Response to the reviewer

Dear Rogier Floors,

We thank you for your useful comments and the time invested in reviewing our manuscript. We have addressed each of your comments as detailed point by point below, which we believe has significantly improved the quality of the manuscript.

Main Comments

Reviewer Point P1 — The paper presents a case study of different models for the Perdigao campaign. It is a noble goal to quantify the accuracy of a model and the resources used to run that model, but unfortunately for me to trust the conclusions provided in the paper I lack important details about both the model setups and the way the 'costs' are calculated. I think the paper should stay clear from drawing too general conclusions about what model is most 'promising' and instead only present the very specific cases for which the models are validated (e.g. from mast to mast). So in it's current form I cannot recommend the paper for publication. I think major revisions are required to rethink the structure and/or allow other researchers to reproduce the results.

Reply: Regarding the point "I think the paper should stay clear from drawing too general conclusions about what model is most 'promising' and instead only present the very specific cases for which the models are validated", we have decided to restructure the paper accordingly. Indeed, we agree that the analysis of "costs" and "skill" is very difficult to generalise, and there were many assumptions made in this work. The new goal of the work is therefore now "to build a simulation database that will help to develop and improve WRA decision tools". We have adjusted the introduction and the analysis section accordingly. The analysis now avoids specific recommendations regarding the suitability of each model; rather, it now focuses on the difficulties and learning outcomes for WRA decision systems, which could be used in the future.

Reviewer Point P2 — I am afraid that with so many models it will be very hard to describe all of the setups, without making the paper 50 pages long. Perhaps a possibility is to put detailed model setups in the appendix. Alternatively, the text should be adapted so that the model setups are provided with enough level of detail to redo the simulations (see detailed comments below).

Reply: It is indeed very difficult to find a balance between describing a simulation setup in a detailed enough manner without stretching a paper to 50+ pages. Especially with the amount of different tools and work flows, as presented within this work. We have added more details and/or links to further resources describing the setups and hope this is sufficient to reproduce the simulations.

Reviewer Point P3 — I also think that In a study like this where no new theory is being presented (which is fine) it is particularly important that the data are openly available so that others can still benefit from the study. So I would suggest expanding the "Data availability." section, with more than just the repository of all Perdigao data.

Reply: As there was no further data preparation and filtering done on the Perdigao data itself, we simply referred to the repository, which provides all the necessary data used within this work.

However, in order to be able to compare simulation results, we made results of all simulations available on ZENODO `https://doi.org/10.5281/zenodo.8005621`.

Further Comments

Reviewer Point P4 — l16: upper heights sounds a bit strange. I suggest higher heights or something similar.

Reply: Has been changed accordingly.

Reviewer Point P5 — l26: long-term wind resource extrapolation: I would call this long-term wind resource correction, the way it is written it seems like the long term wind resource needs extrapolation, but is the shorter term measurements that need to be extrapolated to a longer term climate.

Reply: Agreed, wrong term was used here. Has been changed accordingly.

Reviewer Point P6 — l94: Will this turbine cause any wake effects? This is not discussed.

Reply: This has not been considered due to the very low frequency of wind speeds from this direction. Info added to the paper (l. 95).

Reviewer Point P7 — l149: Are all model outputting time series? Is this generic or which time series are you talking about here?

Reply: Here we talk about the measurement data only. This has been made more clear in the text on line 150.

Reviewer Point P8 — l160: topography -¿ orography (topography is usually defined to include the roughness of the terrain)

Reply: Thank you for the clarification. Has been changed accordingly.

Reviewer Point P9 — Table 1: Common application range: I would rather call this complex or non-complex instead of flat or non-flat. The linearized model will probably work fine in non-flat terrain as long as no flow separation occurs.

Reply: As the term "complex terrain" can be interpreted very differently depending on the used simulation tool, we tried to refrain from using it, see https://doi.org/10.5194/wes-7-2231-2022 Instead we mentioned flow separation as a limiting factor. This has been added to line 179.

Reviewer Point P10 — l177: Who is 'our' here?

Reply: Has been changed to "UTD in-house code"

Reviewer Point P11 — l187: terrain topography - orography

Reply: Has been corrected.

Reviewer Point P12 — l190: Corine Land Cover (European Union, 2018) database: Did you use raster or vector data? Which projection was used? Which datum?

Reply: Vector data of the Corine Land Cover (CLC) 2018, Version 2020_20u1 (`https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=metadata`) was used to obtain vegetation data about the area. See line 214.

