

VALIDATION OF REANALYSIS AND MESOSCALE WIND DATASETS AGAINST DOPPLER LIDAR MEASUREMENTS AT HUB HEIGHT IN THE EASTERN BALTIC REGION

N. Dmitrijevs, M. Mironova, S. Orlova*

Laboratory "Energy Research Centre",
Institute of Electronics and Computer Sciences,
14 Dzerbenes Str., Riga, LV-1006, LATVIA
*e-mail: sorlova@edi.lv

This study assesses the performance of four reanalysis datasets, ERA5, NORA3, CERRA, and EMD WRF Europe+, by comparing them to Doppler LiDAR measurements taken at 192 m in Riga, Latvia. The analysis focused on wind speed, wind direction, Weibull distribution, turbulence intensity (TI), and its diurnal variation, using multiple statistical metrics. All datasets showed strong agreement with the LiDAR data, with correlation coefficients between 0.88 and 0.90. NORA3 performed best overall, while ERA5 had the lowest systematic and relative errors. Wind direction exhibited larger discrepancies, likely due to differences between short-term measurements and climatological reanalysis fields. Both CERRA and EMD underestimated turbulence intensity relative to LiDAR observations. Overall, the results indicate that reanalysis datasets can reliably estimate hub-height wind speed but require careful validation for turbulence intensity and site-specific wind resource evaluations.

Keywords: *Baltic Sea region, reanalysis datasets, turbulence intensity, WindCube LiDAR, wind energy, wind resource assessment.*

1. INTRODUCTION

Accurate wind-resource characterisation is essential for the economic viability and structural integrity of wind energy projects. Traditionally, wind speed has been the

main focus during site assessments, but turbulence intensity (TI), a key factor influencing energy variability and structural fatigue, remains underrepresented in standard eval-

uation methods [1]. This oversight matters because manufacturer power curves and Annual Energy Production (AEP) estimates are usually based on ideal atmospheric conditions that rarely match real-world inflow [1], [2]. While international standards such as IEC 61400-1 [3] provide turbulence-aware design frameworks, including Normal Turbulence Models (NTMs), their use is primarily limited to structural compliance rather than performance forecasting. As a result, the ongoing discrepancy between model predictions and actual turbine output continues to challenge the wind energy industry.

Recent research has significantly revalued wind farm site-selection criteria. While traditional methods focus on locations with high average wind speeds, the study [4] shows that such sites may be less economically viable when turbulence levels are high. Increased structural loads, accelerated component wear, and reduced operational efficiency ultimately negate the benefits of stronger wind resources. These insights emphasise the importance of treating turbulence intensity as a key factor in site assessment, not a secondary one. To address this, measurement methods must be capable of capturing turbulence throughout the entire rotor-swept area.

Study [5] indicates that for large turbines (≥ 5 MW), hub-height measurements alone are inadequate; ground-based LiDAR is necessary to accurately map the wind field and turbulence across the entire rotor plane. This underscores LiDAR's expanding role, not just for wind resource assessment, but also for understanding inflow conditions affecting turbine loading and performance.

Recent advances in atmospheric measurement technologies, such as LiDAR and SoDAR, enable detailed characterisation of wind flow, including wind speed, direction, and turbulence intensity (TI) [6]. In par-

ticular, Doppler wind LiDAR has become an increasingly established tool for wind resource assessment, offering high temporal resolution and the ability to resolve vertical wind profiles across the full rotor swept area without the logistical constraints of meteorological masts. In parallel, reanalysis datasets, such as the ECMWF Reanalysis v5 (ERA5), the Copernicus European Regional Reanalysis (CERRA), the Norwegian Reanalysis version 3 (NORA3), and EMD WRF Europe+, are widely used for wind resource assessment due to their extensive spatial and temporal coverage. However, their coarse spatial resolution and reliance on parameterised turbulence schemes raise questions about their ability to accurately represent local atmospheric variability, particularly turbulence intensity.

Previous studies have extensively validated reanalysis datasets against observations from LiDAR systems and meteorological masts, as well as through intercomparisons among different reanalysis products. In most cases, wind speed has been the primary evaluation metric due to its direct relevance to wind resource assessment and energy applications. However, a broader range of atmospheric variables has also been examined to assess the overall performance and reliability of reanalysis datasets under different climatic and terrain conditions. As summarised in Table 1, the reviewed literature employs a variety of validation parameters, including wind speed, wind direction, air temperature, atmospheric pressure, turbulence-related quantities, and derived statistical performance metrics. This diversity of evaluation parameters reflects the increasing emphasis on comprehensive validation approaches that consider multiple aspects of atmospheric conditions rather than relying solely on wind speed.

