Response to the comments from referee 1

We thank Elliot Simon and Mathieu Pellé for their positive and constructive comments on our manuscript. We greatly appreciate the detailed feedback. Below, we provide point-by-point responses, with our replies marked in blue.

Response to specific comments

Line 15: *"0.03∘-0.04∘"* Change to 0.03° to 0.04° for consistency.

Has been corrected

Line 49: *"precise alignment of the laser"* Better to define as the positioning of the laser beam or laser light emission.

Thank you for the suggestion. The text has been revised to: *"precise positioning of the laser beam"*

Line 55: *"A target accuracy of about 0.0255° has been shown to be achievable"* The achievable pointing accuracy will also depend on lidar manufacturer and mechanical/optical tolerances within the specific lidar unit.

Thank you for this valuable comment. A subordinate clause specifying the limitations has been added: *"A target accuracy of about* 0.02° *has been shown to be achievable (Rott et al., 2022) and may serve as a benchmark here, although the achievable accuracy depends on the specific lidar system and its mechanical and optical tolerances."*

Line 60-61: *"drone positional uncertainty which can be significant at longer ranges"* This may be true if scanning the drone with CNR mapping as Oldroyd et al. 2024 demonstrated, but when following the method in Thorsen et al. 2023, where the lidar is in staring (LOS) mode and the drone if flown into the beam path, the further away the drone is flown, the smaller the positioning uncertainty becomes.

Thank you for pointing out this study. We have added a paragraph discussing this offshore drone work and its reported uncertainty levels in pitch, roll, and elevation. We have added the following part in the manuscript (Line 63): *"Moreover, Thorsen et al. (2023) demonstrated a complementary drone-based calibration approach in which the lidar operates in staring (LOS) mode and the drone is flown directly into the beam path. Using this technique, substantially lower uncertainties in the estimation of pitch, roll, and elevation offset were achieved, with reported uncertainty values of approximately* 0.05° *for pitch and roll and up to* 0.09° *for the elevation offset during an offshore campaign."*

Line 61-62: *"In recent tests, Oldroyd et al. (2024) reported uncertainties of up to 0.17° under high wind conditions, though values closer to 0.05° may be achievable in calmer conditions"* The following work has demonstrated significantly lower uncertainties in elevation, pitch and roll estimations in both the onshore and offshore campaign, similar to the values suggested as achievable. Thorsen, G. R., Simon, E., & Clausen, E. H. (2023). Drone-based scanning lidar pointing calibration (D4.4). DTU Wind and Energy Systems.

As noted in our response to the previous comment, we have added a paragraph on Thorsen et al. (2023), highlighting this study and its reported uncertainties.

Line 95-100: *"The structure of the paper is as follows..."* This may be a personal opinion, but I find this paragraph unnecessary in a journal paper.

We appreciate the reviewer's comment. Although the co-authors have differing opinions on the necessity of this paragraph, we have decided that, given the length and complexity of the manuscript, providing readers with a brief outline of the paper's structure is helpful for most readers.

Line 111: *"Possible causes include.."* Also due to following errors in the scan head's positioning due to e.g. issues with the motor encoder, hall-effect sensor, mechanical misalignment, or motion control software or software parameters.

Thanks for the additional information. We have added it to the text (Line 115): *"Possible causes include misalignment of the mirrors within the scan head, as well as errors in the scan head's positioning, which may result from issues with the motor encoder, hall-effect sensor, mechanical misalignment, motion control software, or related software parameters."*

Line 136: *"in the context of scanning lidar devices, pitch and roll are usually small enough"* Could you provide a range of values where this assumption is valid?

Thanks for pointing on this, since it is always helpful to give guidance on possible ranges. However, with all possible scanning angles, measurement distances and values for pitch and roll there are too many degrees of freedom to just provide a range of values where the assumptions can be assumed as valid. It also strongly depends on the use case whether a certain deviation of a measurement point is still acceptable or not. We therefore decided to give an example which can be seen as conservative, since quite large values for pitch and roll (absolute values below 1°) and a measurement point far away (10 km) were used and the resulting differences using the two possible definitions for $\mathbf{R_{tilt}}$ are still in acceptable ranges for many use cases. We have added the following (Line 143): *"In fact, if the absolute value of pitch and roll are both below 1°, the two definitions lead to a difference of at most 3.05 m of horizontal and less than 0.03 m of vertical deviation at a measurement point in 10 km distance."*

Line 166: *"SSL is a method of determining the beam alignment calibration of a scanning lidar"* I wouldn't necessarily refer to this as a calibration method, as there is no comparison against an agreed standard reference, e.g. a calibrated theodolite also measuring the distance to the water surface. I would propose calling it a pointing verification method instead.

