https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Improving offshore wind data from reanalyses using ship-based lidar measurements

Hugo Rubio[a,b,c], Ville Vakkari[d], Martin Kühn[b,c], and Julia Gottschall[a, e]

[a]Fraunhofer Institute for Wind Energy Systems IWES, 27572 Bremerhaven, Germany
[b]Carl von Ossietzky Universität Oldenburg, School of Mathematics and Science, Institute of Physics
[c]ForWind - Center for Wind Energy Research, Küpkersweg 70, 26129 Oldenburg, Germany
[d]Finnish Meteorological Institute, Helsinki, Finland
[e]University of Bremen, Faculty of Geosciences, 28359 Bremen, Germany

Correspondence: Hugo Rubio (hugo.rubio1@uni-oldenburg.de) and Julia Gottschall (julia.gottschall@iwes.fraunhofer.de)

Abstract.

This study addresses the challenge of integrating ship-based lidar measurements with numerical weather prediction models to improve offshore wind characterisation. Accurate wind measurements are vital for the development of offshore wind energy, yet traditionally used fixed devices, such as meteorological masts and platform- or buoy-based lidars, are expensive and scarce. Ship-based lidar systems offer a flexible, cost-effective alternative by collecting wind data over large areas; however, the non-stationarity of ships results in low data density at any specific location. To overcome this challenge, we propose a novel calibration methodology to assimilate ship-mounted lidar observations into the ERA5 reanalysis by statistically adjusting its wind speed outputs. Inspired by observational nudging, which influences model state variables over time to match observational data, our approach applies a weighted correction directly to the model's wind speed output, preserving the model's underlying physics while ensuring computational efficiency and flexibility. The calibration parameters, including calibration strength, temporal window, and spatial radius of influence, were optimised to maximise the impact and accuracy of the calibration process. The comparison between ERA5 before and after the calibration demonstrates that the methodology effectively reduces the systematic underestimation of wind speeds, particularly in coastal regions where ERA5 struggles with complex flow dynamics. The methodology has been validated against independent measurements from a fixed Doppler lidar system deployed on an island in the northern Baltic Sea, demonstrating the calibration's effectiveness in reducing bias and error spread at this location as well. However, it highlights that the calibration effect is strongly dependent on the distance between the ship and the lidar station, with a bias reduction of 0.2 m s^{-1} when the ship is within 60 km, compared to 0.05 m s^{-1} when considering data within 90 km, as a consequence of the intermittent influence of the ship-based lidar data.

1 Introduction

In recent years, wind energy has emerged as a key player in the development of a green, sustainable energy system. In particular, the prospects of higher and more consistent winds and, therefore, higher energy potential, the large-scale scalability of wind farm projects, the greater availability of space, and the reduced interference with land and human life have fostered the growth

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

of offshore wind energy projects. However, the development of offshore wind farms requires accurate and reliable wind data to support efficient planning, design, and operation. Traditionally, stationary mast-mounted sensors and platform- or buoy-based lidar systems (Clifton et al., 2015; Gottschall et al., 2017; Carbon Trust, 2018) have been used to provide high-quality, site-specific measurements. Despite their accuracy, offshore wind measurements remain scarce due to the high costs and logistical challenges of deploying and maintaining these stationary devices.

A potential alternative to these traditional wind measurement technologies is ship-based lidar systems, due to their simple configuration, cost-efficiency compared to other systems, and flexibility to measure at any desired offshore location, independent of water depth or seabed conditions. In addition, ship-mounted lidars offer a unique advantage: their capability to capture wind data across vast regions, providing comprehensive and accurate information on the spatial and temporal variation of wind conditions. For this reason, these devices have been employed in various applications, including offshore wind farm wake measurements (Wolken-Möhlmann and Gottschall, 2014), validation of numerical models and satellite data (Gottschall et al., 2018; Savazzi et al., 2022; Rubio et al., 2025), and characterisation of mesoscale phenomena such as low-level jets (Pichugina et al., 2017; Rubio et al., 2022). However, the mobility of ship-based lidars, which is a key advantage, also poses a major challenge when applied to offshore wind resource characterisation.Since the ship is in continuous motion, only a limited amount of data is recorded at each location, making it difficult to obtain a complete characterisation of the wind flow at each measurement point. To address this limitation, alternative methodologies are needed to integrate ship-based lidar data into wind resource assessments effectively.

The expensive and time-consuming nature of wind measurement campaigns has motivated the employment of numerical weather prediction (NWP) models as an alternative tool to characterise offshore wind conditions, especially when in-situ observations are unavailable. These datasets typically have a spatial resolution ranging from 1 km to tens of kilometres, providing long-term time series of atmospheric parameters over extensive spatial areas and several vertical levels within the boundary layer (Witha et al., 2019b). However, NWP models have inherent limitations driven by factors such as their temporal and horizontal grid resolution, the physical modelling employed, and the chosen parameterisation options, leading to deviations and uncertainties in the quantities modelled. A well-known NWP model, often used by the wind energy industry, is the ECMWF Reanalysis 5th generation, ERA5 (Hersbach et al., 2020).

In this paper, we address the challenge of integrating ship-based lidar measurements with NWP models, specifically ERA5, to leverage their complementary strengths while overcoming their individual limitations. On the one hand, NWP models provide broad spatial and temporal coverage, mitigating the limited data availability of ship-based lidar measurements. On the other hand, ship-based lidar measurements offer highly accurate wind data that can be used to improve and calibrate NWP model outputs. To achieve this, we propose, test, and validate a novel methodology inspired by data assimilation techniques commonly used in NWP models. Data assimilation integrates observational data into numerical models to improve the accuracy of model predictions (Reen and Stauffer, 2010; Skamarock et al., 2019). Specifically, we build on observational nudging, a form of four-dimensional data assimilation (FDDA) (Liu et al., 2008) by which each grid point within the radius of influence and time window is nudged toward the observations using a weighted average of the difference between the model and the observations.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Unlike conventional observational nudging, which introduces tendency terms in the model equations to adjust the prognostic variables (Stauffer et al., 1991; Deng et al., 2007), our approach directly applies a weighted correction to the model outputs. This method influences ERA5 wind speed predictions over time and space without altering the model's underlying physical equations. Consequently, we refer to our methodology as a multidimensional calibration or simply calibration, rather than a nudging process, as it statistically adjusts ERA5 outputs using ship-based lidar measurements as the reference dataset.

The paper is structured as follows. Following this introduction, Section 2 describes the measurements and numerical model used in the study, the implemented data processing methods, and the calibration methodology. Section 3 presents the main results of this investigation, divided into the optimisation of calibration model parameters, the quantification of the impact of ERA5 calibration, and the validation of the results using an additional dataset not included in the calibration process. Section 4 discusses the main findings and summarises the conclusions of this study, respectively.

2 Data and methods

2.1 Calibration model employed

The calibration methodology employed in this study builds on the observational nudging technique implemented in the Weather Research and Forecasting (WRF) model (Skamarock et al., 2019), an open-source and widely used numerical weather prediction model developed at the National Center for Atmospheric Research (NCAR). Observational nudging in WRF, described in detail by Reen (2016) and Skamarock et al. (2019), introduces tendency terms to guide the numerical model towards observations. In contrast, our approach adapts this concept for post-simulation model output correction rather than adjusting the governing model equations, assimilating observations by applying a weighted correction to model outputs. This enhances model estimates while preserving the underlying physical equations of the numerical model. In this study, only the wind speed was calibrated using the methodology presented below.

The calibrated model values, denoted as Q_{cal}, are computed as:

$$Q_{cal}(x,y,z,t) = Q_{raw}(x,y,z,t) + K\frac{\sum_{i=1}^{N} W_q^2(i,x,y,z,t)[q_0(i) - q_m(x_i,y_i,z_i,t_i)]}{\sum_{i=1}^{N} W_q(i,x,y,z,t)} \quad (1)$$

where Q_{raw} represents the raw and calibrated numerical model variable Q for a specific location, height, and time, K is the dimensionless calibration strength which defines the magnitude of the calibration applied, N is the total number of observations incorporated in the calibration, i is the index of the current observation, W_q is the spatio-temporal weighting function that determines the relative influence of each observation based on spatial and temporal distance between the observation and the current model location (Deng et al., 2007; Dudhia, 2012), q_0 is the observed value of Q, and q_m is the model value of Q interpolated at the observation location.

The weighting function W_q is derived from the product of the horizontal $W_{xy}(x,y,z,t)$, vertical $W_z(x,y,z,t)$, and temporal $W_t(x,y,z,t)$ weighting factors. W_{xy} and W_z define the spatial calibration strength of the model, while the temporal weighting

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

function W_t describes the time period relative to the valid time of the individual observation where the observation is applied and how the strength of this application varies with time.

W_t linearly ramps from 0 to 1 and back to 0 depending on the calibration time window defined by the user (Reen, 2016). The vertical influence has been set very small, so that observations only affect their own height. The horizontal weighting factor W_{xy} is calculated as a function of the radius of influence R_{xy} and the distance between the observations and the grid location D. This study applies the Cressman scheme as the horizontal weighting function, defined as:

$$W_{xy} = \begin{cases} \frac{R_{xy}^2 - D^2}{R_{xy}^2 + D^2} & 0 \leq D \leq R_{xy} \\ 0 & \text{else} \end{cases} \tag{2}$$

The correct tuning of the different parameters that define the temporal and spatial distribution (i.e. the calibration time window for W_t and the radius of influence for W_{xy}), as well as the calibration strength K, is crucial to ensure the accuracy of the model. An inadequate selection of these parameters may compromise the calibration process, either by overfitting the model, leading to excessive deviations, or by applying insufficient corrections, resulting in only a weak calibration effect. To determine the optimal configuration, Section 3.1 systematically evaluates these parameters and establishes the values used in the subsequent analysis.

2.2 Ship-based lidar measurements

In this study, ship-based lidar measurements from a measurement campaign in the northern Baltic Sea have been used. These measurements were obtained in a measurement campaign with a vertical profiler Doppler lidar installed on board the ferry ship *Stena Gothica*, which operates a regular route between Nynäshamn (Sweden) and Hanko (Finland) through the northern Baltic Sea. Figure 1a shows the mean ship route during the measurement campaign. The ferry follows a consistent schedule, completing a full round trip every two days: travelling in one direction on the first day and returning the next. Typically, it remains docked in the harbours during the daytime hours, while most of the transit occurs overnight. Figure 1b displays a time-longitude diagram of the ferry's running schedule, showing the usual longitude at each hour of the day. While minor variations in the route occur, this figure represents the typical schedule.

