Received: 9 October 2020 | Revised: 9 March 2021 | Accepted: 12 March 2021 | Published on: 29 March 2021

DOI: 10.1002/joc.7103

RESEARCH ARTICLE

International Journal of Climatology RMetS

Comparison of ERA5 surface wind speed climatologies over Europe with observations from the HadISD dataset

María O. Molina[1] | Claudia Gutiérrez[1] | Enrique Sánchez[2]

[1]Instituto de Ciencias Ambientales, University of Castilla-La Mancha (UCLM), Toledo, Spain

[2]Faculty of Environmental Sciences and Biochemistry, University of Castilla-La Mancha (UCLM), Toledo, Spain

Correspondence
María O. Molina, Instituto de Ciencias Ambientales, University of Castilla-La Mancha (UCLM), Avenida Carlos III s/n, Toledo 45071, Spain.
Email: mofelia.molina@uclm.es

Funding information
European Social Fund (ESF)

Abstract

Understanding space–time features of wind speed is of high interest in meteorology and several applied sciences. Accurate wind speed measurements in combination with reliable gridded products, such as reanalyses, are needed to address the main characteristics of the wind field. Hourly 10 m wind speed from the European Centre for Medium-Range Weather Forecasts (ECMWF) latest reanalysis (ERA5) is compared with HadISD wind observations from 245 stations across Europe. Averaged ERA5 hourly data is able to reproduce the annual cycle of monthly wind speed in Europe. ERA5 presents slightly larger (shorter) monthly medians in winter (summer) than observations. ERA5 is compared against observations for each station using a frequency distribution-based score (score, from 0 to 1). Most of the stations exhibit hourly scores ranging from 0.8 to 0.9, indicating that ERA5 is able to reproduce the wind speed spectrum range, from light to strong relative frequencies, for any location over Europe. Ranges of mean values, variability, distribution function parameters and high or low wind thresholds frequencies are shown for this ensemble of European stations, allowing for an overall description of wind features. Generally, there is no clear relationship between scores and the variables analysed. The correlation and scores between ERA5 and HadISD is even further increased at longer time frequencies (6–24 hourly), together with centred root-mean-square error (RMSE) and standard deviation decreases. Hourly wind data from ERA5 reanalysis is, despite some shortcomings, valuable information to perform further detailed studies with a regular spatial and time wind distribution, from the climatological or renewable energy perspectives, for example.

KEYWORDS

ERA5, HadISD, observations, score, validation, wind speed

1 | INTRODUCTION

Wind speed is an atmospheric variable that fluctuates over a wide range of time and spatial scales, from the planetary scale, or regional winds, down to smaller local turbulent eddies (from thousands of km to cm or less). Spatial and time scales are closely related, from sub-hourly to decadal values. Therefore, gaining insight into wind dynamics and its associated space–time variability is of high interest for several branches of meteorology, climate, applied sciences and engineering (Pryor *et al*., 2006; Andersson *et al*., 2007; McVicar *et al*., 2012; Lorente-Plazas *et al*., 2015a; Bett and Thornton, 2016; Staffell and Pfenninger, 2016; Bett *et al*., 2017; Jones *et al*., 2017).

Although the lack of reliable measured data in many areas of the world (Staffell and Pfenninger, 2016; Harris *et al*., 2020) is reflected in the limited amount of literature about wind speed, when compared with other climate variables (typically temperature and precipitation), several studies have analysed wind speed variability over different continents (Vautard *et al*., 2010; Wan *et al*., 2010; Kim and Paik, 2015; Hansen *et al*., 2018; Gruber *et al*., 2019; Tian *et al*., 2019; Zeng *et al*., 2019), and over Europe in particular (Azorin-Molina *et al*., 2014; Lorente-Plazas *et al*., 2015b), using different observational data sources and methodologies. In the classical European wind Atlas (Troen and Petersen, 1989), it is observed that wind regimes over Europe mainly depend on the different regional climates and sea-land distributions. One example of this complexity is that the North Sea region presents higher wind speed variability than the rest of Europe (Bett *et al*., 2013) and that the spatial variability can be related to local or regional aspects. Another example of regional aspects involved is the Iberian Peninsula wind patterns, which is characterised by its complex orography (Lorente-Plazas *et al*., 2015a; 2015b) and where it has been studied how wind variability can be related to general circulation patterns such as NAO (Jerez *et al*., 2013; Jerez and Trigo, 2013; Lorente-Plazas *et al*., 2015a; 2015b), capable to describe wind storms over Central Europe (Guenard *et al*., 2005; Donat *et al*., 2010; Brayshaw *et al*., 2011; Cortesi *et al*., 2019). Generally over Europe, yearly temporal variability is characterised by an annual cycle with maximum seasonal wind speed in winter, and a daily cycle in which sunny hours are windier than night ones (Pryor *et al*., 2006; Sinden, 2007; Kiss *et al*., 2009; Bett and Thornton, 2016; Marcos *et al*., 2019).

Several wind data sources have been used for wind speed studies. Highly scattered (in time and space) wind speed measurements, from public and private institutions, have been obtained along many years in meteorological stations at airports or elsewhere. In terms of gridded data, observed temperature and precipitation climatologies (covering several decades, on daily or monthly scales) are available at relatively high resolution (50 km, 25 km), globally or at continental scales. This is the case of the well-known worldwide Climatic Research Unit (CRU) Time-Series (TS) dataset (CRU [Harris *et al*., 2020], http://www.cru.uea.ac.uk/), with monthly resolution, or the E-OBS dataset (Cornes *et al*., 2018) from the European Climate Assessment & Dataset project (ECA [Klok and Klein Tank, 2009], https://www.ecad.eu/), on daily scales, over Europe. The spatial density of wind speed observations is too sparse, making it difficult to apply accurate gridding procedures (Harris *et al*., 2020). The reasons behind this are the strong local character of wind and its high dependence on orography (Stohl *et al*., 1995; Brinckmann *et al*., 2016). Besides, freely available high quality wind speed observations (usually from National Meteorological Services networks, [Thorne *et al*., 2017; Dunn *et al*., 2019]) are more difficult to find than temperature or precipitation data. CRU TS database then just offers a monthly 1961–1990 wind climatology (New *et al*., 2000), while ECA has no gridded wind fields available. Tall towers have recently started to be an additional source of wind data, which is regularly collected in wind farms by private energy companies and usually with limited public access (Ramon *et al*., 2020).

In spite of this, the lack of good enough spatial wind data networks resolution can be partly compensated with the existence of modelling capabilities. Reanalyses products and climate models (global or regional) obtain physically coherent atmospheric fields, and wind in particular, in a continuous domain. An example of how these procedures are applied to wind features is the New European Wind Atlas (Petersen *et al*., 2013), which provides a high-resolution wind field, using the Weather Research and Forecasting regional model. For evaluation purposes, reanalyses databases are currently widely used as a good option to provide a gridded dataset of surface wind and many other atmospheric magnitudes, to complement the available wind observations. They obtain long-term climate data sets based on the assimilation of all the available observations from different sources and solving the main atmospheric evolution equations, with the aim to represent past or current climate on a regular grid (Compo *et al*., 2011; Dee *et al*., 2011). Wind values from reanalyses have been used in evaluation of marine surface wind fields (Swail and Cox, 2000), investigate the multi-decadal trends of wind speed (Vautard *et al*., 2010; Kaiser-Weiss *et al*., 2015; Torralba *et al*., 2017; Zeng *et al*., 2019) or validate simulations when there is an absence of observational wind data collected (Brands *et al*., 2013; Staffell and Pfenninger, 2016). Reanalyses products are also widely used as perfect boundary conditions to force regional climate models dynamical downscaling experiments (Kotlarski *et al*., 2014; Jacob *et al*., 2020), being wind fields a commonly needed input. Also, global reanalyses products are used as boundary conditions to produce regional reanalyses, which are benefited from additional data assimilation in the regional domain (Kaspar *et al*., 2020).

In this work, we make use of the newest reanalysis from the European Centre for Medium-Range Weather

Forecasts (ECMWF, https:// www.ecmwf.int/): ERA5. ERA5 is starting to be validated, being compared with previous reanalyses and with observations at local or regional scales, mostly over Europe (Belmonte Rivas and Stoffelen, 2019; Piasecki *et al*., 2019; Ramon *et al*., 2019; Cucchi *et al*., 2020; ERA5 data documentation, 2020; Jourdier, 2020; Minola *et al*., 2020). Their hourly frequency allows for an analysis of daily cycles with respect to observations (Jourdier, 2020). Intercomparisons with MERRA-2 reanalysis (Olauson, 2018) or an ensemble of them (Ramon *et al*., 2019; Hersbach *et al*., 2020) have also been performed. Some improvements have already been highlighted, related to the former ERA-Interim (Dee *et al*., 2011) reanalysis, although several underestimations against observational values remain.

