Evaluating effects of the terrain on modelled winds in multiple atmospheric model datasets

Maksims Pogumirskis[1, 2], Tija Sīle[1], Lasse Svenningsen[2], and Andrea N. Hahmann[3]

[1]Institute of Numerical Modelling, Faculty of Science and Technology, University of Latvia, Riga, Latvia
[2]EMD International A/S, Aalborg, Denmark
[3]Department of Wind and Energy Systems, Technical University of Denmark, Roskilde, Denmark

Correspondence: Maksims Pogumirskis (maksims.pogumirskis@lu.lv)

Abstract. Numerical atmospheric models are widely used as a source of meteorological data when planning the locations of new wind farms. However, before relying on model output for decision-making, it must be verified against observations. Due to commercial restrictions on the availability of observation data, previous studies on atmospheric model validation for wind energy applications are often limited to a single model or a small geographical region. In this work we perform a large-scale bias assessment of modelled winds at heights spanned by turbine rotors from seven model datasets against data from more than 500 observation campaigns across Europe. Principal component analysis is used to identify spatial, diurnal, and seasonal patterns of wind speed and direction biases. The results of the analysis show that all seven model datasets exhibit similar spatial and temporal patterns of wind speed bias. Models generally show a more positive wind-speed bias in the Central European Plain and a more negative bias in mountainous regions, namely Southern Europe and the Scandinavian Mountains. Moreover, the temporal patterns of biases differ between these regions, and wind direction bias exhibits the same temporal and spatial patterns as wind speed bias. We show that the difference in terrain elevation between the models and the real world is an important statistical factor in explaining both wind speed and direction biases. The magnitude of the wind speed bias ranges from 0.1 to 0.9 ms^{-1} per 100 m of elevation difference, depending on the season and time of day. Two WRF model simulations with different terrain source data are performed, and the modelled winds are compared to provide more robust support for the hypothesis. The results of this work suggest that improving terrain representation in mesoscale models can help improve their performance.

1 Introduction

Global warming, which is primarily driven by anthropogenic greenhouse gases, is a major threat to both humans and ecosystems. In 2016, 196 countries signed the Paris Agreement with the goal of limiting global warming to 2 °C, preferably 1.5 °C, compared to pre-industrial levels (United Nations Climate Change (UNFCCC)). In pursuit of implementing the agreement, the European Union has set the goal of reaching a 45 % share of renewable energy by 2030. To meet the goal, it is estimated that wind turbines with a cumulative capacity of at least 37 GW must be installed each year (European Commission, 2023).

Wind farms require high upfront investment, which is offset by significantly lower operating costs than fossil-fuel plants. Therefore, in the planning stages, it is important to ensure they provide stable, reliable power output throughout their lifetime.

This is needed to ensure profitability, keep electricity prices stable, and maintain a reliable, predictable electricity supply (Backwell et al., 2025; Williams et al., 2025). When planning the locations of new wind farms, it is necessary to acquire precise information about the wind climate at their potential locations. Wind power density is proportional to the cube of wind speed. Therefore, a small error in wind speed estimates can lead to substantially larger errors in wind power production estimates.

Direct measurements are commonly used for wind resource assessment. While the World Meteorological Organization (WMO) observation station network is broad and uniformly covers most of the land, it provides only surface wind observations at 10 m height, which are less useful for estimating the wind at heights spanned by turbine rotors (typically between 50 m and 250 m above ground level). The wind measurement network at heights spanned by turbine rotors is much scarcer than the synoptic network (Ramon et al., 2020). According to MEASNET (2022), for assessing site-specific conditions, wind speed measurements are required at a minimum height equal to 2/3 of the turbine hub height. Preferably, measurement heights should span the entire rotor area. The heights of modern wind turbine rotors exceed 200 m, making tall masts impractical for wind resource assessment at higher elevations (Cheynet et al., 2025). Although lidars allow one to acquire wind data at higher levels more easily, they are unreliable in case of fog or heavy rain (Knoop et al., 2021). Moreover, to acquire a meaningful amount of data, the measurement campaign must be conducted for at least a year, which requires regular equipment maintenance. Wind measurements are highly sensitive to the terrain and other nearby obstacles. According to MEASNET (2022), under homogeneous surface-roughness conditions, most observations are representative over a radius of 2 km in complex terrain and 10 km in simple terrain. Therefore, a separate measurement campaign has to be conducted at each potential wind farm location.

Numerical modelling is an alternative way to obtain information about wind conditions at a site and has several advantages over direct measurements. It allows to produce spatially and temporally continuous wind data. Moreover, running numerical models is much faster and cheaper than performing measurements. However, the results of numerical models should not be relied on without verifying that they generally agree with observations, and their limitations are well understood.

Previous studies that have compared atmospheric model output against observations for wind energy applications are usually limited to a single model, e.g. CERRA (Jourdier et al., 2023; Rouholahnejad et al., 2024), NEWA (Dörenkämper et al., 2020; Meyer and Gottschall, 2022), NORA3 (Solbrekke et al., 2021), or multiple models in a limited geographical region, e.g., France (Jourdier, 2020), Sweden (Kuru, 2023; Zhi, 2023), North Sea (Kalverla et al., 2020; Cheynet et al., 2022; Spangehl et al., 2023). It has been identified that atmospheric models have systematic issues in representing the observed wind climate.

The model's ability to resolve flows affected by terrain is largely dependent on its spatial resolution. Synoptic-scale models usually have horizontal grid spacings of several tens of kilometres, cover the entire globe, and are designed to capture large-scale synoptic features such as anticyclones, cyclones, and fronts. The typical effective resolution of synoptic scale models is approximately 3–4 times their grid spacing (Klaver et al., 2020). Processes that occur at a scale smaller than the effective resolution might not be properly simulated. Studies have shown that, due to their coarse resolution, synoptic scale models generally perform worse in complex terrain (Das and Baidya Roy, 2024) or near complex coastlines (Gualtieri, 2021).

To improve modelled winds in regions with complex land properties, mesoscale modelled datasets are created by dynamically downscaling synoptic scale model output. Mesoscale models usually have horizontal grid spacings of a few kilometres,

cover only a limited area, and can capture winds on finer scales. The effective resolution of the WRF mesoscale model is approximately 7 times its grid resolution (Skamarock, 2004). Studies have shown that mesoscale models outperform synoptic scale models, especially when modelling extreme wind speeds (Solbrekke et al., 2021) or winds in complex terrain (Dörenkämper et al., 2020).

Both synoptic-scale and mesoscale models commonly use terrain-following coordinates, which can introduce numerical artefacts in steep terrain (Shaw and Weller, 2016) and therefore might require additional smoothing of the model terrain (Bouëdec et al., 2025). Smoothing the terrain can degrade the representation of fine features, such as narrow valleys (Schmidli et al., 2018). To resolve the effects of complex terrain, microscale modelling can be performed. The typical horizontal grid spacing of microscale models is on the order of tens of metres. In addition, in microscale models, fine terrain features can be resolved by performing local mesh refinement with unstructured meshes (Yamaguchi et al., 2024) or by using immersed boundary methods (Jafari et al., 2011). The major downside of microscale models is that they are computationally expensive and cannot be directly used to produce long-term wind time series.

We perform a large-scale bias assessment of seven atmospheric model datasets for wind energy applications. Winds at heights relevant for wind energy from three global reanalyses and four mesoscale model datasets are compared with observations from 508 observation campaigns covering the entire European continent. The aim of the analysis is to identify key systematic differences between modelled and observed winds and physically explain their potential causes. For this purpose, model wind-speed biases are analysed alongside wind-direction biases. Principal component analysis is used to identify the principal diurnal, seasonal, and spatial patterns of wind speed and direction biases. We attempt to statistically explain these wind biases by the terrain elevation differences between models and the real world. For brevity, further in the work we refer to the terrain elevation difference between the model and the real-world as "elevation bias".

Studies have evaluated model performance at observation sites in complex terrain, with results heavily dependent on the model dataset (Cheynet et al., 2025). Fernández-González et al. (2018) finds that overall model performance in complex terrain increases with increasing model resolution. On the other hand, Solbakken et al. (2021) notes that increasing model resolution from 3 km to 1 km does not improve model performance metrics. However, it allows the model to better represent certain features of the wind climate. It has also been shown that model performance largely depends on geographical location, atmospheric stability, and wind conditions (Weigel et al., 2008). Therefore, studies that use data from only a few observation sites cannot be reliably used to draw general conclusions about model performance in complex terrain, even if they cover various terrain types (Jourdier, 2020).

Large scale comparisons are needed to draw general conclusions regarding model performance in different terrain types. Previous studies that performed large scale comparisons between modelled and observed wind speeds have identified that models generally show stronger biases in complex terrain. However, most of these large-scale studies focus on 10 m winds rather than on heights relevant to wind energy. Laurila et al. (2021) found that ERA5 underestimates the wind speed in complex terrain. Le Toumelin et al. (2024) observed that the same is true for the AROME model. Dörenkämper et al. (2020) has shown that both ERA5 and NEWA show more negative biases in locations with a higher terrain ruggedness index (RIX). The magnitude of bias in NEWA is smaller than that in ERA5. Hu et al. (2023) has shown that ERA5 generally shows stronger

biases in complex terrain. Das and Baidya Roy (2024) has compared models with wind speed observations in India. It was shown that multiple global reanalyses significantly overestimated wind speeds in complex terrain, while showing lower bias in simpler terrain.

Studies have found that wind speed biases in models vary with relative topographic position, i.e., valleys, plains, ridges, etc. Winstral et al. (2016) observed that COSMO models show a positive wind speed bias in valleys and a negative bias on upper slopes and ridges. Jiménez and Dudhia (2012) observed that the WRF model overestimates wind speeds in valleys and on plains, while underestimating them on hilltops. On the other hand, Gómez-Navarro et al. (2015) has shown that the WRF model generally overestimates wind speeds regardless of the relative topographic position.

Several possible causes of wind speed biases in synoptic scale and mesoscale models in complex terrain have been documented in the literature. Das and Baidya Roy (2024) attributes wind speed biases in complex terrain to model inability to resolve small scale terrain features. Cheynet et al. (2025) has explained model biases by their inability to resolve processes characteristic of complex terrain, such as recirculation and detached flows. Sandu et al. (2017) explains model biases by their errors in estimating orographic drag. Gómez-Navarro et al. (2015) attributes model biases in complex terrain to errors in estimating the effects of unresolved topography. The results of Jiménez and Dudhia (2012) suggest that wind speed biases in complex terrain are caused by errors in estimating the orographic speed-up. Moreover, they noted that a representativeness bias is present in the models. That is, the nearest model grid cell is not always the most representative of the actual observation site, and model bias can be significantly reduced by choosing a nearby grid cell instead. Choosing a grid staggering configuration (Collins et al., 2013) or interpolating winds between staggered and unstaggered grids can reduce representativeness bias.

Previous small scale studies (Carvalho et al., 2012; Schmidli et al., 2018; Saadatabadi et al., 2024) suggested that models perform better if the absolute difference between the model and the real-world terrain elevation is lower. We attempt to statistically explain the model biases related to the relative topographic position (hilltops, plains, valleys) of the observation site using elevation bias. Namely, due to terrain aggregation and smoothing in atmospheric models, terrain elevation in models is generally higher than real-world elevation for observation sites in valleys, and generally lower for observation sites on hilltops. It is further shown that the elevation bias has a statistical effect not only on the average wind speed bias but also on the wind direction bias and their diurnal and seasonal cycles. We perform two WRF model simulations using different terrain source data, and compare the modelled winds to show that model wind biases may be improved by adjusting terrain elevation.

Further in the work, we discuss which physical processes may be affected by differences in terrain elevation between models and the real world. However, regardless of the exact physical cause of the bias, the correlation between model bias and elevation bias suggests that improving terrain representation in atmospheric models could enhance their performance in modelling winds.

The work is structured as follows. Section 2.1 describes the seven model datasets and the observation dataset we used to investigate the wind biases in these model datasets. Section 2.2.1 describes the methods we used when interpolating model data to represent the observations and how we compared model data with the observations. Section 2.2.2 describes the principal component analysis (PCA) method and how we use it to assess model biases. Section 2.2.3 describes how we use terrain elevation data in explaining model biases. Section 2.2.4 describes the WRF numerical model setup that we used, and the numerical experiments that we performed to investigate model biases related to the terrain representation. Section 3.1 shows

that the elevation difference between the model and the real world can statistically explain part of wind speed and direction model biases, and the magnitude of the effect of the elevation difference depends on the time of day and the month of the year. Section 3.2 compares the winds of two WRF model simulations and identifies similar effects to the ones present when comparing model datasets with observations. Section 4 discusses the results and possible physical mechanisms of wind biases induced by the terrain discrepancies. The summary and conclusions are presented in Section 5.

2 Data and methods

2.1 Data

2.1.1 Model datasets

We compare the winds from seven different model datasets against the observations. This includes three global reanalyses: ERA5, MERRA2, and JRA-3Q, and four mesoscale model datasets derived from the ERA5 reanalysis through dynamical downscaling: NORA3, CERRA, NEWA, and EMD-EUR+. General information about model datasets is compiled in Table 1.

The ERA5 (Hersbach et al., 2020) is a global reanalysis developed by the European Centre for Medium-range Weather Forecasts (ECMWF). ERA5 is based on the Integrated Forecasting System (IFS) modelling system. It provides hourly data of the estimated state of the atmosphere with a resolution of 0.25° or ≈28 km in latitudinal direction. Such a resolution is one of the finest for global reanalyses. Since its release, the strong performance of ERA5 in modelling winds has been noted, with ERA5 outperforming the widely used reanalysis MERRA2 across all analysed aspects (Olauson, 2018).

Wind speed data at constant height levels in ERA5 are limited to only 10 m and 100 m, which is fewer levels than in mesoscale model datasets. Additionally, near-surface wind data from ERA5 are available on pressure levels or native model levels. Due to temporal variations in atmospheric pressure, the geopotential height of a constant pressure level can vary by more than 100 m and sometimes go below ground level, making the use of wind data on constant pressure levels inconsistent for our purposes. Access to the ERA5 data on native model levels is much slower. Therefore, for simplicity and consistency, we used only 10 m and 100 m winds from the ERA5.

Land variables from ERA5 have been further processed to a resolution of 0.1° (≈9 km in latitudinal direction) using a land surface model and published as the ERA5-Land dataset (Muñoz Sabater et al., 2021). However, the ERA5-Land dataset provides winds only at 10 m height, making it less suitable for studying winds at heights spanned by wind turbine rotors. Therefore, we decided not to use the ERA5-Land dataset for our work.

There is a known systematic problem with wind speeds at lower model levels in the ERA5. Jourdier (2020) shows that 10 m, 100 m, and lower model level wind speeds in ERA5 experience a sudden artificial drop in value at the start of the assimilation window at 9:00 UTC. Wind speeds slowly recover by the end of the assimilation window at 21:00 UTC. Downscaled mesoscale products can potentially inherit the issue, as they use initial and boundary conditions from ERA5. While this issue is not the main focus of the study, it was briefly examined whether or not it is inherited in downscaled mesoscale model datasets.

The NEWA (Hahmann et al., 2020) is a mesoscale atmospheric dataset based on the open-source Weather Research and Forecasting (WRF) modelling system. It has been developed as an ERANET+ project. The model is run on 27–9–3 km nested domains, where the outer low-resolution domain provides boundary conditions for the inner domains. The outer 27 km domain extends beyond Europe. A total of ten overlapping, nested 9–3 km domains are used to cover the region of interest. The final dataset with 3 km grid spacing is obtained by spatially stitching together results from all ten 3 km domains and covers the European Union, the United Kingdom, Norway, Switzerland, the Balkans, Turkey, and nearby offshore areas. NEWA provides output every 30 minutes. However, to match other model datasets, we used only data for full hours.

The WRF model used in NEWA is reinitialised every 7 days and run for 8 days. The first 24 hours are used as a spin-up time and, therefore, discarded. Meanwhile, the remaining 7 days contribute to the dataset. Spectral nudging to the ERA5 reanalysis is performed only in the outer domain, above the planetary boundary layer. The NEWA simulations use boundary conditions and nudging data derived from 6-hourly ERA5 data.

The EMD-WRF EUROPE+ (or EMD-EUR+ for short) (Svenningsen et al., 2020) mesoscale dataset was developed by EMD International A/S. It is based on the WRF modelling system and runs on a single large 3 km domain without nesting, covering the entire Europe. The model is reinitialised every 7 days and requires a 24-hour spin-up. Spectral nudging to the ERA5 reanalysis is performed above the planetary boundary layer. The EMD-EUR+ simulations use boundary conditions and nudging data derived from hourly ERA5 data.

The CERRA (Ridal et al., 2024) is a mesoscale reanalysis developed by ECMWF. It is based on the HARMONIE-ALADIN modelling system configuration. It has a coarser grid spacing (5.5 km) than the other mesoscale model datasets used. The CERRA dataset provides atmospheric analysis using 3D-VAR data assimilation every 3 hours. Additionally, after each data assimilation, a 6-hour forecast with hourly output is provided. The forecast is extended to 30 hours twice a day: at 0:00 and 12:00 UTC. Dataset documentation leaves the user free to choose which data to use when constructing hourly series (Jourdier et al., 2023). In this work, hourly series were constructed using only forecasts with 1, 2, and 3-hour lead times.

The NORA3 (Haakenstad et al., 2021) is a mesoscale model developed by the Norwegian Meteorological Institute. It is based on the HARMONIE-AROME modelling system configuration and has a spatial grid spacing of 3 km. While other mesoscale datasets used in this work cover all of Europe, NORA3 covers only the northern part. NORA3 is run as a hindcast, which is initialised every 6 hours using data from ERA5 reanalysis. The hindcast is run for 9 hours, using the first 3 hours as a spin-up time. NORA3 assimilates only temperature and air humidity observations. The upper-air temperature, humidity, and wind components are interpolated from the host reanalysis ERA5.

The MERRA2 (Gelaro et al., 2017) is a global reanalysis developed by the National Aeronautics and Space Administration (NASA). MERRA2 is based on the Goddard Earth Observing System (GEOS) modelling system. It provides hourly atmospheric data with $0.5^\circ \times 0.625^\circ$ resolution. For similar reasons as with ERA5, we used winds from MERRA2 only at constant-height levels of 10 m and 50 m AGL.

The JRA-3Q (Kosaka et al., 2024) is a global reanalysis developed by the Japanese Meteorological Agency (JMA) and is based on the JMA global Numerical Weather Prediction (NWP) system. The spatial resolution of the JRA-3Q is 0.375° or $\approx$42 km in latitudinal direction. JRA-3Q provides atmospheric analysis every 6 hours. Additionally, forecast data with 3 and

Table 1. Model datasets and their characteristics.

