Answers to the comments of the reviewers for the manuscript

„Airborne measurements for investigating offshore wind farm wakes

and modifications of sea state - benefits and limitations “

Submitted to Wind Energy Science

by

Astrid Lampert et al.

The authors would like to thank the three anonymous reviewers for their detailed comments and suggestions to improve the manuscript. In the following, the comments are provided in italic letters. Each point is answered in normal letters. Changes to the text are indicated by quotation marks. Lines refer to the original manuscript.

__Review 1:__

Line 51: Too absolute, since LES are often used to validate lower-order models and, in theory, can still be employed despite their high computational cost.

We changed the sentence “At the same time, LES cannot be used for resource assessment due to the large computational costs.” to

“At the same time, the deployment of LES for ressource assessment is limited due to the large computational costs.”

Line 91: A brief explanation of the sensor systems would enhance the clarity and interpretability of the presented results.

We added the following information about the sensors:

“The meteorological instrumentation for measuring the high resolution (100 Hz) wind vector, temperature and humidity is installed in the nose boom. For the wind vector, this includes a five-hole probe (Rosemount, USA) and sensors for differential and total pressure (Setra, USA) to determine the airflow angles. The position and orientation of the aircraft are determined with a combination of an inertial measurement unit (IMU, different systems for the Dornier128: iNAV-RQH-1003, iMAR, Germany, and for the CessnaF406: iNAT-RQT-4001, iMAR, Germany) and a Global Navigation Satellite System (GNSS, receiver OEM6, NovAtel, Canada). The wind vector can then be derived by subtracting the air speed vector from the ground speed vector. The temperature is measured by two complementary sensors, one with high accuracy, but slower response time (102DB1AG, Rosemount, USA), one with lower accuracy, but higher response time (102E4AL, Rosemount, USA). Humidity is measured with three different sensors: A capacitive sensor Humicap H233 (Vaisala, Finland), a dew point mirror TP 3-S (Meteolabor, Switzerland), and a fast, but drifting optical sensor (for the Dornier128: Lyman-alpha sensor L-6/HMS-2, Buck Research, USA, and for the CessnaF406: KH-20, Campbell Scientific, USA).

Additional instruments for measuring surface temperature, sea state and solar and terrestrial radiation are installed in the aircraft cabin: an infrared radiation thermometer (KT15.82D in the Dornier128 and KT19.85 in the CessnaF406, Heimann, now Heitronics, Germany), a laser scanner VZ-

1000 (Riegl, Austria, only for the Dornier128), two pyranometers for measuring upward and downward solar irradiance (CMP22, Kipp and Zonen, The Netherlands), and two pyrgeometers for measuring upward and downward terrestrial irradiance (CGR4, Kipp and Zonen, The Netherlands). The radiation sensors are only available for the CessnaF406. More details on the sensors and the accuracies are provided in Lampert et al. (2020) and Lampert et al. (2024)."

Line 113: Why is it suitable to use an already simplified model and the small-angle approximation in this context?

The sentences and the simplified model concerns concern the estimation of the sensitivity of the derived wind speed to a small constant sideslip angle error. As both the sideslip angle and the assumed installation error are small (around 0.3 deg), first order terms dominate and higher order terms are negligible. Furthermore, the use of 30 s median filtering reduces the influence of short-term fluctuations. Hence, the simplified approach is sufficiently accurate for quantifying the magnitude of the systematic error while maintaining a transparent derivation.

We clarified the sentence by writing "The systematic wind speed correction was derived by analytically estimating wind speed pertubation caused by a static installation offset and a small angle-of-sideslip error using a first-order linearisationof the wind derivation equations (see e.g., Bärfuss et al., 2023)."

Line 116: Unit

We corrected the unit of wind speed.

Line 131: How was vertical mixing quantified? How was the airborne campaign designed to capture measurements at multiple altitudes, or were the data obtained from a different sensor?

The results presented here were obtained by in-situ measurements at hub height, plus vertical profiles up to an altitude of 1 km.

