"Ship-based lidar measurements for validating ASCAT-derived and ERA5 offshore wind profiles"
Rev v2
Hugo Rubio, Daniel Hatfield, Charlotte Bay Hasager, Martin Kühn, and Julia Gottschall
Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2024-11, 2024

Authors response to reviewer comments

We would like to thank the referees once more for their time and effort in reviewing our work. We appreciate their feedback and comments. We have carefully considered their recommendations and concrete suggestions to enhance and clarify our work.

Below, we addressed the additional referees' comments and reply to them point by point. First, the referee's comment is included in italics and bold font, followed by our answer, and when applicable, the new excerpt of the revised version of the manuscript (highlighted in blue).

Anonymous Referee, Referee #1

Referee #1 general comments

1) ***First of all, it is still difficult to understand why the grid points with land contamination are included. To remove them would make the rest of the results much more useful. But it has helped that the results are separated into coastal and offshore areas.***
   We decided to include all grid points, including those with land contamination, to provide a more comprehensive view of how both ERA5 and ASCAT perform in different regions of the study area, especially when compared against the reference ship-based lidar measurements. One of the primary advantages of using non-stationary, ship-borne lidar measurements is that they allow for validation of dataset performance across a variety of regions. By including these points, we aim to show the performance variability both nearshore and offshore, offering a complete picture of the datasets under test accuracy in different environments.

   However, we agree that isolating the effect of land contamination is important to ensure clearer conclusions, particularly since the statistics can differ significantly between coastal and offshore regions. In response to this concern, we made an effort to modify several plots during the previous review round, which helped differentiate the performance of ASCAT and ERA5 in both areas more distinctly. We have also addressed these differences in the results and conclusions. As an additional effort in this direction, we have updated Fig. 13 to exclude nearshore grid points (within 20 km of the coastline), removing their influence from the overall results. The associated text and discussion have been adjusted accordingly to reflect these changes.

   We believe that for the remaining figures, further differentiation is not required, as the effect of coastal grid points is already effectively separated. Additionally, we want to emphasize the challenge of clearly defining a threshold for nearshore versus offshore regions in this particular study area. The region is densely populated with small islets, which, while not always considered true "land" regions, may still impact the performance of both ASCAT and ERA5.

2) ***The conclusion has improved, but it could still be more balanced on the usefulness of the approach. Secondly, it is true that satellite data are valuable measurements as an independent data source. However, when they are mixed with model data and the result, as shown here, is less precise than the model, then it is difficult to see the benefit. It is relatively easy to correct the consistent bias in the ERA5 profile, while the ASCAT-extrapolated wind profiles seem to add noise.***
We have further revised the conclusion to emphasize the main findings of this study, as well as to clarify the potential applications and limitations of the methodology used.

We also want to highlight that drawing conclusions such as "… when they are mixed with model data and the result, as shown here, is less precise than the model…" requires results supporting them. The results presented on this study quantify the performance of ASCAT-derived and ERA5 wind profiles, thus providing insights into the limitations and potential of these datasets and proving the usefulness of the study.

Furthermore, we do not completely agree with the assertion that ASCAT adds only noise, as the results show that both ASCAT and ERA5 exhibit comparable agreement with ship-based lidar measurements, although depending on the location and height under consideration. In fact, the inclusion of ASCAT data may offer a valuable supplementary dataset in cases where model-based outputs alone do not suffice for offshore wind characterization. The uncertainties inherent in numerical models can be mitigated by using multiple datasets, such as combining extrapolated ASCAT and ERA5, to examine the variability of the estimated mean wind speeds across different sources, and therefore helping to, for instance, quantify the risk of wind farm development prior to any available in-situ dataset. The results of this study also provide clear guidance on the limitations to be considered when using ASCAT data, enabling a more informed application of this dataset in future studies.

The following numbering follows the numbering used by the referee in their report, which follow the first author's response (submitted on 14/05/2024):

1 ***Regarding comparison of vertical wind profiles from ERA5, lidar and ASCAT-extrapolation: While comparing with ERA5 is not intended as a validation, there is still a discussion on what is more correct against lidar data in line 430 and 460 and onward which I find problematic in the form it is now, through the whole paper. "Overestimate", "underestimate" and "outperform" implies that a "truth" is known which is not the case in all places where these formulations are used.***
The manuscript has been reviewed to reformulate this wording, and maintaining terms such as "overestimate", "underestimate" and "outperform" in comparisons involving the lidar reference. In fact, we consider the ship lidar data as reference (as they come with an – in principle – well defined uncertainty).

Also: line 12: "...ERA5 displays a consistent bias of approximately 0.5 m s-1 along the profile, whereas ASCAT exhibits a smaller bias within the lower 200 m of the profile." please add a comment about the variability, as this is quite important.

The abstract in general, and this part in particular have been revised.

line 169: Please add: "It should be noted that ASCAT winds are used in data assimilation in ERA5." or something similar.
Included in line 184 of the new version of the manuscript.

5 Separating onshore and offshore in Figure 11 is helpful. But rather than more text on problems of ERA5 in a complex coastal site it would be nice to make sure that land contamination is avoided altogether. The justification from the authors of including the coast-affected data is that some of the points "show a significant agreement". Is that a good argument when the mix of surfaces within the scatterometer footprint may add up to values that are in the range of backscatter from water? This way it is not known what is measured and thus we don't know why there is an agreement.
We have addressed some of the concerns regarding coastal contamination in our response to Comment 1, and we fully acknowledge the referee's point on this challenge.

While we agree that coastal contamination can introduce uncertainty, the aim of our study is to compare the performance of ERA5 and ASCAT under different environments. Leveraging non-stationary ship-based lidar measurements allows us to assess performance both in far-from-shore and nearshore regions, which is a unique strength of this dataset. However, fully investigating the underlying reasons for ASCAT's limitations in coastal regions falls outside the scope of this paper, as such an investigation would require a much more in-depth study into the generation and processing of ASCAT wind fields.

We recognize that removing points potentially affected by coastal contamination would eliminate the discussion of these issues altogether. However, we believe that doing so would not substantially affect the primary discussion of far-from-shore regions, as we have clearly differentiated between coastal and offshore effects in our analysis. Additionally, discussing coastal contamination does not overshadow the other key findings of the study. On the contrary, we believe that highlighting these limitations contributes to a more comprehensive understanding of the dataset's capabilities and shortcomings in various environments.

Also: line 169: "By applying the IQR outlier detection, the impact of coastal contamination on the wind speed data is mitigated by removing unusually high wind speed retrievals in nearshore areas." -> "By applying the IQR outlier detection, the impact of coastal contamination on the wind speed data is mitigated." Only removing the high values through filtering may not remove all land effects. Is there a land flag in the ASCAT data set and did you use it? Please add information in the text about this.
Sentence rewritten for clarity and further context added on this regard.

line 369: "The reason for this is that, despite the filtering process for the ASCAT dataset, the coastal contamination still affects ASCAT measurements,..." and "The stronger impact of coastal contamination at 100 m can be attributed to the inaccurate characterization of stability conditions by ERA5 in nearshore locations due to its coarse horizontal resolution and limited ability to resolve fine-scale atmospheric features in these regions.": In my view the problem of ERA5 in coastal areas is a secondary problem if ASCAT is measuring some land instead of only wind. Please remove the last sentence.