Model dataset	Horizontal grid spacing	Height levels used, m	Temporal resolution, h	Vertical levels	Region	Modelling system
ERA5	0.25°	10, 100	1	137	Global	IFS
MERRA2	0.5° × 0.625°	10, 50	1	72	Global	GEOS
JRA-3Q	0.375°	8 native model levels	3	100	Global	JMA global NWP
NORA3	3 km	50, 100, 250	1	65	Northern Europe	HARMONIE
NEWA	3 km	50, 75, 100, 150, 200	0.5 (1 used)	61	Europe	WRF
CERRA	5.5 km	50, 75, 100, 150, 200	1	106	Europe	HARMONIE
EMD-EUR+	3 km	50, 75, 100, 150, 200	1	50	Europe	WRF

6-hour lead times are provided. For consistency, we constructed 3-hourly time series using only forecast data. The only height level at which JRA-3Q outputs wind speed is 10 m. However, the JRA-3Q reanalysis provides data on 8 native model levels in the lowest approximately 300 m of the atmosphere. Therefore, wind data together with geopotential data on native model levels were used. As the model is run at hybrid vertical levels, the exact AGL heights of these levels vary across model time steps and geographic locations.

Although JRA-3Q was released in 2018, studies assessing its performance in wind energy applications are limited to China (Mi and Liu, 2024). This is likely due to the fact that it is only licensed for non-commercial use. However, its predecessor, JRA-55, has been widely validated for wind energy applications, e.g., Miao et al. (2020); Das and Baidya Roy (2024). While JRA-55 has generally shown poorer performance than ERA5 in modelling wind speeds at heights spanned by turbine rotors (Ramon et al., 2019), JRA-55 uses significantly different approaches when modelling winds than other global reanalyses. Firstly, based on the results of Carvalho (2019); Belmonte Rivas and Stoffelen (2019), JRA-55 is the only global reanalysis that did not show a systematic clockwise wind-direction bias over the Northern Hemisphere oceans, a long-known issue in atmospheric models (Hollingsworth, 1994). Secondly, as reported by Torralba et al. (2017), JRA-55 is the only global reanalysis that assimilates surface wind speed observations over land, which allows it to show good performance in modelling interannual wind speed variations and trends on decadal timescales.

2.1.2 Observation dataset

We evaluate the model biases by comparing the model output with the EMD's internal database of wind-mast data. The database contains time series of wind observations from wind masts and remote sensing devices since the 1980s at different heights up to 300 m above ground level (AGL). The database contains data from over 1000 observation locations around the globe.

This work focuses on Europe. Our results show that model biases vary across parts of Europe based on geographic and terrain characteristics. To simplify the discussion of our results, we introduce a descriptive terminology for regions of Europe,

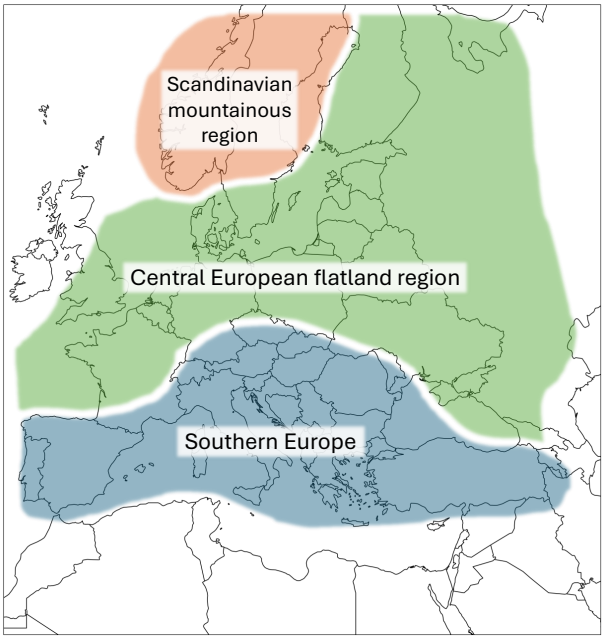

Figure 1. Map showing the geographical regions used for discussion of spatial patterns in the results.

as shown in Figure 1. Due to data confidentiality, the exact locations of observation sites cannot be disclosed. The region of interest and the approximate number of observation sites per country are shown in Figure 2 (a). It should be noted that offshore measurement campaigns were also included in the work, which are counted towards the country in which exclusive economic zone it is located. In subsequent figures, results are aggregated into hexagonal bins to present spatial patterns of biases while obscuring the exact locations of observations. Figure 2 (b) shows installed wind power capacity by country as of 2025. Much of the observational data used in our work come from countries that are currently only expanding their wind power capacity, namely Poland, Lithuania, Romania, and Turkey. It highlights the importance of our work for the future development of wind energy infrastructure.

We use observational datasets spanning more than a year, as this work focuses on analysing diurnal and seasonal patterns of model biases. We include observations at heights of 50 m AGL or higher. In total, data from 508 observation locations were included in the work; 504 of these locations provided both wind speed and direction data, while four provided only the wind speed data. Figure 3 (a) shows the distribution of observation heights. In most observation locations, the winds were measured at heights between 50 m and 125 m. However, the highest-altitude observation campaigns were conducted at heights of nearly 300 m. 289 observation campaigns lasted between 1 and 2 years, 118 lasted between 2 and 3 years, 49 lasted between 3 and 4 years, and 52 lasted 4 years or more. Most of the campaigns cover some subperiod between 2000 and 2024. Not all observation campaigns were compared against all model datasets due to differences in their spatial and temporal coverage.

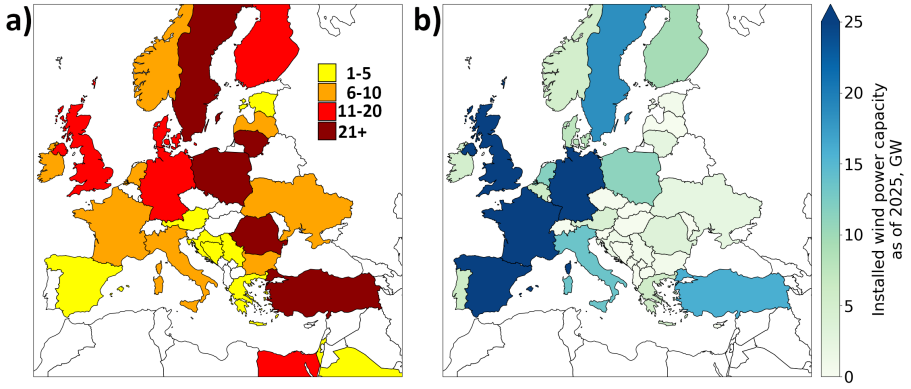

Figure 2. (a) region of interest and approximate number of observation locations included in the study by country. Countries are categorised by the number of observation locations included in the study. No observation data were used from countries coloured in white (including microstates). (b) map showing installed wind power capacity as of 2025, based on Wikipedia (2026).

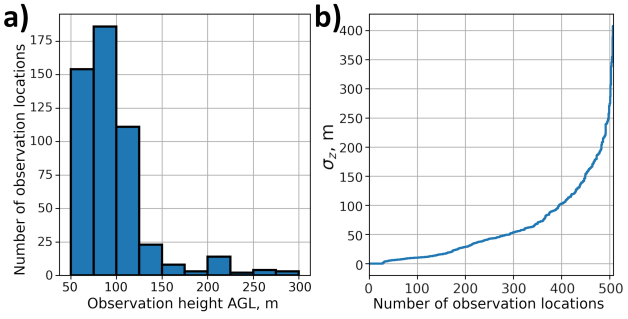

Figure 3. Information about the observational dataset used in the work. (a) histogram of observation heights. The left side of the bin is inclusive, and the right side is exclusive. (b) ranked values of σ_z in metres for observation locations.

We quantify the local complexity of the terrain around the observation locations using the standard deviation of elevation, denoted as σ_z, within a 5 km radius of each location. The value is derived using the gridded elevation data from Copernicus 235 $3''$ (3 arcsecond) DEM (digital elevation model). Figure 3 (b) shows the distribution of σ_z between the observation locations. The entire range of σ_z values is well represented, including offshore locations with $\sigma_z = 0$ and highly complex terrain with σ_z exceeding 200 m. Figure 4 shows the spatial distribution of values of σ_z. In the central European flatland region, the values of σ_z are smaller, indicating locally simpler terrain near the observation sites. Meanwhile, in the Scandinavian mountainous region and Southern Europe, the values of σ_z are higher, indicating locally more complex terrain near the observation sites.

240 Figure 5 shows the distribution of elevation differences between the model and the real-world at observation locations in all seven model datasets. The majority of the observation sites are located above the model grid-cell terrain elevation. For global reanalyses and NORA3, only approximately 28% of the observation locations are located below the model grid-cell terrain elevation. For other mesoscale model datasets, the same is true for only approximately 22% of observation locations. The

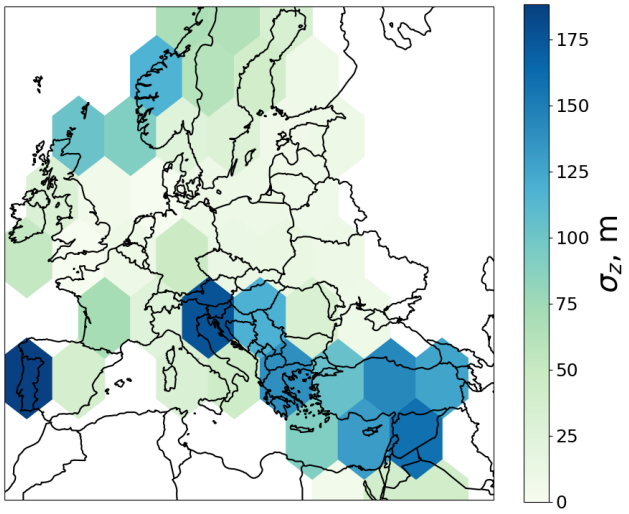

Figure 4. Spatial distribution of values of σ_z at observation locations. Colour of a hexagonal bin represents the average value of σ_z between all masts located inside the bin.

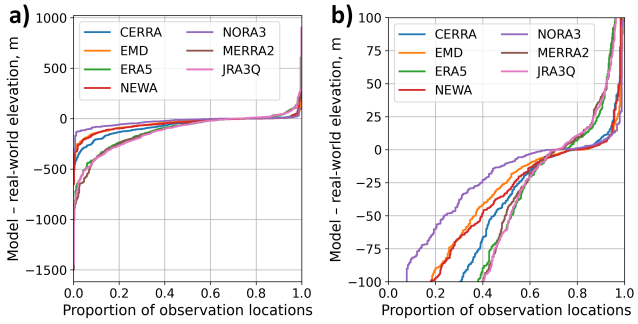

Figure 5. Empirical cumulative distribution functions of the terrain elevation difference at observation locations between model and real-world for different model datasets. (b) zoomed version of (a).

magnitude of the negative differences between the model and the real-world elevation is generally greater than the magnitude of the positive ones. Depending on the model dataset, between 8% and 41% of the observation sites are located more than 100 m above the model terrain elevation. On the other hand, across all models, fewer than 5% of observation sites are located more than 100 m below the model terrain elevation.

If observation sites were randomly distributed within a grid cell, the distribution of differences between the model and the real-world terrain would be expected to be symmetric. In areas with larger σ_z these differences would be greater in magnitude. However, in the case of the EMD mast database, the distribution is skewed towards locations where the real-world terrain elevation exceeds that of the model. This is consistent with the idea that measurements are carried out at the potential locations of wind farms, i.e., locally windier locations expected to be at higher elevations. Moreover, in complex terrain regions, where

σ_z is higher, observation locations are generally positioned higher above the model terrain elevation than in simple terrain regions. This must be taken into account when interpreting the results of this work, as wind speeds are generally higher at higher elevations locally.

The magnitude of the elevation biases appears to be related to the model grid spacing. Models with finer grid spacing tend to show smaller differences between the model and real-world terrain elevation. All global reanalyses show a similar magnitude of the elevation biases. The magnitude of the differences is smaller in CERRA (5.5 km grid spacing) and even smaller in EMD-EUR+ and NEWA (3 km grid spacing). Although NORA3 has the same resolution as EMD-EUR+ and NEWA, it shows a significantly smaller magnitude of elevation biases, likely because it covers only Northern Europe.

Observation data in the EMD's mast database were previously quality controlled by the individual data providers using in-house methods, making them suitable for wind resource assessment. Handling of raw wind measurement data follows the MEASNET (2022) guidelines as closely as possible. However, the exact quality control procedure cannot be disclosed. In addition, the data were further checked for obvious errors. Obviously erroneous values were excluded: wind direction outside the range $[0°, 360°)$ and wind speed outside the range $[0, 100]$ m s^{-1}. Periods with more than three consecutive 0 m s^{-1} wind speed measurements were excluded, as those are likely to be improperly labelled missing observations.

The observation frequency is either 1 hour or 10 minutes, depending on the specific observation campaign. On the other hand, model output is provided hourly or every three hours. Using only 10-minute data points corresponding to model output might introduce excessive aliasing relative to the model resolution. Averaging data to 3-hour intervals would unnecessarily reduce the number of available data points. Therefore, the observed values were averaged at the hourly resolution for comparison.

2.2 Methods

2.2.1 Comparing models with observations

The bias metric was used to quantify the differences between the models and the observations. Bias is defined as the average difference between modelled and observed values. For wind energy applications, accurate predictions of wind speed are generally considered more relevant than those of wind direction. However, in this work, a holistic analysis is performed, and wind speed biases are considered alongside wind direction biases to seek a potential physical explanation.

A positive wind speed bias indicates that the model overestimates wind speed, while a negative bias indicates that it underestimates it. As the wind direction is a circular variable, the bias metric cannot be applied directly. To address this issue, 360° was added to the observed wind direction when necessary to ensure that the difference between the modelled and observed values is between -180° and 180°. Positive wind direction bias indicates that modelled wind direction is more clockwise than the observed one, while negative wind direction bias indicates that modelled wind direction is more anti-clockwise than the observed one.

We analyse spatial, diurnal, and seasonal patterns of wind speed and direction biases. For this purpose, each observation location was described by a 24×12 variable – mean differences between modelled and observed winds aggregated by the hour of day and month of the year. In the case of the JRA-3Q reanalysis, 8×12 variables were used, since it provides output every

3 hours. Principal component analysis was applied to identify temporal and spatial patterns in model biases (as described in Section 2.2.2).

When comparing models with observations, it should be considered how to spatially collocate model output and observations to avoid representativeness errors. Our analysis was designed to examine the connection between model biases and differences between the model's terrain elevation and the actual terrain elevation. Therefore, when comparing the model with the observations, the model data from the closest horizontal grid cell to the observation campaign location were used without considering any land properties, which is the simplest approach. Our findings can be used in future studies to inform decisions about which spatial collocation method to use between model output and observations.

Models generally output winds at different heights than observations. For comparison purposes, the model output was vertically interpolated to the observation height. We use a similar vertical interpolation procedure as in Solbrekke et al. (2021) for both wind speed and direction. Interpolation was performed independently at each model time step. Wind speed was interpolated using the power law relation. The power law exponent is calculated using wind speed data u_i at the time-step i at two heights z_1 and z_2:

$$\alpha = \frac{\ln u_i(z_1) - \ln u_i(z_2)}{\ln z_1 - \ln z_2}. \tag{1}$$

Wind speed at the target height is calculated:

$$u_i(z) = u_i(z_2) \cdot \left(\frac{z}{z_2}\right)^{\alpha}. \tag{2}$$

The wind direction was interpolated considering it as a unit vector and linearly interpolating its components between two heights. If observations were performed between two adjacent model levels, the modelled winds were interpolated to the observation height between these two levels. If observations were performed above the highest model height level, the modelled winds were extrapolated to the observation height from the two highest model levels.

When interpolating wind speeds between different heights, it should be considered that models use different methodologies to calculate 10 m winds and winds at higher atmospheric levels. The 10 m wind speed is usually a diagnostic variable and is not modelled directly. For example, in ECMWF forecasts 10 m winds are interpolated from the lowest model level (Owens and Hewson, 2018). The roughness length of the short grass (0.03 m), which matches recommended WMO observation conditions, is used instead of the actual roughness length of the model grid cell. Therefore, the use of modelled 10 m winds should be generally avoided when interpolating winds to higher levels.

To interpolate the wind speed at any given height, wind speed values at at least two different heights are needed. All mesoscale models provide wind data at multiple heights, covering the entire range of wind mast observation heights. Therefore, using 10 m wind speeds can be avoided. However, ERA5 provides wind speeds only at 10 m and 100 m heights, and MERRA2 provides wind speeds at 10 m and 50 m heights, which makes using 10 m winds unavoidable. Potential effects of using these winds on the identified model biases in ERA5 or MERRA2 are described in Section 4.

2.2.2 Principal component analysis

Due to high maintenance costs, observation campaigns are kept as short as possible, typically lasting just over a year and rarely exceeding three years. Such duration is usually sufficient for decision-making regarding wind farm installation. However, it might not be long enough to identify systematic biases in models when comparing model output to observations from an individual mast. In climate sciences, a 30-year period is usually used to define systematic weather patterns, i.e., climate normals.

The motivation for the analysis is based on the assumption that the model output differs from real-world conditions because certain physical phenomena in the model are improperly represented. Each improperly represented phenomenon contributes to the total model bias by amplifying or cancelling the effects of other phenomena. To improve future atmospheric models, it is important to decompose the total model bias into contributions from each improperly represented physical phenomenon and to fix them individually.

Principal component analysis (PCA) is a statistical technique commonly used to identify the main patterns in large, noisy multidimensional datasets (Jolliffe, 2002). The method considers correlations among the original variables and derives new variables (principal components, or PCs for short) that are linear combinations of them. New variables are created so that the first principal component (PC1) explains as much of the data's variance as possible. Each subsequent principal component (i.e., PC2, PC3, ...) is uncorrelated with all previous PCs and explains the maximum possible remaining variance in the data. Therefore, only a few of these new variables are needed to describe the main modes of variability in the dataset. The values of the principal components describe how prominent each mode is for a particular data point. The principal components can be interpreted as unit vectors that describe the axes along which the data points differ from the dataset's mean. The linear coefficients of each principal component are known as "loadings", which can be interpreted as the vector components. Previous work has successfully applied PCA to identify seasonal patterns in meteorological variables, such as wind direction (Pogumirskis et al., 2021) and temperature and precipitation (Bethere et al., 2017).

