Response to the comments about the submitted paper

Modelling wind farm effects in HARMONIE-AROME (cycle 43.2.2) – part 2: Wind turbine database and application to Europe

We would like to thank the reviewers for their useful comments and suggestions. Our detailed answers follow.

We note that a new Figure 2 has been added in response to Reviewer 1's comments. As a result, the figure numbers in the text and our responses have been updated, and we have indicated this where appropriate.

The reviewers' comments are in *italics* while our answers are not. Additions to the original manuscript are indicated in blue.

Answers to Reviewer 1

Comment R1.1 ***Figure 1 caption:*** *Readers encounter the abbreviations NEA and DINI for the first time in this caption, and it would be helpful to provide their full names in advance. In addition, it is unclear whether the notation "Fischereit et al. (modified)" is a common convention; it might be preferable to use a format such as "Fischereit et al. (2022c')" or "modified–Fischereit et al. (2022c)."*
→ **R1.1** The explanation of the acronyms has been added to the caption. Due to the structure of the manuscript, first the description of the wind turbine database and then of the simulation, it is not possible to introduce the acronyms beforehand in the text, and we think that reusing the same figure to also show the domain extent is very useful. Furthermore, the abbreviation is explained now at the first location where it appears in the text (in the last paragraph of Section 2.1). We followed the reviewers' suggestion and changed the legend "Fischereit et al. (modified)" to "Modified Fischereit et al. (2022c)"

Comment R1.2 ***L125:*** *Since the EMODnet polygon is not familiar to many readers, it would be helpful to provide an intuitive map in the supplementary material so that readers can easily understand it without having to visit the website.*
→ **R1.2** We followed the suggestion of the reviewer and added a spatial depiction of the EMODnet data with a small description to Appendix A2.

Comment R1.3 ***Chapter 2:*** *Although building the wind turbine database is the second goal of this paper, it would be beneficial if the process were described more clearly and explicitly. The authors might consider adding a schematic flowchart to better organize and clarify the complexity of the procedure.*
→ **R1.3** We follow the suggestion of the reviewer and added a flow chart diagram as new Figure 2.

Comment R1.4 ***L258:*** *Since the Danish Meteorological Institute is later referred to as DMI (Figure 1 and L289), the abbreviation should be introduced when it is first mentioned.*
→ **R1.4** Thank you for noticing this. We have added the abbreviation, where it was first mentioned, i.e. in the introduction, as well as in the caption of Figure 1.

Comment R1.5 ***L260:*** *With a horizontal resolution of 2.5 km, it seems likely that multiple turbines could fall within a single grid cell. How such cases are handled in the WFP?*
→ **R1.5** If multiple turbines are present within a single grid cell, the effects of individual turbines on the velocity (and TKE) tendency are summed up. We have added it to the text: "Due to the horizontal grid spacing of 2.5 km, multiple turbines can be present within a single grid cell, depending on the distance between individual turbines within the farm. If multiple turbines are present, their individual effects on the velocity tendency, and, depending on the parameterization, the turbulent kinetic energy (TKE) tendency, are summed up. Some more details on the WFP implementations are given below.".

Comment R1.6 ***Figure 9 and Table 5:*** *It would be helpful to distinguish masts and lidars more clearly, as the current color strategy does not appear sufficiently clear. Adding a table line could also improve readability.*
→ **R1.6** We have implemented a different color strategy: masts are blue and lidars are red (now Figure 10). In addition, vertical lines have been added in Table 5 between masts and lidars.

Comment R1.7 ***L400:*** *Since the current wording may cause confusion, it might be clearer to write "1) FITCH, 2) EWP, and 3) OFF and NWF."*
→ **R1.7** We changed it according to your suggestion.

Comment R1.8 ***Figure 15:*** *The lead time on the x-axis should be made consistent with the ranges used in the text (1–11, 12–23, 24–35, and 36–47).*
→ **R1.8** Thank you for spotting this. We have updated the figure (now figure 16).

Comment R1.9 ***Figure 20 in text:*** *It is unclear whether Figure 20 is correctly referenced in lines 483–484.*
→ **R1.9** Thanks for spotting this. We forgot to update the reference when we switched from row numbering of subplots to numbering each individual subplot. The references have now been updated.