We thank the reviewer for this suggestion. While SSL does not constitute a traceable calibration against a standard reference, it provides a practical method to verify and adjust the beam pointing of a scanning lidar under conditions where the sea-surface range is effectively known. We have expanded the description in the manuscript to clarify the use of the term "calibration" (Line 174): *"SSL is a method for verifying the beam pointing of a scanning lidar in offshore or nearshore environments. In this sense, it can also be regarded as a practical, in-situ calibration of the laser beam alignment."*

Line 199: *"neglecting h_{earth}"* I don't see this defined earlier. Could this be z_{earth}?

Right, thank you very much! We've corrected it.

Line 216-18: *"The difference is that in RHI scanning, the beam is continuously moved through elevation angles, and the returned signal is an integration between two elevation steps. Therefore a small angular resolution should be chosen. In contrast, PPI scanning is performed at discrete elevation angles in step-stare mode."* Both methods (RHI and PPI) can be performed in continuous scan or step-stare mode. Is this reasoning due to a requirement for multiple elevation angles to be measured? i.e. multiple PPIs at different elevation angles vs. a single RHI at a single azimuth angle?

Thank you for pointing this out. The description in the manuscript was not correct and misleading. In this study, both RHI and PPI scans were operated in continuous scanning mode. The relevant distinction here is not the scan mode itself, but the direction in which the beam is continuously moved and thus integrated. We have adapted the text as follows (Line 228): *"Therefore a small angular resolution should be chosen to achieve high accuracy in the determination of the elevation offset. However, since Eq. (13) is optimised with respect to the elevation angle in degrees, the resulting error is limited by the elevation step size used during scanning. In contrast, PPI scanning is performed at discrete elevation angles, while the laser beam is continuously moved in azimuth. This azimuthal integration has a very small influence on φ_E and thus the distance determination, as long as azimuth intervals are chosen small enough (and pitch and roll angles are small)."*

Line 228-29: *"The inflection point of this function is then considered to be the point at which the laser beam enters the water surface"* Has this been verified experimentally? It would be interesting to carry out a small experiment using a calibrated reference e.g. theodolite/total station to verify the actual distance to the water surface, and compare against the estimate obtained via processing the lidar's CNR values. DTU would be interested in partnering on this exercise.

In fact, the University of Oldenburg has conducted an experimental investigation on this. The results were analysed in cooperation with some of the authors of this publication and will be presented in a paper at the Torque Conference 2026. We added a link to the submitted conference paper in Line 652: *"An improved model of the drop of the CNR-signal could further reduce uncertainties. A more detailed investigation of this will be presented in Meyer et al. (2026)."* (Meyer, P., Rott, A., Schneemann, J., Pauscher, L., Gramitzky, K., and Kühn, M.: Experimental validation of the Sea Surface Calibration for scanning lidar static elevation offset determination, submitted to the Torque Conference, 2026.)

Line 247: *"approx."* would be better written out as approximately. This also appears on Lines 420, and 629.

Has been changed.

Line 249: *"middle of probe volume"* should be "middle of the probe volume"

Has been corrected.

Line 284: *"his model"* should be "This model"

Has been corrected.

Line 295: *"s the water depth"* Is the assumption in the wave model that this is deep water? I cannot find a mention of the water depth used. It appears that the site used in this study has a water depth of 35m.

