"Evaluation of low-level jets in the Southern Baltic Sea: a comparison between ship-based lidar observational data and numerical models"

Hugo Rubio, Martin Kühn, and Julia Gottschall

Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2022-40, 2022

Authors response to reviewer comments

First of all, we would like to thank the referees for their time and effort in reviewing our work. We appreciate their feedback and comments, and we have carefully considered their criticisms to improve and clarify our work.

Below, we addressed all the 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 the new excerpt from the revised version of the manuscript (highlighted in blue) when applicable. Line numbers in the comments refer to the preprint version of the paper, unless the contrary is mentioned explicitly.

Stefan Emeis, Referee #1

Referee #1 major issues:

1) ***Astonishingly, the seminal works of Smedman and co-workers on LLJs over the Baltic Sea in the 1990s have completely been ignored.***
   The references (Högström and Smedman-Högström 1984; Smedman et al. 1996) have been added regarding this issue.

2) ***Two mechanisms are given in the manuscript for the formation of LLJs over the Baltic: (1) advection of nocturnal jets formed over land, and (2) baroclinicity. But one decisive mechanism is missing: the flow transition taking place when air moves from the land to the sea. Especially when warm air moves from rough land to a colder and much smoother sea, a sudden acceleration due to the sudden reduction of surface friction sets in. Smedman and co-workers based their data interpretation on this mechanism.***
   A specific mention to this formation mechanism (and the corresponding reference to Smedman´s work) has been included.
   In addition, frictional decoupling may also appear when relatively warm air flows out over colder waters (Smedman et al. 1993).

3) ***Evaluation of the lidar data in this manuscript is very much biased by two facts: (1) by the limited height of 300 m of the lidar measurements, and (2) by the ferry time schedule which allows for measurements at certain sections of the ship track at very few hours of the day only. Due to the second deficiency, the advantage of a moving lidar (compared to those in fixed positions) nearly completely disappears.***
   The authors agree with the mentioned limitations of the used datesets. However, we consider that these limitations are still compatible with the employment of this data for a comparison against numerical models, as long as a proper consideration of these limitations is involved in the derivation of the results and conclusions. Furthermore, the specific limitations of the employed data is one of

the motivations of the presented work, due to the absence of previous similar investigations trying to address to what extent can we extract meaningful information from such a particular dataset.

Regarding the first constraint, preceding literature (Tuononen et al. 2017; Baas 2009; Pichugina et al. 2017) have proved that the vast majority of LLJ events are located below the 300 m limit used in this study. Additionally, considering the focus of this paper in wind energy applications and the current size of offshore wind turbines, the employed vertical extension of the wind profile used in this study allows evaluating the jet phenomena within the relevant environment (an even higher) in which present wind turbines operate. In order to clarify this to the readers, a further explanation of this limitation has been included in the last paragraph in Section 2.4 of the new version of the manuscript and highlighted in the discussion of the new version of the paper.

As mentioned in the second limitation, the relation time-position of the ship track does not allow deriving the jets´ occurrence and properties at a particular location during a complete daily cycle. However, the high reliability of the observational data retrieved by the ship-based lidar, that covers a wide region (differently to the point-located measuring devices), provides an opportunity to compare models outputs against a reference dataset at different locations. The relevance of this fact is highlighted in Section 3.2, where a comparison of the LLJs characteristics retrieved by the models and the observations is performed, and in Section 3.4, where the potential influence of temporal and spatial shift in models´ performance is evaluated. It must be noticed that the evaluations performed in these sections are not achievable through the employment of a single fixed measurement device, pointing out the usefulness of an observational dataset able to provide wind information as a function of both, time and space. In order to highlight the importance of this fact, this has been clearly pointed out in the discussion of the new version of the paper.

4) The numerical models used in the manuscript have their own intrinsic deficiencies (in this context, the work of Sandu et al. (doi:10.1002/jame.20013) should be read and cited).