PCA is a suitable technique for tackling the data issues described above. Firstly, it allowed us to examine biases at a single observation location in the context of other locations. It allowed us to identify systematic biases even when only one year of observational data was available at that location. Secondly, PCA enabled the identification of the principal orthogonal modes of the dataset. These modes are determined purely by statistical considerations. Further numerical and statistical analyses were performed to consider the diurnal, seasonal, and spatial patterns of bias associated with these modes. Based on the analysis, we linked these statistical modes to physical phenomena.

The PCA was performed separately for each model dataset and for wind speed and wind direction. Bias at a particular location was described by $24\times12=288$ ($8\times12=96$ in the case of JRA-3Q) parameters – bias values for the particular month or hour of the day. Let us denote the bias table at a location i as X_i. By taking the average value of the tables between all observation locations, we can compute the average bias of the model – $\bar{X}$, which is also a 24×12 table. These bias tables at each observation location can be viewed as points in a 288-dimensional space centred at zero. Initially, data are defined in the canonical orthonormal basis where each of 288 unit vectors $\boldsymbol{e_j}$ has 1 at position j and 0 at the other 287 positions. By applying PCA, a coordinate transformation is performed that centres the coordinate axes at the mean bias $\bar{X}$ and rotates the unit

vectors $\boldsymbol{e_j}$ so that they align with the principal components, as described before. After the coordinate transformation, model biases can be expressed in new coordinates where principal components (i.e., PC1, PC2, ...) are the unit vectors, "loadings" are components of these vectors, and $y_{i,j}$ are the coordinates of points in the new coordinate system (known as "values" of principal components), where i denotes the observation location and j denotes the rank of the principal component. The bias table at a particular location i in the new coordinates can be expressed as:

$$X_i = \bar{X} + y_{i,1} \cdot PC1 + y_{i,2} \cdot PC2 + \ldots + y_{i,288} \cdot PC288. \tag{3}$$

After the PCA transformation, the principal component loadings describe modes of variation in the temporal patterns of biases. Meanwhile, the principal component values correspond to observation locations and enable the identification of spatial patterns of bias. Since principal components are ranked by the variance they explain, the first few components capture the main modes of model bias. In further analysis, we attempted to explain the most significant principal components using physical phenomena that are poorly represented in models.

When performing PCA, an important question is whether normalisation of the initial variables is necessary. The choice of normalisation can have a significant impact on the results. During the analysis, we attempted to link the principal components with elevation bias. We wanted to quantify the magnitude of the effects of elevation bias on other biases, depending on the time of day and year. Therefore, we chose the normalisation process that would keep the original variances and units. As wind speed and direction have different units, PCA was applied to each of them separately. Therefore, the comparison between biases was performed in absolute values, i.e., in $\mathrm{m\,s^{-1}}$ for the wind speed bias and in degrees for the wind direction bias.

As PCA, at its core, is an eigenvector problem, the scaling of eigenvectors (principal components) and the new coordinates of variables (values of principal components) are arbitrary. The standard practice in normalisation is to set the principal components to unit vectors, so that the sum of the squares of the loadings for each principal component is 1 and the loadings are dimensionless. In that case, the values of the principal components are in $\mathrm{m\,s^{-1}}$ or degrees. However, such a normalisation approach is difficult to interpret physically, since the magnitude of those values is affected by the number of data dimensions, as follows from the fact that the sum of the squares of the loadings equals 1.

To provide a clearer physical interpretation, the values of each principal component were divided by the standard deviation of that component across observation locations, thereby making them dimensionless. On the other hand, loadings were multiplied by the standard deviation of the corresponding principal component across observation locations, thereby setting their units to $\mathrm{m\,s^{-1}}$ or degrees. The scaling procedure can be expressed mathematically as:

$$\widetilde{y}_{i,j} = y_{i,j} / \sigma_j, \tag{4}$$

$$\widetilde{PCj} = PCj \cdot \sigma_j, \tag{5}$$

where $y_{i,j}$ is the value of the principal component j at the i-th observation location before scaling, and $\widetilde{y}_{i,j}$ is the value after scaling. PCj are loadings of the principal component j before scaling and $\widetilde{PCj}$ are loadings of the principal component j after scaling. σ_j is the standard deviation of the values of the principal component j:

$$\sigma_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(y_{i,j} - \bar{y}_j\right)^2}. \tag{6}$$

After normalisation, the principal component's value indicates the prominence of the bias associated with that component at a particular location. The standard deviation is chosen as an interpretable measure, since in the case of the normal distribution 99.7 % of the data lie within 3 standard deviations of the mean. On the other hand, loadings describe a characteristic magnitude of the bias in the dataset – the magnitude of the bias at a location where the value of a principal component is 1 standard deviation from the mean.

2.2.3 Terrain influence on biases

Previous studies, i.e., Dörenkämper et al. (2020); Hu et al. (2023); Das and Baidya Roy (2024), have identified that synoptic scale and mesoscale models show stronger wind speed biases in complex terrain. This work attempts to investigate potential physical explanations for the negative wind speed bias on hilltops and the positive wind speed bias in valleys. Our hypothesis is that the terrain elevation mismatch between the model and the real world has a significant statistical effect on the wind speed and direction biases in atmospheric models.

The terrain in models is spatially aggregated and a smoothed version of the real-world terrain. For points that are locally at a higher elevation, i.e., hilltops, terrain smoothing generally decreases elevation. It not only reduces the hilltop's absolute elevation but also lowers the hill's height relative to its surroundings. Meanwhile, for points that are locally at a lower elevation (i.e., in valleys), smoothing generally increases the elevation. It not only increases the valley's absolute elevation but also decreases its depth relative to its surroundings. In addition, smoothing the terrain generally reduces the magnitude of terrain slopes and the variability of the resolved terrain, thereby reducing orographic drag.

We use elevation bias as a statistical explanatory variable for model biases. Based on physical considerations, the elevation bias can potentially explain model biases. For example, if a model underestimates the hill's height, it would also underestimate the orographic speed-up caused by the hill. Moreover, the elevation bias is a variable that can be easily modified in the input data or accounted for during post-processing of model output to potentially mitigate model biases. The elevation bias does not directly describe changes in slopes or resolved orographic drag. However, a larger absolute elevation bias would indicate stronger local changes in the terrain due to smoothing, leading to a greater reduction in slopes and orographic drag from reduced terrain variability.

In this work, we attempted to statistically explain model wind biases by the terrain elevation bias. We wanted to ensure that the statistical relationship between wind bias and elevation bias appears in the data without a priori imposing it. Moreover, we wanted to ensure that the statistical explanatory power of the elevation bias is stronger than the established result that models generally show stronger wind biases in complex terrain, i.e., at higher elevations. Therefore, we analysed how well model bias can be statistically explained by the difference between the model and real-world elevation, compared with the case where only the real-world elevation is considered. For such a purpose, we performed stepwise linear regression. In the first case, only

the real-world elevation was used to linearly predict values of the principal component of model bias. In the second case, both the model and the real-world elevation were considered as predictors. The F-test was used to evaluate whether the increase in correlation is significant when both the model and the real-world elevation are used to explain the model biases, versus when only the real-world elevation is used as the explanatory variable. In all cases, the regression coefficients for the real-world and model terrain had approximately the same magnitude and opposite signs, indicating that bias is explained by their difference. In further work, the linear combination of model elevation and real-world elevation, which predicts model biases using the linear regression, is referred to as the "weighted difference between model and real-world elevation".

Matiu et al. (2024) has shown that terrain representation in regional climate models has a significant effect on local patterns of the modelled precipitation and temperature. Model biases can arise not only from smoothing and scale mismatches between model and real-world terrain, but also, in some regions, from inaccuracies in the underlying terrain datasets themselves. For example, Gelaro et al. (2017) reported that the MERRA reanalysis used an outdated terrain source compared to the MERRA2 reanalysis, which contained large terrain biases of up to 600 m in Greenland. These errors in terrain representation have significantly affected the modelled surface mass balance in Greenland between these two reanalyses.

Digital terrain models (DTM) are datasets that provide spatially gridded terrain elevation data. On the other hand, digital elevation models (DEM), which are used in this study, also consider other objects on the Earth's surface, such as trees and buildings, in their elevation data. All mesoscale model datasets analysed in this work used the GMTED2010 elevation dataset (Danielson and Gesch, 2011), which compiles terrain data from 11 DEM sources. ERA5 uses its own elevation dataset compiled from 5 different DEM sources (ECMWF, 2016). MERRA2 uses the USGS GTOPO30 elevation dataset, which is compiled from 8 different source DEMs (Bosilovich et al., 2015). JRA-3Q uses the elevation dataset inherited from the JMA operational NWP system (December 2018 configuration). However, the public JRA-3Q documentation reviewed here does not explicitly identify the original DEM from which the model orography was derived.

For the region of interest, both ERA5 and GMTED2010 use SRTM DEM data south of 60°N, which was collected in 2000. North of 60°N ERA5 uses GLOBE DEM data, which was made publicly available in 1999, while GMTED2010 uses DTED1 DEM data, which was collected in 2000. For the region of interest, MERRA2 uses DTED1 DEM data and DCW DEM data, which were collected in 1992. While the quality assessment of the elevation datasets themselves is out of the scope of this study, we acknowledge that wind biases can be influenced by errors in terrain datasets. Therefore, as a source of real-world terrain, we use data from Copernicus $3''$ ($\approx$90 m in the latitudinal direction) DEMs. Data in the Copernicus DEM have been collected between 2011 and 2015 and have shown higher accuracy than older DEMs (Meadows et al., 2024). The absolute vertical accuracy of GMTED2010 at 90% confidence level is 10 m (Danielson and Gesch, 2011), while for Copernicus DEM it is 4 m (Fahrland, 2022).

It should be kept in mind that both wind and elevation data from global reanalyses are regridded to the orthogonal latitude-longitude grid.

This work focuses on identifying local model biases that are directly influenced by terrain differences. In other words, model bias at a particular grid-point is present because the elevation of a grid-cell is different from the terrain elevation at the observation location. These effects can be identified by considering the correlation of model bias and elevation bias.

However, mountains can influence the wind field tens to hundreds of kilometres downstream (Draxl et al., 2021). Therefore, local differences in terrain elevation between the models and the real world can theoretically cause wind biases further away. Identifying these biases is outside the scope of this work.

2.2.4 WRF modelling

The results of the comparison between models and observations suggested that the terrain height difference between models and the real world has a significant statistical effect on the differences between the modelled and observed winds. To test whether the modelled winds can be adjusted in a controlled manner by varying the model terrain elevation, we performed two one-year WRF model simulations using different terrain source datasets, with otherwise identical setups. One run used GMTED2010 terrain input data at 30″ (≈900 m in latitudinal direction) resolution, while the other used 5′ (5 arcminute) (≈9 km in latitudinal direction) resolution. GMTED2010 is the default terrain dataset included with the WRF model installation, available at different resolutions. The motivation for using these two terrain datasets to test the hypothesis lies in how the terrain is represented at different resolutions.

In the real world, terrain height is a continuous variable, i.e., its resolution is infinitely fine. In the Copernicus dataset used in this work, the terrain is aggregated on a 3″ grid (≈90 m in latitudinal direction). In the WRF model run that uses 30″ terrain input data, the terrain is aggregated to a 3 km grid because the model grid is coarser than the input data. Meanwhile, in the WRF model run using 5′ input data, the terrain is smoothed because the input terrain resolution is coarser than the model resolution.

When observations are compared with the 3 km model output, the flow over the real-world terrain is compared with the flow over the 3 km terrain, which is a smoothed version of the real-world terrain. Similarly, when modelled winds over the 30″ terrain are compared with modelled winds over the 5′ terrain, the flow over a detailed terrain is compared with the flow over a smoothed version of the same terrain. An illustrative example of a topographic profile in the Copernicus dataset, as well as in both WRF model runs, is shown in Figure 6. It can be seen that in the Copernicus 3″ DEM, hills are generally higher and valleys are generally deeper than in the WRF model run using 30″ terrain input. Similarly, in the WRF model run that uses 30″ terrain input, hills are generally higher and valleys are generally deeper than in the WRF model run that uses 5′ terrain input.

Models and observations generally show differences due to numerous factors. On the other hand, the two model runs used in this work differ only in their terrain. Therefore, their differences are solely due to differences in terrain representation. The model run that uses 30″ terrain is used as an analogy for the flow in the real world, i.e., with fully resolved terrain. On the other hand, the model run that uses 5′ terrain serves as an analogy for the flow in the model, i.e., with aggregated and smoothed terrain. It is hypothesised that terrain-smoothing differences between these two model runs would induce differences in modelled winds similar to those between a model and observations. When comparing models with observations, observed winds were subtracted from the modelled winds. The order in which model outputs were subtracted was chosen to represent the same effects as when comparing models and observations. The output of the 30″ terrain model run was subtracted from the output of the 5′ terrain model run. It should be noted that the WRF setup in our experiments did not use any subgrid terrain properties, e.g., subgrid terrain variance, as inputs. Even if we used a WRF model setup that considers subgrid terrain properties,

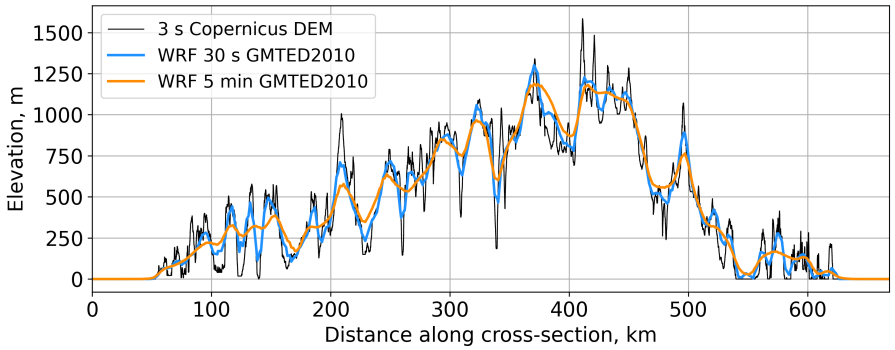

Figure 6. An example of a topographic profile in Copernicus 3″ DEM and both WRF model runs used in this work. The exact location of the line is shown on the map in Figure 7. Terrain elevation in metres is shown on the y-axis, and distance along the line in kilometres is shown on the x-axis.

these properties are not dynamically derived from the input terrain dataset but are pre-calculated on a grid. Therefore, changing only the terrain input dataset in the WRF model run would not directly affect subgrid terrain properties.

Both WRF model runs were performed on an 810×1620 km domain covering Sweden, consisting of 270×540 grid cells. To avoid domain boundary effects of the low resolution boundary conditions, 10 outer grid cells of both domains were excluded from the comparison. Model domains and differences in their terrain are shown in Figure 7. The average difference between elevations in both model runs is 0.4 m, the maximum absolute difference is 490.2 m, and the standard deviation of differences is 35.6 m. Both model runs used the same WRF model setup described in Table A1. Hourly winds at 100 m AGL were used for comparison, as they are close to the typical observation height in the EMD mast database. The analysis was performed using PCA in a similar manner as that applied when comparing models with observations: each of the model grid-points was described by 24×12 variables – average differences in mean wind speed or direction between model runs for the particular hour of the day or month of the year.

Sweden was chosen as the simulated region for several reasons. First, choosing all of Europe as a domain would significantly increase the computational time. Second, Sweden has a diverse terrain — complex terrain in the north and simple terrain in the south. Moreover, large areas of open sea are present. Third, the EMD mast database has high observation coverage in Sweden, which allows for comparison of these model simulations with observations in future work. Due to the higher availability of observation data, 2011 was chosen as the year for both runs.

3 Results

3.1 Comparison between models and observations

The results of the principal component analysis suggest that all seven model datasets exhibit similar principal spatial and temporal patterns in wind speed and direction biases. Temporal patterns of biases in the Central European flatland region differ

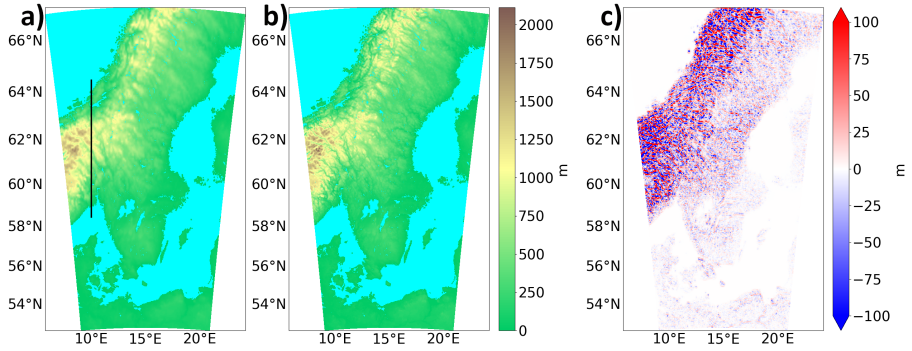

Figure 7. Terrain elevation height in metres and domain extent for both WRF model runs (including the buffer zone). (a) terrain using $5'$ resolution input data, (b) terrain using $30''$ resolution input data and (c) their difference. The line of the topographic profile from Figure 6 is included.

significantly from those in mountainous regions — Southern Europe and the Scandinavian mountainous region. PC1 values for wind speed bias and PC2 values for wind direction bias show similar spatial patterns across all seven model datasets. However, for brevity, detailed results only for the ERA5 reanalysis are presented in the main part of the paper. For other model datasets, mean biases and loadings of principal components are included in the Appendix B.

The temporal patterns of the spatially averaged wind speed bias vary across datasets. In ERA5 (Figure 8, (b)) and MERRA2, the average wind speed bias is more positive during the day and more negative during the night. In NORA3, EMD-EUR+, CERRA, NEWA, and JRA-3Q, the average wind speed bias is more positive during the winter and more negative during the summer. In NORA3, EMD-EUR+, CERRA, and JRA-3Q, the bias is more negative during the day and more positive during sunrise and sunset. In NEWA, the wind speed bias is more negative during the day, and its magnitude is significantly higher than in other model datasets. However, NEWA shows a more positive bias only at sunset, not at sunrise.

