

Response to 2nd round of reviews: Our responses are given in green.

Report #1

Submitted on 03 Jun 2025 Anonymous referee #1	
Checklist for reviewers	
ATTENTION: before filling this section, please check if are you reviewing a normal submission or a data description article. For a normal submission, please fill the top part of the form entitled "Non-data description articles". For data description articles, fill out the relevant part of the form entitled "Data description articles". PLEASE DO NOT FILL BOTH PARTS.	
Non-data description articles	
All types of manuscripts (except for data description articles) need to be evaluated by you according to the following three criteria.	
1) Scientific significance: Does the manuscript represent a substantial contribution to scientific progress within the scope of Wind Energy Science (substantial new concepts, ideas, methods, analyses, or data)?	Excellent Good Fair Poor n/a
2) Scientific quality: Are the scientific approach and applied methods valid? Is sufficient information given so other researchers (in principle) can repeat the work? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?	Excellent Good Fair Poor n/a
3) Presentation quality: Are the scientific results and conclusions presented in a clear, concise, and well-structured way (abstract conveys efficiently the essence of the paper; number and quality of figures/tables; appropriate, fluent and precise use of English language)?	Excellent Good Fair Poor n/a
Data description articles	
For data description articles , you are asked not just to assess the manuscript but, more importantly, the data set itself. If you are reviewing another manuscript type, please select "n/a" for the questions below.	

1) Scientific significance: Is there any potential of the data being useful? This is clearly the most important decision. There are at least three sub-criteria to evaluate:	
Uniqueness: it should not be possible to replicate the experiment, observation or data generation on a routine basis. This is also the case for cost-intensive data sets that might not be replicated due to economic reasons.	Excellent Good Fair Poor n/a
Usefulness: it should be plausible that the data, alone or in combination with other data sets, can be used in future investigations, for the comparison to model outputs or to verify other experiments or observations.	Excellent Good Fair Poor n/a
Completeness: a data set must not be intentionally split, for example, to increase the possible number of publications. It should contain all data that can be reviewed without unnecessary increase of workload and that can be reused in another context by a reader.	Excellent Good Fair Poor n/a
2) Data quality: The data must be ready and accessible for inspection and analysis to make the reviewer's task possible. Even if a submitted data set is the first ever published, its claimed accuracy, the instrumentation employed, and methods of processing should reflect the "state of the art" or the current "best practices". Reviewers will then apply their expert knowledge and experience to perform tests (e.g. statistical tests) and judge whether the data and any possible claimed findings are plausible and do not contain detectable faults.	Excellent Good Fair Poor n/a
3) Presentation quality: The article should describe in a clear, concise and well-structured way the data set and how it was obtained, using an appropriate, fluent, and precise use of the English language. The article text and references should contain all information necessary to evaluate all claims about the data set, whether the claims are explicitly written down in the article, or implicit, through the data being published or their metadata. The authors should point to suitable software or services for simple and free visualization and analysis.	Excellent Good Fair Poor n/a
For final publication, the manuscript should be accepted as is. accepted subject to technical corrections. accepted subject to minor revisions. reconsidered after major revisions: rejected.	

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Suggestions for revision or reasons for rejection

(visible to the public if the article is accepted and published)

Response: We thank the reviewer for their original comments and this positive assessment.

Report #2

Submitted on 04 Jul 2025
Anonymous referee #2

Checklist for reviewers

ATTENTION: before filling this section, please check if are you reviewing a normal submission or a data description article. For a normal submission, please fill the top part of the form entitled "Non-data description articles". For data description articles, fill out the relevant part of the form entitled "Data description articles". **PLEASE DO NOT FILL BOTH PARTS.**

Non-data description articles

All types of manuscripts (except for data description articles) need to be evaluated by you according to the following three criteria.

1) Scientific significance:

Does the manuscript represent a substantial contribution to scientific progress within the scope of Wind Energy Science (substantial new concepts, ideas, methods, analyses, or data)?

Excellent **Good** Fair Poor n/a

2) Scientific quality:

Are the scientific approach and applied methods valid? Is sufficient information given so other researchers (in principle) can repeat the work? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?