We thank the reviewer for pointing this out. The wave model does not assume deep-water conditions. The wave celerity is computed using the general linear dispersion relation for surface gravity waves, which is valid for arbitrary water depth. Deep water represents only a limiting case of this formulation (Eq. (17)) and no approximation is applied here. The classification into shallow-, transitional-, or deep-water regimes depends on the ratio between water depth and wavelength. For the approximately 40 m water depth at the study site, the wave conditions encountered during the measurement campaign fall within the transitional to near-deep-water regime, depending on wavelength. The employed dispersion relation remains valid across this entire range. To clarify this point, the water depth used during the measurement campaign has now been explicitly stated in the manuscript (Line 336): *"The water depth during this measurement campaign was approximately 40 m."*

Line 316-17: *"This minimizes the effects of platform inclination caused by the thrust of the wind turbine during the measurement"* We have observed persisting platform motion long (i.e. hours) after the turbine is stopped. Mainly vibrations in higher energy cycles with a zero-mean displacement, although this depends on foundation type, stiffness and sea state. These get averaged out on longer timescales but fundamentally do influence the instantaneous lidar beam positioning. It would be nice to provide guidance on how long the turbine should be stopped for, and an indication of the higher frequency motion within the 10-minute inclinometer data (e.g. standard deviation)

Thank you for this insightful comment. We agree that platform motion may persist after turbine shutdown and that higher-frequency vibrations can affect the instantaneous lidar beam positioning. In our measurements, the inclinometer signals are derived from acceleration-based sensors and are therefore strongly influenced by higher-frequency horizontal/ transverse motions rather than quasi-static tilt. As a result, the standard deviation of the inclinometer data may not a reliable indicator of high-frequency platform inclination and may be misleading. Instead, turbine standstill was verified using the recorded power output, confirming that the turbine had been shut down during the measurements. While general guidance on times scales would be valuable, this is highly site-specific and beyond the scope of the present study.

Line 322-323: *"(RHI 1) ranged from start elevation angle* Φ *start* $-1.5°$ *to end elevation angle* Φ *end* $-0\circ$*"* This may be a typo, RHI 1 appears to end at -0.3 degrees (e.g. Line 378).

Right, has been corrected.

Line 323-24: *"The azimuth resolution for both scans was 5°."* Please define the number of RHI scans in one sequence. This was slightly confusing as I wouldn't expect an RHI scan to have an azimuth resolution.

Thank you very much for the suggestion. The text has been changed as follows (Line 339): *"For one RHI measurement routine, a sequence of RHI scans was performed at discrete azimuth angles between* $180°$ *and* $45°$*, with an azimuthal step size of* $5°$*. The laser beam with azimuth angles outside this range were blocked by the tower."*

Line 328: *"pulse length of 75 m"* This appears throughout the paper. I assume you mean probe length or range resolution as Vaisala call it. Pulse length is the duration of the emission of the laser pulse (normally in nanoseconds). A probe length of 75 m represents a pulse length of 500 ns. Additionally, you may want to state that this is the smallest probe volume (pulse length) option on the 400S. Since other lidar systems (e.g. 100/200S or other manufacturers) support shorter pulse lengths, which may be preferable in the SSL approach.

Thank you for pointing this out and for your suggestion. We should indeed use "probe length" in the text, which we have corrected throughout the script. We have also added to the text that this is the smallest possible pulse length of the WLS200S (Line 346): *"a probe length of* 75 m*, which is the smallest probe volume of a WLS 400S."* and added in Line 643: *"In this study, the smallest available probe length of 75 m for the WLS400S was used. Other lidar systems, such as the WLS100S/200S, offer shorter probe lengths, which may allow improved range determination and potentially reduce uncertainty in the SSL approach. However, shorter probe lengths may also reduce the achievable measurement range."*

Line 335: *"Data filtering"* It would be helpful to indicate how much of the data is being removed in each step of the filtering process.

We thank the reviewer for this suggestion. We have now added quantitative information on the impact of each filtering step. In Line 358: *"In our measurement campaign, the PPI scans cover the entire azimuth range (*$0°$*–*$360°$*), which means that the laser hits the wind turbine and many measurements are filtered out. In contrast, only 2–3% of the RHI measurements are affected, e.g. by the railing."*, in Line 362: *"Depending on the scan type, this step removed about 0.03–2.8% of the measurements remaining after the initial filter."* and in Line 369: *"This filter had a significant effect on our measurement data: on average, around 36% of RHI 1, 29% of PPI and 9.3% of RHI 2 data were removed due to the growth rate filter."*.

Line 346: *"This range should be adjusted depending on meteorological conditions"* What conditions do you expect to influence the performance of this filter? Aerosol concentration?