Answers to Reviewer 2

Comment R2.1 ***L. 77:*** *When presenting the structure of the paper it would help clarify that there are appendices and what topics they cover.*
→ **R2.1** We have added a respective section after presenting the structure of the main body of the paper.

Comment R2.2 ***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)*
→ **R2.2** The abbreviations have now been introduced in the caption and the first mention in the text.

Comment R2.3 ***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?*
→ **R2.3** For the offshore wind farm splitting algorithm you are referring to, we do see disagreement. We have added this to the text: This happens in a couple of wind farms (13 out of 62 offshore farms) in Finland, the UK, Ireland, the Netherlands, and Norway. However, relative to the total installed capacity associated with the 62 farms, this equates to less than $\tilde{5}$% (if we use the Wind Power data as a reference), thus, it affects mostly smaller farms either in terms of rated power per turbine or number of turbines. Due to the generic nature of the algorithm, this is an acceptable deviation.

Comment R2.4 ***L. 181:*** *Do you mean look-up-table instead of table look-up?*
→ **R2.4** Yes. Fixed

Comment R2.5 ***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, ...)*
→ **R2.5** Added ", described further in Sec. 2.2.7." where we describe the generic classes in more detail. It is a generic class for each dimension (Rotor diameter, power norm) and combines with the hub height in the database of characteristics.

Comment R2.6 ***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.*
→ **R2.6** When the technical database does not provide a clear number we can look up, or the spread between the minimum and maximum hub height is larger than 9 m (we don't do this averaging for other variables), we simply don't use any look-up value. Instead, we let the gap-filling method estimate a value for that variable.

Comment R2.7 ***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?*
→ **R2.7** No, we did not do that, but it's a good suggestion for future iterations of the dataset. We updated the discussion with this segment:

Finally, improvements are possible to fill the gaps in missing information on the rated power, rotor diameter, hub height, and year of commission. In the current version of the database, the sequential application of random forest regression models for gap-filling can lead to overfitting and physical inconsistencies. Because the sequential regression approach treats the imputation of each variable independently, it fails to fully respect their joint distributions. Furthermore, the regression models currently output continuous values—such as estimating a rated power of 9.75 MW—rather than snapping to the discrete, real-world specifications of existing turbine models, and the current workflow does not include post-imputation outlier detection and correction.

To mitigate these effects and ensure only realistic turbine models are generated, future iterations of the dataset will utilize a three-pronged gap-filling approach:

1. Maximized look-ups: We will first increase direct turbine model look-ups wherever possible by utilizing additional data resources.
2. Classification: We will treat the gap-filling primarily as a classification problem, estimating the exact turbine model based on the available known information.
3. Constrained Imputation: For the remaining truly ambiguous records, we will utilize Multiple Imputation by Chained Equations (MICE). This chained process iteratively estimates missing variables rather than treating them as independent problems, which better preserves their joint distributions. Crucially, the imputation will be constrained to the space of real turbine configurations, snapping each simulated record to the nearest feasible, physically consistent turbine in our dataset.

Additionally, the generic power and thrust curves assigned in the database currently assume a turbulence intensity (TI) of 0.1 for all locations following the default value in the PyWake function PyWake-Team (2022). This value is close to the TI value of 0.12 defined for the lowest turbine class in IEC 61400. Future versions of the database could further improve accuracy by varying the assumed TI between offshore and different onshore environments.

Comment R2.8 ***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. 9.75 MW or are the values corrected to the closest real rated Power found at any turbine??*
→ **R2.8** We don't correct the predictions from the random forests models except for physical consistency between variables. The model averages estimates across trees, resulting in a smoother distribution of predictions than the training data. See also the answer to Comment R2.7 for what was added to the discussion which relates to this.

Comment R2.9 ***L. 232:*** *I think the "(relatively)" should go before the "large" here*
→ **R2.9** Fixed.

Comment R2.10 ***L. 247:*** *I am not familiar with the PyWake package. Is "standard_power_ct_curve" a function defined in that framework?*
→ **R2.10** Yes, we have added a reference (PyWake-Team, 2022) to the manuscript at several places

Comment R2.11 ***L. 249:*** *Is the turbulence intensity you use for turbine specification varied between offshore and different onshore locations for the generation of power curves?*

→ **R2.11** We used the default TI of 0.1 for all the generic power curves, but you gave a valid suggestion for future versions. Please see the answer to Comment R2.7 for an update to the discussion relating to this.