Certainly, numerical models have inherent limitations that impede them to more accurately model the wind and atmospheric parameters. However, the goal of our work is not to further investigate the physical deficiencies of the models, but to compare these sorts of datasets against reference observations. This allows evaluating to what extent these models can be applied to derive those LLJs properties relevant for the operation and development of wind turbines under different spatial constraints, as well as analyzing the differences between the two employed reanalyses.
The reference mentioned by the referee has been included in Section 3.2.1 of the new version of the manuscript.

5) At the end of the day this leads to a comparison between limited measurement data and limited model data which does not really makes sense.

We disagree with this statement. Ship-based lidar systems provide reliable and accurate wind measurements, but differently to fixes devices, within a spatially extended region. Therefore, the comparison of the models against this reference dataset allows evaluating the variability in the performance of the numerical models when different spatial effects are involved. Additionally, to the authors knowledge, it is the first comparison of numerical models against ship-based lidar measurements focused in the retrieval of LLJs.

6) ***Given the three above mentioned issues, it is not clear to the reviewer what is the actual purpose of this publication? This publication merely gives a record of lidar measurements onboard a ferry. The above mentioned limitations are partly addressed in the manuscript, but no conclusions are drawn from these facts.***
The abstract and certain parts of the introduction have been modify to more clearly state the goals of the work presented in this paper. Furthermore, the conclusion of the paper has been also rewritten.

Referee #1 minor comments:

1) ***Line 38: „as“ instead of „us“***
We have adjusted this.

2) ***Lines 44/45: extension: lateral or vertical? (If lateral, it seems very small; if vertical, it seems very large)***
We have adjusted this.

3) ***Lines 8-61: the paper Wagner et al. (2019) should be mentioned here again as it is already listed in the list of references***
This reference is actually mentioned in line 52 of the preprinted version.

4) ***Lines 138-149: “newest”: at least a year must be given or even better a citation in order to properly identify the version of ERA5 data (the hint to the ECMWF webpage does not help either as webpages may be updated in future)***
“Newest” refers to the version of the latest reanalysis produced by the ECMWF. Since it may be more clarifying, we updated to word “newest” by “latest” in section 2.2.1. Additionally, a better citation has been included (i.e. (Hersbach et al. 2020)).
The reference to the website leads the reader to the ERA5 Documentation official website, where the ERA5´s known issues are listed, and in particular, the mismatch in the wind speeds between the end of one assimilation cycle and the beginning of the next mentioned in our paper.

5) ***Lines 182-205: the LLJ detection algorithm can only work, if the height of the LLJ core is much lower than 300 m. What happens, if the core height is closer to the uppermost measurement level? This issue has to be discussed. Fig. 8c proves this problem.***
A discussion regarding this issue has been added in the second paragraph in Section 3.2.2 of the new version of the manuscript.

6) ***I suggest that in any new version of this manuscript the section on LLJ formation mechanism is re-written starting with the papers and ideas of Smedman et al. Also a look at a very recent overview paper (most probably it came out after the authors finalized their manuscript) by Schulz-Stellenfleth et al. (2022, DOI: 10.1127/metz/2022/1109) might be useful.***
The missing formation mechanism and the suggested references has been added.

<u>Anonymous Referee, Referee #2</u>

Referee #2 general comments:

1) ***This data set is tricky to analyze. It is too short to do climatological studies or analyze seasonal variation. Additionally, the variation in space is also a challenge. Although the authors make a good effort to address the latter, the data-set is biased in the way the measurements always seem to be from the same location at the approximately at the same time of day. This makes analyzes of temporal variation from one point not possible.***
We would like to thank the referee for recognizing authors´ efforts in this work. However, we consider that the relation time-position of the ship route does not impede a meaningful comparison between the observational dataset and the numerical models. Certainly, ship-based lidar measurements are limited and cannot be used for the derivation of a complete temporal characterization of the wind in a specific location. Nonetheless, and differently to fixed devices, this technology provides an opportunity to evaluate the performance of reanalyses within different spatial constraints though the comparison against a highly reliable observational dataset. Thanks to this, Section 3.2 and 3.4 of this manuscript are focus on evaluating the capabilities of the numerical models in several locations within the ship route and the effect of the temporal and spatial shift.

