



Statistical downscaling of GCMs wind speed data for trend analysis of future scenarios: a case study in the Lombardy region

Lucia Ferrarin¹ · Leonardo Stucchi¹ · Daniele Bocchiola¹

Received: 2 November 2023 / Accepted: 1 March 2024
© The Author(s) 2024

Abstract

Near-surface wind speed is a key climatic variable, affecting many sectors, such as energy production, air pollution, and natural hazard. Lombardy region of Italy is among the European areas with lowest average wind speed, leading generally to low air quality and wind energy potential. However, it is also one of the most affected area by tornadoes in Italy. Here we investigate possible changes in wind circulation as due to prospective global warming. We analysed wind speed WS under future scenarios (SSP1-2.6 and SSP5-8.5) from six Global Climate Models (GCMs) until 2100, tuned against observed WS data. We employed a statistical downscaling method, namely Stochastic Time Random Cascade (STRC) to correct locally GCMs outputs. Three statistical tests, i.e. Linear Regression, Mann Kendall, Moving Window Average, were carried out to analyse future trends of: annual WS averages, 95th quantile (as an indicator of large WS), and the number of days of calm wind per year (NWC). The proposed STRC algorithm can successfully adjust the mean, standard deviation, and autocorrelation structure of the GCM outputs. No strong trends are found for the future. The chosen variables would all display non-stationarity, and the 95th percentile display a positive trend for most of the stations. Concerning NWC, notable discrepancies among GCMs are seen. The STRC algorithm can be used to successfully adjust GCMs outputs to reflect locally observed data and to then generate credible long-term scenarios for WSs as a tool for decision-making.

1 Introduction

1.1 Wind forecast and future scenarios

Prediction of wind speed is relevant in several crucial fields. It is vital for issuing of alerts, to prevent potential catastrophes represented by hurricanes (Pant and Cha 2019) and storm surge (Barbariol et al. 2022), and limit economic damages. Forecast of wind speed is useful in the air/sea transport sector, and one can e.g. reduce fuel consumption and CO₂ emissions (e.g. Prats et al. 2022; Sun et al. 2022). In agriculture, wind forecast helps scheduling spraying, and

dusting procedures (Zhang et al. 2017). In wind power production, a forecast with lead times of even few second may help to guarantee the stability of the electrical grid (Peng et al. 2016). From the times of Romans to WWII, forecasting wind patterns was even a key to success for military operations (e.g. Lowe 2000, Gross 2022). Thanks to the availability of large datasets, machine learning, or similarly conceived models could soon outdate physical modelling (Schultz et al. 2021). However, weather models are quite limited on the temporary scale, ca. 10 days (Zhang et al. 2019), making them unsuitable to predict long term wind variations, which is again a key variable.

In turn, feasibility studies of wind farms (Koletsis et al. 2016), insurances (Collier et al. 2021), and environmental policies (Zhao et al. 2021) require prediction of medium/long term wind fields, that can be affected by climate changes, cascading into changes in weather, and wind trajectories. For this reason, the IPCC panel uses Global

✉ Leonardo Stucchi
leonardo.stucchi@polimi.it

¹ Department of Civil and Environmental Engineering, Politecnico di Milano, L. da Vinci, 32, Milano 20133, Italy

Circulation Models (GCMs) that physically simulate multiple phenomena, including wind, and project climate scenarios based on various Social Economic Pathways (SSPs) until 2100.

1.2 GCMs downscaling

Because they operate at the global scale, GCMs have a large space resolution, ca. 200 km, which is mostly unsuitable for local studies. Indeed, climate variables, and particularly wind speed, can vary significantly, even in flat areas, within the span of few kilometres (Hubbard 1994). To tune the output of GCMs to local weather condition, one can implement downscaling, which provides weather outputs at a given/chosen resolution. Several techniques of downscaling were developed hitherto, and can be roughly folded into two main approaches, i.e. dynamic, and statistical (e.g. Keller et al. 2022).

Dynamical downscaling uses outputs from GCMs as input to regional/local climate models, in turn providing in output regional/local climate variables. Accordingly, dynamic downscaling considers the complexity, and topography of the study area, and may provide better results in terms of spatial and temporal variability (Vaithinada Ayar et al. 2016), at the cost of computationally extensive /expensive simulations.

Statistical downscaling on the other hand uses local climate series, to modify GCMs outputs, basing on statistical, data driven relationships. These statistical rules are then applied for future scenarios to produce local climate series from large scale projections.

While being faster, statistical downscaling requires long as possible historical data to provide robust parameterization. Both methodologies, and hybrid versions of the two, were used to assess future wind scenarios (e.g. Nolan et al. 2012; Gonzalez-Aparicio et al. 2017; Reyers et al. 2015). While dynamical downscaling models may provide consistent results to assess overall wind variations on a regional scale, at point scale, due to the extreme spatial variability of wind, a statistical model may provide better results.

Here we analysed daily wind speed, from 10 climate stations during 25 years in Lombardy region, Italy, and we projected future scenarios of wind using a statistical approach (namely, Stochastic Time Random Cascade STRC, properly introduced here, and tuned locally). Downscaled daily values were then aggregated in the three variables of interest: average wind speed, number of days per year of calm wind, and 95% quantile wind speed. To these variables we then applied test statistics (Linear Regression, Mann Kendall, Moving Window) to detect possible trends or non-stationary processes.

1.3 Study area

Lombardy is the most populated region of Italy, with almost 10 million people located in a 23,863 km² area. It is also the largest industrial area of the country, and one of the largest in Europe. The region is almost equally divided between flatlands, and mountain regions, and the latter provides large water resources, used for industrial and agricultural purposes, in the Po Plain. Landlocked by the Alps in the North, and the Apennines in the South, wind recirculation in the Po Plain is mainly occurring from the Adriatic Sea on the East, leading to low wind velocity values. Wind power production in Lombardy is very modest, with only 12 wind farms, producing 0.06 MW (Terna 2023), but that is only a minor consequence of the scarce wind velocities in the region. The combination of the latter, with high urbanisation and industrialisation, makes Lombardy a most polluted area in Europe (Maranzano 2022), with several deaths yearly attributed to PM10 concentration (Baccini et al. 2011). Still, Po Plain is also the area of Italy most affected by tornadoes (Miglietta and Matsangouras 2018), as witnessed in July 2023, when a wind storm in Milan reached velocities over 100 km/h (Arpa 2023) causing severe damages.