The campaign was conducted from 28 June 2022 to 21 February 2023, lasting in total 238 days. Data was collected using the Fraunhofer IWES in-house ship-based lidar system (Gottschall et al., 2018), which integrates a WindCube WLS7 v2 Doppler lidar, from the manufacturer Vaisala. The lidar was configured to measure wind speeds at twelve different elevations ranging from 60 m to 270 m above sea level (ASL). The system also includes a motion-tracking unit consisting of an attitude and heading reference sensor and a satellite compass to monitor the vessel's movement and position. Additionally, a meteorological station was also installed to record temperature, pressure, humidity, and precipitation data.

To compensate for the ship's motion, a correction algorithm was applied based on the approach described by Wolken-Möhlmann and Gottschall (2014); Rubio and Gottschall (2022). This algorithm accounts for translational vessel movement while disregarding its tilting, as its effect on wind measurements is typically minimal. To ensure data quality, lidar obser-

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

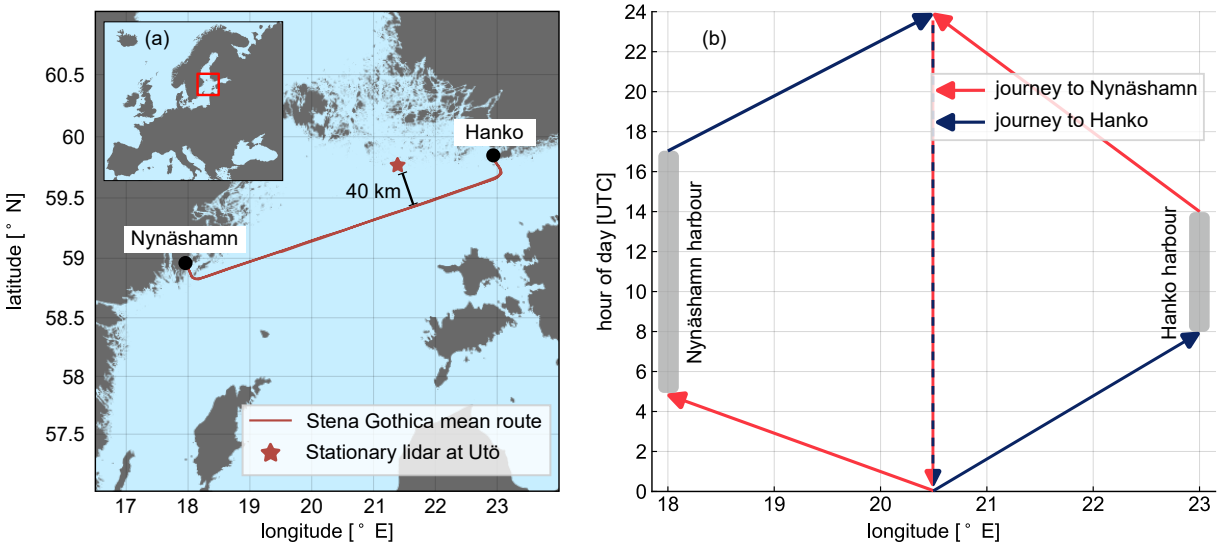

Figure 1. (**a**) Average route of the *Stena Gothica* during the measurement campaign. The location of the stationary lidar at Utö (see Subsection 2.3) is also indicated. (**b**) Time-longitude diagram of the ferry schedule, with the average longitude of the ship at each hour of the day. The periods when the ship is at the harbours are indicated by the grey area.

vations with a carrier-to-noise ratio (CNR) below -23 dB were excluded, following the manufacturer's recommendations to maintain an optimal compromise between data availability and accuracy (Vaisala, 2022). The resulting 10-min ship-based lidar observations were used for the calibration of ERA5 data. Table 1 presents the mean wind speed and the final availability for each measurement height. Additional details and statistical analyses of the measurement campaign are provided in Rubio et al. (2025).

Table 1. Mean wind speed and data availability of processed wind time series per measurement height.

Measurement height (ASL) [m]	**60**	**80**	**90**	**100**	**110**	**120**	**130**	**150**	**170**	**190**	**220**	**270**
Mean wind speed [m s^{-1}]	7.7	8.0	8.1	8.3	8.4	8.5	8.6	8.9	9.2	9.5	9.9	10.5
Data availability [%]	72.4	77.2	79.7	81.7	82.5	82.5	81.4	77.0	71.1	64.2	53.4	35.4

2.3 Stationary lidar

A Halo Photonics Stream Line scanning Doppler lidar (Pearson et al., 2009) is located permanently at Utö island at 59.779 °N, 21.374 °E, 8 m ASL as part of the Finnish ground-based remote sensing network (Hirsikko et al., 2014). The radial resolution of the Halo Doppler lidar is 30 m; the three lowest range gates are discarded due to effects by outgoing pulse. In October 2017, the Halo Doppler lidar laser and amplifier were upgraded to the more powerful XR version.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The Halo Doppler lidar is scheduled to operate two velocity azimuth display (VAD) scans at 4° and 15° elevation angles, respectively, at every 15 minutes. The radial measurements were post-processed according to Vakkari et al. (2019) and filtered with an SNR-threshold of 0.005. Here, we utilise horizontal wind speed profiles retrieved from the 15° elevation angle VAD following Browning and Wexler (1968), which leads to 7.8 m vertical resolution with lowest usable measurement at 35 m ASL. The mean wind speed vertical profile and availability up to 300 m height during the campaign period are shown in Figure 2.

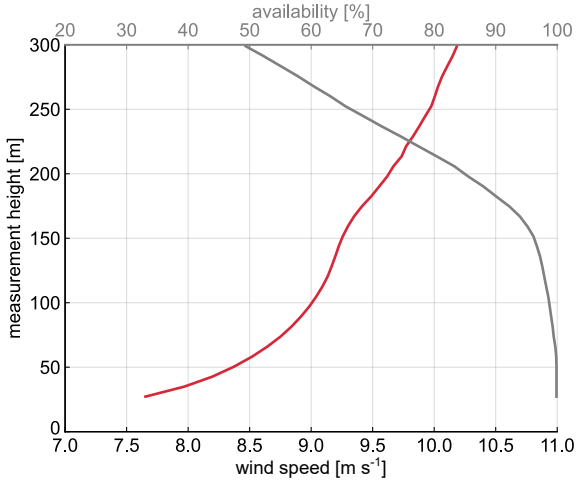

Figure 2. Mean wind speed (red) and availability (grey) vertical profiles from the stationary lidar at Utö during the ship-based lidar campaign period.

2.4 ERA5

ERA5 is, at the time of writing, the latest global atmospheric reanalysis dataset developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach et al., 2020). It replaces the ERA-Interim reanalysis (Dee et al., 2011) and is based on ECMWF's Integrated Forecasting System (IFS) Cycle 41r2. ERA5 provides hourly data for a wide range of atmospheric, land surface, and oceanic parameters on a 0.25° × 0.25° grid (approximately 17 × 31 km in the Baltic Sea), covering the period from 1950 to the present. The dataset includes 137 vertical model (pressure) levels, spanning from the surface up to approximately 80 km (0.01 hPa). The reanalysis is produced using a four-dimensional variational (4D-Var) data assimilation system (Bonavita et al., 2016), which blends numerical model outputs with observations from multiple sources, such as satellite-based instruments, radiosondes, and aircraft measurements.

For this study, the lowest 10 model levels of the horizontal wind speed components were extracted to determine the speed and direction of the wind for the entire study region. It should be noted that only ERA5 data within the time frame of the ship-based lidar measurement campaign have been used in this study.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

2.5 Comparison of datasets

To account for differences in temporal, vertical, and spatial resolution among the datasets used in this study and to provide a fair comparison, several data processing steps were implemented. First, ship-based and stationary lidar observations were averaged to hourly values using a block-average approach with a 1 h time window centred at each hour, so that each hourly value was calculated from measurements recorded half an hour before and after the corresponding timestamp, matching the temporal resolution of ERA5. For each measurement height, hourly values with availability below 80 % were rejected.

After time-averaging, ERA5 data were interpolated to lidar measurement heights using a piecewise cubic Hermite interpolating polynomial (PCHIP) (Fritsch and Carlson, 1980; Brodlie and Butt, 1991). This interpolation methodology preserves the shape of the wind profile by concentrating the curvature of the interpolated line near the data points, thereby preventing the common oscillations associated with spline interpolation. Consequently, the calibration was applied directly to the interpolated ERA5 data without the need for any vertical weighting.

For comparing ship-based lidar measurements with ERA5 data along the ship's route, the collocation strategy outlined by Rubio et al. (2022) was adopted. This method selects the nearest ERA5 grid point for each hourly ship position, making sure that each lidar measurement is paired with the closest available ERA5 wind data. Similarly, for the validation presented in Section 3.3, the stationary lidar measurements were directly compared against both the raw and calibrated ERA5 data from the nearest grid point.

To assess the calibration's effectiveness and compare ERA5's performance before and after calibration, several well-established metrics (Yang et al., 2021) were used. These included the mean absolute error (MAE), the root mean square error (RMSE), and the mean bias. Each metric provides particular insights, enabling a comprehensive analysis and evaluation of the calibration's performance.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|u_{i,m} - u_{i,o}| \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(u_{i,m} - u_{i,o})^2} \tag{4}$$

$$\mathrm{Bias} = \frac{1}{N}\sum_{i=1}^{N}(u_{i,m} - u_{i,o}) \tag{5}$$

Where N is the size of the dataset, $u_{i,m}$ represents the i-th wind speed from the ERA5 model, and $u_{i,o}$ is the corresponding observed (measured) wind speed.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

3 Results

This section contains the main results of this study. First, 3.1 investigates the values to use for the different weighting factors that provide optimal performance for the calibration model and are used for the subsequent results presented. Then, 3.2 highlights differences between the raw and calibrated ERA5 model outputs and the effect of calibration, depending on factors such as location, height, or time of day. Finally, 3.3 presents a validation of the calibration by comparing ERA5 before and after calibration with measurements from the stationary vertical profiler lidar, which was not used to calibrate the model itself.

3.1 Optimization of model parameters

This section analyses the performance of the optimal model configuration by evaluating different values of the main weighting factors: K, W_t, and W_{xy}. To achieve this, a subset of 720 ERA5 timestamps, representing one month of data, was randomly selected. These timestamps were calibrated according to the methodology described in Section 2.1 using different values of the weighting parameters. Subsequently, several statistical metrics (defined in 2.5) were calculated and compared across the different model configurations.

