Response to reviewers

Ida M. Solbrekke

August 2020

1 General comment to reviewer # 1

We would like to thank reviewer #1 for his comment on the manuscript with the working title of "Norwegian hindcast archive (NORA3) - A validation of offshore wind resources in the North Sea and Norwegian Sea". Please see the response below.

1.1 Response to comments from reviewer #1

RC1: The authors do not convincingly show what is novel and/or more advanced in their model compared to others. E.g., from Table 1 I cannot see what is new/better compared to NEWA. The authors discuss that NEWA is not validated offshore; however, this is different from having something in the model which is novel/advanced compared to other models. If the authors could show that NORA3 fits measured data significantly better than competing models/datasets, it could be considered an advancement; however, this is not done.

AR1: Thank you for pointing this out. The novelty of NORA3 over other similar data sets is that NORA3 is created by another numerical weather prediction model (in this case the HARMONIE-AROME, Cy 40h1.2, which have spectral dynamics which is very different from the WRF model which solves the momentum equation on a staggered grid). Currently all available wind resource data sets are created by the Weather Reasearch and Forecasting model (WRF). The creation of NORA3 by a different NWP model will contribute to a diversity in the available wind resource data sets. When a multi-model ensemble of these data sets are considered for wind power planning the wind resource uncertainty can be quantified and reduced in regions like offshore which often have very few and short time in situ observational references. We know from inter-comparison of numerical weather prediction models, that there is no such ting as a overall "best" model. Quality depends strongly on which spatial and temporal scales are of interest, geography, the selected measure of quality and the application the data will be used for. Thus, the usefulness of multi model ensembles has become increasingly clear over the last few decades in research field such as weather prediction and climate change. We also see this slowly trickling down into wind resource research community. In addition, this study contains novel validation measures not seen in the assessment of other wind resource data sets

covering the North Sea, Norwegian Sea and Barents Sea. Finally, we tend to agree that papers focusing on verification of new simulations seldom are extremely novel. However, they fill a very important role by documenting quality. This is a prerequisite to ensure a sound and critical use of the simulations by stakeholders and fellow researchers.

As several new wind resource data set are becoming available a model comparison is indeed of great interest and something we will pursue in future work. However, given the page limitations and the fact that this is the first paper to evaluate the wind resource estimates from this simulation set we have prioritized to put the focus on a detailed comparison against observations instead of introducing a second objective to the paper.

AC1: Chapter 3 is now added to the paper, and encompasses the comparison of NORA3 and ERA-5 towards observations and the result of the downscaling process.

RC2: The statements such as "...NORA3 data is rather well suited...", "...slightly conservative..." and "...slightly underestimated..." are very vague. What are they based on? E.g., "The model is relatively good...": relative to what?

AR2: Thanks for pointing this out. We believe that you found these statements in the Abstract where we write "NORA3 data is rather well suited for wind power estimates, but gives slightly conservative estimates on the offshore wind metrics". The next sentences in the abstract explain what we mean by these phrases: "Wind speeds are typically 5 % (0.5 m/s) lower than observed wind speeds, giving an underestimation of offshore wind power of 10 %-20 % (equivalent to an underestimation of 3 percentage point in the capacity factor),for a selected turbine type and hub height. The model is biased towards lower wind power estimates because of overestimation of the frequency of low-speed wind events ($< 10m/s$) and underestimation of high-speed wind events ($> 10m/s$). The hourly wind speed and wind power variability are slightly underestimated in NORA3. However, the number of hours with zero power production (around 12 % of the time) is fairly well captured, while the duration of each of these events is slightly overestimated,leading to 25-year return values for zero-power duration being too high for four of the six sites. The model is relatively good at capturing spatial co-variability in hourly wind power production among the sites. However, the observed de-correlation length was estimated to be 426 km, whereas the model-based length was 16 % longer".

AC2: The "abstract" and "summary" have been rewritten to clarify these vague statements.

RC3: The biggest issue I have with the presented results is that they are not compared to other models. I thus consider it impossible to judge whether the model results are good or bad (see also the previous comment related to this). As NEWA mesoscale data are available (https://map.neweuropeanwindatlas.eu/), they should be compared to the NORA3 model results. I also consider comparison to ERA5 to be required, as the downscaled data should be shown to outperform the data they are based on.

