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

Hugo Rubio[a], Martin Kühn[b], and Julia Gottschall[a]

[a]Fraunhofer Institute for Wind Energy Systems (IWES), 27572 Bremerhaven, Germany
[b]ForWind, Institute of Physics, Carl von Ossietzky Universität Oldenburg, Küpkersweg 70, 26129 Oldenburg, Germany

Correspondence: Hugo Rubio (hugo.rubio@iwes.fraunhofer.de)

Abstract. ~~Ship-based lidar measurements from the NEWA Ferry Lidar Experiment have been used together with ERA5 and numerical data from the NEWA wind atlas to analyze~~

In contrast to fixed measuring devices, ship-based lidar systems provide highly-reliable wind observations within extensive regions. Therefore, this kind of reference dataset provides a great potential for evaluating the performance of mesoscale numerical models in resembling mesoscale flow phenomena such as low-level ~~jets over the Southern Baltic Sea with the following main objectives: (1) to evaluate the performance of~~ jets, essential for an optimal development and operation of wind turbines. This paper presents a comparison between numerical output data from two state-of-the-art reanalyses (ERA5 and NEWA) and the ship-mounted lidar ~~observations to investigate low-level jet properties along the region~~ measurements from the NEWA Ferry Lidar Experiment. The comparison has been performed along the route covered by the ~~ship track and (2) to compare these observations with numerical simulations to better understand their capacities and limitations in retrieving the characteristics of this mesoscale phenomenon. Most in situ measurement devices can only retrieve wind characteristics in a fixed location. However, the~~ vessel, as well as in specific locations within this route to better understand the capabilities and limitations of the numerical models to precisely resemble the occurrence and main properties of low-level jets under different spatial constraints. The findings of this study show that the non-stationary nature of ship-based lidar systems allows them to capture the variability of the jets' characteristics due to both temporal and spatial effects. ~~The models struggle with accurately capturing the jet features, although they can properly identify their trend in the different locations along the ship track~~Numerical models underestimate the occurrence of the jets and they struggle to accurately describe their main characteristics, with a particularly large underestimation of the fall-off. The found results are strongly influenced by the characteristics of the observations, such as the data availability, the relation time-position of the selected vessel´s route, or the profile height limitation, as well as by the features of the jets, with a particular relevance of core height and fall-off. Additionally, the results illustrate the temporal and spatial shift between the jets detected by the measurements and the models and the potential benefit of considering such deviations when studying low-level jets' climatology through numerical modes.

1 Introduction

The constantly growing demand for carbon-free energy has fostered the increase of wind power generation systems. Although 93 % of the worldwide installed wind capacity is onshore (International Renewable Energy Agency, 2022), the higher and more stationary wind resources available in offshore regions have stimulated an increasing interest in developing new wind farms in these locations (Sempreviva et al., 2008). Particularly in Europe, the cumulative installed wind capacity is expected to grow from 28 GW at the end of 2021 to 79 GW by 2030 (WindEurope, 2022). Nevertheless, the higher cost of grid connection compared to onshore, the challenging logistics of these sorts of projects, and the lack of high quality and accurate measurements at these sites hinder a faster development of offshore wind power plants.

In situ observations are essential for the optimal design of future wind farms, both for evaluating available wind resources and for appropriately selecting wind turbines to withstand the harsh atmospheric and oceanographic conditions. Wind lidars (light detection and ranging) provide an attractive alternative to traditional meteorological (met) masts for providing on-site wind data and retrieving high-quality measurements of the wind profile up to higher heights than met masts (Kindler et al., 2007; Mann et al., 2010), in addition to minimizing the constructional restrictions in deeper waters. Lidar devices can be employed in various configurations, such as their installation on wind turbine nacelles to investigate the wind inflow conditions upstream turbines (Held, 2019) or mounted on floating platforms such as buoys or ships (Gottschall et al., 2017). While buoy-based lidars are a straightforward replacement to the traditional met masts typically used by the wind industry, the implementation of ship-mounted lidars is more intricate due to the non-stationary position of the ship, and thus, a too sparse data coverage that complicates assessing the site-specific wind resources. However, the installation of lidar devices onboard vessels offers attractive advantages compared to both met masts and buoy-based lidars. On the one hand, its relatively simple setup and its installation on already existing floating platforms allow reducing the restrictions, cost and complexity of offshore measurement campaigns. On the other hand, ship-mounted campaigns cover extensive regions, providing wind data from diverse areas of interest. Nonetheless, the availability of highly reliable offshore wind observations is still scarce. Consequently, the extensive temporal and spatial coverage of mesoscale numerical models and their ability to resolve the most significant features of the marine boundary layer has stimulated the employment of numerical data to investigate local wind resource conditions in offshore sites. However, the limitations of the models due to factors such ~~us~~ as a too coarse horizontal and vertical resolution, or the incomplete representation of the physical processes results insufficient for the accurate description of mesoscale phenomena.

The Baltic Sea is a relatively small semi-enclosed sea with a short average distance to shore. Therefore, the land-sea interaction has a relevant influence on the wind characteristics of the region, causing unusual mesoscale conditions (Hallgren et al., 2020) such as a significant higher probability of low-level jets (LLJ) events; see Figure 1. LLJs are a mesoscale-flow phenomenon that can be defined as a relative maximum in the wind speed profile in the lower part of the atmosphere, typically situated between 100 and 500 meters above the surface (Baas et al., 2009) and being able to span ~~an extension of several kilometers width~~ a width of hundreds of kilometers (Banta et al., 2002; Pichugina et al., 2004). LLJs increase the wind shear and turbulence compared to standard wind profiles (commonly described using a logarithmic or power-law profile), affecting

the performance and loads of wind turbines (Gutierrez et al., 2016; Sathe et al., 2013) and their wake recovery rates (Gutierrez et al., 2017). For this reason, the assessment of the relevant wind conditions in an offshore region such as the Baltic Sea requires a comprehensive understanding of the site-specific properties of LLJs.

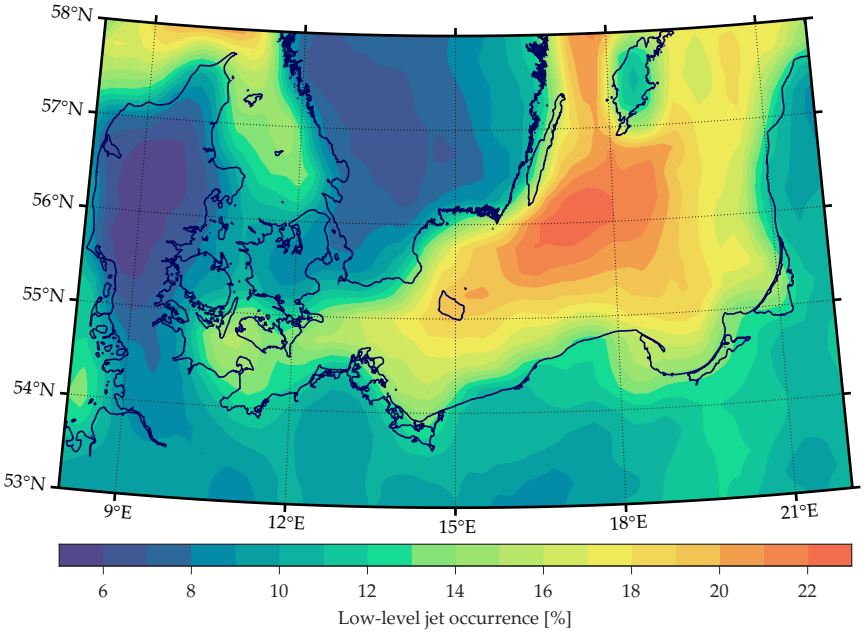

Figure 1. Low-level jet occurrence over the Baltic Sea based on ERA5 data from 2017 up to 500 m height.

Low-level jets have been intensively studied in previous investigations focused on diverse regions worldwide, both in onshore and offshore locations like the Baltic Sea ~~(Hallgren et al., 2020; Svensson et al., 2019a, b)~~(Högström and Smedman-Högström, 1984; Smed , the North Sea ~~(Kalverla et al., 2019; Wagner et al., 2019)~~(Kalverla et al., 2019; Wagner et al., 2019; Schulz-Stellenfleth et al., 2022) , North America (Bonner, 1968; Parish et al., 1988), or the Northern hemisphere's polar regions (Tuononen et al., 2015). According to former studies, there are two main mechanisms that explain the formation of jets in the wind velocity profiles. One of these forcing mechanisms is inertial oscillation (Blackadar, 1957; van de Wiel et al., 2010). In the hours close to sunset, the development of stable stratification leads to a turbulence reduction in the lower part of the boundary layer, resulting in a frictional decoupling between the different horizontal layers. Consequently, the wind accelerates, triggering the development of nocturnal jets. In addition, frictional decoupling may also appear when relatively warm air flows out over colder waters (Smedman et al., 1993).

The second major forcing mechanism for the formation of LLJs is baroclinicity. It causes a reduction of the geostrophic wind speed with height as a consequence of horizontal temperature gradients, which combined with the slowing wind in the near-surface layers due to friction, can result in a maximum on the wind speed profile at intermediate heights (Baas, 2009). Baroclinicity can occur as a consequence of several factors. For instance, a ~~sloppy~~ sloping topography can generate horizontal

gradients of temperature over the daily cycle ~~(Stensrud, 1996)~~(Holton, 1967; Stensrud, 1996). Besides, areas with different surface characteristics, such as coastal sites, where there are strong temperature gradients between sea and land, can lead to baroclinicity and, ultimately, to the formation of the so-called coastal LLJs ~~(Svensson et al., 2019b; Savijärvi et al., 2005)~~ (Baas et al., 2009; Svensson et al., 2019b; Savijärvi et al., 2005). Apart from this, other studies have concluded that sea breezes (Fisher, 1960) and ice edges (Tuononen et al., 2015) may also favor the formation of LLJs.

The ~~large spatiotemporal extent of numerical models and the ability of~~ capability of ship-based lidar systems to provide highly reliable wind data over extensive regions ~~suggest that these two datasets may be attractive alternatives to investigate the temporal and spatial variations of~~ provides an unique opportunity to evaluate the performance of mesoscale numerical models when resembling certain mesoscale effects such ~~as LLJs . This~~ as LLJs within diverse regions and spatial constraints. The work presented in this paper addresses this ~~hypothesis~~ hypotheses by employing the ship-based lidar measurements from the NEWA Ferry Lidar Experiment (Gottschall et al., 2018) in the Southern Baltic Sea; and two ~~of the most frequently used numerical datasets~~well-known and freely available mesoscale numerical models, namely ERA5 and NEWA. ~~First, we aim to~~ For this, we define and implement a comparison methodology according to the different temporal and spatial characteristics of the datasets and evaluate the capabilities ~~of the ship-based lidar technology to retrieve LLJ properties~~ and limitations of the aforementioned reanalyses for modeling the main properties of LLJs along the vessel~~'~~'s route and in specific locations along that route. ~~Then, measurements are compared against the numerical models to assess their performance for LLJ characterization as well as their limiting and influencing factors.~~ This study does not aim to describe in detail either the characteristics of the ship-mounted lidar observations or the physical models applied by the simulations but to ~~understand how these~~ contribute to evaluate how these particular datasets can be used ~~to investigate LLJs and the challenges and restraints for this application~~for the derivation of meaningful information about the jets phenomena.

The manuscript is structured as follows. It starts with a detailed description of the observations and reanalysis datasets used in this study and the definition of the data processing sequence and methodology employed. In particular, a methodology of comparison of the several employed datasets and a LLJ detection algorithm are introduced. Section 3 contains the main results obtained in this investigation. First, an evaluation of the comparison between the wind speed retrievals of the three used datasets is performed. Secondly, LLJ properties along the ship course are analyzed, comparing the obtained properties values for each used dataset. Afterwards, we investigate the sensitivity of the models on the different LLJ features, and finally, the influence of models´ temporal and spatial shift on their capabilities is assessed. Section 4 discusses the implications of the results highlighted in the previous section. Section 5 completes the contribution with our concluding remarks.

2 Materials and methods

In this section, a description of the lidar observations and the used reanalysis datasets is presented. Additionally, the methodology employed for comparing the different datasets and the LLJs detection algorithm are defined in detail.

