BOREAL ENVIRONMENT RESEARCH 30: 181–193
ISSN 1797-2469 (online)

Helsinki 24 October 2025

Wind speeds and gusts in Finland: a comparison of observations and ERA5 reanalysis

Terhi K. Laurila*, Mikko Laapas, Taru Olsson and Kirsti Jylhä

Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland
**corresponding author's e-mail: terhi.laurila@fmi.fi*

Received 02 Jul. 2025, final version received 17 Sep. 2025, accepted 23 Sep. 2025

Laurila T.K., Laapas M., Olsson T. & Jylhä K. 2025: Wind speeds and gusts in Finland: a comparison of observations and ERA5 reanalysis. *Boreal Env. Res.* 30: 181–193.

In this study, we compare hourly wind speed and wind gust observations from 143 weather stations in Finland to ERA5 reanalysis data from 2014–2023. The weather stations are classified into inland, coast, lake, and mountain stations based on their elevation and terrain type. Overall, the correlation between ERA5 and observed winds is strong. However, both the hourly wind speeds and wind gusts in ERA5 are underestimated except for the weakest winds, which are overestimated. The bias is larger in wind speeds than in wind gusts. Among the four station classes, the coast and inland stations have in general the best agreement between ERA5 and observed winds, while ERA5 performance in the mountain stations is relatively poor.

Introduction

Extreme winds can cause large impacts on society by causing e.g. forest damage, power outages, destruction of buildings and property and even losses of life (Láng *et al.* 2021, Jasiūnas *et al.* 2023, Romagnoli *et al.* 2023, Laurila *et al.* 2025, Virman *et al.* 2025). Accurate wind information is essential for a range of applications, including wind energy production and site assessment (e.g. Soares *et al.* 2020, Gualtieri 2021), transportation safety (e.g. Vajda *et al.* 2014, Taszarek *et al.* 2020), air quality modelling (e.g. Ottosen *et al.* 2019), emergency preparedness (e.g. Haakana *et al.* 2024) and structural design (e.g. Jafari & Alipour 2021). Wind data are also critical for evaluating coastal flooding risks due to sea level and waves (e.g. Leijala *et al.* 2018) and for assessing operational safety in facilities such as industrial plants, airports, and power generation sites, including nuclear power plants (Jylhä *et al.* 2018). In both urban and rural settings, strong winds can influence, for example, the spread of wildfires and airborne pollutants, underscoring their broad societal relevance. Winds are therefore an important topic of research needed in many sectors across society.

Windiness can be described by two variables: wind speed and wind gust. Wind speed is a 10-min average speed of wind flow measured ideally at 10-m height, and wind gust is a 3-second maximum at 10-m height. While the mean wind speed gives a good overlook of the wind climate and can represent large-scale systems, wind gust information is essential for describing extreme wind events. Therefore, both the wind speed and wind gust are investigated in this study.

Reanalysis is a gridded dataset that combines a weather model and observations. Reanalysis datasets are commonly used in climate assessments due to their long-time coverage, usually

Editor in charge of this article: Veli-Matti Kerminen
https://doi.org/10.60910/ber2025.q6kf-qp39

of multiple decades, and large spatial coverage (there are both global and regional reanalyses). In addition, reanalysis data is homogeneous on time while observations are influenced by many additional factors, such as changes in instruments and their locations, and environmental changes in the surroundings of the observation stations. While the benefits of a reanalysis are evident, there are also disadvantages. The grid size of a reanalysis is relatively large (e.g. around 31 km in ERA5 reanalysis used in this study), which means that the value in one grid box is a smoothed result, representing average conditions within the grid box, and may differ from the observed value. Especially regarding wind speed, the averaging of grid boxes may smooth the most extreme values. In addition, the turbulent nature of wind gusts is of a too small scale for reanalyses to be resolved directly. Therefore, wind gusts need to be parametrized, i.e., represented by other grid size variables that are resolvable.

Multiple studies comparing ERA5 near-surface i.e. 10-m wind speeds to different observational datasets have been made in recent years. Among others, Chen *et al.* (2024) studied ERA5 wind speeds associated with extratropical cyclones over central and eastern North America and compared them to the in-situ station data during 2005–2019. They found that ERA5 performs well in estimating wind speed, but it tends to overestimate low winds and underestimate high winds. The authors observed seasonal and regional variations with the best performance in winter and consistent underestimation over complex terrain like the Rockies. Jiang *et al.* (2021) evaluated ERA5 near-surface wind speeds using the ground automatic meteorological observation data from 3 April to 31 October 2020 over Hainan Island and the South China Sea. They also found that ERA5 generally overestimated low wind speeds and underestimated high wind speeds, and that the accuracy of ERA5 varied with terrain; it performed better over offshore islands than over the land areas of Hainan Island, with the poorest performance over mountainous regions. Regarding Europe, Molina *et al.* (2021) compared ERA5 near-surface wind speeds with wind observations including 245 stations across Europe during 1979–2018. Their finding was that ERA5 effectively captured the seasonal cycle of monthly wind speeds, but it again tended to overestimate the frequency of low wind speeds and underestimate higher wind speeds, particularly in complex terrains such as mountainous regions. Therefore, Molina *et al.* (2021) concluded that overall ERA5 winds demonstrate good performance across most European locations, but the accuracy varied with terrain complexity and wind speed ranges.

Compared to near-surface wind speeds, much less ERA5 evaluation studies have considered wind gusts. Minola *et al.* (2020) evaluated ERA5 near-surface wind speed and wind gust data in Sweden in 2013–2017 based on observations from 90 weather stations. They divided the stations to coast, inland and mountain stations based on the terrain characteristics around the stations and found that ERA5 exhibits the best performance at coastal stations, closely capturing both the seasonal and diurnal cycles of wind speeds and gusts. However, ERA5 showed terrain-dependent biases; it overestimated wind speeds in inland areas and substantially underestimated both wind speeds and wind gusts in mountainous regions. They additionally shared the commonly found result that ERA5 has a consistent tendency to overestimate low wind speeds and gusts and underestimate high wind speeds and gusts.

In Finland, both reanalysis and observations have been used in wind climate research (e.g. Laapas and Venäläinen 2017, Gregow *et al.* 2020, Laurila *et al.* 2021). However, to the authors' knowledge, previously there have not been any ERA5 wind evaluation studies specifically for Finland. The central goal of this study is to determine how well ERA5 represents observed near-surface wind speeds and wind gusts in Finland, and to clarify the extent to which ERA5 can be used as a reliable basis for wind-related research and applications.