AR3: Thanks for your comment. Please see AC1. I understand that Table

1 and section 2.1.1 creates an expectations towards a comparison of NORA3 and the other data sets listed in the table. Therefor, Table 1 and section 2.1.1 will be removed from the revised manuscript to make the intention of this study more clear, namely a thorough validation of NORA3 towards offshore observations.

AC3 The "abstract", "introduction" and "summary" have been rewritten to clarify the intention and the novelty of the study. Please see the revised manuscript.

RC4: As a large portion of the paper is comparing wind speed distributions, I suggest the authors to compare the distribution results also to microscale resolution data, which are available both from NEWA and GWA (https://globalwindatlas.info/).

AC4: Thank you for the suggestion of comparing NORA3 with microscale resolution data. This is of great interest since comparing NORA3 with microscale resolution data will determine how well NORA3 is capturing meso- and small scale features. However, as mention in AC1 we have prioritized to focus on a detailed comparison against observations instead of introducing several new aspects to the paper. Hence, comparing to microscale resolution data will be further work.

RC5: The authors write: "... underestimation of 3 percentage point in the capacity factor". This would seem to me a large error, e.g., when considering profitability of an offshore wind power plant, and does not seem to be in line with the other statements of the model being quite good (see the comment 2) above). Please comment.

AR5: This is an interesting point raised by the reviewer and we tend to agree. It highlights how rather small wind errors translates into sizable wind power errors. Since a cost-function like the levelized cost of energy(LCOE) is inverse proportional to the wind power produced during the lifetime of the wind farm, using NORA3 as wind power production estimator will result in a wind farm profitability being on average 6% higher than first expected. Given that the required rate of return when planning offshore wind projects typically is 5-10 %, a deficiency of 3 percentage point in the CF is an sizable error. This highlight the need for building up archives of simulations like the NORA3, to be able to conduct informed uncertainty calculations for the production in regions where observations are not existing or are very sparse.

However, a comparison between NORA3 and ERA5 in terms of the wind power capacity factor (CF) shows that the ERA5-based CFs are on average 5 percentage point (approximately 10% difference in CF) lower than the observational based CFs. Hence, the improvements using NORA3 over ERA5 gives more realistic wind power profitability measures.

AC5 The second last paragraph in chapter 3 (line 207-213) is added to the paper and encloses the wind power profitability.

RC6: I do not understand what this sentence means: "The corresponding wind speed variability is given by the Weibull standard deviation (std). The Weibull std is used instead of the Gaussian std because of the shape of the wind speed distribution (see Fig. 2a for an example wind speed distribution)." Standard deviation (or variance) gives the 2nd moment information of any (well-defined continuous) distribution (https://en.wikipedia.org/wiki/Standard_deviation);

it is not related to the shape of the distribution. I find the sentence very confusing, please clarify and give references to what a "Weibull std" is and why it should be used.

AR6: We are sorry for the confusion regarding the choice of standard deviation (std). What was meant was that the std was not calculated directly from the observations, but by first fitting a Weibull distribution to the data before the std was calculated. The idea was that this would be somewhat more robust then calculating it directly because the extreme tail of the empirical distribution was reduced with the Weibull fit and therefor made a smaller impact on the std.

AC6 We have now removed Weibull std and are instead using a traditional std calculated directly from the empirical data.

RC7: If you consider standard deviation to be an inadequate measure, why not use higher moment measures (such as kurtosis, skewness) and/or compare the entire PDFs/CDFs of the distributions?

AR7: Figure 2 is showing the wind distribution and differences in the simulated distribution, compared to the observed. In the discussion of the figures we discuss the quality of the higher moments in the simulation.

RC8: Why only 2004 to 2018 are simulated? ERA5 data are available for tens of years; why not downscale more years? Would not using longer time series provide a better estimate especially of events occurring very rarely? (the authors focus on rare events in specific sections)

AR8: Thanks for referring to the somewhat short model period of NORA3 compared to ERA5. We have only validated 2004-2016 (study period changed after exclusion of FINO1 from 2010-2018) due to the currently available data from NORA3. NORA3 is continuously being generated, and at the time of writing it covers 1994-2020. When the model integration is finalized (in autumn 2021) the NORA3 data will cover the time period from 1979 to present, and will be regularly updated in the coming years when ERA5 data becomes available.