The model's performance was first evaluated across different values of the calibration strength parameter K. Figure 3 presents the key performance metrics at the 12 available lidar measurement heights for varying values of K.

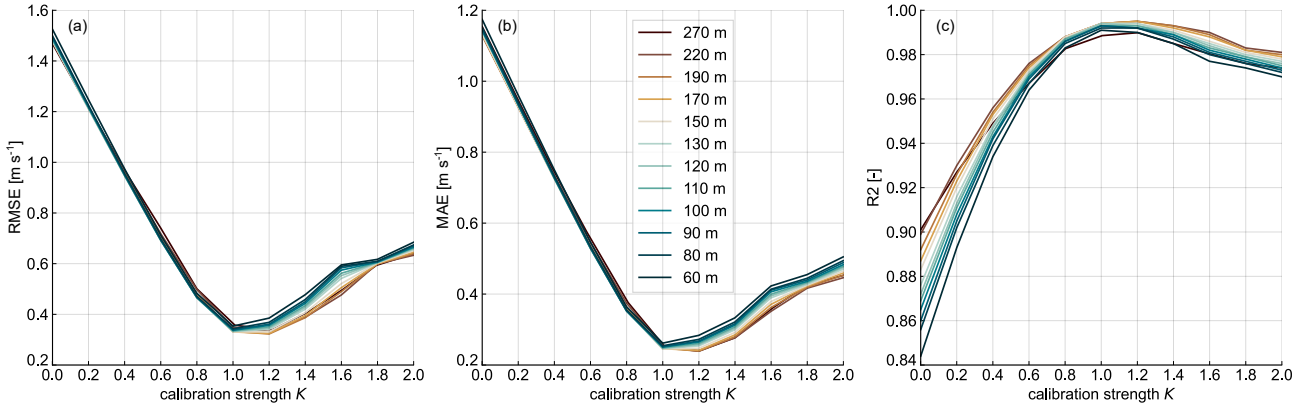

Figure 3. Primary error metrics: (**a**) Root Mean Square Error (RMSE), (**b**) Mean Absolute Error (MAE), and (**c**) Coefficient of Determination (R^2), evaluated for different values of calibration strength (K) at all lidar measurement heights. Metrics for K = 0 correspond to the case without any calibration applied ($ERA5_{raw}$ vs lidar).

The results show a consistent trend across all height levels, regardless of the metric considered. RMSE and MAE are highest at lower calibration strengths, highlighting the still relatively high discrepancy between the calibrated ERA5 data and the lidar observations, due to a very weak calibration effect. R^2 values are also low, reflecting the poor correlation between the two datasets. The optimal performance occurs between K = 1.0 and K = 1.2. In this range, the RMSE stabilises around 0.38 m

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

s^{-1}, the MAE around 0.26 m s^{-1}, and R^2 values also reach their maximum values for this range of K, indicating an optimal and well-balanced integration of the observational and model information.

Beyond K = 1.2, both the error metrics and R^2 gradually deteriorate. For such values of K, the influence of lidar measurements on the calibrated data is strongly amplified, including any noise present in the data, resulting in overfitting of the calibrated ERA5 dataset to the measured data. This overfitting introduces excessive variability and discrepancies, leading to higher RMSE and MAE values and lower R^2, indicating the degradation of model performance and generalisation ability.

The optimal value of K appears to vary slightly with height. Heights below 150 m tend to perform best at K = 1.0, while higher elevations achieve slightly better results at K = 1.2. However, this difference is minimal and does not significantly affect the overall performance. Given these findings, we adopt K = 1.0 as the preferred value, as it provides the best overall performance across a wider range of heights, particularly those most relevant to wind energy applications.

After setting the calibration strength to K = 1, we investigated how varying the temporal and spatial weighting parameters affects the model's calibration effectiveness. Figure 4 presents colour maps for the three statistical metrics evaluated, showing their variations across different temporal windows (ranging from 0.5 h to 10 h) and spatial radii of influence (spanning from 15 km to 300 km). The results are presented at measurement heights of 60 m, 110 m, and 170 m.

Increasing the spatial radius of influence to 90 km significantly improves model performance across all three statistical metrics. RMSE and MAE values decreased sharply, indicating reduced error, while R^2 values increased, reflecting a better fit between the calibrated ERA5 data and the ship-based lidar observations. However, further increasing the radius of influence beyond 90 km yields no substantial improvement in any of the metrics and heights evaluated. This behaviour is mainly explained by the characteristics of the Cressman spatial weighting function employed in this study. Initially, increasing the radius of influence includes additional lidar data points that are closer and more strongly correlated with the target grid point of the ERA5 data being calibrated, therefore improving the calibration effect. However, as the radius of influence expands beyond 90 km, including more distant data points that are spatially decorrelated with the wind measurements at the target calibration location has a limited impact on the calibration. Since these decorrelated points contribute less meaningful information, their inclusion leads to a plateau in the performance of the statistical metrics.

In contrast, the temporal weighting exhibits a different pattern. The model achieves optimal performance with a temporal window of 0.5 h, whereas higher values result in rapid deterioration in calibration effectiveness. This happens because increasing the temporal window introduces observations that are less representative of the current wind conditions and their temporal variations at the target locations, leading to temporal decorrelation of the calibrated dataset and additional noise that degrades the model's performance. The importance of temporal representativeness and the predominantly local-in-time impact of lidar-based corrections have also been reported in previous studies (Ivanova et al., 2025).

Therefore, the calibration model used in this methodology has been investigated to perform optimally with K = 1, a temporal window of 0.5 h, and a horizontal radius of influence R_{xy} of 90 km.

3.2 Effect of calibration on ERA5

This section compares ERA5 raw and calibrated data with ship-based lidar measurements during the campaign period.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

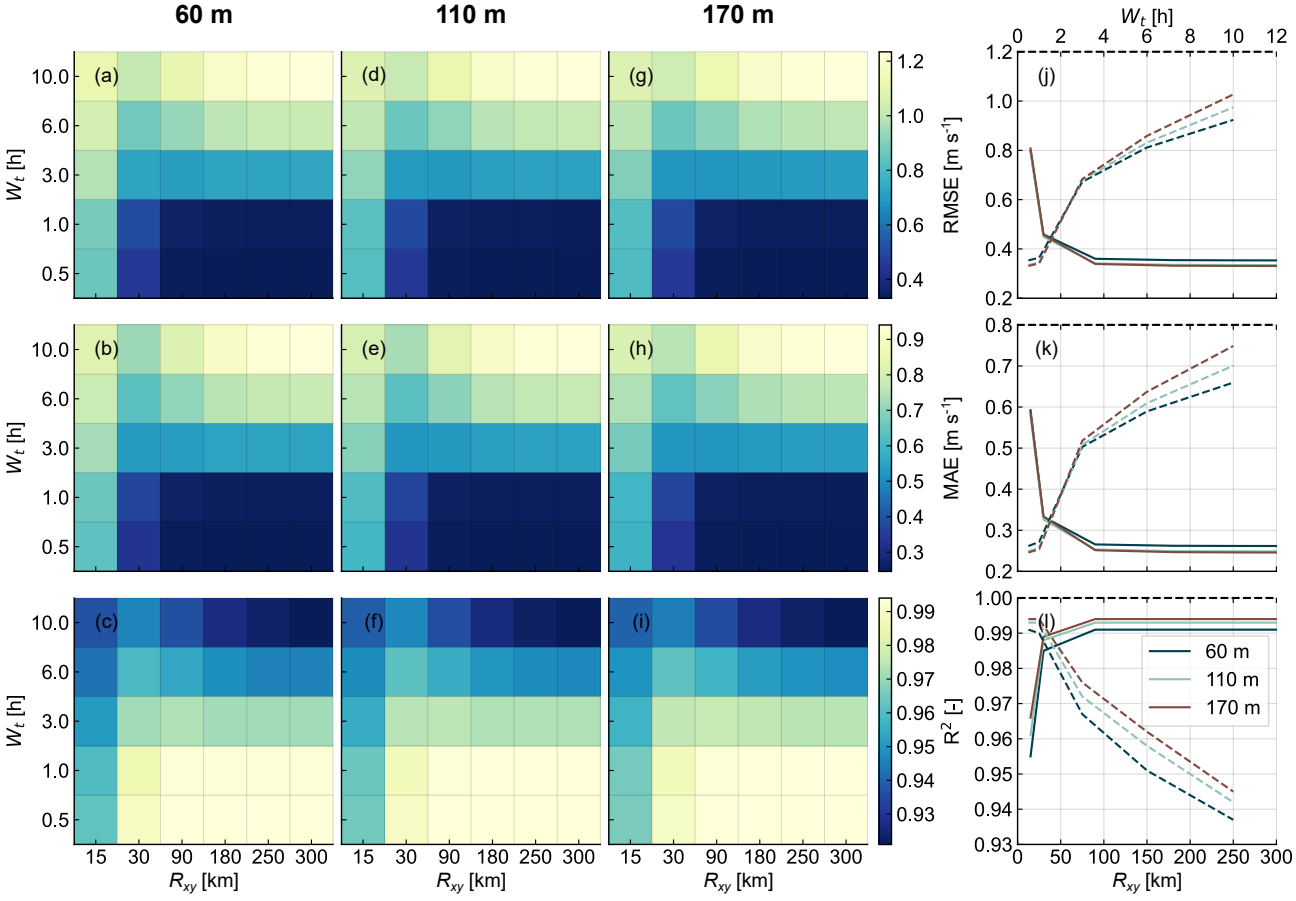

Figure 4. (**a**-**i**) Heatmaps showing the RMSE (**a**,**d**,**g**), MAE (**b**,**e**,**h**), and R^2 (**c**,**f**,**i**) at heights of 60 m (**a**,**b**,**c**), 110 m (**d**,**e**,**f**), and 170 m (**g**,**h**,**i**) for different values of temporal windows (W_t) and spatial radii of influence (R_{xy}). Panels (**j**,**k**,**l**) illustrate the variation of these metrics with a fixed W_t (solid lines) or a fixed R_{xy} (dashed lines) for the corresponding heights.

Figure 5 presents scatter plots comparing lidar-measured and ERA5 horizontal wind speeds at three different heights, covering the entire duration of the measurement campaign along the ship route. The continuous red line represents the linear regression, with its output parameters and the corresponding coefficient of determination displayed in the bottom right corner of each subplot. The dashed line indicates a 1:1 relationship for reference, while the colour of the scatter points reflects the frequency of occurrence.