2) ***However, my main concern is that the main results of the study is the comparison with the reanalyses products. It is not clear what is really novel here that hasn't already been published in similar studies from the same region using the same reanalyses products, which you also cite in the manuscript e.g. Witha et al. 2019 and Hallgren et al. 2020. To be able to accept this manuscript I would like to see some more, other type of analysis trying to get a deeper understanding of the results from the comparison such as: During what conditions do the models perform better/worse? Also adding more evaluation metrics could be useful in this sense. How can one use these results to improve the models?***
Compare to previous similar literature, LLJs modelled by the reanalyses are compared against an observational dataset retrieved through the employment of a non-stationary device. On the one hand, the definition and implementation of a proper comparison methodology is a challenge not previously addressed by the preceding literature. On the other hand, the comparison against non-fixed measurements allows the evaluation of the models performance considering different spatial effects, what is specially interesting for mesoscale phenomena such us LLJs.

Regarding the literature specifically mentioned by the referee, we would like to point out that in (Witha et al. 2019), even though the same observational dataset is used, the LLJ phenomena is not specifically addressed. Additionally, in (Hallgren et al. 2020) numerical models are compared against observational data retrieved at different locations using diverse fixed measuring systems and within non-overlapping time periods.

Finally, we agree with the referee suggestion about the great potential for extending this paper by performing deeper investigations to understand better the results obtained through the presented comparison, that to the authors knowledge, it is the first LLJs comparison using this sort of observational data. However, we also believe that this is not within the scope of this paper, but a potential topic to evaluate in future work. Therefore, we have included a further discussion about the outlook and future research regarding this topic in the conclusion of the new version of the manuscript.

3) ***Discussion of the benefits of using ferry based Lidar would be useful and give examples of these.***
Compare to the traditionally employed fixed devices, the main benefit of the ship-based lidars is its capability of providing highly reliable wind data within extensive regions. In the new version of the manuscript this has been highlighted in the abstract and the introduction.
Additionally, from a technological point of view, the main advantages of the technology (such us its cost efficiency or flexibility) are also discussed in the introduction.

4) ***Illustrative case studies could also be useful e.g .perhaps for some specific synoptic situation where the analysis would benefit from a moving platform. Is it possible to use this type of platform to evaluate models for internal boundary layer? These are just some examples, but this study would require some more along these lines.***
These are definitely interesting suggestions for future work, but we have not considered them within the scope of our paper. Instead, our work focused in the implementation of a first-of-its-kind comparison methodology between the ship-based lidar measurements and the numerical models, in order to analyze the capabilities of two state-of-the-art reanalyses for retrieving LLJs properties under diverse spatiotemporal features.

Referee #2 specific comments:

5) ***Line 3: it is stated that the objective is to evaluate performance of the ship-mounted lidar to investigate LLJ properties along the ship track. However, I can't see that this is presented in the manuscript. The LLJ properties from Lidar measurement are presented, but the performance is not evaluated in any formal sense.***
In order to clarify the aim of our work, we have modified the abstract and the introduction in the new version of the manuscript.

6) ***Line 39: "results are insufficient"***
Sentence has been rewritten for clarification:
However, the limitations of the models due to factors such as a too coarse horizontal and vertical resolution, or the incomplete representation of the physical processes lead to an insufficiently accurate description of mesoscale phenomena.

7) ***Line 51: LLJs in the Baltic Sea have been studied also before the mentioned references. 1984 Högström and Smedman present a first paper where the LLJs formation mechanism is described as an "analogy in space to the classical Blackadar nocturnal jet frequently observed in continental areas". This mechanism is missing in this section. Other studies also followed from the group e.g. Smedman et al. 1995: Spectra, variances and length scales in a marine stable boundary layer dominated by a low level jet, BLM, 76(3):211–232.***
Further references have been added mentioning previous studies focused in LLJs in the Baltic Sea. Additionally, the generation of frictional decoupling due to spatial related frictional decoupling has

been included: In addition, the frictional decoupling may also occur when relatively warm air flows out over colder waters (Smedman, 1993).