We added in Sect. 2: "Based on these in-situ measurements, vertical mixing in the atmospheric boundary layer is identified by fluctuations of the vertical component of the wind speed vector. This allows calculating turbulent kinetic energy, momentum fluxes and heat fluxes."

Figure 1: Explanation for higher wind speeds on coast?

We added in the text: "For a potential explanation of higher wind speed close to the coast, the reader is referred to Sect. 3.4."

Line 149: It seems to suggest a comparison regarding data that the SAR is unable to collect.

We agree that the methods are very different, and motivate the comparison in the text:

"Although the methods of airborne in-situ measurements of wakes and satellite borne remote sensing observations based on synthetic aperture radar (SAR) are highly different and sample on other spatial and temporal scales, the quantification of wakes based on the regular SAR overpasses are an important monitoring instrument, which will be compared to almost simultaneous airborne measurements in the following."

Line 154: Quantification of agreement?

Thanks for the constructive comment. We have added quantitative measures of agreement between the SAR-derived hub-height wind speed and the airborne running-mean wind speed along transects T6–T8. The revised manuscript reports the correlation, the standard deviation of the difference and bias.

"For transects T6–T8, the SAR-derived and airborne wind speeds show an overall agreement with an rms of 0.8 m/s and a correlation coefficient of 0.8. The bias varies between the individual transects (including a change in sign between T6 and T7/T8), indicating that the differences are not systematic across all transects and may depend on local atmospheric boundary-layer conditions. The standard deviation of the SAR and airborne wind-speed differences is about 0.3 m/s."

Figure 3: At what height were the airborne measurements taken? Was it also at hub height?

We included in the figure caption: "Comparison of airborne wind measurements of D-IBUF at hub height"

Line 217: LaTex Citation Error

Thank you, we corrected the error.

Line 297: Are the black lines determined analytically or by manual placement?

We changed the text to: "The area influenced by the wind farm is depicted within the manually inserted black lines"

<u>Review 2:</u>

The manuscript summarizes flight measurements around wind farms. This is a highly relevant and exciting study program with clear added value for the scientific community.

The authors would like to thank the reviewer for this positive judgement.

Regarding the manuscript specifically, it states (lines 64-65) "that are already published, and adds new aspects, which have not yet been published so far." However, it is not clear which results are new and which have been previously published (there is no central place where this is summarized).

It is unclear to what extent it should be seen as a review of the authors' previous work and how much is new.

For figures 5, 6, and 7, it is clearly stated that they are obtained from previous work. Is it correct to assume that Figures 1, 2, 3, 4, 8, and 9 have not been previously published?

We added in the text:

"It summarizes results that are already published, and adds new aspects, which have not yet been published so far: Such new material includes

- the direct comparison of almost simultaneous airborne and satellite-derived wind field in the wake of wind farms (Sect. 3.1)
- a study of changes of sea surface properties (Sect. 3.6)

- limitations in identifying the global blockage effect (Sect. 3.7)"

What does "Powered by Esri" in Figures 1,2, and 4 mean?

This is a requirement by Copernicus for applications that contain a map:

Esri attribution is the requirement to display ***Powered by Esri*** text in *all* applications that use Esri technology. This includes applications that use any ArcGIS API, SDK, service, content, or data. The attribution is required for applications that either contain a map or do not contain a map.

You need to include Esri attribution if your application uses:

- An ArcGIS Maps SDK, ArcGIS scripting API, or ArcGIS software API.
- An open source or third-party library with an ArcGIS service, content, or data.
- ArcGIS services, content, or data directly or indirectly with a REST API.
- ArcGIS services, content, or data but do not display a map.

https://developers.arcgis.com/documentation/esri-and-data-attribution/#esri-attribution

** line 34: including a different visual appearance (Ladenburg, 2009),*

We do not understand this point. The mentioned text and reference are the same as in the manuscript, and it is not indicated what should be changed here.