Table 1. Literature Analysis Based on Reanalysis Models and Parameters Studied

Reanalysis model	Study scope/metrics	Reference
ERA5, NORA3, NEWA	Wind speed profiles at 100-500 m validated against Doppler wind LiDAR. Including mean wind speed, bias, correlation, RMSE.	[7]
ERA5, CERRA, WRF	Mean and variance of wind speed, correlation, RMSE. Validated at 5 offshore LiDAR locations in the Dutch North Sea and the FINO3 platform.	[8]
NORA3	Wind speed profiles validated against LiDAR. Mean wind speed, wind rose, Weibull distribution, correlation, AEP estimates.	[9]
CFSv2, ERA5, MERRA2	Mean wind speed, wind power density, Weibull distribution, AEP, and LCOE using the IEA-15 MW turbine.	[10]
ERA5 (both atmospheric and wave data)	Turbulence intensity, TI expressed as a function of wind speed (U), height (z), wave phase velocity (cp), and atmospheric stability (z/L, Monin-Obukhov length) at 10 m. Compares directly against the IEC 61400-1 (2019) TI formulation as well as ISO and other empirical TI models, validated against measured TI at named sites: Høvsøre, Horns Rev, FINO1/2/3	[11]
NORA3	Fatigue loading study with veer gradients, turbulence intensity, shear profiles, wave states as inputs. Consequently, damage equivalent loads (DELs) on the tower, blades, and main bearings as output.	[12]
EMD-WRF, ERA5	Wind profile parameters such as shear, mean speed, and direction. Based on vertical wind shear, Monin-Obukhov length, and Richardson number as indicators, mechanical and thermal atmospheric stability was determined. All data verified with onsite measurements.	[13]
ERA5, NOW23	Wind shear and veer across the rotor layer under different stability regimes (Obukhov length with temperature, relative humidity, and pressure), statistics. Wind speed droughts, frequency and duration of low-output events.	[14]

Therefore, this study aims to validate multiple reanalysis datasets against high-frequency LiDAR measurements using an extended parameter set, with particular emphasis on turbulence intensity, to provide a more comprehensive assessment of their suitability for wind energy applications in the Baltic Sea region.

This study addresses three specific research questions: How accurately do

NORA3, ERA5, CERRA, and EMD WRF Europe+ reproduce hub-height wind speed distributions observed by LiDAR in a Latvian environment? Do reanalysis-derived turbulence intensity estimates agree with high-frequency LiDAR measurements, and what are the primary sources of discrepancy? Which dataset provides the most reliable representation of wind conditions for wind energy applications in the Eastern Baltic Sea region?

2. DATA AND METHODOLOGY

2.1. Reanalysis Datasets

Reanalysis datasets provide physically consistent reconstructions of atmospheric conditions by assimilating observational

data into numerical weather prediction models. In this study, multiple reanalysis products with varying spatial and temporal

resolutions are used to assess their performance against LiDAR observations. The selected reanalysis datasets are evaluated for their suitability in wind resource assessment in the Baltic Sea region, with a particular focus on the Latvian environment.

The **ERA5** dataset is the fifth-generation atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The system utilises the Integrated Forecasting System (IFS) Cycle 41r2, incorporating a 12-hourly 4D-Var data assimilation scheme that integrates billions of historical observations into a physically consistent framework. ERA5 offers a horizontal resolution of 31 km, an increased vertical resolution of 137 levels, and hourly temporal outputs for all parameters [15].

The **NORA3** dataset provides high-resolution records of weather patterns for Northern Europe, using a 3 km horizontal grid and 65 vertical levels based on a terrain-following pressure system that provides data from 10 m to 750 m. It is based on the non-hydrostatic, convection-permitting (CP) HARMONIE-AROME model, which solves the fully compressible Euler equations [16].

The **CERRA** dataset is designed specifically for the European domain. With a horizontal resolution of 5.5 km it captures small-scale meteorological phenomena and complex topographical influences. The system is built on the ALADIN-HARMONIE

numerical weather prediction model and utilises a 3-hourly data assimilation cycle [17].

EMD WRF Europe+ model is specifically optimised for wind energy and atmospheric research across the European domain. By downscaling ERA5 reanalysis boundary conditions using the Weather Research and Forecasting (WRF) model, it achieves a high spatial resolution of 3 km with hourly temporal outputs. This modelling framework provides a consistent vertical profile of the atmosphere, offering data across 127 parameters and multiple heights up to 4,000 meters [18].

These four datasets represent a range of spatial resolutions and modelling approaches, from global reanalysis (**ERA5**) to high-resolution regional downscaling (**NORA3**, **CERRA**, **EMD WRF Europe+**). All datasets were retrieved from the WindPro software (version 4.1.292) METEO object for the coordinates specified in Table 1 and processed according to the methodology described in the following section. As the reanalysis datasets have discrete spatial resolution, we extracted data from the nearest grid points to the LiDAR measurement site, resulting in small spatial mismatches between the observation location and the model grid points. Table 2 provides approximate mismatch distances. The ERA5 model, with its coarser grid, exhibits the largest mismatch, while the EMD-WRF model has the closest available grid point.

Table 2. Geographic Coordinates for Reanalysis Data Extraction

Reanalysis model	Longitude	Latitude	Distance to LiDAR
NORA3	24.192	56.980	~0.345 km
ERA5	24.000	56.908	~14.17 km
CERRA	24.228	56.931	~5.545 km
EMD WRF Europe+	24.200	56.978	~0.301 km

To integrate the differing datasets, processing was required to align varying struc-

tures and temporal resolutions. The study period was restricted to the period from 18

September 2025 to 1 December 2025, using a synchronised 1-hour interval. The study period of approximately 10.5 weeks covers only the autumn–early winter season, which limits the climatological representativeness of the results. Therefore, the findings were interpreted in the context of a measurement campaign rather than a long-term wind resource assessment.