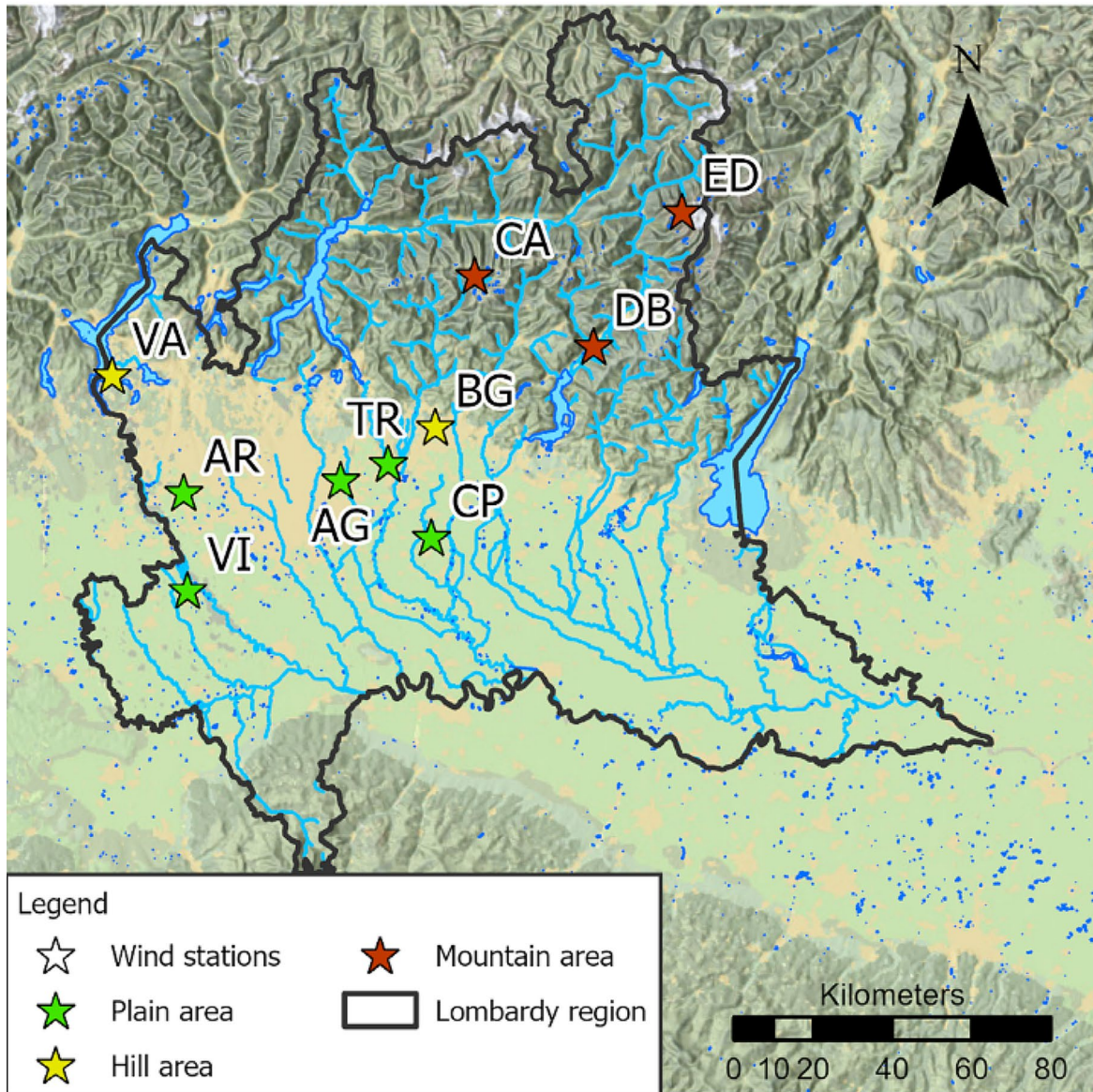
2 Data and methods

2.1 Wind stations data

The wind speed (WS) observed data analysed in the study were measured by automatic weather stations operated by the Regional Environmental Protection Agency, Arpa Lombardia, which provides WS data every 10, or 60 min in more than 500 stations, since 1989. However, only 230 of such station operated for more than 10 years, and all of them have frequent data gaps. Here, we selected 10 stations, with the longest periods of data collection, and the fewer gaps. These stations display a wide range of altitude, from 96 to 2108 m a.s.l. (Fig. 1; Table 1). Each station was assigned to a specific elevation class based on criteria outlined by the Central Statistics Institute of Italy (ISTAT 2024), distinguishing between plain (below 300 m a.s.l.), hill (300 to 600 m a.s.l.), and mountain (above 600 m a.s.l.) regions. The elevation class is not strictly determined by the station's elevation, but rather by the predominant elevations of the entire municipality. This decision was made to account for the topographic characteristics of the surrounding environment in our analysis, as these factors significantly influence the local wind patterns.

Starting from daily average data, we computed at monthly scale: the average WS, the number of days of calm wind (NCW), i.e. with WS < 1 m/s, and the 95th quantile (Fig. 2).

(a)



(b)

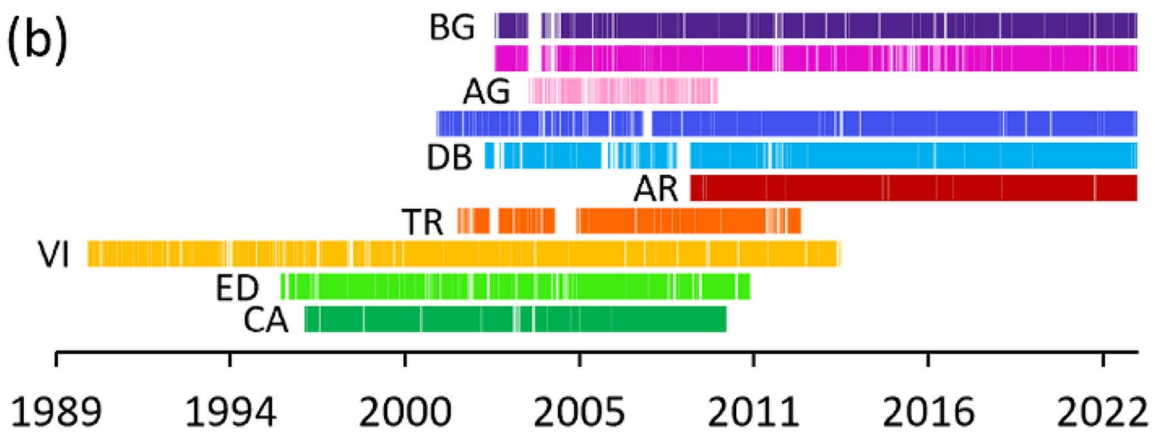


Fig. 1 (a) Map of Lombardy region with considered wind station clustered for elevation class. (b) Data availability for each station

Table 1 ID and features of the weather stations

| Station ID | Municipality | Altitude (m a.s.l.) | Location | | Municipality elevation class | Period of activity | |
|------------|--------------------|---------------------|-----------|----------|------------------------------|--------------------|------------|
| | | | Longitude | Latitude | | Start | End |
| BG | Bergamo | 249 | 9.661 | 45.696 | Hill | 01/01/1997 | 31/03/2010 |
| CP | Capralba | 96 | 9.646 | 45.445 | Plain | 21/03/1996 | 29/12/2010 |
| AG | Agrate Brianza | 162 | 9.354 | 45.576 | Plain | 13/03/1990 | 30/10/2013 |
| VA | Varese | 193 | 8.615 | 45.814 | Hill | 27/10/2001 | 01/08/2012 |
| DB | Darfo Boario Terme | 222 | 10.177 | 45.875 | Mountain | 12/02/2009 | 22/02/2023 |
| AR | Arconate | 182 | 8.847 | 45.549 | Plain | 28/08/2002 | 22/02/2023 |
| TR | Trezzo sull'Adda | 199 | 9.508 | 45.614 | Plain | 23/02/2001 | 22/02/2023 |
| VI | Vigevano | 107 | 8.860 | 45.326 | Plain | 24/01/2004 | 31/12/2009 |
| ED | Edolo | 2108 | 10.470 | 46.177 | Mountain | 01/01/2003 | 21/02/2023 |
| CA | Carona | 1954 | 9.794 | 46.039 | Mountain | 01/01/2003 | 22/02/2023 |