2.1 Ship-based lidar observations

The observations used in this study were obtained during the execution of the NEWA Ferry Lidar Experiment that took place between February and June 2017 (Gottschall et al., 2018). In this campaign, a wind lidar profiler was installed onboard a ferry boat to measure the winds along the ship track, covering a region of several hundred kilometers in the Southern Baltic Sea from Kiel (Germany) to Klaipeda (Lithuania). Each trip from one destination to the other took around 20 hours, and the ship spends about 4 hours in the harbor after each journey before returning. Figure 2a shows the hourly averaged ship position during the execution of the campaign.

The lidar device used in this campaign was a vertical profiling Doppler lidar from the manufacturer Leosphere (model WindCube WLS7), and it was configured to measure winds at 12 different height levels ranging from 65 m to 275 m above sea level (see Figure 2b). This device has a sampling resolution of about 0.7 s per line-of-sight (LoS) measurement, obtaining wind values from radial-velocity measurements at four azimuth positions, each separated by 90° with a half-opening angle of 28° and followed by a fifth vertical beam. Each LoS velocity is converted to wind speed and direction using a Doppler Beam Swinging (DBS) technique (Peña et al., 2015), reconstructing the 3-dimensional wind vector after each new LoS measurement.

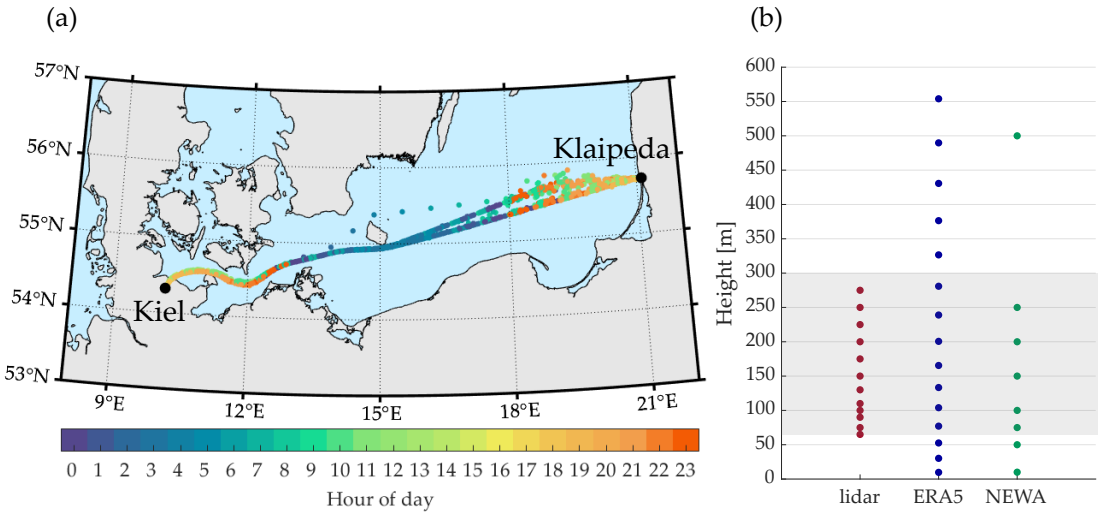

Figure 2. (**a**) Hourly averaged ship position during the execution of the measurement campaign. The hour of the day for each position is indicated by the color scale. (**b**) Retrieved heights for each dataset. For ERA5, the shown heights are the mean heights of the model levels. The shadowed area represents the bottom and top limit heights of vertical profiles used for LLJ detection in this study.

Apart from the lidar device, the integrated measurement system is composed of an xSens MTi-G attitude and heading reference sensor (AHRS) and a Trimble SPS261 satellite compass used to record the high-resolution motion and positioning information. Additionally, a weather station by the manufacturer Vaisala was installed to collect atmospheric data (air pressure, temperature, relative humidity, and precipitation). Further specifications about the ship-mounted lidar system, its components, and

pictures of its installation can be found in ~~(Wolken-Möhlmann et al., 2014; Gottschall et al., 2018)~~Wolken-Möhlmann et al. (2014); Gottsch
.

2.1.1 Lidar data motion compensation and quality check

An indispensable element of the ship-mounted lidar systems is the compensation of vessel motion effects on lidar observations. The ship velocity, tilting, and heading influence the geometry of lidar beam projections contaminating the radial-velocity measurements retrieved by the device. Consequently, each single LoS velocity measurement requires a correction in order to provide reliable wind data. This correction can be either done by using a motion-stabilizing platform to avoid lidar tilting (Achtert et al., 2015), by a post-processing motion compensation algorithm (Zhai et al., 2018), or by a combination of both.

During the execution of the NEWA Ferry Lidar experiment, no motion-stabilizing platform was used, requiring the implementation of a motion correction algorithm. For this, vessel motion data combined with lidar measurements were used, and a simplified motion correction algorithm (Wolken-Möhlmann et al., 2014) was implemented. This algorithm considers the translational ship velocity and orientation, ignoring vessel tilting due to its negligible influence on the results (Wolken-Möhlmann et al., 2014).

Additionally to the motion compensation post-processing, ~~likewise in any other measurement campaign,~~ a quality check of the lidar observations has been ~~a fundamental step~~ implemented to assure the reliability of the output data. In this study, observations with a carrier-to-noise ratio (CNR) lower than -23 dB were rejected from the final database. Then, we averaged the lidar observations into hourly values using a block average with a 1-hour time window centered at each hour. This way, each hourly value was calculated from the measurements recorded half an hour before and after the corresponding timestamp. ~~Hourly~~ For each measurement height, hourly values with an availability below 80 % were ~~filtered, and~~ rejected. Additionally, wind profiles with a missing measurement at 100 m height ~~were deleted from the database. Additionally, those profiles~~ and with more than 70 % of the data missing ~~(over~~ in the whole profile ~~)~~ were excluded from the database. After this process, the total lidar availability was 89.6 % and 83.3 % at 100 m and 200 m height, respectively.

2.2 Numerical model datasets

Numerical mesoscale models are able to ~~capture~~ simulate wind conditions within large-scale areas, being especially useful in offshore environments with limited measurements available. Because of this, evaluating these models against in situ observations is vital to assess their performance under different conditions. From a wind energy application perspective, numerical models must be able to characterize not only the average wind features, but also the variable conditions resulting from mesoscale effects as well as wind shear and turbulence.

Different from the observations, which can be assigned to a single point, numerical models retrieve the average conditions of each grid box covering the spatial domain, restricting their capacity to retrieve extreme wind features. Additionally, their horizontal resolution limits their ability to resolve the rapid spatial wind variations in coastal areas.

The investigations presented in this study were accomplished employing two ~~numerical models datasets well-known within the wind energy industry~~state-of-the-art numerical models, i.e., ERA5 ~~(ECMWF)~~ (Hersbach et al., 2020) and NEWA (Hah-

mann et al., 2020; Dörenkämper et al., 2020). Both datasets are open access and have a suitable temporal and spatial coverage for their application in this study. Table 1 shows the main characteristics of both numerical models, and a more detailed description is included in the following lines.

Table 1. Mean characteristics of used numerical models

	ERA5	**NEWA**
Complete name	ECMWF Retrospective Analysis 5th generation	New European Wind Atlas
Time coverage	1950 - present	1989 - 2018
Spatial Domain	Global	Europe
Horizontal resolution (Baltic Sea)	17 x 31 km	3 x 3 km
Vertical resolution	137 levels up to 0.01 hPa	61 levels up to 50 hPa
Temporal resolution	1 h	0.5 h
Data assimilation	12 hr 4D-Var	-
Boundary conditions	-	ERA5 (9, 6 and 3 km nested domains)
Model	IFS Cycle 41r2	WRF v3.8.1 (modified)

2.2.1 ERA5

ERA5 (ECMWF Reanalysis 5th Generation) is the ~~newest~~ latest reanalysis dataset produced by the European Center for Medium-Range Weather Forecast (ECMWF)~~(ECMWF)~~(Hersbach et al., 2020). It integrates modeled data with observations in sites widespread across the world using a 10-member ensemble 4D-var data assimilation together with the ECMWF Integrated Forecasting System (IFS Cycle 41r2). It offers a large amount of atmospheric, land, and oceanic variables covering the Earth from January 1950 to the present and utilizes 137 pressure (model) levels which go from surface level to the top of the atmosphere, up to 80 km height. These output variables are available in hourly resolution using a 0.25°- 0.25° latitude-longitude grid, or in other words, with a horizontal resolution of around 30 km (17 x 31 km in the Baltic Sea). The assimilation scheme used by ERA5 uses 12-hourly windows in which observations are used from 09:00 to 21:00 (inclusive) UTC and from 21:00 to 09:00 (inclusive) UTC of the next day. It is known that the current version of this reanalysis dataset has a mismatch in the wind speed between the end of one assimilation cycle and the beginning of the following (ECMWF).

For this paper, only the 21 lowest model levels (up to approximately 1 km height) were used. For each level, u and v wind components were employed to asssess the horizontal wind speed and direction.

2.2.2 NEWA

The New European Wind Atlas (NEWA) was generated to provide a high-resolution dataset of wind resource parameters covering the whole of Europe and Turkey (Hahmann et al., 2020). This wind atlas is based on 30 years of model simulations employing a modified version of the open access Weather Research and Forecast (WRF) model Version 3.8.1 (Dörenkämper et al., 2020) over a grid with a 3x3 km spatial resolution. The NEWA database was collected by running the WRF model simulations for 7 days plus a 24h ~~spin-off~~ spin-up period and using ERA5 as initial and boundary conditions (Hahmann et al., 2020). All simulations ran using three nested domains with a 3, 9, and 27 km horizontal grid resolution for the innermost, intermediate, and outer domain, respectively. The whole region covered by NEWA was divided into 10 independent regions, and then all simulations were welded together along their borders. The WRF settings used for the generation of NEWA comprise a modified Mellor-Yamada-Nakanishi-Niino (MYNN) planetary boundary layer (PBL) scheme and the sea surface temperature was obtained from OSTIA. For further details we refer the reader to (Hahmann et al., 2020).

Parameters can be downloaded in 30 minutes time steps between 1989 and 2018 at eight levels between 10 to 500 meters height.

2.3 Comparison of datasets

The different temporal, vertical and spatial resolution of the datasets used in this study requires the definition of a common framework for comparison. For this, lidar observations were averaged to hourly values, as explained in Section 2.1.1. Analogously, an overlapping block average was used to determine NEWA hourly data, using the previous and subsequent 30-min recordings in addition to the value at the corresponding hour. Finally, ship position information has been employed to calculate the mean hourly ship position of the vessel.

After the time-averaging process, the adjacent grid point for each hourly ship position was selected (for both numerical databases), assuring that every hourly lidar measurement is compared against wind values retrieved by the models in the nearest grid point. Consequently, the different spatial resolutions of the models' grids are a limiting factor in their capacity to correctly feature the conditions at the site where the observation was made. This fact can be observed in Figure 3, where the coarser horizontal resolution of ERA5 leads to a worse ability to resolve the geographical and coastal features, as well as to higher distances with regards to the corresponding vessel position.

Additionally, hourly data where the measured profile was incomplete were not considered for the analysis. In order to compare wind speed profiles and the presence of LLJs, wind speed has been interpolated for every 10^{th} meter height, starting from the lowest lidar measurement height (i.e., 65 m) and limited up to 300 m elevation (for moth the measurements and the numerical models). For this, a piecewise cubic Hermite interpolating polynomial (PCHIP) (Fritsch and Carlson, 1980; Brodlie and Butt, 1991) has been employed. This interpolation methodology concentrates the curvature ~~in the nearby~~ of the interpolated line closer to the interpolating points, providing a continuous description of the wind profile and preventing the common swings that can be produced when using a spline interpolation.

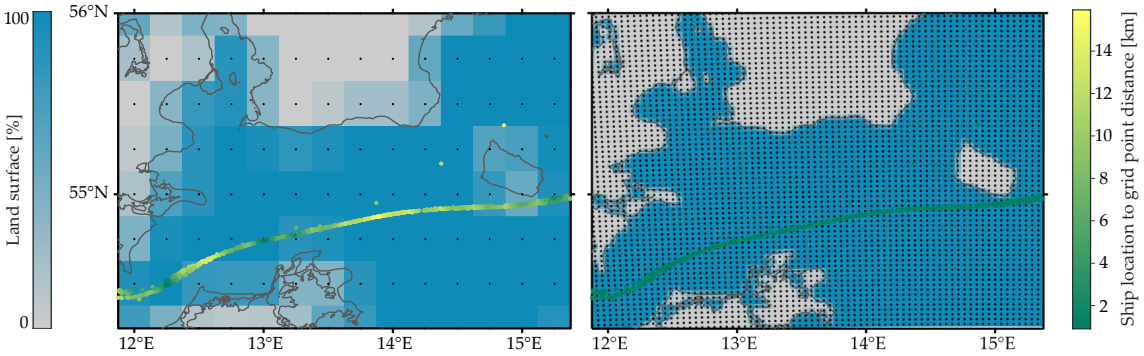

Figure 3. Land/sea mask for ERA5 (left) and NEWA (right) grids. For each grid point (black dots), the ratio between land and water in the corresponding grid box is shown. Hourly ship positions are included in a green to yellow color scale, indicating the distance between each hourly vessel position and the nearest grid point.