The objective of this work is twofold. Firstly, to describe the wind behaviour and its distribution among the different regions of Europe, taking advantage of the availability of hourly surface wind measurements from the HadISD dataset. For that purpose, we make use of the wind speed frequency distribution and inspect several statistical computations, focusing both on temporal (seasonal cycle, time variability and correlations) and non-temporal (means or extreme values) aspects, from hourly to daily scales. The analysis is made for each meteorological station for the whole 40 years period (1979-2018). Then, the second objective is to validate ERA5 wind field reanalysis against such observed statistics, including scores to quantify the degree of agreement with the available observations. Some verifications of ERA5 winds have already been published, here these analyses are complemented using a database with hourly observations that covers the entire European domain. The study of individual sites at different time scales makes the results a step forward compared to other reanalyses and more local studies. If the proposed assessment is able to show the capability of ERA5 reanalysis to describe surface wind fields, several applications could be considered then for further studies, from climate models validation to energy resources analysis. The hourly frequency availability, the regular and full spatial distribution, and the consistency with other atmospheric variables inside ERA5, would help to advance in the understanding of wind variability and their processes and mechanisms.

2 | DATA

2.1 | Meteorological observations

The meteorological observations used for this study are provided by the Met Office Hadley Centre's Integrated Surface Database, HadISD. HadISD, version 3.1.0.2019f (Smith *et al*., 2011; Dunn *et al*., 2012, 2016, 2019) is a global subdaily dataset based on the ISD dataset from NOAA's NCDC distributed by the UK Met Office Hadley Centre (https://www.metoffice.gov.uk/hadobs/hadisd/). It offers freely and automatically quality-controlled hourly estimates data of, among other variables, wind speed data from global weather stations. The quality check controls procedures such as duplicate check, distributional gap check and neighbour outliers, to remove bad data and keep the extremes (Dunn *et al*., 2012). The station selection and merging procedures result in a database with 8,139 stations worldwide, most of them concentrated in North America and Europe (2,786 stations), which is updated annually and now covers the period 1931–2020, inclusive (Dunn, 2019).

The stations selected for this study are those with valid values in at least the 90% hourly of time steps for the period 1979–2018, a more strict criteria than the used in previous studies (Azorin-Molina *et al*., 2014; Lorente-Plazas *et al*., 2015a; 2015b; Ramon *et al*., 2020). The use of a very restrictive threshold increases the reliability of the results but, in return, reduces the spatial representation required to analyse spatial variability in depth. The choice of threshold is somewhat arbitrary, but the number of stations obtained presents a reasonable balance between both aspects. This results in 245 hourly averaged wind speed series at 10 m height, from the initial available stations across Europe (see Figure 1 and Table S1 for more details about meteorological stations). The meteorological stations are not homogeneously distributed over Europe. The highest station density is seen in the United Kingdom, Belgium, the Netherlands and Denmark. Due to the orographic, land and coastal complexity of Europe, the representativity for the whole of Europe is therefore limited. It is important to point out that, despite the limited spatial representation of HadISD observations, it is one of the databases with the largest number of collected available observations for Europe. The proposed analysis for the reanalysis data will be focused on the grid points corresponding to the locations of the observations.

2.2 | Reanalysis data: ERA5

ERA5 (Copernicus Climate Change Service, 2017), DOI:10.24381/cds.adbb2d47, is the fifth generation reanalysis developed at the ECMWF. It provides hourly estimates for a large amount of atmospheric, ocean wave and land surface variables. The information from observations is extracted from many satellite or conventional instruments (Hersbach *et al*., 2020), but surface wind

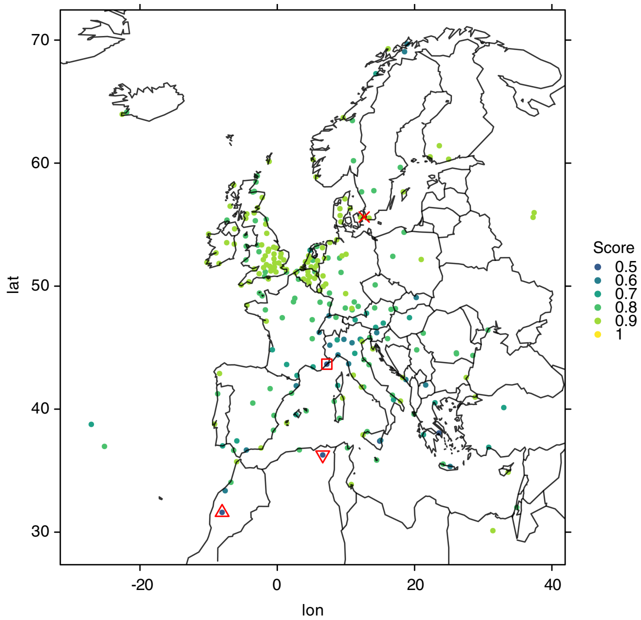

FIGURE 1 Situation in the map of each meteorological station. Colour scale represents the score of each location. The frequency distribution of the stations marked with a red symbol are represented in Figure 5: station 88 (× symbol), station 132 (square), station 240 (triangle) and station 242 (upside-down triangle) [Colour figure can be viewed at wileyonlinelibrary.com]

from land stations are not assimilated. The atmospheric component is interpolated to 37 pressure levels from the surface up to 1 Pa. It has a horizontal grid resolution 0.25°, corresponding to around 31 km. This represents a step forward with respect to the previous ERA-40 (Uppala *et al.*, 2005) and ERA-Interim (Dee *et al.*, 2011) reanalyses. ERA-40, covered 1958–2001 period with a spatial resolution of ~125 km, and ERA-Interim covered 1979 to present with a resolution of 79 km. Currently, ERA5 is freely available through the EU-funded Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/) from 1950 to present. More information about ERA5 characteristics can be found in (Hersbach *et al.*, 2019).

Components of the horizontal wind field at 10 m above ground, u (eastward) and v (northward), for the 1979–2018 period are used to calculate wind module as $\sqrt{u^2+v^2}$ for each hour at the cells where observational data is located. The method used to select the data from ERA5 is by extracting the values for the cells in which the specific points (latitude and longitude) in which the observation stations are located fall. Using this technique, the same cell could correspond to different stations. In this case, almost all stations correspond to different locations except in two cases. Time steps with no observational data are not considered also for ERA5 data.

3 | METHODS

3.1 | Overall statistical evaluation

Monthly averages of hourly data (1979–2018) are computed for each meteorological station and cell from the reanalysis, obtaining the annual cycle for both databases, and are compared through a Student-t test. Some previous studies have indicated a poor representation of the observed hourly variability of wind speed at individual locations by reanalyses, and better ones at 6-hourly or longer (Cannon *et al.*, 2015; Kaspar *et al.*, 2020). Here, hourly, 6- and 24-hourly time frequencies (usually used scales on previous studies) are assessed to study this behaviour. The Pearson's correlation coefficient, centred root-mean-square error (CRMSE) and standard deviation (*SD*) between ERA5 and HadISD data are computed at each location for the period 1979–2018 at those three temporal scales, to analyse the degree of agreement of ERA5 with respect to HadISD observations. CRMSE is defined as follows:

$$CRMSE=\sqrt{\frac{\sum_{i=1}^{n}\left((r_i-\bar{r})-(o_i-\bar{o})\right)^2}{N}} \tag{1}$$

being N the number of observations, r reanalysis values and o observational data. The mean of the time series is represented by $\bar{r}$ and $\bar{o}$ for the reanalysis and observations, respectively.

These statistics are represented by a Taylor's diagram (Taylor, 2001), a graph that shows how the spatiotemporal correlation, the CRMSE and the ratio of the variances of reanalysis match with observations. The CRMSE of the reanalysis appears as the radial distance from the position of a perfect model. The Pearson correlation coefficient is represented by the exterior arc and the ratio of standard deviations between the model and the observation is represented by the interior arcs. In the normalised diagram, the CRMSE and the two standard deviations of each time series are normalised by the standard deviation of the corresponding observed field and so, the perfect model always lies at standard deviation = 1 and CRMSE = 0. This allows the comparison of data sets with different variances in the same diagram.

3.2 | Wind parameters

Several specific wind parameters are also calculated for each station for the whole period, merging scores and a matrix of other magnitudes that are likely to be relevant to describe wind features, with the aim to study their combined relations. The proposed statistics are the following:

- The coefficient of variation (*CV*), as an adimensional measure of wind variability, using the standard deviation of wind normalised with the mean wind, to make numbers more comparable from any location. Their mathematical formula is:

$$CV = \frac{\sigma}{\mu} \tag{2}$$

where σ is the standard deviation and μ the mean of the wind speed.