The analysis demonstrates that, at all heights, the calibration process substantially improves the agreement between ERA5 and lidar measurements along the ship-based lidar route. Specifically, prior to calibration, ERA5 shows moderate agreement, with R^2 values between 0.84 and 0.88 and regression line slopes between 0.90 and 0.95. These values indicate a systematic underestimation of wind speeds, consistent with previous studies (Kalverla, 2019; Knoop et al., 2020). After the calibration, R^2 values approach near-perfect levels (0.986 to 0.992), and the regression slopes are closer to 1, effectively minimising bias

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

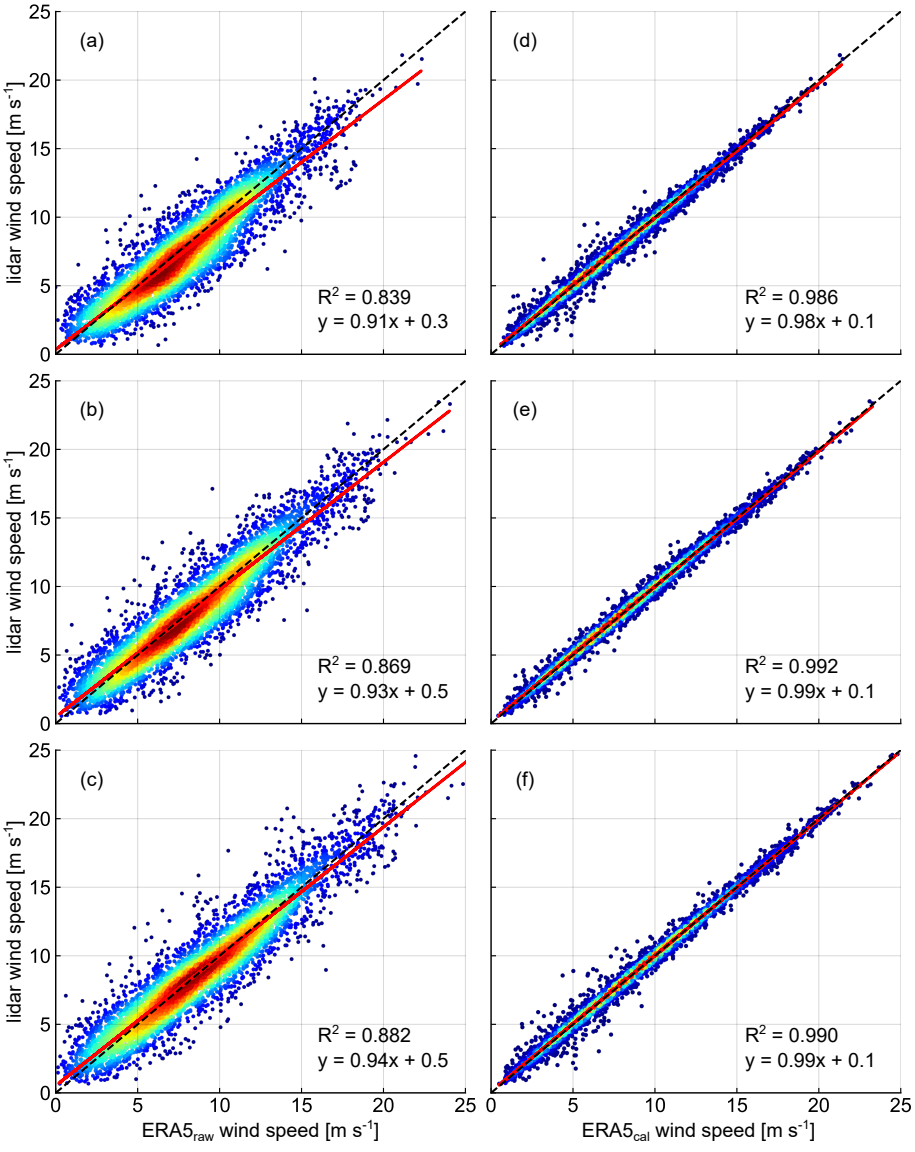

Figure 5. Linear regression of ship-based lidar wind speed against ERA5 raw (**a**, **b**, **c**) and ERA5 calibrated wind speed(**d**, **e**, **f**) at 60 m (**a**, **d**), 110 m (**b**, **e**), and 170 m (**c**, **f**).

and enhancing the accuracy of ERA, as expected, since the same observational dataset was used for both the calibration and the presented comparison.

The spatial impact of the calibration model on ERA5 data is illustrated in Fig. 6, which displays colour maps depicting the average wind speed difference between ERA5 raw and calibrated ($ERA5_{raw}$ - $ERA5_{cal}$) for the duration of the campaign at heights of 60 m, 110 m, and 170 m. Furthermore, Fig. 6a shows the mean wind speed bias per longitude bin between the $ERA5_{raw}$ data and the ship-based lidar measurements ($ERA5_{raw}$ - ship-based lidar).

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

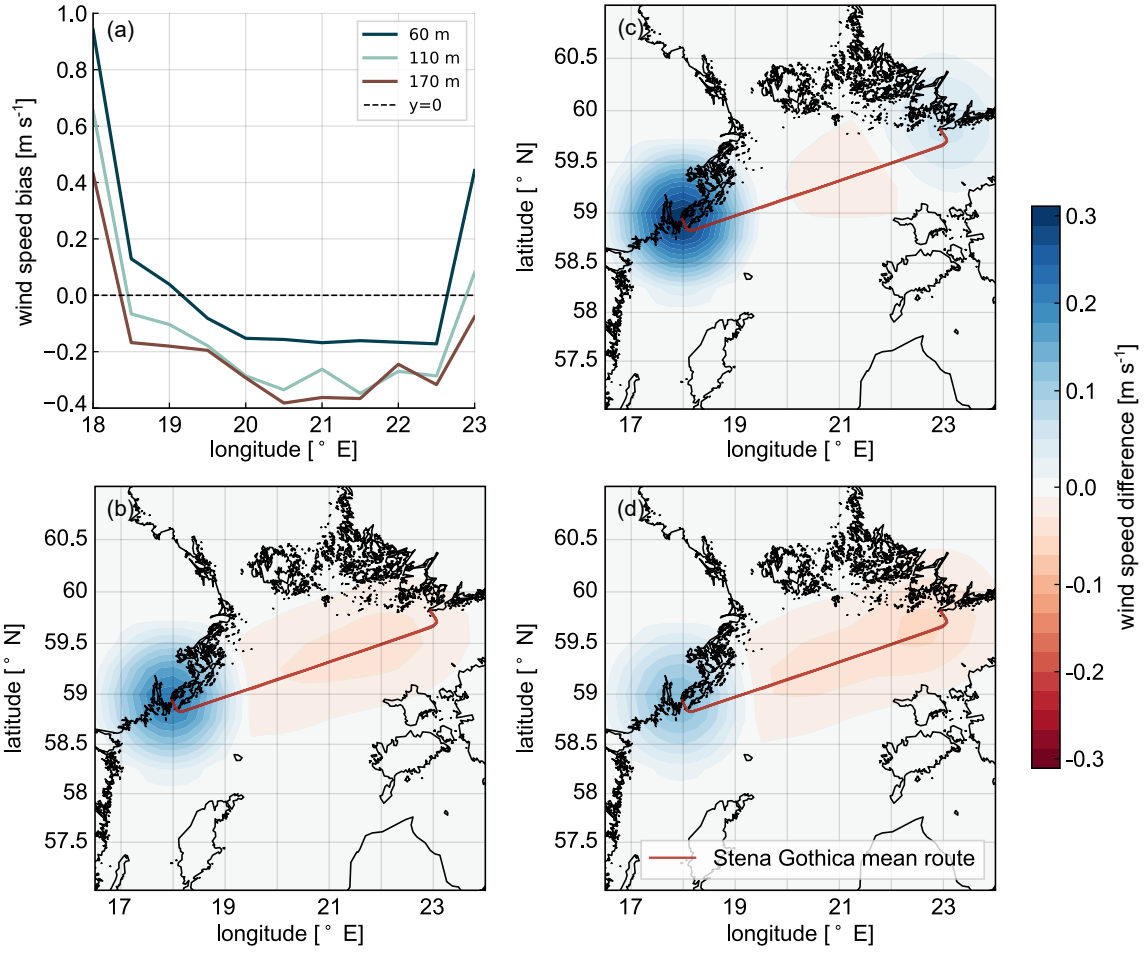

Figure 6. (**a**) Mean wind speed bias (ERA5 - ship-based lidar) per longitude bin. (**c**, **b**, **d**)Mean wind speed difference between $ERA5_{raw}$ and $ERA5_{cal}$ ($ERA5_{raw}$ - $ERA5_{cal}$) across the area of study for the entire duration of the measurement campaign at 60 m, 110 m, and 170 m, respectively.

The calibration effect exhibits considerable spatial differences. The westernmost region of the study area, near the Nynäshamn harbour, shows the most significant calibration effect at the three heights presented, mainly due to the higher deviation of the raw ERA5 data compared to the lidar measurements as observed in Fig 6a, and the larger amount of data in this area due to the ship´s stationary periods in the harbour. In this region, the raw and calibrated ERA5 difference follows a circular pattern, with the calibration effects diminishing while moving further from the Nynäshamn harbour. This pattern is caused by the Cressman horizontal weighting function used in this study, which reduces the calibration strength at ERA5 grid points farther from the measurement location, and by the regular track followed by the ship on every trip, which concentrates the impact of the calibration in the near surroundings of the ship track.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The overall calibration effect is strongly correlated with the bias trend observed when comparing $ERA5_{raw}$ and lidar measurements at different longitudes. In the western region near the coast, $ERA5_{raw}$ consistently overestimates wind speed relative to ship-based lidar measurements. This discrepancy is mainly due to ERA5's inability to accurately simulate the complex coastal atmospheric dynamics and small-scale wind-flow variations in these regions, largely because of its coarse resolution (Dörenkämper et al., 2015; Gualtieri, 2021). Consequently, the calibration model decreases the mean ERA5 wind speed, showing a positive difference in the ERA5 before and after correction comparison, with maximum differences reaching values of approximately 0.3 m s^{-1}, 0.2 m s^{-1}, and 0.1 m s^{-1} at 60 m, 110 m, and 170 m, respectively. This indicates a greater influence of surface heterogeneities and small-scale processes at lower altitudes, penalising ERA5's accuracy while allowing the calibration model to have a greater impact through the inclusion of high-resolution, accurate measurements. A different trend is observed for longitudes east of 19°, where $ERA5_{raw}$ tends to underestimate lidar measurements, resulting in negative values of the mean wind speed difference in the colour maps. These negative values persist along the route to Hanko harbour in the easternmost region of the ship route at 110 m and 170 m. For 60 m, there is a positive difference between ERA5 raw and calibrated around Hanko harbour, coinciding again with the more prominent positive bias observed in ERA5 raw against lidar at this altitude and location.