We chose these additional indicators because Lombardy is both (i) the Italian region with lowest WSs, and thereby highest NCW, and (ii) one of the regions most affected by windstorms (Miglietta and Matsangouras 2018). The station with the highest WS values is CA, in the mountains, while lowest values are seen in BG and CP, on hilly, and flat turf, respectively. Seasonality of WS is evident only for DB station, while more relevant cyclical behaviour is detected for NCW, where 4 stations (AR, TR, VI, ED) similarly show lowest values in summer.

As a further preliminary analysis, we assessed spatial correlation of WS through Pearson coefficient. In Fig. 3 it is shown (a) the correlation of WS between stations, against horizontal distance, (b) the minimum, average and maximum correlation values for each station against others. As expected, correlation is strongly related to distance, and higher for stations within the Po Valley (BG, CP, AG, AR, TR). This seemingly indicates an overall homogeneity, and generally calm wind of the area. On the other hand, the stations placed in the mountains (CA, ED, DB) show lower correlation values, likely because wind fields depend upon local microclimate conditions. It seems clear here that WS is a complex variable, not only affected by topography, but likely by other important factors, playing a role in wind circulation.

2.2 Global circulation models

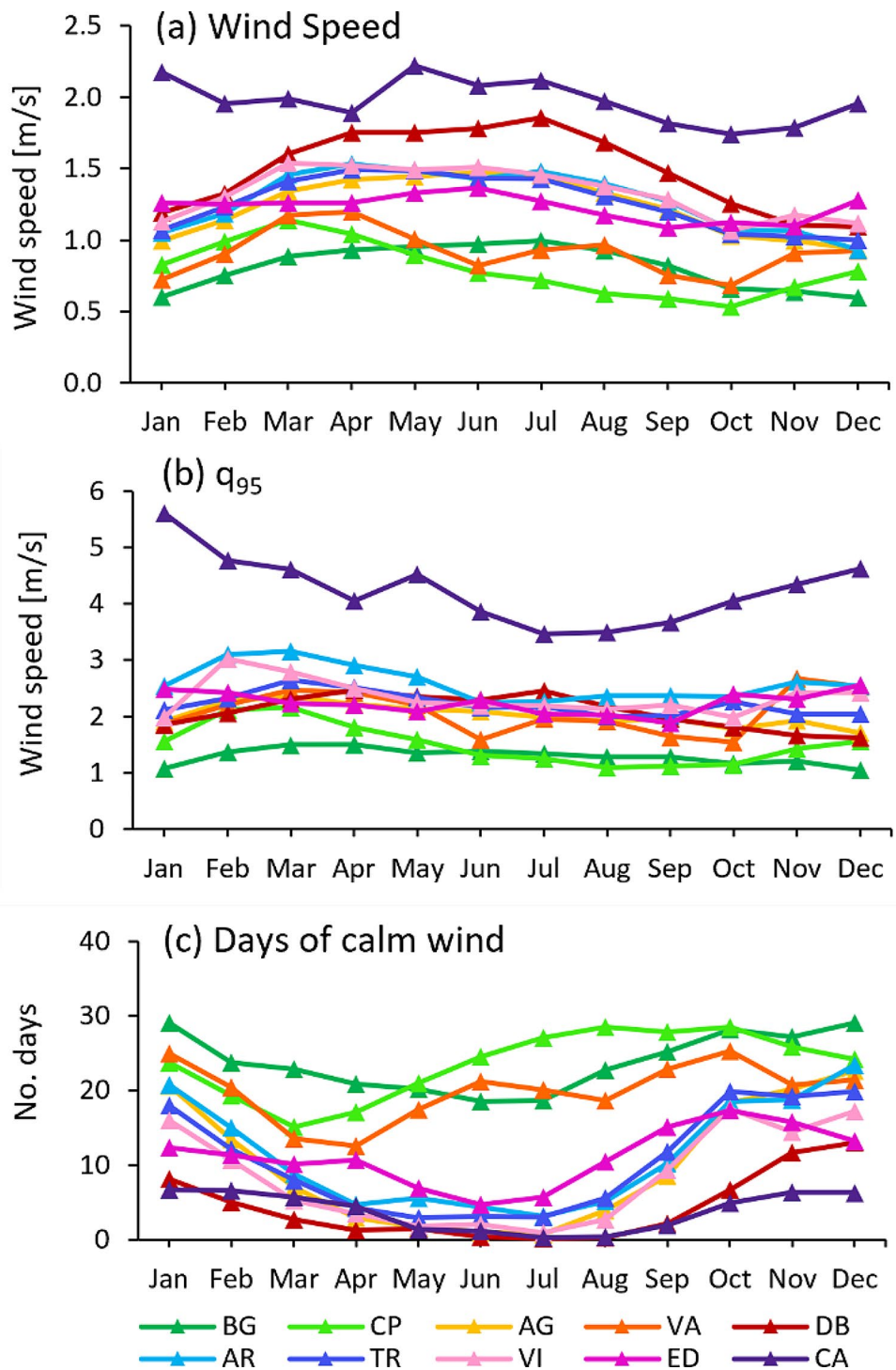
Six Global Circulation Models (GCMs), EC-EARTH3, CESM2, MIROC6, CMCC-CM2, MPI-ESM, and HadGEM3, were considered. These come from the Assessment Report 6 of the Intergovernmental Panel on Climate Change (IPCC). We assessed potential wind scenarios in Lombardy according to two Shared Socioeconomic Pathways: SSP1-2.6 (cut to net zero CO₂ emissions around 2075 and estimated warming of +1.4 °C by 2100) and SSP5-8.5 (CO₂ emissions triple by 2075 and estimated warming of +4.4 °C by 2100), giving somehow the expected extreme conditions in term of potential emission patterns, and climate evolution

thereby. The GCMs considered have spatial resolution between 80 and 130 km. A first comparison between ground observations, and GCMs showed a low correspondence (Fig. 4). Specifically, this discrepancy is seen as:

- i) A discernible bias in the calculated averages (Fig. 4a), with most GCMs tending to overestimate the wind speed (MPIESM being the one with the highest bias). Conversely, HADGEM3 and MIROC6 display relatively smaller biases, but they exhibit a tendency towards underestimation, particularly pronounced when considering station CA, which ranks second in altitude and records the highest wind speed values (both in terms of average and the 95th quantile) and exhibits the lowest NCW values (Fig. 2).
- ii) Distinct differences of the coefficient of variation (CV). Most GCMs exhibit a negative bias in their CV values (Fig. 4b), indicating a smoother signal compared to the observed data. This disparity in wind variability is a somewhat common occurrence when comparing GCMs with local observational data, likely attributable to coarse spatial resolution of the former. However, it is noteworthy that two specific models, MIROC6 and MPI-ESM, exhibit a higher level of variability with respect to five, out of the ten monitoring stations. Station DB, in particular, stands out displaying lowest variability, despite being located in a mountain region, with high wind speed values, second only to the high-altitude station CA.
- iii) Substantial differences in the autocorrelation structure (Fig. 5). In the case of 8 out of the 10 monitoring stations, the GCMs, with the exception of MIROC6, display autocorrelation values with very rapid decline, when compared against observed data. GCMs thus fail to faithfully capture periodicity in wind speed variations.