Data and methods

Wind observations

Observations of wind speed and wind gust are obtained from the weather stations maintained by the Finnish Meteorological Institute (FMI). Wind gust observation network reached its cur-

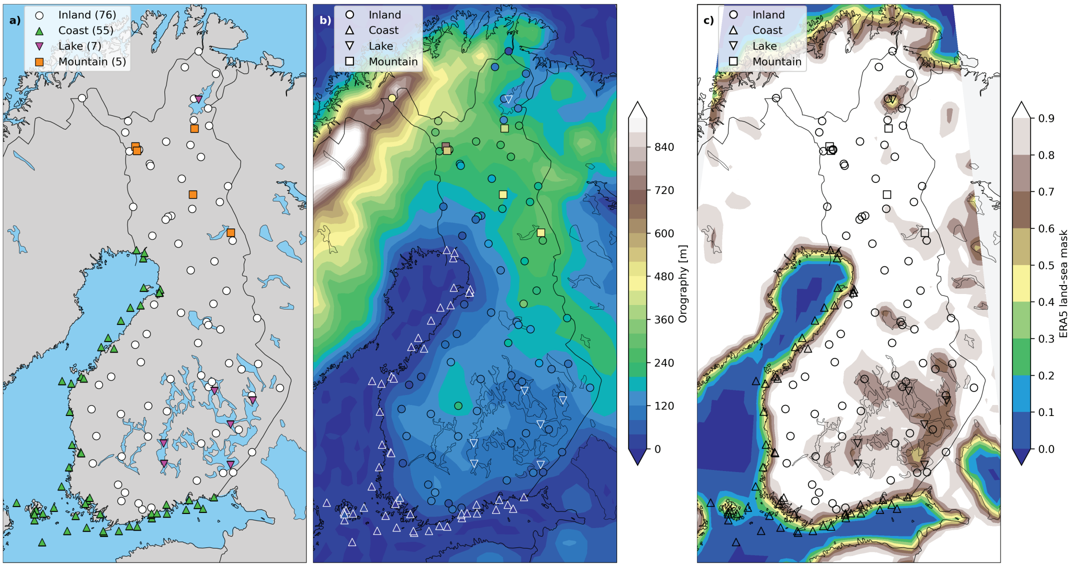

Fig. 1. (**a**) FMI's weather stations from which the wind observations are obtained. The symbols indicate the station classification: inland stations (76 in total) as circles, coast stations (55) as upward triangles, lake stations (7) as downward triangles and mountain stations (5) as squares. (**b**) Elevation of FMI's weather stations (classes as different symbols, filled color indicates the elevation value) and ERA5 orography (background map colors). (**c**) ERA5 land-sea mask, giving fractional values in the range 0 (sea) to 1 (land). Sea areas and major lakes are shown in Figure 1a by blue shading. In Figure 1c the major lakes are drawn with black contours in the background map.

rent extent about a decade ago. Therefore, we examine the period of 2014–2023 when both wind speed and wind gust observations are available from most of the weather stations.

First, wind speed and wind gust observations from FMI's weather stations were retrieved within a 10-minute time interval. Then, those stations were removed which have below 95% temporal data coverage, meaning that a station to be included can miss in total only 7 days' worth of data. Based on this criterion, 36 stations were rejected. The final dataset of wind observations includes 143 weather stations. Lastly, the observation data were converted to hourly values to be comparable to ERA5 data (which has a 1-hour time interval). The 10-minute wind speed values were averaged to hourly mean wind speeds, while the hourly maximum wind gusts were retrieved from 3-second wind gust observations.

ERA5 reanalysis

ERA5 reanalysis is the newest reanalysis dataset of the European Centre for Medium Range Weather Forecasts (ECMWF; Hersbach *et al.* 2020). It has a spatial resolution of around 31 km, and it covers a time period from 1940 onwards. The data fields are available every hour. In this study, we used the variables of instantaneous 10-m wind speed and hourly maximum 10-m wind gust.

In the ECMWF weather model, which is the basis for ERA5 reanalysis, wind gust is calculated as a sum of three components: 10-m wind speed, turbulent factor (describes surface friction and boundary layer stability) and convective factor (describes convective downdrafts) (ECMWF 2016). Therefore, ERA5 wind gusts are strengthened from 10-m wind speed by turbulent and convective factors.

Classification of weather stations

The final set of 143 weather stations were divided into four classes based on their location and elevation: inland (76 stations), coast (55 stations), lake (7 stations) and mountain (5 stations) classes (Fig. 1a). The classification was made manually to ensure the correct class for each weather station.

As the resolution of ERA5 is 31 km, it is apparent that it has a limited ability to represent sub-grid orography and surface roughness. When comparing the ERA5 orography and the elevation of the weather stations (Fig. 1b), we can see that the elevations of inland, coast and lake stations are mostly well comparable. However, ERA5 clearly underestimates the elevation of mountain stations (Fig. 1b). In addition, there is one inland station in the most northwestern Finland (the Käsivarsi area) in the mountainous region where ERA5 orography is higher than the actual elevation of the station since ERA5 does not differentiate the small-scale variation in the terrain.

Obviously, surface roughness is larger over land, with e.g. grass, forest, and buildings, than over water areas, which are smooth. High surface roughness slows down the wind speeds by blocking the free air flow and thus, the modelled surface roughness in a reanalysis has a large effect on the resulting winds. The spatial resolution of 31 km in ERA5 leads to challenges in representing the sub-grid variations in surface roughness, especially over the sub-grid-scale lakes. In the ERA5 land-sea mask, which indicates the proportion of land in each grid (Fig. 1c), the large lakes in Finland have somewhat lower land proportion and therefore lower surface roughness than the land areas. However, the ERA5 land-sea mask values for the lake station grids, varying from 0.42 to 0.85 with an average of 0.70, do not represent the real conditions. In reality for the lake stations, wind fetches over open lake can be up to 10 km long (depending on the direction) and therefore the realistic roughness values would be close to zero. Hence, the lakes in Finland tend to have too high surface roughness in ERA5.

ERA5 gives the wind speeds and wind gusts at 10-m height above ground. The wind instruments in weather stations are, however, in practice installed to various heights depending on e.g. the surrounding environment and how that is assessed to affect the wind. In selecting the installation height for the instruments, the aim is that the surroundings would not affect the speed or direction of wind velocity. In this study, the environmental representativeness of the weather stations was not a selection criterion (but only the data coverage as described earlier). Therefore, in addition to the elevation of the location, the height level of the wind measurement may differ between ERA5 and observations.