The height limitation in the vertical profiles up to 300 m avoid the detection of jets located higher in the atmosphere. However, preceding literature where higher observational wind profiles were employed shows that the majority of the LLJs are located at heights below 250 m height. Therefore, the scarce occurrence of these events prevent them to significantly influence the calculated statistics. In Tuononen et al. (2017), the distribution of LLJ´s core heights measured with a Doppler lidar reaching up to several kilometers heights shows that the vast majority of jets measured in Utö (Northern Baltic Sea) are below 200 m. In Baas et al. (2009), from the same distribution it can be derived that LLJs are usually locate between 140 and 260 m height. And in Pichugina et al. (2017), a ship-mounted lidar measuring profiles up to around 2.5 km proved that most of the detected jets were located at heights below 200 m. Moreover, it must be recalled that this paper is focused in wind energy applications, and thus, due to the current size of offshore wind turbines currently reaching tip heights up to around 220 m (International Energy Agency, 2019), the employed extension of the wind profile used in this study provides wind information about the relevant environment in which present wind turbines operate.

2.4 LLJ detection algorithm

Although LLJs can be identified as wind speed maximums in the lower part of the atmosphere, the criterion used to discern whether a jet is considered as an LLJ or not is currently neither rigorously nor objectively defined. In most cases, the difference between the maximum wind speed and the minimum above it (the so-called fall-off) is used as the primary criterion for determining if a maximum is considered an LLJ. In ~~(Bonner, 1968)~~Bonner (1968), several types of LLJ are established according to both the core speed of the jet and the minimum fall-off value required above them. In ~~(Stull, 1988; Andreas et al., 2000)~~ Stull (1988); Andreas et al. (2000), LLJs are defined as a maximum in the wind speed profile that arises in the lowest 1500

m of the atmosphere, and that is at least 2 m s^{-1} faster than the wind speed values beneath and above. In ~~(Baas et al., 2009)~~ Baas et al. (2009), a maximum on the wind speed profile within the lowest 500 m of the atmosphere is considered as an LLJ if the fall-off is at least 2 m s^{-1} and 25 % faster than the wind speed of the subsequent minimum above.

The selected criteria have a decisive influence on the amount of LLJs detected. As can be observed in Table 2, increasing the absolute (difference between the maximum and minimum above in m s^{-1}) and relative (difference between the maximum and minimum above in percentage) fall-off threshold drastically decreases the total amount of events, both in the reanalyses and the observations. Therefore, the selection of the criteria must be made within a compromise between the availability of a sufficient number of event to obtain meaningful information about the jets, and the labeling of too weak jets as LLJs. In this study, as in ~~(Hallgren et al., 2020)~~Hallgren et al. (2020), an LLJ is defined as a maximum in the wind speed profile that is more than 1 m s^{-1} faster than the minimum above. When no minimum is present above the LLJ, the wind speed at the top height in the wind profile is considered as the minimum value. If more than one jet is detected in the same wind profile, the one with the strongest fall-off prevails. Additionally, as in ~~(Baas et al., 2009)~~Baas et al. (2009), local minimums are neglected if the wind speed increases less than 1 m s^{-1} before dropping again below the minimum studied. An illustration of the criteria for LLJ detection is shown in Figure 4.

Table 2. Number of detected low-level jet events for different criteria (wind profiles up to 300 m)

Criteria	lidar	ERA5	NEWA
fall-off larger than 1 m s^{-1}	139	52	81
fall-off larger than 1 m s^{-1} and 20 %	65	28	43
fall-off larger than 2 m s^{-1}	54	8	24
fall-off larger than 2 m s^{-1} and 20 %	44	7	18

~~The vertical extension of the wind profile analyzed also influences the detection of jets. In (Kalverla et al., 2019), it is observed that considering wind profiles extended~~ Previous studies (Svensson et al., 2018; Kalverla et al., 2019) observed that reanalyses generally overestimate the elevation of the jets in the profile and that extending the height of the scanned profiles up to 500 m ~~rather than 300 m considerably raises the number of LLJs events in the ERA5 dataset. For this reason~~considerably increases the frequency of the jets. Nevertheless, and in order to ensure a fair comparison between the three different datasets, wind profiles scanned to detect LLJs were restricted to start at the lowest measurement height (65 m above sea level) and up to 300 m ~~maximum~~ height for the three employed datasets.

3 Results

After introducing the employed methodology in the preceding section, we now present the main results obtained in this study. First, a comparison between the obtained wind speeds, wind distributions, and vertical profiles of the different datasets is presented to justify the comparison methodology employed. Next, an evaluation of the main LLJ properties along the ship

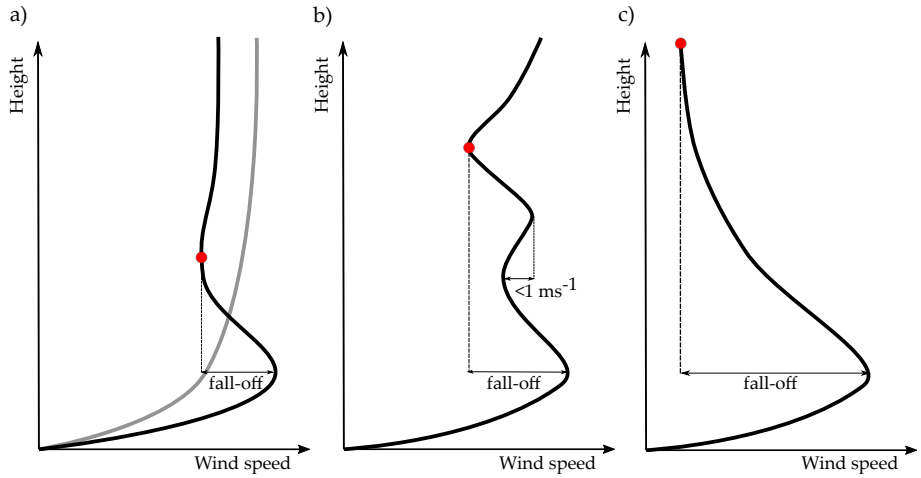

Figure 4. Schematic view of LLJ detection criteria. Red dots indicate the minimum in the wind profile used for calculating the fall-off. The grey line on (a) represents a "standard" logarithmic wind profile. (**a**) A maximum on the wind profile is considered as a LLJ if it is 1 m s^{-1} faster than the minimum above. (**b**) A minimum is neglected if the wind speed upwards increases less than 1 m s^{-1} before decreasing again. (**c**) If no minimum is detected, the wind speed at the top of the profile is considered as minimum.

course is performed and a comparison between the values retrieved by the observations and the reanalyses. Later, the sensitivity of the models on the different LLJs properties is assessed. Finally, the influence of the models' temporal and spatial shifts on their performance is investigated.

3.1 Wind speed comparison

Before analyzing the morphology of the jets, a comparison between the wind speed retrievals of the three used datasets is presented. Figure 5 shows the scatter plot and regression lines of the hourly averaged values retrieved by the lidar and the numerical models. As can be observed, the coefficient of determination (R^2) reaches values of 0.798 and 0.897 for NEWA and ERA5, respectively. These amounts are in line with the results found in ~~(Witha et al., 2019)~~Witha et al. (2019), where several numerical models were compared against the measeurements from the NEWA Ferry Lidar Experiment and obtaining coefficients of correlation (R) of 0.899 for NEWA and 0.946 for ERA5. The small differences between the coefficients found in our paper and in Witha et al. (2019) are due to the different filtering and data quality approaches implemented as well as the measurement-models co-location procedures.

The data suggest that there is a fair agreement between both reanalyses and the observations, although ERA5 performs slightly better than NEWA, as it was also found in ~~(Witha et al., 2019)~~Witha et al. (2019). The better performance of ERA5 can be a consequence of its more frequent updates of the analysis fields, in contrast to the long-term forecasts used in an atlas-like model such as NEWA. Additionally, and even though models with a high resolution are capable of more realistically

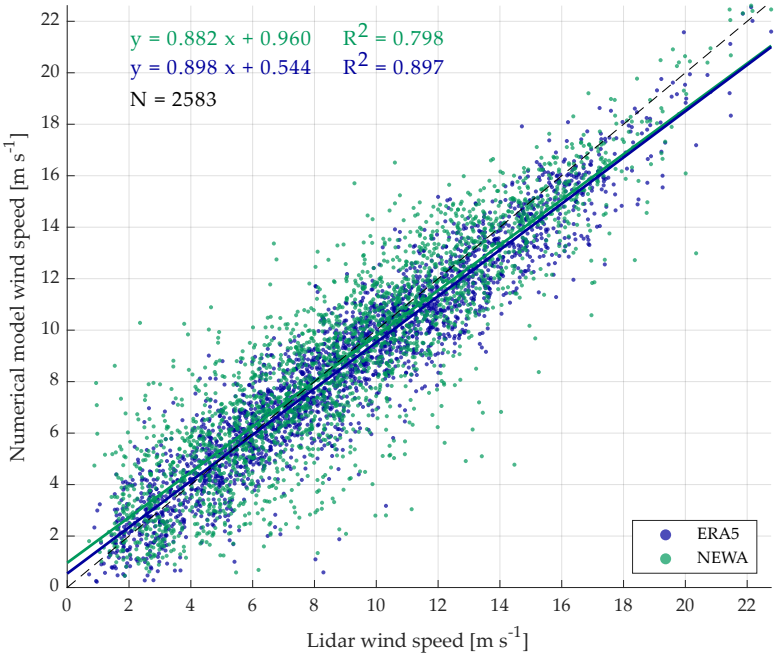

Figure 5. Scatter plot of wind speeds measurements at 100 m from lidar observations and numerical models. Linear regression lines are included with corresponding colors, as well as the linear regression equations and coefficients of determination. The dashed black line represents the $y = x$ line. N indicates the total number of data points considered in the comparison.

capturing the local features of the wind field, it is known that models with a coarser resolution can achieve better standard verification metrics (Murphy, 1988; Warner, 2010).

Figure 6a shows the wind speed kernel distributions at 100 m height. Both numerical models satisfactorily capture the wind speed distribution, although ERA5 shows a considerable overestimation of the frequency in the most common wind speed range. Furthermore, both models underestimate the frequency of higher wind events. Regarding the wind speed profiles, shown in Figure 6b, ERA5 underestimates the wind speed by a nearly constant amount of 0.3 m s^{-1} along the entire vertical profile. In contrast, the NEWA profile is approximately unbiased at heights close to the surface, but the disparity with the measurements progressively increases with the height, reaching a bias of aproximately 0.5 m s^{-1} at the upper part of the profile. Therefore, on average, NEWA has a smaller wind shear than ERA5 and the observations, with slightly higher wind speeds at the bottom of the wind profile but lower speeds at the top. Apart from this, it can also be observed that both models retrieve wind profiles with slower wind velocities. This overall underestimation of the wind is consistent with the results found in ~~(Hallgren et al., 2020; Kalverla et al., 2020)~~Hallgren et al. (2020); Kalverla et al. (2020), where similar biases were found in different locations of the Baltic and the North Sea.