To further examine the alignment between GCMs and observed data, we employed Pearson correlation coefficient

Fig. 2 (a) Average, (b) 95th quantile and (c) number of days of calm wind computed at monthly scale for each station



as a quantitative measure. Correlation coefficients were calculated between observations, and GCMs data within the corresponding geographical tiles, at daily (Fig. 6a), and monthly scale (Fig. 6b). Correlation values are notably low, particularly at daily scale, failing to attain values above 0.1 for any combination of GCM and monitoring station. Considering monthly values, correlation is generally better, with a substantial improvement for HADGEM3, and yet not

good enough to consider GCMs data as well representative of local (measured) values.

Collectively, the findings from our analysis thus far demonstrate that the GCMs under scrutiny cannot generate daily WS time series aligning closely (in a statistical fashion) to the observed data. Therefore, one has to adjust GCMs outputs, to better conform to local wind conditions in the study

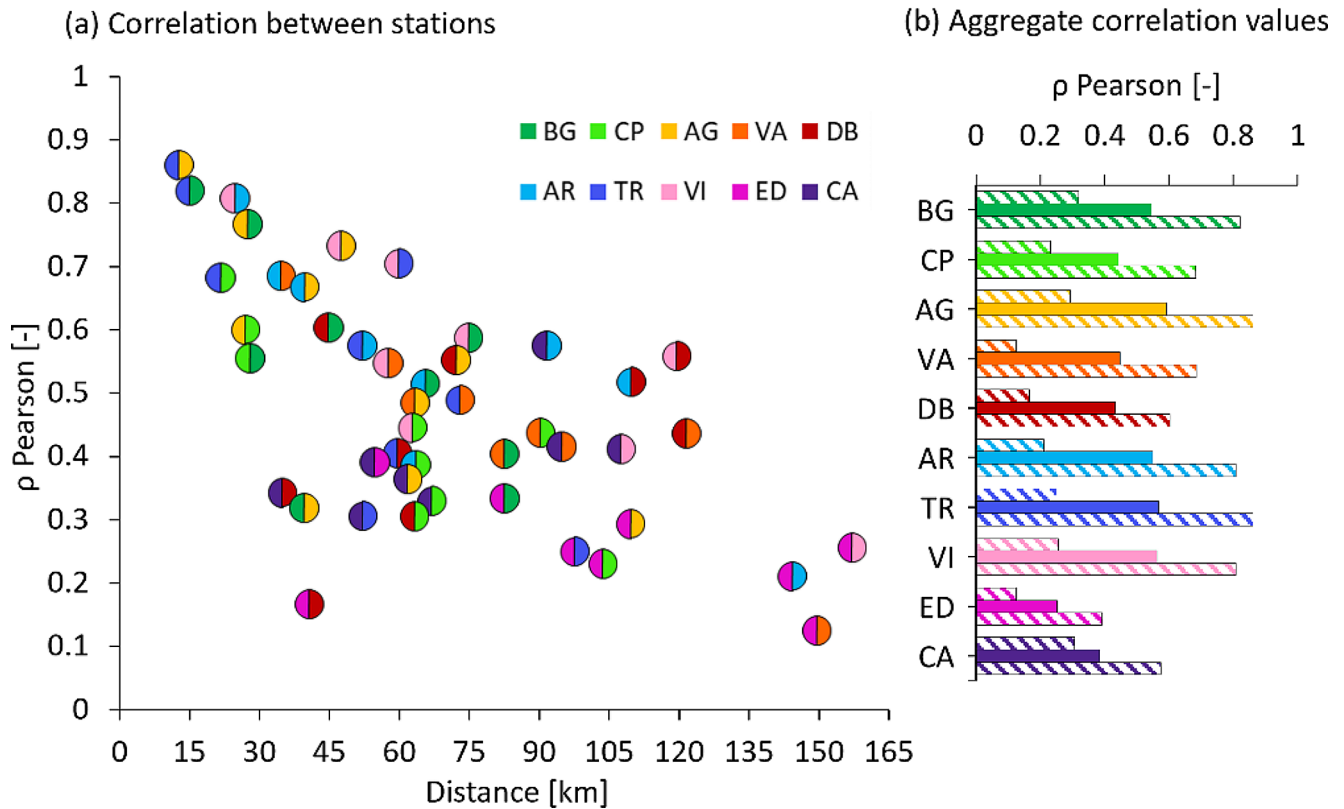


Fig. 3 (a) Pearson correlation values of WS for each station as a function of the distance. (b) Minimum, average and maximum correlation value for each station with all the others. Colour code represents station considered in the comparison

area, which is pivotal to understanding wind fields' dynamics within the study area.

2.3 The Stochastic Time Random Cascade downscaling

To correct the discrepancies in the GCMs data, we employed a Stochastic Time Random Cascade (STRC) algorithm. This method has been widely used in the field of meteorology, primarily for modelling the spatial distribution of meteorological variables, with a notable focus on precipitation (Schertzer and Lovejoy 1987; Gupta and Waymire 1993; Over and Gupta 1994; Deidda 2000). Moreover, it has been specifically applied for spatial downscaling of GCMs precipitation outputs (Bocchiola 2007; Groppelli et al. 2011). Importantly, the STRC algorithm structure allows to extend its application to also model the temporal distribution of a variable, as we have done in this study with daily WS values.

The STRC algorithm is based on a branching tree structure (Fig. 7), where each layer of the tree represents the variable at a different scale (e.g. Bocchiola and Rosso 2006). The node positioned at the top of the tree, namely the root node, represents the coarsest resolution data (i.e. the input of the algorithm). Moving downwards, each subsequent layer is characterized by a progressively finer resolution,

ultimately ending in the nodes at the bottom, which corresponds to the variable at the finest resolution (i.e. the output of the algorithm).

This method enables to increase the spatial or temporal resolution of a signal while concurrently preserving the spatial or temporal correlation structure observed in the measured data. In our specific application, it allowed us to downscale GCMs data from monthly scale, which we considered to be more consistent than daily data, to synthetic daily series, where it is preserved the time-wise autocorrelation structure identified in the observations.