Excellent **Good** Fair Poor n/a

3) Presentation quality:

Are the scientific results and conclusions presented in a clear, concise, and well-structured way (abstract conveys efficiently the essence of the paper; number and quality of figures/tables; appropriate, fluent and precise use of English language)?

Excellent Good Fair Poor n/a

Data description articles

For **data description articles**, you are asked not just to assess the manuscript but, more importantly, the data set itself. **If you are reviewing another manuscript type, please select "n/a" for the questions below.**

1) Scientific significance:

Is there any potential of the data being useful? This is clearly

the most important decision. There are at least three sub-criteria to evaluate:	
Uniqueness: it should not be possible to replicate the experiment, observation or data generation on a routine basis. This is also the case for cost-intensive data sets that might not be replicated due to economic reasons.	Excellent Good Fair Poor n/a
Usefulness: it should be plausible that the data, alone or in combination with other data sets, can be used in future investigations, for the comparison to model outputs or to verify other experiments or observations.	Excellent Good Fair Poor n/a
Completeness: a data set must not be intentionally split, for example, to increase the possible number of publications. It should contain all data that can be reviewed without unnecessary increase of workload and that can be reused in another context by a reader.	Excellent Good Fair Poor n/a
2) Data quality: The data must be ready and accessible for inspection and analysis to make the reviewer's task possible. Even if a submitted data set is the first ever published, its claimed accuracy, the instrumentation employed, and methods of processing should reflect the "state of the art" or the current "best practices". Reviewers will then apply their expert knowledge and experience to perform tests (e.g. statistical tests) and judge whether the data and any possible claimed findings are plausible and do not contain detectable faults.	Excellent Good Fair Poor n/a
3) Presentation quality: The article should describe in a clear, concise and well-structured way the data set and how it was obtained, using an appropriate, fluent, and precise use of the English language. The article text and references should contain all information necessary to evaluate all claims about the data set, whether the claims are explicitly written down in the article, or implicit, through the data being published or their metadata. The authors should point to suitable software or services for simple and free visualization and analysis.	Excellent Good Fair Poor n/a
For final publication, the manuscript should be accepted as is. accepted subject to technical corrections . accepted subject to minor revisions . reconsidered after major revisions: rejected.	

Were a revised manuscript to be sent for another round of reviews:

I would be willing to review the revised manuscript.

I would not be willing to review the revised manuscript.

Suggestions for revision or reasons for rejection

(visible to the public if the article is accepted and published)

The reviewers' efforts to enhance the readability of the text and figures are appreciated. I would like to offer a few concluding observations.

It is unfortunate that the authors are unwilling to explore an alternative roughness parameterization scheme, especially since such a scheme has been shown to better predict wind fields under wind–wave misalignment, which is directly related to the research question of this manuscript. Considering that the simulations are relatively short, incorporating an additional scheme would likely have been feasible within the time frame.

If this is truly outside the scope of the current project, it is imperative to acknowledge the limitations of the present model setup in addressing wind–wave misalignment at an earlier stage in the manuscript. This acknowledgement should be made prior to any mention of future work. Furthermore, the discussion section would benefit from a more extensive examination of the limitations of the model. This should encompass not only the roughness parameterization but also the uncertainty associated with wind farm parameterization.

Specific remarks:

Line 71: Please clarify that this refers to wave height.

Line 95: Consider adding "during hurricane events" to clarify the context, similar to the phrasing in line 108 referring to extreme events.

Line 118: Rephrase the storyline for clarity.

Could the uncertainty of the observations be addressed or quantified, at least qualitatively?

Figure suggestions:

To improve readability, consider adding titles such as "Irene" and "Sandy" above the respective columns.

Response:

It is not really our “unwillingness” to make new simulations with other roughness length schemes. It is an unfortunate reality of finite and fully exhausted resources (both personnel and computational) that has become more acute in recent months.