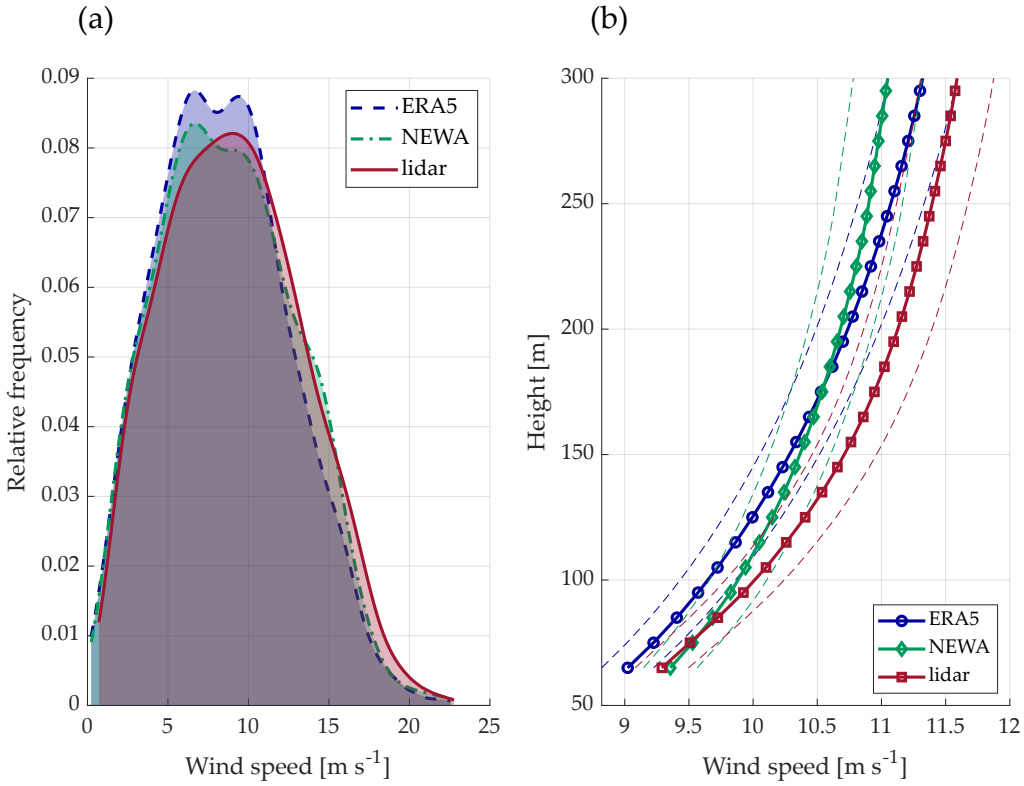

Figure 6. (**a**) Wind speed kernel distribution at 100 m height. (**b**) Average wind speed profiles. The dashed lines represent the 95 % confidence intervals.

3.2 Low-level jet properties along the ship track

3.2.1 Daily cycle of low-level jets

Figure 7a presents the diurnal frequency cycle of LLJs for the three datasets. Figure 7b shows the hourly average distance to shore and fetch length (horizontal distance, in the direction of the wind directions, over which the wind has blown without obstruction) for each hour of the day. The vessel route from Klaipeda to Kiel and vice-versa takes around 20 hours, and after each journey, the ship is in the harbor for approximately 4 hours (see Figure 2a). Therefore, the ship location follows a cycle of 24 h, meaning that its hourly position is approximately the same every day. Consequently, the diurnal cycle in this figure represents the different occurrence of LLJs over the day while passing through the several regions covered by the ship track.

The two numerical models considerably underestimate the frequency of LLJs in the vast majority of hours. ERA5 is the model with the lowest occurrence during the period under study, with jet events in the 3.6 % of the hours compared to the 5.5 % and 9.4 % of NEWA and the measurements, respectively. These results agree with the findings from previous studies (Hallgren et al., 2020; Kalverla et al., 2020). To a large extent, this ~~misestimation~~ underestimation is caused by

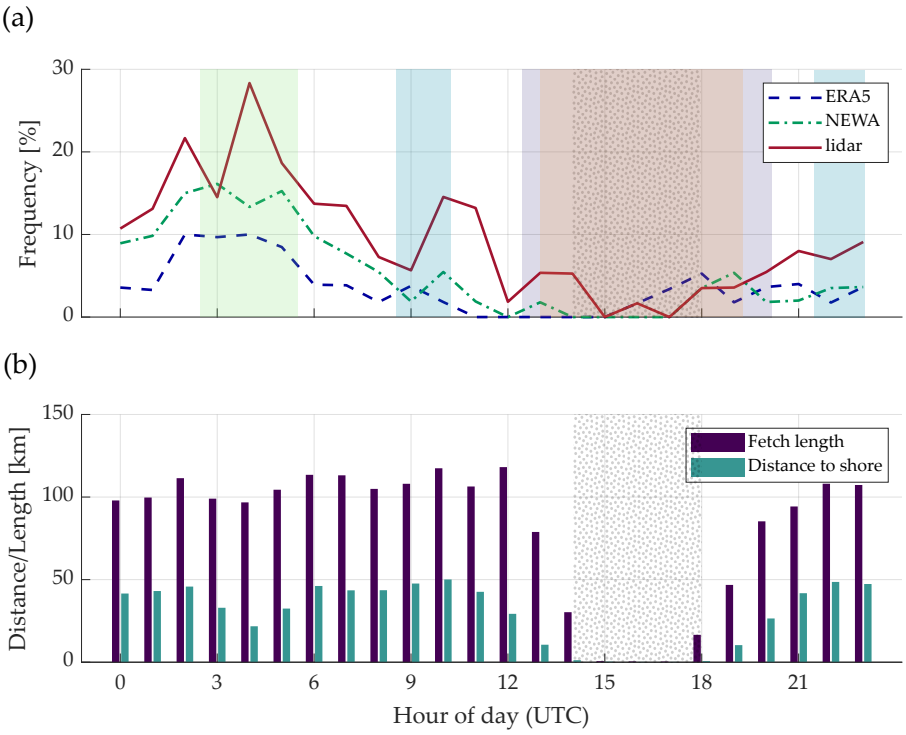

Figure 7. Hourly frequency of low-level jet occurrence (**a**) and mean distance to shore/fetch length for each hour (**b**). The grey dotted area indicates the time interval when the ship is in harbor. The coloured shadowed areas refer to the periods where the ship is at the locations indicated in Figure 8a.

the parameterization of turbulence in the models, which tend to overestimate the turbulent mixing during stable conditions (Cheinet et al., 2005; Sandu et al., 2013; Holtslag et al., 2013) typically appearing during spring in the Baltic Sea, and thus, to disguise anomalies such as LLJs in the wind profiles~~(Cheinet et al., 2005; Holtslag et al., 2013)~~. Additionally, previous studies (Kalverla et al., 2019) concluded that numerical models typically locate LLJs too high in the atmosphere, which together with the profile limitation of 300 m used in this study (due to the lidar device height range) results in fall-offs above the jet core that are too weak to be considered as LLJ.

Analogously to the results exposed in (Svensson et al., 2019a, b), most of the LLJs develop during the nighttime, with a maximum frequency of around 30 % at 0400 UTC according to the lidar measurements. These LLJs usually appear as a consequence of the development of stable stratification conditions, the advection of warm-air started during the preceding day, or LLJs transported from land that are generated as the results of nocturnal cooling over the land surface (Svensson et al., 2019a). On the contrary, the period with the lowest amount of LLJ occurs between 1400 and 1800 UTC, concurring with the period when the vessel is in the harbor and thus, when onshore microscale phenomena characterize the local wind conditions. Although jets can also form in ~~nearshore locations~~ the locations near the harbor, their development occurs typically at upper heights than in offshore sites (Nunalee and Basu, 2014), and considering the maximum heights of 300 m of the profile, may

account for the absence of LLJ incident in these hours. Additionally, onshore LLJs usually develop during the nighttime due to the reduction of the turbulence and the consequent development of stable stratification; nonetheless, the vessel is usually far away from the ports during the night, which contributes to ~~explains~~ explain the absence of jet events during the central part of the day. This onshore daily cycle (from 1400 to 1800 UTC) agrees with the one obtained through a fixed onshore met mast in (Baas et al., 2009), where a minimum LLJ occurrence is observed during the daytime with a progressive increase starting at approximately 1700 UTC.

3.2.2 Jet properties at different fixed locations

Figure 8 includes the average values of the LLJs frequency, core height, and core speed at four different locations along the ship track ~~. Two of these locations~~ using co-located values of models and observations in both time and space. These locations have been selected aiming to evaluate the datasets in sites with predictably different LLJs´ characteristics (locations A and D can be classified as onshore ~~(A and D in Figure 8a) , whereas the other two have mainly offshore characteristics~~whereas B and C as offshore) and assuring the existence of a certain amount of jets for the derivation of consistent statistics. The mean values in this figure have been calculated using wind profiles up to 300 m for the three datasets, although values with profiles up to 500 m have also been assessed in the reanalyses to evaluate the effect of profile upper limit in LLJs frequency and properties.

As can be observed in Figure 8b, there is an apparent disparity in the occurrence of LLJ between the offshore and ~~inshore~~ onshore areas. While their frequency is below 3 % in sites A and D for the three datasets, B and C show higher occurrences and exceeding 15 % frequencies according to lidar measurements. As mentioned above, both numerical models underestimate the frequency of jets when only 300 m profiles are considered, except for position D, where ERA5 has a slightly higher occurrence. When increasing the top limit of the models profiles up to 500 m, the frequency raises ~~considerably~~ substantially in all locations, with an exceptionally remarkable increase in offshore positions. ~~With regard to the~~ This increase can be explained by three main reasons. First, the tendency of numerical models to position the jets too high in the atmosphere, as observed in (Svensson et al., 2018; Kalverla et al., 2019) and thus, the consideration of jets that are not seen when only 300 m profiles are scanned. The second potential explanation is the excessive flattering of the wind profiles modelled by the reanalyses during stable conditions (Cheinet et al., 2005; Sandu et al., 2013; Holtslag et al., 2013), that leads to a too weak negative shear above the jet core and the resulting requirement of a higher profile top height to exceed the fall-off threshold value. And finally, the inherent characteristic of the LLJ detection algorithm that hinders the detection of weak jets located close to the upper limit of the profile top height.

In onshore areas, the extension in the wind profile height has a considerably more pronounced effect in ERA5 than in NEWA, resulting in an increase of the events from 1.9 % to 9.5 % in position D, for instance. Nonetheless, considering the offshore sites, the impact of the profile extension on the two numerical models differs depending on the location examined. On the one hand, it causes that both reanalyses have frequencies higher than the corresponding measurements in B, with an overestimation in the ~~appearance~~ occurrence of 140 % and 80 % for ERA5 and NEWA, respectively. On the other hand, at location C, the frequency of ERA5 is slightly increased compared to the 300 m profiles case, maintaining a value still below the frequency

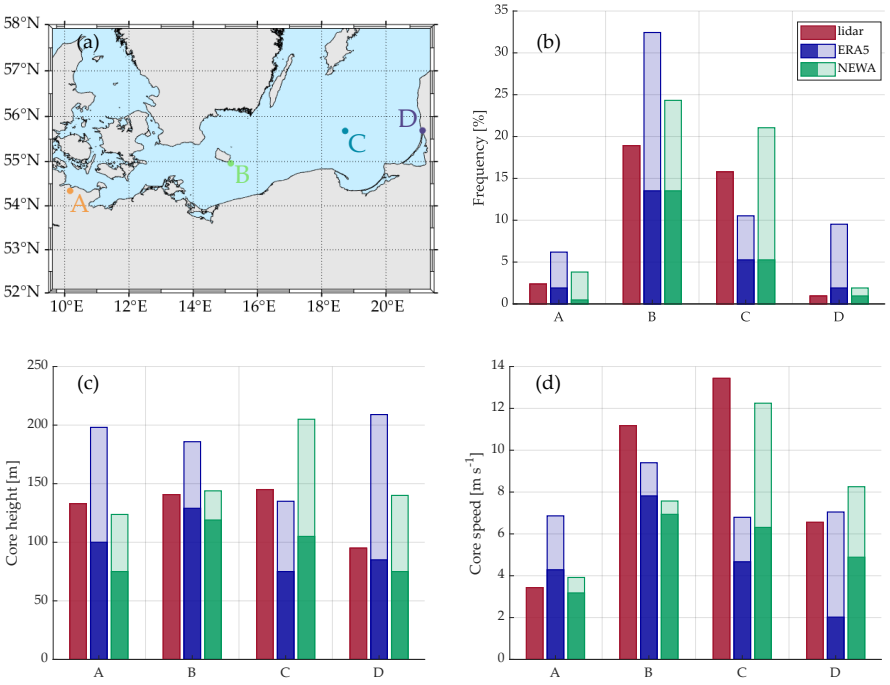

Figure 8. (**a**) Map with the four locations (A, B, C and D) where LLJ properties were calculated. Average values of LLJ frequency (**b**), core height (**c**) and core speed (**d**) at the four location. The plain filled bars indicate the values obtained when a 300 m profile is considered. The semi-transparent areas represent the increase when a profile of up to 500 m is used for the calculation (only for the numerical models)

of the observations. For NEWA, the jet occurrence reaches a value of 21 %, which is 5.2 percentage points beyond the lidar measurements.