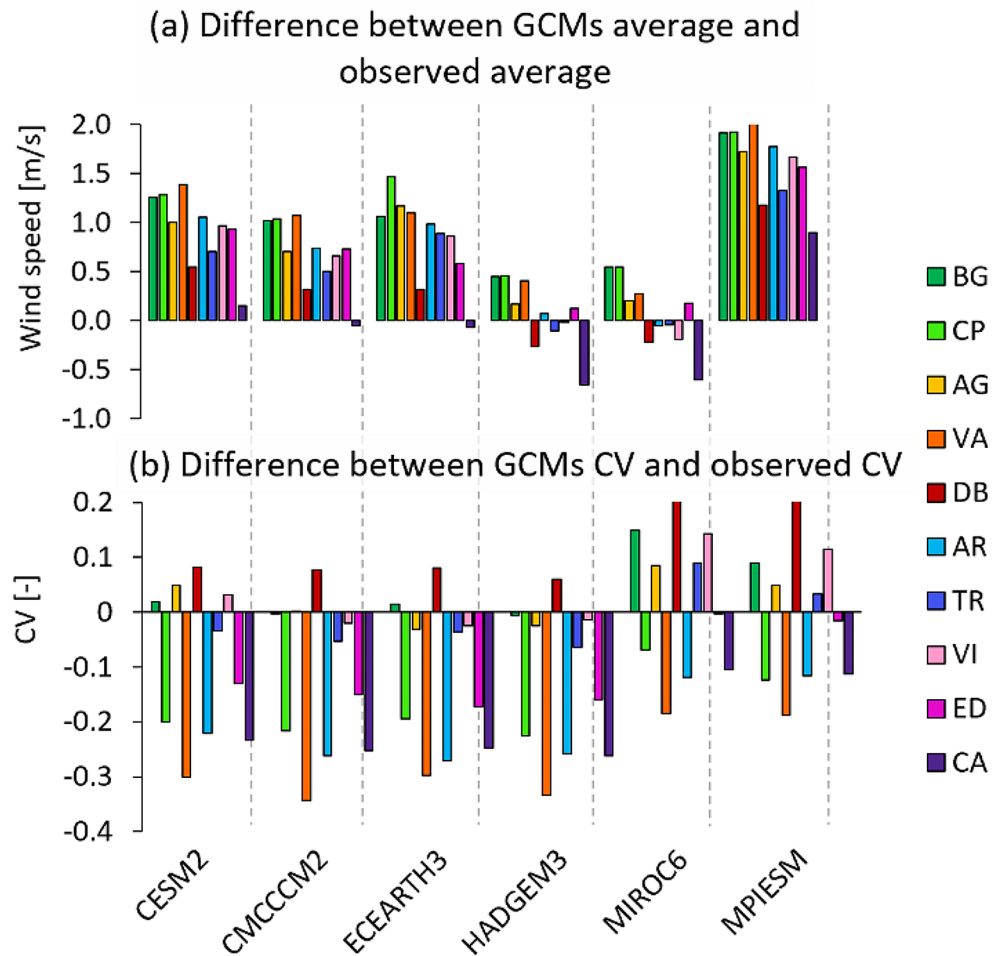
2.4 STRC algorithm structure

Within the framework of the STRC algorithm, the first step requires to correct the GCMs data with a month specific correction factor α_i obtained by the following:

$$\alpha_i = \frac{\overline{WS}_{OBS,i}}{\overline{WS}_{GCM,i}} \quad (1)$$

Where $\overline{WS}_{OBS,i}$ is the average value of the WS observations of the month i , and $\overline{WS}_{GCM,i}$ is the average value of the GCM WS data for the same month.

Fig. 4 (a) Difference of average WS between GCMs estimates and observation. (b) Difference of the coefficient of variation (CV) between GCMs estimates and observation. Positive values point to overestimate of WS/variability of GCMs with respect to observed data



Then, the values in each layer are generated, by multiplying the corresponding value of the preceding layer by distinct multiplicative factors (MFs). The MFs are stochastically drawn from a probability distribution. Notably, the parameters of the distribution vary for each layer-to-layer downward step, and are estimated from the observations. In this study, we decided to employ the lognormal distribution to model the MF. This fits with established practice in the application of the STRC algorithm, basing on the following consideration:

- i) The product of two variables, each following a logarithmic distribution, is itself described by a logarithmic distribution. This property applies well in the field of meteorology, as numerous variables have been empirically observed to exhibit lognormal distribution characteristics (Baran and Lerch 2015; Bocchiola 2007; Tustison et al. 2002; Marsan et al. 1996; Tapley and Waylen 1990).
- ii) The two parameters of the lognormal distribution (i.e. m , s), are directly linked to the average value of the distribution. Consequently, if the average value is known,

the two parameters are uniquely interrelated, reducing to one (i.e. variance) the number of calibration parameters.

The choice of the lognormal distribution was also tested against another typically adopted distribution, i.e. Weibull (Ozay and Celiktas 2016; Pishgar-Komleh et al. 2015), the former giving better performance (see Sect. 3.1).

The generic WS value i at the scale s can be computed starting from the value at the root node as follows:

$$WS_s^i = WS_0 \prod_{j=1}^s MF_j^i \tag{2}$$

Where WS_0 is the input value (i.e. the data at the coarsest resolution) and MF_j^i is the multiplicative factor for the value i at the layer j . The statistics of the lognormal distribution of the factor y are the following:

$$E[MF_s] = 1 \tag{3}$$

$$VAR[MF_s] = \vartheta_{MF_s}^2 \tag{4}$$

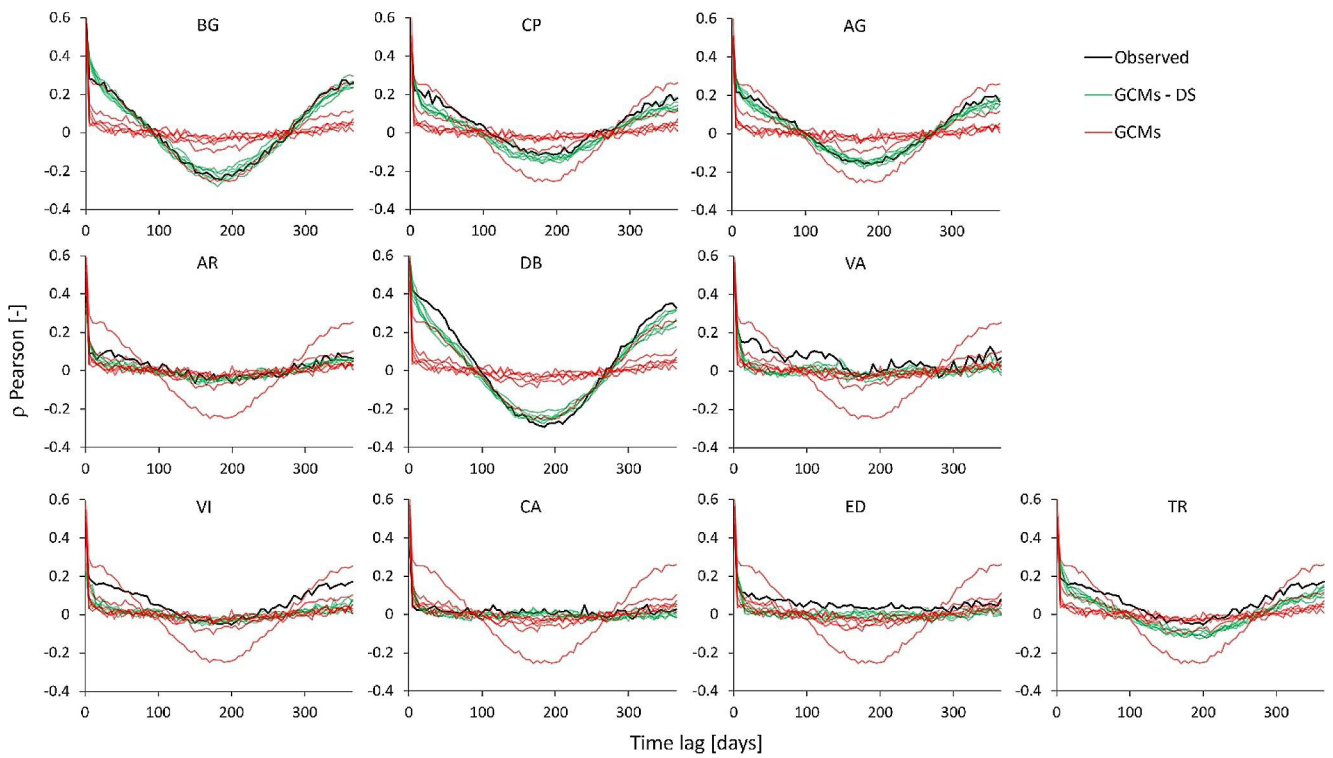
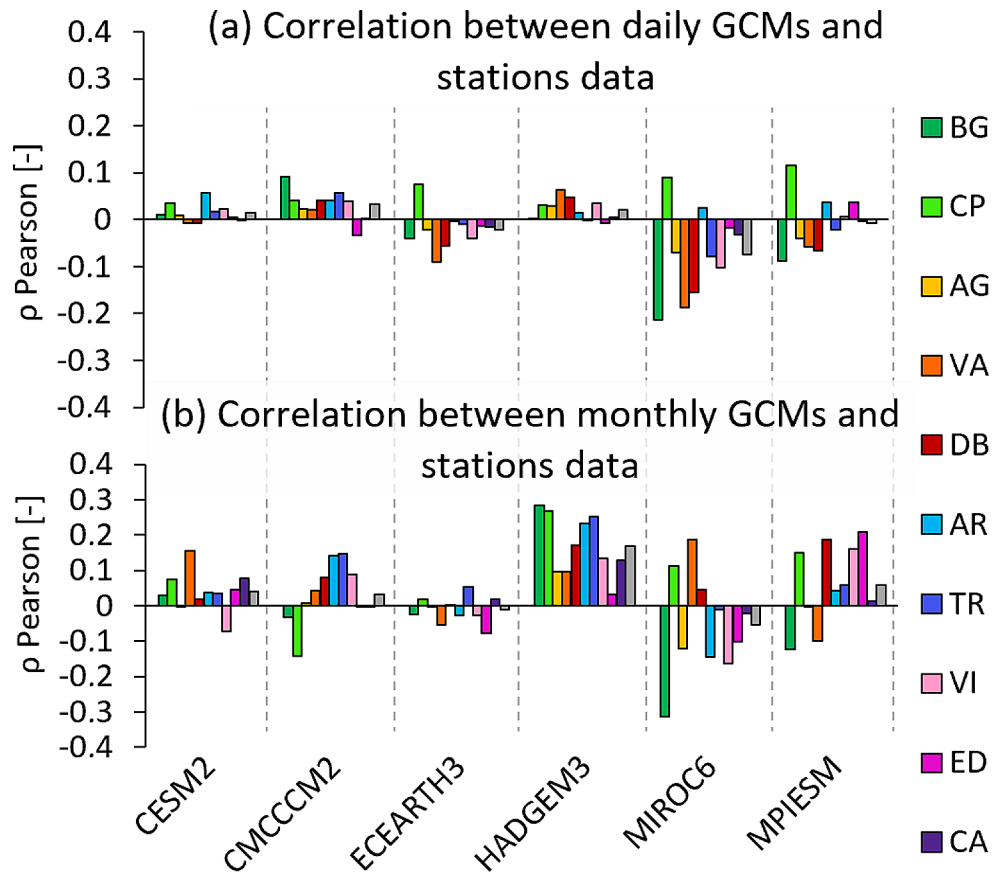


Fig. 5 Comparison between autocorrelation of observed WS (black line) and simulated WS by GCMs: uncorrected (red line) and downscaled (DS, green line), for the 10 monitoring stations

Fig. 6 Pearson linear correlation between (a) daily observed WS and daily simulated WS and (b) monthly observed WS and monthly simulated WS



Equation (2) yields the average conservation of the mass in the cascade and it is a direct consequence of the application of the correction factor in Eq. 1. Since the expected value of a lognormal distribution is given by:

$$E [MF_s] = e^{\mu_{\ln MF_s} + \frac{1}{2}\sigma_{\ln MF_s}^2} \tag{5}$$

From Eqs. (2) and (4) we can write the parameters of the lognormal distribution as follows:

$$MF_s \sim LN\left(-\frac{1}{2}\sigma_{\ln MF_s}^2, \sigma_{\ln MF_s}\right) \tag{6}$$

As mentioned before, from Eq. (5) one only needs to estimate one parameter, $\sigma_{\ln MF_s}$. We employed an iterative estimation algorithm, identifying $\sigma_{\ln MF_s}$, by variance fitting, i.e. minimizing the difference of the standard deviation between the observation and the GCMs downscaled data, at each time scale (i.e. for every layer of the tree structure).

In this study, we apply the algorithm for each station, by starting from the mean monthly value of the corresponding GCM cell, which is then downscaled at daily scale. The tree structure consists of four layers (branches), then split into two, three and five new branches, respectively (see Fig. 7). The upper layer represents the monthly average data from the GCM, while the layers below represent 15 days, 5 days and daily averages. Every value at each layer is obtained

multiplying one value of the layer above by a MF. Due to the inherent structure of the algorithm, it follows that values derived from the same “parent” branch exhibit a higher degree of correlation with each other, compared to values obtained from a different “parent” branch. For the same reason, values originating from neighbouring “parents” display stronger correlations than values originating from “parents” farther apart in the tree structure. This structure aids production of downscaled data characterized by autocorrelation patterns that diminish as the time lag increases, with a rate that closely approximates the observed autocorrelation behaviour. Note that the calibration of $\sigma_{\ln MF_s}$ is done for every station individually, since ground based WS data showed a marked difference in the autocorrelation structure.

2.5 Trend analysis tests

To pursue a comprehensive trend analysis on the corrected GCMs data, we employed three distinct trend assessment methods:

- i) The Linear Regression F-test was utilized to investigate the presence of a linear trend within the time series data.
- ii) The Mann-Kendall test (Kendall 1975), a non-parametric test, was applied to examine whether there exists a monotonic trend in the data, without making any assumptions about the rate of the trend.