Comment R2.12 ***Figure 7:*** *Here it says "(a) power curves", but should be "power coefficient"*
→ **R2.12** Fixed.

Comment R2.13 ***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.*
→ **R2.13** "3-hour-cycles" is commonly used in weather forecasts to express the method and time interval of producing initial conditions for later forecasts. We have added the following to the text to make it clearer: "This means that between 48-hour forecasts, a sequence of 3-hour forecasts is performed, where each forecast is initialized from the final hour of the preceding one. The purpose of this is to have more frequent data assimilation, making the NWP model stay close to the atmospheric state."

Comment R2.14 ***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.*
→ **R2.14** More information has been added, and the respective report has been added to the reference list. In fact, VAD-1660m and VAD-1500m measure wind in the opposite direction of Gode Wind 1, which explains the different profiles: The lidar datasets were downloaded from the data hub of Preliminary Investigation of Sites (PINTA portal) provided by the Bundesamt für Seeschifffahrt und Hydrographie (BSH) and were part of preliminary site investigations for the new wind farm development areas N-3.7 and N3.8 east and west of Gode Wind 1. Two different lidar types were used. Two Streamline XR lidar were placed on a wind turbine in the west of Gode Wind 1 (VAD-1500m) and on a transmission station in the east of Gode Wind 1 (VAD-1600m, VAD-3000m), respectively. A Velocity Azimuth Display (VAD) scanning pattern with several azimuth positions and different distances from the lidar position (1500 m towards west, 1600 m and 3000 m towards east) and defined elevation angles between 40 m and 200 m was performed to retrieve radial wind speeds and derive wind speeds and wind directions from those. In addition, a Windcube V2 lidar was placed on the transmission system, which measured the vertical wind profile in 80 m to 280 m height close to development area N-3.7 (N-03-07-UW_V2). More lidar measurements were performed, but were not analysed here, since they probed the atmosphere in similar areas. More details on the performed measurements can be found in Neumann et al. (2020).

Comment R2.15 ***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?*
→ **R2.15** The answer is more towards your second interpretation. But we see that this caused some confusion, so we rephrased the sentence as follows "For each measurement location, any measurements at timestamps missing at one height are removed from all heights and from the cor-

responding forecasts to ensure consistent wind profiles with the same timestamps in all heights."

Comment R2.16 ***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?*
→ **R2.16** This is almost what we did. We did not choose a threshold relative to all other data sources, but only within one wind profile measurement. We then calculated the maximum and minimum availability across all heights and removed heights with availability below 90 %. The reason why we defined the availability relative to other heights and not overall is the low overall availabilities of the lidar sites (Table 5). We tried to optimize data availability and profile consistency. We also rephrased this sentence: "If a measurement height has 10 % lower data availability than the others, it is removed entirely to maximize the overall mean data availability rate."

Comment R2.17 ***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.*
→ **R2.17** We have implemented the reviewer's suggestion.

Comment R2.18 ***L. 359:*** *In my mind it should be "at a height of 100m" instead of "at 100m height".*
→ **R2.18** We have changed this throughout the text.

Comment R2.19 ***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?*

- *The same applies for Figures 16-18.*

→ **R2.19** Yes, the colorbars are depicted correctly. However, we should have mentioned that the results are normalized by the no-wind-farm scenario (NWF). This has been added to the caption of now Fig. 14, 15 and 17. Figure 18 and 19 (new numbering) show non-normalized differences.

Comment R2.20 ***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.*
→ **R2.20** The focus here was to compare wind turbine-affected areas both over land and over water with a non-wind turbine-affected area. This was the reason why we chose points 1-4. We acknowledge the difference in the north-west of the domain between the simulation with 65 and with 90 levels. However, as we already discussed in the text, these are likely due to some changes in the structure functions used in the data assimilation when switching from 65 to 90 levels. They will likely disappear over longer simulation periods and are not expected to be physically significant. Since the focus is on the wind farm impact, which is expected to be close to the wind farms, we did not analyze other points here.