The principal components describe how the temporal patterns of model biases in different observation locations differ from the temporal pattern of the spatially averaged model bias. For example, the temporal patterns of spatially averaged wind speed bias in ERA5 are shown in (Figure 8, (b)). The loadings of PC1 of the wind speed bias in ERA5 are shown in Figure 9, (a). The temporal patterns of wind speed bias at a particular location can be calculated by combining the spatially averaged temporal pattern of the model wind speed bias with the sum of the values of all principal components at that location, each multiplied by the corresponding principal component. PCA is performed with the goal of linking each principal component to some improperly represented atmospheric phenomena. By setting all principal component values except one to zero, it is possible to analyse how model biases across different regions are affected by the improper representation of a particular atmospheric phenomenon. Further in the work, PC1 of the wind speed bias (explains 82.5 % of variance in the wind speed bias in the ERA5 reanalysis) and PC1 and PC2 of the wind direction bias (explain 61.1 % and 5.7 % of variance in the wind direction bias in the ERA5 reanalysis, respectively) are analysed. For other model datasets, PC1 of the wind speed bias explains between 61.6 %

and 81.3 % of the variance, PC1 of the wind direction bias explains between 52.4 % and 68.3 % of the variance, and PC2 of the wind direction bias explains between 3.4 % and 6.3 % of the variance.

Figure 10 shows spatial patterns of PC1 values of the wind speed bias. Higher PC1 values are observed at observation sites in the Central European flatland region. Meanwhile, more negative PC1 values are present at observation sites in mountainous regions — namely, the Scandinavian mountainous region and Southern Europe. It indicates that the temporal patterns of the wind speed biases generally differ between these regions. Characteristic temporal patterns of biases for regions with positive (negative) values of PC1 can be calculated by taking PC1 multiplied by the standard deviation of PC1 values between all observation locations and adding (subtracting) it to (from) the spatially averaged table of biases.

In mountainous regions, the PC1 values for wind speed bias are more negative. Therefore, if we subtract PC1 of the wind speed bias multiplied by the standard deviation of values of PC1 between all observation sites from the average bias table, we obtain a characteristic bias table for mountainous regions (Figure 8, (a)). In mountainous regions, all model datasets show an overall more negative wind-speed bias. The magnitude of the negative wind speed bias is greater during the night and lower during the day.

In the Central European flatland region, the PC1 values for wind speed bias are more positive. Therefore, if we add PC1 of the wind speed bias multiplied by the standard deviation of the PC1 values between all observation sites to the average bias table, we obtain a characteristic bias table for the Central European flatland region (Figure 8, (c)). It can be seen that in the Central European flatland region, the bias is more positive in winter than in summer and more negative during the day than at night.

Temporal patterns of spatially averaged wind-direction bias vary across datasets. In EMD-EUR+, CERRA, ERA5 (Figure 11, (b)) and MERRA2, the wind direction bias is more clockwise during the day and more anti-clockwise during the night. In NORA3 wind direction bias is more clockwise in winter and more anti-clockwise in summer. In NEWA the wind direction bias is more clockwise during the night and more anti-clockwise during the day. In JRA-3Q the wind direction bias is more clockwise closer to sunset.

Similarly to values of PC1 of the wind speed bias, the values of PC2 of the wind direction bias show differences between parts of Europe – the Central European flatland region and mountainous regions (Figure 12). The difference is that the sign of PC2 of the wind direction bias is the opposite to the sign of PC1 of the wind speed bias. That is, the PC2 values for wind direction bias are negative at observation sites in the Central European flatland region and positive at those in the mountainous regions. Characteristic patterns of the wind direction biases for these regions can be acquired by adding (subtracting) the loadings of PC2 of the wind direction (Figure 9, (c)) multiplied by the standard deviation of PC2 values between all observation sites to (from) the spatially averaged temporal pattern of the wind direction bias. In the Central European flatland region, the PC2 values for wind direction bias are negative. Therefore, in the Central European flatland region, the temporal patterns of the wind direction bias are closer to the ones shown in Figure 13 (a). In the Central European flatland region, all model datasets show more clockwise wind direction bias in winter and more anticlockwise wind direction bias in summer. In mountainous regions, PC2 of the wind direction bias shows positive values. Therefore, in mountainous regions the characteristic temporal patterns of the wind direction bias are closer to the ones shown in Figure 13, (c). In mountainous regions, the wind direction

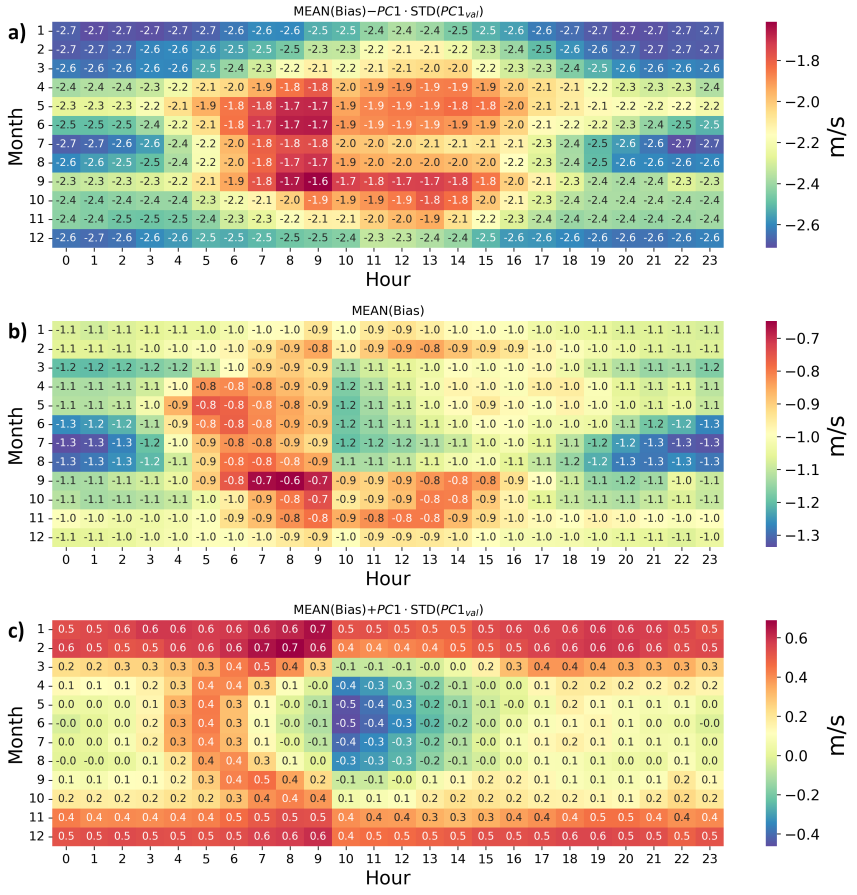

Figure 8. Wind speed biases ($\mathrm{m\,s^{-1}}$) in ERA5 reanalysis dependent on month (y-axis) and time of day (x-axis). (a) spatially averaged bias minus 1 standard deviation of values of the PC1 between observation locations (such temporal pattern of biases is characteristic to Scandinavian mountainous region and Southern Europe) (b) spatially averaged bias for the entire domain of study. (c) spatially averaged bias plus 1 standard deviation of values of the PC1 between observation locations (such temporal pattern of biases is characteristic to Central European flatland region). Note the differences in scale.

bias is more clockwise during the day and more anticlockwise during the night. PC2 of the wind direction bias describes only differences in the temporal patterns of the wind direction bias, while PC1 of the wind speed bias describes both differences in temporal patterns and the overall magnitude of the bias (more positive or negative wind speed bias).

The overall magnitude of the wind direction bias (overall clockwise or anticlockwise wind direction bias) is described by the PC1 of the wind direction bias. The loadings of PC1 of the wind direction bias are shown in Figure 9. Positive values of PC1 of the wind direction indicate that wind direction bias at a particular location is overall more clockwise, while negative values indicate that the bias is overall more anticlockwise. The PC1 loadings for wind direction bias vary across months and times of day. However, when adding (Figure 11, (c)) or subtracting (Figure 11, (a)) 1 standard deviation of values of PC1 of

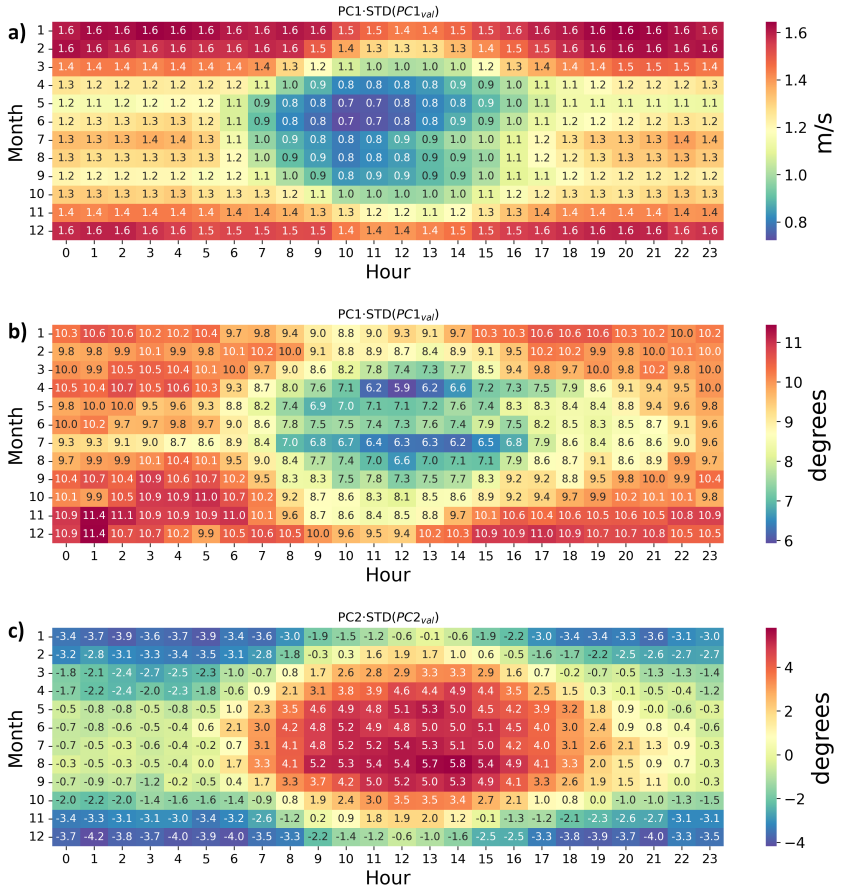

Figure 9. Comparison of ERA5 reanalysis and observations. (a) loadings of the PC1 for wind speed bias, (b) PC1 of wind direction bias and (c) PC2 of wind direction bias. Loadings are normalised by multiplying them by the standard deviation of values of the particular principal component between all observation locations. Units are m s^{-1} for wind speed and $^{\circ}$ for wind direction.

the wind direction bias to the spatially averaged wind direction bias has a much smaller effect on temporal patterns than in the case for PC1 of the wind speed bias and PC2 of the wind direction bias.

PC1 values for wind speed bias and PC2 values for wind direction bias showed regional differences across parts of Europe. On the other hand, we could not identify meaningful spatial patterns in the values of PC1 of wind direction bias. We acknowledge that, for confidentiality reasons, the exact locations of observations cannot be disclosed; therefore, the lack of evident spatial patterns cannot be shown. However, we demonstrate that PC1 values for wind direction bias show similar spatial patterns across all datasets when considering correlations. We considered correlations in PC1 values of wind direction bias across all observation locations for each pair of model datasets. The correlation for each model pair exceeds 0.70. If we consider only mesoscale models, the correlation for each model pair exceeds 0.90. This suggests strong agreement in the spatial patterns of PC1 values for wind direction bias across all model datasets.

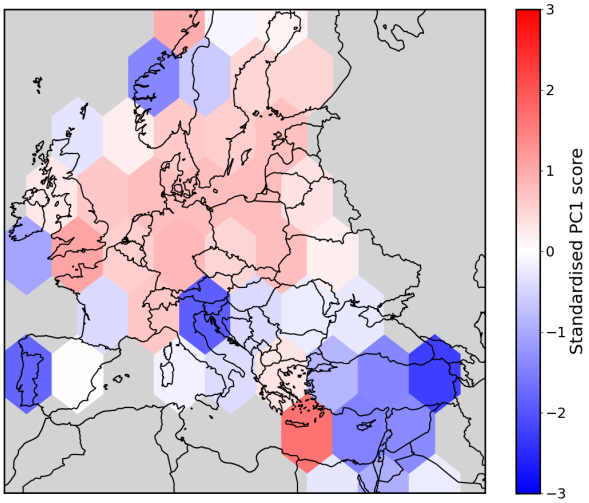

Figure 10. Spatial distribution of values of the PC1 of the wind speed bias in the ERA5 reanalysis. Colour of a hexagonal bin represents the average value of PC1 between all masts located inside the bin. Values are normalised by dividing a value of PC1 at each location by the standard deviation of all PC1 values between all observation locations.

The standard deviation of the average wind-direction bias across observation locations ranges from 8–10°, depending on the model dataset. The magnitude of the bias varies only slightly across model datasets, whereas the elevation biases in mesoscale models are more than two times smaller than those in global reanalyses, indicating that PC1 of the wind direction bias cannot be significantly explained by elevation bias. Nevertheless, in Figure 14, the elevation bias exhibits a weak explanatory signal in PC1 values, with a weak but statistically significant correlation of approximately 0.2 across most model datasets.

According to Möhrlen et al. (2022) a wind direction measurement bias of 5° is acceptable for the wind resource assessment, which is significantly lower than the one observed in our dataset. Jiménez and Dudhia (2013) compared the observed and modelled winds at 2 km resolution over Iberian peninsula. They have observed biases in mean wind direction on the order of tens of degrees. Significantly stronger wind-direction biases were present in complex terrain, which could only be partially mitigated by choosing a more representative nearby grid point. They explained these biases by the models' inability to sufficiently represent local terrain features at a 2 km resolution.

Since in our dataset strong wind direction biases were also present in a simple terrain, it suggests that the mean wind direction bias is likely a combination of measurement biases, biases induced by the local unresolved terrain features, biases caused by the elevation discrepancy between models and the real-world, and possibly biases caused by other physical processes that were not considered in this work. Identifying the exact causes of wind direction biases at each observation location would require manual inspection of the surrounding area. As our analysis focuses on identifying patterns of bias at a larger scale, manual inspection of the observation sites is outside the scope of this study.

To test the hypothesis of whether the biases are induced by the differences in the terrain elevation between models and the real world, correlations between principal components and terrain variables – real-world elevation, model elevation and

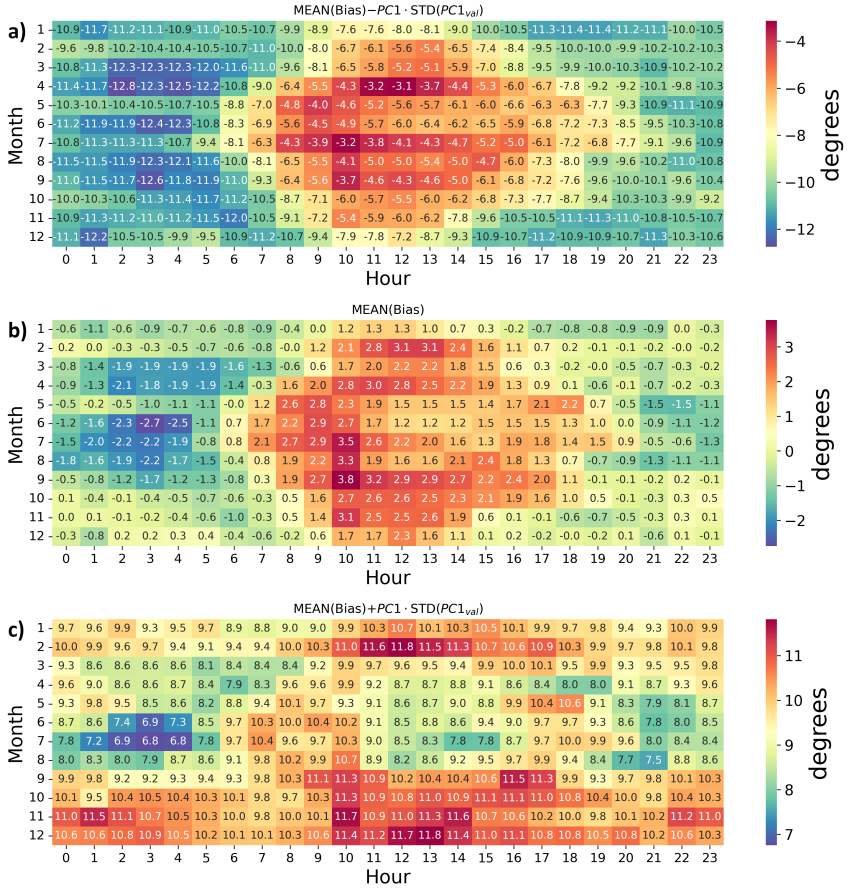

Figure 11. Wind direction biases (°) in the ERA5 reanalysis dependent on month (y axis) and time of day (x axis) for: (a) spatially averaged bias minus 1 standard deviation of values of the PC1 between observation locations, (b) spatially averaged bias across the entire study domain, and (c) spatially averaged bias plus 1 standard deviation of values of the PC1 between observation locations. Note the differences in scale.

their weighted difference were considered. The correlations between elevation variables and the principal components of wind speed and direction biases are summarised in Figure 14. Since the sign of principal components is arbitrary when applying PCA to different model datasets, correlations are multiplied by -1 when necessary to match the sign with the ERA5, as presented above. For all model datasets and all bias variables, the correlation becomes noticeably stronger when elevation bias is included, compared to a case where only real-world elevation is considered. The results of the F-test indicate that the increase in correlation is statistically significant in all cases at the 0.05 significance level.

While correlations between elevation bias and model bias variables are statistically significant, in some cases, they are weak and cannot support a hypothesis about a potential physical cause. However, it should be kept in mind that the atmosphere is chaotic and that model biases arise from a range of physical and numerical phenomena. Therefore, if causation can be

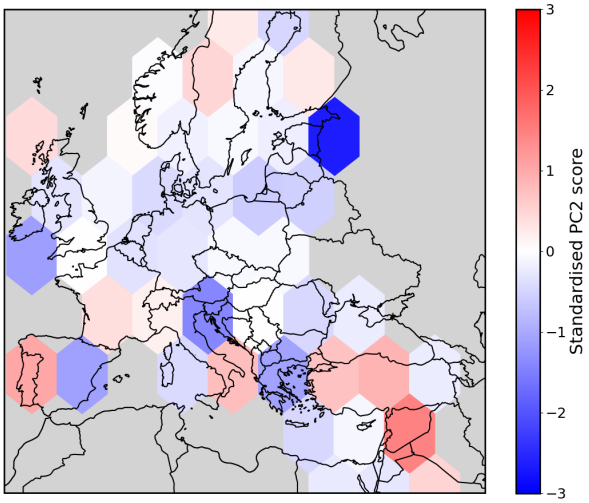

Figure 12. Map showing spatial distribution of values of the PC2 of the wind direction bias in ERA5 reanalysis. The colour of a hexagonal bin represents the average PC2 value across all masts within the bin. Values are normalised by dividing a value of PC2 at each location by the standard deviation of all PC2 values between all observation locations.

established, a low correlation implies the presence of noise in the data. In that case, correlation is used to measure the strength of the effect rather than to indicate its potential presence. For example, a correlation (R) of 0.3 between model bias and elevation bias indicates that elevation bias explains 9 % of the variance (R^2) in model biases. The numerical experiment that establishes the link between changing the elevation and the modelled winds is described in Section 3.2. In Section 4, it is described which physical phenomena might theoretically cause wind biases in the case when the model elevation differs from the real-world elevation.