** Line 115-116: Does the uncertainty depend on the main wind speed, and if so, how strong is this dependence?*

We changed the text to:

"The systematic wind speed correction was derived by analytically estimating wind speed pertubation caused by a static installation offset and a small angle-of-sideslip error using a first-order linearisation of the wind derivation equations. Assuming an absolute angular error of 0.3◦ on the angle of sideslip and a measurement speed of 65 m/s, the wind speed error of the 100 Hz data is ±0.35 m/s on flight legs perpendicular to the prevailing wind direction."

** Line 115-116: What is the uncertainty in the velocity fluctuations?*

The uncertainty for the wind speed is stated in the answer above.

** Figure 2/4: Can you indicate the main wind direction? Is it correct to assume the flight paths are always perpendicular to the shown flight paths?*

We added an arrow to indicate median wind direction in both Fig. 2 and Fig. 4.

We added in the figure captions: "Wind speed measurements along the flight path perpendicular to the main wind direction are colour coded and plotted for every 15 s."

** Figure 2/4: Can you indicate some characteristic times of the measurements in the figure*

We added in the figure captions: "One flight leg of around 45 km length requires a flight time of around 12 min."

** Figure 2/4: How much is the variation in the wind direction over the measurement duration of the flight? What is the potential impact of wind direction variations during the measurements?*

Wind direction variations are potentially affecting the wake, but this dynamic behaviour is beyond the presentation within this paper. The wind measurements therefore denote 4D in-situ data in the form of time series with all its inherent properties.

Basic statistics on wind direction and speed during the measurement legs were added in Fig. 2 and Fig. 4 to illustrate the variability of wind conditions during the flight.

We added in the figure captions: "The arrow indicates the main wind direction. The variability of the wind field during the measurement flight is illustrated by the standard deviations of wind speed and wind direction."

** Figure 2/4: Can you also indicate the wind farm names in this figure? It is a bit confusing that the view of this figure differs from that in Figure 1.*

We added the wind farm names in Fig. 2 and Fig. 4.

** Line 162: Can you comment on the uncertainty in the SAR data due to the extrapolation towards hub-height?*

We thank the reviewer for the comment. We have added a short note on the uncertainty due to the extrapolation of SAR-derived 10 m wind speed to hub height. This uncertainty depends on the assumed wind profile, atmospheric stability, and surface roughness. Although it is reduced by using stability information from the airborne profiles, it remains relevant in and near wind-farm wakes, where turbulence, vertical shear, and sea-surface roughness can be modified.

"One should note that extrapolating from SAR-derived 10 m wind speed towards hub height introduces uncertainty related to atmospheric stability, surface roughness, and the vertical wind profile. This uncertainty is reduced here by using stability information from the flight profiles, but the stability functions used in the extrapolation are still idealised (Garrat, 1994) ."

** Figure 3: How long does each measurement (T6, T7, T8) individually take?*

We added in the figure caption: "One flight leg of around 45 km length requires a flight time of around 12 min."

** Figure 3: Some SAR data seems missing; i.e. the yellow lines have some gaps. Can you comment and explain this?*

We thank the reviewer for pointing this out. The gaps in the SAR-derived curves are caused by quality filtering of unrealistic SAR wind-speed values. Some transects, especially T6, pass close to the Amrumbank West wind farm, where turbine-related SAR artifacts can affect the backscatter and produce unrealistic wind-speed spikes. These spikes were removed before the comparison, which leads to missing points in the SAR-derived curves.

We added in the figure caption: "The gaps in the SAR-derived data are caused by quality filtering of unrealistic SAR wind speed values due to turbine-related SAR artifacts close to the wind farm Amrumbank West."

** Figure 3: What is the turbulence intensity of the wind signal measured by the airborne measurements? What is the approximate uncertainty in this?*

The importance of this figure is to illustrate a comparison of satellite-derived wake data and airborne in-situ data. We further quantified: "For transects T6–T8, the SAR-derived and airborne wind speeds show an overall agreement with an rms of 0.8 m/s and a correlation coefficient of 0.8. The bias varies between the individual transects (including a change in sign between T6 and T7/T8), indicating that the differences are not systematic across all transects and may depend on local atmospheric boundary-layer conditions. The standard deviation of the SAR and airborne wind-speed differences is about 0.3 m/s."