We appreciate the reviewer's comment and agree that the primary cause of the stronger deviations in coastal areas is due to ASCAT's coastal contamination. The intent of this part of the manuscript was not to compare the overall differences between ASCAT and ERA5, but rather to explain why these differences are more pronounced at 100 m compared to 10 m. That said, we understand the referee's concern, and a more detailed discussion of the potential reasons for the variation in ASCAT and ERA5 differences with height is provided a few lines below (lines 390–394). Therefore, as suggested, we have removed the second sentence to avoid confusion and to highlight the primary issue.

19 ***Line 286: From first review: "Please comment on the fact that there is land in two of the grid boxes and argue why they can be included in the analysis."***
Authors response: "A comment regarding this fact has been added in lines 312-315 of the new version of the manuscript."
This review: Please add the comment in the text when describing figure 5, e.g in line 297, and explain why you choose to keep them (see also point 5 above).
We added a comment on this specifically referring to Figure 5 and Figure 6. However, we did refer to the wrong lines. This is within lines 308 and 312: "Furthermore, Figs. 5 and 6 show that the surface of certain ASCAT grid boxes, particularly those at or near the two harbours, is partially covered by land. This situation may lead to coastal contamination and excessively high wind speed retrievals within these grid boxes. The influence of this effect is discussed in the results section of this study, where its potential impact on the presented findings becomes apparent."

Other comments:

A. ***line 308-316: This part is a description of what is coming and can be removed or moved to the section at the end of the introduction.***
We acknowledge the referee's comment. While we have included a concise summary of the manuscript's structure at the end of the introduction, we believe that the additional details provided in these lines enhance the reader's understanding of the distinct subsections within the Results section. This text serves to clarify the specific analyses and comparisons that follow, offering essential context for the findings presented. Furthermore, given its brevity, we believe that this excerpt does not significantly lengthen the manuscript, yet it does provide valuable guidance for the reader navigating through the results.

B. ***Figure 8: please specify what is subtracted from what in the figure caption.***
Clarification added.

C. ***line 330: How can we conclude that the collocated approach is underestimating? Should it just be "The lower values associated with …"?***
Yes, this has been corrected according to referee comment.

D. ***line 370: "The stronger impact of coastal contamination…" -> "An additional source of uncertainty leading to stronger impact of the coastal contamination…"***
As suggested by the referee in comment 5, the sentence referred to in this comment has been removed from the manuscript.

E. ***line 374-383: This paragraph is a mix of results and discussion. It would be better to move it to chapter 4.***
   In this paragraph, we highlight the key conclusion derived from results presented in Figure 10, and as suggested by Referee #2, briefly compare these findings with previous literature. We believe this brief discussion helps interpret the results and fits well within the flow of this section.

F. ***line 423: In my opinion, "Excellent" is a too strong word for this visual comparison. Please modify, fex: "For the remaining locations, both datasets demonstrate comparable agreement with the lidar wind profile."***
   Modified according to referee suggestion.

G. ***The paragraph starting at line 475: The machine-learning approach was already mentioned in the section 2.4, there is no need to repeat it here since it was not applied.***
   ML approaches are mentioned twice in the manuscript—once in Section 2.4 and once in the Discussion (as referred to by the reviewer). However, the context and focus of these mentions differ. In Section 2.4, the mention is brief and serves as part of a general review of satellite wind extrapolation methods used in prior literature. In contrast, the mention in the Discussion highlights the potential advantages of ML approaches, suggesting that they may offer improved results over the mean stability correction method applied in this study. However, we also clarify that ML approaches are not feasible in this case due to the limited data available for training such models.
   Given that both mentions address distinct aspects of ML methods—one as part of a broader literature review, and the other as a reflection on the study's limitations and potential—we believe that both references are relevant and valuable to the manuscript. Therefore, we kindly suggest retaining this mention in the Discussion.

H. ***line 494: Please replace "observed region" with "open ocean" or "far from the coast" in the sentence: "The comparison between ASCAT and ERA5 winds reveals an overall good agreement when assessing the mean wind speed values across the observed region of the Baltic Sea".***
   Done.

I. ***line: 210: Here it is stated that the statistical MOST (applied in the paper) can be applied up to turbine operating heights, but in line 415 it is said that the heights above 200m are well beyond the range of applicability. Please write more clearly what range the method should be valid in.***
   To the authors' knowledge, there is no explicitly defined height limit for the applicability of the extrapolation methodology used in this study. However, as explained in the paper, it is expected to perform better than the instantaneous MOST at heights above the surface layer. There is currently no literature that defines or investigates a specific height for this methodology. In fact, this study represents the first comparison of ASCAT-extrapolated profiles using this methodology against in situ measurements at heights above 200 m.

   Given that most wind turbines operate below the 200 m threshold, we believe that stating the methodology is applicable "up to turbine operating heights" is effectively very similar as saying "up to below 200 m.". This aligns with the results presented, which show that extrapolated ASCAT profiles demonstrate a decreased performance above 170 m, and this is mentioned in the discussion presented results. However, establishing a more precise and definitive threshold for the validity of this methodology would require additional analyses.

J. ***There are typos throughout the manuscript.***
The full manuscript has been reviewed to correct typos and enhance its clarity and readiness.

Ine Wijnant, Referee #2

Before addressing Referee´s #2 comments, we would like to note that we believe Referee #2 may have reviewed the wrong version of the manuscript in this revision round. This is evidenced by comments referring to content that was either altered or removed in the latest version. Below are some specific examples:

- **Line 31**: Referee 2 suggested removing the word "therefore." However, this paragraph was removed entirely during the previous (first) review round. The comment corresponds to Line 31 in the first submitted version of the manuscript, where the word "therefore" was present.
- **Line 241**: Referee 2's comment discusses the "C± constants" being set to 6 and 4 for stable and unstable portions, respectively. This text is found in the first submitted version of the manuscript, but the corresponding section was revised in the second version (Line 241) submitted for review.
- **Figure 6**: Referee 2 mentions "Six locations used for the comparison of the datasets." This refers to Figure 6 in the first submitted version. After a new figure was added during the first reviewing round, this content now corresponds to Figure 7 in the updated version.

These examples suggest that Referee 2 may have reviewed the first submitted version of the manuscript, which is inconsistent with the version we intended for that review. As a result, some comments may no longer be relevant or applicable.

In response to such comments, we have highlighted where they may no longer apply to the current version of the manuscript or have already been addressed. However, most comments are still relevant, and we have worked to address them in the newly revised manuscript submitted after the second (current) review round to improve the paper.

Referee #2 general comments

1) ***There are a few things that I think need to be addressed in the paper:***
    a. ***Uncertainty in the lidar measurements. are the differences that you find with ERA5 and/or modified ASCAT significant? See e.g. page 14: TNO report - DOWA validation against offshore mast and LiDAR measurements | Report | Dutch Offshore Wind Atlas***
    This was already addressed during the first review round, in which we added some information on this regard in the Discussion section: ***"…***Additionally, it is acknowledged that lidar measurements, like any other observational data, are subject to inherent uncertainties that may impact the results (Duncanet al., 2019b; Rubio and Gottschall, 2022). Nevertheless, the observed deviations between the lidar measurements and both extrapolated ASCAT and ERA5 significantly exceed the magnitude of potential discrepancies attributable to floating lidar uncertainties, which can be up to approximately 2 % with mast-mounted anemometers as lower limit reference (Wolken-Mohlmann et al., 2022)".

    b. ***Your method is not robust with more/larger WFs (there are no wind farms yet in the Baltic according to https://map.4coffshore.com/offshorewind/,***

but you expect significant growth). Ship-based lidar measurements may be affected by wind farms (WF), ERA5 definitely does not take WF effects into account and ASCAT is too coarse to measure WF effects (at least in detail: then you need SAR).