We thank the reviewer for the comment. The growth rate is influenced by meteorological conditions that affect the range-dependent drop of the CNR signal, in particular aerosol concentration or fog. We have clarified this point in the manuscript (Line 368). *"This range should be adjusted depending on meteorological conditions, that influence the drop of the CNR signal with range, e.g. aerosol concentration or fog."*

Line 349: *"Figure 7a"* If possible, change the way this information is shown because the markers are stacked atop each other and difficult to interpret.

Thank you for the suggestion. In Figure 7a, the markers appear stacked because the effect of an incorrect distance estimation on pitch is extremely small. As indicated by the axis scale (0°–0.001°), the differences between the individual settings are negligible and can effectively be disregarded. Changing the axis scaling would not improve interpretability, as even the current resolution already represents very small angular deviations. We therefore decided to keep the figure unchanged, as it is intended to illustrate the absence of a relevant effect rather than to resolve differences between the individual cases.

Line 436-437: *"It should be noted that SSL measurements are generally not performed under such conditions"* It may not be straightforward for the lidar operator to identify periods with low wave and turbine motion to run the SSL scans. This could be done manually or in an automated manner using forecasts or live measurements if available, however I have often seen SSL scans programmed to execute on a routine schedule regardless of conditions present at the site. There are also sites where these conditions are rarely met, even during installation and decommissioning.

We agree with the reviewer. We have rephrased the sentence in Line 471 to: *"It should be noted that SSL measurements should generally not be conducted under such extreme conditions, where high waves, strong winds, or turbine operation could induce platform tilting and other effects associated with high sea states."*

Line 488: *"Figure 11, "Right axes and dashed line display an external reference measurement"* I would be very interested to see the sub 10-minute variability (e.g. error bars of standard deviation) of the lidar/platform motion. I assume this data is taken from a period where the turbine is shut down for a longer period but this may not be the case.

Thank you for this comment. As discussed in our response to Lines 316–17, the standard deviation of the inclinometer signal may be misleading in this context. The inclinometer data are derived from acceleration-based sensors and are therefore strongly influenced by higher-frequency horizontal and transverse motions in addition to the high-frequency platform tilt. Consequently, short-term variability or standard deviation does not directly reflect residual inclination relevant to the lidar tilting. Over longer periods (e.g. 10 minutes) the horizontal/transverse movements are averaged out. The data shown in Figure 11 were acquired during a period when the turbine was fully shut down, as confirmed by the recorded power output (written in Line 332). For this reason, and due to the limited interpretability of high-frequency inclinometer statistics, we did not include sub-10-minute error bars in the figure.

Line 522: *"several months earlier"* I agree that the elevation offset should not change with time, although I have observed this happening due to mechanical or software faults. Repeating this as a post-campaign check would strengthen the belief that it has not changed.

We agree that repeating the check would strengthen confidence in the stability of the elevation offset. Unfortunately for us, a reliable post-campaign check for this measurement campaign was not performed.

Line 592: *"up to 4000 m a lidar height of 20 m."* "At a lidar height of 20 m?"

Has been corrected.

Line 642: *"The extended SSL analysis code and an example dataset are being prepared for public release."* "It will be greatly appreciated to share the validated processing code with the community and include it together with this publication."

We are very pleased that this has been recognized. The repository with the validated SSL analysis code and a representative example dataset can now be found *https://github.com/kiralg/xssl*. The manuscript has been updated accordingly.

Further changes

- In addition, we have standardised the spelling throughout the manuscript to the British English versions of words such as "minimise," "characterised," "idealised," and "prioritise."
- We have updated the notation of of $\|3°\|$ to $|3|°$.
- The symbol A in the linear term of Eq. (14) was renamed to a, since A is used later to denote the amplitude.
- Renamed s in Eq. (17) to z_{d} for water depth, to avoid confusion with s in Eq. (15).

Response to the comments from referee 2

We thank the reviewer for their positive and constructive comments on our manuscript. We greatly appreciate the detailed feedback and are pleased that the reviewer recognizes the relevance and potential impact of our work. We also welcome the interest in providing readers with the script and example dataset. The repository containing the validated SSL analysis code along with a representative example dataset is now available at *https://github.com/kiralg/xssl* . Below, we provide point-by-point responses to the reviewer's comments, with our replies indicated in blue.