Reviewer Point P13 — l190: "The stem size and distribution reproduce the same canopy frontal solidity (Monti et al., 2019; Nepf, 2012) of the actual vegetation at the site extracted from the Corine Land Cover (European Union, 2018) database". This is not clear: what is the actual vegetation at the site? How can you get that from CORINE data which is just a satellite based product? How can that match the stem size and distribution?

Reply: Additional information has been added to the section.

Reviewer Point P14 — PCE-LES: which source did you use for the terrain elevation?

Reply: Additional information has been added to the section.

Reviewer Point P15 — Section 2.3: after reading this I was expecting all modellers use the same terrain elevation, but based on l198 I start to doubt that because there SRTM is mentioned. In the LES section no source is mentioned.

Reply: The choice of how and what to use for modelling the terrain elevation was left to the modeller. As this is highly dependent on the used tools. This has been clarified in the text, see line 172.

Reviewer Point P16 — l215: Scaled how? You mean assuming the wind distribution for the three months is representative for the whole year?

Reply: This is correct. A remark has been added for clarification.

Reviewer Point P17 — l218: That reference does not really describe the WAsP stability model. Better to cite Troen and Petersen (1989). What stability setting are used in the end?

Reply: Thank you for the reference. This has been changed. WAsP wind modelling variables were set by default according to the atmospheric stability effects obtained.

Reviewer Point P18 — l219: What was the source of this tiff data?

Reply: Information has been added to the text.

Reviewer Point P19 — l220: What is in the .map file? What was the source?

Reply: Information has been added to the text.

Reviewer Point P20 — l221: There is several roughness tables in that reference? Which one was used?

Reply: Information has been added to the text.

Reviewer Point P21 — l223: The direction variable: you mean wind direction? From which height? What is NCA?

Reply: Sentence has been changed for clarification. Reference was fixed as well. Meant was the Earth Observing Laboratory.

Reviewer Point P22 — l225: Which 'data'?

Reply: Information has been added to the text.

Reviewer Point P23 — l225: References for MERRA and ERA5 missing

Reply: Information has been added to the text.

Reviewer Point P24 — l226: Coefficient of determination between what and what?

Reply: Information has been added to the text.

Reviewer Point P25 — l226: How do you define "very similar"?

Reply: Information has been added to the text.

Reviewer Point P26 — l226: "some basic filtering", "constant line values", "some variables": specify what filtering, what is a constant line value, which variables?

Reply: This has been made more clear in the text.

Reviewer Point P27 — l229: Section on stability: this can also be made quantitative.

Reply: Information has been added to the text.

Reviewer Point P28 — l229: What kind of adjustment?

Reply: Information has been added to the text.

Reviewer Point P29 — l230: What was optimized with respect to what?

Reply: Information has been added to the text.

Reviewer Point P30 — l232: Please specify in more detail what kind of long-term correction you did.

Reply: Information has been added to the text.

Reviewer Point P31 — l246: what was the upper domain boundary?

Reply: We specified the height of the E-Wind domain (6 km) in the manuscript. Furthermore, we added some details on the used boundary conditions.

Reviewer Point P32 — l246: Is the first bin centered around north or from 0 to 15 degrees?

Reply: In the RANS simulation, we do not consider sectors, but discrete wind directions. The first geostrophic wind direction is 0°. We added this information to the manuscript.

Reviewer Point P33 — l247: How do you define the wind shear?

Reply: The definition was added to the manuscript.

Reviewer Point P34 — l255: Similar as my previous comment: so then you assume the 3 month wind measurements are representative of the full year? That is fine, but it is inconsistent with the previous model setup (Windpro), where you apply a long term correction.

Reply: For the wind speeds no long term corrections were performed. Only the energy production values that were calculated for three months was scaled to obtain the annual energy production.

Reviewer Point P35 — l261: roughness height -¿ roughness length

Reply: Thanks for spotting. Has been corrected.

Reviewer Point P36 — l265: Is the roughness length varying with wind direction sector?

Reply: This has been made more clear by adding "for all wind directions".

Reviewer Point P37 — l267: grid independence study -¿ I assume the conclusion of the grid independence study was that the simulations were not dependent on resolution? Why was the resolution of 15 m optimal? In which sense was it optimal?

Reply: The sentence "For resolutions below 15 m no change in wind speed profiles compared to the 15 m resolution was observed." has been added

Reviewer Point P38 — l285-l289: Taking the mean of a RMSE is mixing different errors metrics. You should calculate the squared errors from each sector and do the root-mean in the last step?