2.2. LiDAR Observational Data

Observation datasets used for analysis and comparison were obtained from a Vaisala WindCube LiDAR deployed in Riga (56°58' 38.71"N 24°11'42.60"E) in the urban environment. Vertical profiles were measured at 18 levels from 52 to 250 m AGL. Below 200 m, the vertical resolution alternates between 8 m and 12 m. Above 200 m, the spacing increases to 20–30 m. The 192 m level was selected for the primary comparison because it represented the upper part of the measurement range and was comparable to the hub heights of modern large wind turbines.

LiDAR data were collected from 18 September 2025 to 1 December 2025. Data were extracted at 192 m AGL, the measurement level closest to the typical hub heights of modern multi-megawatt wind turbines.

2.3. Data Processing and Analytical Methods

Model performance was evaluated using the Pearson correlation coefficient (r), coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), mean bias error (Bias), mean absolute percentage error (MAPE), Willmott's index of agreement (d-index), and the standard deviation ratio (SD Ratio). These metrics were used to quantify the agreement between LiDAR observations and the reanalysis datasets with respect to correlation, prediction error, systematic bias, and variability.

For vertical consistency, all data were extracted at a common height of 192 meters. While primary parameters included wind speed, direction, temperature, and turbulence, the latter two were not available across all sources. The comparative analysis for temperature and turbulence was limited to the CERRA and EMD datasets.

Wind measurements were recorded at a temporal resolution of approximately 1 minute; however, to match the reanalysis model timestamps, data were aggregated to 1-hour intervals. Subsequently, the reanalysis and observational datasets were compared at corresponding time points, and records with missing data were removed. This control procedure yielded 1720 complete data points for further analysis. All datasets were temporally synchronized to a common hourly resolution prior to statistical analysis.

Quality control procedures were applied to the observational data. The LiDAR device performs automated self-assessment of measurements based on carrier-to-noise ratio (CNR) metrics, which are calibrated and validated by the manufacturer.

Scatter plots were used to evaluate the relationship between measured and modelled wind speeds at 192 m AGL. Wind roses were generated to compare the directional distribution of wind, while Weibull probability density functions were used to assess the statistical distribution of wind speeds. Because reanalysis products have inherent temporal-resolution constraints, this study deviates from standard recommendations and uses a 1-hour averaging period to validate reanalysis-derived TI values.

Wind resource assessment commonly follows IEC 61400-1:2019, which recommends a 10-minute averaging period for estimating turbulence intensity. Because the temporal resolution of the reanalysis products differs from this recommendation, a common 1-hour averaging period was adopted for all datasets to ensure consistent comparison. A key element of IEC 61400-1:2019 is its classification of wind turbines into classes and categories based on turbulence characteristics to ensure safe operation. These categories are defined by the reference value $I(\text{ref})$, which represents the characteristic longitudinal turbulence intensity at a hub-height wind speed of 15 m/s. The standard describes turbulence intensity using the Normal Turbulence Model (NTM) [2].

To assess the impact of the adopted 1-hour averaging period on turbulence intensity estimates, both 10-minute and 1-hour TI values were computed from the LiDAR observations. The mean TI derived from 1-hour averages was found to be approximately 19 % lower than the corresponding 10-minute values, consistent with the expected reduction in wind speed standard deviation at longer averaging intervals. This systematic difference should be considered when comparing LiDAR-derived TI with reanalysis products, as it implies that the true discrepancy between reanalysis and IEC-compliant TI estimates may be larger than reported in this study.

3. RESULTS AND DISCUSSION

To characterise the vertical structure of the wind field, the mean diurnal variation of wind speed was analysed at all available LiDAR measurement heights between 52

Turbulence intensity (TI) values were obtained directly from both the LiDAR observations and the reanalysis datasets. In the wind energy sector, TI is calculated from raw data by Eq. (1):

$$TI = \frac{\sigma}{\bar{U}}. \quad (1)$$

This ratio of the standard deviation of wind speed (σ) to the mean wind speed ($\bar{U}$) over a specified time interval provides insight into wind variability and atmospheric stability conditions [19].

Wind speed is commonly calculated by its longitudinal, lateral, and vertical wind components:

$$U = \sqrt{u^2 + v^2 + w^2}. \quad (2)$$

Turbulence kinetic energy (TKE) is an overall measure of atmospheric turbulence [20]. It is directly related to the transport of momentum, heat, and moisture through the boundary layer. It is expressed by the equation:

$$TKE = 0.5 * [\overline{(u')^2} + \overline{(v')^2} + \overline{(w')^2}]. \quad (3)$$

TKE represents the absolute energy of turbulent motion, whereas TI describes the relative magnitude of wind-speed fluctuations normalised by the mean wind speed. Assuming isotropic turbulence, TI can be related to TKE by Eq. (4).

$$TI = \sqrt{\frac{2 * TKE}{3}} / \bar{U}. \quad (4)$$

and 192 m AGL (Fig. 1). This analysis provides the basis for understanding the evolution of the vertical wind profile and the associated wind shear throughout the day.