All datasets agree on a mean core height smaller in the ~~nearshore~~ onshore areas than the offshore ones when looking at the profiles up to 300 m (except ERA5 in C). However, both numerical models underestimate the mean height in all locations, with NEWA as the dataset with the lower jet height values in most sites. ~~The increase in the~~ Analogously to the jets frequency, increasing models wind profiles up to 500 m ~~increments~~ results in a substantial increment in the mean core height in all locations, although it is striking that particularly for ERA5, this rise is higher in the ~~inshore~~ onshore locations than in the far-offshore ones.

The core speed is considerably lower in ~~nearshore~~ the harbors places as a result of the weaker mean wind speed and the lower mean core height in these locations. Although both numerical models obey this trend, they show different mean values compared to those given by the observations. According to the measurements, the more offshore, the higher the average core speed, with a maximum value at location C of 13.4 m s^{-1}. However, numerical models show their maximum values close to Bornholm (location B), where again, the influence of the island may affect the performance of these datasets. The increase in the wind profile results in a rise in the jet velocity proportional to the variation in the core height, which confirms the strong

relationship between the core height and velocity. Both models show mean values of core speed lower than those by the lidar even when considering 500 m and offshore location, highlighting the systematic underestimation of the wind speed by the models and their difficulties to retrieve extreme cases with higher wind speeds.

In order to separate the temporal and the local effects, Figure 9 presents the daily cycle of these LLJ characteristics at the four locations aforementioned. For this, we used data from the two numerical models at the corresponding nearest grid point, considering the entire period of the measurement campaign and profiles up to 300 m. The four locations present different patterns in the daily cycle of jets, being possible to discern between two distinct trends. On the one hand, onshore sites (Figures 9a and 9j) show no LLJs during the central hours of the day (from around 0600 to 1600 UTC) and maximum frequency values during the night and the early morning. On the other hand, ~~nearshore~~ onshore locations do not show such an obviously defined daily cycle, although on average, a considerable higher mean occurrence than for onshore sites.

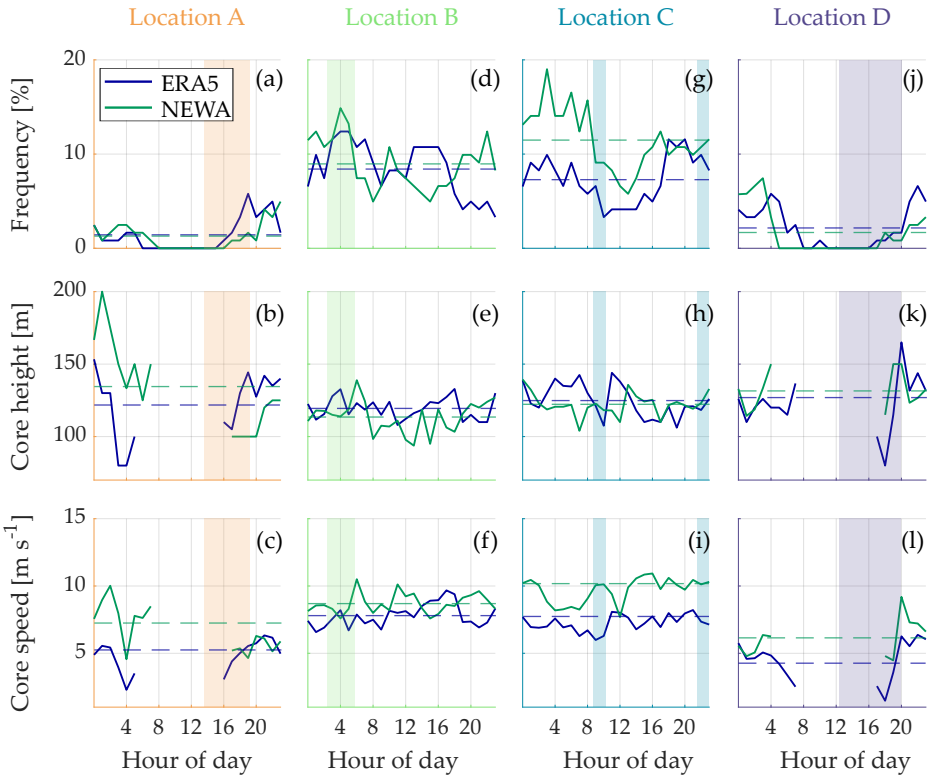

Figure 9. Frequency, core height and core speed daily cycles at the selected four location based on ERA5 and NEWA. The shadowed areas indicate the time intervals when the vessel is close (within 10 km distance) to the corresponding location. Dashed lines are the mean values.

Coloured areas in Figure 9 mark the periods when the vessel is next to the respective location (within a distance of 10 km). Interestingly, the trends in LLJ frequencies observed in these shadowed zones can also be identified in the daily cycle presented in Figure 7a. From 1400 to 1900 UTC, when the ship is near the harbor, the LLJ frequency is null during the early afternoon

and progressively increases up to values of around 5 % from 1600 UTC. Additionally, the frequency peak situated at 0400 UTC in Figure 9d, also coincides with the maximum in Figure 7a, although in this last plot, the maximum is less pronounced than the one in Figure 9d. Regarding location C, the overlapping period in which the ship is close is divided into two regions. One in the morning between 0900 and 1000 UTC; and the other in the late night from 2200 to 2300 UTC. In the first period, the frequency of LLJs is reduced compared to the one in the early morning but still more significant than the occurrence during the afternoon. In the second one, the frequency is higher compared to the afternoon and lower than in the early morning. These trends are correctly reproduced by the daily cycle presented in Figure 7a, although it can be highlighted that the frequency values shown there are lower than those indicated in Figure 9. The atmospheric features in offshore areas that lead to the generation of marine LLJs have a considerably weaker daily cycle compared to those in onshore locations (Liu and Liang, 2010), inducing a more constant amount of jet events throughout the day in offshore territories as observed in Figure 9. However, daily frequencies in Figure 7a exhibit a drastic variability throughout the day, partly caused by the variable ship position during the day, and partly by the temporal variance of wind conditions during the different times of the day.

The core height and speed plots do not show a clearly defined daily cycle in the marine locations. However, there are pronounced oscillations in the onshore locations that may be a consequence of the smaller number of events detected during the morning and evening. Differently to what can be observed in Figure 8c, offshore locations present slightly lower mean core heights compared to the onshore ones, as a consequence of the notable increase in the mean core height in locations A and D. With regard to the core speed, and analogously to the mean values presented in Figure 8d, offshore sites have higher mean jet velocities, with NEWA showing values above those from ERA5.

Additionally, the mean values of these characteristics for the shadowed areas (considering only the hours where the ship is in that location) are presented in Table 3. The occurrence values are lower than those shown in Figure 8b for both models and locations except in C. Thus, separating the temporal effects on the diurnal cycle increases the models' underestimation in these regions compared to the occurrence obtained when considering the ship track. It is striking that NEWA frequencies in ~~near-shore~~ onshore locations are lower than those from ERA5, contrary to sites B and C, where NEWA shows higher values. Comparing the results from this table with those from Figure 8c, ERA5 shows a higher mean core height than NEWA in most positions (A, B, and C). However, both reanalyses show increased values in this table for all locations, except for the moderate decrease in location B. This means that the mean values for the jet height presented in this table are closer to the lidar measurements than those in Figure 8c. Regarding the jet velocities, ERA5 shows very similar values to those in Figure 8d for locations A and B, but for sites C and D the table shows higher values, showing that the consideration of the ship track emphasises the underestimation of this variable. NEWA gives faster jet velocities in the four locations compared to ERA5 and the table values are higher than those in Figure 8d.

3.2.3 Frequency bias

Models' ability to accurately describe the marine boundary layer features allows them to perform better in offshore regions than for on- or nearshore areas. For this reason, a better characterization of LLJs in far-offshore locations is expected. To evaluate how the ability of the models to detect jet events varies with respect to the distance to shore, Figure 10 presents the frequency

Table 3. Mean values of main low-level jets characteristics according to ERA5 and NEWA. Wind profiles up to 300 m and only data from those hours where the ship is in this location (within 10 km distance) are used (corresponding to shadow areas in Figure 9)

	Location A		**Location B**		**Location C**		**Location D**	
	ERA5	**NEWA**	**ERA5**	**NEWA**	**ERA5**	**NEWA**	**ERA5**	**NEWA**
Frequency [%]	0.8	0.2	10.5	12.4	7.9	11.4	0.4	0.3
Core height [m]	115.0	100.0	125.4	115.5	122.6	121.2	98.3	132.5
Core speed [m s^{-1}]	4.2	5.3	7.6	8.2	7.1	9.9	2.5	4.6

bias (FBIAS) depending on the coastal distance, calculated as the ratio between the number of LLJs predicted by the numerical models and the observations:

$$\mathrm{FBIAS} = \frac{hits + alarms}{hits + misses} \frac{hits + false\ alarms}{hits + misses} \tag{1}$$

As can be observed, FBIAS shows values close to the perfect score value of 1 for distances below to 60 km.Nevertheless, when considering further distances a deterioration of the FBIAS indicates that both reanalyses struggle with capturing LLJ events that occur far away from the shore. This fact can emerge from different factors, like the site-specific properties of LLJs in the different regions, the time windows when the ship is in the area of interest, and the limitation of observations up to 300 m height. Additionally, it can be noted that both models present FBIAS smaller than 1 for all distances, showing again the systematical underestimation of the number of events.

Apart from the FBIAS, the ratios between models and observations have been assessed for the core height and speed. Nonetheless, no correlation has been found between the distance to the coast and the score of the ratio.

3.3 Sensitivity of models performance on low-level jet characteristics

In order to evaluate how low-level jet characteristics influence the capability of numerical models to capture these phenomena, Figure 11 is presented. In this figure, the boxplots of the different LLJ features (i.e., core height, core speed, and fall-off) are included, being classified according to whether an LLJ is detected by both the numerical model and the observations (hit); whether it is identified by the observations but not by the numerical model (miss); or whether it is present in the reanalyses but not in the lidar dataset (false alarm). The occurrence of each kind of these events is indicated in Table 4.

The average core height is considerably well predicted by the two models for those LLJs classified as hits, although ERA5 is capable of more accurately assessing this property. ~~.~~However, as can be seen for both models, the mean core height of the lidar observations is more prominent for those events classified as misses than for the hits, meaning that LLJs situated closer to the upper limit of the vertical profile are frequently missed by the numerical models. As previ-

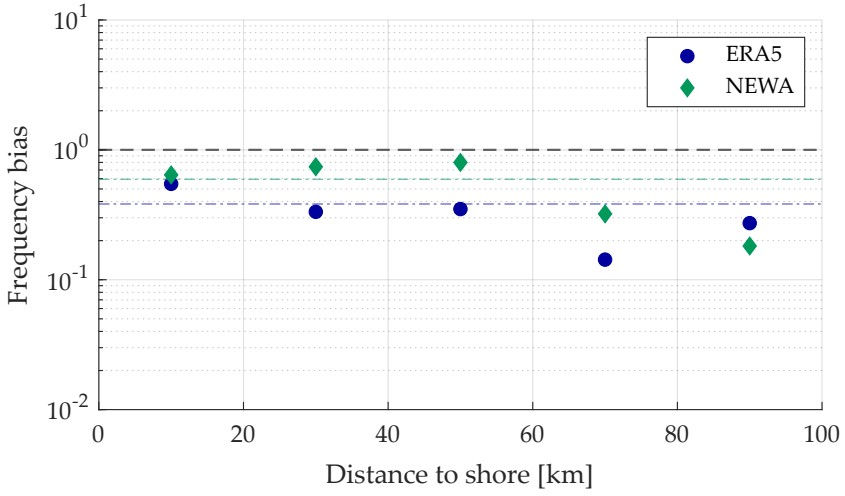

Figure 10. Bin-averaged frequency bias (FBIAS) of low-level jets for the two numerical models against distance to shore (20 km bins width have been used). The dash-dotted lines represent the average frequency bias for the complete dataset. Black dashed line represent a perfect score (FBIAS = 1).

Table 4. Number of hits, misses and false alarms for each numerical model.

	ERA5			NEWA		
	hits	**~~mises~~misses**	**false alarms**	**hits**	**misses**	**false alarms**
Number of cases	28	111	24	38	101	43

ously mentioned in this study, this can be a consequence of the models' tendency to place LLJs too high in the atmosphere ~~(Kalverla et al., 2019; Svensson et al., 2018)~~(Svensson et al., 2018; Kalverla et al., 2019), leading to fall-off values above the core too weak to be considered as LLJ or to directly disregarding wind maxima appearing in upper heights of the models wind profiles. Additionally, there is a clear underestimation of the core height by the models in the jets classified as false alarms, suggesting that both models identify LLJs at the bottom part of the profile, although they are not present in the measurements.