RC9: The authors write: "Hence, the observed wind speed is somewhat more intermittent and variable than the modeled wind speed, indicating that HARMONIEAROME is missing some of the high-frequency variability embedded in the wind field." Why do you say that especially high-frequency variability is missing? Variance (and thus standard deviation) is a sum of all spectral components (https://en.wikipedia.org/wiki/Spectral_density) and thus a mismatch in standard deviations could be caused by mismatch in the spectra at any frequencies.

AR9: Thank you so much for noticing this badly explained phrase. I agree that standard deviation and variance encloses all temporal scales. During the generation process of the paper we performed a spectral filtering of the observational and modeled data using a Butterworth filter. This result was not implemented as a part of the paper. Performing this temporal-filtering illustrated that NORA3 was missing out some high-frequency time variability.

AC9: I have rewritten the sentence in the following way: "Hence, the observed wind speed is somewhat more intermittent and variable than the modeled wind speed, indicating that HARMONIE-AROME is missing some of the variability embedded in the wind field."

RC10: I find it difficult to understand what exactly is presented and done in sections 2.6, 4.4 and 4.5. I understand that for the estimation of extreme (very high) wind speeds, extreme value theory may be needed to extrapolate beyond the limited time range of data (however, please note comment 8) above related to this). But I find it very confusing that two very different causes for "zero wind power events" are mixed: one being quite frequent (wind speed below cut-in speed), which does not seem to need extreme value theory for estimating the likelihood; and one being relatively rare (wind speed above cut-out speed). But even the latter should happen offshore many times during a multi-year dataset. Why does it need extreme value theory? I would consider extreme value theory to be needed to estimate values like 50-year maximum wind speed, which may be needed for the design of the turbine (but that is very different from wind speeds causing storm shutdowns).

AR10: Thanks for your comment of the extreme value analysis on zero power events. I agree that performing an extreme value analysis on the number of zero events will not work since the the occurrence of these events are not a rare event. However, we have performed an extreme value analyses on the maximum duration of a zero-event and not the occurrence of the events. We calculate the maximum duration (in hours) of a zero event expected to happen at least once during the lifetime of a wind turbine (25 years) due to either too low ($u < u_ci$) or too high ($u \geq u_{co}$) wind speeds. Knowing about the expected extreme duration of zero-event is useful information for planning the amount of regulating power needed in a grid to ensure the needed electricity supply at all times, as well as for planned turbine maintenance and for the profitability of the wind park. Therefor, we have validated NORA3's ability to reproduce the observational based maximum duration of a zero power event.

RC11: I find sections 4.4. and 4.5 confusing. a) Why do you apply extreme value theory? (see the previous comment); b) How is the extreme value theory exactly used to give you estimate of how long the "zero-events" last? Why does this matter, e.g., for LCOE? (I think LCOE is impacted by the entire generation distribution, not just the "zero-event" likelihood). I do not understand what the last paragraph on 4.5 is saying, consider a thorough rewrite.

AR11: Thank you for this comment. Please see AC10. I agree that the LCOE is determined from the entire generation distribution. However, the generation distribution also includes the part of the distribution caused by either too low or too high wind to produce wind power, hence the zero power generation. Since, the overarching goal of an operating wind farm is to produce as much as possible, it is crucial to know about how frequent a zero power event is, and also the maximum expected length of these events during the turbine lifetime due to scheduled maintenance and as information on how reliable the production from the wind farm is when large amounts of wind energy is planned to be integrated into the grid. The extreme value theory is conducted in the following way:

- Firstly, the Kolmogorov-Smirnov p-test is performed to test the null-hypothesis. The null-hypothesis states: the data is **not** drawn from the

chosen data distribution (GEV or Pareto). Testing the null-hypothesis is done by calculating the distance between the empirical (either "block maxima" (BM) or "peak over threshold" (POT)) and theoretical (extreme value (GEV) and generalized Pareto distribution (Pareto)) cumulative distribution functions.

- Secondly, since the result from the Kolmogorov-Smirnov test tell us that we cannot exclude the possibility that the data is drawn from either of the two data distributions we fit the observational based and model based extreme zero-event durations to GEV and Pareto and find the corresponding 25-year return values for the following five sites: Ekofisk, Sleipner, Gullfaks C, Draugen and Heidrun. Fino1 is excluded from the extreme value analysis due to the shorter time series (2004-2009).