We note, in case the reviewer missed it, that we did discuss this matter in section 2.2 (much before the discussion of the simulations):

“Variation of wave state and z_0 with WS is an important determinant of extreme, near surface WSs and turbulence intensity (Zambon et al., 2014b; Porchetta et al., 2019; Porchetta et al., 2020; Porchetta et al., 2021; Wang et al., 2024). The COAWST simulations are configured using the Taylor Yelland formulation (Taylor and Yelland, 2001) to calculate z_0 following past research (Zambon et al., 2014a) that found use of this parameterization resulted in better fidelity for Hurricane Sandy track, intensity, SST, and Hs than alternatives (Oost et al., 2002; Drennan et al., 2005). Use of the MYNN surface layer with WRF and the DRAGLIM_DAVIS drag limiter option with COAWST, means all simulations implement a maximum ocean roughness drag coefficient of 2.85×10^{-3} , consistent with research that has shown asymptotic behavior of drag at high WSs (Davis et al., 2008).”

Naturally we fully acknowledge:

- Fu et al. were able to achieve greater agreement with observations of z_0 with an ANN model trained using; surface wind speed (WSPD), the angle between the peak wave direction and the wind direction (DIR_wav_wind), wave speed (Cp), friction velocity (u^*), significant wave height (Hs), wind direction (DIR_wind), and peak wave direction (DIR_wav) than with any physics-based scheme, which is interesting but it is not a formulation that has yet been adopted within the WRF community and could not readily be adopted for our research. Though we note the great potential for machine-learning emulators.
- That the work of Porchetta et al yielded very interesting results in terms of the impact of wind-wave misalignment on surface roughness length. The version of COAWST we employed does not (yet) have this roughness length scheme available as an option.

Accordingly, we have added additional text to the paragraph from 2.2 so it now reads:

“Variation of wave state and z_0 with WS is an important determinant of extreme, near surface WSs and turbulence intensity (Zambon et al., 2014a; Porchetta et al., 2019; Porchetta et al., 2020; Porchetta et al., 2021; Wang et al., 2024). Further, wind-wave misalignment also plays a key role in near-surface WSs and z_0 (Porchetta et al., 2019) and machine learning tools have suggested z_0 prediction accuracy can be improved by inclusion of wind-wave directional misalignment as a predictor (Fu et al., 2023). The COAWST simulations are configured using the Taylor Yelland formulation (Taylor and Yelland, 2001) to calculate z_0 following past research (Zambon et al., 2014b) that found use of this parameterization resulted in better fidelity for Hurricane Sandy track, intensity, SST, and Hs than alternatives (Oost et al., 2002; Drennan et al., 2005). Use of the MYNN surface layer with WRF and the DRAGLIM_DAVIS drag limiter option with COAWST, means all simulations implement a maximum ocean roughness drag coefficient of 2.85×10^{-3} , consistent with research that has shown asymptotic behavior of drag at high WSs (Davis et al., 2008).”

We have also added this comment (last sentence below) at line 260:

“Three-dimensional and joint occurrences of HH WS, Hs, and Tp, in WT-containing grid cells from the COAWST simulations are presented along with histograms of estimated wind-wave misalignment at the LA cluster centers in HH WS classes of 3 to $< 10.6 \text{ m s}^{-1}$, 10.6 to 25 m s^{-1} , and $> 25 \text{ m s}^{-1}$, to represent high thrust, moderate thrust, and above rated

WS (Fig. 2c). We caution that the specific model set-up may play a critical role in dictating wind-wave misalignment and feedback via the z_0 (Porchetta et al., 2019).”