The ERA5 reanalysis shows the strongest correlations between elevation bias and model bias variables. The correlation is 0.71 for PC1 of the wind speed bias, and 0.41 and -0.50 for PC1 and PC2 of the wind direction bias, respectively. In the case of wind speed, the effects of the elevation bias on the mean bias and temporal patterns of biases are combined into PC1. Meanwhile, in the case of wind-direction bias, the effects of the elevation bias are split between PC1 for the mean bias and PC2 for temporal patterns of bias. When elevation bias is accounted for, EMD-EUR+ and NEWA show the strongest increases in correlation compared with the case where only real-world elevation is considered. For PC1 of the wind speed bias, the correlation with real-world elevation is only -0.16 for EMD-EUR+ and -0.15 for NEWA. However, when elevation bias is accounted for, the correlations increase to 0.37 and 0.46, respectively. The correlations between real-world and model elevation are 0.93–0.94 for global reanalyses and 0.98–0.99 for mesoscale model datasets.

All correlations between the principal components of model biases and weighted elevation differences between the model and the real world are statistically significant at the 0.05 significance level. The smallest number of masts was used when assessing biases in the NORA3 dataset, only 254. For 254 data points, a correlation coefficient with an absolute value greater than 0.12 is statistically significant at the 0.05 significance level. For other datasets, the correlation threshold is even lower.

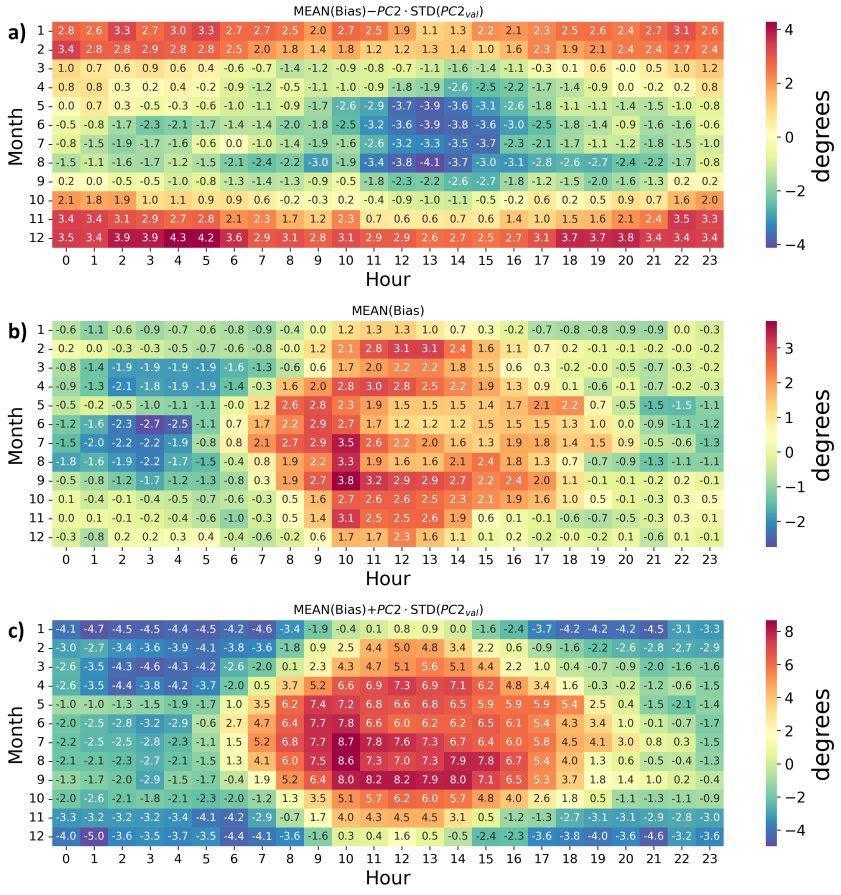

Figure 13. Wind direction biases (in degrees) in the ERA5 reanalysis as a function of month (y axis) and time of day (x axis) for: (a) spatially averaged bias minus 1 standard deviation of values of the PC2 between observation locations (such temporal pattern of biases is characteristic to Central European flatland region), (b) spatially averaged bias for the entire domain of study, and (c) spatially averaged bias plus 1 standard deviation of values of the PC2 between observation locations (such temporal pattern of biases is characteristic to Scandinavian mountainous region and Southern Europe). Note the differences in scale.

When the model elevation or the real-world elevation is considered individually, the correlation is not always statistically significant.

The correlations suggest a statistical relationship between model wind biases and the elevation bias. If observations are performed above the model grid-cell elevation, models tend to underestimate wind speed. The underestimation is stronger in winter than in summer, and stronger at night than during the day. Additionally, the model shows an overall anti-clockwise wind direction bias. The bias is more anti-clockwise in winter than in summer, and more anti-clockwise at night than during the day. On the other hand, if observations are made below the model grid-cell elevation, models tend to overestimate wind speed. The overestimation is stronger in winter than in summer, and stronger at night than during the day. Additionally, the model exhibits

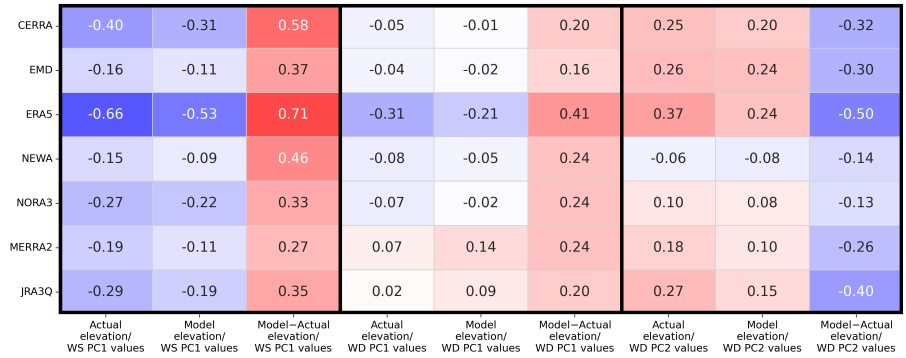

Figure 14. Correlation between values of principal components that describe model biases and terrain variables (x-axis) for different model datasets (y-axis). WS denotes wind speed, and WD denotes wind direction biases.

an overall clockwise wind-direction bias. The bias is more clockwise in winter than in summer, and more clockwise at night than during the day.

3.1.1 100 m wind speed bias in ERA5

One of the objectives of the work was to assess how the 100 m wind speed bias in ERA5 is carried over into downscaled mesoscale model datasets. In Figure 8, it is clearly visible that in all months, speed bias becomes more negative from 9:00 UTC to 10:00 UTC in the ERA5 reanalysis. The magnitude of the change is 0.1-0.3 $\mathrm{m\,s^{-1}}$ depending on the month, which is significantly higher than the usual hourly change in the mean bias for the particular month. It is consistent with the description of the bias in Jourdier (2020), that the modelled wind speed shows an unrealistic decrease from 9:00 UTC to 10:00 UTC. Jourdier (2020) also mentions that the decrease is stronger during convective atmospheric conditions, which is consistent with the weaker magnitude of the decrease in winter. The decrease is also visible in the data from the CERRA reanalysis. However, it is less prominent. Meanwhile, there is no noticeable unrealistic change in the wind speed bias between 9:00 and 10:00 UTC in the plots for NORA3, NEWA, and EMD-EUR+.

The prominence of a 100 m wind-speed bias in the mesoscale model output can be explained by model initialisation strategies. CERRA reanalysis assimilates observations every three hours, and in this work, data with a spin-up time of only one hour are used. Moreover, both CERRA and ERA5 are developed by the same organisation, ECMWF. Therefore, the CERRA reanalysis is likely to be closer to the host reanalysis, ERA5, than to other downscaled products.

The NORA3 dataset uses a longer spin-up time of 3 hours than CERRA, which is likely to reduce the effect of bias. NEWA and EMD-EUR+ reinitialise the model once a week. Moreover, they are always initialised at 0:00 UTC, during the 21:00–9:00 UTC assimilation window, which does not experience 100 m wind speed issues in the ERA5. Therefore, these model datasets are less affected by the issue. However, it has to be kept in mind that all mesoscale models use boundary conditions from ERA5, which are regularly updated. In addition, NEWA and EMD-EUR+ are nudged towards ERA5. Therefore, they are

likely to experience a performance drop due to the 100 m wind speed issue in ERA5, even though it is not directly observable in the plots of this work.

3.2 Comparison of the WRF model simulations

The results of the model dataset comparison with observations showed that wind speed and direction biases are correlated with the elevation bias. Moreover, the magnitude of the bias depends on the time of day and the time of year. As model terrain elevation is something that can be easily adjusted, we made a hypothesis that the bias can be potentially improved if the model terrain is adjusted to better match the real world terrain.

To test the hypothesis, we conducted an experiment by running two numerical simulations using different terrain input datasets and comparing the modelled winds. The goal of the experiment was to identify how the diurnal and seasonal patterns of the modelled winds change when the terrain is smoothed in the model. We anticipated that the main mode of wind differences between model runs with $5'$ and $30''$ input terrain resolution would be similar to the main mode of differences between modelled and observed winds. It would indicate that by adjusting terrain elevation in the model run, it is possible to adjust modelled winds to reduce the bias caused by the elevation discrepancy.

While there is potential to reduce model bias caused by elevation discrepancies by adjusting the model elevation, it is unclear how this would affect overall model performance. Smoothing of the terrain affects both slopes and the resolved orographic drag caused by terrain variability, neither of which is directly addressed by our analysis. Moreover, in models, the terrain is smoothed for practical reasons — to reduce steep elevation gradients that might cause numerical instabilities.

When comparing two model simulations, PCA was applied in a similar way as it was applied when comparing model datasets and observations. However, rather than using only 508 observation locations, WRF model simulations were compared across all 250×520 grid cells.

The loadings for the PCA of the model run comparison are shown in Figure 15. The loadings of PC1 of the wind speed difference when comparing two model simulations show similar diurnal and seasonal patterns as the loadings of PC1 of the wind speed when comparing model datasets with observations (Figure 9). In case of wind direction difference, there is a difference from the case when comparing model datasets with observations. When comparing model datasets with observations, PC1 describes the overall magnitude of the wind direction bias, and PC2 captures its diurnal and seasonal patterns. When comparing two WRF model runs, the PC1 loadings for the wind direction difference show similarities to those for wind speed. PC1 of the wind direction difference between WRF model runs captures both the overall magnitude and the temporal patterns of that difference. A potential explanation for this difference is that, when comparing two model runs, there is no overall wind-direction bias due to improper reference-direction calibration or subgrid obstacles.

For both wind speed and wind direction differences, values of PC1 show similar spatial patterns (Figure 16). PC1 shows slightly negative values over the sea and in Southern Sweden, where the terrain is simple. In the northern parts of Sweden, where the terrain is complex, PC1 shows a high contrast in values: strongly positive and strongly negative. The patterns of these values follow the terrain. Strongly positive values are present in valleys, while strongly negative values are present on top of hills.

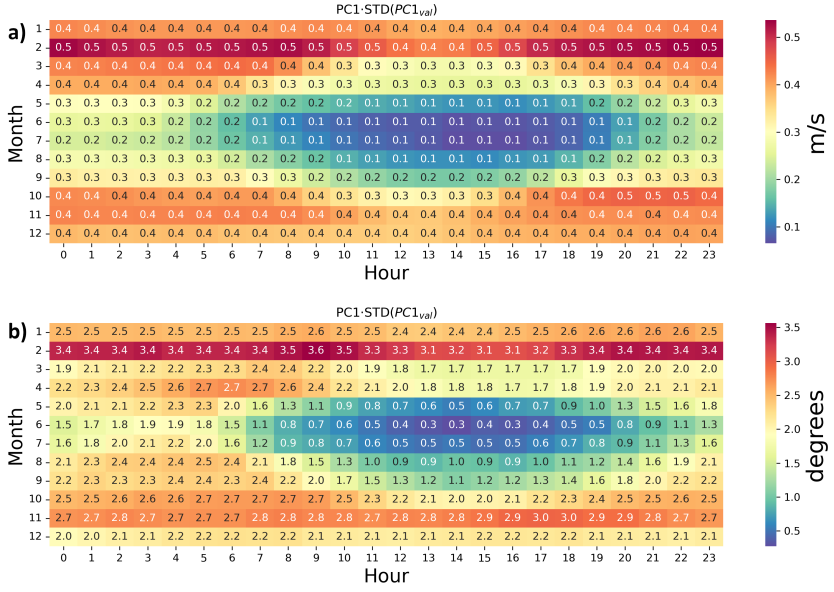

Figure 15. Comparison of the two WRF model simulations: (a) Loadings of the PC1 for wind speed difference and (b) wind direction difference multiplied by the standard deviation of the PC1 values between all grid-points. Units are $\mathrm{m\,s^{-1}}$ for wind speed and $^\circ$ for wind direction.

PC1 of the wind speed difference explains 75 % of the variance in wind speed differences between both WRF simulations. PC1 of the wind direction difference explains 22 % of the variance in wind direction differences between both WRF model simulations.

Similarly to observations, the principal components of the differences between two model simulations show correlations with the elevation difference between these two model runs. Values of PC1 of the wind speed difference between model simulations show a 0.35 correlation with the elevation in the 30″ model run and a 0.42 correlation with the elevation in the 5′ model run. When the weighted difference between elevations is considered, the correlation increases to 0.69. The PC1 values of the wind direction difference between model simulations show a 0.05 correlation with the elevation in the 30″ model run and 0.12 correlation with the elevation in the 5′ model run. When the weighted difference between elevations is considered, the correlation increases to 0.56. The correlation between elevation in the 30″ model run and the 5′ model run is greater than 0.99.

These correlations suggest a similar statistical relationship between elevation and wind differences when comparing two model simulations, as when model datasets are compared with observations. If the terrain elevation of a grid cell is the same in two model runs, model simulations show, on average, small differences in wind speeds and directions. On the other hand, if the terrain elevation of a grid cell differs between two model runs, the model simulation with the higher average elevation shows higher wind speed and a more clockwise wind direction. The wind speed difference is higher in winter than in summer

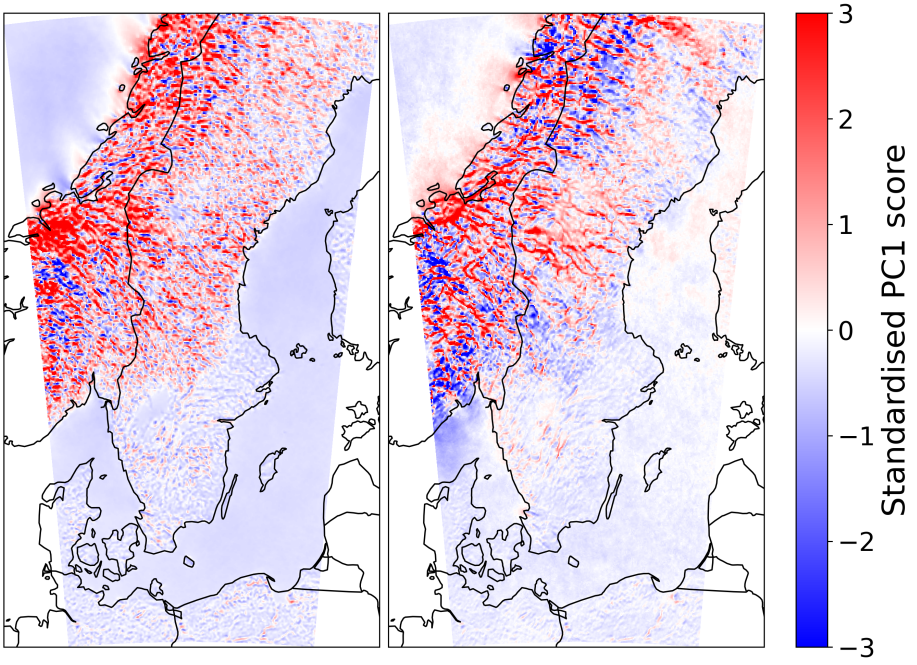

Figure 16. Values of PC1 for differences between both model runs of wind speed (left) and wind direction (right). Values are normalised by dividing a value of PC1 at each grid-point by the standard deviation of all PC1 values between all grid-points.

and higher during the night than during the day. The difference in wind direction is more clockwise in winter than in summer and more clockwise during the night than during the day.

By multiplying the regression coefficients that explain the bias with the elevation difference by the principal component loadings, it is possible to estimate the magnitude of the bias induced by the elevation difference. The elevation difference of 100 m between two model runs induces a difference in wind speed between 0.1 $\mathrm{m\,s^{-1}}$ in summer during the day and 0.9 $\mathrm{m\,s^{-1}}$ in winter. When comparing model datasets with observations, the elevation difference showed an effect on the wind speed bias of approximately the same magnitude. The elevation difference of 100 m induces a difference in the wind direction between 0.4° in summer during the day and 5.6° in winter. However, the effects of the elevation difference on the wind direction bias cannot be clearly isolated from the observations, since there is a large contribution to the wind direction bias from improper reference-direction calibration or obstacles in close proximity to the observation location. It is likely that the magnitude of the effect is similar to that observed when comparing two model simulations. It should be noted that the magnitude of the effect of elevation on the modelled wind speed is averaged for the entire region and does not consider land use or any other factors.