Results

In this section, first the 10-year average wind speeds and wind gusts are presented on a map to show the spatial distribution. Then, the evaluation of the hourly values is investigated in regard to correlation, bias, and frequency distribution. Finally, the monthly mean values of ERA5 and observations are examined.

Average over the whole time period (2014–2023)

The average wind speed (Fig. 2a) and wind gust (Fig. 2b) over the 10-year study period of 2014–2023 show a clear land-sea difference with stronger wind speeds and gusts over sea than over land. This is visible in ERA5 as well as observations when comparing inland and coast stations. The mean winds in inland stations are in general slightly overestimated in ERA5 compared to observations, especially for wind gusts (Fig. 2b). In coast stations, the observed winds are largely similar to ERA5 values. In ERA5, large lakes of Finland, that also stand out from ERA5 land-sea mask (Fig. 1c), do not stand out with higher winds although the observed wind speeds and gusts in lake stations are stronger than in inland stations. The biggest difference is seen in mountain stations where ERA5 largely underestimates the mean wind speeds and gusts.

Correlation of observed and ERA5 winds

The correlations between observations and ERA5 for hourly mean wind speed and hourly maximum wind gust during the period of

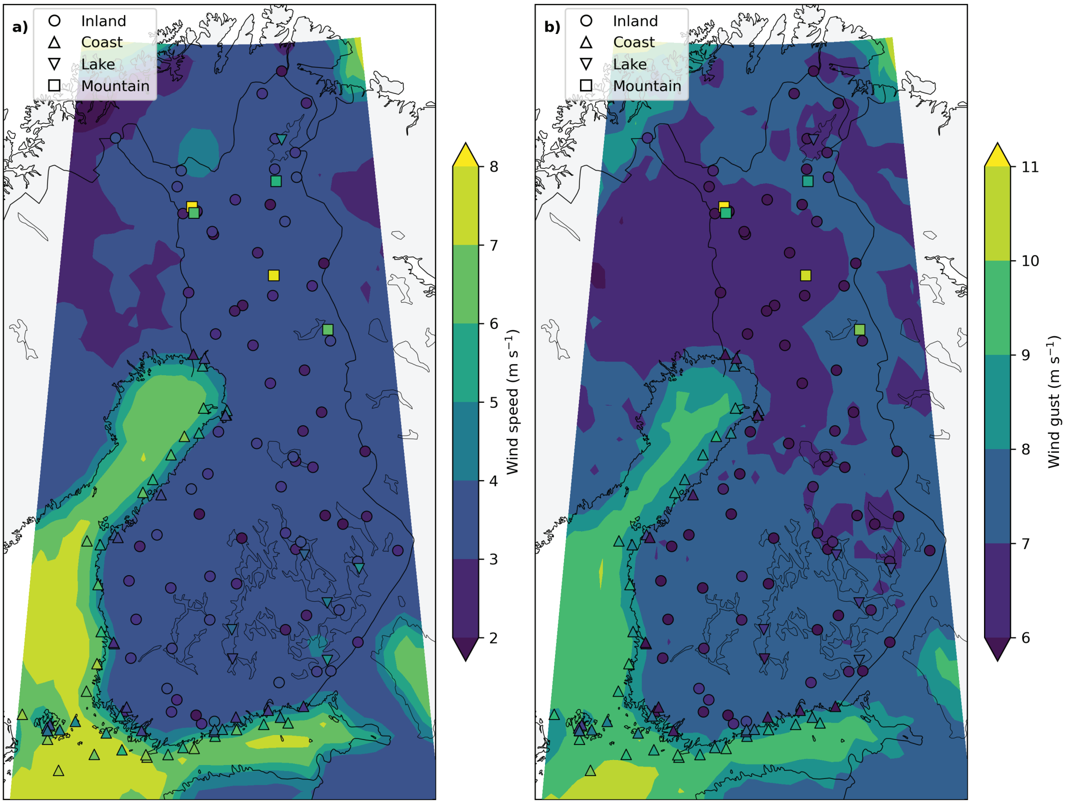

Fig. 2. (**a**) Wind speed and b) wind gust of FMI's weather stations (classes as different symbols) and ERA5 (map colors) as averages over the whole study period 2014–2023.

2014–2023 are presented in Figs. 3 and 4. To address autocorrelation, we subsampled the hourly data using the first lag for which the autocorrelation dropped below 0.1, this lag ranging up to 24 hours. Overall, the correlations are strong or very strong (autocorrelation-adjusted correlation coefficients between 0.74 and 0.86, *p*-values $p < 10^{-16}$) in all station classes for both the wind speed and wind gust. The correlation is somewhat better for wind gust (Fig. 4) than for wind speed (Fig. 3), being the highest ($r = 0.86$) for wind gust in the coast stations. Although the correlations are strong, it is apparent that in general the weak winds are overestimated, whereas the strong winds are underestimated. The clearest underestimation in ERA5 is found in mountain stations. The elevation difference in mountain stations between ERA5 and observation stations (see Fig. 1b) likely affects the large underestimation in ERA5.

Because many studies consider extreme winds using high percentiles, also the correlations between observations and ERA5 for the 95th and 98th percentiles of wind speeds and wind gusts were calculated and are attached to the Supplementary Information (see Figs. S1 and S2 in Supplementary Information). The percentiles are calculated from the 2014–2023 period so that each station has one observed value and one ERA5 value for the whole period. Regarding wind speeds, the average high percentile wind speed (see triangles in Fig. S1 in Supplementary Information) in ERA5 is the most comparable to observations inland and the least in mountainous areas. Regarding wind gusts, the average high percentile wind gust (see triangles in Fig. S2 in Supplementary Information) in ERA5 performs the best in coast stations. However, the performance of ERA5 varies between stations and the spread is large, especially in coast and mountain stations.