AC11 I have added some additional and explanatory text to the relevant sections regarding the extreme value analysis (sections 2.6, between line 170 and 175 and between line 178 and 181).

2 General comment to reviewer # 2

We would like to express our gratitude to referee #2 for his/her time and effort revising the manuscript with the working title of "Norwegian hindcast archive (NORA3) - A validation of offshore wind resources in the North Sea and Norwegian Sea". With the comments from referee 2 we strongly believe that the manuscript now is better and more precise. Please see our response to the comments below (new text added to the paper is in italic font).

3 Respons to major comments from reviewer # 2

RC1 FINO1 results: The FINO1 results differ strongly from the other sites. A quick online search could have given the authors the answer why this is the case. Since 2010 the wind farm alpha ventus started operating only about 400m east of the mast. Since 2015 the mast is located in the center of a large wind farm cluster. This change in time is by the way nicely reflected in Figure 6c. The deviations discussed in the manuscript thus mainly originate from the fact that FINO1 is in most of the years investigated here not measuring free wind conditions but winds that are strongly impacted by the turbines surrounding the mast.

AR1: Thank you very much for pointing out the close proximity of Alpha Ventus wind farm to the FINO1 met mast. The commissioning of Alpha Ventus explains why the observation-based capacity factor (CF) strongly deviates form the modeled CF after 2009 (Alpha ventus was operational from August 2009), and generally why the validation results for FINO1 are odd.

AC1: To cope with this we now only use FINO1 data between 2004 and 2009 (until August) in the validation of NORA3 for FINO1. Figures changed due to the shorter time series of FINO1 are: Fig 2b, Fig 6-10. Also Figures in appendix are influenced: Fig A1 and Fig B1. In addition, the following tables have changes: Tables 2-4, and Tables 6-8. The discussion of the affected figures and tables are rewritten accordingly.

RC2: Comparison to other data sets: In the introduction several other data sets are discussed but no comparison is made, not even to the ERA5 data. Thus, it is not even evaluated if the NORA3 is at all an improvement over ERA5. Thus, the study in my opinion does not show that NORA3 really is a suitable data set for wind energy applications in comparison to all the other data products that are available.

AR2: Thanks for addressing this relevant and important issue regarding comparison of NORA3 against the host data set (ERA5) and other wind resource data sets. NORA3 are extensively validated towards observations and compared against the host data set in the paper of [3]. Nevertheless, we have now performed a validation of NORA3 and ERA5 towards observations. As several new wind resource data set are becoming available a model comparison is indeed of great interest and something we will pursue in future work. However, given the page limitations and the fact that this is the first paper to evaluate the wind resource estimates from this data set we have prioritized to put the focus on a detailed comparison against observations instead of introducing a second objective to the paper.

AC2: Chapter 3 is now added to the paper, and encompasses the comparison of NORA3 and ERA-5 towards observations and the result of the downscaling process.

RC3: Novelty of NORA3: After reading the whole manuscript, I have not understood what the novelty and improvement of NORA3 data are. The resolution in space is similar to NEWA, the resolution in time of 1 hour is poor. The wind industry would benefit from 30 min or even 10min resolutions. Doesn't this would also have an impact on the extremes discussed in the study? A mesoscale model does provide sub-hourly fluctuations and could improve the ramping. Is the novelty just the fact that NORA3 will be available for the time period from 1979 and further continued?

AR3: Thank you for pointing this out. The novelty of NORA3 over other similar data sets is that NORA3 is created by another numerical weather prediction model (HARMONIE-AROME, Cy 40h1.2, which in contrast to WRF has spectral dynamics and very different physical schemes (turbulence, radiation etc.)). All other wind resource data sets are created by the Weather Reasearch and Forecasting model (WRF). The creation of NORA3 by a different NWP model will contribute to a diversity in the available wind resource data sets.When a multi-model ensemble of these data sets are considered for wind power planning the wind resource uncertainty can be quantified and reduced in regions like offshore which often have very few and short time in situ observational references. We know from inter-comparison of numerical weather prediction models, that there is no such ting as a overall "best" model. Quality

depends strongly on which spatial and temporal scales are of interest, geography, the selected measure of quality and the application the data will be used for. Thus, the usefulness of multi model ensembles has become increasingly clear over the last few decades in research field such as weather prediction and climate change. We also see this slowly trickling down into wind resource research community. In addition, this study contains novel validation measures not seen in the assessment of other wind resource data sets covering the North Sea, Norwegian Sea and Barents Sea. Finally, we tend to agree that papers focusing on verification of new simulations seldom are extremely novel. However, they fill a very important role by documenting quality. This is a prerequisite to ensure a sound and critical use of the simulations by stakeholders and fellow researchers.