Loadings of the main principal components and correlations with terrain for the two model-simulation comparisons and the model-observation comparison look similar. However, spatial patterns of principal component values at first glance appear different. When comparing model datasets and observations, PC1 values for wind speed bias are more positive in simple terrain and more negative in complex terrain. On the other hand, when comparing two model simulations, the PC1 values for the wind

speed difference are more negative on top of the hills, more positive in valleys, and closer to zero in simple terrain. However, differences in the spatial patterns of principal component values can be explained by the fact that observation campaigns in our dataset are more often conducted on hilltops. As observation campaigns are more likely to be performed on top of the hill, the strongly negative PC1 values of the wind speed bias are more abundant than the strongly positive values when comparing model datasets to observations in the complex terrain. This explains the more negative PC1 values of wind speed bias in complex terrain in observations.

4 Discussion

Differences in model biases between parts of Europe have been observed before in the NEWA and ERA5 model datasets (Dörenkämper et al., 2020). They have shown that wind speed biases in Europe become more positive towards the north and more negative towards the east. Previous studies have found that both the MERRA and MERRA2 reanalyses underestimate wind speeds in Northern and Southern Europe, while overestimating them in Central Europe (Staffell and Pfenninger, 2016; Monforti and Gonzalez-Aparicio, 2017). They have noted that reanalyses underestimate wind capacity factors in Southern Europe by up to 30% and overestimate them in Central Europe by up to 60%. This work confirms the presence of the bias in MERRA2 and shows that the same bias is present in the ERA5 and JRA-3Q global reanalyses. Moreover, the bias is present in mesoscale datasets which perform dynamic downscaling. However, the magnitude of the bias is lower in mesoscale model datasets. In synoptic scale models, the standard deviation of the wind speed bias between all observation locations is 1.4-1.8 $\mathrm{m\,s^{-1}}$ in winter during the night and 0.5-0.7 $\mathrm{m\,s^{-1}}$ in summer during the day. On the other hand, in mesoscale models, the standard deviation of the wind speed bias across all observation locations is 1.1-1.3 $\mathrm{m\,s^{-1}}$ at night in winter and 0.3-0.6 $\mathrm{m\,s^{-1}}$ during the day in summer.

We acknowledge that the use of 10 m winds and only two levels in ERA5 and MERRA2 reanalyses might have introduced additional biases, especially for the highest observation campaigns exceeding 200 m AGL. However, most of the campaigns used in this work reported winds at heights between 50 m and 125 m, which are closer to the upper height levels of the models and therefore are less likely to be significantly affected by errors in vertical extrapolation. The identified wind biases in ERA5 and MERRA2 are similar to those in the other five model datasets. Moreover, these results are supported by the WRF modelling. Therefore, it is unlikely that biases in ERA5 and MERRA2 reanalyses related to differences in terrain elevation were artificially introduced by using 10 m winds and only two model levels.

More negative wind speed bias on hilltops and more positive wind speed bias in valleys has been observed before, e.g. Jiménez and Dudhia (2012); Winstral et al. (2016). Previous studies (Carvalho et al., 2012; Schmidli et al., 2018; Saadatabadi et al., 2024) have suggested that the model's performance in modelling winds is generally better when the model elevation more closely matches the real-world elevation. In this work, we establish a potential link between these two results. Namely, the wind speed bias is correlated with the elevation bias and modelled wind speeds can be adjusted in a controlled way by changing the model terrain elevation. Due to terrain smoothing in models, the absolute elevations of the valleys are higher, and

the absolute elevations of the hilltops are lower than the real-world elevations. It explains the differences in wind-speed biases between valleys and hilltops due to elevation bias.

In the observation dataset used, measurements in the Scandinavian mountainous region and Southern Europe are generally conducted in more complex local terrain with a higher σ_z. Therefore, their elevation is likely to differ significantly from that of the model grid cell. Meanwhile, in the Central European flatland region, observations are performed in simpler terrain where σ_z is lower. Therefore, observations in the Central European flatland region are generally performed closer to the elevation of the model grid cell. The distribution of elevation biases is skewed towards locations where the real-world elevation exceeds the model elevation. Therefore, in simpler terrain, observation campaigns are more likely to be located slightly above the model elevation, whereas in complex terrain, they are more likely to be located significantly higher. This may explain the generally more negative wind speed bias in the Scandinavian mountainous region and Southern Europe, and the generally more positive bias in the Central European flatland region.

The results of this work show that wind-speed and wind-direction biases in models are correlated with elevation bias. The correlation between wind speed bias and elevation bias ranges from 0.27 to 0.71, depending on the model dataset. The correlation between wind direction bias and elevation bias ranges from 0.13 to 0.50, depending on the model dataset. When two WRF model simulations that used different terrain input data were compared, we observed a similar response of the modelled winds to terrain smoothing, as was present when the model datasets were compared with the observations. It suggests that wind biases can be mitigated by adjusting the model's elevation to better match real-world elevation. However, the implications of modifying the model elevation on the overall model performance require further research. On the one hand, modifying the model elevation can improve the representation of other terrain features beyond the local elevation differences. These features include terrain slopes, terrain variability that affects resolved orographic drag, and large mountains that affect winds hundreds of kilometres downstream. Effects of some unresolved (subgrid) terrain features, such as slopes and terrain variability, can be parametrised in atmospheric models (Helbig et al., 2017).

The results of Schmidli et al. (2018) suggest that decreasing the smoothing intensity of the terrain can improve the model performance in modelling winds in complex terrain. On the other hand, terrain smoothing in atmospheric models is performed to avoid sharp terrain gradients, which can cause non-physical pressure gradients (Mahrer, 1984) or numerical instabilities (Bouëdec et al., 2025). Moreover, smoothing can potentially help average out uncertainties in the DEMs themselves. Therefore, the smoothing intensity of the terrain must be balanced to preserve the main features of the terrain while avoiding numerical artefacts.

The results of this work do not directly provide any information on the exact underlying physical processes that cause biases. It is likely that multiple physical processes are affected by differences in how terrain is represented between models and the real world. These processes are discussed in further paragraphs.

Since the terrain in the models is smoothed out, the heights of the hills and the depths of the valleys are generally underestimated. If the hill's height is underestimated in the model, it understates orographic speed-up, leading to lower wind speeds. Errors in estimating orographic speed-up have been described as a potential cause of wind speed bias before (Jiménez and Dudhia, 2012).

Wind speed generally increases with height, and wind direction tends to become more clockwise (in the Northern Hemisphere). If the height of the hill is underestimated in the model, it is exposed to a lower wind speed and a more anti-clockwise wind direction. The relationship between topographic exposure and model wind speed biases has previously been observed by Pauscher et al. (2025). Topographic exposure is defined as the angle between the horizontal and the horizon when looking in a particular direction. It describes how exposed to or sheltered from winds the location is. A positive exposure angle indicates the location is sheltered by the terrain, while a negative exposure angle indicates it is exposed to wind.

Differences in topographic exposure can explain why wind speed and direction biases are lower during the day than at night. If we assume that above the planetary boundary layer (PBL) wind flow is purely geostrophic with some wind speed and direction. Due to friction, near-surface winds are slower and more anti-clockwise (in the Northern Hemisphere). Within the PBL, both wind speed and direction change continuously between these two values. The shapes of vertical wind speed and wind direction profiles depend on atmospheric stability, which varies with season and time of day (Kelly and Troen, 2016). If the atmosphere is more stable, vertical mixing is reduced, and wind speed and direction change more uniformly throughout the boundary layer. On the other hand, if the atmosphere is unstable and mixing is strong, vertical wind gradients are pronounced near the surface, and the change is slower at higher levels. If we assume that an observation mast that measures wind speed at 100 m height above the ground level is located on top of a hill, which is 50 m above the model grid cell level. In the model, the mast is exposed to winds at 100 m height, while in the real world, it is exposed to winds at 150 m height. During stable atmospheric conditions, the wind speed and direction differences between 100 m and 150 m heights are greater than during unstable conditions. Therefore, the difference between model output and observations is likely to be higher during stable conditions. Differences in topographic exposure can also explain the relationship between wind speed and wind direction biases.

The results of the analysis show that elevation differences have a stronger statistical effect on both wind speed and wind direction biases in winter than in summer. In winter, a 100 m elevation difference has a statistical effect of up to 0.9 $\mathrm{m\,s^{-1}}$ on wind speed and up to 5.6° on wind direction. Meanwhile, in summer, a 100 m elevation difference has a statistical effect of up to 0.5 $\mathrm{m\,s^{-1}}$ on wind speed and 3.9° on wind direction. In Europe, wind speed is generally higher in winter than in summer. Differences in topographic exposure and orographic speed-up have a stronger absolute effect at higher wind speeds, which may help explain differences in wind speed biases between summer and winter. Regarding the wind direction, a more positive wind-speed bias can cause stronger turbulent mixing, leading to a more clockwise wind-direction bias.

Another possible cause of wind biases related to the terrain representation is differences between the model's terrain slopes and those in the real world. Spatial aggregation of the terrain generally smooths out the slopes, which can affect multiple processes. Duine et al. (2024) has shown that terrain smoothing in models significantly affects modelled convective boundary layer heights. Moreover, slope smoothing affects the dynamics of slope-related flows. Differences in terrain slope affect the solar radiation received by the surface, which in turn affects vertical momentum transfer due to convection. If we assume that the winds at higher levels of the atmosphere are modelled correctly, in the Northern Hemisphere models overestimating (underestimating) vertical transport of momentum can cause clockwise (anticlockwise) near-surface wind direction bias and positive (negative) wind speed bias (Sandu et al., 2020).

Our analysis shows that elevation bias statistically explains wind bias locally, i.e., within the same grid cell. However, differences in terrain elevation can influence modelled winds further downstream, a potential effect not considered in our analysis.

Jiménez and Dudhia (2012) has previously noted that the WRF model overestimates wind speeds in valleys and underestimates wind speeds on top of hills. The study suggested and implemented wind correction using the Laplacian of the elevation. The Laplacian of the elevation on a regular orthogonal grid is defined as:

$$\Delta h_{i,j} = \frac{h_{i-1,j} + h_{i+1,j} + h_{i,j-1} + h_{i,j+1} - 4 \cdot h_{i,j}}{dx^2}, \tag{7}$$

where dx is the horizontal grid spacing and $h_{i,j}$ is the elevation at the position indexed i, j. Positive Laplacian indicates that the elevation of a grid-point is lower than the average of its neighbours, while negative Laplacian indicates the elevation of a grid-point is higher than the average of its neighbours. The Laplacian of the elevation was considered a statistical explanatory variable for model biases during the exploratory data analysis in this work. However, the elevation Laplacian showed weaker explanatory power than the elevation difference. When comparing principal components of differences between model simulations, weighted differences in the elevation Laplacian showed correlations of 0.43 for PC1 values of both wind speed and wind direction differences. Compared to 0.69 and 0.56 for the case when the weighted elevation difference is considered.

Jiménez and Dudhia (2012, 2013) have noted that the model performance in modelling winds in complex terrain can be significantly improved by choosing a more representative model grid point both horizontally and vertically. These studies have used a subjective approach that cannot be generalised to larger areas when assessing wind resources. Choosing the appropriate model grid point to represent the observation location should be physically justified. The results of this work suggest that choosing a nearby model grid point that better matches the observed elevation can reduce wind speed bias. However, it cannot be guaranteed that there exists a model grid point that matches the observed elevation. These studies also suggested choosing the appropriate grid point not only horizontally but also vertically. However, for wind energy applications, knowledge of the entire wind profile in the lowest few hundred metres is necessary, which would make the choice of procedure more complicated.

Microscale modelling has the potential to improve modelled winds in complex terrain. Due to their higher resolution, microscale models can represent features of the terrain that mesoscale models cannot. Moreover, finer terrain resolution can reduce the elevation difference between the model and the real world. However, an important factor in capturing the terrain elevation at the point of interest is the quality of the terrain input data and its processing. A significant limiting factor for microscale simulations is the computational cost. Mesoscale simulations can provide multiple decades of wind time series over large geographical regions. However, computational fluid dynamics (CFD) simulations at microscale resolution are not feasible to such an extent. Commonly, two microscale CFD modelling frameworks are used for wind resource modelling: Reynolds-averaged Navier-Stokes (RANS) and large eddy simulations (LES). While the RANS approach is computationally faster (Wagenbrenner et al., 2019), it struggles to model unsteady flows or cases where flow separation occurs (Sumner et al., 2010). Moreover, the results of the RANS simulations vary with the turbulence model used (Zou et al., 2022). On the other hand, LES provides more accurate and consistent results but is significantly more computationally expensive, especially under stable atmospheric conditions (Bechmann and Sørensen, 2010).

Typically, an approach conceptually similar to measure-correlate-predict (MCP) (Woods and Watson, 1997) is used to produce long-term time-series from microscale CFD simulations. Only a few cases of characteristic wind conditions are fully simulated using the CFD approaches (Milla-Val et al., 2024). Later, simple statistics (Goyal and Ojha, 2011) or machine learning (Dujardin and Lehning, 2022) can be applied to the reference series, such as mesoscale model output, to construct the complete time-series that considers microscale effects. The accuracy of such an approach depends significantly on the accuracy of mesoscale model output (Travaglini et al., 2024), which is used both as the inlet conditions for modelling reference cases and as the reference time series.

Microscale CFD approaches for wind resource assessment can be reasonably applied only to small geographical regions on the order of kilometres. For large-scale wind resource assessment, the WAsP linear model can be applied, e.g., (Dörenkämper et al., 2020). However, the study reports that the linear model largely overestimates wind speed in highly complex orography because its assumptions do not hold. Moreover, the approach works only with temporally averaged wind speed, which depends on the wind direction and does not directly produce the complete time series.

An important aspect to consider when downscaling mesoscale wind data to microscale is how to harmonise the elevation description between the two approaches. Before applying the WAsP linear model, the wind climate is generalised, i.e., microscale terrain effects resolved by the mesoscale model are removed (Floors et al., 2025). Similarly, when applying CFD downscaling, terrain differences between mesoscale and microscale models must be considered. Our results suggest that terrain discrepancies affect both wind speed and direction, and the magnitude of the effect depends on the time of day and the month.

5 Conclusions

A large dataset of wind observations from tall masts and remote sensing devices was used to address systematic biases across multiple atmospheric model datasets. To the authors' knowledge, such large scale comparisons between modelled and observed winds at heights spanned by turbine rotors have not been conducted before. It was shown that all seven different atmospheric model datasets share the same bias when modelling wind speeds at heights spanned by turbine rotors. Models generally underestimate wind speeds in mountainous regions — such as the Scandinavian mountainous region and Southern Europe — while overestimating them in the Central European flatland region. Moreover, it was shown that the diurnal and seasonal cycles of wind speed biases also differ between these parts of Europe. It was observed that wind direction biases share the same diurnal and spatial patterns as wind speed biases.

The elevation difference between the models and the real-world terrain is a statistically significant factor associated with wind speed and certain wind-direction biases across multiple model datasets. The PC1 of wind speed biases explains between 61.1% and 82.5% of the variance in model biases across observation locations, depending on the model dataset. The PC1 values of wind speed biases show a correlation with the elevation difference between 0.27 and 0.71, depending on the model dataset. Comparison between two different model simulations allowed to estimate that the elevation difference of 100 m has a statistical effect between 0.1 and 0.9 $\mathrm{m\,s^{-1}}$ on the wind speed bias and an effect between 0.4° and 5.6° on the wind direction bias. The magnitude of the effect depends on the time of day and the month of the year. The effect of the elevation difference on

bias when comparing model datasets and observations is similar in magnitude to that when comparing two model simulations. It should be noted that we considered only the average magnitude of the diurnal bias in the region. However, the intensity of the diurnal bias cycle at a particular location can also depend on land-cover type, which we did not consider in our analysis.

The spatial differences in biases between the Central European flatland region and mountainous regions (the Scandinavian mountainous region and Southern Europe) can be explained by the dataset being biased towards observation campaigns located at locally higher elevations, where real-world elevation exceeds the model elevation. In mountainous regions, the magnitude of differences between real-world and model terrain elevations is greater, leading to stronger biases. These height differences influence how several physical processes are represented in the model, thereby impacting modelled winds.

We used the elevation difference between the model and the real-world terrain as a measure of how terrain differs between models and the real-world. It directly describes the effects of the elevation differences on changes in topographic exposure and orographic speed-up due to terrain smoothing. Moreover, it indirectly contains information on how strong slopes are locally affected, thereby influencing modelled slope flows and resolved orographic drag. However, it cannot fully describe the complicated nature of flow changes in complex terrain due to terrain smoothing. The results of this work cannot distinguish the contribution of different physical processes to this wind bias. However, the results suggest that the bias may be reduced by modifying the model terrain to more closely match real-world elevation. On the other hand, terrain smoothing in numerical atmospheric models is done for practical reasons. Overly steep slopes in atmospheric models can lead to numerical instabilities, especially when running a model with a dense vertical-level distribution near the surface. Ensuring the numerical stability of the model is crucial to avoid unnecessary computational costs when performing operational model runs. Therefore, further research is needed to balance the terrain smoothness level to keep terrain features well represented while ensuring the model's numerical stability.

Data availability. Observation data from the EMD internal mast database and EMD-EUR+ dataset are not publicly available. Data from NORA3 model dataset is available from https://thredds.met.no/thredds/projects/nora3.html. Data from CERRA reanalysis is available from https://cds.climate.copernicus.eu/datasets/reanalysis-cerra-height-levels. Data from ERA5 reanalysis is available from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels. Data from MERRA2 reanalysis is available from https://disc.gsfc.nasa.gov/datasets?project=MERRA-2. Data from NEWA model dataset is available from https://map.neweuropeanwindatlas.eu/. Data from JRA-3Q reanalysis is available from https://gdex.ucar.edu/datasets/d640001/.

Appendix A: WRF setup

Table A1. Setup of the WRF simulations used in this work.

WRF version	4.5.1.
Domain	Single 810×1620 km domain centred on Sweden.
Grid	3 km grid spacing. 270×540 grid cells. 10 cell wide buffer zone at the boundary.
Vertical levels	50
Simulation length	Simulation covering the entire 2011 stitching together 48 h runs with 6 h spin-up.
Terrain data	GMTED2010 with 30″ and 5′ resolution
Dynamical forcing	Hourly ERA5 data on surface and pressure levels
Time step	12 s
PBL scheme	MYJ
Surface layer scheme	Eta Similarity
Land surface model	Noah
Cloud microphysics scheme	Ferrier Eta
Radiation scheme	RRTMG
Cumulus scheme	None

Appendix B: Results of PCA for all model datasets

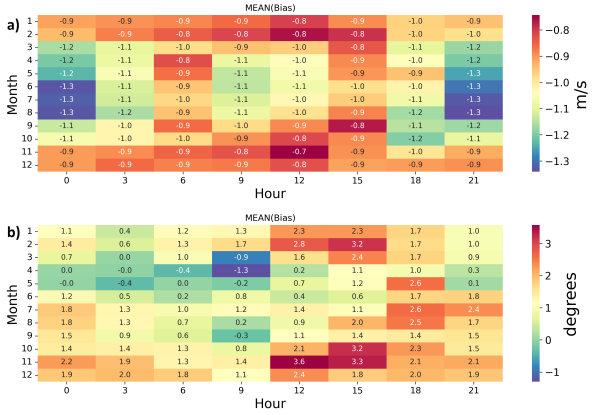

Figure B1. Spatially averaged biases for the entire domain of study for wind speed (in $\mathrm{m\,s^{-1}}$) (a) and wind direction (in degrees) (b) in JRA-3Q reanalysis dependent on month (y axis) and time of day (x axis).

