Overall feedback

This manuscript provides an extensive evaluation of the ADDA-v2 climatological wind dataset for North America. The manuscript text is logically structured and well-written and the figures are clear. A large set of measurement stations and different aspects of the model performance are considered over a range of timescales. The analysis goes into sufficient depth and an uncertainty analysis is included. It is also valuable that a comparison is done to the widely used dataset used to drive the model runs, i.e. ERA5 to see where the downscaled dataset significantly improves on the driving data. However, some choices or sections in the manuscript would benefit from changes or additions, which I will detail below.

General comments

1. It would be valuable if the findings of this evaluation (and the ADDA-v2 dataset itself) would be compared to other evaluation studies (and other datasets), preferably for the region. See also specific comment nr. 2. Hence, some references should be added in the discussion section (or somewhere else) to better situate this wind product in relation to other datasets and also how the evaluation/performance differs from other studies.

Specific comments

1. Page 2, line 40: I recommend adding that fine-scale surface variations (topography, land cover, …) are also better represented.

2. Page 2, Line 60: Considering that this wind dataset of Draxl et al. (2015a,b) exists and has an even higher spatial resolution, could you add somewhere in the manuscript why this new data product ADDA-v2 is perhaps a better choice going forward?

3. Page 3, line 72: 20 years is indeed a substantial length, but climate variability at 20-year timescales still exist – it could be valuable to mention somewhere in the manuscript how the ADDA-v2 dataset can be supplemented to account for this in wind resource assessments.

4. Page 4, line 98: Could you provide motivation for the choice of two months of spin-up? It is rather limited for the soil component as says a publication of Jerez et al (2020) [https://doi.org/10.1029/2019MS001945], definitely if simulations start in the cold months.

5. Page 4, line 122: On such a large domain, why is no spectral nudging employed for small wavenumbers? Has it been validated at some point that the properties of synoptic systems are indeed adequately reproduced? It would be good to comment on this.

6. Page 5, table 1: were there any model options activated to account for subgrid-scale orography (e.g. topo_wind for YSU PBL scheme or the GWDO scheme) – would be good to mention this.

7. Page 5: table 1: can you motivate the choice for 49 vertical levels? Low-level winds are usually quite sensitive to this choice. Perhaps sensitivity tests were conducted?

8. Page 5, line 138: are the observations also corrected for mast flow distortions? Perhaps good to comment on this and the implication on observational uncertainty.

9. Page 11, line 267: "As discussed in (Section 3.1)" should be corrected.

10. Page 11, line 277: I would not use "improvement" here.

11. Page 13, line 293: So is this the r between the seasonally-averaged values? Or is it the seasonally-averaged value of daily r values? Sometimes in the manuscript this is not very clear.

12. Page 19, section 3.1.3 (wind roses): I agree that a good approximation of the wind rose is a first indication that synoptic winds are well captured. However, the manuscript would benefit from any additional analysis (or references to related studies) which looks at this in more detail. I mainly say this because the domain is very large and no nudging is used. If this cannot be provided, it would be good to mention that this has not been inspected in full detail.

13. Page 19, line 377: section index 3.1.3 is repeated here.. should be 3.1.4 I think?

14. Page 21, line 409: your prior analysis of seasonally-average diurnal cycles shows that statistically (not for specific days) the diurnal cycle is captured well. Here you seem to suggest to the reader not to use the sub-daily information of ADDA-v2. Is this not too critical? If you would feed the diurnal cycles of 500 winter days from ADDA-v2 to a wake model to check performance of a wind farm, I don't feel like the output would not be trustworthy. So maybe rephrase this a bit or provide additional clarification.

15. Page 22, figure 6: As surface wind speeds are usually well below 10 ms-1 on average, RMSE does not give a good idea of how good the agreement actually is. Could you use the rRMSE here?

16. Page 22, figure 6: In the introduction you mention that wind datasets are important for risk assessments of high winds. Yet, for this surface wind speed evaluation, you focus on seasonal RMSE's. Why not also look at extremes? Is there perhaps a reason why this dataset is not intended for looking at extreme winds? Logically, for wind energy purposes one would be interested in winds higher up, so I'm curious what the surface wind speed output is for. You could add an analysis, or clarify why the presented analysis is sufficient.

17. Page 24, figure 7: These indicators '7-2', '5-1' – where do these come from? These should be explained somewhere.

18. Page 25, line 498: Instead of saying that high friction velocities correspond to weaker winds, could you explain a bit better to the reader why this is the case?

19. Page 26, section 3.4: An interesting addition. However, I am very interested to see the importance of the two components of model uncertainty: initialization and physics choices. I would expect that the lateral forcing would lead the initialization to not matter too much on annual statistics compared to physics parametrization choices. Please add this.

20. Page 28, figure 10: Is the inter-annual variability over a specific 2-week period something that wind farm operators are concerned about? As expected, this variability is very large. I think that the inter-annual variability on seasonal timescales might be good to add as well: e.g. how good or how bad a winter period can be in terms of wind resource over the lifetime of a wind farm. You could motivate your choice or add also the seasonal timescale to this analysis.

21. Page 29, line 553: Is the lower inter-annual variability in summer not simply a consequence of lower wind speeds? Relatively speaking, the variability could be as large as for winter or even larger.

22. Page 30, line 601: I get a "404 not found" when pasting the link to reach the hub height wind data. Please make sure that a reliable pointer is available to access the data and that some documentation is available there. Perhaps also for the full ADDA-v2 data: include an e-mail / link where to request access.