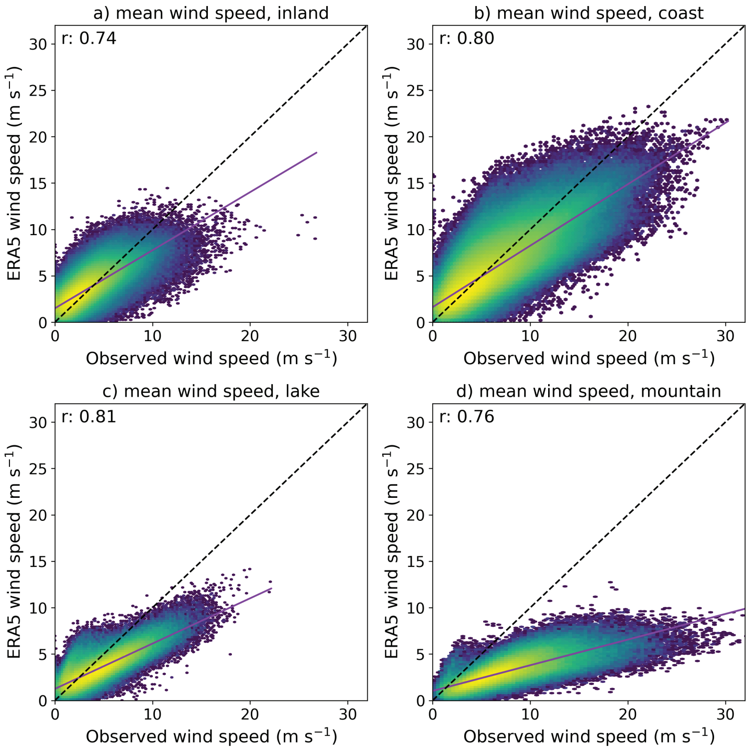

Fig. 3. Correlation of hourly mean wind speed between observations and ERA5 in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations. The black dotted line is the 1:1 line, and the purple solid line is the fitted linear regression line. The *r* values are autocorrelation-adjusted correlation coefficients.

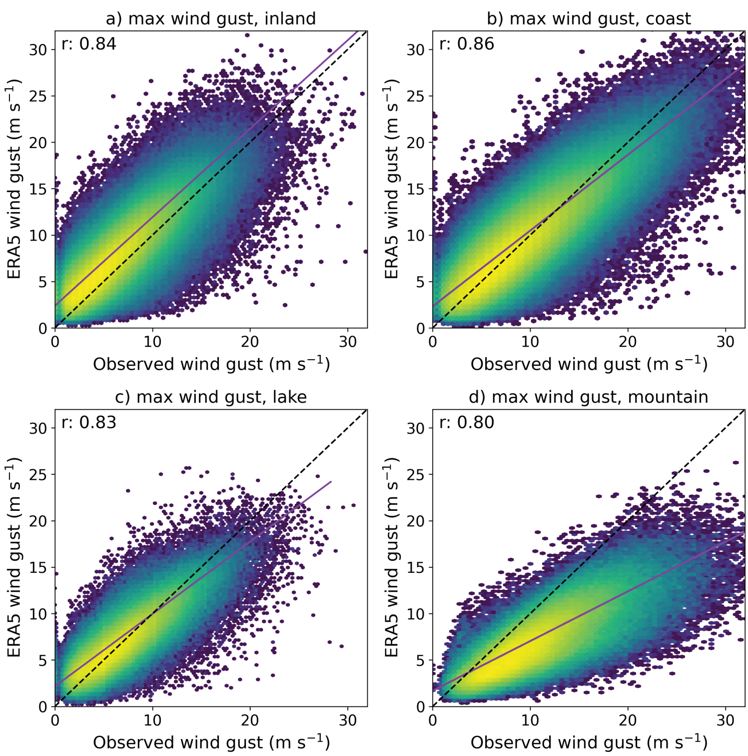

Fig. 4. Correlation of hourly maximum wind gust between observations and ERA5 in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations. The black dotted line is the 1:1 line, and the purple solid line is the fitted linear regression line. The *r* values are autocorrelation-adjusted correlation coefficients.

Bias of ERA5 winds compared to observations

Here, bias is calculated as the difference between hourly ERA5 and observed wind value (i.e., ERA5 minus observed wind). The winds are divided into 5 m s^{-1} intervals to see if and how the bias differs between weak and strong winds. It must be noted that there are considerably more occurrences, i.e. hourly observations, with lower wind speed intervals than with higher ones. For example, there are only 26 occurrences in total (of which 25 are in mountain stations) when observed wind speed is over 30 m s^{-1} and only 26 occurrences when observed wind gust is over 40 m s^{-1} (compared to about 7.6 million occurrences of observed wind speeds and 3.7 million occurrences of observed wind gusts below 5 m s^{-1}).

Regarding the hourly wind speed, the bias increases towards stronger wind speeds (Fig. 5). In inland and coast stations, the weakest winds (< 5 m s^{-1}) have a positive bias, i.e., ERA5 overestimates the weak winds compared to observations, while wind speeds over 5 m s^{-1} are mostly underestimated in ERA5. In lake and mountain stations, even weak wind speeds show a negative bias, i.e. underestimation of ERA5 wind speeds. Overall, the variation of the biases (the height of the gray box and the distance between whiskers) within the wind intervals is relatively small. Comparing the different station classes, the bias is the smallest in coast stations and the largest in mountain stations. For example, observed strong wind speeds between 20 to 25 m s^{-1} are underestimated in ERA5 for around 30 % in coast stations, 45 % in lake stations, 55 % in inland stations and 70 % in mountain stations.

The hourly wind gust shows an increasing bias towards stronger wind gusts (Fig. 6) resembling that of the hourly wind speed bias. However, the increase is not as steep for the wind gust as for the wind speed. The bias is relatively small with wind gusts below 20 m s^{-1} in all station classes except mountain stations. Observed wind gusts between 20 to 25 m s^{-1} are underestimated in ERA5 by around 15 % in inland and coast stations, 25 % in lake stations and 45 % in mountain stations. The ERA5 bias is smaller for the wind gusts than for the wind speeds in most of the wind intervals and station classes. An exception is found in one coast station where the strongest observed wind gust of over 40 m s^{-1} has the largest bias of all, but this is only one occurrence and therefore may not be generalized. Moreover, since there are considerably fewer cases of extreme winds than weak winds, the sample sizes differ between the wind gust classes, and this reduces comparability of the categories.