Certainly, wind farm wake losses are a relevant consideration when investigating the potential energy yield of a site. However, as noted in our previous response to the referee, the primary focus of this study is not on evaluating the potential of wind farms, but rather on validating ASCAT and ERA5 datasets against reference lidar measurements, with special attention to wind speeds at turbine operating heights, providing insights regarding their potential and limitations in characterizing offshore winds.

While, in some parts of the paper (in the Discussion), we mention the potential application of paper's findings to wind resource assessment, this is not an objective of the paper. Instead, we believe the validation of these datasets helps to clarify their limitations and accuracy, offering insight into the contexts in which they can be effectively used or not. Although they are not usable for highly detailed and advanced wind resource assessments, these datasets are still useful for other applications, such as large-scale planning of wind (or energy) potential or preliminary site screening studies helping to identify regions with promising resources, particularly in areas with limited or no in-situ measurements. It is in these cases that we believe, these datasets can be used, and where the methodology and findings presented in our study are relevant, meaningful and applicable.

Lastly, as the referee correctly pointed out, wind farm wake losses do not influence the results presented here. First, no nearby wind farms currently exist in the study area to induce wake effects. Second, ERA5 does not account for wake losses of any type, and neither ERA5 nor ASCAT have the spatial resolution required to reliably capture internal wind farm wake effects.

However, in order to address referee's concern, we have revised the Conclusions section to clarify potential applications of the finding from this study and to highlight the limitations of ERA5 and ASCAT for characterizing wake effects: "…This methodology is particularly beneficial in scenarios where more complex extrapolation methods are impractical or when in situ measurements are limited, providing an additional source of wind information and thereby improving the reliability of offshore wind characterization studies. However, the application of both ERA5 and ASCAT must be approached with caution due to their inherent characteristics, including insufficient spatial resolution and the inability to adequately capture wind farm wake effects, which limit their utility for detailed wind farm energy yield assessments. Despite this, these datasets are still valuable for other applications, such as large-scale planning of wind potential or preliminary site screening studies, helping to identify regions with promising resources. In such cases, the findings of this study provide valuable insights into the conditions under which these datasets and methodology can be applied and the level of reliability that can be expected. Nonetheless, it is crucial to also acknowledge the primary limitations of this approach, such as excessive wind speed deviations in nearshore locations and the increased expected error at higher altitudes."

c. ***How does your method compare to assimilating ASCAT into the NWP reanalysis like it was done in DOWA (point 1 in [Innovations in the DOWA project | DOWA project | Dutch Offshore Wind Atlas])?***
   Data assimilation methodologies are out of the scope of this paper and therefore, these methods have not been investigated or discussed in this work.

d. ***You basically show in your paper that ASCAT and ERA5 should not be used closer than 40 km from the coast (validation results based on ship-based lidar). That is a conclusion that I miss in your paper. As far as I know the ASCAT coastal product is only valid 15 km away from the coast and ERA5 has problems with abrupt changes in surface roughness, such as on the coast. A model (such as ERA5) assumes a grid box average surface roughness for a combination of land and water whereas the wind feels land or water. The larger the grid box size, the larger the problem (ERA5 grid box size 31 km). So basically ASCAT and ERA5 have quality issues near the coast and this is what you find confirmed in your paper.***
   We acknowledge the referee's reflection. However, a key conclusion we draw from the study is that gridded datasets, like ASCAT and ERA5, are not really comparable to microscale measurements systems, such as, in this case, a floating lidar system. Consequently, the application of these datasets requires caution, especially in nearshore areas. It is important to manage expectations regarding their accuracy and reliability in these regions. Additionally, we did not define a strict threshold of 40 km, as doing so is challenging with the available data. The study area is densely populated with small islands and islets, which, while not always classified as "land" regions, can still influence the performance of both ASCAT and ERA5.

   In any case, the limitations of both ASCAT and ERA5 in nearshore regions have been highlighted throughout the entire manuscript, as this is a key conclusion shown by the presented results, and explicitly mentioned in the conclusion of the manuscript.

2) ***Comments from earlier review that have not been addressed yet are:***
   a. ***Are there other measurements that you can compare to lidar measurements in harbor (where ASCAT and ERA5 are particularly inaccurate)?***
   To the authors' knowledge, no additional measurements were available near the ship's track during the measurement campaign.

   b. ***Have you considered triple (or quadruple collocation) to assess uncertainties (there are also uncertainties in your lidar measurements! What are they)?***
   Triple collocation requires a large amount of data to effectively derive the uncertainty levels of ship-based lidar measurements. Unfortunately, due to the relatively short duration of the campaign and the limited availability of

ASCAT data during this period, insufficient data was collected to implement triple collocation.

In addition, while the application of triple collocation is indeed an interesting approach for further investigating the uncertainties in lidar measurements, it falls outside the scope of this paper and would be more suitable for future research.

c. ***Have you considered using other wind climatology's such as NEWA [GMD - The Making of the New European Wind Atlas – Part 2: Production and evaluation (copernicus.org)]?***

The temporal coverage of the NEWA does not align with the period of the ship-based lidar measurement campaign utilized in this study. Since our analysis and results are specific to this time period, incorporating NEWA would not have been applicable nor comparable. Therefore, we chose not to use it in this analysis.

Referee #2 specific comments

The referee has included additional comments in a bundled format within a PDF file. These "subcomments" are presented below each corresponding main comment in grey font, accompanied by a grey highlight in the main comment to indicate the specific text to which the subcomment refers, consistent with the referee's formatting in her reply.

3) ***Line 3: typo: observations***
   Corrected.

4) ***Line 9/10: The comparison reveals a close agreement between ASCAT and ERA5 beyond 40 km distance from the coast. Unclear what you mean: close agreement between two different approaches (account for stability)? At 10m height or also extrapolated to hub heights?***
   This part of the abstract has been modified for clarity, as requested by the referee.

5) ***Line 10/11: (Extrapolated) ASCAT tends to significantly overestimate the mean wind speed derived from lidar measurements, while ERA5 exhibits a consistent underestimation. I assume the difference between lidar measurement and (Extrapolated)ASCAT/ERA5 is larger than the lidar measurement uncertainty?***
   A comment regarding the uncertainty level expected for the floating ship-based lidar system has been included in the Discussion section of the paper.

6) ***Line 21: However, in situ …***
   Corrected.

7) ***Line 26/27: Floating lidar systems can be moved to different locations, but generally measure at one location for a certain period of time. With profiling lidar systems installed on cruising ships it is possible to provide reliable wind profile measurements over larger areas.***
   Manuscript modified according to referee suggestion.