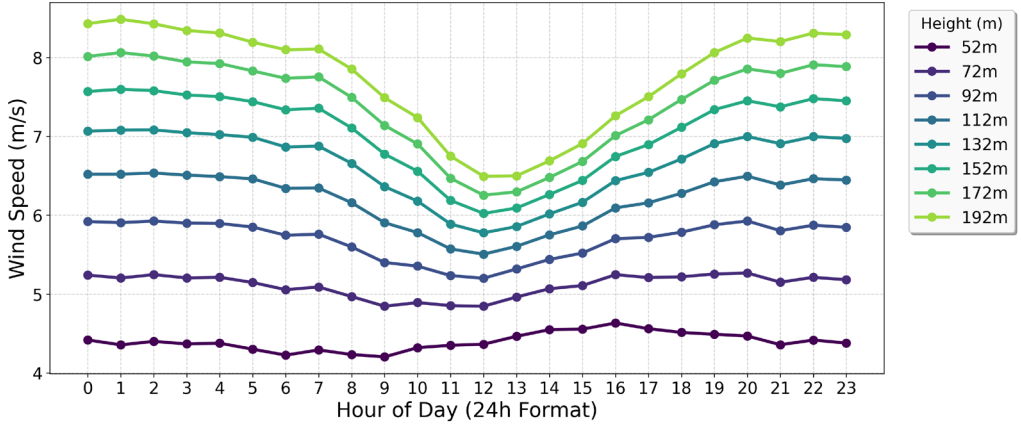


Fig. 1. Mean diurnal variation of wind speed at eight measurement heights (52–192 m AGL).

Although all levels show a similar temporal evolution, the diurnal cycle magnitude increases markedly with height. Wind speeds decrease from early morning and reach a minimum around midday, then gradually recover during the afternoon and evening. The largest daily variation occurs at **192 m**, while the weakest variation is observed at **52 m**, suggesting that the upper part of the rotor-swept layer is more strongly affected by atmospheric boundary-layer dynamics than the lower levels.

The consistent increase in wind speed with height throughout the day confirms the persistence of vertical wind shear across the measurement period, although the magnitude of the shear varies with the diurnal evolution of the atmospheric boundary layer. To further investigate the vertical structure of the wind field, the diurnal variation of the wind shear exponent was evaluated using LiDAR measurements and the corresponding reanalysis datasets (Fig. 2).

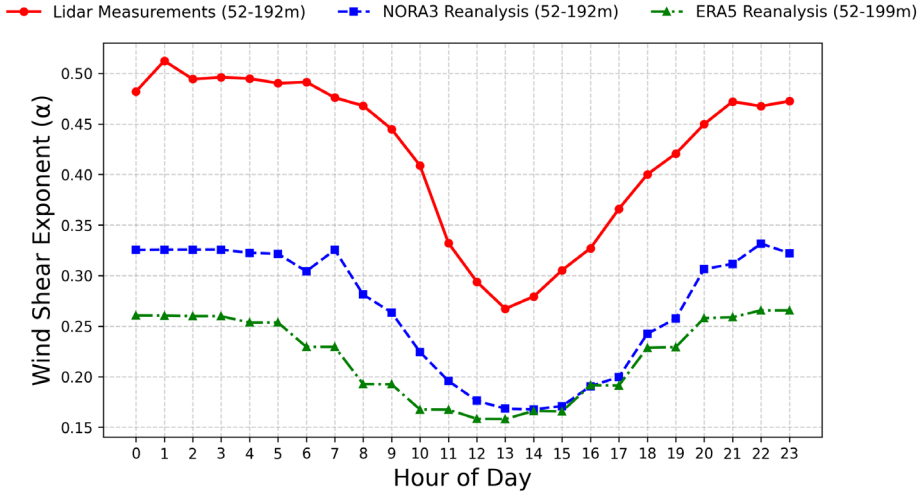


Fig. 2. Diurnal variation of the wind shear exponent (52–192 m AGL) derived from LiDAR observations and reanalysis datasets.

Figure 2 shows the diurnal variation of the wind shear exponent (α) between 52 and 192 m AGL, derived from LiDAR measurements and the NORA3 and ERA5 reanalysis datasets. All datasets follow a similar daily pattern, with peak wind shear at night and early morning, and lowest values around midday. This pattern reflects boundary-layer development: increased daytime turbulence reduces the vertical wind-speed gradient, while stable nighttime conditions increase wind shear. Although both reanalysis datasets match the overall temporal trend, they consistently underestimate the wind shear exponent compared to LiDAR data throughout the day. The greatest gap occurs at night, with LiDAR values around 0.50, while NORA3 and ERA5 stay below 0.33 and 0.27. During the day, dif-

ferences lessen but remain, showing that reanalysis models underestimate the wind shear gradient. NORA3 aligns more closely with LiDAR observations than ERA5, reliably capturing higher wind shear values across the diurnal cycle. Wind shear exponent values approaching 0.50 indicate a strong increase in wind speed with height, reflecting a pronounced vertical wind-speed gradient under stable nocturnal atmospheric conditions.

Wind speed agreement between LiDAR observations and reanalysis datasets was assessed through Pearson correlation analysis at 192 m AGL. Figure 3 presents the resulting scatter plots, with each panel comparing a single reanalysis product against LiDAR observations and reporting the Pearson correlation coefficient (r).