8) ***Line 61: "sloping topography" (not sloppy)***
We have adjusted this.

9) ***Line 63: Concerning the Stensrud 1996 reference: I think this was first presented in Holton 1967: The diurnal boundary layer wind oscillation above sloping terrain. Tellus***
This additional reference has been added.

10) ***Lines 69-70: a detail but is there support to say that NEWA is one of the most frequently used re-analyzes products? ERA-5 is for sure one them though.***
We decided to compare ERA5 against NEWA in order to evaluate if the ERA5 downscaling process executed for the generation of NEWA brings further benefits for this application. This has been clarified in the new version of the paper.

11) ***Line 116: "likewise in any" replace with something like "and like any"***
Sentence has been rewritten for clarification:
Additionally to the motion compensation post-processing, a quality check of the lidar observations has been implemented to assure the reliability of the output data.

12) ***Line 118: why was -23 DB limit chosen?***
This is the threshold value recommended by the lidar manufacturer for the used device to maintain an optimal compromise between the data availability and its accuracy.

13) ***Line 121: I suggest replacing "filtered" with "rejected"***
We have adjusted this.

14) ***Line 122: how is this 70% limit different from the 80% limit mentioned on line 121?***
The 80% limit refers to the availability of each hourly-averaged data point, evaluated independently for each height. The 70% refers to the availability over the whole profile, this is, the mean hourly availability considering all the measurement heights. If this mean is below 70%, all hourly values (for all the heights) are excluded from the database. This has been rewritten for clarification.

15) ***Line 125: replace "capture" with e.g. "simulate"***
We have adjusted this.

16) ***Lines 134 and 139: correct reference for ERA-5 Hersbach et al. 2020 https://doi.org/10.1002/qj.3803***
We have adjusted this.

17) ***Lines 146-147: how did you deal with this (mismatch between cycles)***
We did not take any particular measure regarding this, since it is an inherent characteristic of the ERA5 reanalysis dataset.

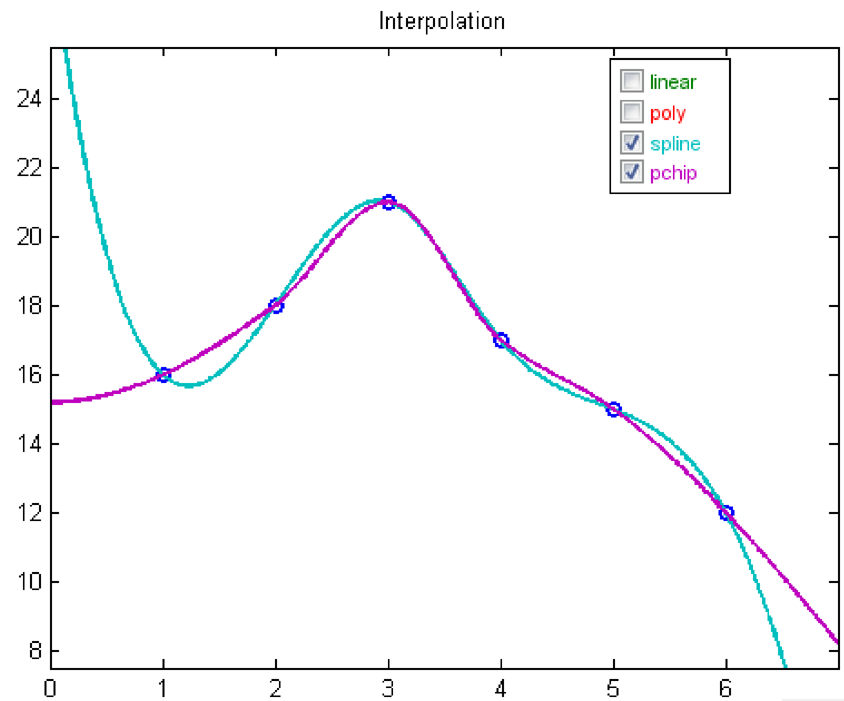
18) Line 155: “spin-off” replace with “spin-up”

We have adjusted this.