** line 177-178: "In addition, a wind farm not in operation does induce downstream turbulence, even if the wind speed is not reduced significantly" --> This statement seems out of context here. Do you know how large the effect is?*

We agree that the statement does not substantially contribute to the topic. Therefore, we removed the sentence.

** Line 185: "Particularly, the engineering model coupled with WRF—" --> this is unclear: please specify further.*

We agree with the reviewer that this description was too brief. We have revised the text to explicitly clarify the coupling mechanism of the EM-WRF model:

"Particularly, the engineering model coupled with WRF (EM-WRF, Fig. 7) - where a heterogeneous, undisturbed WRF background simulation is used as input to initialize each turbine locally, and turbine wakes are propagated as passive scalars along the atmospheric flow streamlines derived from the WRF data (zum Berge et al., 2024) - proved to be most consistent with the airborne observations, especially under stable stratification."

** Line 188: Statements here are not really in a logical order.*

We changed the sentence to: "All models (see Fig. 7) showed a reduced accuracy in predicting wake position and strength at distances beyond 30 km from the wind farm cluster. Therefore, accurate inflow conditions and proper atmospheric stability characterization are required to improve the performance of all model results over such large scales."

** Section 3.2: This section overall can benefit from better structuring and division in some paragraphs, and some more details to provide context to the reader.*

We sharpened the focus of the section and organised the text as several paragraphs. In particular, we added in the text: "Airborne observations provide the opportunity to focus the measurement strategy on covering overlapping wake effects of multiple wind farms, and to confirm simulations of the wind field."

** Figure 5: Do you know why the effect is observed for only part of the x-range of the wind farm?*

We added in the text:

“The distance of the flight trajectory to the wind farms varies due to wind park geometry and wind direction, resulting in visible effects of the latent heat fluxes only for part of the areas indicated as downstream of the wind farms in Fig. 5.”

** Line 217: A citation is missing.*

Thank you – we corrected the latex citation error.

** Line 232: "measurements are a suitable tool to study these effects of coastal transition." --> Can be removed here: This is more what is generally outlined in abstract, conclusion, and introduction, but does not need to be repeated throughout the manuscript.*

We modified the sentence and put it into the conclusion section: “Airborne measurements can resolve the complex and height-dependent phenomena of wind field modifications for coastal transitions, and such data can help to understand and properly simulate mesoscale processes in offshore wind assessments.”

** Line 254-255: The two lines are very vague and do not form a proper paragraph.*

Thank you for the valuable point. We reviewed the sentence which now sounds as follows:

“Simulation progress has been achieved by coupling atmospheric and oceanic models. Nevertheless, current simulation approaches for representing wind-turbine generated turbulence do not fully reproduce observations, particularly regarding the vertical transport of turbulent kinetic energy and its impact on the sea surface (Larsen et al., 2024)”

** Figure 7: Please increase the font size of labels in this figure.*

Thank you for the hint – we increased the font size of the labels in Fig. 7.

** Figure 8: Can you explain the "red lines" at North displacements of 6 km (49 km panel), 8 km (70 km panel), and 22 km (76 km panel)*

The red lines visible at these three displacement positions are artifacts caused by intermittent disturbances captured during the LiDAR measurements — likely originating from passing vessels or similar moving objects in the field of view. The fact that similar disturbances appear across three different range panels (49, 70, and 76 km) at spatially consistent positions suggests that the source was advected or moving during the measurement period, leaving a coherent signature across the three multiple panels.

We changed the figure caption to: “The horizontal, narrow lines with high energy density can be attributed to boats or other disturbances near the sea surface. The fact that similar disturbances appear across three different range panels (49, 70, and 76 km) at similar spatial positions suggests that the source was advected or moving during the measurement period.”