Transition areas can also be identified where inflexion points in the ERA5 raw and calibrated differences occur, transitioning from positive to negative (or vice versa) biases. In these zones, the difference between ERA5 raw and correlated is minimal, as grid points are influenced by both positively and negatively biased lidar measurements, leading to an overall mean difference close to zero. However, although the averaged calibration effect is small, it does not necessarily imply the calibration effect is negligible, as it may result from an averaging effect that compensates the ERA5 upwards and downwards calibrated data.

To evaluate the temporal effect of the calibration, Fig. 7 illustrates the daily cycle of wind speed differences between ERA5 raw and calibrated ($ERA5_{raw}$ - $ERA5_{cal}$) at three different heights. The mean distance from ship to shore per hour is depicted by the grey bars.

As shown in the figure, all three heights exhibit very similar trends, with larger, positive deviations during the central part of the day, from 05:00 UTC to 18:00 UTC. During this period, mean differences reach 0.6 m s^{-1} and 0.8 m s^{-1}, depending on the height considered. The largest difference between the two ERA5 datasets is observed at 60 m, indicating a stronger calibration effect at this height compared to the others. This central part of the day corresponds to the periods when the ship is mostly near the two harbours, in the westernmost and easternmost areas of the ship track. As previously mentioned, these are regions where ERA5's coarse resolution struggles to capture the complex coastal flow phenomena. This is consistent with the results presented in Fig. 6, which showed a stronger effect of the calibration in these coastal areas, where ERA5 raw showed the largest deviations relative to the ship-based lidar measurements and where more lidar data are available and included in the calibration model.

During the rest of the day, predominantly at night, the deviation changes sign, showing negative values (or almost zero at 60 m) from 19:00 UTC to the end of the day, with smaller differences. This coincides with the period when the ship is primarily travelling between the two harbours, resulting in smaller differences between ERA5 raw and the lidar (see Fig. 6a).

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

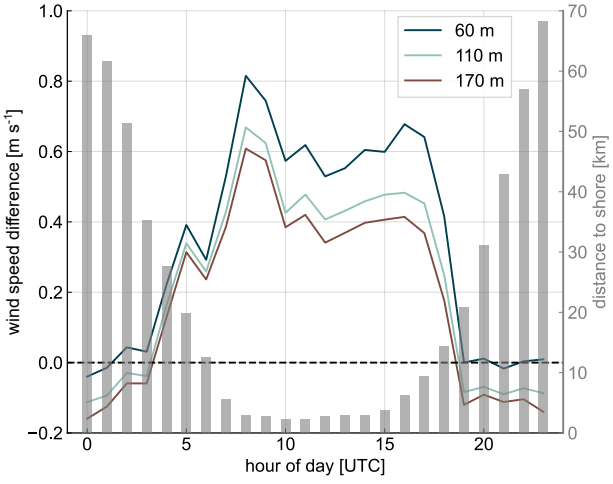

Figure 7. Daily cycle of $ERA5_{raw}$ minus $ERA5_{cal}$ difference at 60 m, 110 m, and 170 m height. The mean ship distance to shore per hour is indicated by the grey bars.

3.3 Validation against stationary lidar observations

In contrast to the previous section, this section compares ERA5 data before and after calibration with measurements from the fixed lidar at Utö, rather than with ship-based lidar measurements. This comparison aims to validate the methodology used by assessing the effects of the calibration against an independent dataset not used in the calibration process. In doing so, we try to demonstrate the generalisation capabilities of the calibration methodology. For this comparison, only ERA5 data from the grid cell closest to the Utö lidar location are used, both before and after calibration.

Figure 8 illustrates the effect of ERA5 calibration at Utö island as a function of (a) the distance between the ship-based lidar and the stationary lidar, and (b) the time of day. Due to the non-stationarity of the ship-based measurements used to calibrate ERA5, both spatial proximity and the ferry schedule are expected to modulate the magnitude of the calibration effect.

Figure 8a presents the mean wind speed bias at 110 m between ERA5 raw and calibrated compared to Utö lidar measurements. The bias is calculated considering several ship-to-Utö distance thresholds. For each threshold, only timestamps when the ship is closer than the specified distance to the Utö lidar are included in the bias calculation. The "all data" case includes data from the entire campaign duration, regardless of the ship's distance from the Utö lidar.

As can be seen, ERA5 raw systematically underestimates wind speeds at Utö, as also reported in Hallgren et al. (2022). The bias ranges from -0.36 m s^{-1} for the 50 km threshold to -0.25 m s^{-1} when considering all available data. After applying the calibration, the bias is consistently reduced across all distance thresholds, demonstrating the effectiveness of the nudging approach. The most substantial improvements occur for the shortest distance thresholds, with the bias decreasing from -0.36 m s^{-1} to -0.14 m s^{-1} for the 50 km case and from -0.3 m s^{-1} to -0.14 m s^{-1} for the 60 km case. The magnitude of these bias reductions is comparable to improvements reported in lidar-assisted mesoscale simulations employing observation nudging or four-dimensional data assimilation, despite the present approach being applied as a post-processing correction (Sommerfeld

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

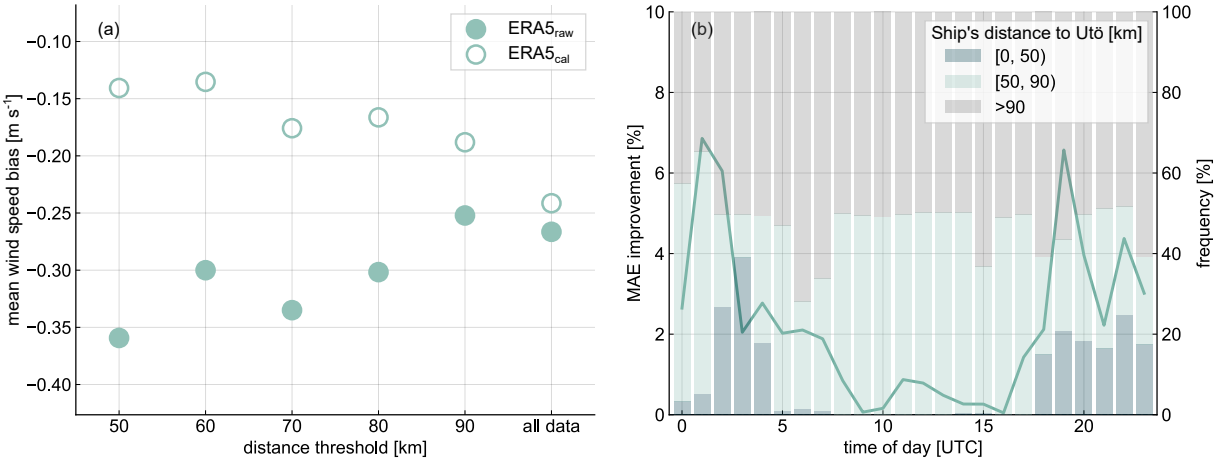

Figure 8. (**a**) Mean wind speed bias at 110 m between ERA5 raw and calibrated compared to stationary lidar measurements (ERA5 - lidar) for different ship-to-Utö lidar distance threshold. For each threshold, only timestamps when the ship is closer than the specified distance to the Utö lidar are included in the bias calculation. (**b**) MAE improvement calculated as *MAE $ERA5_{raw}$ - MAE $ERA5_{cal}$)/MAE $ERA5_{raw}$ for each hour of the day. The stacked bars represent the frequency distribution of the ship-to-Utö distance.

et al., 2019; Ivanova et al., 2025). However, there is a clear decrease in the strength of the calibration effect with increasing distance thresholds. Specifically, deviation improvements of approximately 0.22 m s^{-1}, 0.06 m s^{-1}, and 0.02 m s^{-1} are observed for the 50 km, 90 km, and all data cases, respectively. However, there is a clear decrease in the strength of the calibration effect with increasing distance thresholds.

Figure 8b presents the MAE improvement from calibration as a function of hour of day, along with the frequency distribution of ship-to-Utö distances for each hour. The calibration results in a positive reduction in MAE at all hours; however, the magnitude of the improvement is strongly dependent on the ferry schedule. The largest improvements occur between approximately 17:00 and 07:00 UTC, coinciding with the periods when the ship is closest to the stationary lidar. During these hours, a greater fraction of ship-based lidar observations lie within 50 km of Utö, for which the horizontal weighting function assigns a stronger influence to observations, thereby enhancing the calibration effect. During the central hours of the day, the MAE improvement is more limited, despite a substantial proportion of ship-based lidar records falling within the 90 km radius of influence. In these cases, the ship lidar observations still contribute to the calibration, but their impact is considerably weaker due to the greater separation distance and the corresponding reduction in weighting strength. Overall, the temporal variability of the calibration effect at Utö directly reflects the combined impact of ship trajectory, spatial weighting, and the intermittent proximity of the moving observational platform.

The wind speed difference error distributions for ERA5 raw and calibrated data relative to Utö lidar measurements at 60 m, 110 m, and 170 m are presented in Fig. 9, considering only timestamps when the ship was within 50 km of Utö.

At all three heights, the distributions reveal a systematic negative bias, reaffirming ERA5's tendency to underestimate wind speeds at this location. After applying the calibration, the distribution peak shifts toward zero, indicating a partial correction

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

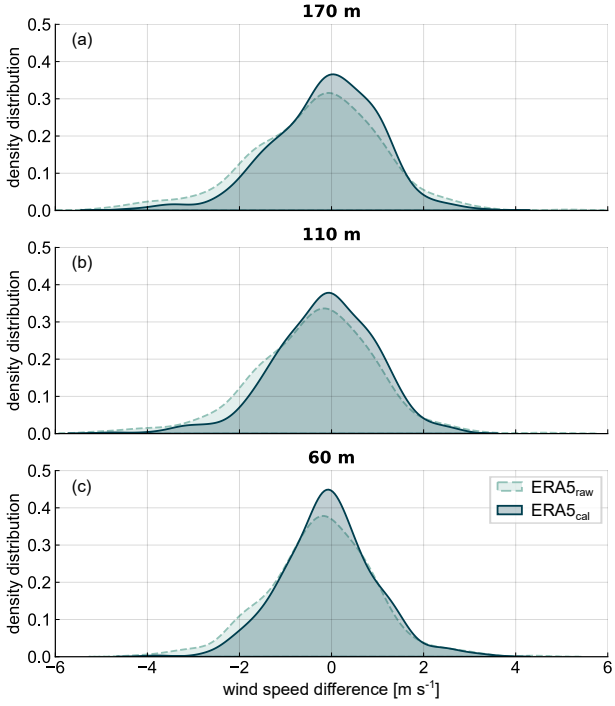

Figure 9. Probability density distribution of the wind speed difference for ERA5 raw and calibrated versus Utö lidar (ERA5 - stationary lidar) at 60 m (**c**), 110 m (**b**), and 170 m (**a**). Only timestamps when the ship is closer than 50 km to the Utö lidar are included in the bias calculation.

of this bias. In addition to bias reduction, the calibration also contributes to a slight decrease in error spread, particularly by mitigating extreme underestimations. This is evidenced by the negative tail of the distribution, where occurrences of substantial ERA5 underestimation are less frequent after the calibration.