Comment R2.21 ***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*

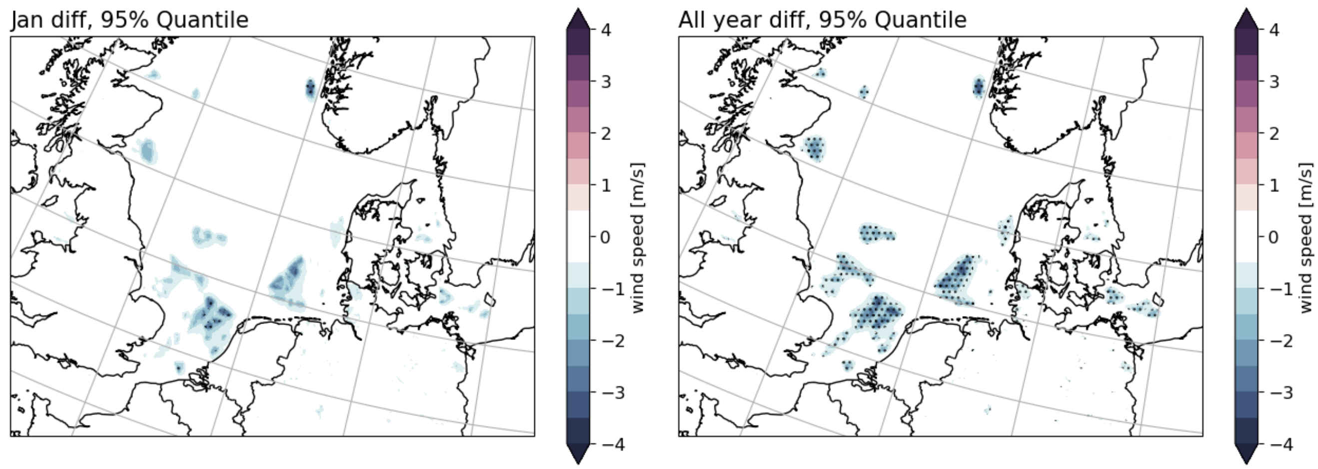

Figure R1: Differences in mean wind speed (m/s) at 150 m above ground and the significance of these differences (dotted where above the 95 % significance level) between two simulations with and without wind farms, as expected in a 2030 wind farms scenario. Left: analysis for a single month (January); Right: analysis for a whole year.

algorithm than the random forest approach?
→ **R2.21** See the answer to Comment R2.7, which covers this issue.

Comment R2.22 ***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?*
→ **R2.22** Yes, we anticipate that the areas will grow and become better defined as the simulation time increases. We have tested this using another NWP model, the Weather Research and Forecasting (WRF) model and the FITCH WFP, with longer-duration simulations. In the figure below, we show the difference in wind speed at 150 m between two simulations: one without wind farms and the other with wind farms, in a 2030 scenario of wind farm build-up. The differences in the left-side plot, which shows the difference for January, are as large as −3.25 m/s, but significance at the 95 % level is realized at only 306 grid points. For the same two simulations over a longer period of a full year, the differences in wind speed are similar, up to −3.42 m/s, but larger areas are significant at the 95 % level (5147 grid points). This is consistent with the argument that the signal-to-noise ratio increases with increased sample size.

We have added this to the text "Overall, the significant areas are quite small. However, if comparable results were found for forecasts longer than one month, the size of the statistically-significant areas would grow.".

Comment R2.23 ***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.*
→ **R2.23** Indeed, however, our approach was to show geographical maps in Plate-Carree, while showing HARMONIE results in the native rotated pole grid. Therefore, also Figure 1 and Figure 9 (now Figure 10) are using the Plate-Carree projection. However, we made a mistake in the projection of Figure 19 (now Figure 20), where we also used Plate-Carree.

References

Neumann, T., Cañadillas, B., Trujillo, J., and Frühmann, R.: Report on Meteorological Measurements, Tech. rep., URL https://pinta.bsh.de/N-3.7?lang=en/1000&tab=daten&typen=Bericht&themen=Wind, 2020.

PyWake-Team: PyWake: standard_power_ct_curve, URL https://gitlab.windenergy.dtu.dk/TOPFARM/PyWake/-/blob/cd5ff8363ae2615a92860d409e748b4a0431f33d/py_wake/utils/generic_power_ct_curves.py#L7, 2022.