** Figure 8: What is the black in panels 70 and 76 km?*

The black regions in panels 70 and 76 km correspond to areas of very low spectral energy density, particularly at angular frequencies below approximately 1.5 rad/s. Overall, the spectra at these large distances from the wind park exhibit a more homogeneous distribution, with energy levels predominantly in the green range of the color scale, leaving low-energy regions appearing black due to the applied range.

We included in the text: “At downstream distances of 70 km and 76 km, very low spectral energy density are observed (black colour), particularly at angular frequencies below approximately 1.5 rad/s. Overall, the spectra at these large distances from the wind farm exhibit a more homogeneous distribution, with energy levels predominantly around 0.2 $m^2rad^3s^{-3}$ (green range of the color scale), leaving low-energy regions appearing black due to the applied range.”

** Figure 9: What confuses me in this figure is the mismatch in locations shown in the first panels. I.e Panel (b) indicates latitudes: 54.2 to 54.65, which does not seem to clearly correspond to what is shown in panel (a). Similar for the longitudes in panel c, which do not correspond to panel a.*

Indeed - the mixing of decimal and deg/minute notation is confusing. We changed the tick format as well as the tick/grid positions to enhance readability.

Review 3:

The manuscript gives an impressive overview of airborne measurements. Unfortunately, it is too much descriptive. This may be fine with respect to meteorology, for the wind energy community (wind farm developers, wind turbine designers, wind energy researchers) more specifications are required, e.g. with respect to the airborne measurements:

We would like to thank the reviewer for the judgement. We are happy to improve the manuscript based on the comments below, and include more quantitative information.

• *what is the averaging period of the wind speed (e.g. as shown in Fig. 2)? And what is the typical distance travelled by the aircraft during this time interval?*

The wind data are generally available at 100 Hz. For Figure 2 and 4, the wind speed is averaged over 15 s to reduce fluctuations. The aircraft is operated at an air speed of 65 m/s. Therefore, the aircraft travels a distance of around 1 km for each measurement point.

We added in the figure caption: “Wind speed measurements along the flight path perpendicular to the main wind direction are colour coded and plotted for every 15 s, corresponding to a displayed horizontal resolution of around 1 km.”

• *How accurate are these measurements? (e.g. on basis of the calibration of the instruments and/or comparing with sonic measurements on a static met-mast, when the aircraft has been flown close to it)*

We included a paragraph about sensors and accuracies:

“The meteorological instrumentation for measuring the high resolution (100 Hz) wind vector, temperature and humidity is installed in the nose boom. For the wind vector, this includes a five-hole probe (Rosemount, USA) and sensors for differential and total pressure (Setra, USA) to determine the airflow angles. The position and orientation of the aircraft are determined with a combination of an inertial measurement unit (IMU, different systems for the Dornier128: iNAV-RQH-1003, iMAR, Germany, and for the CessnaF406: iNAT-RQT-4001, iMAR, Germany) and a Global Navigation Satellite System (GNSS, receiver OEM6, NovAtel, Canada). The wind vector can then be derived by subtracting the air speed vector from the ground speed vector. The temperature is measured by two complementary sensors, one with high accuracy, but slower response time (102DB1AG, Rosemount, USA), one with lower accuracy, but higher response time (102E4AL, Rosemount, USA). Humidity is measured with three different sensors: A capacitive sensor Humicap H233 (Vaisala, Finland), a dew point mirror TP 3-S (Meteolabor, Switzerland), and a fast, but drifting optical sensor (for the Dornier128: Lyman-alpha sensor L-6/HMS-2, Buck Research, USA, and for the CessnaF406: KH-20, Campbell Scientific, USA).

Additional instruments for measuring surface temperature, sea state and solar and terrestrial radiation are installed in the aircraft cabin: an infrared radiation thermometer (KT15.82D in the Dornier128 and KT19.85 in the CessnaF406, Heimann, now Heitronics, Germany), a laser scanner VZ-1000 (Riegl, Austria, only for the Dornier128), two pyranometers for measuring upward and downward solar irradiance (CMP22, Kipp and Zonen, The Netherlands), and two pyrgeometers for measuring upward and downward terrestrial irradiance (CGR4, Kipp and Zonen, The Netherlands). The radiation sensors are only available for the CessnaF406. More details on the sensors and the accuracies are provided in Lampert et al. (2020) and Lampert et al. (2024)."