8) ***Line 27/30: (can be formulated shorter/clearer): Before profiling lidar systems on cruising ships can become a generally accepted alternative for offshore met masts and floating lidar, specific challenges have to be overcome such as validation against reference data and quantifying the associated uncertainty (Rubio and Gottschall, 2022). Still, ship based lidar has already been used in different wind energy related studies. In Wolken-Möhlmann…***
   i. **You used industry standard, but industry standard for what? These measurements are only useful for validation of reanalyses which can be used wind resource assessments if the growing effect of wind farms is accounted for.**

   Manuscript modified according to referee suggestion.

9) ***Line 34-38: However, while numerical models have demonstrated good performance in shallow-water offshore regions compared to in situ measurements (Witha et al., 2019b), they often fail to describe the spatial and temporal variability of wind with***

sufficient accuracy and detail*. I suggest an alternative text: <u>Numerical weather prediction (NWP) models in re-analyses mode</u> are commonly used … spatial coverage. However, while numerical models have demonstrated good performance in shallow-water offshore regions compared to in situ measurements (Witha et al., 2019b; <u>Wijnant et al, 2019), they have problems with areas with large changes in surface roughness, such as the coast. The larger the grid box size, the larger the problem because the model assumes a grid box average surface roughness for a larger area (whereas the wind feels land or water, not a combination). Also most re-analyses do not take into account the (changing) effect of wind farms on the atmosphere (except: https://wins50.nl/</u>

   i. **'… they often fail to describe the spatial and temporal variability of wind with sufficient accuracy and detail' = very general conclusion which I do not agree with. The spatial and temporal variability is pretty well captured in the DOWA/WINS50 re-analyses (see validation section of KNMI Technical report - The Dutch Offshore Wind Atlas (DOWA): description of the dataset | Report | Dutch Offshore Wind Atlas). Witha just used one weather model (WRF). Abbreviate 'Numerical weather prediction models' to NWP models (not numerical models).**
   ii. ***<u>KNMI Technical report - The Dutch Offshore Wind Atlas (DOWA): description of the dataset | Report | Dutch Offshore Wind Atlas</u>***

Manuscript modified according to referee suggestion and reference suggested added.

10) ***Line 38-41: This limitation arises from factors such as the inaccurate parameterization of the model variables or the insufficient temporal and spatial resolution of the models' output data. Furthermore, the lack of in situ measurements in deeper offshore regions hinders the validation of these datasets, leading to increased uncertainties in derived wind statistics for such locations. I suggest an alternative text: <u>Each NWP model has its own limitations (caused e.g. by grid and domain size and physical modelling and parametrisation choices). This results in uncertainties in wind statistics based on these NWP models and these uncertainties can be quantified when validation measurements (incl. measurement uncertainties) are available. This is however often a problem for hub heights, especially for far-offshore locations with deep water.</u>***

   i. **Again a bit over simplified and I do not agree with what you write. I do not think that you can say 'inaccurate parametrisation of model variables'. Choices in NWP models always are a trade-off (optimise which forecast lead-time? Optimise which parameters?).**

Text modified according to referee suggestion.

11) ***Line 42-44: To overcome the limitations of in situ measurements and numerical models, satellite remote sensing devices have emerged as a potential alternative for characterizing ocean winds and climate over large areas, capturing the wind variability with a temporal coverage of over 15 years. I suggest an alternative text: <u>Scatterometer (wind) measurements from satellites are a welcome additional source of information in these data sparse areas.</u> Several studies …***

   i. **I would leave this 'model/measurement limitations' out because ASCAT has its own limitations. Basically you have scatterometer data from satellites as additional information in data sparse areas (it is no replacement for models** or other measurements). Note that another way of using ASCAT is data-assimilation in the NWP model, like in DOWA/WINS50 (see [TNO report - DOWA validation against ASCAT satellite winds | Report | Dutch Offshore Wind Atlas](#)).

The manuscript has been updated according to the referee's suggestion.

12) *Line 47-53: Fluffy writing: does not make it clearer and there are some mistakes in it. I suggest an alternative text: The ASCAT coastal product is available since 2007 and provides high quality offshore wind measurements on a 12.5 km grid spacing for locations further than 15 km from the coast. The ASCAT wind speed bias is less than -0.23 ms-1 in coastal areas (15- 50 km from the coast) and -0.29 ms-1 elsewhere (TNO report - DOWA validation against ASCAT satellite winds | Report | Dutch Offshore Wind Atlas). However, ASCAT has its limitations: only available twice a day (around 09:30 and 21:30 UTC) and stability dependent assumptions have to be made to derive turbine height winds from the ASCAT 10m winds.*

   i. Assume you used that?
   ii. Better. Someone might otherwise read this in 10 years time and think ASCAT is available since 2016
   iii. That is not the same as resolution!!! Ask Ad Stoffelen KNMI.
   iv. If I am correct: please check
   v. You write: 'Lastly, the trustworthiness of satellite retrievals remains a knowledge gap, due to the lack of available in situ datasets for validation especially in deep water regions'. I left this out because I think it is incorrect: ASCAT has been extensively validated (besides: its quality does not depend on water depth). Ask Ad Stoffelen KNMI.

As suggested by the referee, we have modified "12.5 km resolution" by "12.5 km grid spacing" across the manuscript. However, we have retained other details as originally presented since additional information regarding ASCAT—such as the year it became available, the specific product utilized in this study, and other relevant details—can be found in Section 2.2 of the manuscript. Furthermore, the text excerpt referenced by the referee in sub-comment v. was removed during the previous review round and was no longer present in the latest submitted version of the manuscript.

13) *Line 54: The Baltic Sea is an area of great interests for offshore wind development…*

   i. And yet: you do not mention the effect of wind farms (WF) on the atmosphere. ERA5 is without WF effects, ASCAT is too course, at least for detail (you need SAR for that), but your ship based lidar may measure the effects up to 100 (?) km from a WF. I think you should at least mention WF effects in the paper and tell what the consequences of these WF effects are for your method.
   Can you quanlify what you mean with near shore (I assume > 15 km from shore otherwise ASCAT not valid)?

In response to comment 1b, we have already argument why we believe wake effects are not relevant in this study. First, datasets used are not suitable/capable of capturing wake effects. Secondly, due to the absence of nearby wind farms in the period considered in this study, their effects in the results and methodology presented are none. The fact that there is interest in the development of wind farms in the Baltic Sea does not imply that our study needs to contemplate those. Especially, when at the time this study is realized, there are none. Furthermore, we believe that the inability to characterize wake effects is not a limitation of the methods applied in this study, but rather a limitation inherent to the nature of ERA5 and ASCAT themselves. Thus, such considerations are beyond the scope of this paper.

The quantification of nearshore regions is addressed in several sections of the paper, particularly during the discussion of the results. Specifically, we refer to two distinct thresholds (20 km and 40 km), as significant changes in the observed biases are noted when applying these distance values.

14) ***Line 64-71: I suggest that you change sequence of what you write to make it clearer, e.g. To derive wind profiles from the ASCAT coastal product 10 m measurements, we employ the long-term stability correction approach presented in Kelly and Gryning (2010) and implemented in Badger et al. (2016). For this, we utilize the stability information from ECMWF Reanalysis 5th generation (ERA5)and compare two different collocating methods to evaluate the potential influence of the limited temporal resolution of satellite overpasses in the ASCAT extrapolated profiles. Not only the ASCAT derived wind profiles, but also the wind profiles from ERA5 are then compared to the lidar profiles.***
Manuscript modified according to referee suggestion.