A statistical analysis of the wind speed deviation between $ERA5_{raw}$ and $ERA5_{cal}$ with respect to the Utö lidar observations is presented in Fig. 10 in the form of a box plot for several heights, considering only the wind speed differences of those timestamps when the ship is closer than 50 km from the Utö lidar. The coloured boxes mark the 25th and 75th percentiles; the white lines correspond to the median; and the whiskers indicate the data extremes, calculated as 1.5 times the interquartile range.

The boxplots reveal systematic biases and variability before and after the calibration. For raw ERA5, the median wind-speed difference is consistently negative up to 220 m, highlighting ERA5's general tendency to underestimate the measurements from the stationary lidar at Utö. This underestimation is more pronounced in the lower to mid-altitude range, with values around -0.27 m s^{-1} at 60 m to 130 m, where atmospheric dynamics are more complex and challenging to represent for ERA5, as was also reported in (Hallgren et al., 2022). The median wind speed difference turns slightly positive at 270 m (0.06 m s^{-1}).

The interquartile range (IQR) increases slightly with height, from 1.38 m s^{-1} at 60 m to 1.80 m s^{-1} at 270, suggesting greater variability in ERA5 deviations at higher altitudes. One possible explanation is the reduced availability of lidar data at these

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

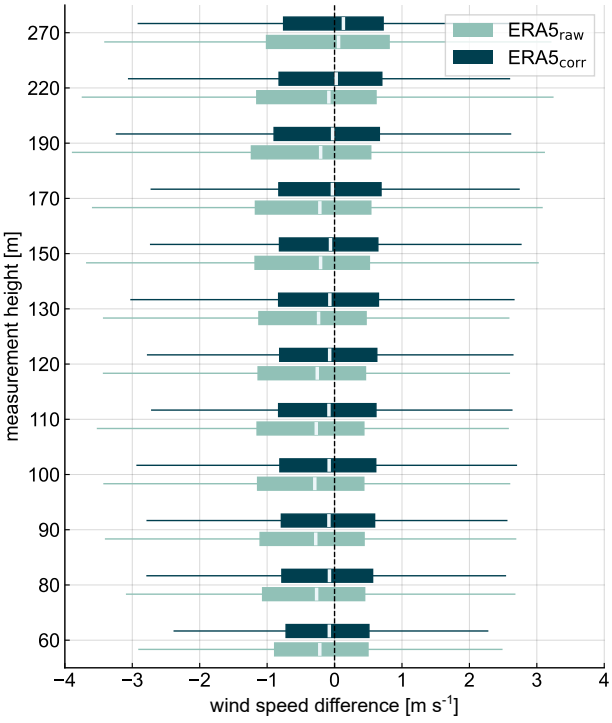

Figure 10. Box plots of the wind speed difference from ERA5 raw and calibrated minus stationary lidar at all ship-based lidar measurement heights. The coloured boxes extend from the first to the third quartiles of the data, and the medians are indicated by white lines. The whiskers extend to the data extremes, defined as 1.5 times the interquartile range (IQR) above and below the upper and lower quartiles, respectively. Only timestamps when the ship is closer than 50 km to the stationary lidar are included in the bias calculation.

heights, leading to increased statistical uncertainty and a wider distribution of differences. This effect is also reflected in the whiskers, with both the lower and upper whiskers extending farther at higher altitudes, indicating more extreme underestimations and overestimations than at lower altitudes.

After calibration, the median wind speed differences shift closer to zero at all heights, with improvements between 36 % and 29 % observed and the largest improvement at lower heights. This significantly reduces the systematic bias, except at 270 m, which shows a slight increase in the median bias. We do not have a clear explanation for this; however, it could be due to the limited availability of ship-based lidar measurements at this height, reducing the usable data for calibration and its effect. The IQR also narrows at all heights, including the 270 m level, indicating a reduction in error variability with the calibration. For instance, at 60 m, the IQR decreases from 1.39 m s^{-1} to 1.23 m s^{-1}, and at 270 m, it reduces from 1.82 m s^{-1} to 1.48 m s^{-1}. The whiskers are generally shorter after the calibration, especially on the negative side, demonstrating a reduction in extreme underestimations. These improvements confirm that the calibration effectively enhances the agreement between ERA5 and lidar observations, particularly at lower heights, where the initial underestimation was most pronounced.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Finally, we compare the 1 h horizontal wind speeds for a period of approximately 72 hours from 14 September 2022 to 17 September 2022, as shown in Fig. 11, to visualise the effect of the calibration on the ERA5 data at the Utö location. Figure 11a presents the wind speed time series at a height of 110 m for the Utö lidar, $ERA5_{raw}$, and $ERA5_{cal}$. The grey bars indicate the distance of the ship from the Utö lidar, and the dashed horizontal line represents the 90 km threshold used in the Cressman scheme, beyond which the calibration strength is null. Figures 11b and 11c display the time-height cross-sections of the wind speed differences between $ERA5_{raw}$ and $ERA5_{cal}$, respectively, compared to the stationary lidar. The white spaces in these two panels indicate missing or filtered-out Utö lidar data.

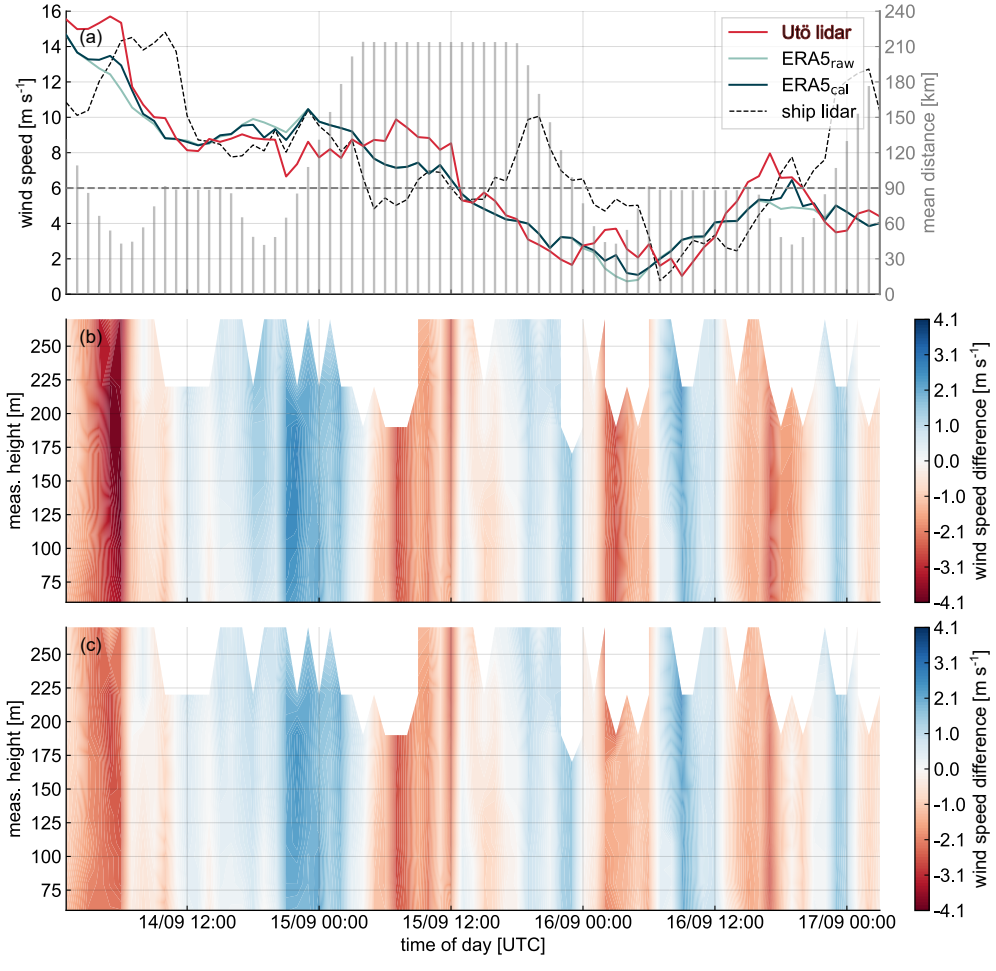

Figure 11. (**a**) Time series of wind speed at 110 m height for the Utö lidar, the ship-based lidar, and ERA5 raw and calibrated for the period from 14/05/2023 to 17/09/2023. The ship-to-Utö lidar distance is indicated by the grey bars, and the 90 km horizontal radius of influence, as defined in the Cressman horizontal weighting function, is depicted by the horizontal grey dashed line. Panels (**b**) and (**c**) show the time-height cross-section of the wind speed differences between $ERA5_{raw}$ and $ERA5_{cal}$ minus the Utö lidar, respectively, for the same time period. Blank areas indicate periods and heights where Utö lidar data is missing.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The calibration has demonstrated a positive effect on ERA5's statistical performance. Regarding the effect on the time series, the calibration brings ERA5 wind speed estimates closer to the observations from the stationary lidar at Utö. Nevertheless, the total impact on the time series remains subtle, primarily due to the ship's non-stationarity and the consequent intermittent influence of ship-based lidar measurements. The calibration effect is most noticeable when the ship is within approximately 60 km of the stationary lidar, where it is sufficiently influential to reduce discrepancies.

Although calibration improves the alignment of $ERA5_{cal}$ with reference observations, some discrepancies remain. In certain cases, the calibrated ERA5 data, though closer to the lidar measurements, do not fully capture the observed wind speed variations. This can be attributed to the minimum 40 km separation between the stationary lidar and the ship, meaning that the wind conditions measured by the ship-based lidar—and subsequently introduced into the calibrated ERA5 dataset—may differ from those at Utö. In addition, advection effects can cause temporal shifts in wind speed variations between the two locations. For instance, between 16 September 12:00 and 17 September 00:00, when the wind direction was approximately 325° (not shown in the figure), the wind speed increase observed by the ship-based lidar occurred after a similar increase was recorded by the Utö lidar, highlighting the influence of transport effects caused by the relatively large distance between the calibration data (ship-based lidar measurements) and the location of the validation data.

