Review of *"Modelling wind farm effects in HARMONIE-AROME (cycle 43.2.2) – part 2: Wind turbine database and application to Europe"* by Jana Fischereit, Bjarke T.E. Olsen, Marc Imberger, Henrik Vedel, Kristian H. Mølller, Andrea N. Hahmann and Xiaoli Guo Larsén

The provided manuscript provides a description of the generation of a large wind turbine database covering most of the European on- and offshore wind turbines. Here, the authors make use of different data sources, to collect all placed turbines, filling data gaps using a random forest approach. Subsequently, the authors perform sequential 12-hour forecasts using the HARMONIE-AROME model, with a lead time of 48 hours for the months of February 2020 and August 2022 where both months combined represent the long term wind speed and direction statistics quite well. Within the numerical weather prediction model, the authors implemented two different wind farm parametrizations, based on the Fitch parametrization and explicit wake parametrization (EWP). The implementation of those into the HARMONIE-AROME model is described in part one of this paper series. The authors show that especially the Fitch parametrization provides good agreement with weather masts and lidar measurement at different on- and offshore locations across Northern Europe. Further, the simulations show that the incorporation of wind turbines into the NWP significantly changes wind speeds and temperatures in close proximity to wind farms and hence should be included in these models.

From my point of view the manuscript presents a very interesting study, contributing much to the scientific community and provides a great deal of novelty. It is well written and the structure is good to follow. The figures are well prepared and mostly contribute a lot to the understanding of the study. However, I identified several points that need further clarification before publishing the study. The specific comments are listed below.

Based on the review, I would recommend the manuscript to be published subject to minor revision.

Specific comments:

1. L. 77: When presenting the structure of the paper it would help clarify that there are appendices and what topics they cover.
2. Figure 1: In the caption of the figure you mention two abbreviations which are not introduced beforehand which I couldn't associate with anything right away (NEA, DINI)
3. L. 144f: You say that in some cases the algorithm to determine the wind turbine positions does not converge. Could you clarify in how many cases that happens. Is this a significant part of the installed capacity or would you consider this as negligible?
4. L. 181: Do you mean look-up-table instead of table look-up?
5. L. 184: Could you please clarify on the concept of generic classes here. Does this mean there is one generic class of turbines that shares all the same dimensions and characteristics or is there a generic class for each dimension (Rotor diameter, Hub height, Turbine model, …)
6. L. 189: You mention that the database uses e.g. an average height per turbine model (if the difference is smaller or equal to 9m. I.e. for larger differences, the same turbine can be in two different database entries? And how is this spread handled for other parameters, such as the rated power, which can also be altered for a turbine by e.g. applying a power boost.
7. L. 229: Considering the outliers observed on the 20% test data set. Did you perform an outlier detection and correction after applying your gap-filling algorithm?
8. L.229: Also, I am not too familiar with the random forest approach, but from your results it looks like the output you use for gap-filling are not discretized in any way. Is it true, i.e. can

turbines have rated powers of e.g. 9.75MW or are the values corrected to the closest real rated Power found at any turbine?

9. L. 232: I think the "(relatively)" should go before the "large" here
10. L. 247: I am not familiar with the PyWake package. Is "standard_power_ct_curve" a function defined in that framework?
11. L. 249: Is the turbulence intensity you use for turbine specification varied between offshore and different onshore locations for the generation of power curves?
12. Figure 7: Here it says "(a) power curves", but should be "power coefficient".
13. L. 271: I do not fully grasp the workflow here, what is meant by "3-hour-cycles"? If you could elaborate, that would be great.
14. L. 300: I think met masts are self-explanatory, but a little bit of information on the lidar scans would be nice. The description must not be extensive, but it would be good to know the resolution, scanning pattern and maximum height and what the difference between VAD-1600m and VAD 1500m? Especially considering Fig. 11 where the profiles differ quite a bit the differences between the scans would be good to know, since from Fig. 9a it looks like the location is the same.
15. L. 316: Could you elaborate a little bit on the data removal process. In "any missing time stamps at one height are removed both from all heights and from the forecasts." Does this mean that if from one observation, one point is missing, you delete all points from that height from the other data sources as well, or the other way around, if one height is missing, you remove that specific data source for that time stamp?
16. L. 316: Why did you choose a threshold relative to all other data sources here, instead of defining a threshold based on the entirety of the time steps, e.g. when availability for one source at one height lies below 80% we remove that specific height at that position?
17. L. 339: Here, maybe start the sentence with "A height of 90m […]". This way the sentence doesn't start with a number right away. This comment is purely based on personal preference, so no need in answering this if you keep it your way.
18. L. 359: In my mind it should be "at a height of 100m" instead of "at 100m height".
19. Fig. 13 & 14(b)-(d): Are the colorbars depicted correctly here? In the caption you mention difference and the Figure titles suggest, you subtract one wind field from another, but the colorbar is given in percentage. Which one is correct?
    a. The same applies for Figures 16-18.
20. L. 469: Is there a specific reason on why you did not choose a point further away from the coast (e.g. at the locations or to the west of Sandbank and DanTyks wind farms)? Comparing a,c,e, to b,d,f it seems that in regions far away from the coast the 2m Temperature differences seem to be quite strongly into the opposite direction as compared to closer to the coast. Here a time series evaluation of the profiles (similar to Fig. 20) would also be interesting.
21. L. 527: How do you aim at mitigating the effects overfitting might have on the turbine specifications? Do you consider shrinking the dataset or using a different machine learning algorithm than the random forest approach?
22. L. 532ff: In your current study, the statistically significantly differing areas are quite small. Do you recon that with longer simulation/forecasting times these areas might grow in size?
23. Figure A1: The map projection here seems to be skewed. It would be nice if you would align this with the Figures in the main part of the manuscript.