- Cut-in and cut-out threshold rates. Cut-in and cut-out rates are the thresholds in which wind turbines operate. Cut-in is the wind speed at which the turbine starts to generate electricity and cut-out is the velocity at which production needs to be shut down to prevent damage on wind turbines (Manwell *et al.*, 2010). These values depend on the wind turbine, here 3 and 25 m/s, respectively are considered as are commonly used values for standard wind turbines. In this way, correlations are performed using the percentage of hours below (above) these fixed thresholds for each meteorological station.
- Elevation. As surface wind climatology is strongly dependent on topographic features of the terrain (Jiménez *et al.*, 2008), the elevation parameter, together with all previously mentioned parameters, can be used to evaluate if the geographical situation and the characteristics of wind distribution in each observational station can affect the score results obtained (see Section 3.3 for score definition).

All the computational processes have been carried out with the R free software (R Core Team, 2019). The specific package used for the Taylor diagrams is plotrix-package (Lemon, 2006).

3.3 | Evaluation method: score

Perkins *et al.* (2007, 2013) developed a score method to evaluate temperature and precipitation results from climate models against available observations. This approach has been already used in wind speed studies (Lorente-Plazas *et al.*, 2015a; 2015b; Gómez *et al.*, 2016; González *et al.*, 2017; Santos *et al.*, 2018; Nogueira *et al.*, 2019), mostly comparing modelling results against reanalysis. The method is based on the amount of overlap between the frequency distribution of the wind speed for observations (the reference data) and ERA5 (looking for their grade of matching related to the reference data). If ERA5 compares observed values perfectly, both frequency distribution areas will coincide, and so the score will be 1. Any other situation will lead to values between 0 and 1. The formal expression is:

$$\text{score} = \sum_{1}^{n} \min(Z_m, Z_o) \tag{3}$$

where n is the number of bins used to calculate the frequency distribution for a given location (here, 0.5 m/s has been used as bin size), Z_m is the fraction (or frequency) of values in a given bin from the reanalysis and Z_o is the fraction (or frequency) of values in a given bin from the observed data. The sum of all Z_m is 1, and the same for Z_o sum.

Therefore, the frequency distribution of all wind observation values (hourly data) is calculated at each location together with the same computation for the ERA5 corresponding cell. A score between 0 (poor score) and 1 (good score) is obtained, as a measure of the common area between the two frequency distributions. Some remarks about this method of comparison is that the frequency distribution of the observations refer to a specific location, while the frequency distribution of the reanalysis represents an area of 31 km. Also, a misrepresentation of extremes could be obtained even with good scores (Perkins *et al.*, 2013), as the area around tails is quite small, and so large differences there would not be reflected on score.

4 | RESULTS

4.1 | Statistical comparison of ERA5 against observations

A first overall evaluation of the spatial variability of wind is made through the representation of the mean annual cycle of the 245 locations (meteorological stations and corresponding reanalysis grid cells) across Europe in Figure 2. Each box represents the monthly mean data distribution of hourly data at station location in the period 1979–2018. HadISD observations (green boxes) show the annual cycle of European wind speed, with maximum values in winter months and a minimum in summer, in agreement with previous analysis made with reanalyses covering the whole Europe or just some regions (Pryor *et al.*, 2006; Sinden, 2007; Kiss and Jánosi, 2008; Bett and

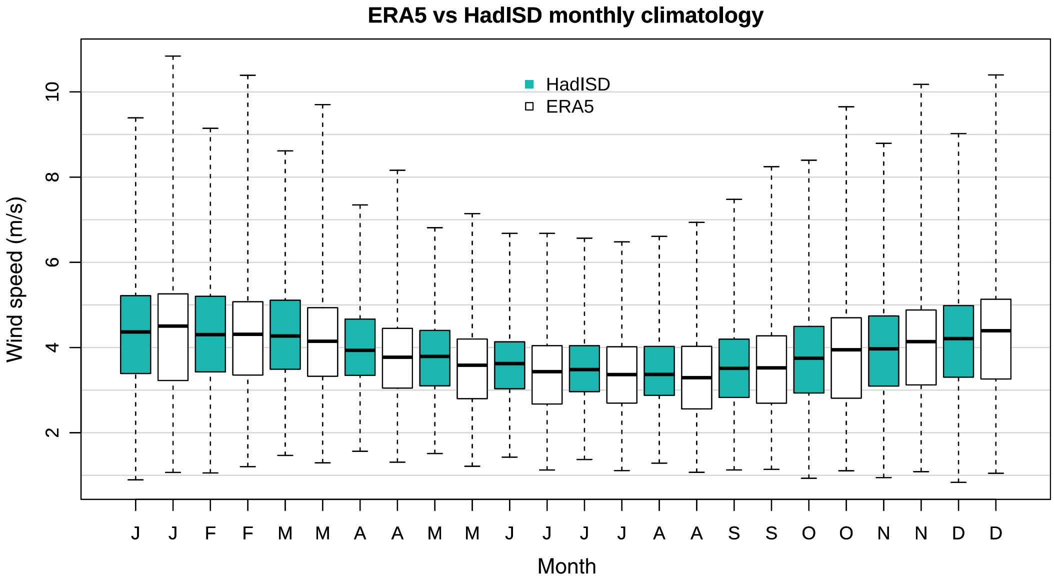

FIGURE 2 Monthly wind speed (1979–2018) of HadISD meteorological stations (green colour) and ERA5 cells (transparent colour) box plots. Limits of the boxes represent the locations in the 25th and 75th percentile, and the black line in the middle represents the 50*th* percentile. The upper whisker is located at the maximum value, whereas the lower whisker is located at the minimum value [Colour figure can be viewed at wileyonlinelibrary.com]

Thornton, 2016; Gómez *et al*., 2016; Minola *et al*., 2020). Winds in the summer months present a median value (p50) of around 3.5 m/s, whereas the median wind speed in the winter months is around 4–4.5 m/s. The month-to-month comparison of ERA5 reanalysis with observations indicate a similar representation of the annual cycle. However, ERA5 presents median values that are slightly larger in winter and slightly smaller in summer. These monthly differences are not statistically significant (*t*-Student) with a confidence interval (*p*-value) of 0.05. In terms of the wind speed annual cycle representation by the reanalysis compared with observations, there is a mixture of results in previous literature for different reanalyses products in different areas of the globe. Thus, Ramon *et al*. (2019) point that reanalyses in the Iberian Peninsula tend to show weaker seasonal mean winds than observed, Decker *et al*., (2012) reveals that reanalyses products tend to overestimate the monthly wind variability in the Northern Hemisphere, and Minola *et al*. (2020) shows an overestimation of ERA5 in the coastal and inland regions and an underestimation of the seasonal cycle in the mountainous regions of Sweden in the 2013–2017 period.

Upper extreme wind values (whisker ends) are generally larger for ERA5 than for observations, which means that the reanalysis gives a wider range of monthly wind values for the analysed period. This is especially the case for autumn-winter months (DJF and SON), when differences between extreme values of reanalysis and observations are more noticeable. ERA5 seems to overestimate higher winds more clearly in those months. Cold months (SON, DJF) also exhibit a clear asymmetry in ERA5 boxes, with lower percentiles width (25–50) larger than higher ones (50–75), which is not seen in observations (green) boxes, although interquartile range (percentile 75 vs. percentile 25) is quite similar for both observations and reanalysis. Summer ERA5 boxes are, on the contrary, symmetric, although lower box values (percentile 25) seem to give smaller values than observational ones. Differences among locations seem to be larger in autumn-winter months, as the interquartile range and whiskers are bigger in DJF and SON, in both series. In Minola *et al*. (2020), ERA5 shows discrepancies in its capability to represent different annual cycle patterns, from coastal to inland or high elevation stations. Here, the average of all the 245 locations in Europe is presented and an overall well agreement is found, however, due to the strong local character of wind, discrepancies in the monthly annual cycle of wind between observations and reanalysis should be inspected locally in order to find differences related to each station characteristics.