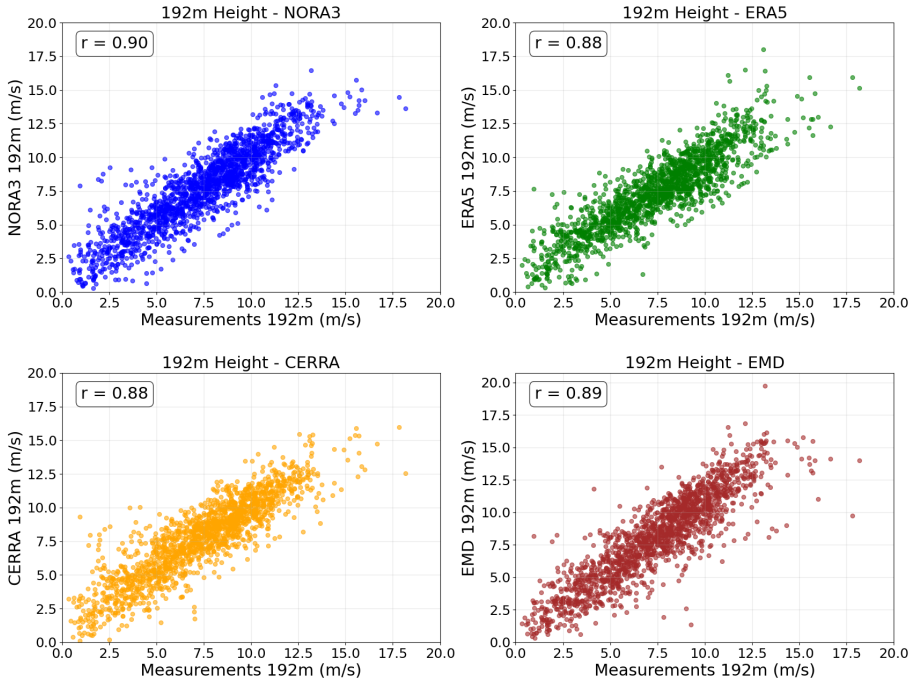


Fig. 3. Correlation of wind speeds at 192 m between LiDAR measurements and reanalysis models.

Overall, the scatter plots demonstrate good agreement between LiDAR obser-

vations and all reanalysis datasets, with NORA3 exhibiting the tightest clustering

around the regression trend.

At low wind speeds (<4 m/s), all datasets tend to slightly overestimate the observations. As wind speed rises, the data points scatter more, reflecting increased uncer-

tainty in modelling stronger winds. After about 17 m/s, the number of observations decreases, and the scatter widens, showing reduced model accuracy during extreme wind conditions.

Table 3. Statistical Error Metrics for Reanalysis Wind Speed Validation against LiDAR at 192 m AGL

Models	r	R ²	MAE	RMSE	Bias	MAPE	d-index	SD Ratio
NORA3	0.90	0.78	1.08	1.42	+0.31	21.5%	0.94	1.00
ERA5	0.88	0.78	1.09	1.42	+0.01	20.8%	0.94	0.93
CERRA	0.88	0.76	1.12	1.49	+0.40	23.4%	0.93	0.95
EMD	0.89	0.70	1.30	1.67	+0.67	22.8%	0.93	1.00

Overall, all datasets performed well, with r values from 0.88 to 0.90 and d-index scores above 0.93, indicating strong agreement with measurements. Considering all statistical indicators, NORA3 provides the best overall agreement with the LiDAR measurements, achieving the highest correlation coefficient ($r = 0.90$), the lowest MAE (1.08 m/s), and an SD Ratio of 1.00. ERA5 performs similarly in terms of RMSE while

exhibiting the lowest Bias (+0.01 m/s) and MAPE (20.8 %), indicating the smallest systematic and relative errors. CERRA shows slightly lower overall accuracy, whereas EMD reproduces the observed variability (SD Ratio = 1.00) but yields the largest prediction errors, as reflected by its highest RMSE (1.67 m/s) and Bias (+0.67 m/s).



Fig. 4. Wind rose comparison between LiDAR measurements and reanalysis models at 192 m.

Wind direction exhibits larger discrepancies than wind speed among the evaluated datasets. Long-term meteorological observations in Latvia identify south-west (SW) winds as the prevailing regional wind direction [21]. In contrast, the LiDAR measurements and the reanalysis datasets reveal notable differences in both directional frequency and overall wind distribution patterns.

The short-term LiDAR observations, covering approximately 10.5 weeks, indicate that south-south-east (SSE) is the most frequently occurring wind direction during the measurement campaign. However, all reanalysis products identify either south (S) or south-southwest (SSW) as the dominant

direction, although with varying occurrence frequencies. This discrepancy likely reflects the fact that reanalysis models are optimised to reproduce broader climatological and large-scale atmospheric tendencies rather than localised short-term or seasonal variability.

These findings suggest applying reanalysis datasets with caution when assessing short-duration wind regimes or site-specific directional characteristics. Nevertheless, among the evaluated products, ERA5 and CERRA demonstrate the closest agreement with the LiDAR observations, showing the most comparable directional patterns and the strongest overall similarity to the measured wind rose.