Response to specific comments

Equation 13: The scale analysis leads to a useful and elegant equation. In commercial measurement campaigns, elevation-angle errors are often modeled as a sine curve of azimuth angle[1]. It would be beneficial to link the SSL approach with the traditional hard-target calibration method. This relationship can be derived by taking the partial derivatives of a planar equation in polar coordinates.

Thank you for the suggestion. We have added a link to the established sine curve equation in the paper in Line 216: *"The functional form of Eq. (13) corresponds to the established sinusoidal representation of elevation-angle deviations ($\varphi_{\mathrm{E}} - \varphi_{\mathrm{L}}$) as a function of azimuth commonly used in traditional hard-target calibration (Vasiljevic et al., 2016; Oldroyd et al., 2024)"*.

L172-177: It is interesting that the pitch and roll angles can be determined by the SSL using RHI scans instead of PPI scans. As the authors also analyze, the impact of range detection must be influenced by the ratio between the Lidar height (h_{lidar}) and the range (r_w). In this sense, the accuracy of the range detection appears to depend on the range itself when results from different distances with RHI scans are used.

We thank the reviewer for this helpful comment. The accuracy of the range detection for RHI-scans indeed depends to some degree on the range itself. This is due to the fact that the RHI approach scans over a range interval for each. We use 0.02°. Nevertheless, Eq. (13) is optimised with respect to the elevation angle, which compensates for distance-dependent effects when combining measurements at different ranges. To emphasise this aspect, we have added the following in Line 229: *"Because the beam integrates over a finite elevation interval, the returned signal contains contributions from multiple ranges. However, since Eq. (13) is optimised with respect to the elevation angle in degrees, the resulting error is limited by the elevation angle step size used during scanning but likely to be much smaller."*

L310-334: The measurement campaign settings should be listed in a table for brevity.

Thank you for the suggestion. The table has been added (Table 1).

L347-349: A manual filter was used for data filtering. However, this procedure appears ambiguous and makes it difficult for readers to clearly follow the SSL workflow. To remove outliers more systematically, one possible approach would be to apply CNR-based range detection to the bin-averaged CNR values.

Thank you for your helpful comment. We have revised the section to clarify the visual outlier detection method. We agree that an automatic method for detecting outliers would contribute to standardisation. However, in our workflow, this was an effective method. Establishing reliable and transferable recommendations for data filtering should be based on a larger set of SSLs during varying environmental conditions. To address these points, we have revised the text in the manuscript as follows in Line 371: *"In a final step, outliers were manually screened and removed using a visually guided, iterative reverse filtering procedure. Deviations of individual measurements from a model fitted to the optimised parameters were examined, and data points exhibiting the largest residual errors were successively excluded. In our workflow, this approach proved effective. In the present measurements, this procedure affected approximately 5–10 discrete azimuth angles, depending on the setting. However, a more systematic and automated approach would be advantageous, particularly to improve transparency, reproducibility, and standardization of the SSL workflow, especially for datasets with higher variability. Ideally, this should be evaluated using a larger dataset of SSLs during varying environmental conditions."*

Figure 9: The impact of waves is clearly displayed. However, during scanning Lidar measurement campaigns, the significant wave height is rarely measured. On the other hand, the information of wind speed would be available. Therefore, if the relationship between H_s and 10-minute mean wind speed is provided as reference information, it would make the results easier to understand. For instance, is a 10-minute mean wind speed of less than 4 m/s considered suitable for SSL conditions?

Thank you for the suggestion. The relationship between wind speed and significant wave height is complex. We have added the following to the manuscript in Line 454: *"The relationship between significant wave height and wind speed is complex and non-linear and is influenced by both local wind forcing and remotely generated swell. Under idealised conditions, namely stationary wind, open-sea conditions, and a sufficiently long fetch, a significant wave height of approximately $H_\mathrm{s} = 1$ m can typically be expected at 10-minute wind speeds of about 6-7 m/s at a reference height of 10 m (Holthuijsen, 2010). At real measurement sites, however, additional factors such as water depth, tidal state, wave shoaling, and pre-existing swell substantially affect wave development. As a result, deriving the significant wave height solely from wind speed can be misleading for the SSL user. In practice, measurements or observational data of H_s from oceanographic monitoring systems should preferably be used."*

Figure 11: Φ obtained from the RHI 2 result shows lower values than the other results, whereas the RMSE for RHI 2 is similar to the others. In this case, is the RMSE truly a good indicator for assessing the uncertainty of SSL?