With that aim of looking into the differences among locations, Figure 3 presents a scatterplot of wind speed 40-year averaged monthly means. Dots represent each ERA5-cells/stations monthly pairs, with colours to distinguish each month. A close overall correspondence between the two data sets is obtained, since most of the points are close to the diagonal line. This relation is stronger in summer months (red dots) than in winter ones (blue dots), which present larger values (more than

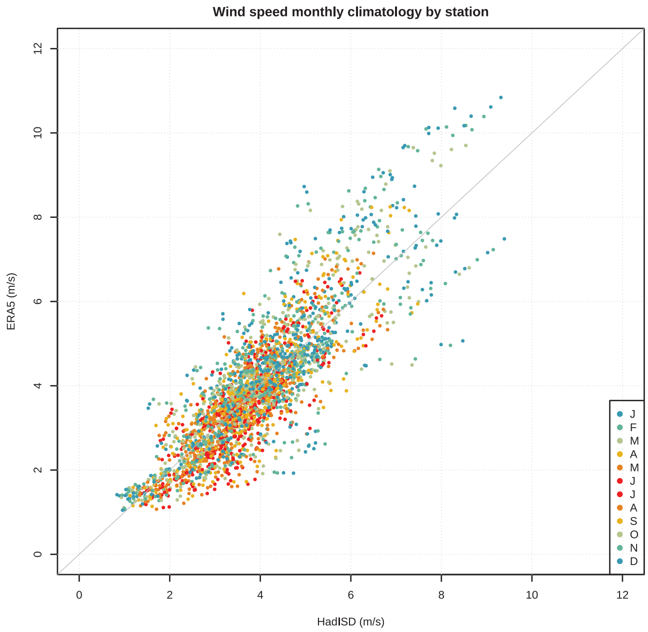

FIGURE 3 Comparison between the wind speed monthly climatology (1979–2018) of the 245 HadISD meteorological stations across Europe and ERA5 reanalysis data. The diagonal line indicates a perfect fit [Colour figure can be viewed at wileyonlinelibrary.com]

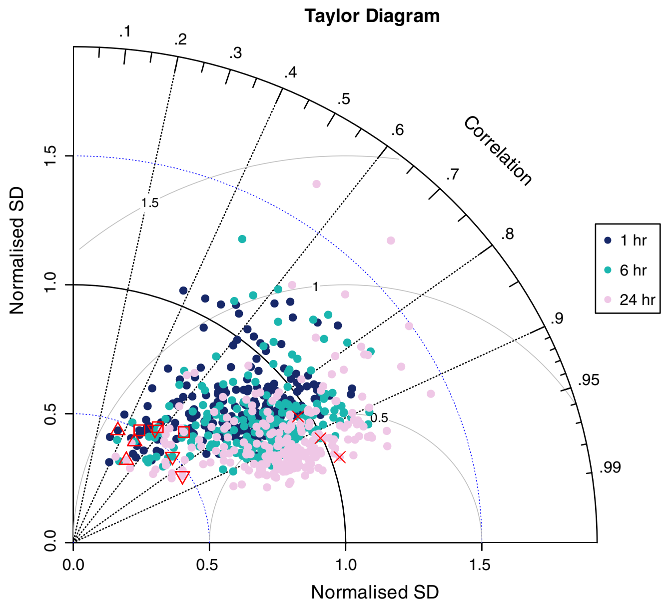

FIGURE 4 Normalised Taylor diagram of the ERA5 reanalysis and HadISD observations comparing hourly (blue points), 6-hourly (green points) and 24-hourly (pink points) data of each location studied for the period 1979–2018. The Pearson correlation coefficient is represented by the exterior arc and the ratio of standard deviations between the model and the observation is represented by the interior arcs (blue semicircles). The CRMSE of the reanalysis appears as the radial distance from the position of the perfect model (grey semicircles). In that case, reanalysis will represent better observations data when points will be nearer to the perfect model: HadISD SD = 1, CRMSE = 0 and Pearson's correlation = 1. The stations marked with a red symbol represent station 88 with the best score (× symbol), and the stations 132 (square), 240 (triangle) and 242 (upside-down triangle) with the lowest scores [Colour figure can be viewed at wileyonlinelibrary.com]

6 m/s) and higher biases, although the relative error is similar all months (not shown). Both results are consistent with the averaged (to all the locations) result of the Figure 2. Related to larger wind values, it is important to notice that some authors (Ramon *et al.*, 2019) claim that reanalyses have difficulties to accurately represent those high winter wind speeds, since they are the result of not a specific location, as it is the case of observational data, but of a larger area (a grid cell of several km) and so, some smoothing of wind variability could be obtained in that area.

Previous studies (Cannon *et al.*, 2015) indicate a poor representation of the observed hourly variability of wind speed at individual locations by MERRA reanalysis. To evaluate whether ERA5 is able to reproduce the temporal variability of wind for these time scales, Figure 4 shows ERA5 reanalysis compared with HadISD observations for hourly, 6- and 24-hr averaged intervals, comparing Pearson's correlation, CRMSE and standard deviation into a Taylor diagram. The closer CRMSE to 0 and SD and correlation to 1, the better ERA5 wind speed time series will be to observational data. The results reveal important differences comparing hourly against daily statistics. Hourly data has a coefficient of correlation that varies from 0.6 to 0.85 for most of the locations, although some points are found with low correlation values down to 0.3, and a CRMSE higher than 0.5 in all data points. On the other hand, the 24-hourly data present higher correlation coefficient, with most of the points between 0.9 and 0.95, together with smaller CRMSE, below 0.5, and standard deviation closer to the observations than for hourly results. For the 6-hourly averaged intervals, most of the data have a correlation about 0.8–0.9, which is also better than the hourly scales correlation, with also lower CRMSE values (around 0.5) and smaller standard deviations. These results are in accordance with other studies (Decker *et al.*, 2012; Cannon *et al.*, 2015; Kaiser-Weiss *et al.*, 2015; Rose and Apt, 2016; Minola *et al.*, 2020), where wind speeds from different reanalyses products are compared against flux tower or meteorological data in different areas of the globe. But in contrast with others (Toledo *et al.*, 2015; Coburn, 2019), probably mainly related to the studied area. Coburn (2019) in the upper Midwest of the United States, showed that correlations between reanalyses and observations improve with higher temporal reanalysis resolution (monthly correlations are smaller than daily correlations). Note that all these previous studies have a more limited studied area compared to the work presented here, in which the whole European area is analysed and so it would include all those different and mixed regional behaviours.

Seasonal Taylor diagrams have also been computed (Figure S1) with the aim of seeing if sub-yearly scales could affect the obtained results. The analysis reflects a similar picture for any season between ERA5 and HadISD in terms of correlation, CRMSE and SD to what is observed in Figure 4 on an annual basis. If mean bias error or RMSE are computed, absolute biases range from ±2 m/s (see Figure S2), with a relation with time frequency and better and worse represented stations that is consistent with the normalized Taylor diagrams seen in Figure 4.

4.2 | Wind parameters analysis

In order to analyse different wind parameters and features, scatter plots of several combinations for the 245 stations are shown in Figure 6. Figures on the lower triangle present the matrix of scatter plot combinations (CV, percentage of data below/above the cut-in and cut-out threshold and station elevation). Scores (ERA5 vs. observations) are also shown with the colour palette range. On the upper panel triangle, a summary of these scatterplots is indicated with the Pearson correlation coefficient. This means that scatterplots that exhibit a cloud of points near to a straight line will show values near to 1 (blue colour) or −1 (red colour). The more spread the points are, the lower the correlation coefficient is obtained.

Scatter plot matrix figures allow for a complete analysis of combinations of wind statistics. Results show that CV is highly (0.76) correlated with low wind percentage of days (cut-in) but just slightly with strong wind ones (cut-out, 0.17). Locations where light winds are more frequent also tend to present larger CV, that is, more variable wind conditions, whereas a small amount of light winds seem to be related to more stable (less variable) wind values. The elevation of the meteorological station, as mentioned before, does not seem to play a relevant role on the obtained statistics, being the larger correlation (0.43) with cut-in fraction of data. That is, the higher the elevation, the more frequent low wind data are measured, although correlation is not very high. Finally, cut-in (with 10–79% of hours below 3 m/s) and cut-out (from 0.5 to 12.8% of the hours above 25 m/s) frequencies do not seem to be highly correlated (0.35), so larger light wind amounts are not only related to a smaller amount of strong wind data.

The comparison of hourly (Figure 6) with 6- (Figure S4) and 24-hourly (Figure S5) time-scales, allows for an inspection of the effect of data frequency on the analysed parameters. Results indicate that, as time averaging increases, frequencies above the cut-out decrease and the cut-in frequencies increase: there is a larger percentage of single hours above the 25 m/s threshold than days that, on average, do exceed that value. The opposite occurs in the case of cut-in frequencies. It is also important to point out that magnitude does not seem to change linearly from hourly data to 24-hourly averages (Figure S5). Differences in the 6-hourly performance compared with hourly were already pointed by (Cannon *et al.*, 2015).