AC3: The introduction, abstract and summary is rewritten to clarify the scope and the novelty of the study.

RC4: Wind Speed Interpolation: The wind speed interpolation is done using a power law relation. This is an empirical method with deficiencies in unstable conditions. I suggest to use a logarithmic interpolation here.

AR4: Thanks for your comment. Both logarithmic wind profile (log law) and power law wind profile (power law) assumes neutral atmospheric stability of the atmosphere and will have deficiencies in stable and unstable conditions. [2] reviewed different extrapolation methods (logarithmic models, Deaves and Harris, and power law) for 96 different locations worldwide. He concluded that the power law was the most reliable and widely most used extrapolation method. In addition, according [5] the usage of log-law is most suited near the surface.

AC4: Despite the aforementioned results from Gualitieri and Sill we have compared the performance of the log law and the power law (with time varying power exponent) for the six offshore sites used in this study. For the majority of the sites, interpolating NORA3 wind speed to the sensor height using the two different methods (log-law and power law) we see that the model bias using log law is larger than using the power law method for the majority of the sites. Therefor, we stick to the power law method with a time dependent power exponent when wind speed interpolating is conducted.

We add a clarifying text to section 2.3 (line 125 to 131).

4 Response to minor comments from reviewer # 2

RC1: Line 80: 31 x 31 km → I think you mean 3x3km here.

AR1/AC1:: Thanks for noticing the error, and you are totally right. I have now changed it to 3x3km.

RC2: Line 83: ... and research by many European countries ... → I think European weather services and research institutes is better suited here.

AC2: Done.

RC3: Line 108: ... are hourly wind observations ... → I think it is better

to write something like "10 min average values provided at every hour" directly in the text than with the footnote

AC3: Done.

RC4: Line 139: ... the turbine park ... → "wind farm" fits much better here.

AC4: Done.

RC5: Line 181: ... exploiting NORA3 as a planning tool ... → How is NORA3 a planning tool. Isn't this just a dataset and a wind farm planning tool could make use of these data?

AR5/AC5: Thanks for commenting on this badly formulated sentence. You are complete right. I have changed it to the following: "Prior to exploiting NORA3 as a wind resource data set in the planning-phase of future offshore wind power installations the model performance has to be validated and verified against observational data"

RC6: Table 3: Weibull parameters are much more commonly abbreivated by lambda and k than b and a

AR6/AC6: a and b are changed to λ and k.

RC7: Line 230: ... wind park configuration and internal position... → better: wind farm layout

AC7: Done.

RC8: Line 242: But we cannot rule out... → better: "conclude"

AR8/AC8: I changed the sentence in line ... to the following: "But, we cannot exclude the possibility that the platform design at Sleipner affects the flow field more than the design of the other platforms." (line 279)

RC9: Section 3.4: I would rather put this section as 2.2.1

AC9: Thank you for your suggestion. As section 3.4 is a part of the results and not a method-section we would like to keep the section where it is placed today.

RC10: Line 283: ... an hour-to-hour variability typically 7%-9% (Table 7) of the installed capacity → ... variability OF typically ...

AC10: Done

RC11: Line 377-378: The general picture is that the NORA3 data is rather well suited for wind power estimates in the absence of in situ data → afterwards all deficiencies are discussed. This conclusion does not fit to the sentences before.

AR11/AC11: Thanks for pointing this out. I have now rewritten the sentence in line 414 the following way: "Nevertheless, there is a tendency towards the model generating slightly conservative estimates, and the results are summarized below."

RC12: References: Please add availability for ALL publications that are not books

AC12: Unfortunately, we do not understand what the reviewer mean by RC12.

5 General comment to reviewer # 3

We would like to express our gratitude to referee #3 for his/hers time and effort revising the manuscript with the working title of "Norwegian hindcast archive (NORA3) - A validation of offshore wind resources in the North Sea and Norwegian Sea". With the comments from referee 3 we strongly believe that the manuscript now is better and more precise. Please see our response to the comments below (new text added to the paper is in italic font).

