

In this document, the reviewer's comments are in black, the authors' responses are in red.

We thank the reviewer for their thoughtful comments, which gave us an opportunity to revisit our analysis.

Scope

The scope of the manuscript is well-suited for this journal.

Originality

The question of originality is not critical for a data journal. However, this is very important, and it would encourage other groups to publish the description of their datasets in this manner.

Scientific rigor

In general, the scientific rigor is adequate. However, I have a few questions and clarification points below.

1. The wind time series uses a 5-minute temporal resolution. Have you checked that the spectra contain energy at this time scale? If not, this should be mentioned somewhere.
The choice of the 5-minute temporal resolution comes from the needs of the grid integration community. We have now added a sentence to mention this ("The choice of the 5-minute temporal resolution is also dictated to accommodate needs of the grid integration community."). Regarding spectra, we have not checked those for NOW-23, but we have relied on the analysis that was completed for the previous-generation WIND Toolkit (see Fig. 9 in <http://dx.doi.org/10.1016/j.apenergy.2015.03.121>), which did not show a significant peak at 5 minutes.
2. L104-109: The ERA5 reanalysis uses the OSTIA SST product. This should be mentioned, and unsurprisingly, forcing the WRF model with the same SST is often advantageous.
We mention this aspect in Section 2: "The first SST product we consider is the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) data set produced by the UK Met Office, which provides data at 1/20 deg horizontal resolution and is the standard product included in both ERA5 and MERRA-2."
3. How can you explain the observed wind profile in Figure 13? By the way, what is the source of SST for the Great Lakes? Are these points treated as lakes or seas?
Unfortunately, the relatively limited (in time) and old nature of the Great Lakes data set limits our ability to be fully confident about the observed data: the shape of the mean wind profile could be due to poor QC of the raw lidar data (not performed by us and not accessible to us) or it could be a physical mechanism (e.g., a low-level jet observed on average at that location over the period of record). Regarding the SST data for the region, the same data set (OSTIA) is used. We have clarified this in the text.
4. L380. It is not clear what "overestimates atmospheric stability" means. More stable? Values of temperature gradients? Please clarify.
We have changed it "overestimates the frequency of stable conditions".

Writing

The writing is clear and well-structured.

Length

The manuscript is a bit long. Also, the structure is tedious, looking more like a report than a scientific paper. I suggest removing the sub-sections for each zone, which often contain only one or two sentences. The eight regions could be combined according to each ocean basin.

We have removed all sub-sections as suggested. We have also merged most tables into two (see answer to one of the comments below). As a result, the article is now 6 pages shorter. We still kept the current divisions into 8 regions to follow the division of the NOW-23 data set into 8 modeling domains.

Figures and tables

I suggest redrawing most of the wind profiles and Taylor diagrams. It is not possible to distinguish the various simulations. I suggest using a logarithmic y scale in the profiles, which is standard in wind energy. For the Taylor diagrams, it is possible to zoom in the relevant region so that the differences between the runs are highlighted. See, for example, <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017JD027504>.

We have updated the figures as suggested.

Are tables 2 and 12 the same? Could the tables be compressed to shorten and facilitate reading? I suggest one table with all the possible ensemble members and a final column stating which domain they are used in.

We have merged all model setup tables into two tables (now Tables 2 and 3).

I also suggest that all figure captions contain the height and time period used in the validation (e.g., Figs 3, 4, 8, 9, 11). The period needs to be added in the caption of Figure 18.

We have updated the figure captions.

Title

The title is short and informative.

Abstract

The abstract is concise but contains most of the relevant information. The years and the model used in the simulations should also be included to facilitate future searches.

We have added the year and model information to the abstract.

References

The references are relevant and comprehensive.

Recommendation

I recommend publishing the article once the suggestions regarding the structure and the figures' redrawing are considered.

Minor corrections:

1. Please substitute "WRF" with "WRF model" or "the WRF model" when appropriate.
Changed.
2. L79: consider -> considered. Maybe the rest of the sentence should also be in the past?
Changed.