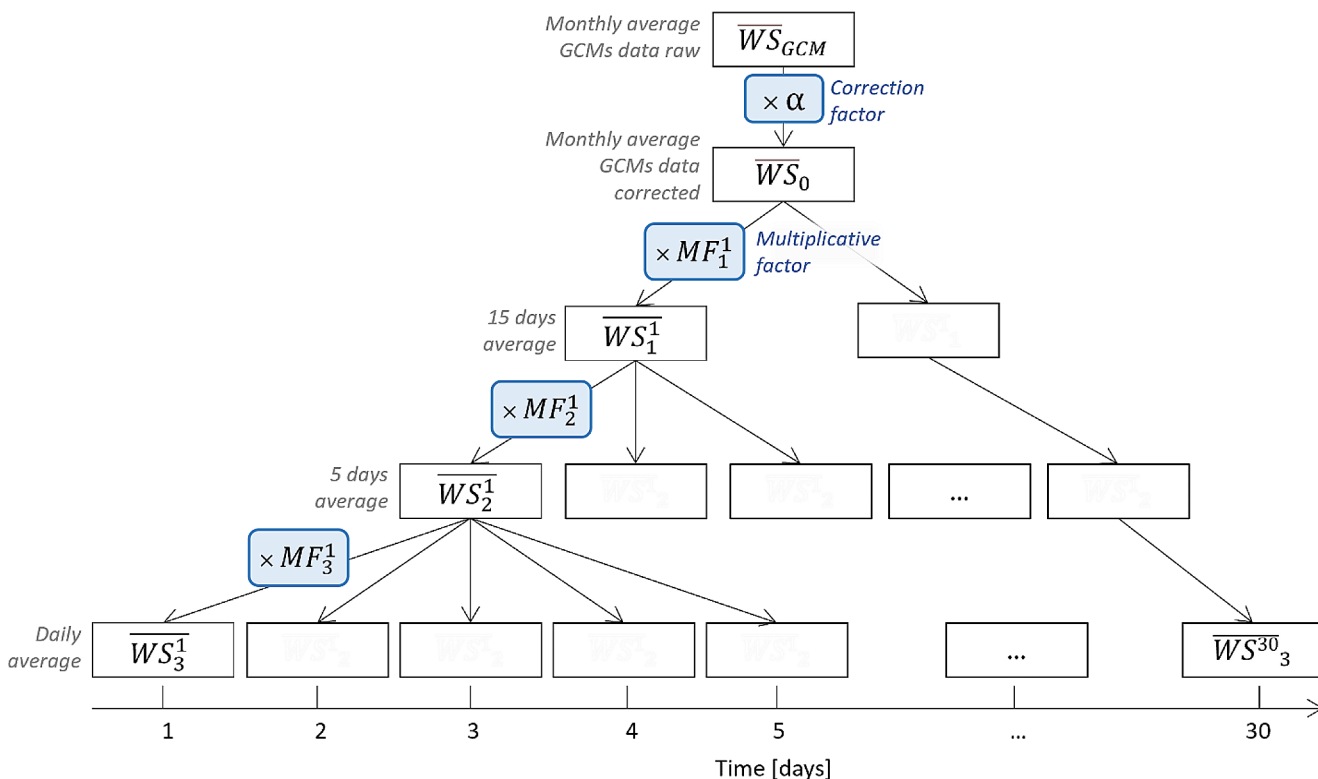


Fig. 7 Tree structure of the STRC algorithm employed to downscale monthly average WS data from the GCMs to a daily time scale

iii) The Moving Window Average test (De Michele et al. 1998) was employed to assess whether the data series exhibit stationarity or exhibit significant deviations from the mean. This was accomplished by comparing the moving average over 30-year periods (WMO, 2017).

These trend tests were pursued on the same variables previously analysed (i.e. yearly averages, 95th quantiles and NCW). The data series subject to testing pertained to the future projections of WS under the SSP1-2.6 and SSP5-8.5 scenarios, during 2015–2100. These projections were down-scaled for each monitoring station. All tests were executed with a 95% confidence level.

3 Results and discussion

3.1 Distribution of observed multiplicative factors

The initial assumption requiring evaluation pertains to the distribution of the MF, derived computing the ratio between daily WS observations and corrected GCMs data (i.e. multiplied by the factor α_i). Quantile-Quantile (QQ) plots in Fig. 8 (Marden 2004), compare quantiles of the natural logarithm of observed MF values (Y axis), against those of a normal distribution (X axis). For most stations, MF values are well approximated by a lognormal distribution.

However, there are exceptions, particularly evident in the higher MF values of specific stations (e.g. CP, VA, and ED). Notably, the QQ plot shapes suggest that these stations exhibit MF value distributions with a fatter right tail, as compared to a lognormal distribution. Consequently, for such stations, we expect a decreased reliability of the STRC algorithm for high WS values. The insights from the QQ plots find confirmation in the Goodness of Fit tests conducted on the MF (Fig. 9). A substantial number of stations pass the Chi-Squared test with the majority of the GCMs. However, fewer GCMs pass the Kolmogorov-Smirnov and Anderson-Darling tests, primarily because these tests place greater emphasis on the tails of the distribution, which corresponds to a vulnerability in the MF of specific stations. In conclusion, we believe these findings provide enough justification for employing a lognormal distribution to characterize the bulk of the MF distribution, while showing potential limitations in representing high-extreme WS events.

3.2 STRC downscale calibration

The STRC algorithm is calibrated for each weather station and every GCM. $\sigma_{\ln MF}$ is computed for every layer of the tree structures with an iterative algorithm, which minimizes the difference on the standard deviation between the observations and the downscaled GCMs data, at every time scale (i.e. every layer of the tree). In this study $\sigma_{\ln MF}$ was computed specifically for individual months. This decision