19) Line 179: “concentrates” do you mean “conserves”?

No, we mean that the inflexion points of the interpolated line are located closer to the interpolating points. The picture below (extracted from [Splines and Pchips » Cleve’s Corner: Cleve Moler on Mathematics and Computing - MATLAB & Simulink (mathworks.com)]) shows a comparison of pchip against an spline interpolation, with a clear example of this fact for instance, between the points x = 1 and x = 2 and between x = 5 and x = 6.

In the case of a spline interpolation, the maximum curvature is found between the interpolating points, whereas the pchip interpolation results in a line with its curvature at (or closer to) the interpolating points.



20) Line 201: Reference Kalverla: The handling of references should be to place the parenthesis around the year only. This needs to be checked at several places in the manuscript

We have adjusted this along the whole document.

21) Line 202: “extended” replace with “extending”

We have adjusted this.

22) Line 217: comparison with Witha et al: you are using essentially the same data set as Witha et al., please comment on why the results are different.

This has been included in the new version of the paper.

The small differences between the coefficients found in our paper and in (Witha et al. 2019) are due to the different filtering and data quality approaches implemented as well as the specific measurement-models co-location procedures.

23) Line 247: “misestimation” replace with “undererstimation”
We have adjusted this.

24) Line 249: Cheinet at al. year missing
We have adjusted this.

25) Lines 263-264: the onshore daily cycle is well studied as mentioned previously in the manuscript.
Unfortunately, we do not completely understand this comment.
In this lines we highlight that the LLJs´ daily cycle retrieved by the ship lidar measuring while being onshore (in the harbor) agrees with the one founded in previous literature.

26) Line 267: Can you motive the choice of these four locations?
This has been included in Section 3.2.2 of the new version of the manuscript: These locations have been selected aiming to evaluate the datasets in sites with predictably different LLJs´ characteristics (locations A and D can be classified as onshore whereas B and C as offshore) and assuring the existence of a certain amount of jets for the derivation of consistent statistics.

27) Figure 8: does this figure show comparison of co-located model-observations pair in time also, or just the location co-location?
This figure compares the observation and the models when the ship is in any of the considered locations (this is, co-located both in time and space). This has been clarified in the new version of the manuscript.

28) Line 281: “appearance” replace with “occurrence”
We have adjusted this.

29) Section 3.2.2. you use the term “inshore”, “near shore” and “onshore” to describe the same locations, please be consistent.
We have adjusted this.

30) Lines 282-283: not following here, in the previous sentence it is stated that the frequency is overestimated in ERA 5 140% and then it is stated that it is underestimating in this sentence (?). Please clarify.
We missed to state that these lines refer to location C. This has been clarified.

31) Lines 294-295: “The increase in the wind profile” do you mean the extension of the analyzed wind profile height in the models?
Yes, we have addressed this for clarification.
The consideration of the extended wind profiles results in a rise…

32) ***Figure 9: Here you could add the lidar measurements in the shadowed areas. Additionally, this type of analysis would benefit from some more statistics: do the means differ from each other significantly? What is the spread around each averaged point?***
We decided not to add the lidar measurements here due to the fact that for this plot, we directly used the output data from the models (here, there is no filtering process according to lidar quality check or availability). Therefore, we considered that including this data in the same plot with different preprocessing approaches may be confusing for the reader.

33) ***Table 3: Why not add the Lidar measurements in the comparison and present a similar analysis as in Figure 9?***
Same as in comment above.

34) ***Figure 10: Is there any correlation between frequency bias and the forecast length? Are the means statistically significantly different?***
The frequency bias has been evaluated against both the fetch length and the forecast length. However, no correlation has been found so we skipped this comparison from the manuscript.

35) ***Line 359: "alarm" the correct term is "false alarm", you need to correct this at several instances in the manuscript.***
This has been addressed.

