wes-2023-177

Improving Wind and Power Predictions via Four-Dimensional Data Assimilation in the WRF Model: Case Study of Storms in February 2022 at Belgian Offshore Wind Farms

Tsvetelina Ivanova, Sara Porchetta, Sophia Buckingham, Jeroen van Beeck, and Wim Munters

This manuscript investigates the utilization of offshore in situ observations, specifically from SCADA and lidar devices, to improve wind speed, direction and power prediction of the WRF model near and at a wind farm. The study focuses on the application of an observational nudging approach to assimilate the measurements into WRF and considers also the inclusion of wake effects through the parameterization of wind farms.

General comments

- The manuscript is excessively convoluted and lacks clarity, with a presentation of concepts and findings that is somewhat fragmented, making it difficult to follow the logical progression of the study. Authors should consider revising the manuscript readability.
- According to what it stated in the manuscript, the main aim of this study is to improve the forecasting accuracy of a numerical model during extreme events. However, the results and methodology presented do not really address forecasting issues. Instead, they use WRF (past) output information nudged at the same time when (past) observations were done. A reformulation of the objectives and motivation of the paper would better align with the findings presented.
- Although the paper outlines the data assimilation methodology employed, there is a lack of in-depth description regarding the selection criteria for the nudging parameter and the configuration of the nudging algorithm. Key aspects such as vertical and temporal weighting, as well as horizontal nudging weighting function, require proper definition.
- The results section primary outlines the different modelling scenarios considered and presents the accompanying figures and plots. However, there is not enough analysis, interpretation and discussion of the paper's findings. More discussion is required to justify how the paper addresses its objectives. For instance, the discussion provided regarding Fig. 10 provides only a superficial description of the figure itself.
- In situ measurements play a crucial role in the study since they are used for both the assimilation of the WRF model and the validation of the results. However, the description of the measurement devices and corresponding processing is insufficient.
- Finally, it is strongly recommend to align the paper with the specific writing guidelines from WES: WES - Submission (wind-energy-science.net), and to revise the consistency of units, format and styling used. Some examples: using ms^{-1}, instead of $m\ s^{-1}$; using 20h30 instead of 20:30 h, consistency in time format (21h and then 20h00 in page 8), figure captions should be included in the text and not in the figure itself.

Specific comments:

Line 20: "…although these can have strengths especially in post-processing techniques". What strengths are you referring to?

Line 86: additional relevant reference: https://doi.org/10.1016/j.jcp.2007.01.037

Lines 93-95: The text enumerates all the parameters of the equation, but it does not provide the values for these parameters used in the study. Moreover, some crucial elements such as the weighting functions (vertical, horizontal and temporal) are not defined. In addition, for clarity, it would be beneficial to provide the description of the nudging methodology subsequent to the description of the WRF model and its configuration.

Line 125: "…with sufficient points across the wind turbine rotor for a typical offshore wind turbine in the Belgian-Dutch cluster." How is the resolution of the WRF model levels within the turbine rotor area? Are these vertical points coinciding with the heights of the measurement devices or there is any vertical interpolation?

Line 133: "… our study investigates different configurations for FDDA with respect to data sources…". In which regard has the FDDA being configuring depending on data sources?

Line 139: Does it refer to a vertical profiler lidar?

Line 142: "The LiDAR (ZX 300M) is installed at that platform since August 2021 and has been collecting wind speed and wind direction information since". It would be beneficial to specify the specific time period during which data was utilized.

Line 143: "…located upstream from the farm…". The use of "upstream" and "downstream" can be ambiguous without a clear reference direction. Consider specifying the direction, such as "southwest" to provide clarity. Idem for the "downstream" used in line 158. Additionally, if only data from the southwest direction were used in the study, it would be relevant to mention this to explain why no wind direction filtering was required.

Line 144: The eleven measurements heights are referenced above the instrument or above the sea level? information on any lidar data processing or quality checks implemented would be valuable.

Line 148: Further details about the SCADA data are necessary. Include information on the quantity of data used, which wind turbines it was collected from, and whether any data processing was conducted. In line 151 it is mentioned that lidar and SCADA data are pre-processed, but these processes are not defined in the manuscript. Then, in line 152 a "standard procedure" to translate to horizontal wind components is mentioned, but not defined.

Line 155: "…LEG platform is positioned at approximately 30 km South-West from Hoek van Holland…". Where is Hoek van Holland?? It is mentioned but not indicated in any map, or specified if it is a region/city/… in (I guess) The Netherlands?

Line 159: It is important to include information about the lidar at the Europlatform. Consider adding details such as the lidar model, measurement heights, and any data processing procedures employed for this particular instrument.

Figure 2: Adding panel labels to figures with multiple panels enhances clarity (applicable to all figures in the paper). In addition, when referring to "two subsets of 5 wind turbines each" in

the figure caption, make sure this concept is introduced and explained in the text prior to the figure. Also, what is the distance between the cluster and the EPL and LEG lidars?

Line 170: "Gradually transient" is unclear and could be interpreted in different ways. Consider revising this phrase to provide a clearer description of the wind flow.

Line 178: "…LiDAR observations in Fig. 3 evaluate…" This is an awkward statement. Instead clarify how the mentioned wind direction (247.52 deg) was calculated.

Line 184: "…contains a subset of 5 wind turbines considered as front row with respect…" Is there any reasoning regarding which 5 turbines are considered? Which specific turbines within the cluster were used?

Figure 3: the top panel plots normalized wind speed, but what is the reference used for normalization? Additionally, it seems there are some data gaps. What is the cause of these gaps?