5.1 Respons to major comments from reviewer #3

RC1: It is not clear that why the authors use NORA3 for their study. They gave Table1 which shows the related datasets and also,they emphasized that the NEWA dataalso prepared similar to NORA3. But they did not make neither qualitive nor quantitive comparison between these two datasets and explained whythey used or what is the superiority of NORA3. So, it is not clear why the authors prefer NORA3 instead of NEWA.

AR1: We had hoped that it was clear from the paper that we use NORA3 because it is a newly produced wind data set, that has not been validated for wind resource and wind power estimations. Thus the objective of the paper is an evaluation of this new simulation. This is a prerequisite to ensure a sound and critical use of the simulations by stakeholders and fellow researchers.

NORA3 compares to e.g. NEWA in terms of model configuration like spatial and temporal resolution, initial- and boundary information. We would like to emphasize that we do not "prefer NORA3 instead of NEWA". As several new wind resource data set are becoming available a model comparison is indeed of great interest and something we will pursue in future work. However, given the page limitations and the fact that this is the first paper to evaluate the wind resource estimates from NORA3 we have prioritized to put the focus on a detailed comparison against observations instead of introducing a second objective to the paper.

AR1: Chapter 3 is now added to the paper, and encompasses the comparison of NORA3 and ERA-5 towards observations and the result of the downscaling process.

RC2: I believe that the length of the simulation does not provide enough information about the climatology of the region to select the best offshore wind farm areas. In my opinion, if the authors want to define the best offshore wind production areas, they should keep the simulation period as long as possible to include extreme atmospheric conditions and inter annual variability. Moreover, If I would have preferred NORA3 instead of longer simulation and finer temporal resolution NEWA, I would rather simulate the model for fine horizontal simulation such as 1km. Thus, it might help to solve the mesoscale circulations such as sea-land breezes at near-coastal wind farm that can affect the power productions.

AR2: Thank you for pointing out the length of the model time series. I agree - using the data set for mapping the area most suited for wind power

production the times series should be as long as possible. The NORA3 data set is continuously being generated and will when finalized cover the time period from 1979 and onward. Hence, the possibility to download and analyze a much longer NORA3 time series will be possible in the near future.

RC3: It is not clear to me that how the model downscaled 31km ERA5 data to 3km HARMONIE-AROME domain. According to the authors explanations, I think ERA5 is directly downscaled to 3km that I have concern about this sharp interpolation affects the performance of the model

AR3: There is no clear recommendation from the literature on the best strategy for downscaling, regarding nesting down gradually or direct leap to a finer resolution. Recent papers indicate that there is little quality difference between the two approaches, but a tendency for better results using no nesting [4, 1].

RC4: Wind speed interpolation should be conducted by using log-profile wind calculation. The authors did not give information about model general tendency to atmospheric stability or any characteristic of the domain stability. Therefore, logarithmic wind profile can be used for the interpolation

AR4: Thanks for your comment. Both logarithmic wind profile (log law) and power law wind profile (power law) assumes neutral atmospheric stability of the atmosphere and will have deficiencies in stable and unstable conditions. [2] reviewed different extrapolation methods (logarithmic models, Deaves and Harris, and power law) for 96 different locations worldwide. He concluded that the power law was the most reliable and widely most used extrapolation method. In addition, according [5] the usage of log-law is most suited near the surface.

AC4: Despite the aforementioned results from Gualitieri and Sill we have compared the performance of the log law and the power law (with time varying power exponent) for the six offshore sites used in this study. For the majority of the sites, interpolating NORA3 wind speed to the sensor height using the two different methods (log-law and power law) we see that the model bias using log law is larger than using the power law method for the majority of the sites. Therefor, we stick to the power law method with a time dependent power exponent when wind speed interpolating is conducted.

We added a clarifying text to section 2.3 (line 125 to 131).

6 Response to minor comments from reviewer # 3

RC1: Line 52: WRF model running 10 separate model runs for 10 independent regions" should be ...running independently for 10 domains.

AR1/AC1:: Done.

RC2: Line 80 "model runs with a horizontal resolution of 31 x 31km" should be 3x3km

AC2: Done.

RC3: line 83 "by many European countries" instead of the countries Weather Services should be proper.