15) ***Line 75-76: … of the reliability and accuracy of satellite measurements derived wind statistics for offshore wind characterization at wind energy relevant heights.***
Corrected according to referee comment.

16) ***Line 86: What is the accuracy of your lidar measurements? If you want to compare your measurements to model data, you will have to be able to tell whether the difference that you find is significant (outside the measurement uncertainty). See e.g. TNO report - DOWA validation against offshore mast and LiDAR measurements | Report | Dutch Offshore Wind Atlas***
This was already addressed during the first review round, including a citation to the suggested reference:
"Consequently, the mean values derived from lidar measurements may exhibit biases that vary depending on the time slots during which measurements were acquired at particular locations. Additionally, it is acknowledged that lidar measurements, like any other observational data, are subject to inherent uncertainties that may impact the results (Duncanet al., 2019b; Rubio and Gottschall, 2022). Nevertheless, the observed deviations between the lidar measurements and both extrapolated ASCAT and ERA5 significantly exceed the magnitude of potential discrepancies attributable to floating lidar uncertainties, which can be up to approximately 2 % with mast-mounted anemometers as lower limit reference (Wolken-Mohlmann et al., 2022)".

17) ***Line 104-105: the motion (take the s out) effects***
Corrected.

18) ***Line 117: fig 2b is the daily cycle the ship (lidar) experiences because it is connected to the location of the ship. It is not how the wind depends on the hour in the day (which is what normally is meant by 'daily cycle'). Maybe use a different name to avoid confusion (wind speed daily cycle plots normally give highest wind speeds during the day), e.g. Wind speed ship daily cycle.***
Corrected as suggested by referee using "Wind speed ship daily cycle", both in figure caption and main text.

19) ***Line 31: Therefore (?), the ….***
The paragraph containing the "Therefore" mentioned by the referee was removed from the manuscript during the previous review round, as we felt it would be more appropriately placed in the conclusion or discussion section.

20) ***Line 155/156: … available horizontal grid spacings of 12.5 km and 25 km***
    i. **Not the same as resolution**

Corrected according to referee comment.

21) ***Line 159: what do you mean by Both of these (?) are implemented (?) at…***
We are referring to the two services mentioned in the sentence before. This has been clarified in the manuscript.

22) ***Line 168-170: By applying the IQR outlier detection, the impact of coastal contamination on the wind speed data is minimized, leading to more accurate and reliable results in nearshore areas.***
    i. **I assume this is part of the ASCAT coastal product? Is nearshore more than 15 km from the coast?**

The IQR outlier detection was specifically implemented in our study to filter out 10m ASCAT mean wind speeds with values excessively high near the coast. The relevant section of the manuscript has been revised to clarify this.

23) ***Line 189: Several methodologies <u>for vertical extrapolation of satellite measurements</u> …***
Corrected according to referee comment.

24) ***Line 260-264: As observed, considering the stability information from the full campaign results in a better theoretical distribution compared to the collocated approach. Although the difference is minimal at the harbor site, it is more pronounced at the offshore location, where a significant underestimation of unstable stability occurrence is observed. The harbor site presents a rather symmetric distribution around zero, meaning that both unstable and stable atmospheric conditions are equally represented. However, the offshore site exhibits a higher occurrence of unstable conditions, compared to the stable side of the curve.***
    i. **There are a few things to consider here: (1) at the harbors you have on average 'land behaviour' of stability which means a daily cycle with more stable in the night and more unstable during the day; offshore there is no such daily cycle (stability depends more on the season); (2) I assume that in your collocated harbor graphs you look at around 09:30 UTC and in your collocated offshore graphs at around 21:30 UTC. I assume for the whole**

measuring campaign you looked at day and night for the period 28 June 2022 to 21 February 2023?
Harbor: in the collocated set (around 09:30 UTC), you would expect unstable to be slightly underrepresented if it is day time, but you see the opposite (more unstable than the theoretical line which is derived for all stability classes). A lot depends on what is really day time off course (in winter well after 07:00). If you take the whole dataset you have apparently a higher percentage stable (?) and a better match with the theoretical line.
Offshore: there is no daily cycle effect on stability here. The only thing that affects stability is that you miss 3 months of the year where the sea is particularly cold (but that is the case for both the collocated and whole dataset). So that probably explains the overestimation of unstable. But why do we only see that for collocated, not for the whole campaign?

Section 3.1 of the manuscript was modified during the first reviewing round in order to better explain how the stability daily cycle may influence the differences between the collocated and full campaign approaches. Therefore, we would like to refer the referee to the latest version of the manuscript, since this issue is properly discussed there.

25) Line 192-193: …. performance at different vertical and horizontal constraints.

i. What do you mean?

This sentence was removed from the manuscript during the first review round.

26) Figure 6. Six locations used for the comparison of the datasets. The approximate distance to the nearest shore is indicated, in km, below of each site. Please add: Location A is the harbour of Nynäshamn (Sweden) and location D the harbour of Hanko (Finland).

We assume the referee is referring to Figure 7 of the new version of the manuscript, since Figure 6 was added after the first reviewing round and shows number of lidar samples recorded at each ASCAT grid cell.
The suggested text has been added to Figure 7 caption.

27) Line 241: In this study, the values for the C± constants have been set to 6 and 4 for the stable and unstable portions, respectively.

i. So basically the stability correction has only 2 values for C which are the same for the whole Baltic Sea and all heights, one for stable and one for unstable: correct? Did you consider other values, like the ones from literature? Can you show why these are the best values? Does it not depend on the season which values are…

Correct, as specified in the manuscript: "…C± constants have been set to 6 and 4 for the stable and unstable portions, respectively. … identical values of C± were applied to all ASCAT grid points."

As an example, the comparison between the ERA5 NPD of 1/L and the theoretical distribution obtained with values of C+=5 and C-=12 (used in (Optis et al. 2021)) is shown below. As can be observed, even with the inclusion of ERA5 data from the full campaign, the theoretical distribution struggles to closely resemble the NPD of the 1/L parameter calculated by ERA5.

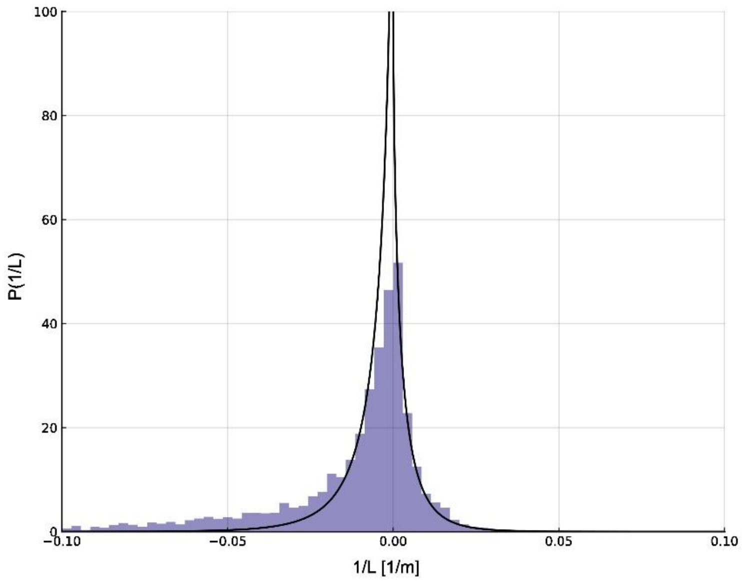

Figure 1: Normalized probability density function of 1/L from ERA5 and theoretical distributions calculated from Eq. (2) using 5 and 12 for C+ and C-, respectively. The same offshore location as in Figure 4d and full campaign data was considered for this example.