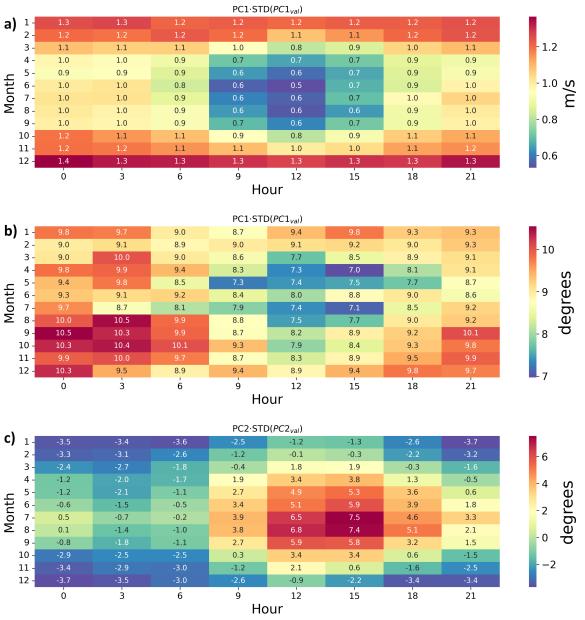

Figure B2. JRA-3Q reanalysis and observation comparison. Loadings of the PC1 for wind speed bias (a), PC1 of wind direction bias (b) and PC2 of wind direction bias (c). Multiplied by the standard deviation of values of the corresponding PC.

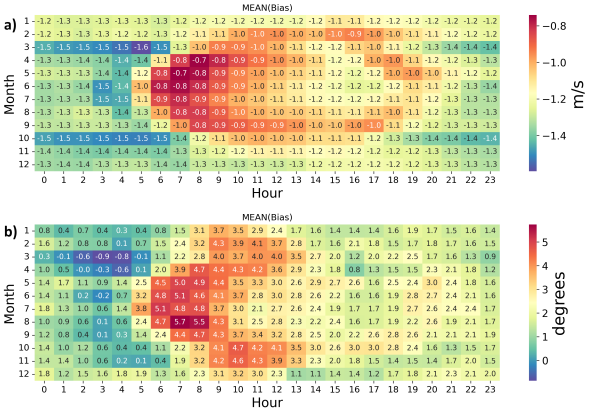

Figure B3. Spatially averaged biases for the entire domain of study for wind speed (in $\mathrm{m\,s^{-1}}$) (a) and wind direction (in degrees) (b) in MERRA2 reanalysis dependent on month (y axis) and time of day (x axis).

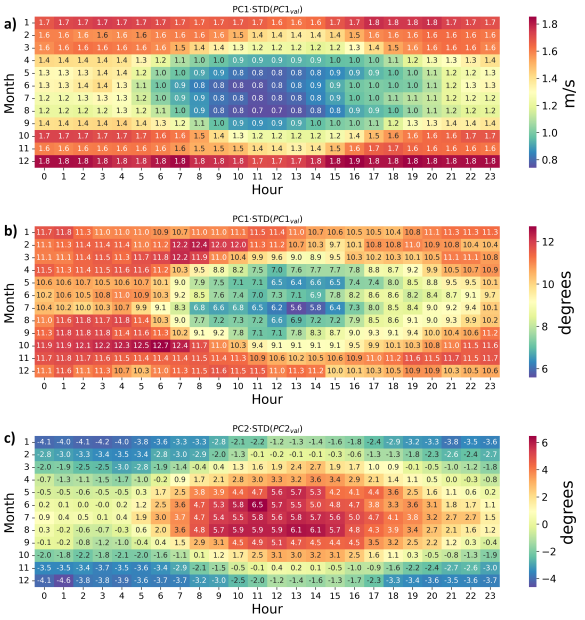

Figure B4. MERRA2 reanalysis and observation comparison. Loadings of the PC1 for wind speed bias (a), PC1 of wind direction bias (b) and PC2 of wind direction bias (c). Multiplied by the standard deviation of values of the corresponding PC.

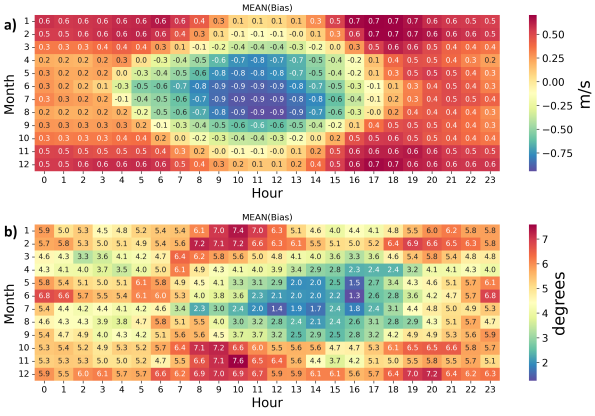

Figure B5. Spatially averaged biases for the entire domain of study for wind speed (in $\mathrm{m\,s^{-1}}$) (a) and wind direction (in degrees) (b) in NEWA model dataset dependent on month (y axis) and time of day (x axis).

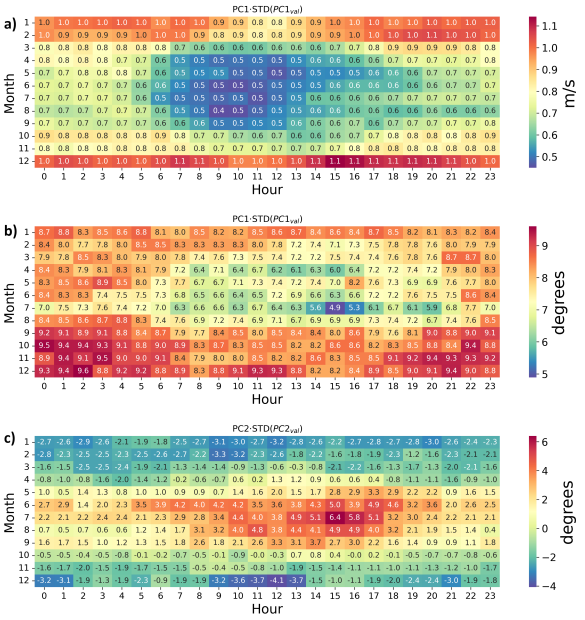

Figure B6. NEWA model dataset and observation comparison. Loadings of the PC1 for wind speed bias (a), PC1 of wind direction bias (b) and PC2 of wind direction bias (c). Multiplied by the standard deviation of values of the corresponding PC.

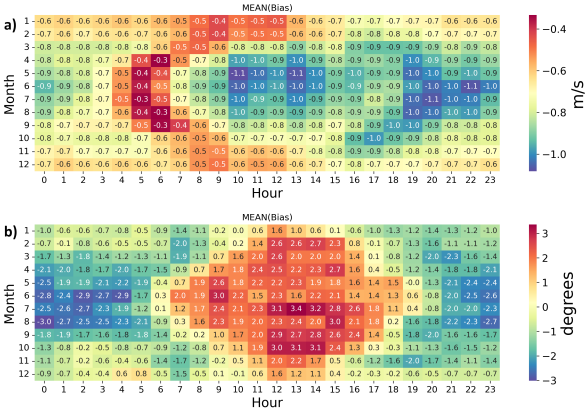

Figure B7. Spatially averaged biases for the entire domain of study for wind speed (in $\mathrm{m\,s^{-1}}$) (a) and wind direction (in degrees) (b) in CERRA model dataset dependent on month (y axis) and time of day (x axis).

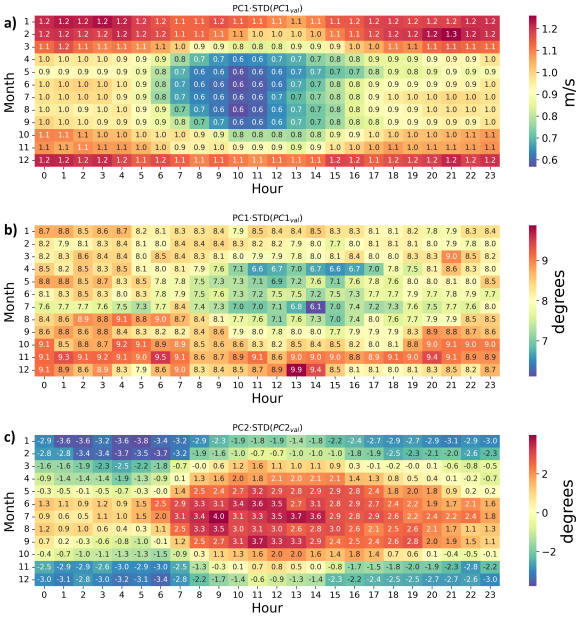

Figure B8. CERRA model dataset and observation comparison. Loadings of the PC1 for wind speed bias (a), PC1 of wind direction bias (b) and PC2 of wind direction bias (c). Multiplied by the standard deviation of values of the corresponding PC.

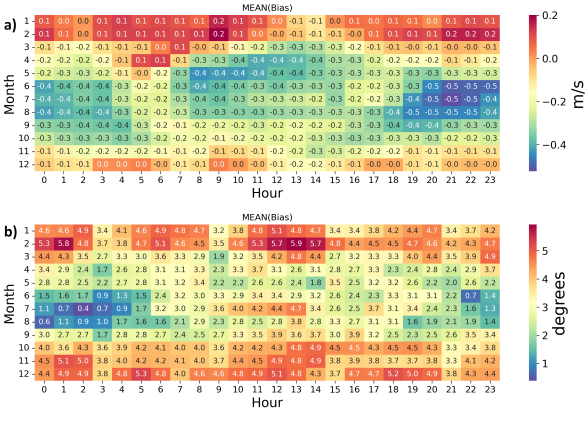

Figure B9. Spatially averaged biases for the entire domain of study for wind speed (in $\mathrm{m\,s^{-1}}$) (a) and wind direction (in degrees) (b) in NORA3 model dataset dependent on month (y axis) and time of day (x axis).

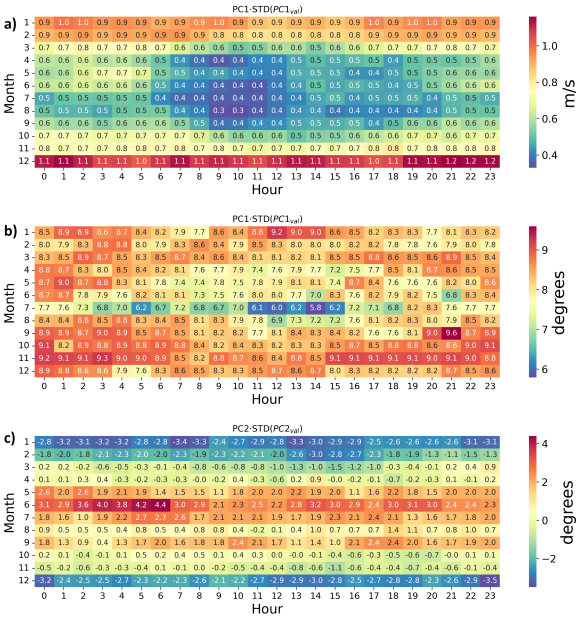

Figure B10. NORA3 model dataset and observation comparison. Loadings of the PC1 for wind speed bias (a), PC1 of wind direction bias (b) and PC2 of wind direction bias (c). Multiplied by the standard deviation of values of the corresponding PC.

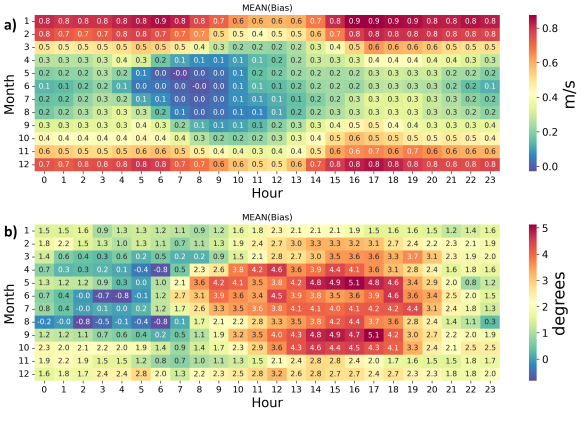

Figure B11. Spatially averaged biases for the entire domain of study for wind speed (in m s^{-1}) (a) and wind direction (in degrees) (b) in EMD-EUR+ model dataset dependent on month (y axis) and time of day (x axis).

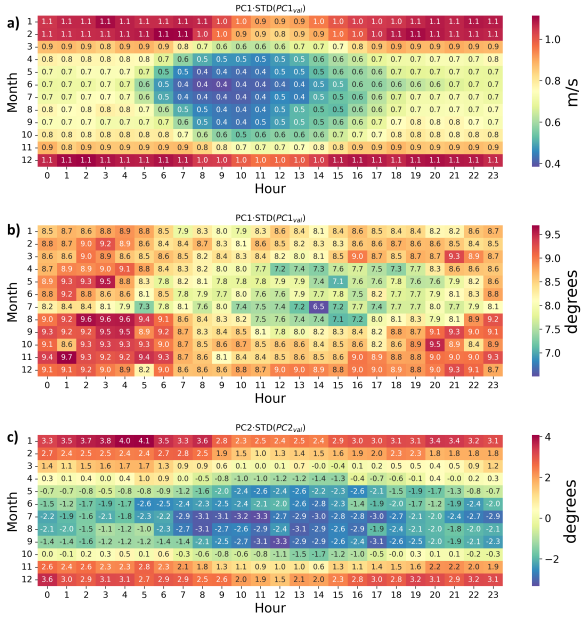

Figure B12. EMD-EUR+ model dataset and observation comparison. Loadings of the PC1 for wind speed bias (a), PC1 of wind direction bias (b) and PC2 of wind direction bias (c). Multiplied by the standard deviation of values of the corresponding PC.

Author contributions. MP designed the study, performed model dataset comparison against measurements, performed WRF simulations, analysed the results, and prepared the figures. TS, LS and AH provided scientific supervision. MP wrote the draft for the manuscript. All authors contributed to the final version of the manuscript.

Competing interests. MP and LS work at the EMD International A/S, which commercially provides data from the EMD-EUR+ model dataset and EMD's internal mast database. AH is a member of the editorial board of Wind Energy Science.

Acknowledgements. Maksims Pogumirskis has received funding from the European Union's Horizon Europe research and innovation programme (HORIZON-MSCA-2022-DN-01-01) under the Marie Sklodowska-Curie grant agreement no. 101119550 (AptWind).

References

Backwell, B., Clarke, E., Ladwa, R., Mullin, S., Jayasurya, F., and Hutchinson, M.: Global Wind Report 2025, Tech. rep., Global Wind Energy Council, Lisbon, Portugal, 2025.

Bechmann, A. and Sørensen, N. N.: Hybrid RANS/LES method for wind flow over complex terrain, Wind Energy, 13, 36–50, https://doi.org/https://doi.org/10.1002/we.346, 2010.

Belmonte Rivas, M. and Stoffelen, A.: Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT, Ocean Science, 15, 831–852, https://doi.org/10.5194/os-15-831-2019, 2019.

Bethere, L., Sennikovs, J., and Bethers, U.: Climate indices for the Baltic states from principal component analysis, Earth System Dynamics, 8, 951–962, https://doi.org/10.5194/esd-8-951-2017, 2017.

Bosilovich, M. G., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Todling, R., Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2: Initial Evaluation of the Climate, NASA Technical Report NASA/TM-2015-104606, Vol. 43, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, 2015.

Bouëdec, E. L., Chemel, C., and Staquet, C.: Dealing with steep slopes when modeling stable boundary-layer flow in Alpine terrain, Quarterly Journal of the Royal Meteorological Society, 151, e4835, https://doi.org/https://doi.org/10.1002/qj.4835, 2025.

Carvalho, D.: An Assessment of NASA's GMAO MERRA-2 Reanalysis Surface Winds, Journal of Climate, 32, 8261–8281, https://doi.org/10.1175/JCLI-D-19-0199.1, 2019.

Carvalho, D., Rocha, A., Gómez-Gesteira, M., and Santos, C.: A sensitivity study of the WRF model in wind simulation for an area of high wind energy, Environmental Modelling & Software, 33, 23–34, https://doi.org/https://doi.org/10.1016/j.envsoft.2012.01.019, 2012.

Cheynet, E., Solbrekke, I. M., Diezel, J. M., and Reuder, J.: A one-year comparison of new wind atlases over the North Sea, Journal of Physics: Conference Series, 2362, 012 009, https://doi.org/10.1088/1742-6596/2362/1/012009, 2022.

Cheynet, E., Diezel, J. M., Haakenstad, H., Breivik, Ø., Peña, A., and Reuder, J.: Tall wind profile validation of ERA5, NORA3, and NEWA datasets using lidar observations, Wind Energy Science, 10, 733–754, https://doi.org/10.5194/wes-10-733-2025, 2025.

Collins, S. N., James, R. S., Ray, P., Chen, K., Lassman, A., and Brownlee, J.: Grids in Numerical Weather and Climate Models, in: Climate Change and Regional/Local Responses, edited by Zhang, Y. and Ray, P., chap. 4, IntechOpen, London, https://doi.org/10.5772/55922, 2013.

Danielson, J. J. and Gesch, D. B.: Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010), Open-File Report 2011–1073, U.S. Geological Survey, https://doi.org/10.3133/ofr20111073, 2011.

Das, A. and Baidya Roy, S.: JRA55 is the best reanalysis representing observed near-surface wind speeds over India, Atmospheric Research, 297, 107 111, https://doi.org/https://doi.org/10.1016/j.atmosres.2023.107111, 2024.

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, Geoscientific Model Development, 13, 5079–5102, https://doi.org/10.5194/gmd-13-5079-2020, 2020.