AR3/AC3: Done. The sentence is rewritten in the following way: "... and it is used in short-range operational forecasting and research by many European weather *services and research institutes"*

RC4: Line 95-96 "from the National Centers for Environmental Protection (NCEP) (1oresolution)" there are different horizontal resolution NCEP data be specific for that. I think it should be NCEP-FNL data.

AR4/AC4: The section will be removed in the revised version of the manuscript.

RC5: Line 190 "Hence, the observed wind speed is somewhat more intermittent and variable than the modeled wind speed, indicating that HARMONIEAROME is missing some of the high-frequency variability embedded in the wind field". I think this sentence conflict with the previous sentence, the authors stated that the observation wind speed is between 4.7-6.5 m/s. I understand that the modeled wind speed is more variable but wind speed between 4.7-6.5 m/s is less seen from the model outputs

ACR: Thanks for commenting on this badly explained phrase. During the generation process of the paper we performed a time-filtering of the observational and modeled data using a Butterworth filter. This result was not implemented as a part of the paper. Performing this filtering procedure illustrated that NORA3 was missing out some high-frequency variability. **AC5:** We have removed the "high frequency" from the sentence. See line 226.

RC6: Line 191 "HARMONIEAROME" should be HARMONIE-AROME.

AC6: Done.

RC7: Line 193-196 Weibull distribution notation I recommend to use common parameters terminology for Weibull distribution which the readers are familiar instead of a and b notations.

AR7/AC7: Thank you for this comment. I have now changed a and b to λ and k.

RC8: It is interesting to see the dominant direction missed by the model especially at Sleipner where the flow is channeled. I would expect the almost bidirectional wind as we can see on the observation in Figure A0 for Sleipner (NW-SSE direction). How does the authors explain this? Do they believe the model capture the common flow patterns?

AC8: Thank you for bringing attention to the wind rose at Sleipner and for this interesting comment. Looking at Figures A1, B1 and B2 we can conclude that the model is somehow struggling with winds in the whole sector SSE-NWW for Sleipner, underestimating the wind speed and also the occurrence of these wind events in this sector. Since an oil- and gas platform is a large structure some flow distortion is expected to happen. It is difficult to state whether the difference between the observations and NORA3 values at Sleiper is due to the model or due to flow distortion caused by interaction between the flow and the platforms. However, Sleipner is the site with largest deficiencies between the observational and modeled data. For that reason a flow distortion of unknown magnitude and importance is probably present at Sleipner.

RC9: I recommend the authors use meteorological wind direction notation instead of given the direction in a degree unit. For example, 0:29 NW, 30:59 NNW in the text and also in Figure B2.

AR9: Thanks for your suggestion for changing the xlabels in Fig. B2 to meteorological wind direction notation instead of given the direction in a degree unit.

AC9: We have changed the labels on the x-axis in Fig. B2 to abbreviations of the corresponding wind direction interval (SE, NNW osv). I have also changed the text in sections related to wind directions accordingly, and the caption of fig B2.

RC10: Also I saw "wind park" many places in the manuscript. I recommended to use "wind farm" (line 146,228,230).

AC10: Done.

References

[1] A. Beck. Impact of nesting strategies in dynamical downscaling of reanalysis data. *Geophysical Research Letters*, 31(19):L19101, 2004.

[2] G. Gualtieri. A comprehensive review on wind resource extrapolation models applied in wind energy. *Renewable and Sustainable Energy Reviews*, 102:215–233, 2019.

[3] H. Haakenstad, Ø. Breivik, B. R. Furevik, M. Reistad, P. Bohlinger, and O. J. Aarsnes. NORA3: A non-hydrostatic high-resolution hindcast for the North Sea, the Norwegian Sea and the Barents Sea. *Submitted to Journal of Applied Meteorology and Climatology*, 2021.

[4] M. Raffa, A. Reder, M. Adinolfi, and P. Mercogliano. A Comparison between One-Step and Two-Step Nesting Strategy in the Dynamical Downscaling of Regional Climate Model COSMO-CLM at 2.2 km Driven by ERA5 Reanalysis. *Atmosphere*, 12(2):260, 2 2021.

[5] B. L. Sill. Turbulent boundary layer profiles over uniform rough surfaces. *Journal of wind engineering and industrial aerodynamics*, 31:147–163, 1988.