4.3 | ERA5 evaluation: score

The score for each station allows for an overall analysis of how close ERA5 is to observations in terms of the whole frequency distribution of the wind values (Perkins *et al.*, 2013), whereas the analyses of the previous section were focused on statistical evaluation that could help on the analysis of time variability and bias, through the correlation coefficients and the CRMSE. While correlations analyse the temporal evolution of wind speed, the score parameter evaluates the whole frequency distribution, where time aspect is not considered. Then, it could happen that for certain stations a good correlation (comparable evolution of both time series) and large and regular biases throughout the series are presented at once, leading to bad scores. This kind of mixed behaviour could be detected with the proposed analysis. However, this extreme situation is not the case in any of the analysed stations. A first overview of the scores, computed from hourly data at each of the 245 locations, is shown on the map of Europe in Figure 1, with a colour gradient scale from 0.5 to 1. The 86.53% of the locations have scores higher than 0.7, which indicates the general ability of ERA5 to reproduce the observed frequency distribution. Scores do not show any clear geographical dependence, although many locations with scores below 0.7 are found close to mountainous areas around the Alps. Highest scores (almost 1) are located over Denmark, and lowest ones of 0.5 more frequently in the southern Mediterranean regions. Locations with higher scores tend to correspond to a higher correlation coefficient and lower CRMSE and SD values over the Taylor diagram (Figure 4), which is consistent with the performance of the Perkins scoring method, although the statistics do not analyse exactly the same features. If mean biases are computed (Figure S3), it can be seen that they are generally related to smaller scores, with absolute biases up to 2 m/s when scores are around 0.5, and less than 1 m/s when they are 0.7 or better.

To illustrate specific cases with small and, more interestingly, large differences between wind distributions, Figure 5 presents hourly frequency distribution performance of the best (a: station 88) and some of the worst (b: station 132, c: station 240 and d: station 242) stations. In the best one, reanalysis and observation distributions perfectly fits. For the poorer scored stations, it is seen that the reanalysis tends to largely overestimate lower wind speed frequencies (0–3 m/s range) and underestimate the higher ones (4–8 m/s range). This reanalysis overestimation of light winds is a common and known feature, already seen in previous studies (Larsén and Mann, 2009; Carvalho *et al.*, 2014; Cannon *et al.*, 2015). Both overestimation of light wind frequency and underestimation of strong wind frequency, have been also seen in previous works when models are used (Frank, 2001; Stopa and Cheung, 2014; Schewe *et al.*, 2019). This underestimation of extremes from model analysis can be related to the fact that solving the dynamics equations is applied to each point continuously, to grids of defined size. This fact can cause an inaccurate description of real elevation, especially over regions with complex orography, leading to some degree of smoothing (Niermann *et al.*, 2019). It is relevant to notice that the higher resolution of the reanalysis, the smaller this issue is expected to be. Also, parameterizations in models, which means a limitation in several ways to fully describe atmospheric mechanisms, can lead to an underestimation of the most extreme events (Larsén and Mann, 2009; Larsén *et al.*, 2013; Carvalho *et al.*, 2014; Cannon *et al.*, 2015). This idea is not contradictory with some overestimation of strong wind values seen on Figures 2 and 3. Firstly, only some specific months and stations present this behaviour. Secondly, on Figures 2 and 3 monthly extreme values are seen whereas here, direct hourly data are used to compute scores, and so time averaging and scales are quite different. Finally, strong winds deviations from ERA5 related to observations are almost not reflected on scores, due to the fact that tail frequencies contribute just slightly to the final number (Perkins *et al.*, 2013).

4.3.1 | Scores dependence on wind parameters

A more detailed analysis of the scores relation with other statistics that represent the whole 1979–2018 wind data series at each location, and the comparison of ERA5 and observational results are shown in Figure 6 with the aim to study the relevance of that wind parameters analysis in the scores obtained. Scores correlation with other parameters (as computed from hourly data) is shown in the first row of Figure 6. It is seen that scores are negatively correlated (near −0.4) with the coefficient of

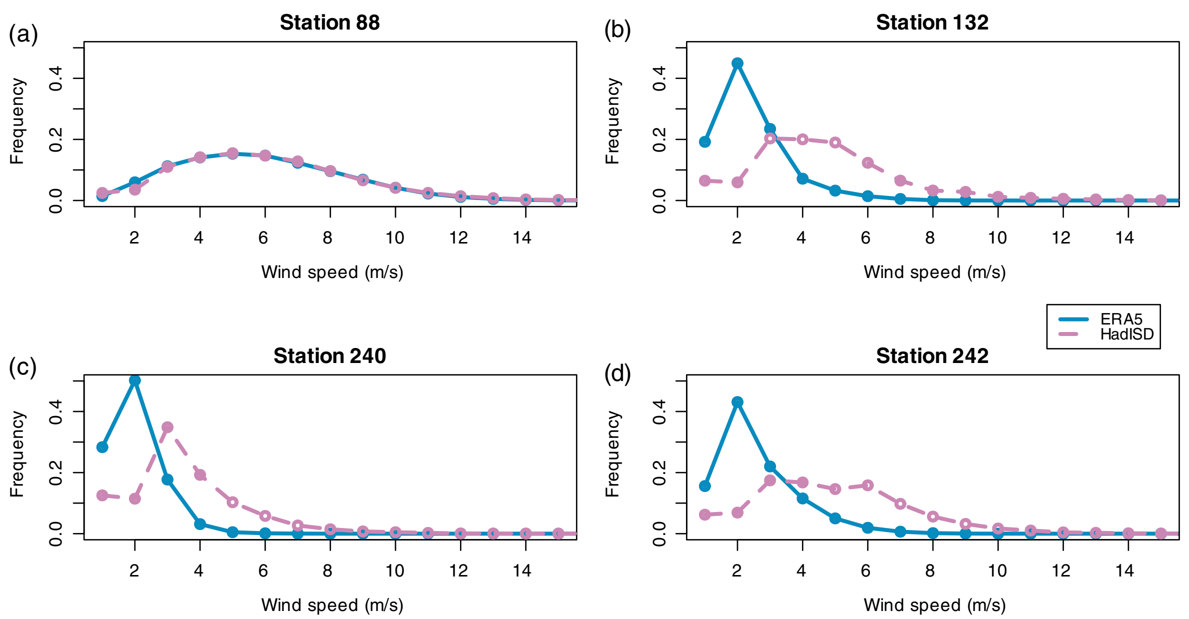

FIGURE 5 Frequency distribution of the hourly ERA5 reanalysis vs HadISD observations illustrating the total score in (a: station 88) the best score test (0.97) and (b: station 132, c: station 240 and d: station 242) the poorest score (0.45, 0.45 and 0.5, respectively). The location of each station can be seen marked with a red symbol in Figure 1: × symbol for station 88, square for 132, triangle for 240 and upside-down triangle for 242 [Colour figure can be viewed at wileyonlinelibrary.com]

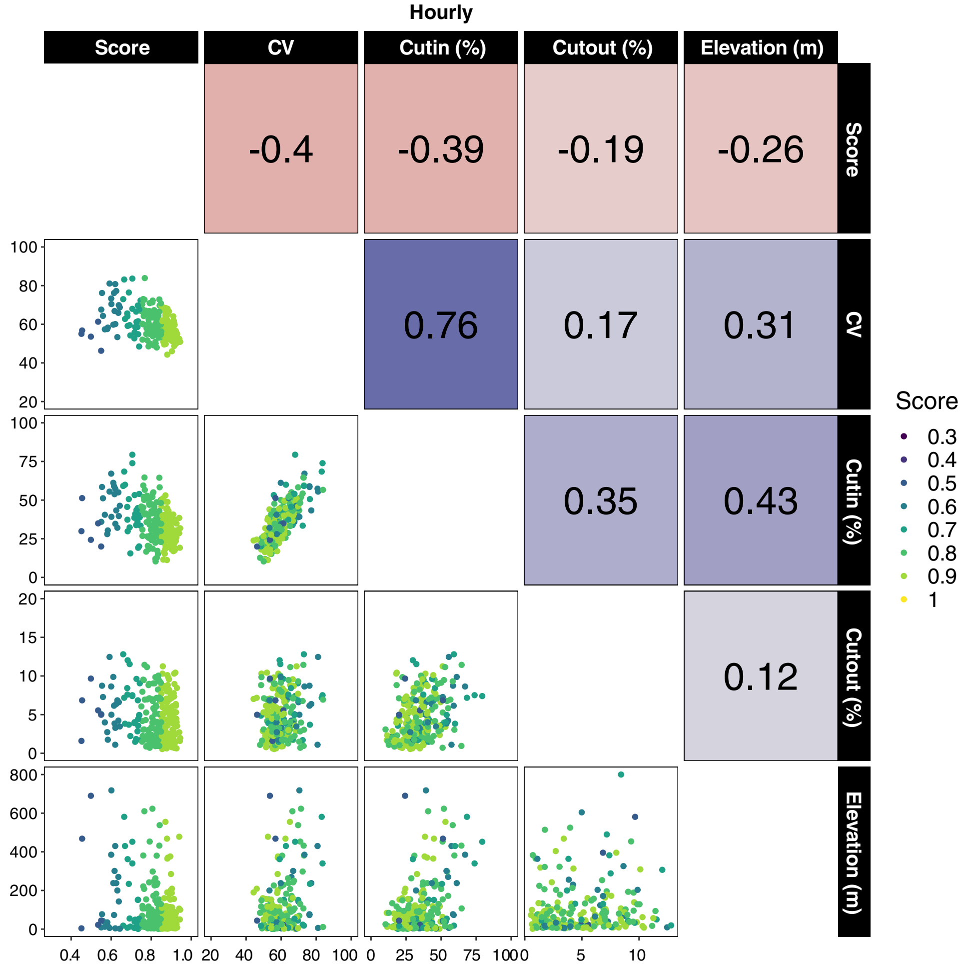

FIGURE 6 Scatter plot matrix of the HadISD hourly wind speed observations data of each location comparing: score (score), coefficient of variation (CV), frequency of values under the cut-in (3 m/s) and above the cut-out (25 m/s) rates and elevation of the meteorological station. Lower panels show scatter plots and upper panels the Pearson's correlation for each variable combination (from −1, red colour, to 1, blue colour). Dots make reference to meteorological stations and colours gradient ranges from lower (green) to higher (yellow) scores. That way, the agreement of ERA5 with observations is represented by the colour scale [Colour figure can be viewed at wileyonlinelibrary.com]