Draxl, C., Worsnop, R. P., Xia, G., Pichugina, Y., Chand, D., Lundquist, J. K., Sharp, J., Wedam, G., Wilczak, J. M., and Berg, L. K.: Mountain waves can impact wind power generation, Wind Energy Science, 6, 45–60, https://doi.org/10.5194/wes-6-45-2021, 2021.

Duine, G.-J., De Wekker, S. F. J., and Knievel, J. C.: The Influence of Terrain Smoothing on Simulated Convective Boundary-Layer Depths in Mountainous Terrain, Atmosphere, 15, https://doi.org/10.3390/atmos15020145, 2024.

Dujardin, J. and Lehning, M.: Wind-Topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning, Quarterly Journal of the Royal Meteorological Society, 148, 1368–1388, https://doi.org/https://doi.org/10.1002/qj.4265, 2022.

ECMWF: IFS Documentation CY41R2 - Part IV: Physical Processes, 4, European Centre for Medium-Range Weather Forecasts, https://doi.org/10.21957/tr5rv27xu, 2016.

European Commission: Commission sets out immediate actions to support the European wind power industry, https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5185, accessed: 2025-03-24, 2023.

Fahrland, E.: Copernicus DEM, Copernicus Digital Elevation Model, Product Handbook, Tech. rep., AIRBUS, https://dataspace.copernicus.eu/sites/default/files/media/files/2024-06/geo1988-copernicusdem-spe-002_producthandbook_i5.0.pdf, 2022.

Fernández-González, S., Martín, M. L., García-Ortega, E., Merino, A., Lorenzana, J., Sánchez, J. L., Valero, F., and Rodrigo, J. S.: Sensitivity Analysis of the WRF Model: Wind-Resource Assessment for Complex Terrain, Journal of Applied Meteorology and Climatology, 57, 733 – 753, https://doi.org/10.1175/JAMC-D-17-0121.1, 2018.

Floors, R. R., Davis, N., Olsen, B. T., Badger, J., and Hansen, B. O.: Global Wind Atlas v4: Generalized Wind Climates, https://doi.org/10.11583/DTU.28945943.v1, 2025.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)", Journal of Climate, 30, 5419 – 5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Gómez-Navarro, J. J., Raible, C. C., and Dierer, S.: Sensitivity of the WRF model to PBL parametrisations and nesting techniques: evaluation of wind storms over complex terrain, Geoscientific Model Development, 8, 3349–3363, https://doi.org/10.5194/gmd-8-3349-2015, 2015.

Goyal, M. K. and Ojha, C. S. P.: Evaluation of linear regression methods as downscaling tools in temperature projections over the Pichola Lake Basin in India, Hydrological Processes, 25, 1453–1465, https://doi.org/https://doi.org/10.1002/hyp.7911, 2011.

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

Haakenstad, H., Øyvind Breivik, Furevik, B. R., Reistad, M., Bohlinger, P., and Aarnes, O. J.: NORA3: A Nonhydrostatic High-Resolution Hindcast of the North Sea, the Norwegian Sea, and the Barents Sea, Journal of Applied Meteorology and Climatology, 60, 1443 – 1464, https://doi.org/10.1175/JAMC-D-21-0029.1, 2021.

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, Geoscientific Model Development, 13, 5053–5078, https://doi.org/10.5194/gmd-13-5053-2020, 2020.

Helbig, N., Mott, R., van Herwijnen, A., Winstral, A., and Jonas, T.: Parameterizing surface wind speed over complex topography, Journal of Geophysical Research: Atmospheres, 122, 651–667, https://doi.org/https://doi.org/10.1002/2016JD025593, 2017.

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., De 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., de 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/https://doi.org/10.1002/qj.3803, 2020.

Hollingsworth, A.: Validation and diagnosis of atmospheric models, Dynamics of Atmospheres and Oceans, 20, 227–246, https://doi.org/https://doi.org/10.1016/0377-0265(94)90019-1, forecasting weather and climate since, 1994.

Hu, W., Scholz, Y., Yeligeti, M., Bremen, L., and Deng, Y.: Downscaling ERA5 wind speed data: a machine learning approach considering topographic influences, Environmental Research Letters, 18, https://doi.org/10.1088/1748-9326/aceb0a, 2023.

Jafari, S., Chokani, N., and Abhari, R. S.: Terrain Effects on Wind Flow: Simulations With an Immersed Boundary Method, Volume 1: Aircraft Engine; Ceramics; Coal, Biomass and Alternative Fuels; Wind Turbine Technology, 869–878, https://doi.org/10.1115/GT2011-46240, 2011.

Jiménez, P. A. and Dudhia, J.: Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model, Journal of Applied Meteorology and Climatology, 51, 300 – 316, https://doi.org/10.1175/JAMC-D-11-084.1, 2012.

Jiménez, P. A. and Dudhia, J.: On the Ability of the WRF Model to Reproduce the Surface Wind Direction over Complex Terrain, Journal of Applied Meteorology and Climatology, 52, 1610 – 1617, https://doi.org/10.1175/JAMC-D-12-0266.1, 2013.

Jolliffe, I. T.: Principal Component Analysis, Springer Series in Statistics, Springer New York, NY, 2 edn., ISBN 978-0-387-95442-4, https://doi.org/10.1007/b98835, springer Science+Business Media, Published: 01 October 2002, 2002.

Jourdier, B.: Evaluation of ERA5, MERRA-2, COSMO-REA6, NEWA and AROME to simulate wind power production over France, Advances in Science and Research, 17, 63–77, https://doi.org/10.5194/asr-17-63-2020, 2020.

Jourdier, B., Diaz, C., and Dubus, L.: Evaluation of CERRA for wind energy applications, in: EMS Annual Meeting 2023, Bratislava, Slovakia, https://doi.org/10.5194/ems2023-311, 2023.

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, Quarterly Journal of the Royal Meteorological Society, 146, 1498–1515, https://doi.org/https://doi.org/10.1002/qj.3748, 2020.

Kelly, M. and Troen, I.: Probabilistic stability and 'tall' wind profiles: theory and method for use in wind resource assessment, Wind Energy, 19, 227–241, https://doi.org/https://doi.org/10.1002/we.1829, 2016.

Klaver, R., Haarsma, R., Vidale, P. L., and Hazeleger, W.: Effective resolution in high resolution global atmospheric models for climate studies, Atmospheric Science Letters, 21, e952, https://doi.org/https://doi.org/10.1002/asl.952, 2020.

Knoop, S., Bosveld, F. C., de Haij, M. J., and Apituley, A.: A 2-year intercomparison of continuous-wave focusing wind lidar and tall mast wind measurements at Cabauw, Atmospheric Measurement Techniques, 14, 2219–2235, https://doi.org/10.5194/amt-14-2219-2021, 2021.

Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T. Y., Tokuhiro, T., Sato, Y., Matsushita, Y., and Onogi, K.: The JRA-3Q Reanalysis, Journal of the Meteorological Society of Japan. Ser. II, 102, 49–109, https://doi.org/10.2151/jmsj.2024-004, 2024.

Kuru, S.: Evaluating the accuracy of NEWA, ERA5 and NORA3 in predicting onshore wind conditions: a comparative study using ICOS meteorological mast data in Sweden, Master of science dissertation, Uppsala University, Department of Earth Sciences, Campus Gotland, Uppsala, Sweden, 2023.

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

Le Toumelin, L., Gouttevin, I., Galiez, C., and Helbig, N.: A two-fold deep-learning strategy to correct and downscale winds over mountains, Nonlinear Processes in Geophysics, 31, 75–97, https://doi.org/10.5194/npg-31-75-2024, 2024.

Mahrer, Y.: An Improved Numerical Approximation of the Horizontal Gradients in a Terrain-Following Coordinate System, Monthly Weather Review, 112, 918 – 922, https://doi.org/10.1175/1520-0493(1984)112<0918:AINAOT>2.0.CO;2, 1984.

Matiu, M., Napoli, A., Kotlarski, S., Zardi, D., Bellin, A., and Majone, B.: Elevation-dependent biases of raw and bias-adjusted EURO-CORDEX regional climate models in the European Alps, Climate Dynamics, 62, 9013–9030, https://doi.org/10.1007/s00382-024-07376-y, 2024.

Meadows, M., Jones, S., and Reinke, K.: Vertical accuracy assessment of freely available global DEMs (FABDEM, Copernicus DEM, NASADEM, AW3D30 and SRTM) in flood-prone environments, International Journal of Digital Earth, 17, 2308 734, https://doi.org/10.1080/17538947.2024.2308734, 2024.

MEASNET: Evaluation of Site-Specific Wind Conditions, Tech. rep., Campus de Montegancedo. EDIFICIO “CIDA”., Madrid, Spain, 2022.

Meyer, P. J. and Gottschall, J.: How do NEWA and ERA5 compare for assessing offshore wind resources and wind farm siting conditions?, Journal of Physics: Conference Series, 2151, 012 009, https://doi.org/10.1088/1742-6596/2151/1/012009, 2022.

Mi, X. and Liu, P.: A Comparative Study of Near-Surface Wind Speed Observations and Reanalysis Datasets in China Over the Past 38 Years, International Journal of Climatology, https://doi.org/https://doi.org/10.1002/joc.8738, 2024.

Miao, H., Dong, D., Huang, G., Hu, K., Tian, Q., and Gong, Y.: Evaluation of Northern Hemisphere surface wind speed and wind power density in multiple reanalysis datasets, Energy, 200, 117 382, https://doi.org/https://doi.org/10.1016/j.energy.2020.117382, 2020.

Milla-Val, J., Montañés, C., and Fueyo, N.: Economical microscale predictions of wind over complex terrain from mesoscale simulations using machine learning, Modeling Earth Systems and Environment, 10, 1407–1421, 2024.

Möhrlen, C., Zack, J., Browell, J., Wilbert, S., Wirth, I., and Felder, M.: Recommended Practice for the Implementation of Renewable Energy Forecasting Solutions, Tech. rep., IEA Wind, https://iea-wind.org/wp-content/uploads/2022/06/IEAWind_Task36_Recommended_Practice_Part4_1st_Edition_public.pdf, 2022.

Monforti, F. and Gonzalez-Aparicio, I.: Comparing the impact of uncertainties on technical and meteorological parameters in wind power time series modelling in the European Union, Applied Energy, 206, 439–450, https://doi.org/https://doi.org/10.1016/j.apenergy.2017.08.217, 2017.

Muñoz Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Olauson, J.: ERA5: The new champion of wind power modelling?, Renewable Energy, 126, 322–331, https://doi.org/https://doi.org/10.1016/j.renene.2018.03.056, 2018.

Owens, R. G. and Hewson, T. D.: ECMWF Forecast User Guide, ECMWF, https://doi.org/10.21957/m1cs7h, 2018.

Pauscher, L., Geiger, D., Wirth, M., Yuan, D., Weber, H., and Callies, D.: Downscaling reanalysis models using a simple data-driven approach based on topography and atmospheric conditions, in: Proceedings of the Wind Energy Science Conference (WESC 2025), Wind Energy Science Conference, Nantes, France, 24–27 June 2025, 2025.

Pogumirskis, M., Sīle, T., Seņņikovs, J., and Bethers, U.: PCA analysis of wind direction climate in the baltic states, Tellus A: Dynamic Meteorology and Oceanography, 73, 1–16, https://doi.org/10.1080/16000870.2021.1962490, 2021.

Ramon, J., Lledó, L., Torralba, V., Soret, A., and Doblas-Reyes, F. J.: What global reanalysis best represents near-surface winds?, Quarterly Journal of the Royal Meteorological Society, 145, 3236–3251, https://doi.org/https://doi.org/10.1002/qj.3616, 2019.

Ramon, J., Lledó, L., Pérez-Zanón, N., Soret, A., and Doblas-Reyes, F. J.: The Tall Tower Dataset: a unique initiative to boost wind energy research, Earth System Science Data, 12, 429–439, https://doi.org/10.5194/essd-12-429-2020, 2020.

Ridal, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Schimanke, S., Andrae, U., Berggren, L., Brousseau, P., Dahlgren, P., Edvinsson, L., El-Said, A., Glinton, M., Hagelin, S., Hopsch, S., Isaksson, L., Medeiros, P., Olsson, E., Unden, P., and Wang, Z. Q.: CERRA, the Copernicus European Regional Reanalysis system, Quarterly Journal of the Royal Meteorological Society, 150, 3385–3411, https://doi.org/https://doi.org/10.1002/qj.4764, 2024.

Rouholahnejad, F., Meyer, P. J., and Gottschall, J.: Collocating wind data: A case study on the verification of the CERRA dataset, Journal of Physics: Conference Series, 2875, 012 016, https://doi.org/10.1088/1742-6596/2875/1/012016, 2024.

Saadatabadi, A. R., Hamzeh, N. H., Kaskaoutis, D. G., Ghasabi, Z., Penchah, M. M., Sotiropoulou, R.-E. P., and Habibi, M.: Optimization and Evaluation of the Weather Research and Forecasting (WRF) Model for Wind Energy Resource Assessment and Mapping in Iran, Applied Sciences, 14, https://doi.org/10.3390/app14083304, 2024.

Sandu, I., Zadra, A., and Wedi, N.: Impact of orographic drag on forecast skill, ECMWF Newsletter, 150, published January 2017, 2017.

Sandu, I., Bechtold, P., Nuijens, L., Beljaars, A., and Brown, A.: On the causes of systematic forecast biases in near-surface wind direction over the oceans, https://doi.org/10.21957/wggbl43u, 2020.

Schmidli, J., Böing, S., and Fuhrer, O.: Accuracy of Simulated Diurnal Valley Winds in the Swiss Alps: Influence of Grid Resolution, Topography Filtering, and Land Surface Datasets, Atmosphere, 9, https://doi.org/10.3390/atmos9050196, 2018.

Shaw, J. and Weller, H.: Comparison of Terrain-Following and Cut-Cell Grids Using a Nonhydrostatic Model, Monthly Weather Review, 144, 2085 – 2099, https://doi.org/10.1175/MWR-D-15-0226.1, 2016.

Skamarock, W. C.: Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra, Monthly Weather Review, 132, 3019 – 3032, https://doi.org/10.1175/MWR2830.1, 2004.

Solbakken, K., Birkelund, Y., and Samuelsen, E. M.: Evaluation of surface wind using WRF in complex terrain: Atmospheric input data and grid spacing, Environmental Modelling & Software, 145, 105 182, https://doi.org/https://doi.org/10.1016/j.envsoft.2021.105182, 2021.

Solbrekke, I. M., Sorteberg, A., and Haakenstad, H.: The 3 km Norwegian reanalysis (NORA3) – a validation of offshore wind resources in the North Sea and the Norwegian Sea, Wind Energy Science, 6, 1501–1519, https://doi.org/10.5194/wes-6-1501-2021, 2021.

Spangehl, T., Borsche, M., Niermann, D., Kaspar, F., Schimanke, S., Brienen, S., Möller, T., and Brast, M.: Intercomparing the quality of recent reanalyses for offshore wind farm planning in Germany's exclusive economic zone of the North Sea, Advances in Science and Research, 20, 109–128, https://doi.org/10.5194/asr-20-109-2023, 2023.

Staffell, I. and Pfenninger, S.: Using bias-corrected reanalysis to simulate current and future wind power output, Energy, 114, 1224–1239, https://doi.org/https://doi.org/10.1016/j.energy.2016.08.068, 2016.

Sumner, J., Watters, C. S., and Masson, C.: CFD in Wind Energy: The Virtual, Multiscale Wind Tunnel, Energies, 3, 989–1013, https://doi.org/10.3390/en3050989, 2010.

Svenningsen, L., Thøgersen, M., and Sørensen, T.: Validation of EMD-WRF EUROPE+ (ERA5) mesoscale dataset, Tech. rep., EMD International A/S, https://help.emd.dk/knowledgebase/content/TechNotes/TechnicalNote7_EMDWRF_EUROPE_Validation.pdf, 2020.

Torralba, V., Doblas-Reyes, F. J., and Gonzalez-Reviriego, N.: Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison, Environmental Research Letters, 12, 114 019, https://doi.org/10.1088/1748-9326/aa8a58, 2017.

Travaglini, R., Melani, P. F., Balduzzi, F., Orlando, S., Viti, V., and Bianchini, A.: A critical insight into the most effective CFD settings to model atmospheric stability in wind energy applications, Journal of Physics: Conference Series, 2767, 092 031, https://doi.org/10.1088/1742-6596/2767/9/092031, 2024.

United Nations Climate Change (UNFCCC): The Paris Agreement, https://unfccc.int/process-and-meetings/the-paris-agreement, accessed: 2025-03-24.

Wagenbrenner, N. S., Forthofer, J. M., Page, W. G., and Butler, B. W.: Development and Evaluation of a Reynolds-Averaged Navier–Stokes Solver in WindNinja for Operational Wildland Fire Applications, Atmosphere, 10, https://doi.org/10.3390/atmos10110672, 2019.

Weigel, A. P., Liniger, M. A., and Appenzeller, C.: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?, Quarterly Journal of the Royal Meteorological Society, 134, 241–260, https://doi.org/https://doi.org/10.1002/qj.210, 2008.

Wikipedia: Wind power in Europe, https://en.wikipedia.org/wiki/Wind_power_in_Europe, accessed: 2026-09-24, 2026.

Williams, R., Zhao, F., Melkonyan, N., Hutchinson, M., and Cheong, J.: Global Offshore Wind Report 2025, Tech. rep., Global Wind Energy Council, Lisbon, Portugal, 2025.

Winstral, A., Jonas, T., and Helbig, N.: Statistical Downscaling of Gridded Wind Speed Data Using Local Topography, Journal of Hydrometeorology, 18, https://doi.org/10.1175/JHM-D-16-0054.1, 2016.

Woods, J. and Watson, S.: A new matrix method of predicting long-term wind roses with MCP, Journal of Wind Engineering and Industrial Aerodynamics, 66, 85–94, https://doi.org/https://doi.org/10.1016/S0167-6105(97)00009-3, 1997.

Yamaguchi, A., Tavana, A., and Ishihara, T.: Assessment of Wind over Complex Terrain Considering the Effects of Topography, Atmospheric Stability and Turbine Wakes, Atmosphere, 15, https://doi.org/10.3390/atmos15060723, 2024.

Zhi, M. H.: Accuracy of Energy Estimation for Small Wind Farms Based on Mesoscale Wind Data, Bachelor of science dissertation, Uppsala University, Department of Earth Sciences, Campus Gotland, 2023.

Zou, Y., Yue, P., Liu, Q., He, X., and Wang, Z.: Wind Field Characteristics of Complex Terrain Based on Experimental and Numerical Investigation, Applied Sciences, 12, https://doi.org/10.3390/app12105124, 2022.