28) ***Line 307-311: First, the coastal contamination of near shore areas leads to the removal of some ASCAT overpasses for data quality reasons, leading to a reduced number of ASCAT observations in these areas. Consequently, the insufficient number of valid wind speed measurements obtained from the collocated approach introduces a biased representation of the prevailing stability conditions during the campaign period.***

    i. **ASCAT and ERA5 both have problems in coastal areas (see general remarks). So the uncertainty in the wind data that you use in fig 7 is large in these areas (larger than further offshore). Uncertainties in ERA5 are probably larger near the Swedish coast with prevailing W-SW winds. Insufficient number of measurements does not necessarily have to lead to a biased…**

We would like to refer the referee to the new amendments made in Section 3.1 during the first reviewing round.

29) ***Line 313-315 (from previous review): 'pronounced instability in the morning?' Why would ERA5 produce stronger unstable conditions (lower 1/L) in the morning at Nynashamn? What do we know about the water temperature near Nynashamn and how it is modelled by ERA5 (shallower/warmer water between Bedaron and the mainland maybe?)? ERA5 has grid boxes of 31 km2 so model values are probably very land-contaminated in that area: can you make a plot of the ERA5 grid boxes near the harbours? What is the prevailing wind direction? Basically ERA5 and ASCAT are not very good in coastal area: maybe you should take them out of your analyses?***

As in the previous comment, we would like to refer the referee to the amendments made in Section 3.1 during the first review round. These revisions now explicitly discuss the daily stability cycle at the different locations considered and highlight the significant role that the ERA5 land mask (or land contamination) plays in this context.

30) ***Line 315-316: This results in a lower wind speed compared to the full campaign approach, as can be derived from Eq. 4.***

    i. **So what you say is that the wind at 100m height is lower in more unstable conditions? If that is what you mean, it is wrong.**

This sentence has been modified for clarity and precision.

31) ***Line 316-317: In contrast, the other locations do not exhibit such pronounced daily stability cycles, and therefore, smaller differences are reported between the two approaches.***

    i. **I explained why this is the case for offshore. For the Finnish harbour the prevailing W-SW'ly wind is the reason for a reduced daily stability cycle. Please add the why to your paper.**

We would like to refer the referee to the new amendments made in Section 3.1 during the first reviewing round.

32) ***Line 317-320: Finally, as mentioned in Section 2.4, the same values of the semi-empirical constant C± are assumed for the entire region, instead of using a site-specific definition of these constants. Therefore, the suitability of the selected values may not be optimal for certain locations, leading to an anomalous theoretical representation of the empirical atmospheric distribution.***

    i. **I think your conclusion should be different. You can conclude that your method works well for offshore, but not near the coast (so not for the entire ship track) because of poor quality ASCAT and ERA5 near the coast (less than 31 km from the coast actually).**

This section was modified during the first reviewing round, so we would like to refer the referee to the new section. In the new version of the manuscript the text excerpt is not used as any conclusion, but as a reason that potentially contributes to explaining the differences between the collocated and full campaign approach.

33) ***Figure 9 basically shows you that ASCAT winds look unrealistic near the coast at 10 and (more so) at 100m. Especially near the Swedish coast where the wind blows predominantly from land to sea, wind near the coast should be lower than further offshore. So this figure proves that you cannot use your method near the coast for 2 reasons: (1) quality of ASCAT, (2) grid size of ERA5 (averages surface roughnesses of land and sea in grid box, therefore wrong for both wind from land and from sea). Small scale effects such as sea breeze and low level jets (you mention these in line 341) don't have a significant effect on your mean values.***

The limitations of both ASCAT and ERA5 in coastal regions due to coastal contamination and coarse grid resolution are now explicitly addressed in the manuscript: "The effect of coastal contamination in the ASCAT map is particularly visible in the 100 m height map, where the highest mean wind speeds are located along the perimeter of the region with available data. An additional source of uncertainty contributing to the stronger impact of the coastal contamination at 100 m is the inaccurate characterization of stability conditions by ERA5 in nearshore locations due to its coarse horizontal resolution and limited ability to resolve fine-scale atmospheric features in these regions."

In response to the comment on sea breezes and low-level jets, these are now only mentioned in the introduction (line 62).

Regarding the statement that "*this figure proves that you cannot use your method near the coast*", it is important to clarify that this limitation stems from inherent characteristics of the ASCAT and ERA5 datasets, rather than from the methodology employed in this study.

34) Line 341-342 (fig 9 10m validation): (from previous review) compare to Validation of DOWA ('undisturbed wind' = HARMONIE without WFP) with ASCAT (too coarse to measure wind farm effects) at 10 m height: TNO report - DOWA validation against ASCAT satellite winds | Report | Dutch Offshore Wind Atlas. Because you use ERA5 stability info to calculate ASCAT-derived wind speeds at 100m height, the difference you see at 100m should mainly be because of differences at 10m, right?

In the first review round, a comparison of the results presented in Figure 10 (previously Figure 9 in the initially submitted version of the manuscript) with the reference provided by the referee was already included, and the corresponding reference was incorporated into the manuscript at that time.

The fact that biases between the two datasets at 10 m and 100 m differ are basically explained by three main factors, which have been now added in the manuscript: "The differing biases between these two datasets at the two heights levels (10 m and 100 m) can be attributed to three key factors: first, the inherent difference between the datasets at 10m, second, the mean stability correction approach used to extrapolate ASCAT; and finally, as illustrated in Figure 8, the impact of the collocation strategy applied for the theoretical stability characterization".

35) Line 342: Figure 10a illustrates the difference in wind speed between ASCAT and ERA5 at 10 m and 100 m

i. **Wrong use of the word 'disparity' (nothing unfair about this difference).**

Corrected.

36) Lines 347-350: This discrepancy in the nearshore areas can be explained by the combination of too high wind speeds retrieved by ASCAT due to coastal contamination and ERA5's inability to properly resolve the coastal atmospheric phenomena and its coarse horizontal resolution that leads to the omission of the flow phenomena variations caused by the small islands present in these coastal regions.

i. **It has nothing to do with coastal atmospheric phenomena or flow phenomena variations (do you mean sea breezes?). It has everything to do with 'land roughness contamination' of the roughness in the coastal grid cells**

Sentence rewritten for clarity: "…ERA5's inability to properly resolve the coastal atmospheric phenomena and small-scale wind flow variations due to its coarse horizontal resolution."