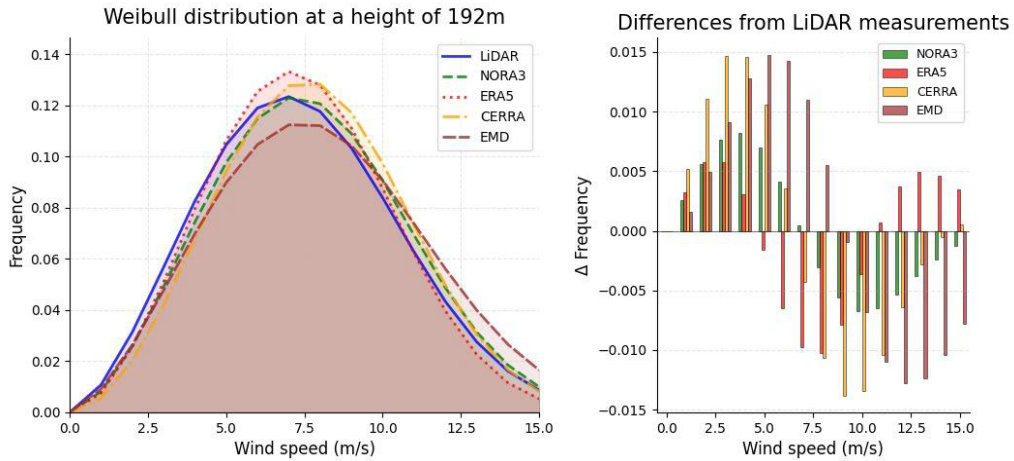


Fig. 5. Weibull distribution at 192 m for LiDAR measurements and reanalysis models.

The Weibull distribution analysis indicates that all reanalysis datasets reproduce the general wind-speed characteristics observed by LiDAR at 192 m (see Fig. 5). The peak frequency for all datasets occurs around 6.5–8.0 m/s, suggesting this is the most common wind speed range at 192 m. However, systematic differences remain, particularly in the low- and high-wind-speed tails. NORA3 shows the closest

overall correspondence with the measured distribution, while ERA5 and CERRA accurately capture the dominant mid-range wind speeds. EMD exhibits the broadest distribution and the largest deviations at elevated wind speeds.

These findings indicate that although reanalysis products are suitable for representing long-term wind climatology, dataset-specific biases should be considered

when assessing wind resource potential or turbine performance.

For the following comparison of TI, for reanalysis models pre-calculated characteristics from windPRO software were used. To calculate LiDAR, Eq. (1) was used.

LiDAR measurements recorded at 1-minute resolution were aggregated to 1-hour intervals prior to comparison. Since NORA3 and ERA5 provide turbulence intensity parameter once in two hours, both models were excluded from further analysis.

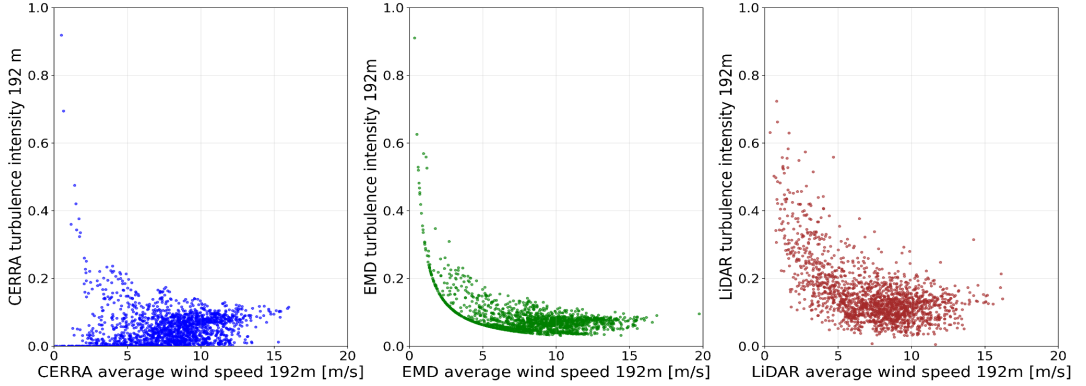


Fig. 6. Turbulence intensity as a function of mean wind speed: A comparison of LiDAR measurements and reanalysis models.

As shown in Fig. 6, the LiDAR measurements indicate average turbulence intensities generally ranging from 5 % to 15 % for wind speeds between 5 and 10 m/s, with occasional higher values of 15–20 %. This suggests moderate turbulence conditions within the wind speed range most relevant for turbine operation.

In contrast, the CERRA and EMD datasets exhibit systematically lower turbulence estimates, typically within the 5 % to 10 %

range at mean wind speeds of approximately 7 to 12 m/s. In particular, CERRA shows a pronounced concentration of values near 0 % turbulence intensity at wind speeds between 5 and 10 m/s, which is inconsistent with the variability observed in the LiDAR measurements. As shown in Table 4, both CERRA and EMD exhibit negative biases in turbulence intensity, confirming the systematic underestimation visible in Fig. 6.

Table 4. Statistical Error Metrics for Turbulence Intensity Validation against LiDAR

Models	MAE	RMSE	Bias
CERRA	0.099	0.132	−0.097
EMD	0.072	0.098	−0.067

The lower turbulence levels reported by EMD and CERRA than by LiDAR observations suggest a systematic underestimation, which may relate to the model’s parameterisation approach but requires further inves-

tigation. This may be associated with the turbulence parameterisation employed by the reanalysis models [22]. However, further investigation is required to determine the exact causes of these discrepancies.