Reply: The sentence has been changed for clarification. What is meant here is that the RMSE value for each simulation was taken and then averaged to make a relative comparison between the simulations. The calculated RMSE value for each simulation, however, was calculated as you indicated.

Reviewer Point P39 — l292: m/s should be in normal font

Reply: Thanks for spotting. Has been corrected.

Reviewer Point P40 — l302: it would be good to mention here that this is the calibration point.

Reply: Sentence has been added.

Reviewer Point P41 — l318: This is not that surprising because both masts are located on top of a hill. It would be useful to relate this to the "most similar predictor" discussion in https://wes.copernicus.org/articles/5/1679/2020/.

Reply: Thank you very much for this reference. However, as the performance differences, i.e. either better or worse, between met mast 25 and met mast 20 vary quite a lot for the various simulations we would refrain from relating this to the "most similar predictor".

Reviewer Point P42 — Fig 4: I am bit confused how big the errors are at 80 m. Wasn't that used for calibration? How can there be already RMSE of up to 0.4 m/s?

Reply: For the OpenFOAM and the Fluent simulations the calibration was done at a height of 100 m. E-Wind did the calibration based on wind shear, whereas for the LES simulations only 8 wind directions were simulated and then interpolated to obtain the desired results for 12 sectors. WAsP performs internal corrections, especially in complex terrain, which can slightly vary the results. These information, where missing or wrong, have been added to the respective model descriptions.

Reviewer Point P43 — Sect 3.1.2: This section is hard to understand ; what is the main message? The relative costs and skill score appear suddenly in Fig 7, but it feels like some background on the numbers should be available (appendix?). As discussed the costs and skill are extremely hard to quantify, so you could end up with any ranking of the models here. I would avoid drawing conclusions like "Taking the cost scores into account, E-Wind is the most suitable tool for the Perdigão site.".

Reply: The methodology for this was developed and presented in Barber et al. (2022b), whereas we applied this methodology to our case study and simply presented the results. Based on and in the context of this methodology the terms 'cost', 'skill', 'most promising', 'most suitable tool', etc. was used. Having said that, in the processing of restructuring the paper and changing the goal, as mentioned above, this section has now been removed.

Reviewer Point P44 — l370: mast 19? You mean 29?

Reply: Thanks for spotting. Has been corrected.

Reviewer Point P45 — l380: end of line: height.

Reply: Thanks for spotting. Has been corrected.

Reviewer Point P46 — Fig 8: Are we comparing AEP at the same heights here? That should be added somewhere.

Reply: The AEP was calculated and compared at height 80 m. In the section "Skill and cost scores" we changed and clarified some text about the AEP calculation, now also stating that it is determined at a height of 80 m.

Reviewer Point P47 — l386: What is an AEP by sector? I only know about an AEP as the production for a year, i.e. for all sectors combined.

Reply: This is simply supposed to show how important a simulated wind direction is in terms of the contribution to the overall AEP. For some wind directions the wind profiles and the AEP values were underestimated and for some others overestimated, which in some cases can cancel each other out. So

by just looking at the overall AEP one might get the impression that the prediction of the overall AEP is fine, however, is in fact a result of a combination of under- and overestimations for different wind directions. A sentence has been added to make the purpose of the plot more clear, see line 431.

Reviewer Point P48 — l423: It is for me again not quite clear how this is quantified. I would leave out generalizations like this and just discuss model differences. How does a single AEP prediction from one mast to the other make this the model the best for entire Perdigao site?

Reply: We agree that this statement is too bold and quite an extrapolation. Sentence has been deleted.

Reviewer Point P49 — Fig 11: Where is mast 20?

Reply: Unfortunately, the simulation results for the AEP at met mast 20 are not available. This is due to the original design of this case study, with the original goal to only compare met masts 25 and 29 for the AEP predictions. This information has been addded to the text, see line 394.

Reviewer Point P50 — Sect 3.3: I agree there is so many differences in the difference model chains to calculate AEP that it is impossible to say what it is the exact reason. If mast 29 is used for calibration one would not expect any model error in AEP? So I would just leave this section out.

Reply: This section has been reformulated to be more clear and precise. Also see the next point for further details.

Reviewer Point P51 — l454: It would be a very surprising conclusions if the AEP error did not depend on wind speed error. What about air density? How has that been calculated in the different model chains?

Reply: The statement of this section is that the overall wind speed prediction accuracy is not the only variable that influences the overall AEP prediction accuracy, but that looking at sector values, especially for the most frequent wind direction, is important. This is also the reason for introducing the "AEP by sector" plots and how they contribute to the overall AEP.