The time-height cross-section plots further illustrate the calibration effect, showing that both $ERA5_{raw}$ and $ERA5_{cal}$ exhibit similar overall trends compared to the stationary lidar. However, when the ship is sufficiently close, a reduction in bias is observed across the entire wind profile. This effect is particularly evident during the early hours of 14 September, where the largest deviations occur in the raw ERA5 data. Although the overall impact of the calibration is relatively modest and is primarily visible at specific time steps, it improves the representation of the wind speed time series in ERA5.

4 Concluding Discussion

This study investigated the potential of non-stationary, ship-based lidar measurements to enhance offshore wind estimates from the ERA5 reanalysis, employing a statistical calibration method inspired by observational nudging. This approach has been tested over the northern Baltic Sea using approximately eight months of ship-based lidar data and was assessed using independent observations from a stationary lidar located on Utö island. The discussion below places these findings in the context of existing literature, methodological limitations, and current research needs.

The methodology presented here lies between two established strategies used in offshore wind applications. On the one hand, observational nudging in mesoscale models such as WRF has demonstrated clear benefits for wind and power prediction (Mylonas et al., 2018; Ivanova et al., 2025), but at the cost of increased computational demand, model accessibility, and sensitivity to physical parameterisations (Carvalho et al., 2014). On the other hand, statistical post-processing techniques applied to reanalysis and mesoscale outputs are widely used in wind resource assessment, such as Measure-Correlate-Predict (MCP) methods (Carta et al., 2013), but they typically rely on stationary reference observations and are therefore not directly applicable to mobile platforms. By adapting the conceptual framework of observational nudging to a purely statistical, post-processing formulation, this study expands previous uses of ship-mounted lidar data, which have largely focused on validation

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

and intercomparison applicability, towards a robust calibration model that explicitly accounts for the non-stationarity of the observations.

Optimising the calibration parameters is a fundamental first step in the methodology presented, as this approach does not modify the underlying model physics but instead relies entirely on the statistical representativeness of the observational data used to correct the reanalysis output. Consequently, an inappropriate choice of the calibration strength K and the temporal W_t and horizontal W_{xy} weighting functions can introduce considerable noise into the calibration process, for instance, due to influence windows that are too broad and actually worsen the model's performance by adding decorrelated information. This strong sensitivity to the calibration parameters suggests that, when working with non-stationary ship-based lidar measurements, calibration effectiveness is primarily governed by atmospheric representativeness and decorrelation scales rather than by the total number of available observations.

Results show that the calibration substantially improves the agreement between ERA5 and the vessel-based lidar measurements along the ship route, particularly in coastal and near-coastal areas. This behaviour is consistent with the well-documented tendency of ERA5 to underestimate wind speeds at lower heights and in regions influenced by coastal heterogeneity and unresolved small-scale wind flow variations (Witha et al., 2019a; Duncan et al., 2019). Additionally, this area benefits from a higher density of lidar measurements, enhancing the calibration model's effectiveness. The temporal analysis of calibration effects closely matches the spatial behaviour, showing the greatest improvements during the day when the vessel is near the harbours. This is a direct consequence of the ship-based measurement system's non-stationarity and the fixed route taken by the *Stena Gothica* during the measurement, which concentrates the calibration impact to specific regions and time periods.

Compared to the impact of observational nudging, as when introducing lidar measurements into numerical models such as WRF (as in (Sommerfeld et al., 2019; Ivanova et al., 2025)), the calibration methodology presented here has the potential to produce a more notable effect on the model output. This is because the methodology is not constrained by the equations governing the model's physics. However, it relies heavily on the quality of the measurements used for calibration, resulting in an increased sensitivity to noise and biases in those measurements. Consequently, this highlights the fundamental importance of optimising model parameters prior to the full calibration process to maximise the calibration performance.

The comparison with the stationary lidar at Utö provides insights into the generalisation capability of the proposed methodology. The calibration yields modest but systematic reductions in the mean wind speed bias, with the magnitude of improvement strongly dependent on the distance between the ship and the validation site. The clearest enhancement happens when the ship is within 50-60 km, while the effect becomes weak when considering the full radius of influence. This distance dependence reflects both the spatial weighting applied in the calibration and the intermittent sampling inherent to a moving platform. In addition to reducing the mean bias, the calibration also decreases the error spread and the occurrence of large deviations. The largest improvements are observed at lower heights, where the raw ERA5 data show the strongest underestimation and where a greater availability of ship-based lidar measurements enhances the statistical robustness of the calibration (see Table 1). Nevertheless, the impact of the calibration on the wind speed time series is modest, as evidenced in Fig. 11. This limited impact can be explained by the intermittent influence of ship-based lidar data due to ship motion, as well as by advection effects that introduce temporal mismatches between the calibration (ship-based lidar measurements) and the validation data (Utö lidar).

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Overall, these results demonstrate limited but tangible generalisation of the calibration, while clearly defining the spatial and temporal boundaries within which the method is effective.

Within its validated scope, this study supports several conclusions. A nudging-inspired calibration approach can be formulated to incorporate ship-based lidar measurements into a gridded reanalysis in a computationally efficient and technically straightforward way, provided an adequate definition of the model calibration parameters. When applied to ERA5, the methodology substantially reduces systematic wind speed biases along the ship trajectory, particularly in coastal and near-coastal regions where reanalysis performance is known to be weaker and observational density is higher. When evaluated against an independent stationary lidar, the calibration produces moderate but systematic bias reductions that are intermittent and strongly distance-dependent.

These findings show both the strengths and the limitations of the approach. Compared to observational nudging, the proposed calibration methodology is computationally more efficient, straightforward to implement, and flexible enough to be applied to any gridded numerical datasets. However, its effectiveness depends on the quality and representativeness of the observational data used for calibration, making it more sensitive to noise and potential biases in the observations.

A key constraint identified is that the calibration's impact is limited by the distance between the ship-based lidar and the target location, as mobile platforms continuously move toward and away from a given site. Consequently, improvements at fixed locations are intermittent and decay rapidly with increasing separation. Nevertheless, the widespread availability of ferry routes and commercial ships in regions of interest for offshore wind development, together with the comparatively low cost of ship-based lidar campaigns, suggests that the methodology could be enhanced by combining multiple ship-based lidar datasets. By carefully defining mapping strategies and employing multiple ship-based lidar devices, either simultaneously or sequentially, the spatial coverage and persistence of the calibration influence could be increased at target locations or across multiple sites, although this has yet to be demonstrated.

The results do not imply that ship-based lidar measurements can replace fixed reference systems for site-specific wind resource assessment. Rather, they indicate that ship-based lidar data can complement existing observational and modelling approaches by contributing to regional-scale bias mitigation, particularly in data-sparse offshore regions. Within this context, potential applications include supporting the identification and characterisation of offshore wind development areas, improving offshore wind atlases through numerical modelling, and refining estimates of long-term mean wind speeds in regions lacking permanent measurements. In addition, the ability of ship-mounted systems to sample spatially extended flow features suggests a possible role in informing the representation of wake effects in numerical models, although this application requires further investigation. Most importantly, the methodology is not limited to a single location but can be applied across an entire region influenced by the ship's route, providing broader spatial benefits.

Future work should focus on increasing spatial and temporal persistence by combining multiple ship routes, sequential measurement campaigns, or different mobile platforms, as well as on assessing regime-dependent performance under different atmospheric conditions. Extending the methodology to include additional variables, such as wind direction, vertical shear, or stability-related metrics, would be required for broader wind-energy applications. Finally, systematic benchmarking against

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

stationary-lidar-based statistical corrections and against dynamical downscaling with lidar nudging would help to further clarify the role of ship-based calibration within the wider landscape of offshore wind characterisation methodologies.

Data availability. Data used for this paper were collected from the following sources. Ship-based lidar measurements were provided by Fraunhofer IWES and are available upon request. The ERA5 data are freely available via the Copernicus Data Storage (CDS): https://cds.climate.copernicus.eu/cdsapp#!/home. The stationary lidar data are available from Ville Vakkari (Finish Meteorological INstitute FMI) upon request.

Author contributions. HR performed the investigation, data processing, analysis and visualization and wrote the manuscript. VV provided the stationary lidar data and wrote the description of the dataset. All authors contributed to the conceptualization and methodology and reviewed the manuscript. JG had a supervisory function.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Wind Energy Science.

Acknowledgements. This research received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 858358 (LIKE – Lidar Knowledge Europe). Special thanks to Stena Line for providing the opportunity to conduct the campaign onboard the *Stena Gothica* and the Research Institutes of Sweden RISE for their coordination of the measurement campaign.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

References

Bonavita, M., Hólm, E., Isaksen, L., and Fisher, M.: The evolution of the ECMWF hybrid data assimilation system, Q. J. Roy. Meteorol. Soc., 142, 287–303, https://doi.org/10.1002/qj.2652, 2016.

Brodlie, K. W. and Butt, S.: Preserving convexity using piecewise cubic interpolation, Comput. Graph., 15, 15–23, https://doi.org/10.1016/0097-8493(91)90026-E, 1991.

Browning, K. A. and Wexler, R.: The Determination of Kinematic Properties of a Wind Field Using Doppler Radar, Journal of Applied Meteorology, 7, 105–113, https://doi.org/10.1175/1520-0450(1968)007<0105:TDOKPO>2.0.CO;2, 1968.

Carbon Trust: Carbon Trust Offshore Wind Accelerator Roadmap for the Commercial Acceptance of Floating LIDAR Technology: Technical Report, https://www.carbontrust.com/our-work-and-impact/guides-reports-and-tools/roadmap-for-commercial-acceptance-of-floating-lidar, last access: 21 February 2026, 2018.

Carta, J. A., Velázquez, S., and Cabrera, P.: A review of measure-correlate-predict (MCP) methods used to estimate long-term wind characteristics at a target site, Renewable and Sustainable Energy Reviews, 27, 362–400, https://doi.org/10.1016/j.rser.2013.07.004, 2013.

Carvalho, D., Rocha, A., Gómez-Gesteira, M., and Silva Santos, C.: Sensitivity of the WRF model wind simulation and wind energy production estimates to planetary boundary layer parameterizations for onshore and offshore areas in the Iberian Peninsula, Applied Energy, 135, 234–246, https://doi.org/10.1016/j.apenergy.2014.08.082, 2014.