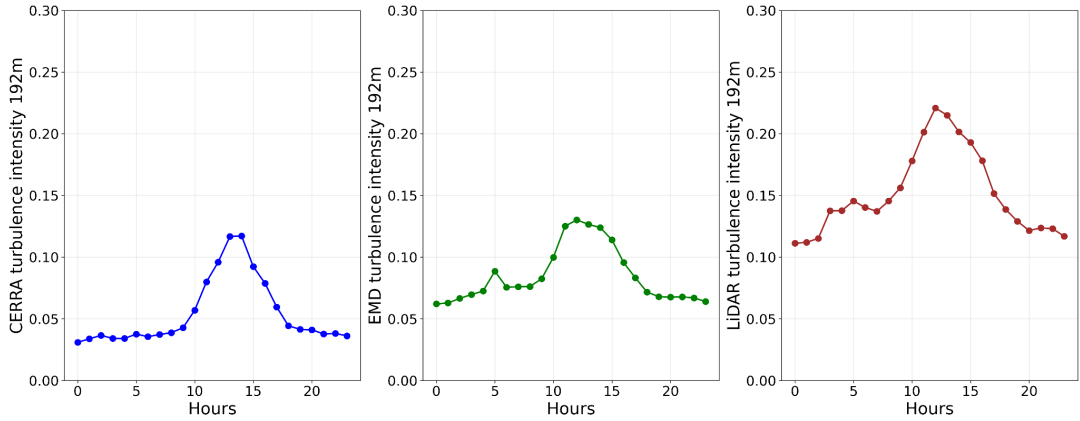


Fig. 7. Turbulence intensity diurnal cycle compared between LiDAR measurements and reanalysis models.

Figure 7 illustrates the diurnal variation of turbulence intensity, showing how wind fluctuations change throughout the day. Both reanalysis models and LiDAR observations exhibit a similar pattern, with turbulence increasing during daytime hours, rising around 09:00–10:00, and then decreasing after approximately 15:00 until stabilising between 17:00 and 20:00.

The LiDAR data indicate consistently higher TI values, with daytime increases of approximately 10 percentage points. Among the reanalysis products, CERRA shows the lowest overall TI values but reproduces a similar daytime increase of about 10 percentage points, whereas EMD displays the smallest daytime rise, around 6–7 percentage points.

These diurnal variations are closely linked to solar heating of the Earth’s surface and the resulting atmospheric boundary layer (ABL) dynamics.

The observed positive TI bias in LiDAR measurements relative to all reanalysis products is partly attributable to the instrument’s urban deployment, as none of the reanalysis datasets explicitly resolve urban surface roughness at the measurement site scale.

Overall, the results demonstrate that high-resolution regional reanalysis datasets reproduce hub-height wind speed reasonably well, whereas larger discrepancies remain for wind direction, wind shear and turbulence intensity.

4. CONCLUSIONS

This study evaluated the performance of four reanalysis datasets (NORA3, ERA5, CERRA, and EMD WRF Europe+) against WindCube LiDAR measurements at 192 m AGL in an urban environment. The comparison included wind speed, wind direction, wind-speed distribution, turbulence

intensity, and the diurnal evolution of atmospheric conditions.

The results demonstrate that all datasets reproduce the general wind-speed characteristics with strong correlations ($r = 0.88$ – 0.90) and high agreement ($d\text{-index} \geq 0.93$). Among the evaluated products, NORA3

achieved the best overall performance, with the highest correlation and lowest prediction error, while ERA5 showed the smallest systematic bias. Both datasets showed better agreement with the LiDAR observations than CERRA and EMD.

The analysis further revealed that reanalysis datasets reproduced the overall diurnal evolution of wind speed and wind shear but systematically underestimated the vertical wind-speed gradient and turbulence intensity, particularly under stable nocturnal conditions. The largest discrepancies were observed for wind direction and turbulence intensity, highlighting the limitations of current reanalysis products in representing

local atmospheric variability.

Overall, the results indicate that high-resolution regional reanalysis datasets, particularly NORA3, provide a reliable basis for wind resource assessment at modern turbine hub heights. However, site-specific LiDAR measurements remain essential for accurate characterisation of wind direction, vertical wind structure, and turbulence intensity, especially in complex urban environments.

Future work will focus on extended validation using long-term datasets covering the full annual cycle, with additional emphasis on wind veer and atmospheric stability classification.