LLJs can develop within an extensive range of jet speeds, with events identified from 2 to 20 m s^{-1} according to the lidar observations. Analogously to the core height, both numerical models correctly predict the average core speed of events classified as hits, with differences of 14.0 % and 9.3 % for ERA5 and NEWA, respectively. This bias is a consequence of the underestimation of this parameter, more evident in ERA5. It is also striking that in contrast to the observations, ERA5 LLJs happen only up to a core speed of around 12 m s^{-1}, a fact that exemplifies the limitations of this dataset for retrieving extreme wind conditions.

Contrary to the core height and velocities, none of the two numerical models is capable of accurately predicting the fall-off values in those events classified as hits, underestimating the mean value by 1.9 m s^{-1} in the case of ERA5 and 1.7 m s^{-1} for

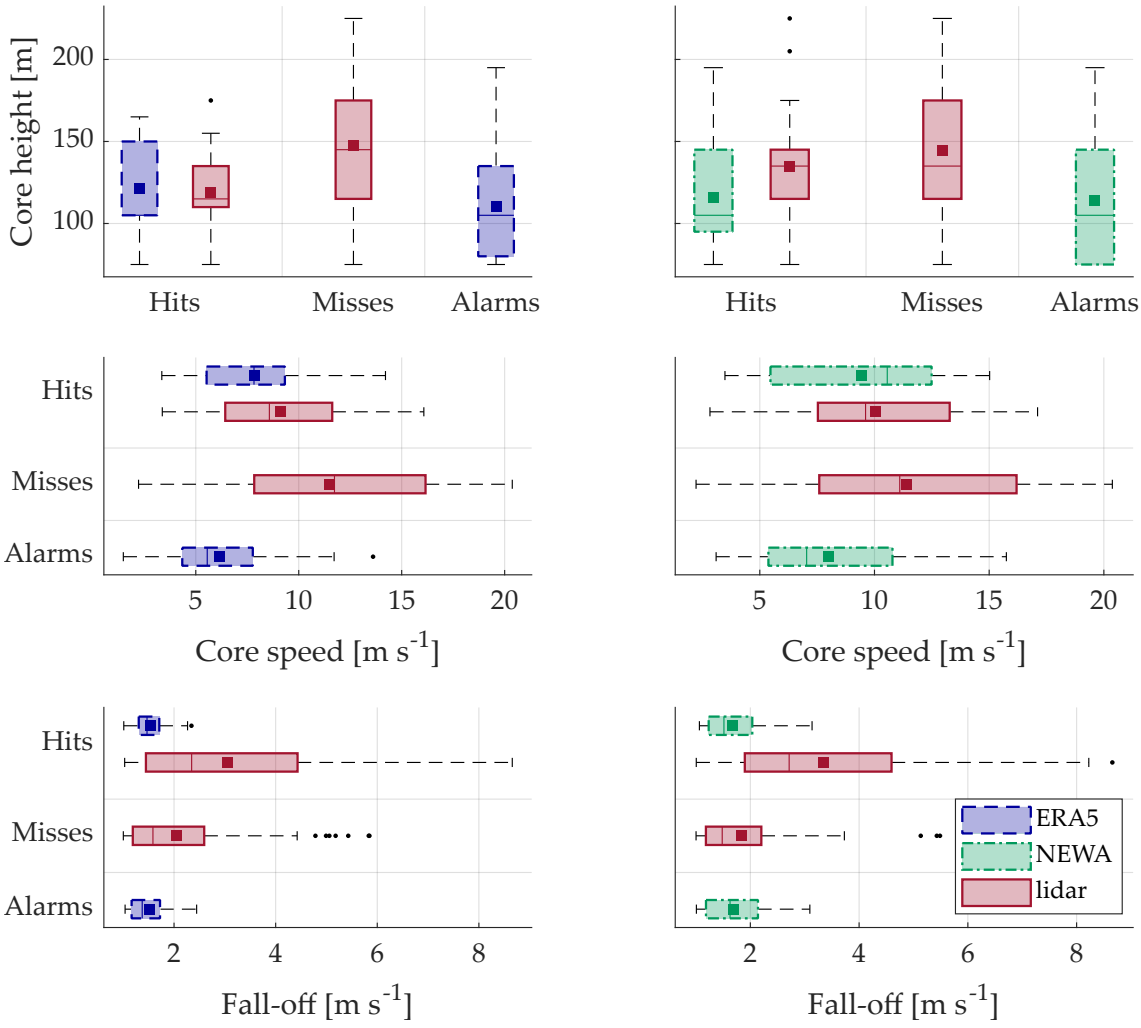

Figure 11. Boxplots showing the core height (top row), core speed (middle row) and fall-off (bottom row) for the different datasets, and classified according to hits, misses and false alarms events. The bottom and the top edges of the boxes indicate the 25th and 75th percentiles, respectively. The line inside the box is the median and the square is the mean. The whiskers extent to the extremes, defined as a distance of 1.5 times the interquartile range (IQR) above and below the upper and lower quartiles, respectively. The outliers are represented by the black dots.

NEWA. The mean fall-off for the hits is considerably higher than the one corresponding to the misses, suggesting that those jets with stronger fall-offs are easier to identify by the numerical models rather than LLJs with fall-off values below 2.5 m s^{-1} in the case of ERA5 and 2.1 m s^{-1} for NEWA. The lower fall-off values for the misses can occur as a consequence of two main factors: Firstly, it can be due to the smearing out of the models' wind profiles and, therefore, a decrease in the possible jets' strength that provokes the models to skip those events with weaker fall-offs. Secondly, ~~as a result of the location of~~ the tendency of numerical models to locate LLJs very high in the ~~vertical profile , leading to weaker jets captured by lidar but missed by the models~~profile may result in weak jets with fall-off values below the considered threshold (see Subsection 2.4). Apart from this, the high number of outliers in the misses events suggest that although numerical models miss jets with lower fall-offs, they can neither retrieve LLJs during higher wind speed conditions.

3.4 Time and spatial shift

When comparing numerical models and observations, their different spatial resolution may result in distinct capabilities to feature wind characteristics at the point where the observation is retrieved. It is possible that the closest model grid point to the observation does not correctly resemble the wind characteristics, for instance, in a coastal location, but another surrounding grid point does it better. Additionally, the various temporal resolutions of the datasets may influence the capabilities of the models to capture the temporal variations of the wind and the phase when a particular phenomenon, such as LLJs, occurs. In order to understand the influence these two considerations have in the models' ability to identify LLJ, their performance has been evaluated in three particular cases, schematically represented in Figure 12:

- **Case a** (reference case): as explained in Section 2.3, for each hourly position of the vessel, the nearest grid point of each reanalysis has been selected. Then, the lidar and models wind profiles are evaluated to determine the presence of LLJs;
- **Case b**: for each hourly position of the ship, the grid points of each model inside a threshold radius R (3 km for NEWA and 17 km for ERA5, according to their minimum resolution) are selected. The existence of jets is evaluated in the lidar profile and in all numerical model´s profiles at the grid points inside the threshold. If an LLJ is identified in any of these points at the considered hour T_0, that timestamp is marked as positive with regards to LLJ occurrence for the corresponding model. In this case, the influence of the potential spatial shift in model performance is evaluated;
- **Case c**: for each hourly position of the ferry boat, the nearest grid point is selected (analogously to Case 1). However, in this case, for the corresponding hour T_0 at that position, the presence of LLJs was investigated in the vertical profiles of the models at the hours T_0, T_0-*1*, and T_0+*1*. When an LLJ is detected in at least one of these three timestamps, we consider there is an LLJ at hour T_0 in the corresponding model. In this case, the effect of temporal shift is considered.

It may be noticed that others approaches are also available for this evaluation, such as interpolating the nearby model data to the measurement location or the combination of a temporal and spatial window. However, the used cases in this study were selected in order to independently evaluate the potential effect of the models' temporal and spatial offset.

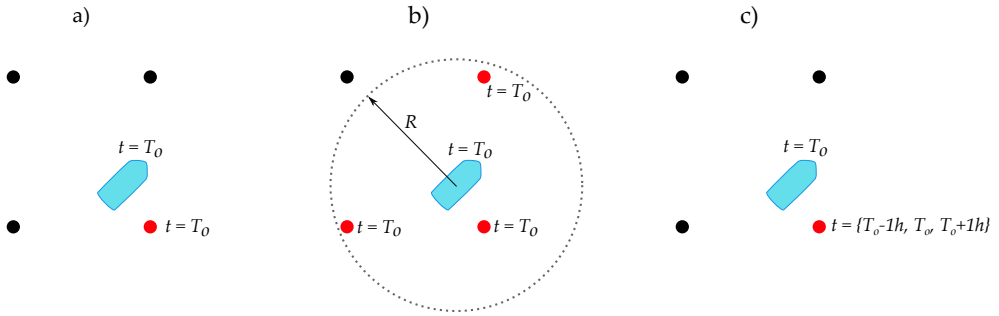

Figure 12. Sketch of the three cases considered to evaluate the temporal and spatial shift of the reanalyses. For each case, the ship and the surrounding grid points are included. The grid point(s) used for comparison against the measurements is/are colored in red. The timestamps used for comparison are also indicated.

For the three aforementioned cases, the number of hits, misses, and false alarms are applied to estimate the performance of models by calculating two skill score indicators. The first indicator used is the symmetric extreme dependency score (SEDS), defined in ~~(Hogan et al., 2009)~~ Hogan et al. (2009) as:

$$\mathrm{SEDS} = \frac{\ln[(hits + alarms)/n] + \ln[(hits + misses)/n]}{\ln(hits/n)} \frac{\ln[(hits + false\ alarms)/n] + \ln[(hits + misses)/n]}{\ln(hits/n)} - 1 \quad (2)$$

where n is the total number of observations included in the datasets. This parameter ranges from -1 to 1, with 1 indicating a perfect estimation by the model (all LLJs events are labeled as hits), 0 indicating a random forecast, and -1 meaning that no LLJ is classified as hit.

The second used indicator is the frequency bias, already presented in Section 3.2.3. A value of 1 of the FBIAS indicates a perfect score. A value greater than 1 implies an overestimation of the number of events and vice versa.

These two parameters give information about the performance of the models in predicting LLJ events. The main difference is that the SEDS penalizes the model performance when there are phase errors (misses and false alarms), whereas the FBIAS only considers the total number of LLJs, ignoring the timing of these events. Because of this, SEDS gives meaningful information about the forecast capabilities of the model, while the FBIAS measures its climatological performance.

Table 5 shows the values of these indicators for the three considered cases. As can be observed, the underestimation of the LLJ events is evidenced by the results of skill scores in any of the three cases, with values of SEDS and FBIAS are smaller than 1. However, considering the potential spatial and temporal shift notably improves the climatological performance of the two models in cases b and c, reaching FBIAS values of 0.56 for ERA5 and 0.94 for NEWA in the last case. Although this increase can be seen both in cases b and c, it is more notable in the latter, suggesting that the temporal shift has a more relevant influence on the performance of both models. It is also interesting that the spatial shift has stronger repercussions in ERA5 than NEWA, with a respective improvement in the FBIAS of 26 % and 9 %. This difference may arise as a consequence of the worse spatial resolution of ERA5, which takes a bigger advantage when considering additional grid points close to the ship

location when evaluating jet events. However, despite the noticeable improvement in the FBIAS in cases b and c, the rise in the number of misses and false alarms compared to the reference case impede the enhancement of the SEDS's score, that remains almost constant for all the cases. NEWA has better FBIAS values for all the considered cases but similar values of SEDS. The reason for this is that although numerical models with finer resolutions are typically able to capture fine-scale structures better, they are more likely to have mismatches in the phase of the events (Kalverla et al., 2020).

Table 5. Skill scores indicators for the two reanalyses and the three considered cases.

	Case a		Case b		Case c	
	ERA5	**NEWA**	**ERA5**	**NEWA**	**ERA5**	**NEWA**
SEDS	0.53	0.52	0.55	0.53	0.52	0.56
FBIAS	0.37	0.58	0.49	0.65	0.56	0.92

Despite SEDS is commonly used as a forecasting score, in this case, it can also be used as a control parameter to assess the meaningfulness of the comparison performed, and thus, the validity of the FBIAS values obtained. For example, considering a case b with a larger threshold radius would increase the number of LLJs identified by the models and hence raise the FBIAS. Nevertheless, this higher number of events is caused by the consideration of LLJs happening far away from the location where the observation is taken, leading to a degraded value of the SEDS. Therefore, the SEDS is a useful parameter that can be used as an indicator of the applicability of the obtained FBIAS when using numerical models to evaluate the climatology of LLJs.