We have also expanded the discussion regarding uncertainty, so the final paragraph of the manuscript now reads:

“Mesoscale simulations performed at convection permitting resolution such as those presented herein allow simulation of the hurricane lifespans and associated power production over large domains and can be used as here to assess whether improved treatment of atmosphere-ocean dynamical coupling alters extreme conditions of relevance to offshore WTs. However, it is important to acknowledge that results from any numerical simulations are subject to uncertainty. For example, the highest structural loading may occur in the cyclone eye-wall (Han et al., 2014) which is of a scale (Marks et al., 2008) that is not fully represented in the simulations presented here. Nevertheless, analyses of the simulations suggest the structure of the hurricanes is reasonably represented in our modeling (Fig. 4 and Fig. S4) and simulations performed at the same grid spacing were shown to represent some aspects of flow in the eye wall (Müller et al., 2024). Further, while the Fitch WFP used herein is the most widely adopted within the wind energy community, such schemes do not comprehensively treat WT wake generation (Fischereit et al., 2022). Past research has suggested that use of alternative schemes such as the Explicit Wake Parameterization tends to lead to substantially lower wake losses in large offshore wind farms and higher power production than are simulated using the Fitch scheme (Pryor and Barthelmie, 2024a). Future work employing mesoscale-microscale coupling (Wang et al., 2024) could be used to evolve further details of geophysical properties of relevance to structural loading. Further, the hurricanes simulated herein were extremely powerful and both tracked within 100 km of offshore wind energy LA cluster centers (C and D for Hurricane Irene, B and C for Hurricane Sandy). However, they do not represent a comprehensive climatology of historical or possible intense tropical/extratropical cyclones (Barthelmie et al., 2021). Undertaking comparable simulations of additional extreme cyclones and simulations with different configurations including alternative z_0 parameterizations (Porchetta et al., 2019; Fu et al., 2023) and a wave boundary layer model within SWAN (Du et al., 2017) would also be useful in determining if findings presented herein are generalizable and to quantify the degree to which the meteorological and oceanic extreme conditions vary according to the precise model formulation.”

Line 71: Please clarify that this refers to wave height.

Done – The manuscript now reads “Equivalent estimates of extreme WSs and Hs from buoy measurements are; 32.6 m s^{-1} and 9.5 m, respectively (Kresning et al., 2024).”

Line 95: Consider adding "during hurricane events" to clarify the context, similar to the phrasing in line 108 referring to extreme events.

Done – The manuscript now reads “Only limited previous research has sought to quantify the degree to which wind-wave coupling improves simulation fidelity and/or intensity for WSs at heights of relevance to offshore WTs particularly during extreme events.”

Line 118: Rephrase the storyline for clarity.

We gather that the reviewer does not like the term storylines. We have rewritten to avoid it. The manuscript now reads “Research presented herein focuses on simulations of two of the most powerful hurricanes that have occurred within the U.S. eastern coastal waters in which offshore wind energy LAs have been auctioned (Fig. 1, see further details in Table S1 and Figs. S1-S2).”

Could the uncertainty of the observations be addressed or quantified, at least qualitatively?

This matter was not raised in the prior review, but we are happy to address it. We have added text to the paragraph that begins at line 225 as shown below:

“These data sets have some inherent constraints. For the HURDAT2 data set this includes use of subjective smoothing to produce representative 6 h track data which does not necessarily equate to a precise storm history (Landsea and Franklin, 2013). HURDAT2 accuracy for landfalling hurricanes has been estimated as $\sim 2.5 \text{ m s}^{-1}$ for maximum WS, 1 hPa for SLP, and the location is correct within approximately 0.1° of latitude and longitude (Landsea and Franklin, 2013). Spatial averaging naturally impacts the spatial distribution of precipitation within the gridded IMERG data and can lead to underestimation of high precipitation rates compared to point measurements (Hu and Franzke, 2020; Nie and Sun, 2020; Huffman et al., 2024). With respect to the NDBC data, there is a limited number and thus spatial coverage of buoys (NDBC, 2009). Efforts to optimize buoy design to enhance NDBC measurement accuracy have been previously documented (Taft et al., 2009) as have data quality control procedures (NDBC, 2023). Best available information suggests the total sensor accuracy for WS is $\pm 1 \text{ m s}^{-1}$, although lower accuracy may arise during high wave states (NDBC, 2023). For Hs the stated accuracy is $\pm 0.2 \text{ m}$ and for SLP it is $\pm 1 \text{ hPa}$.”

Figure suggestions:

To improve readability, consider adding titles such as "Irene" and "Sandy" above the respective columns.

The name of each hurricane is now included in all figures.