We thank the reviewer for this question. It is correct that the RMSE does not directly represent the overall SSL uncertainty. Rather, it describes how well the measured data is represented by the SSL model. To clarify this distinction, we have added the following statement to the manuscript in Line 497: *"The RMSE quantifies how well the measured values are represented by the SSL model and serves as an indicator of uncertainty: a large RMSE indicates that the data are poorly described by the model, implying increased uncertainty in the estimated pitch, roll, and elevation offset parameters. However, RMSE is not the only contributing factor. Uncertainties in distance determination as well as wave effects also influence* *the overall uncertainty of the SSL measurement."*

L525-532: In this study, α, β, Φ and h_{lidar} were derived using the Levenberg–Marquardt algorithm. Among these parameters, h_{lidar} is expected to be measured directly on site. In our experience, during most scanning Lidar campaigns, the longitude, latitude, and height above mean sea level of the devices are routinely measured. Moreover, if a tidal model is used to estimate the astronomical tide, the time-dependent Lidar height can also be obtained. If h_{lidar} were used as an input parameter for the curve-fitting approach, would the optimization results become more stable?

Thank you for this insightful comment. That is indeed an interesting point. In this study, the lidar height above sea level was not measured. Even measuring the water surface introduces a certain degree of uncertainty. Additional uncertainties are likely even when using a reliable tidal model. Whether these uncertainties are greater due to the optimisation process or the measurements themselves is an interesting question, but one that cannot be answered here. However, we have added the following to the discussion (Line 653):*"In this study, h_{lidar} was treated as an unknown parameter and determined within the optimisation, as no direct measurements of this height were available. While using a known h_{lidar}, derived from on-site measurements and tidal models, which could potentially stabilise the optimisation, these estimates are also subject to uncertainty. Prescribing a fixed (measured) h_{lidar} may lead to stronger propagation of range errors ϵ_{r} to the remaining parameters, whereas treating h_{lidar} as a free variable can absorb part of these errors. Whether a measured or optimised h_{lidar} leads to more stable results therefore depends on the relative magnitude of measurement and modelling uncertainties and cannot be conclusively assessed here."*.

L593-602: In the practical recommendations, a maximum measurement range of up to 4000 m is mentioned. However, a minimum distance would also be expected. It would be helpful to indicate an appropriate measurement range for SSL. Such information would be valuable for applying the SSL method in actual measurement campaigns, helping to reduce variability among users.

We thank the reviewer for this suggestion. However, in contrast to the upper distance limit, where the CNR drop is no longer clearly recognisable at greater distances due to the decreasing signal, we had no problem detecting the CNR drop at the minimum distance limit during our measurement. The limitations at short distances are instead primarily caused by uncertainties in distance determination (see Figure 6a). For the RHI 2 configuration, a distance error of about 37.5 m resulted in an elevation offset error of approximately 0.16°. This uncertainty motivated the recommended elevation angle range of −1.5° to −0.3°, as stated in Line 626. This indirectly gave us a lower distance limit, but not an additional one. Therefore, we decided not to specify a separate lower distance limit here.

Response to technical corrections (optional)

Figure 1: Please consider adding a list of all variables in an appendix for clarity and ease of reference.

We appreciate the reviewer's comment. We have added a list of variables covering the main variables that appear repeatedly throughout the script. Variables that are used exclusively in internal functions have not been listed, as they are local to those functions and do not affect the overall understanding of the work.

Table 1: The unit of h_{lidar} should be "m"

Right, thank you for pointing this out. It has been corrected.

Further changes

- In addition, we have standardised the spelling throughout the manuscript to the British English versions of words such as "minimise," "characterised," "idealised," and "prioritise."
- We have updated the notation of of $\|3°\|$ to $|3|°$.
- The symbol A in the linear term of Eq. (14) was renamed to a, since A is used later to denote the amplitude.
- Renamed s in Eq. (17) to z_{d} for water depth, to avoid confusion with s in Eq. (15).