• *How is the turbulence intensity determined? (what is the sampling rate?)*

Actually we do not use the parameter turbulence intensity in the manuscript. We now added information on the sampling rate and measurement uncertainty for the wind vector:

"The meteorological instrumentation for measuring the high resolution (100 Hz) wind vector, temperature and humidity is installed in the nose boom."

"the wind speed error of the 100 Hz data is ±0.35 m/s on flight legs perpendicular to the prevailing wind direction"

• *How is the stratification (heat flux) obtained?*

We added in the text: "Latent heat fluxes were calculated based on the eddy covariance method, correlating the fluctuations of the vertical wind speed component and the fluctuations of the mixing ratio based on measurements of the Lyman-alpha fast humidity sensor. More details of the averaging method are explained in Platis et al. (2023)."

Line 161: 'These include the ability to directly validate satellite-derived wind speed at hub height'; so, specify the (typical) mean and standard deviation of the difference between the SAR and airborne wind speed (Fig. 3)

We thank the reviewer for the comment. Please see the response to the previous comments from reviewer 1 (on the quantification of agreement). For transects T6–T8, the estimated bias between the SAR-derived and airborne wind speeds is about 0.7 m/s. The standard deviation of the difference is about 0.3 m s^{-1}, with an rms of 0.8 m s^{-1} and a correlation coefficient of 0.8.

Line 164: 'Airborne measurements also complement'; how?

We changed the sentence to: "Airborne measurements also extend the spatial coverage provided by long-range scanning wind lidar systems, which are typically fixed at a single location and can retrieve wind field data over distances of up to approximately 15 km, depending on atmospheric conditions"

Line 187: ‘all models showed reduced performance further downstream’; specify

We changed the sentence to: " All models (see Fig. 7) showed a reduced accuracy in predicting wake position and strength at distances beyond 30 km from the wind farm cluster. Therefore, accurate inflow conditions and proper atmospheric stability characterization are required to improve the performance of all model results over such large scales."

Line 188: ‘The study underlines the value of high-resolution airborne data for benchmarking wake models in clustered offshore wind farm’; be more specific: what did airborne exactly add to the traditional way of wake measurements (in other words: what can be done with airborne measurements what is not possible with e.g. a lidar or met-mast).

We changed the sentence to: “Unlike traditional met-masts, which only provide single-point time series, or Doppler lidars, which are limited to a scanning range of about 10 km, in-situ airborne measurements allow capturing the full, large-scale cluster wake developments that can extend beyond 50 km to 70 km under stable atmospheric conditions (Platis et al., 2020). By flying meandering patterns perpendicular to the wind direction at hub height across multiple locations, aircraft data uniquely provide high-spatial-resolution cross-sections of the wind field. This makes it possible to simultaneously evaluate wake deficits, lateral wake extent, and merged wake structures from multiple adjacent wind farm clusters across vast horizontal distances.”

Line 217: Typo: (?Djath et al.

We corrected the latex citation error.

Line 226-228: ‘the dissipation of a land-based LLJ’; ‘a breakdown of the LLJ’; clarify what the difference is.

Thank you for this comment. We agree that the wording was ambiguous and could imply two distinct physical mechanisms. In this section, following the interpretation presented in Section 4.3 of Cañadillas et al. (2023), both expressions were intended to describe the gradual weakening and eventual erosion of the land-based low-level jet as the flow adjusts offshore under changing stability conditions and enhanced vertical mixing. To avoid confusion, we revised the wording and now consistently refer to the “weakening” or “gradual decay” of the LLJ throughout the manuscript.