4 Discussion

This study characterized LLJ distribution, properties, and occurrence over the Southern Baltic Sea by means of ship-based lidar measurements and two mesoscale numerical models. To this end, a methodology has been presented to assure a fair comparison between the different datasets involved and their different temporal and spatial characteristics.

We started with a comparison between the wind speed retrievals of the three datasets used in this paper. Although the statistics scores evaluated in this study show a good performance and are in line with the results found in previous studies (Witha et al., 2019), both numerical models exhibit a systematic underestimation of the wind speed evidenced when comparing the frequency distributions and mean vertical profiles of the three databases.

The weak daily cycle of the atmospheric features that lead to the generation of offshore LLJs results in a constant number of jets during the whole day ~~Liu and Liang (2010)~~(Liu and Liang, 2010). However, the daily occurrence of LLJs calculated along the ship's route evidences a remarkable variability, due to the combined influence of both temporal and spatial effects. On the one hand, the majority of the jet events occur while the ship is offshore during the night and the early morning as a consequence of the generated stable stratification conditions, the advection of warm-air started during the preceding day, or the transport of LLJs generated as the results of nocturnal cooling over the land surface from the mainland to offshore Svensson et al. (2019a). On the other hand, the lowest amount of LLJs is detected during the afternoon, coinciding with the period when

the ship is on or very close to the harbors. Both numerical models ~~are able to~~ can identify the LLJs diurnal cycle correctly. Nevertheless, these datasets manifest a substantial underestimation in the number of LLJs during the vast majority of the hours, and therefore, also along the different regions covered by the ship track.

~~In order to evaluate the~~ The capabilities of the ~~ship-based observations to capture the frequency of jets at four sites distributed along the vessel track, we used the numerical models ' output data to calculate the~~ two numerical models to accurately model the main LLJs characteristics have been evaluated at four specific sites. In this case, numerical models also show a clear underestimation in the number of jet events, although this underestimation is less pronounced in the onshore sites and for ERA5. However, both models agree with the measurement in terms of the locations with the highest and lowest number of LLJs, with B showing maximum value and the onshore sites presenting the lowest number of events. Regarding the core height, reanalyses also show lower values compared to the observations. Numerical models agree with the measurement regarding the core speed, showing considerably higher values in offshore areas than onshore, although the disagreement with the measurements varies depending on the specific site and model considered. Increasing the vertical extension of the models' profiles up to 500 m substantially modifies the obtained values, increasing the mean values for the three properties and four locations evaluated and exposing the relevance of this fact in the characterization of LLJs using the reanalyses.

The daily cycle of ~~jets' main properties~~ the jets' characteristics (occurrence, core height, and core speed) at ~~those specific locations for the whole duration of the measurement campaign~~the aforementioned four locations have been also studied using the models' output data. The daily cycle observed at the four locations in those periods when the ferry is nearby can also be identified at those same periods in the diurnal cycle along the whole vessel route (Figure 7). Therefore, the ship-based lidar technology has ~~a~~ potential applicability for characterizing the occurrence of jet events within ~~a vast region~~ the vast region covered by the ship track. Nonetheless, this applicability is highly limited by the constant translation of the ship ~~, that exclusively allows to characterize the frequency cycle in those time frames~~ and by the relation position-time of the route, since only the characteristics of the jets when the vessel is near the location of interest can be derived.

Generally, models are capable of more accurately retrieve wind conditions in offshore locations, where the microscale phenomena are less relevant. However, the results found in this study show a deterioration of the FBIAS with the increment of the distance to the coast. With regards to the core speed and height, no correlation has been found between the performance of the models and the distance.

~~The fluctuations of the core speed and height of LLJs have also been studied at the aforementioned four locations.When considering 300 m height profiles, offshore sites show slightly higher values of the mean core height than near-shore locations, as well as considerably faster mean core velocities. Both numerical models show difficulties in predicting these features reliably, showing more or less accurate mean values depending on both the property considered and the location under study. In any case, most of the LLJs occurred at heights in the range of modern wind turbine rotors, and with jet strengths between 7 and 15 m s^{-1}, emphasizing their potential impact on wind turbine loads and performance.~~

~~The~~

The limitations of the datasets used in this study (i.e. profile height limitation and relation time-position) must be properly considered when interpreting the obtained results. Additionally, the different characteristics of the LLJs play a fundamental

role in the capacity of the models to identify these phenomena. Both numerical models have more difficulties with resembling those LLJ with core ~~height~~ heights closer to the upper limit of their wind profile, which can result from the tendency of the numerical models to place LLJ excessively height in the atmosphere and a too weak fall-off above the core. According to the measurements, LLJ can occur within a broad range of core velocities, ~~and~~ from very calm to stormy conditions. Furthermore, both numerical models undervalue ~~this parameter~~the mean core speed values, possibly connected to the underestimation of the core height. It is striking that both reanalyses are failing in predicting the mean jets' fall-off velocities, with underestimations in both cases of 2 m s^{-1} ~~. Apart from this~~as a consequence of the excessively flattered wind profiles. Moreover, both models struggle in detecting those jets with lowers fall-offs, as well as extreme events with fall-off values above 4 m s^{-1}.

In the last section, the influence of the spatial and temporal ~~shift~~ shifts between the observations and models' jet events is presented. Both numerical models ~~considerable~~ considerably increase their climatological ~~score when considering this potential phase errors, reaching~~ scores when considering either a spatial or phase shift, with a particularly relevant improvement when the second is considered which leads to maximum FBIAS values of 0.555 and 0.938 for ERA5 and NEWA, respectively. On the contrary, the SEDS remains almost invariable for the three cases considered, as a consequence of the ~~remarkably~~ remarkable increase in the number of misses and false alarms when considering possible shifts in time and space. In consequence, further understanding ~~about the phase~~ of the spatiotemporal errors of the models ~~is required in order to develop~~ would contribute to developing an optimal strategy for applying numerical models for studying LLJs climatology over large regions.

All the results exposed in this study are based on ~~the~~ ship-mounted lidar observations, and thus, several considerations must be highlighted. Firstly, it must be noticed that measurements are subject to systematic and random errors that may influence the results and that lidar ~~systems~~ system are inherently affected by other uncertainties like the exact measuring height or the discard of raw measurements due to unfavorable atmospheric conditions such as low aerosols density, the presence of fog or low clouds. Furthermore, floating lidar systems ~~requires~~ require implementing measures to compensate for the ship motion effects on the measurements. Although a post-processing motion correction algorithm has been implemented, the uncertainty associated with the decontaminated ~~measurement~~ measurements is still unknown. Secondly, it is crucial to consider the pertinence of the mapping strategy and data availability when interpreting the obtained results. On the one hand, the available observations cover a period of around ~~3~~ three months, and therefore ~~,~~ they are unable to completely represent the wind climatology either over the whole region covered by the ship course or in specific areas within it. On the other hand, and due to the intrinsic non-stationarity of ship-based lidar measurements, the availability of the data at each measurement point is low and limited by the time window when the ship is near a considered location. Because of this, the observed values of the LLJ features at the different locations only include the behavior of this phenomenon during the site-specific time window. Therefore, ship-based lidar measurement campaigns require a careful evaluation and design of the mapping strategy to assure the output data~~''~~'s convenience and applicability, both for the general characterization of winds and the study of more specific phenomena.

5 Conclusions

Throughout this study, an effort has been made to ~~better understand how~~ define an implement a comparison methodology between ship-mounted lidar measurements and two state-of-the-art numerical models in order to investigate their accuracy to characterize LLJs over a wide region. The permanent translation of the ship does not allow deriving the complete daily cycle of the jets characteristics in a particular location. Nevertheless, and differently to fixed measuring devices, ship-based ~~lidar measurements can be used to investigate the characteristics of LLJs over a vast region, as well as a comparison with two well-known numerical models within the wind industry to evaluate their capabilities and limitations.~~

~~Ship-mounted lidars are a valuable technology to investigate LLJs along the vast region~~covered by the ship track. ~~This technology~~ systems can provide meaningful information about the jets~~'~~ ' properties and their temporal and spatial variations~~along different sites. Nevertheless, the permanent translation of the ship is a challenge when using these measurements to derive jets characteristics in a particular location. The measurements' interrupted temporal extension may lead to the derivation of jets' information that is only representative from specific time periods but not from the overall climatological features of the LLJs~~, as well as a highly reliable observations to compare numerical models against a reference dataset under different temporal and spatial effects.

~~The~~ It was shown that the incomplete representation of the physical phenomena hinders models from characterizing LLJs~~' properties and occurrence~~ ' features accurately. However, they capture ~~relevant information about the variability and behaviour of~~ the ~~jet characteristics, provided that a proper consideration of their limitations is involved.~~ variability of LLJs properties associated with the different locations where these are evaluated. The occurrence of jets is systematically underestimated, although this is further emphasized in offshore sites. The core height and speed values modelled by the reanalyses are usually also underestimated compared to the measurements, but this difference varies depending on the model and the characteristics of the location considered. NEWA was the model that best captured the occurrence of LLJs along the ship route, despite ERA5 shows an smaller bias regarding the mean core height at the four evaluated locations.

Apart from the physical constraints of the numerical models to resemble wind conditions, we conclude that reanalyses capabilities are strongly restricted by the inherent attributes of the LLJs, the features of the models (i.e., vertical and horizontal resolution), as well as other factors associated with the inherent characteristics of the available observations (e.g., the top height of the vertical profile, data availability, or the relation time-position). Additionally, considering the temporal and spatial shift between models and observations has shown a relevant potential to increase the capabilities of the models to investigate LLJs climatology.

Nowadays, the availability of ship-mounted lidar datasets is still scarce. Therefore, the execution of novel measurement campaigns using different mapping strategies, higher wind profiles, durations, and locations will yield more information about the capabilities of this technology and the numerical models. In addition, the large spatiotemporal extent of the numerical models offer an attractive alternative to counteract these inherent limitations of ship-lidar technology, highlighting the great potential of combining these different datasets to more accurately describe the temporal and spatial characteristics of jets over extensive areas.

Data availability. Data used for this paper was collected from the following sources. Lidar measurements were provided by Fraunhofer IWES and they are available upon request. The ERA5 data are freely available via the Copernicus Data Storage (CDS): https://cds.climate.copernicus.eu/cdsapp#!/home. NEWA data are available from https://map.neweuropeanwindatlas.eu.

Author contributions. Conceptualization, methodology, and project administration, H.R. and J.G.; investigation, software, formal analysis, visualization, writing and editing, H.R.; general guidance, review and supervision, J.G. and M.K.

Competing interests. The authors declare no conflict of interest.

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). The NEWA Ferry Lidar Experiment was founded as part of the NEWA project by the German Federal Ministry for Economic Affairs and Energy (ref. no. 0325832A/B) on the basis of a decision by the German Bundestag with further financial support from NEWA ERA-NET Plus, topic FP7-ENERGY.2013.10.1.2. Special thanks to Martin Dörenkämper (Fraunhofer IWES) for providing the NEWA data and Pedro Santos (Fraunhofer IWES) and Charlotte B. Hasager (DTU) for their suggestions to improve this paper.

References

Achtert, P., Brooks, I. M., Brooks, B. J., Moat, B. I., Prytherch, J., Persson, P. O. G., and Tjernström, M.: Measurement of wind profiles by motion-stabilised ship-borne Doppler lidar, Atmos. Meas. Tech., 8, 4993–5007, https://doi.org/10.5194/amt-8-4993-2015, 2015.

Andreas, E. L., Claffy, K. J., and Makshtas, A. P.: Low-Level Atmospheric Jets And Inversions Over The Western Weddell Sea, Bound-lay. Meteorol., 97, 459–486, https://doi.org/10.1023/A:1002793831076, 2000.

Baas, P.: Turbulence and low-level jets in the stable boundary layer, Ph.D. thesis, Wageningen University, 2009.

Baas, P., Bosveld, F. C., Klein Baltink, H., and Holtslag, A. A. M.: A Climatology of Nocturnal Low-Level Jets at Cabauw, J. Appl. Meteorol. Clim., 48, 1627–1642, https://doi.org/10.1175/2009JAMC1965.1, 2009.

