Review of wes-2026-24 titled "Evaluating effects of the terrain on modelled winds in multiple atmospheric model datasets" by M. Pogumirskis, T. Sile, L. Svenningsen and A. H. Hahmann

Overview: This study assesses seven numerical models of large use in the atmospheric community against over 500 observational datasets collected in Europe. The goal is to connect the models' biases in wind speed and direction with physical properties of the observational sites. A principal component analysis (PCA) of the bias reveals high correlation between the latter and the difference in terrain elevation between models and reality. The topic is of major interest and the large amount of data examined in this work (both numerical and observational) provides solid statistical foundation to the analysis. Thus, I recommend this article for publication on WES. There are, however, few major points to address before publication:

- The authors find compelling evidence correlating the bias in wind speed and direction with the discrepancy in the terrain elevation between numerical models and reality. Then, the authors argue that larger elevation discrepancies are present in complex terrain scenarios, which, in the authors' view, explains why synoptic and mesoscale models under perform in complex terrain. This last passage is not necessarily true and should be discussed on a case-by-case basis, although it may broaden the discussion too much. As an alternative, the authors could limit themselves by saying that elevation discrepancies are a major cause of model biases without relating it to local terrain complexity.
- A major result arising from the current analysis is the dependence of the wind speed (and direction) bias on the hour of the day and month of the year. This emerges from several figures, both in the model vs. observation and WRF environments. In the Discussion section, the authors link the diurnal bias to the cycle of thermal stability, yet they do not provide a comprehensive discussion of the seasonal trend of bias.
- Minor grammar errors are present throughout the manuscript. Please revise the text and address them.

More comments are provided below.

Specific comments

- Line 35: I would not refer to the year-long timescale as "climate" since climate typically spans over 10 years or more. Please consider replacing this word with "meteorology" or something similar.

- Line 71: Please add a short paragraph to summarize the content of the remaining Sections.
- Line 144: The authors claim to focus on the physical mechanisms causing wind speed biases on complex terrain specifically. By looking at the data availability at country level (Figure 1), it is unclear how many of the 580 observational locations are indeed on complex terrain. I recommend that the authors add a metric to quantify the local terrain complexity.
- Lines 159 - 160: I wonder whether the authors discarded temporal periods with too large variability of 10-minutes averaged wind speed values. Any intra-hour variability may cause uncertainties comparable to the biases quantified between observations and models.
- Lines 179-180: Although the authors discuss potential limitations of this interpolation technique, I believe that more details are needed about the vertical interpolation of wind speed. To my understanding, you utilize the power law to determine wind speed at the observed height; do you calibrate reference wind speed and shear exponent on numerical predictions? If so, how reliable are they considering that some models provide only 2 levels?
- Line 185: The roughness length may vary throughout the year depending on the site and vegetation type. Are you modeling this effect as well?
- Lines 240-246: This paragraph is a bit hard to read. Please consider adding further equations to clarify the normalization process you adopted.
- Lines 258 - 259: This point deserves more in-depth discussion. The authors claim that, within a certain area, the terrain variability is not well captured by the model resolution and, thus, the modeled average elevation is erroneous because it results from a smoother terrain – which is true. However, considering that models operate over ~1 km-by-1 km grids, the local wind speed is not only determined by the average elevation but also by flow distortions induced by sub-grid terrain variability. In other words, the authors should not only consider the bias in average elevation, but also the standard deviation of the terrain elevation within the averaging area.
- Lines 267 - 268: I am not sure about this statement. Local topographic features can induce channeling and other speed-up effects downstream a certain location. This can occur for spatial scales smaller than the model grid but relevant for wind energy. In other words, although fixing local biases is definitely valuable, its footprint on non-local model improvement still has to be determined.

- Section 3.1: I found this Section really hard to read, in particular the interpretation of temporal distribution of PCs. Why is there a distinction between different geographical areas if the PCA is not carried on separately for different areas? What is the reason of calculating the spatial average of PC1? My suggestion is to thoroughly review this Section at least up to line 395 and provide more clarity to the scope of this analysis.
- Lines 328 - 329: The entire section is based on distinct biases found for different regions - Central European Plain, Southern Europe and Scandinavian Mountains. However, the precise boundaries of these regions should be directly reported in Fig. 4 for a meaningful correlation with the spatial bias distribution.
- Lines 367 - 369: I am a bit confused by these lines. The caption of Fig. 4 (as well as line 330) describes the spatial distribution of PC1 of wind speed bias, whereas at lines 367 - 369 the authors state that this figure reports the values of PC2 of wind direction bias. Please clarify.
- Line 401: In Figure 9, PC1 of wind speed shows a high correlation coefficient with elevation errors (0.71), whereas this relationship is less obvious for wind direction PC1 and PC2 (correlation coefficients of 0.5 and -0.41). Please make this distinction clearer.
- Line 403: What is *P*?
- Line 425-432: The authors claim to interpret the observed correlations between PCs and elevation differences between models and reality, but the following lines provide no interpretation, rather they are a mere description of what emerges from the figures. Inter-changing result's description and results' interpretation is a recurring theme of this manuscript which I suggest to avoid in the future.
- Line 504: The authors argue that observational campaigns are more often located on top of hills, which would explain the negative PC1 values. This statement is not generally true as it depends both on the observation site and the scope of the field campaign. If this is true for your set of observations, please state it.
- Lines 525-526: This result is shown for synoptic spatial scales in the model vs. observation case, but it is unclear how you can infer it for mesoscales since there is no analysis for individual observational sites.
- Lines 570-572: Again, this is true for flat terrains, while on complex terrain this is not necessarily true and depends on the local topography as well as on the distance between probed locations.

- Line 584: Please provide a brief explanation of the Laplace of the elevation.

Technical comments

- Line 147: Please replace “focusses” with “focuses”.
- Line 223: Please replace “Lets” with “Let us”.
- Figure 3: Please add labels for each panel.
- Line 358: Please replace “shows” with “showed”.
- Line 424: Please replace “lets” with “let us” throughout the manuscript.