36) ***Table 4: spelling "mises" --> "misses"***
We have adjusted this.

37) ***Line 382: Last sentence: please clarify, it's hard to follow the reasoning here.***
We have modified the sentence for clarifying this.
Secondly, the tendency of numerical models to locate LLJs very high in the profile may result in weak jets with fall-off values below the considered threshold (see Subsection 2.4).

38) ***Figure 11: Why not include ERA and NEWA in the same plot? This would make the comparison easier.***
We cannot group the two numerical models in a single plot because the events classified as hits, misses, or false alarms depend on the model selected for performing that comparison. For this reason, the values showed by the lidar are different depending on with numerical model is compared against.

39) ***Figure 12: Other options are also available and should be commented: e.g. interpolate the nearby model data to the measurement location or combining a spatial and time window.***
This has been commented on Section 3.4 of the new version of the manuscript.

40) ***Line 491: One way to investigate how successful the motion correction is would be to study the spectrum of the velocity measurements. If the motion correction algorithm is successfully implemented the peak around the mean wave period should been removed. Although this requires access to the raw turbulence data from the lidar which might not be the case here (?)***

This is an interesting analysis to be conducted when further evaluations of the motion compensation algorithm are performed. However, these investigations are out of the scope of this paper.

Publication bibliography

Baas, Peter (2009): Turbulence and low-level jets in the stable boundary layer. Wageningen University.

Hallgren, Christoffer; Arnqvist, Johan; Ivanell, Stefan; Körnich, Heiner; Vakkari, Ville; Sahlée, Erik (2020): Looking for an Offshore Low-Level Jet Champion among Recent Reanalyses: A Tight Race over the Baltic Sea. In *Energies* 13 (14), p. 3670. DOI: 10.3390/en13143670.

Hersbach, Hans; Bell, Bill; Berrisford, Paul; Hirahara, Shoji; Horányi, András; Muñoz-Sabater, Joaquín et al. (2020): The ERA5 global reanalysis. In *Q. J. R. Meteorol. Soc.* 146 (730), pp. 1999–2049. DOI: 10.1002/qj.3803.

Högström, Ulf; Smedman-Högström, Ann-Sofi (1984): The wind regime in coastal areas with special reference to results obtained from the Swedish wind energy program. In *Boundary-Layer Meteorology* 30 (1-4), pp. 351–373. DOI: 10.1007/BF00121961.

Pichugina, Y. L.; Brewer, W. A.; Banta, R. M.; Choukulkar, A.; Clack, C. T. M.; Marquis, M. C. et al. (2017): Properties of the offshore low level jet and rotor layer wind shear as measured by scanning Doppler Lidar. In *Wind Energ.* 20 (6), pp. 987–1002. DOI: 10.1002/we.2075.

Smedman, Ann-Sofi; Högström, Ulf; Bergström, Hans (1996): Low Level Jets - A Decisive Factor for Offshore Wind Energy Siting in the Baltic Sea. In *Wind Engineering* 29 (3).

Smedman, Ann-Sofi; Tjernstrm, Michael; Hgstrm, Ulf (1993): Analysis of the turbulence structure of a marine low-level jet. In *Boundary-Layer Meteorol* 66 (1-2), pp. 105–126. DOI: 10.1007/BF00705462.

Tuononen, Minttu; O'Connor, Ewan J.; Sinclair, Victoria A.; Vakkari, Ville (2017): Low-Level Jets over Utö, Finland, Based on Doppler Lidar Observations. In *Journal of Applied Meteorology and Climatology* 56 (9), pp. 2577–2594. DOI: 10.1175/JAMC-D-16-0411.1.

Witha, Björn; Dörenkämper, Martin; Frank, Helmut; García-Bustamante, Elena; González-Rouco, Fidel; Navarro, Jorge et al. (2019): The NEWA Ferry Lidar Benchmark: Comparing mesoscale models with lidar measurements along a ship route. In *Wind Energy Science Conference, Cork, Ireland*. DOI: 10.5281/zenodo.3372693.