variation and cut-in rate. This negative correlation would mean that stations with smaller CV or with less frequent light winds (small cut-in fraction of data) would be related to better ERA5 performance, which could be reasonable, as the more spread or the more frequent light winds are present, the hardest is for reanalysis to accurately represent observed data, as partially described on previous statistics. However, there is no clear linear relationship and the large spread of the cloud points, which points out a low negative correlation, prevent from stating a robust conclusion. Frequency of data above the cut-out rate and elevation present even less clear correlation with scores (−0.19 and −0.26, respectively). Altitude then seems to be not very relevant related to better or worse scores. In Minola *et al.* (2020), differences in wind speed and gusts representation between ERA5 and observations were observed between the inland (elevation lower than 750 m above sea level) and mountain (elevation higher than 750 m) regions of Sweden, and they attributed it to differences in elevation and convection. Therefore, it seems that the station's elevation could play a role on the ERA5 performance but, in this overall analysis over Europe, we are not able to obtain any clear conclusion about that aspect.

In Kaspar *et al.* (2020), it is seen that the maximum peak correlation between observations and ERA5 10 m wind speed is reached at a weekly time scale. Here, higher scores with 6-hourly (Figure S4) and 24-hourly data (Figure S5) are also found compared with hourly ones on Figure 5. That overall better scores are consistent with what was already seen on Taylor diagrams (Figure 4 and S1). ERA5 seems to perform better compared with observational data for 6- and 24-hourly averaged data also when using scores. The reason behind can be the fact that small scale (hourly) discrepancies are smoothed (or compensated) when 24-hourly data are obtained. The smoothing effect of time averaging already indicated before, improves the ERA5 representation of extreme or highly variable situations. Previous literature has been mostly based on 24-hourly analysis, so similarities or differences would reinforce or complement those studies. Since ERA5 and HadISD have hourly data as their reference temporal resolution, the comparative analysis of different time frequencies could add value to previous studies with lower temporal and spatial resolutions and provide a new insight on wind features.

5 | CONCLUSIONS

Wind speed observations over Europe from 245 HadISD stations, and ERA5 corresponding reanalysis values, for the period 1979–2018 have been studied here. It is, to our knowledge, the largest observational data analysis combined with reanalysis information, covering the whole Europe. Hourly data availability, hardly found in previous studies, is a quite relevant feature for a better insight on wind analysis, which is even more interesting in combination with the newest ERA5 reanalysis, available also on hourly scales.

The analysis of HadISD wind observations, monthly average of all European stations, show time variability throughout the year, with a clear annual cycle of smaller winds in summer and more windy winter. ERA5 accurately reproduces that observed temporal pattern. A better agreement between ERA5 and HadISD stations monthly means is seen in summer months compared with winter ones. Hourly data time correlations for the whole 1979–2018 period between ERA5 reanalysis and observations are high, with values that range from 0.6 to 0.85 for most of the locations. Such correlations, together with small CRMSE and standard deviations, are further increased as data average changes from hourly to 6-hourly (0.8–0.9) and to 24-hourly (0.9–0.95) results. Different wind parameters have also been studied for each observations station. The ranges of values for the different magnitudes are 44.2–83.9 for CV, from 10 to 79% of hours under 3 m/s and from 0.5 to 12.8% of hours above 25 m/s, allowing an overall description of the wind features in Europe.

Wind speed frequency distributions are computed for each station and used to compare ERA5 results by means of a score. Most of the stations present 0.8–0.9 score, showing that ERA5 is able to reproduce observed wind speed range of values and its frequency over any location in Europe. Scores do not show any clear geographical dependence, although the larger number of higher scores stations are located in the United Kingdom, Belgium, the Netherlands and Denmark. The poorest score obtained is 0.5, with hourly data, and just for a few specific stations. Higher scores are obtained at 6-hourly and 24-hourly data when compared with hourly ones, in accordance with the correlations obtained with the Taylor diagram. This is an expected result, due to the smoothing effect of time averaging, that can mask extremes or compensate limitations of reanalysis, leading to such better performance with respect to observations.

In summary, results obtained here, in the comparison between ERA5 and HadISD, for several statistics and locations from hourly to daily time scales, as well as the annual cycle description, are reasonably good. This conclusion allows for ERA5 reanalysis to be used, for example, to validate climate models simulating present climate conditions, with dynamical spatial consistency. In particular, reanalyses outputs have also great potential and could be used when observational data are not available or the quality is not good enough, as well as for future studies related to the wind field such as wind energy production estimates or inspection of extreme wind events.

ACKNOWLEDGEMENTS

This work has been supported by a pre-doctoral fellowship by the University of Castilla-La Mancha (UCLM), financed by the European Social Fund (ESF). We are grateful to the reanalysis production centre ECMWF and the Copernicus programme for facilitating access to the reanalysis data, and the Met Office Hadley Centre for the observations data.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

The ERA5 data used in this article are freely available on the Copernicus Data Store (CDS) at https://cds.climate.copernicus.eu/cdsapp#/dataset/reanalysis-era5-single-levels?tab=overview (Hersbach *et al.*, 2018), and the HadISD dataset is available under a non-commercial government licence at https://www.metoffice.gov.uk/hadobs/hadisd/v310_2019f/index.html (Dunn *et al.*, 2016, 2019).

ORCID

María O. Molina https://orcid.org/0000-0003-0838-6959
Enrique Sánchez https://orcid.org/0000-0002-7720-4437

REFERENCES

Andersson, C., Langner, J. and Bergstroumm, R. (2007) Interannual variation and trends in air pollution over Europe due to climate variability during 1958–2001 simulated with a regional CTM coupled to the ERA40 reanalysis. *Tellus B: Chemical and Physical Meteorology*, 59(1), 77–98.

Azorin-Molina, C., Vicente-Serrano, S., McVicar, T., Jerez, S., Sanchez-Lorenzo, A., López-Moreno, J., Revuelto, J., Trigo, R., Lopez-Bustins, J. and Coelho, M.F. (2014) Homogenization and assessment of observed near-surface wind speed trends over Spain and Portugal, 1961–2011. *Journal of Climate*, 27, 3692–3712.

Belmonte Rivas, M. and Stoffelen, A. (2019) Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. *Ocean Science*, 15(3), 831–852.

Bett, P. and Thornton, H. (2016) The climatological relationships between wind and solar energy supply in Britain. *Renewable Energy*, 87, 96–110.

Bett, P.E., Thornton, H.E. and Clark, R.T. (2013). European wind variability over 140 yr, *arXiv preprint arXiv:1301.4032*.

Bett, P.E., Thornton, H.E. and Clark, R.T. (2017) Using the twentieth century reanalysis to assess climate variability for the European wind industry. *Theoretical and Applied Climatology*, 127(1–2), 61–80.

Brands, S., Herrera, S., Fernández, J. and Gutiérrez, J.M. (2013) How well do CMIP5 Earth System Models simulate present climate conditions in Europe and Africa? *Climate Dynamics*, 41 (3–4), 803–817.

Brayshaw, D.J., Troccoli, A., Fordham, R. and Methven, J. (2011) The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: a case study over the uk. *Renewable Energy*, 36(8), 2087–2096.

Brinckmann, S., Krähenmann, S. and Bissolli, P. (2016) High-resolution daily gridded data sets of air temperature and wind speed for Europe. *Earth System Science Data*, 8(2), 491–516.

Cannon, D., Brayshaw, D., Methven, J., Coker, P. and Lenaghan, D. (2015) Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. *Renewable Energy*, 75, 767–778.

