw and bias-adjusted EURO-CORDEX regional climate models in the European Alps. *Clim Dyn* 62, 9013–9030 (2024). https://doi.org/10.1007/s00382-024-07376-y

Olauson, J., 2018, ERA5: the new champion of wind power modelling? Renewable Energy, 126, pp. 322-331, https://doi.org/10.1016/j.renene.2018.03.056