As wind speeds in Finland have a seasonal cycle with the strongest winds during winter and the weakest during summer (e.g. Laurila *et al.* 2021), the ERA5 wind bias is additionally examined separately for seasons. In inland and coast stations, the ERA5 bias is mostly lower in winter than in summer, with both wind speed (Fig. 7) and wind gust (Fig. 8). This is more clearly visible with stronger (> 15 m s^{-1}) wind speeds and wind gusts and, therefore, it may be due to the different type of strong winds causing storms in summer and winter. In summer, thunderstorms are typically small-scale (even just 10 km in diameter) and the strong winds last only a short time (minutes) in one location, while wintertime windstorms are large-scale (even thousands of kilometers) and the strong winds in one location can last a long time (hours). Hence, large-scale windstorms are much better represented in ERA5 than small-scale thunderstorms. In lake stations, the bias is relatively similar in all seasons, especially for wind speeds whereas wind gusts show small variations between seasons. On the contrary, in mountain stations the bias is mostly the largest in winter and the lowest in summer. This is likely because ERA5 is not able to replicate the seasonal variation in winds over the mountains and hence the bias is larger in winter when the observed winds are stronger in general.

Frequency distributions of observed and ERA5 winds

The frequency distributions of wind speeds (Fig. 9) and wind gusts (Fig. 10) show the differences in the whole range of winds between ERA5 and observations. In the inland stations (Figs. 9a and 10a), the peak of the ERA5 distribution is shifted towards high values compared to the observed distribution in both wind speeds and wind gusts. This is also visible in the coast

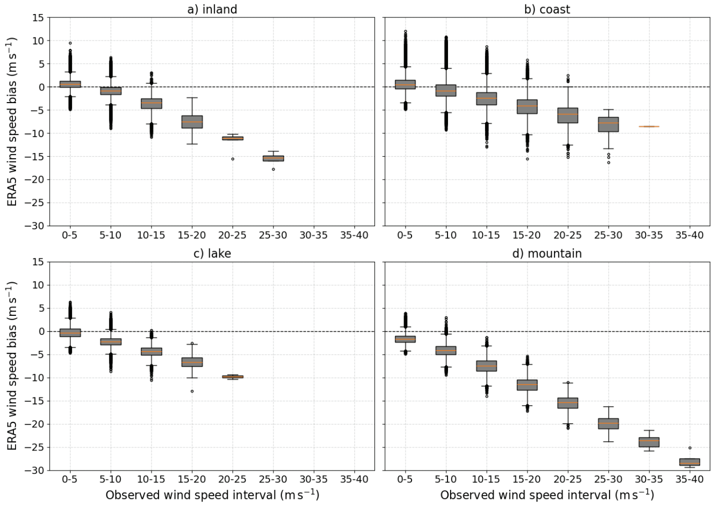

Fig. 5. Bias box plots in ERA5 hourly mean wind speed compared to observations in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations. The box plot displays the median (orange line), interquartile range (box), the most extreme data points within 1.5 times the interquartile range (whiskers), and outliers outside the whisker range (small circles). There are less than 10 data points in the following classes and intervals: inland > 20 m s^{-1}, coast > 30 m s^{-1}, lake > 20 m s^{-1} and mountain > 35 m s^{-1}.

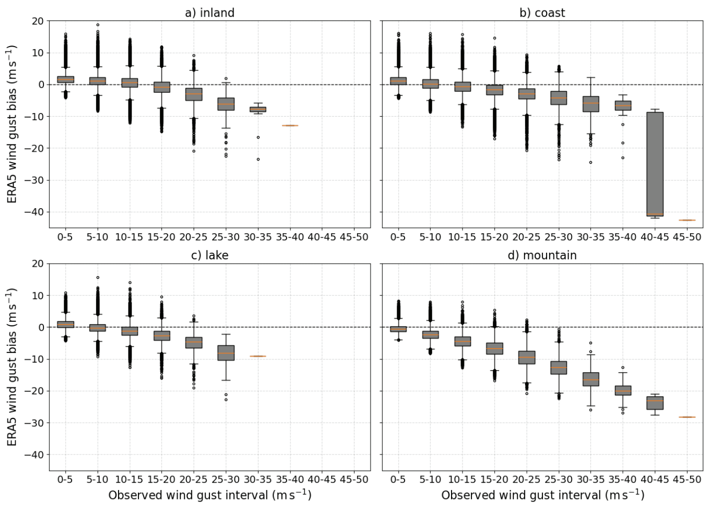

Fig. 6. Bias box plots in ERA5 hourly maximum wind gust compared to observations in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations. The box plot displays the median (orange line), interquartile range (box), the most extreme data points within 1.5 times the interquartile range (whiskers), and outliers outside the whisker range (small circles). There are less than 10 data points in the following classes and intervals: inland > 35 m s^{-1}, coast > 45 m s^{-1}, lake > 30 m s^{-1} and mountain > 45 m s^{-1}.

stations (Figs. 9b and 10b), although to a lesser degree. This means that ERA5 overestimates the majority of the weak wind speeds and wind gusts, which was also seen from the correlation and bias results. In the lake stations, the distribution peaks are at the same location, while in the mountain stations the ERA5 distribution peak is largely shifted towards low values compared to the observations. The shapes of the distributions are more comparable between ERA5 and observations for wind gusts than for wind speeds in all station classes.

The right tail of the distribution, consisting of the most extreme winds, extends to much higher values in observations than in ERA5. In inland stations, the maximum wind speed in ERA5 is 14.5 m s^{-1} while the observed maximum is 26.8 m s^{-1}. For wind gusts in inland stations, the ERA5 maximum is 31.5 m s^{-1} and observed maximum 35.2 m s^{-1}, so the difference in the maximum values is clearly smaller in wind gust (4.0 m s^{-1}) than wind speed (12.3 m s^{-1}). In coast stations, the difference between ERA5 (23.2 m s^{-1}) and observed (30.3 m s^{-1}) maximum wind speed is 7.1 m s^{-1}, and the maximum wind gust difference is 10.4 m s^{-1}. In the lake stations, the differences between ERA5 and observed maximum wind speed and wind gust, 8.0 m s^{-1} and 4.9 m s^{-1},

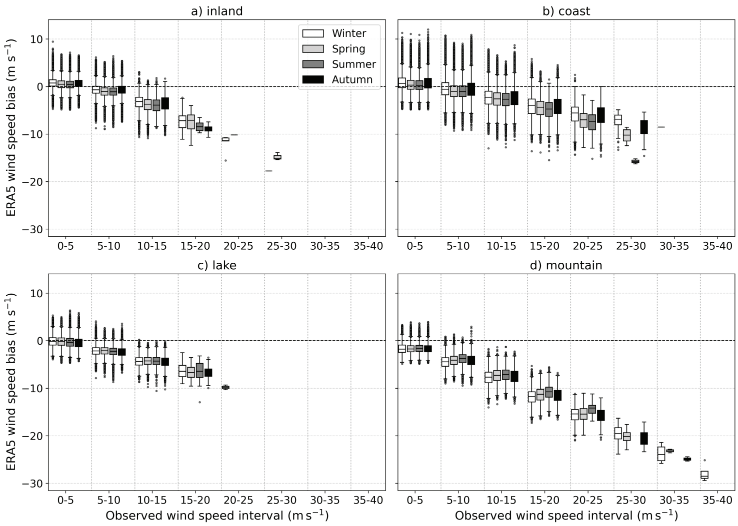

Fig. 7. Seasonal bias box plots in ERA5 hourly mean wind speed compared to observations in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations. The box plot displays the median (black line), interquartile range (box), the most extreme data points within 1.5 times the interquartile range (whiskers), and outliers outside the whisker range (small circles).