37) ***Figure 10 shows you that you should not use your method within about 40 km from the coast (you should expect 31 km because of the grid size of ERA5 and what I explained earlier)***

Results presented in Figure 11 (Figure 10 in first submitted version of the manuscript) show higher discrepancies between ERA5 and ASCAT at both 10m and 100m (within these 40km distance to shore). However, since the extrapolation methodology used in this study does not affect the data at 10m, we cannot conclude that the method itself should not be used within 40 km from the coast. Rather, we believe a more accurate conclusion is that, within this region, higher uncertainty is expected in both ERA5 and ASCAT values, as evidenced by the larger differences observed due to the limitations of these datasets (e.g. ERA5 grid size as mentioned by referee). Therefore, we want to highlight that this is not due to a limitation of the method, but a limitation of these datasets. This has been discussed in the manuscript.

While the extrapolation may contribute to some additional uncertainty, as seen by the consistently larger bias at 100m compared to 10m (also explicitly mentioned in the manuscript), the key limitations regarding the applicability closer or further away from the shore lie in the datasets themselves, not in the methodology employed in this study.

38) ***Line 355-356: Nonetheless, the majority of grid points exhibit wind speed differences below ±1 m s-1. As previously discussed, wind speed differences above this threshold correspond to those of near-shore grid points.***

   i. **This big difference of 1 m/s in mean values is not the bias, but the max difference, right?**

This sentence was revised during the first review round.

To clarify, each point in Figure 11a represents a grid point within the study area, and the y-axis shows the difference between the wind speed obtained from ASCAT and ERA5 (i.e., ASCAT minus ERA5). This difference represents the bias between the two datasets, not the maximum difference.

As a reminder, as explained in Section 2.4, the extrapolation methodology used only provides mean wind speed values. Thus, for each ASCAT grid point, we have a mean wind speed for the entre study period (the duration of the campaign). These mean values are then compared against the corresponding mean wind speeds provided by ERA5, with one mean value per grid point as well. Consequently, the figure illustrates the bias in mean wind speeds between the two datasets, rather than instantaneous (or max or min) differences.

39) ***Line 400: what do you mean with the word 'trend 'here? The word trend is used for change in time (e.g. climate change), but this is not what you mean…***

The second part of the sentence has been omitted, since it does not add any meaningful information: This observation holds true for all three presented elevation levels. Notably, the western area…

40) ***Line 400-403: Notably, the western area of the ship route (longitude below 18.5 degrees) exhibits the largest errors for both ASCAT-derived winds (using ERA5) and ERA5 winds, with maximum differences exceeding 3 m s-1 at all elevation levels. This indicates that wind speed estimation cannot be done accurately enough in these***

areas with ASCAT and/or ERA5 because (1) poor quality of ASCAT coastal product closer than 15 km from the coast and (2) ERA5 grid box size (surface roughness in land-water grid boxes on the coast problematic).

i. **Is it possible to add distance to the nearest coast to fig 13? In this figure we are looking at winds at 60m, 150m and 220 m, so at ASCAT derived winds (with ERA5). The ASCAT coastal product is only valid 15 km or more out of the coast as far as I know...**

Manuscript modified according to referee suggestion.
As suggested by the referee, the approximate distance to shore is now indicated in Figure 14 (Figure 13 in first submitted version of the manuscript).

41) ***Line 404-405: highlighting the different shear resemble obtained from each of the datasets and their different representation of the wind profiles .***

i. **Sentence unclear: shear resemble?**

This sentence was already modified during the first reviewing round: "highlighting the different shear exhibited by each of the datasets and their different representations of the wind profiles".

42) ***Line 406: (mentioned in previous review: seems like a good idea to write that your results are conform what others have found): Bias ERA5 at hub height 0.5 m/s is also what is found on the North Sea in Characterisation of offshore winds for energy applications — Research@WUR and at Cabauw in Energies | Free Full-Text | Dutch Offshore Wind Atlas Validation against Cabauw Meteomast Wind Measurements (mdpi.com). NEWA comparable to ERA5 (at least on the North Sea). Undisturbed winds in DOWA (2008-2018) and WINS50 (2019-2021) are much better than ERA5 (including correlation) and the domain covers most of the Baltic Sea, but hourly data unfortunately not available for 2022 and 2023 when you have the lidar measurements (Home | Dutch Offshore Wind Atlas; WINS50 - Winds of the North Sea in 2050).***

Both references were added, as suggested by the referee, during the first reviewing round. "…ERA5 consistently underestimates the wind speed by approximately 0.5 m s-1 throughout the entire profile,
which aligns with the findings of previous studies (Kalverla, 2019; Knoop et al., 2020; Rubio et al., 2022)".

43) ***Line 408-409: ERA5 usually underestimates the wind speed, this is more pronounced at higher elevations and in the eastern part of the ship track. In contrast, ASCAT mainly overestimates compared to the lidar (typo) measurements.***

i. **If anything: more pronounced in western part of ship track (not eastern) which also makes more sense with prevailing westerly winds (land contamination ERA5 grid surface roughness)**

The mentioned typo was already corrected during the first reviewing round.
It has been corrected though, that the more pronounced underestimation mentioned in the manuscript is visible in the western part (not in the eastern).

44) ***Line 418-419: When comparing the two datasets, ERA5 shows a smaller nRMSE in the majority of the studied region, except in the Eastern area near the harbour in Hanko.***

What is your explanation for this? Does it have anything to do with time of overpass ASCAT, the location characteristics?
An explanation on this has been added in the Section 2.2 of the manuscript: " …except in the eastern area near the harbour in Hanko. This may be attributed to the differing spatial resolutions of the two datasets. In the east of 22 degrees longitude, the finer resolution of ASCAT mitigates the impact of coastal contamination, enabling it to capture local conditions more effectively and consequently leading to a lower average nRMSE in this region. In contrast, the coarser resolution of ERA5 may be insufficient to adequately represent the average wind characteristics in this area. Conversely, in the western part of the studied area, with features more intricate topography and a higher density of small islets within a few tens of kilometres from the mainland shoreline, ASCAT measurements are more susceptible to coastal contaminated. This results in excessively high wind measurements at 10 m, thereby contributing to larger nRMSE values across the whole profile.".

45) ***Line 419-421: When comparing the bias and nRMSE shown by the two datasets, the average absolute bias across the entire region is smaller for ASCAT compared to ERA5 at the three heights considered (see Fig. 13). Differently, as can be observed in Fig. 14, most of the locations reveal a smaller nRMSE for ERA5 than for ASCAT. Bit confusing. I suggest an alternative text: So for all heights considered the bias (compared to the lidar measurements) of the ASCAT-derived wind speeds is smaller than the bias of the ERA5 wind speeds (fig 13), but for most of the region (except for the eastern part of the region near the Finnish coast) the nRMSE of the ERA5 wind speeds is better (fig 14).***
Text excerpt modified for clarity.

46) ***Line 427-428: The objective of this study has been to evaluate the accuracy of ASCAT-derived wind speed profiles for the characterization of offshore wind resources at turbine operating heights in the Northern Baltic Sea.***
   i. **Goal wind resource assessments?**
   As I said before, this work is interesting for wind energy, but only because we can use the ship-based lidar measurements for validation of mesoscale or LES models that include the effect of wind farms. We can then use these models with changed wind farm scenarios to predict the wind resource in the future. Bear in mind that mean values of the wind are not relevant if you want to predict power: you need to look at correlation on a 10 min (or hourly) basis, especially for wind speeds between cut-in and rated (power curve).