Carvalho, D., Rocha, A., Gómez-Gesteira, M. and Silva Santos, C. (2014) WRF wind simulation and wind energy production estimates forced by different reanalyses: comparison with observed data for Portugal. *Applied Energy*, 117, 116–126.

Coburn, J.J. (2019) Assessing wind data from reanalyses for the upper midwest. *Journal of Applied Meteorology and Climatology*, 58(3), 429–446.

Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R.I., Grant, A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, Ø., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D. and Worley, S.J. (2011) The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137 (654), 1–28.

Copernicus Climate Change Service. (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J. and Jones, P.D. (2018) An ensemble version of the e-obs temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17), 9391–9409.

Cortesi, N., Torralba, V., González-Reviriego, N., Soret, A. and Doblas-Reyes, F.J. (2019) Characterization of European wind speed variability using weather regimes. *Climate Dynamics*, 53 (7–8), 4961–4976.

Cucchi, M., Weedon, G.P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H. and Buontempo, C. (2020) WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, 12(3), 2097–2120.

Decker, M., Brunke, M.A., Wang, Z., Sakaguchi, K., Zeng, X. and Bosilovich, M.G. (2012) Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF Using Flux Tower Observations. *Journal of Climate*, 25(6), 1916–1944.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597.

Donat, M.G., Leckebusch, G.C., Pinto, J.G. and Ulbrich, U. (2010) European storminess and associated circulation weather types: future changes deduced from a multi-model ensemble of GCM simulations. *Climate Research*, 42(1), 27–43.

Dunn, R. (2019). Hadisd version 3: monthly updates. Met Office Hadley Centre Technical Note, no. 103.

Dunn, R.J., Willett, K.M. and Parker, D.E. (2019) Changes in statistical distributions of sub-daily surface temperatures and wind speed. *Earth System Dynamics*, 10(4), 765–788.

Dunn, R.J., Willett, K.M., Thorne, P.W., Woolley, E.V., Durre, I., Dai, A., Parker, D.E. and Vose, R.E. (2012). HadISD: a quality-controlled global synoptic report database for selected variables at long-term stations from 1973–2011. arXiv preprint arXiv: 1210.7191.

Dunn, R.J.H., Willett, K.M., Parker, D.E. and Mitchell, L. (2016) Expanding HadISD: quality-controlled, sub-daily station data from 1931. *Geoscientific Instrumentation, Methods and Data Systems*, 5(2), 473–491.

ERA5 data documentation (2020), 'ERA5 data documentation: Known issues', Available at: https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation\#ERA5:datadocumentation-Knownissues [Accessed: 30th June 2020].

Frank, H.P. (2001). Extreme winds over Denmark from the NCEP/NCAR reanalysis.

Gómez, G., Cabos, W.D., Liguori, G., Sein, D., Lozano-Galeana, S., Fita, L., Fernández, J., Magariño, M.E., Jiménez-Guerrero, P., Montávez, J.P., Domínguez, M., Romera, R. and Gaertner, M.Á. (2016) Characterization of the wind speed variability and future change in the Iberian Peninsula and the Balearic Islands. *Wind Energy*, 19(7), 1223–1237.

González, A., Pérez, J.C., Díaz, J.P. and Expósito, F.J. (2017) Future projections of wind resource in a mountainous archipelago, Canary Islands. *Renewable Energy*, 104, 120–128.

Gruber, K., Klöckl, C., Regner, P., Baumgartner, J. and Schmidt, J. (2019) Assessing the Global Wind Atlas and local measurements for bias correction of wind power generation simulated from MERRA-2 in Brazil. *Energy*, 189, 116212.

Guenard, V., Drobinski, P., Caccia, J.-L., Campistron, B. and Bench, B. (2005) An observational study of the mesoscale mistral dynamics. *Boundary-Layer Meteorology*, 115(2), 263–288.

Hansen, J.C., Mortensen, N.G., Cronin, T., Hansen, B.O., Nielsen, M., Hahmann, A.N., Badger, J. and Larsén, X.G. (2018). Wind Atlas for South Africa (WASA)–Best practice guide for application of WASA.

Harris, I., Osborn, T., Jones, P. and Lister, D. (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7, 109.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2018) ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.adbb2d47 [Accessed: 1st March 2020].

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.N. (2020) The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049.

Hersbach, H., Bell, W., Berrisford, P., Horányi, A., Sabater, J.M., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C. and Dee, D. (2019). Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter*. pp. 17–24.

Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., Benestad, R., Boberg, F., Buonomo, E., Cardoso, R. M., Casanueva, A., Christensen, O.B., Christensen, J.H., Coppola, E., de Cruz, L., Davin, E.L., Dobler, A., Domínguez, M., Fealy, R., Fernandez, J., Gaertner, M.A., García-Díez, M., Giorgi, F., Gobiet, A., Goergen, K., Gómez-Navarro, J.J., Alemán, J.J.G., Gutiérrez, C., Gutiérrez, J.M., Güttler, I., Haensler, A., Halenka, T., Jerez, S., Jiménez-Guerrero, P., Jones, R.G., Keuler, K., Kjellström, E., Knist, S., Kotlarski, S., Maraun, D., van Meijgaard, E., Mercogliano, P., Montávez, J.P., Navarra, A., Nikulin, G., de Noblet-Ducoudré, N., Panitz, H.J., Pfeifer, S., Piazza, M., Pichelli, E., Pietikäinen, J.P., Prein, A.F., Preuschmann, S., Rechid, D., Rockel, B., Romera, R., Sánchez, E., Sieck, K., Soares, P.M.M., Somot, S., Srnec, L., Sørland, S.L., Termonia, P., Truhetz, H., Vautard, R., Warrach-Sagi, K. and Wulfmeyer, V. (2020) Regional climate downscaling over europe: perspectives from the euro-cordex community. *Regional Environmental Change*, 20(2), 1–20.

Jerez, S. and Trigo, R.M. (2013) Time-scale and extent at which large-scale circulation modes determine the wind and solar potential in the Iberian Peninsula. *Environmental Research Letters*, 8(4), 044035.

Jerez, S., Trigo, R.M., Vicente-Serrano, S.M., Pozo-Vázquez, D., Lorente-Plazas, R., Lorenzo-Lacruz, J., Santos-Alamillos, F. and Montávez, J.P. (2013) The impact of the North Atlantic oscillation on renewable energy resources in Southwestern Europe. *Journal of Applied Meteorology and Climatology*, 52(10), 2204–2225.

Jiménez, P.A., García-Bustamante, E., González-Rouco, J.F., Valero, F., Montávez, J.P. and Navarro, J. (2008) Surface wind regionalization in complex terrain. *Journal of Applied Meteorology and Climatology*, 47(1), 308–325.

Jones, P., Harpham, C., Troccoli, A., Gschwind, B., Ranchin, T., Wald, L., Goodess, C. and Dorling, S. (2017) Using ERA-Interim reanalysis for creating datasets of energy-relevant climate variables. *Earth System Science Data*, 9, 1–31.

Jourdier, B. (2020) 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.

Kaiser-Weiss, A., Kaspar, F., Heene, V., Borsche, M., Tan, D., Poli, P., Obregon, A. and Gregow, H. (2015) Comparison of regional and global reanalysis near-surface winds with station observations over Germany. *Advances in Science and Research*, 12(1), 187–198.

Kaspar, F., Niermann, D., Borsche, M., Fiedler, S., Keller, J., Potthast, R., Rösch, T., Spangehl, T. and Tinz, B. (2020) Regional atmospheric reanalysis activities at Deutscher Wetterdienst: review of evaluation results and application examples with a focus on renewable energy. *Advances in Science and Research*, 17, 115–128.

Kim, J. and Paik, K. (2015) Recent recovery of surface wind speed after decadal decrease: a focus on South Korea. *Climate Dynamics*, 45(5–6), 1699–1712.

Kiss, P. and Jánosi, I.M. (2008) Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe. *Energy Conversion and Management*, 49(8), 2142–2151.

Kiss, P., Varga, L. and Jánosi, I.M. (2009) Comparison of wind power estimates from the ECMWF reanalyses with direct turbine measurements. *Journal of Renewable and Sustainable Energy*, 1(3), 033105.

Klok, E. and Klein Tank, A. (2009) Updated and extended European dataset of daily climate observations. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 29(8), 1182–1191.

Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K. and Wulfmeyer, V. (2014) Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development*, 7 (4), 1297–1333.

Larsén, X. and Mann, J. (2009) Extreme winds from the ncep/ncar reanalysis data. *Wind Energy*, 12, 556–573.

Larsén, X.G., Badger, J., Hahmann, A.N. and Mortensen, N.G. (2013) The selective dynamical downscaling method for extreme-wind atlases. *Wind Energy*, 16(8), 1167–1182.