Banta, R. M., Newsom, R. K., Lundquist, J. K., Pichugina, Y. L., Coulter, R. L., and Mahrt, L.: Nocturnal Low-Level Jet Characteristics Over Kansas During Cases-99, Bound-lay. Meteorol., 105, 221–252, https://doi.org/10.1023/A:1019992330866, 2002.

Blackadar, A. K.: Boundary Layer Wind Maxima and Their Significance for the Growth of Nocturnal Inversions, B. Am. Meteorol. Soc., 38, 283–290, https://doi.org/10.1175/1520-0477-38.5.283, 1957.

Bonner, W. D.: Climatology of the low level jet, Mon. Weather Rev., 96, 833–850, https://doi.org/10.1175/1520-0493(1968)096<0833:COTLLJ>2.0.CO;2, 1968.

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.

Cheinet, S., Beljaars, A., Köhler, M., Morcrette, J.-J., and Viterbo, P.: Assessing physical processes in the ECMWF model forecasts using ARM SGP observations. ECMWF-ARM Report Series., 2005.

Dörenkämper, M., Olsen, B. T., Witha, B., Hahmann, A. N., Davis, N. N., Barcons, J., Ezber, Y., García-Bustamante, E., González-Rouco, J. F., Navarro, J., Sastre-Marugán, M., Sīle, T., Trei, W., Žagar, M., Badger, J., Gottschall, J., Sanz Rodrigo, J., and Mann, J.: The Making of the New European Wind Atlas – Part 2: Production and evaluation, Geosci. Model Dev., 13, 5079–5102, https://doi.org/10.5194/gmd-13-5079-2020, 2020.

ECMWF: ERA5 Data Documentation. Available online: https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation (accessed on 31 March 2021).

Fisher, E. L.: An observational study of the sea breeze, J. Meteorol., 17, 645–660, https://doi.org/10.1175/1520-0469(1960)017<0645:AOSOTS>2.0.CO;2, 1960.

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.: Energy 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.

Gutierrez, W., Araya, G., Kiliyanpilakkil, P., Ruiz-Columbie, A., Tutkun, M., and Castillo, L.: Structural impact assessment of low level jets over wind turbines, J. Renew. Sustain. Energy, 8, 023 308, https://doi.org/10.1063/1.4945359, 2016.

Gutierrez, W., Ruiz-Columbie, A., Tutkun, M., and Castillo, L.: Impacts of the low-level jet's negative wind shear on the wind turbine, Wind Energy Sci., 2, 533–545, https://doi.org/10.5194/wes-2-533-2017, 2017.

Hahmann, A. N., Sīle, T., Witha, B., Davis, N. N., Dörenkämper, M., Ezber, Y., García-Bustamante, E., González-Rouco, J. F., Navarro, J., Olsen, B. T., and Söderberg, S.: The making of the New European Wind Atlas – Part 1: Model sensitivity, Geosci. Model Dev., 13, 5053–5078, https://doi.org/10.5194/gmd-13-5053-2020, 2020.

Hallgren, C., Arnqvist, J., Ivanell, S., Körnich, H., Vakkari, V., and Sahlée, E.: Looking for an Offshore Low-Level Jet Champion among Recent Reanalyses: A Tight Race over the Baltic Sea, Energies, 13, 3670, https://doi.org/10.3390/en13143670, 2020.

Held, D. P.: Inflow Measurements by Nacelle Mounted Lidars for Wind Turbine and Farm Control, Ph.D. thesis, DTU Wind Energy, 2019.

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, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hogan, R. J., O'Connor, E. J., and Illingworth, A. J.: Verification of cloud-fraction forecasts, Q. J. Roy. Meteor. Soc., 135, 1494–1511, https://doi.org/10.1002/qj.481, 2009.

Högström, U. and Smedman-Högström, A.-S.: The wind regime in coastal areas with special reference to results obtained from the Swedish wind energy program, Boundary-Layer Meteorology, 30, 351–373, https://doi.org/10.1007/BF00121961, 1984.

Holton, J. R.: The diurnal boundary layer wind oscillation above sloping terrain, Tellus, 19, 199–205, https://doi.org/10.1111/j.2153-3490.1967.tb01473.x, 1967.

Holtslag, A. A. M., Svensson, G., Baas, P., Basu, S., Beare, B., Beljaars, A. C. M., Bosveld, F. C., Cuxart, J., Lindvall, J., Steeneveld, G. J., Tjernström, M., and van de Wiel, B. J. H.: Stable Atmospheric Boundary Layers and Diurnal Cycles: Challenges for Weather and Climate Models, B. Am. Meteorol. Soc., 94, 1691–1706, https://doi.org/10.1175/BAMS-D-11-00187.1, 2013.

International Energy Agency: Offshore Wind Outlook 2019, 2019.

International Renewable Energy Agency: Renewable Energy Capacity Statistics 2022, 2022.

Kalverla, P. C., Duncan Jr., J. B., Steeneveld, G.-J., and Holtslag, A. A. M.: Low-level jets over the North Sea based on ERA5 and observations: together they do better, Wind Energy Sci., 4, 193–209, https://doi.org/10.5194/wes-4-193-2019, 2019.

Kalverla, P. C., Holtslag, A. A. M., Ronda, R. J., and Steeneveld, G.-J.: Quality of wind characteristics in recent wind atlases over the North Sea, Q. J. Roy. Meteor. Soc., 146, 1498–1515, https://doi.org/10.1002/qj.3748, 2020.

Kindler, D., Oldroyd, A., MacAskill, A., and Finch, D.: An eight month test campaign of the Qinetiq ZephIR system: Preliminary results, Meteorol. Z., 16, 479–489, https://doi.org/10.1127/0941-2948/2007/0226, 2007.

Liu, S. and Liang, X.-Z.: Observed Diurnal Cycle Climatology of Planetary Boundary Layer Height, J. Climate, 23, 5790–5809, https://doi.org/10.1175/2010JCLI3552.1, 2010.

Mann, J., Peña, A., Bingöl, F., Wagner, R., and Courtney, M. S.: Lidar Scanning of Momentum Flux in and above the Atmospheric Surface Layer, J. Atmos. Ocean. Tech., 27, 959–976, https://doi.org/10.1175/2010JTECHA1389.1, 2010.

Murphy, A. H.: Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient, Mon. Weather Rev., 116, 2417–2424, https://doi.org/10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2, 1988.

Nunalee, C. G. and Basu, S.: Mesoscale modeling of coastal low-level jets: implications for offshore wind resource estimation, Wind Energy, 17, 1199–1216, https://doi.org/10.1002/we.1628, 2014.

Parish, T. R., Rodi, A. R., and Clark, R. D.: A Case Study of the Summertime Great Plains Low Level Jet, Mon. Weather Rev., 116, 94–105, https://doi.org/10.1175/1520-0493(1988)116<0094:ACSOTS>2.0.CO;2, 1988.

Peña, A., Hasager, C. B., Badger, M., BArthelmie, R. J., Bingöl, F., Cariou, J.-P., Emeis, S., Frandsen, S. T., Harris, M., and Karagali, I.: Remote Sensing for Wind Energy, DTU Wind Energy Report, 2015.

Pichugina, Y. L., Banta, R. M., Kelley, N. D., Sandberg, S. P., Machol, J. L., and Brewer, W. A.: Nocturnal low-level jet characteristics over Southern Colorado, 16th Symposium on Boundary Layers and Turbulence, Portland ME, United States, 2004.

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.

Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T., and Balsamo, G.: Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models?, Journal of Advances in Modeling Earth Systems, 5, 117–133, https://doi.org/10.1002/jame.20013, 2013.

Sathe, A., Mann, J., Barlas, T., Bierbooms, W., and van Bussel, G.: Influence of atmospheric stability on wind turbine loads, Wind Energy, 16, 1013–1032, https://doi.org/10.1002/we.1528, 2013.

Savijärvi, H., Niemelä, S., and Tisler, P.: Coastal winds and low-level jets: Simulations for sea gulfs, Q. J. Roy. Meteor. Soc., 131, 625–637, https://doi.org/10.1256/qj.03.177, 2005.

Schulz-Stellenfleth, J., Emeis, S., Dörenkämper, M., Bange, J., Cañadillas, B., Neumann, T., Schneemann, J., Weber, I., zum Berge, K., Platis, A., Djath, B., Gottschall, J., Vollmer, L., Rausch, T., Barekzai, M., Hammel, J., Steinfeld, G., and Lampert, A.: Coastal impacts on offshore wind farms – a review focussing on the German Bight area, Meteorologische Zeitschrift, https://doi.org/10.1127/metz/2022/1109, 2022.

Sempreviva, A. M., Barthelmie, R. J., and Pryor, S. C.: Review of Methodologies for Offshore Wind Resource Assessment in European Seas, Surv. Geophys., 29, 471–497, https://doi.org/10.1007/s10712-008-9050-2, 2008.

Smedman, A.-S., Tjernstrm, M., and Hgstrm, U.: Analysis of the turbulence structure of a marine low-level jet, Boundary-Layer Meteorology, 66, 105–126, https://doi.org/10.1007/BF00705462, 1993.

Smedman, A.-S., Högström, U., and Bergström, H.: Low Level Jets - A Decisive Factor for Offshore Wind Energy Siting in the Baltic Sea, Wind Engineering, 29, 1996.

Stensrud, D. J.: Importance of Low-Level Jets to Climate: A Review, J. Climate, 9, 1698–1711, https://doi.org/10.1175/1520-0442(1996)009<1698:IOLLJT>2.0.CO;2, 1996.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, vol. 13 of *Atmospheric Sciences Library*, Springer Netherlands, Dordrecht, 1988.

Svensson, N., Bergström, H., Rutgersson, A., and Sahlée, E.: Offshore wind speed profiles: Investigating variations using WRF simulations and LiDAR measurements, in: Mesoscale Processes over the Baltic Sea, Department for Earth Sciences, Uppsala University: Uppsala, Sweden, 2018.

Svensson, N., Arnqvist, J., Bergström, H., Rutgersson, A., and Sahlée, E.: Measurements and Modelling of Offshore Wind Profiles in a Semi-Enclosed Sea, Atmosphere, 10, 194, https://doi.org/10.3390/atmos10040194, 2019a.

Svensson, N., Bergström, H., Rutgersson, A., and Sahlée, E.: Modification of the Baltic Sea wind field by land–sea interaction, Wind Energy, 22, 764–779, https://doi.org/10.1002/we.2320, 2019b.

Tuononen, M., Sinclair, V. A., and Vihma, T.: A climatology of low-level jets in the mid-latitudes and polar regions of the Northern Hemisphere, Atmos. Sci. Lett., 16, 492–499, https://doi.org/10.1002/asl.587, 2015.

Tuononen, M., O'Connor, E. J., Sinclair, V. A., and Vakkari, V.: Low-Level Jets over Utö, Finland, Based on Doppler Lidar Observations, J. Appl. Meteorol. Clim., 56, 2577–2594, https://doi.org/10.1175/JAMC-D-16-0411.1, 2017.

van de Wiel, B. J. H., Moene, A. F., Steeneveld, G. J., Baas, P., Bosveld, F. C., and Holtslag, A. A. M.: A Conceptual View on Inertial Oscillations and Nocturnal Low-Level Jets, J. Atmos. Sci., 67, 2679–2689, https://doi.org/10.1175/2010JAS3289.1, 2010.

Wagner, D., Steinfeld, G., Witha, B., Wurps, H., and Reuder, J.: Low Level Jets over the Southern North Sea, Meteorol. Z., 28, 389–415, https://doi.org/10.1127/metz/2019/0948, 2019.

Warner, T. T.: Numerical Weather and Climate Prediction, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511763243, 2010.

WindEurope: Wind energy in Europe: 2021 Statistics and the outlook for 2022-2026, 2022.

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, Cork, Ireland, https://doi.org/10.5281/zenodo.3372693, 2019.

Wolken-Möhlmann, G., Gottschall, J., and Lange, B.: First Verification Test and Wake Measurement Results Using a SHIP-LIDAR System, Energy Procedia, 53, 146–155, https://doi.org/10.1016/j.egypro.2014.07.223, 2014.

Zhai, X., Wu, S., Liu, B., Song, X., and Yin, J.: Shipborne Wind Measurement and Motion-induced Error Correction of a Coherent Doppler Lidar over the Yellow Sea in 2014, Atmos. Meas. Tech., 11, 1313–1331, https://doi.org/10.5194/amt-11-1313-2018, 2018.