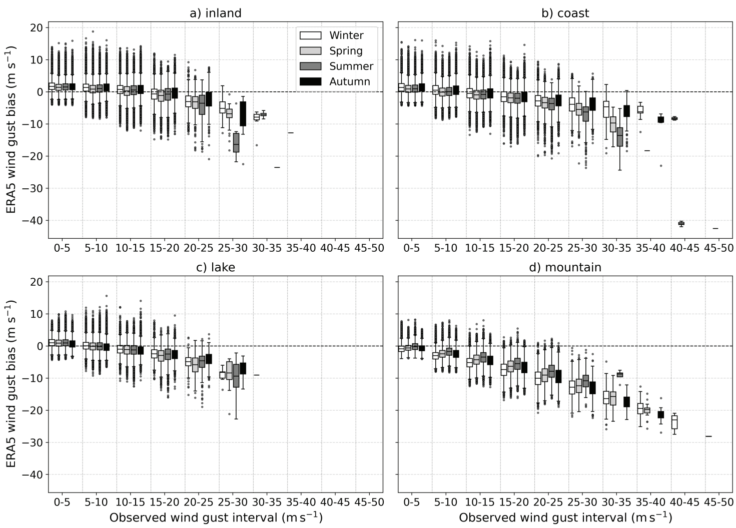

Fig. 8. Seasonal bias box plots in ERA5 hourly maximum wind gust compared to observations in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations. The box plot displays the median (black line), interquartile range (box), the most extreme data points within 1.5 times the interquartile range (whiskers), and outliers outside the whisker range (small circles).

respectively, are comparable to the inland and coast stations. However, in the mountain stations, the maximum values differ largely being over 20 m s^{-1} lower in ERA5 than in observations.

Monthly means of observed and ERA5 winds

The best agreement for the monthly means of wind speed and wind gust between ERA5 and observations is found in the coast stations (Fig. 11b). ERA5 monthly means are only slightly differing from the observed values, and the seasonal variation, with the highest winds during autumn and winter, resembles the observed seasonality. In the lake stations (Fig. 11c), the monthly means of ERA5 wind gusts are likewise well represented. However, wind speeds are distinctly underestimated in ERA5. Nonetheless, the seasonal variation in ERA5 is similar to the observations also in the lake stations.

In the inland stations (Fig. 11a), ERA5 monthly means are overestimated in all months. This is more pronounced with wind gusts than with wind speeds. This finding probably stems from the fact that the wind speeds over inland are

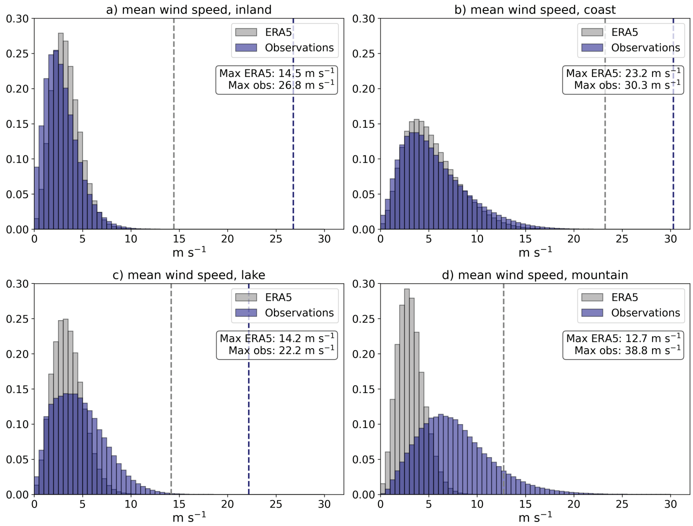

Fig. 9. Distributions of hourly mean wind speed in observations and ERA5 in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations.

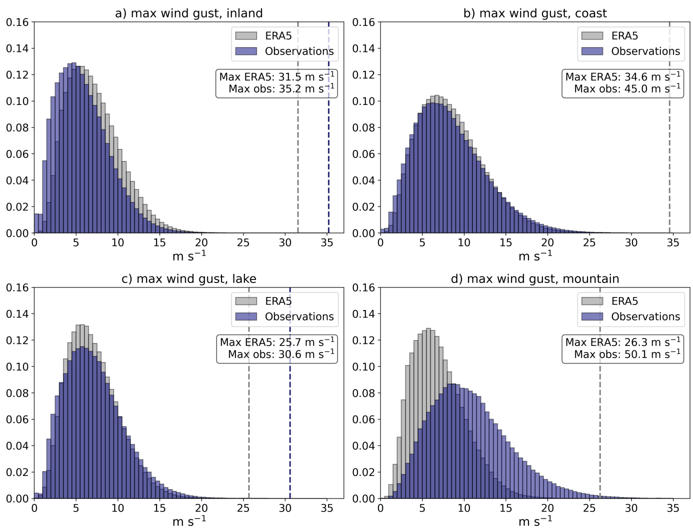

Fig. 10. Distributions of hourly maximum wind gust in observations and ERA5 in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations.

mostly quite weak overall, and the weak hourly winds in ERA5 are in general overestimated (Fig. 3) and hence are also the monthly means. The clearest differences between ERA5 and observed monthly means are again found in the mountain stations (Fig. 11d), and as was noted before, ERA5 does not capture the observed seasonal variation in the mountainous regions.