We changed the text to:” Cañadillas et al. (2023) present two contrasting airborne measurement cases illustrating the offshore adjustment of low-level jets (LLJs) over the German Bight under offshore-flow conditions. In the first case (Fig. 5a; 23 July 2020), warmer sea-surface temperatures offshore (approximately 17,°C at FINO1 compared to about 13,°C over land during the early morning) reduced atmospheric stability over the sea and enhanced vertical turbulent mixing. This promoted the downward transport of momentum and led to a rapid weakening and gradual erosion of the land-based LLJ as the flow adjusted offshore, accompanied by increasing near-surface wind speeds with increasing fetch. In the second case (Fig. 5b; 23 September 2020), the flow over land was initially much more strongly stratified. Although the warmer sea surface gradually reduced the stability offshore, the boundary layer adjustment occurred more slowly, allowing the LLJ to persist farther offshore before weakening. As a consequence, the wind speed decreased with increasing distance from the coast.”

Line 237: 'high-resolution data in both horizontal and vertical directions'; specify

We changed the text to: "Unlike many ground-based methods, they provide high-resolution data in the range of few cm in both horizontal and vertical directions"

Line 250: 'it struggled to accurately reproduce the absolute wind speed'; specify

Thank you for this comment. We agree that the original wording was too general and could suggest an overall poor model performance. However, Cañadillas et al. (2022) showed generally good agreement between the lidar observations and the WRF simulations including wind farm effects (WRF-WF), with mean wind speed differences of approximately 2%. Larger deviations were mainly observed under stable atmospheric conditions. To clarify this point, we revised the text accordingly.

The revised text now reads:

"The WRF model, using the wind farm parameterization by Fitch et al. (2012), was found to capture the horizontal extent of wakes well (Cañadillas et al., 2022). The agreement between lidar observations and WRF simulations including wind farm effects (WRF-WF) was generally good, with mean wind speed differences of approximately 2%, although larger deviations occurred under stable atmospheric conditions."

Line 254: 'The mechanism for creating turbulence for wind turbines does not seem to be constant with height to reproduce observations (Larsen et al., 2024)'; I do not understand this rather odd sentence; is it about ambient turbulence or wind turbine generated turbulence?

Indeed it sounds odd – it is about the mismatch between simulated wind-turbine generated turbulence generation and observations. We reviewed the sentence which now sounds as follows:

"Simulation progress has been achieved by coupling atmospheric and oceanic models. Nevertheless, current simulation approaches for representing wind-turbine generated turbulence do not fully reproduce observations, particularly regarding the vertical transport of turbulent kinetic energy and its impact on the sea surface (Larsen et al., 2024)"

Figure 7: Clarify 'roll mean' and 'running mean values'

The terms 'running mean' and 'roll mean' refer to the same method in this study. Applying a rolling average filters the raw flight data and model outputs, making them highly comparable and significantly improving plot readability. We changed the wording in Fig. 7 accordingly.

Figure 7 top left; it seems odd that the wind speeds in Leg 6 are smaller than in Leg 5;

Specifically, as observed on Leg 6 in Fig. 7, the distinct drop in measured wind speed is not caused by the wind farms, but is rather due to the spatial variability and the closer distance to the coast and/or a temporal change in the weather situation (e.g., cloud or rain formation) during the multi-hour measurement flight, which the steady-state engineering models fail to capture. We added this information at the end of Sect. 3.5.2.

several models (EM-BL, EM-LR and EM-TP) show the same. What is the reason?

The models are initialized with a uniform, undisturbed free-stream wind speed for each flight leg, therefore they simply default to this baseline background velocity at the lateral edges of the flight path or far downstream where the wake has dissipated. Since no wake effect is computed in these regions, the models output the exact same undisturbed background wind speed, causing their curves to overlap.

We added in the text: “Several models provide the same output for specific locations, and the corresponding lines overlap in Fig. 7. As the models are initialized with a uniform, undisturbed free-stream wind speed for each flight leg, they have the same background velocity at the lateral edges of the flight path or far downstream where the wake has dissipated. Since no wake effect is computed in these regions, the models provide the exact same undisturbed background wind speed.”