No, the goal of this paper is not wind resource assessment, understood as the characterization of wind farms potential. This is discussed in several sections of the manuscript, and detailly explained in response to comment 1b.

47) ***Line 431: … obtained from a (typo) novel ship-based lidar campaign***
This typo was already corrected during the first reviewing round.

48) ***Line 435: … that machine learning-based techniques for extrapolating satellite winds could surpass the long-term correction method employed herein. Questionable English. I suggest an alternative text: … that machine learning-based techniques for***

extrapolating satellite winds could work better than the long-term correction method that was used in this study.

Corrected as suggested by referee.

49) Line 436-437: However, the limited amount of data available over the campaign period hinders the implementation of such data-driven approaches.

i. **Not an ML expert, but is the fact that you have a short campaign really the limiting factor? You have ERA5 and ASCAT measurements for a much longer period, so can you not perform your long-term stability correction? What I do know is that ML cannot reproduce events that have not occurred yet (extremes).**

The comparison presented in this study is specifically focused on the campaign period, as the mean values from the lidar measurements are compared against the corresponding ERA5 mean values and the extrapolated ASCAT wind speeds—both derived for the same time frame. The extrapolation of ASCAT data is based on mean 10 m wind speeds and ERA5 stability information during the campaign period. Therefore, even though the total temporal extension of ERA5 and ASCAT is larger, it is not directly applicable to this study, since the average stability conditions are during this longer period would be different to those present specifically during the campaign.

We believe that the limited campaign duration does indeed present a significant constraint, since ML models are data-based models that require large amounts of data to effectively capture the complex relationships between input and output variables. Additionally, the methodology employed in this study to extrapolate ASCAT wind speeds produces only mean values, without providing a one-to-one, time-correlated comparison between the three datasets. This further reduces the available comparable data, making it even more challenging to develop or apply ML models for this purpose (or triple collocation, because of the same reason).

That being said, there are studies with different scopes and data availability that focus on the development and evaluation of ML algorithms for ASCAT extrapolation, such as (Optis et al. 2021; Hatfield et al. 2023). But ML is out of the scope of this study, since an alternative approach is presented and employed.

50) Line 441-442: The methodology revealed a remarkable congruence between these two approaches across most of the area examined, thus underscoring the robustness of the methodology.

i. **Not convinced this conclusion is justified (see earlier comments).**

Sentence clarified: "These two approaches shown a remarkable agreement across most of the area examined, highlighting the robustness of the mean stability correction approach independently of the strategy selected".

51) Line 443-446: This divergence can be attributed to the limited availability of valid wind speed measurements in the collocated approach, the constraints of considering atmospheric conditions solely during morning and evening hours, and the generic definition of the empirical constants C± required for the calculation of the theoretical stability distributions at each site.

i. **Rethink this conclusion also based on earlier remarks**

We believe, as explained in Section 3.1 that the main factors to explain the differences between the collocated and full campaign approach are those mention here. The sentence has been rewritten for conciseness and clarity:

"This divergence can be attributed to the limited availability of valid wind speed measurements in the collocated approach, which slightly affects the 10 m ASCAT mean values extrapolated, differences in the predominant stability conditions between the two approaches due to the temporal discretization of ASCAT overpasses, and the generic definition of the empirical constants C± for calculating the theoretical stability distributions at each site."

52) ***Discussion: please rewrite given all comments given (running out of time to give detailed comments)***

Discussion revised.

53) ***Line 486-492: Finally, it is imperative to highlight that although the disparities in wind speeds between ASCAT and ERA5 relative to lidar are generally small in far-offshore regions, their cumulative impact over a large-scale wind energy project can still have relevant implications for energy production estimates and financial assessments. Therefore, continued efforts to refine both satellite based measurements and numerical models are essential to enhance the accuracy of wind resource assessments for offshore wind energy applications. The diverse characteristics and insights into wind patterns derived from satellite-derived observations, numerical models, and ship-based lidar measurements suggest that an integrative approach, harnessing the collective strengths of these datasets, could yield substantial gains in the accuracy and reliability of offshore wind statistics derivation.***

ii. **The ASCAT measurements extrapolated to 100m with ERA5 are not representative for wind in or near wind farms and therefore do not give accurate wind resource assessments (neither does ERA5 for areas with wind farms or Measure Correlate Predict for areas where the number/size of wind farms is changing). So what we need to do is further develop Numerical Weather Prediction models that include solving the effect of wind farms (for which we need measurements for validation) and run these models for current and future wind farm layouts. ML is a useful tool, but cannot be used to derive extremes in wind climate.**

You should also bear in mind that there is no significant trend in the wind climate (apart from at 10m over land) but a strong Inter Annual Variability (IAV). This is the case for the North Sea, but most likely also for the Baltic? Do you know? If there is a strong IAV, then it is important to assess how representative the period you look at is for the wind climate. For the Dutch part of the North Sea (DEEZ) the IAV is 3.5 and 4% for sites in the northern part of the DEEZ and between 4 and 4.5% in the southern part of the DEEZ (Inter-annual wind speed variability on the North Sea | Report | KNMI Projects). Is any information like this available for the Baltic Sea? How representative is 28-6-22 until 21-2-23 for the wind climate in the Baltic Sea? This you can check e.g. with ERA5 data (compare ERA5 28-6-22 - 21-2-23 to ERA January1940-now).

So what is the added value of having these 100m wind speeds based on ASCAT? Compared to lidar, the ASCAT derived 100m wind are maybe more accurate than those from ERA5, but only available twice a day. Should we just not assimilate ASCAT in ERA5 and focus more on how useful this ship based lidar technique is to get validation measurements for models including wind farm effects (wakes/blockage)? That is what I like about this work.

The conclusion regarding whether ASCAT and ERA5 are representative of wind conditions in or near wind farms cannot be drawn from the results of this study, as the data presented do not include nearshore wind farms. Therefore, this is a consideration that must be addressed separately when using datasets affected by phenomena like wind farm wake effects. However, while ERA5 and ASCAT do not directly account for such effects, it is still possible to estimate them using physical or numerical models, such as eddy viscosity or machine learning (ML) models, in combinations with these datasets. Although the accuracy of this methodology remains relatively limited, they still provide valuable insight when no other data are available, especially in the absence of in situ measurements. This suggests that while it may not be the ideal or preferred method, it can still serve as a viable alternative in certain situations. Nevertheless, wake effects lie beyond the scope of this study.

As previously mentioned in our response to the referee, inter-annual variability is not a relevant factor in this study, as the analysis only covers data from the same time frame (the measurement campaign period) across all three datasets. We have not conducted any assessment of IAV trends for this specific period or region, as it falls outside the intended focus of this work and does not play any role in the findings presented.

Publication bibliography

Hatfield, Daniel; Hasager, Charlotte Bay; Karagali, Ioanna (2023): Vertical extrapolation of Advanced Scatterometer (ASCAT) ocean surface winds using machine-learning techniques. In *Wind Energ. Sci.* 8 (4), pp. 621–637. DOI: 10.5194/wes-8-621-2023.

Optis, Mike; Bodini, Nicola; Debnath, Mithu; Doubrawa, Paula (2021): New methods to improve the vertical extrapolation of near-surface offshore wind speeds. In *Wind Energ. Sci.* 6 (3), pp. 935–948. DOI: 10.5194/wes-6-935-2021.