Lemon, J. (2006) Plotrix: a package in the red light district of R. *R-News*, 6(4), 8–12.

Lorente-Plazas, R., Montávez, J., Jerez, S., Gómez-Navarro, J., Jiménez-Guerrero, P. and Jiménez, P. (2015b) A 49 year hindcast of surface winds over the Iberian Peninsula. *International Journal of Climatology*, 35(10), 3007–3023.

Lorente-Plazas, R., Montávez, J., Jimenez, P., Jerez, S., Gómez-Navarro, J., García-Valero, J. and Jimenez-Guerrero, P. (2015a) Characterization of surface winds over the Iberian Peninsula. *International Journal of Climatology*, 35(6), 1007–1026.

Manwell, J.F., McGowan, J.G. and Rogers, A.L. (2010) *Wind Energy Explained: Theory, Design and Application*. Hoboken, New Jersey: John Wiley & Sons.

Marcos, R., González-Reviriego, N., Torralba, V., Soret, A. and Doblas-Reyes, F.J. (2019) Characterization of the near surface wind speed distribution at global scale: Era-interim reanalysis and ecmwf seasonal forecasting system 4. *Climate Dynamics*, 52 (5–6), 3307–3319.

McVicar, T.R., Roderick, M.L., Donohue, R.J., Li, L.T., Van Niel, T. G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N.M., Mescherskaya, A.V., Kruger, A.C., Rehman, S. and Dinpashoh, Y. (2012) Global review and synthesis of trends in observed terrestrial near-surface wind speeds: implications for evaporation. *Journal of Hydrology*, 416-417, 182–205.

Minola, L., Zhang, F., Azorin-Molina, C., Pirooz, A.S., Flay, R., Hersbach, H. and Chen, D. (2020) Near-surface mean and gust wind speeds in ERA5 across Sweden: towards an improved gust parametrization. *Climate Dynamics*, 55, 887–907.

New, M., Hulme, M. and Jones, P. (2000) Representing twentieth-century space–time climate variability. Part II: development of 1901–96 monthly grids of terrestrial surface climate. *Journal of Climate*, 13(13), 2217–2238.

Niermann, D., Borsche, M., Kaiser-Weiss, A.K. and Kaspar, F. (2019) Evaluating renewable-energy-relevant parameters of cosmo-rea6 by comparison with satellite data, station observations and other reanalyses. *Meteorologische Zeitschrift*, 28, 347–360.

Nogueira, M., Soares, P.M., Tomé, R. and Cardoso, R.M. (2019) High-resolution multi-model projections of onshore wind resources over Portugal under a changing climate. *Theoretical and Applied Climatology*, 136(1–2), 347–362.

Olauson, J. (2018) ERA5: The new champion of wind power modelling? *Renewable Energy*, 126, 322–331.

Perkins, S., Pitman, A., Holbrook, N. and McAneney, J. (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate*, 20(17), 4356–4376.

Perkins, S., Pitman, A. and Sisson, S. (2013) Systematic differences in future 20 year temperature extremes in AR4 model projections over Australia as a function of model skill. *International Journal of Climatology*, 33(5), 1153–1167.

Petersen, E.L., Troen, I., Jørgensen, H.E. and Mann, J. (2013) Are local wind power resources well estimated? *Environmental Research Letters*, 8(1), 011005.

Piasecki, A., Jurasz, J. and Kies, A. (2019) Measurements and reanalysis data on wind speed and solar irradiation from energy generation perspectives at several locations in poland. *SN Applied Sciences*, 1, 1–8.

Pryor, S., Barthelmie, R.J. and Schoof, J. (2006) Inter-annual variability of wind indices across Europe. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 9(1–2), 27–38.

R Core Team (2019), *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Ramon, J., Lledó, L., Pérez-Zanón, N., Soret, A. and Doblas-Reyes, F.J. (2020) The tall tower dataset: a unique initiative to boost wind energy research. *Earth System Science Data*, 12(1), 429–439.

Ramon, J., Lledo, L., Torralba, V., Soret, A. and Doblas-Reyes, F. (2019) What global reanalysis best represents near surface winds? *Quarterly Journal of the Royal Meteorological Society*, 145, 3236–3251.

Rose, S. and Apt, J. (2016) Quantifying sources of uncertainty in reanalysis derived wind speed. *Renewable Energy*, 94, 157–165.

Santos, F., Gómez-Gesteira, M., de Castro, M., Añel, J., Carvalho, D., Costoya, X. and Dias, J. (2018) On the accuracy of CORDEX RCMs to project future winds over the Iberian Peninsula and surrounding ocean. *Applied Energy*, 228, 289–300.

Schewe, J., Gosling, S.N., Reyer, C., Zhao, F., Ciais, P., Elliott, J., Francois, L., Huber, V., Lotze, H.K., Seneviratne, S.I., et al. (2019) State-of-the-art global models underestimate impacts from climate extremes. *Nature Communications*, 10(1), 1–14.

Sinden, G. (2007) Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand. *Energy Policy*, 35(1), 112–127.

Smith, A., Lott, N. and Vose, R. (2011) The integrated surface database: recent developments and partnerships. *Bulletin of the American Meteorological Society*, 92(6), 704–708.

Staffell, I. and Pfenninger, S. (2016) Using bias-corrected reanalysis to simulate current and future wind power output. *Energy*, 114, 1224–1239.

Stohl, A., Wotawa, G., Seibert, P. and Kromp-Kolb, H. (1995) Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. *Journal of Applied Meteorology*, 34(10), 2149–2165.

Stopa, J.E. and Cheung, K.F. (2014) Intercomparison of wind and wave data from the ecmwf reanalysis interim and the ncep climate forecast system reanalysis. *Ocean Modelling*, 75, 65–83.

Swail, V.R. and Cox, A.T. (2000) On the use of NCEP–NCAR reanalysis surface marine wind fields for a long-term north atlantic wave hindcast. *Journal of Atmospheric and Oceanic Technology*, 17(4), 532–545.

Taylor Karl, E. (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192. http://dx.doi.org/10.1029/2000jd900719.

Thorne, P.W., Allan, R.J., Ashcroft, L., Brohan, P., Dunn, R.J.H., Menne, M.J., Pearce, P.R., Picas, J., Willett, K.M., Benoy, M., Bronnimann, S., Canziani, P.O., Coll, J., Crouthamel, R., Compo, G.P., Cuppett, D., Curley, M., Duffy, C., Gillespie, I., Guijarro, J., Jourdain, S., Kent, E.C., Kubota, H., Legg, T.P., Li, Q., Matsumoto, J., Murphy, C., Rayner, N.A., Rennie, J.J., Rustemeier, E., Slivinski, L.C., Slonosky, V., Squintu, A., Tinz, B., Valente, M.A., Walsh, S., Wang, X.L., Westcott, N., Wood, K., Woodruff, S.D. and Worley, S.J. (2017) Toward an integrated set of surface meteorological observations for climate science and applications. *Bulletin of the American Meteorological Society*, 98(12), 2689–2702 Available at: https://journals.ametsoc.org/view/journals/bams/98/12/bams-d-16-0165.1.xml.

Tian, Q., Huang, G., Hu, K. and Niyogi, D. (2019) Observed and global climate model based changes in wind power potential over the northern hemisphere during 1979–2016. *Energy*, 167, 1224–1235.

Toledo, C., Chávez-Arroyo, R., Loera, L. and Probst, O. (2015) A surface wind speed map for Mexico based on NARR and observational data. *Meteorological Applications*, 22(3), 666–678.

Torralba, V., Doblas-Reyes, F.J. and Gonzalez-Reviriego, N. (2017) Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison. *Environmental Research Letters*, 12 (11), 114019.

Troen, I. and Petersen, E.L. (1989). 'European wind atlas'.

Uppala, S.M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V. D.C., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al. (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612), 2961–3012.

Vautard, R., Cattiaux, J., Yiou, P., Thépaut, J.-N. and Ciais, P. (2010) Northern hemisphere atmospheric stilling partly attributed to an increase in surface roughness. *Nature Geoscience*, 3(11), 756–761.

Wan, H., Wang, X. and Swail, V. (2010), 'Homogenization and trend analysis of Canadian near-surface wind speeds', *Journal of Climate* 23(5), 1209–1225.cited By 141.

Zeng, Z., Ziegler, A.D., Searchinger, T., Yang, L., Chen, A., Ju, K., Piao, S., Li, L.Z., Ciais, P., Chen, D., et al. (2019) A reversal in global terrestrial stilling and its implications for wind energy production. *Nature Climate Change*, 9(12), 979–985.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Molina MO, Gutiérrez C, Sánchez E. Comparison of ERA5 surface wind speed climatologies over Europe with observations from the HadISD dataset. *Int J Climatol*. 2021;41:4864–4878. https://doi.org/10.1002/joc.7103