Line 202: Was it considered to test a larger Rxy? Considering the large distance between the lidar and the turbines, it may be interesting to see if a larger Rxy has a more notorious effect in the predictions at the turbine locations. This would help to elucidate the optimal horizontal weighting configuration of the nudging algorithm, which is one of the paper goals.

Lines 207 to 210: The cyclical approach described needs further clarification regarding its differences and benefits over the standard approach. It is unclear why assimilated measurements for one hour would be more effective than continuous assimilation throughout the period. Additionally, are not the assimilated measurements propagated in the standard approach as well? "…data is assimilated and its effect is propagating as the simulation is running"

Line 219: Including the formula used to calculate MAE would be helpful. Additionally, it is surprising that only one metric (MEA) has been used in the discussion of results. Additionally, while MAE is commonly used, incorporating other metrics like root mean squared error could provide a more comprehensive analysis, especially considering the focus on forecasting where outliers are significant.

Table 2: Clarify the meaning of L1-L6 and S1-S3 in the table caption. Additionally, table captions are usually located above the table, not below.

Line 229: It is mentioned that "The power is also normalized by the rated values.". But no power data is presented in Fig. 6.

Line 233: "…reduced MAE values…". Specify what the reduced MAE values are being compared to for clarity.

Line 239: "…both wind speed and wind direction, especially at the wind farm location.". Is it not this statement somewhat predictable, given that you are comparing the nudged results with the observations used for assimilating WRF?

Line 242: Clarify what is meant by "reanalysis of various events".

Figure 5: Comments regarding the results should go in the main texts, not in the figure caption ("…results… show that energy is indeed extracted from the flow").

Line 245: Why lidar and SCADA data were never assimilated simultaneously?

Line 247: Define what is meant by a "lidar data point". A measurement height maybe?

Line 249: "The configuration of these 11 cases are detailed in Table 2". For clarity and brevity, do not need to repeat this statement since it was already mentioned in line 245.

Figure 6: What are the grey lines next to the scatter crosses indicating? It seems that some scatter points have these lines, and some others don't. Additionally, they are missing in the wind direction plot.

Figure 7: any reasoning about why the assimilation of SCADA is not really improving (or even slightly worsening) the results at the EPL and LEG locations? Considering they are downwards the turbines; shouldn't the propagated nudged information be more productive here? Maybe a larger Rxy would be helpful on this regard?

Line 269: Clarify the difference between the nudging and assimilation windows. The manuscript indicates that "The assimilation window τ during which observations are considered by the model…", but how are the observations being considered here if they are assimilated during nudging window (from line 274: "…quantities being assimilated during the nudging window…").

Line 273: Explain what is meant by the lidar being "prominently positioned".

Fig. 10 is presented but never discussed or commented.

Line 280: "…nudging windows to showcase this effect.". Clarify which effect is being referred here.

Line 287: "…we compute the MAEs in height with respect to the Whi LiDAR profiles.". This sentence is unclear and may require rephrasing or further explanation.

Line 289: "The height of the assimilated LiDAR point is solely at 104.5 m, yet improvements with respect to observations are perceived along the whole profile." . Is this a result of the vertical weighting approach used? It would be clarifying to explicitly describe the weighting approach.

Line 290: "A vertical smoothing in wind speed profiles is expressed in this figure which ensures the smooth transition between simulation and observation." Could you further explain this? How was this smoothing implemented?

Line 305: "Performing FDDA of SCADA also enhances (locally) the predictions (the three rows corresponding to 'F2, WFP FDDA, S1-3' with the three different radius of influence 10 km, 4 km, and 2 km." Could you clarify this sentence?

Figure 8: Recheck date format along the manuscript for consistency.

Figure 9: Provide an explanation for the green shadowed areas in the figure and why they change size along the figure. Caption should include a description of the assimilation window indicated with the tau symbol. Additionally, avoid discussing results in figure captions ("MAE values are reduced for the case of cyclic nudging in 'F2, WFP FDDA L6' as compared to when no FDDA is performed (in 'F2, WFP')").

Figure 11: It seems that increasing the Rxy improves the effect of the FDDA. Why not trying even larger values to find the optimal value of Rxy? Same with Gq.

Figure 12: Discuss why the wind speed MAE of "F2, WFP" is larger than the error of "F2, WFP_off" at all locations.

Line 316: The manuscript primarily discusses the potential of the employed approach for forecasting applications, yet focuses on improving past model data by nudging towards the observation timeframe (also past). Further discussion regarding how the methodology could be adapted for forecasting purposes is needed.

Line 331: “the downstream region of interest if the wind direction is not from South-West”. If the wind does not come from South West, the region of interest is not downstream anymore.

Line 332: There are additional and more relevant reasons for the scarcity of offshore in situ data, such as the cost of deployment and maintenance of measurement campaigns, as well as the structural limitations of installing devices in deeper waters.

Technical comments:

Line 13: Only cases when the in-text citation is part of the sentence must be formatted as “Skamarock et al. (2019)”. In any other case, should be “(Skamarock et al. 2019)”. See WES - Submission (wind-energy-science.net).

Line 39: “highlights” instead of “highlight”

Line 245: “reference” instead of “referece”

Line 249: “of” instead of “off”

Line 300: “namely when not performing any nudging ‘F2, WFP_off’ and ‘F2, WFP’, when nudging only LiDAR data ‘F2, WFP FDDA, L1-6’s, and when nudging only SCADA ‘F2, WFP FDDA, S1-3’s, all of which have their parameters detailed in Table 2.” This sentence is redundant, since readers are already referred to Table 2.