Clifton, A., Boquet, M., Des Burin Roziers, E., Westerhellweg, A., Hofsass, M., Klaas, T., Vogstad, K., Clive, P., Harris, M., Wylie, S., Osler, E., Banta, B., Choukulkar, A., Lundquist, J., and Aitken, M.: Remote Sensing of Complex Flows by Doppler Wind Lidar: Issues and Preliminary Recommendations, National Renewable Energy Laboratory (NREL), https://doi.org/10.2172/1351595, 2015.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Deng, A., Stauffer, D. R., , J., Otte, T., and Hunter, G.: Update on analysis nudging FDDA in WRF-ARW, in: Proceedings of the 8th WRF Users' Workshop, 11–15 June 2007, National Center for Atmospheric Research, Boulder, Colorado, p. 35, 2007.

Dörenkämper, M., Optis, M., Monahan, A., and Steinfeld, G.: On the Offshore Advection of Boundary-Layer Structures and the Influence on Offshore Wind Conditions, Bound.-Lay. Meteorol., 155, 459–482, https://doi.org/10.1007/s10546-015-0008-x, 2015.

Dudhia, J.: WRF Four-Dimensional Data Assimilation (FDDA), https://www2.mmm.ucar.edu/wrf/users/tutorial/presentation_pdfs/200801/WRF_FDDA_Dudhia.pdf, (last access: 21 February 2026), 2012.

Duncan, J. B., Marseille, G. J., and Wijnant, I. L.: DOWA validation against ASCAT satellite winds, TNO Report 2018 R11649, 2019.

Fritsch, F. N. and Carlson, R. E.: Monotone Piecewise Cubic Interpolation, SIAM J. Numer. Anal., 17, 238–246, https://doi.org/10.1137/0717021, 1980.

Gottschall, J., Gribben, B., Stein, D., and Würth, I.: Floating lidar as an advanced offshore wind speed measurement technique: current technology status and gap analysis in regard to full maturity, Wiley Interdiscip. Rev.: Energ. Environ., 6, e250, https://doi.org/10.1002/wene.250, 2017.

Gottschall, J., Catalano, E., Dörenkämper, M., and Witha, B.: The NEWA Ferry Lidar Experiment: Measuring Mesoscale Winds in the Southern Baltic Sea, Remote Sens., 10, 1620, https://doi.org/10.3390/rs10101620, 2018.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

Gualtieri, G.: Reliability of ERA5 Reanalysis Data for Wind Resource Assessment: A Comparison against Tall Towers, Energies, 14, 4169, https://doi.org/10.3390/EN14144169, 2021.

Hallgren, C., Arnqvist, J., Nilsson, E., Ivanell, S., Shapkalijevski, M., Thomasson, A., Pettersson, H., and Sahlée, E.: Classification and properties of non-idealized coastal wind profiles – an observational study, Wind Energ. Sci., 7, 1183–1207, https://doi.org/10.5194/wes-7-1183-2022, 2022.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hirsikko, A., O'Connor, E. J., Komppula, M., Korhonen, K., Pfüller, A., Giannakaki, E., Wood, C. R., Bauer-Pfundstein, M., Poikonen, A., Karppinen, T., Lonka, H., Kurri, M., Heinonen, J., Moisseev, D., Asmi, E., Aaltonen, V., Nordbo, A., Rodriguez, E., Lihavainen, H., Laaksonen, A., Lehtinen, K. E. J., Laurila, T., Petäjä, T., Kulmala, M., and Viisanen, Y.: Observing wind, aerosol particles, cloud and precipitation: Finland's new ground-based remote-sensing network, Atmospheric Measurement Techniques, 7, 1351–1375, https://doi.org/10.5194/amt-7-1351-2014, 2014.

Ivanova, T., Porchetta, S., Buckingham, S., Glabeke, G., van Beeck, J., and Munters, W.: Improving wind and power predictions via four-dimensional data assimilation in the WRF model: case study of storms in February 2022 at Belgian offshore wind farms, Wind Energy Science, 10, 245–268, https://doi.org/10.5194/wes-10-245-2025, 2025.

Kalverla, P. C.: Characterisation of offshore winds for energy applications, Ph.D. thesis, Wageningen University, the Netherlands, https://doi.org/10.18174/498797, 2019.

Knoop, S., Ramakrishnan, P., and Wijnant, I.: Dutch Offshore Wind Atlas Validation against Cabauw Meteomast Wind Measurements, Energies, 13, 6558, https://doi.org/10.3390/en13246558, 2020.

Liu, Y., Warner, T. T., Bowers, J. F., Carson, L. P., Chen, F., Clough, C. A., Davis, C. A., Egeland, C. H., Halvorson, S. F., Huck, T. W., Lachapelle, L., Malone, R. E., Rife, D. L., Sheu, R.-S., Swerdlin, S. P., and Weingarten, D. S.: The Operational Mesogamma-Scale Analysis and Forecast System of the U.S. Army Test and Evaluation Command. Part I: Overview of the Modeling System, the Forecast Products, and How the Products Are Used, Journal of Applied Meteorology and Climatology, 47, 1077–1092, https://doi.org/10.1175/2007JAMC1653.1, 2008.

Mylonas, M. P., Barbouchi, S., Herrmann, H., and Nastos, P. T.: Sensitivity analysis of observational nudging methodology to reduce error in wind resource assessment (WRA) in the North Sea, Renewable Energy, 120, 446–456, https://doi.org/10.1016/j.renene.2017.12.088, 2018.

Pearson, G., Davies, F., and Collier, C.: An Analysis of the Performance of the UFAM Pulsed Doppler Lidar for Observing the Boundary Layer, Journal of Atmospheric and Oceanic Technology, 26, 240–250, https://doi.org/10.1175/2008JTECHA1128.1, 2009.

Pichugina, Y. L., Brewer, W. A., Banta, R. M., Choukulkar, A., Clack, C. T. M., Marquis, M. C., McCarty, B. J., Weickmann, A. M., Sandberg, S. P., Marchbanks, R. D., and Hardesty, R. M.: Properties of the offshore low level jet and rotor layer wind shear as measured by scanning Doppler Lidar, Wind Energy, 20, 987–1002, https://doi.org/10.1002/we.2075, 2017.

Reen, B.: A Brief Guide to Observation Nudging in WRF, Tech. rep., Army Research Laboratory, https://doi.org/http://www2.mmm.ucar.edu/wrf/users/docs/ObsNudgingGuide.pdf, (last access: 21 February 2026), 2016.

https://doi.org/10.5194/wes-2026-50
Preprint. Discussion started: 26 March 2026

Reen, B. P. and Stauffer, D. R.: Data Assimilation Strategies in the Planetary Boundary Layer, Boundary-Layer Meteorology, 137, 237–269, https://doi.org/10.1007/s10546-010-9528-6, 2010.

Rubio, H. and Gottschall, J.: Development of an analytical uncertainty model for ship-based lidar measurements, J. Phys. Conf. Ser., 2362, 012 034, https://doi.org/10.1088/1742-6596/2362/1/012034, 2022.

Rubio, H., Kühn, M., and Gottschall, J.: Evaluation of low-level jets in the southern Baltic Sea: a comparison between ship-based lidar observational data and numerical models, Wind Energy Science, 7, 2433–2455, https://doi.org/10.5194/wes-7-2433-2022, 2022.

Rubio, H., Hatfield, D., Hasager, C. B., Kühn, M., and Gottschall, J.: Ship-based lidar measurements for validating ASCAT-derived and ERA5 offshore wind profiles, Atmospheric Measurement Techniques, 18, 4949–4968, https://doi.org/10.5194/amt-18-4949-2025, 2025.

Savazzi, A. C. M., Nuijens, L., Sandu, I., George, G., and Bechtold, P.: The representation of the trade winds in ECMWF forecasts and reanalyses during EUREC 4 A, Atmos. Chem. Phys., 22, 13 049–13 066, https://doi.org/10.5194/acp-22-13049-2022, 2022.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Model Version 4, National Center for Atmospheric Research (NCAR), https://doi.org/10.5065/1DFH-6P97, 2019.

Sommerfeld, M., Dörenkämper, M., Steinfeld, G., and Crawford, C.: Improving mesoscale wind speed forecasts using lidar-based observation nudging for airborne wind energy systems, Wind Energ. Sci., 4, 563–580, https://doi.org/10.5194/wes-4-563-2019, 2019.

Stauffer, D. R., Seaman, N. L., and Binkowski, F. S.: Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model Part II: Effects of Data Assimilation within the Planetary Boundary Layer, Monthly Weather Review, 119, 734–754, https://doi.org/10.1175/1520-0493(1991)119%3C0734:UOFDDA%3E2.0.CO;2, 1991.

Vaisala: WindCube Data Filtering Guidelines, https://doi.org/https://www.vaisala.com/sites/default/files/documents/WindCube-Data-Filtering-Guidelines-v1.3.pdf, (last access: 21 February 2026), 2022.

Vakkari, V., Manninen, A. J., O'Connor, E. J., Schween, J. H., van Zyl, P. G., and Marinou, E.: A novel post-processing algorithm for Halo Doppler lidars, Atmospheric Measurement Techniques, 12, 839–852, https://doi.org/10.5194/amt-12-839-2019, 2019.

Witha, B., Dörenkämper, M., Frank, H., García-Bustamante, E., González-Rouco, F., Navarro, J., Schneider, M., Steeneveld, G.-J., Svensson, N., and Gottschall, J.: The NEWA Ferry Lidar Benchmark: Comparing mesoscale models with lidar measurements along a ship route, Wind Energy Science Conference, 17–20 June 2019, Cork, Ireland, https://doi.org/10.5281/zenodo.3372693, 2019a.

Witha, B., Hahmann, A., Sile, T., Dörenkämper, M., Ezber, Y., García-Bustamante, E., González-Rouco, J. F., Leroy, G., and Navarro, J.: WRF model sensitivity studies and specifications for the NEWA mesoscale wind atlas production runs, Zenodo [presentation], https://doi.org/10.5281/zenodo.2682604, 2019b.

Wolken-Möhlmann, G. and Gottschall, J.: Ship-based lidar measurement in the wake of an offshore wind farm, J. Phys. Conf. Ser., 555, 012 043, https://doi.org/10.1088/1742-6596/555/1/012043, 2014.

Yang, B., Zhong, L., Wang, J., Shu, H., Zhang, X., Yu, T., and Sun, L.: State-of-the-art one-stop handbook on wind forecasting technologies: An overview of classifications, methodologies, and analysis, Journal of Cleaner Production, 283, 124 628, https://doi.org/10.1016/j.jclepro.2020.124628, 2021.