Discussion and conclusions

The results of this study show that wind speed and wind gust in ERA5 in comparison to observations in Finland perform differently depending on the topography and surface type. A similar kind of study has been made by Minola *et al.* (2020) who compared wind speed and wind gust observations in Sweden in 2014–2017 to ERA5 and ERA-Interim reanalysis data. They classified the weather stations to inland, coast and mountain stations like in our study, except that we added a class 'lake' since there are multiple weather stations in Finland where the nearby lake influences the winds. The outcomes of this paper are largely in agreement with the results from Minola *et al.* (2020). They also found that the coast stations have a better agreement

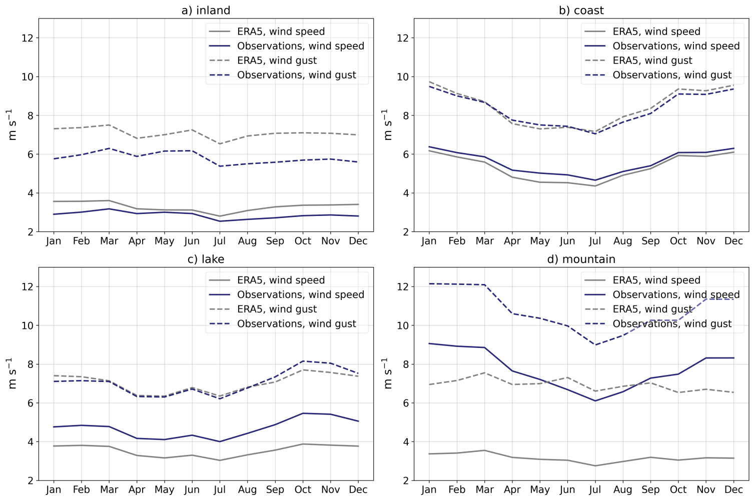

Fig. 11. Monthly means of hourly mean wind speed (solid lines) and hourly maximum wind gust (dashed lines), as derived from observations and ERA5, in different classes: (**a**) inland, (**b**) coast, (**c**) lake and (**d**) mountain stations.

between ERA5 and observations than inland and mountain stations. Our results show the overestimation of weak winds and an increasing bias towards higher winds that have previously been found in many studies (e.g. Chen *et al.* 2024, Jiang *et al.* 2021, Molina *et al.* 2021, Minola *et al.* 2020). However, the performance of ERA5 in representing monthly mean winds had some differences between Sweden and Finland. While our study shows that in the Finnish coast stations ERA5 compares well with the monthly mean wind speeds and wind gusts, Minola *et al.* (2020) found ERA5 slightly overestimating both the wind speeds and wind gusts in the Swedish coast stations. In the inland stations, both our study and Minola *et al.* (2020) show ERA5 overestimating monthly mean wind speeds and wind gusts, but the bias is greater in Sweden than in Finland. The large underestimation of ERA5 monthly mean winds in the mountain stations is seen both in Finland and Sweden.

By investigating two Finnish weather stations, Rantanen *et al.* (2021) found ERA5 wind gust distribution to be in better agreement with observations than ERA5 wind speed distribution. In their case study of storm Aila, Rantanen *et al.* (2021) considered wind speed and wind gust distributions in Septembers 2004–2020 retrieved from ERA5 and from observations at the Rauma and Pietarsaari weather stations in the Finnish coast where record-high winds were observed during storm Aila. They noticed that ERA5 represented better the distribution of wind gusts than that of wind speeds but highlighted that because two stations only were included in their study, as well as only the month of September, further research on the topic was required.

In the current study, 143 FMI weather stations were considered, categorizing them into four classes: inland (76 stations), coast (55 stations), lake (7 stations), and mountain (5 stations). Therefore, our study fills the research gap raised by Rantanen *et al.* (2021) and shares the same result with a larger sample of weather stations covering the whole year. As described in the Data section, wind gust in ERA5 is calculated as a sum of three components: 10-min wind speed, a turbulent factor, and a convective factor. Since the hourly wind speeds in ERA5 are in general underestimated (except for the weakest winds), one or both of the other two components need to overestimate the hourly wind gusts as the end result is closer to observed. A future study could investigate the ECMWF wind gust parameter in more detail to examine how one or both of the factors overestimate the wind gusts. Minola *et al.* (2020) already took steps this way by developing an improved gust parameter for Sweden by adding an elevation dependency and by tuning the convective gust contribution.

Our study shows that the distribution shapes between ERA5 and observations resemble each other quite well, especially in the inland and coast stations. The distributions are better aligned for wind gusts than for wind speeds. Therefore, while it is apparent that the absolute values of high wind speeds and wind gusts are underestimated in ERA5, the high percentiles obtained from ERA5 are sensible to be used in extreme wind studies. Among the four station classes, coast and inland stations showed the best alignment with ERA5. In Finland, most of the land areas are inland or coast types, which is also evident from the station amounts (131 stations in total in inland and coast classes, compared to lake and mountain classes that have only 12 stations in total). This underscores the dataset's suitability for wind-related modeling, assessments, and applications in Finland. However, caution is warranted when interpreting ERA5 data in mountainous regions, where the agreement with observations is notably weaker.

ERA5 is increasingly used due to its global and spatially and temporally homogeneous representation of numerous weather variables. The primary aim of this study was to evaluate the performance of ERA5 in representing observed near-surface wind speeds and gusts in Finland. In doing so, the study provides new insight into the reliability of ERA5 for wind-related research and applications. The results provide valuable new information for a wide range of applications. ,It is important to recognize that ERA5 is a reanalysis rather than an observational dataset, and different reanalyses may yield varying results. Therefore, the evaluation of ERA5's ability to represent observed wind conditions in Finland is essential. Moreover, our findings benefit not only the wind-specific studies but also those involving compound events, such as the co-occurrence of strong winds with heavy snowfall, phenomena that are affected by winds, such as wildfire spread and sea level oscillations, or impact research on, for example, forest damage or coastal flooding.

Acknowledgements: This research was partially funded by the Finnish State Nuclear Waste Management Fund (VYR) through the MAWECLI project (Dnro SAFER 6/2023, 8/2024 and 5/2025) in the SAFER2028 programme.

Supplementary Information: The supplementary information related to this article is available online at: https://doi.org/10.60910/ber2025.q6kf-qp39

References

Chen, T.-C., Collet, F., & Di Luca, A. 2024. Evaluation of ERA5 precipitation and 10-m wind speed associated with extratropical cyclones using station data over North America. *Int. J. Climatol*., 44(3), 729–747. https://doi.org/10.1002/joc.8339

ECMWF (European Center for Medium-Range Weather Forecasts) 2016. IFS Documentation CY45R2—Part IV: Physical processes. IFS doc. https://www.ecmwf.int/node/16648. Accessed 2 May 2025.