REFERENCES

- Kim, D. Y., Kim, Y. H., & Kim, B. S. (2021). Changes in wind turbine power characteristics and annual energy production due to atmospheric stability, turbulence intensity, and wind shear. *Energy*, 214, 119051. <https://doi.org/10.1016/j.energy.2020.119051>
- Teixeira, V. D. L., Hetem Junior, A., Ferreira, L., & Teixeira, J. C. (2021). Wind turbine power curve prediction using artificial neural networks: A SCADA-based approach. *IEEE Access*, 9, 118476–118491. <https://doi.org/10.1109/ACCESS.2021.3106510>
- International Electrotechnical Commission. (2019). *Wind energy generation systems – Part 1: Design requirements* (IEC Standard No. 61400-1:2019). <https://webstore.iec.ch/publication/26418>
- Asadi, M., & Pourhossein, K. (2021). Wind farm site selection considering turbulence intensity. *Energy*, 236, 121480. <https://doi.org/10.1016/j.energy.2021.121480>
- Shi, Y., Hu, F., Li, X., Zhang, Z., & Zhang, K. (2025). Improving large wind turbine power curve by integrating lidar-measured multiple wind parameters: A coastal case study. *Energies*, 18(24), 6398. <https://doi.org/10.3390/en18246398>
- Lang, S., & McKeogh, E. (2011). LIDAR and SODAR measurements of wind speed and direction in upland terrain for wind energy purposes. *Remote Sensing*, 3(9), 1871–1901. <https://doi.org/10.3390/rs3091871>
- Cheyne, E., Diezel, J. M., Haakenstad, H., Breivik, Ø., Peña, A., & Reuder, J. (2025). Tall wind profile validation of ERA5, NORA3, and NEWA datasets using lidar observations. *Wind Energy Science*, 10(4), 733–754. <https://doi.org/10.5194/wes-10-733-2025>
- Rouholahnejad, F., Meyer, P. J., & Gottschall, J. (2024). Collocating wind data: A case study on the verification of the CERRA dataset. *Journal of Physics: Conference Series*, 2875(1), 012016. <https://doi.org/10.1088/1742-6596/2875/1/012016>
- Komasilovs, V., Mironova, M., Dmitrijevs, N., Kamolins, E., & Orlova, S. (2026). NORA3 dataset comparison with observed onshore wind measurements in the Eastern Baltic Sea region. *Energies*, 19(5), 1144. <https://doi.org/10.3390/en19051144>

10. de Assis Tavares, L. F., Shadman, M., Assad, L. P. de F., & Estefen, S. F. (2022). Influence of the WRF model and atmospheric reanalysis on the offshore wind resource potential and cost estimation: A case study for Rio de Janeiro State. *Energy*, 240, 122767. <https://doi.org/10.1016/j.energy.2021.122767>
11. Larsén, X. G., Imberger, M., & Floors, R. (2026). Modelling global offshore turbulence intensity including large-scale turbulence, stability and sea state. *Wind Energy Science*, 11, 1853–1869. <https://doi.org/10.5194/wes-11-1853-2026>
12. Krathe, V. L., Jonkman, J., & Bachynski-Polić, E. E. (2025). The influence of wind veer and drivetrain flexibility on fatigue loading for large floating wind turbines. *Wind Energy Science*, 10, 2903–2923. <https://doi.org/10.5194/wes-10-2903-2025>
13. Ryu, G., Lee, D., Han, S., Kim, J., Kim, Y.-G., & Moon, C.-J. (2025). Comparison and Verification of EMD-WRF and ERA5 Meteorological Data with Offshore Measurement Data for Assessing Offshore Wind Resources in the Southwest Sea of Korea. *Journal of the Wind Engineering Institute of Korea*, 29(1), 3–14. <https://doi.org/10.37109/weik.2025.29.1.3>
14. Qiu, L., & Howland, M. F. (2025). Quantitative evaluation of reanalysis and NWP models for wind resource assessment in offshore environments. *Journal of Geophysical Research: Atmospheres*. <https://doi.org/10.1029/2025JD043490>
15. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
16. Haakenstad, H., Breivik, Ø., Furevik, B. R., Reistad, M., Bohlinger, P., & Aarnes, O. J. (2021). NORA3: A nonhydrostatic high-resolution hindcast of the North Sea, the Norwegian Sea, and the Barents Sea. *Journal of Applied Meteorology and Climatology*, 60(10), 1443–1464. <https://doi.org/10.1175/JAMC-D-21-0029.1>
17. Ridal, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Schimanke, S., Andrae, U., ... & Wang, Z. (2024). CERRA, the Copernicus European Regional Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 150(763), 3385–3411. <https://doi.org/10.1002/qj.4764>
18. EMD International A/S. (2019). *EMD-WRF Europe+ (ERA5)*. Wiki-WindPRO. [http://help.emd.dk/mediawiki/index.php/EMD-WRF_Europe%2B_\(ERA5\)](http://help.emd.dk/mediawiki/index.php/EMD-WRF_Europe%2B_(ERA5))
19. Ren, G., Liu, J., Wan, J., Li, F., Guo, Y., & Yu, D. (2018). The analysis of turbulence intensity based on wind speed data in onshore wind farms. *Renewable Energy*, 123, 756–766. <https://doi.org/10.1016/j.renene.2018.02.096>
20. Stull, R. B. (2017). *Practical meteorology: An algebra-based survey of atmospheric science* (Version 1.02b). University of British Columbia. https://www.coas.ubc.ca/books/Practical_Meteorology/
21. Pogumirskis, M., Sīle, T., Seņņikovs, J., & Bethers, U. (2021). PCA analysis of wind direction climate in the Baltic states. *Tellus A: Dynamic Meteorology and Oceanography*, 73(1), 1–16. <https://doi.org/10.1080/16000870.2021.1962490>
22. Arenas-López, J. P., & Badaoui, M. (2020). Stochastic modelling of wind speeds based on turbulence intensity. *Renewable Energy*, 155, 10–22. <https://doi.org/10.1016/j.renene.2020.03.104>