Gregow, H., Rantanen, M., Laurila, T. K. &, Mäkelä, A. 2020. Review on winds, extratropical cyclones and their impacts in Northern Europe and Finland. *Finnish Meteorological Institute Reports*, 2020(3). http://hdl.handle.net/10138/320298 (accessed 3 June 2025).

Gualtieri, G. 2021. Reliability of ERA5 Reanalysis Data for Wind Resource Assessment: A Comparison against Tall Towers. *Energies*, 14(14), 4169. https://doi.org/10.3390/en14144169

Haakana, J., Räisänen, O., Karhunen, M., Láng-Ritter, I., and Lassila, J., 2024. Impact of wind speed and soil frost on electricity distribution system reliability. *Heliyon*, 10(24), https://doi.org/10.1016/j.heliyon.2024.e40846

Hersbach H., Bell B., Berrisford P. *et al*., 2020. The ERA5 global reanalysis. *Q J R Meteorol Soc*., 146: 1999–2049. https://doi.org/10.1002/qj.3803

Jafari, M., & Alipour, A., 2021. Methodologies to mitigate wind-induced vibration of tall buildings: A state-of-the-art review. *Journal of Building Engineering*, 33, 101582. https://doi.org/10.1016/j.jobe.2020.101582

Jasiūnas, J., Heikkinen, T., Lund, P. D., and Láng-Ritter, I., 2023. Resilience of electric grid to extreme wind: Considering local details at national scale. *Reliability Engineering & System Safety*, 232, https://doi.org/10.1016/j.ress.2022.109070

Jiang, Y., Han, S., Shi, C., Gao, T., Zhen, H., & Liu, X. 2021. Evaluation of HRCLDAS and ERA5 Datasets for Near-Surface Wind over Hainan Island and South China Sea. *Atmosphere*, 12(6), 766. https://doi.org/10.3390/atmos12060766

Jylhä, K., Kämäräinen, M., Fortelius, C., Gregow, H., Helander, J., Hyvärinen, O., Johansson, M., Karppinen, A., Korpinen, A., Kouznetsov, R., Kurzeneva, E., Leijala, U., Mäkelä, A., Pellikka, H., Saku, S., Sandberg, J., Sofiev, M., Vajda, A., Venäläinen, A. & Vira, J. 2018. Recent meteorological and marine studies to support nuclear power plant safety in Finland. *Energy*, 165. https://doi.org/10.1016/j.energy.2018.09.033

Laapas, M. & Venäläinen, A. 2017. Homogenization and trend analysis of monthly mean and maximum wind speed time series in Finland, 1959–2015. *Int. J. Climatol.*, 37: 4803–4813. https://doi.org/10.1002/joc.5124

Láng, I., Mäkelä, A., Korpela, P., and Gregow, H., 2021. Windstorm Aila 2020: societal impacts. *FMI's Climate Bulletin: Research Letters*, 3(1), 15–17, https://doi.org/10.35614/ISSN-2341-6408-IK-2021-05-RL

Laurila T. K., Sinclair V. A., & Gregow H., 2021. Climatology, variability, and trends in near-surface wind speeds over the North Atlantic and Europe during 1979–2018 based on ERA5. *Int. J. Climatol.*, 41: 2253–2278. https://doi.org/10.1002/joc.6957

Laurila, T.K., Rantanen, M. and Punkka, A.-J. (2025), Storm Lyly, the first storm with observed hurricane-force winds in Finland. *Weather*, 80: 124–126. https://doi.org/10.1002/wea.7684

Leijala, U., Björkqvist, J.-V., Johansson, M. M., Pellikka, H., Laakso, L., & Kahma, K. K., 2018. Combining probability distributions of sea level variations and wave run-up to evaluate coastal flooding risks, *Nat. Hazards Earth Syst. Sci.*, 18, 2785–2799, https://doi.org/10.5194/nhess-18-2785-2018

Minola, L., Zhang, F., Azorin-Molina, C. *et al.*, 2020. Near-surface mean and gust wind speeds in ERA5 across Sweden: towards an improved gust parametrization. *Clim. Dyn.*, 55, 887–907. https://doi.org/10.1007/s00382-020-05302-6

Molina M. O., Gutiérrez C., & Sánchez E. 2021. Comparison of ERA5 surface wind speed climatologies over Europe with observations from the HadISD dataset. *Int. J. Climatol.*, 41: 4864–4878. https://doi.org/10.1002/joc.7103

Ottosen, T. B., Ketzel, M., Skov, H. *et al.* 2019. Microscale modelling of the urban wind speed for air pollution applications. *Sci. Rep.*, 9, 14279. https://doi.org/10.1038/s41598-019-50033-2

Rantanen, M., Laurila, T. K., Sinclair, V. & Gregow, H., 2021. Storm Aila: An unusually strong autumn storm in Finland. *Weather*, 76: 306–312. https://doi.org/10.1002/wea.3943

Romagnoli, F., Cadei, A., Costa, M., *et al.* 2023. Windstorm impacts on European forest-related systems: An interdisciplinary perspective. *Forest Ecology and Management,* 541, https://doi.org/10.1016/j.foreco.2023.121048

Soares, P. M., Lima, D. C., and Nogueira, M. 2020. Global offshore wind energy resources using the new ERA-5 reanalysis. *Environmental Research Letters*, 15(10), 1040a2. https://doi.org/10.1088/1748-9326/abb10d

Taszarek, M., Kendzierski, S., & Pilguj, N. 2020. Hazardous weather affecting European airports: Climatological estimates of situations with limited visibility, thunderstorm, low-level wind shear and snowfall from ERA5. *Weather and Climate Extremes*, 28, 100243. https://doi.org/10.1016/j.wace.2020.100243

Vajda, A., Tuomenvirta, H., Juga, I. *et al.* 2014. Severe weather affecting European transport systems: the identification, classification and frequencies of events. *Nat. Hazards 72*, 169–188. https://doi.org/10.1007/s11069-013-0895-4

Virman, M., Rantanen, M., Laurila, T.K., Mäkelä, A., van den Broek, D. and Gregow, H. (2025), Damaging thunderstorms in Finland in June 2021. *Weather*, 80: 111–119. https://doi.org/10.1002/wea.7602