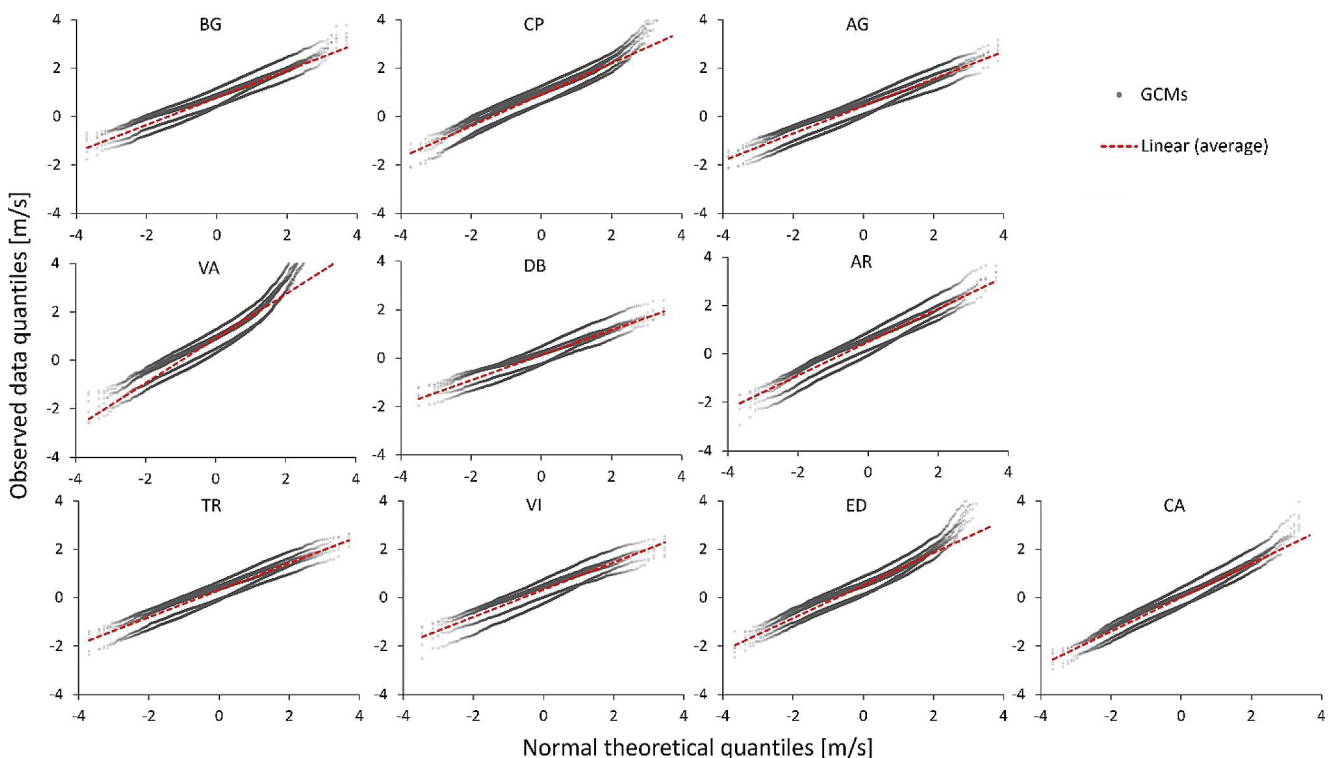


Fig. 8 Quantile-Quantile plots of the MF computed between WS observations and GCMs data, for each station

Fig. 9 Number of GCMs that passes the goodness of fitness tests, divided by test (Chi squared, Kolmogorov-Smirnov and Anderson-Darling) and by station

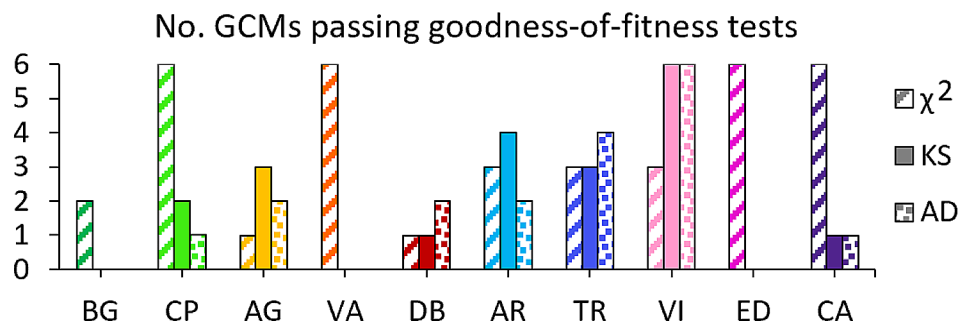
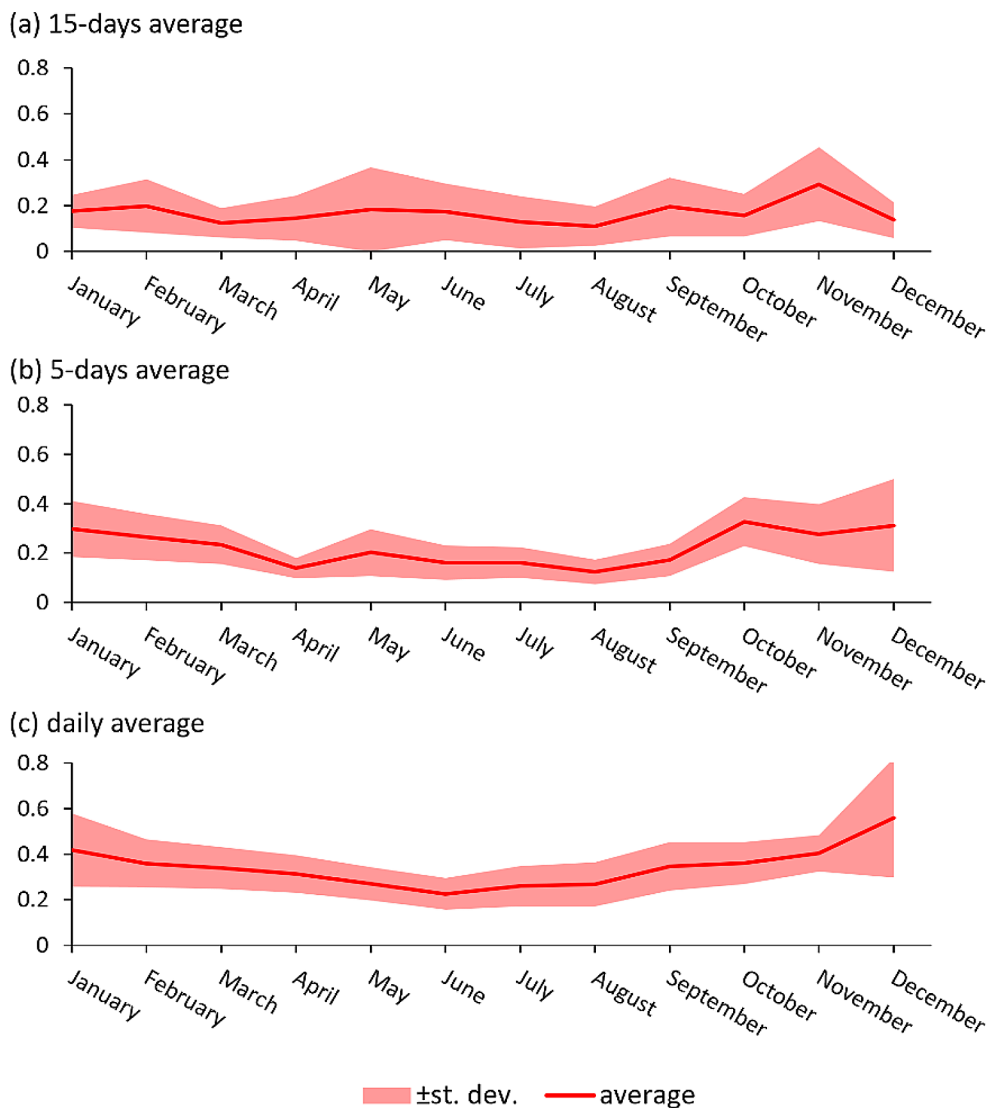


Fig. 10 σ_{lnMF} computed for each month, for the different layers of the STRC algorithm. The value of the parameter and its monthly changes increase with the temporal scale



was made based on the observation of pronounced seasonality in WS data across multiple weather stations, that may imply different distributions parameters of the MF. As expected, the estimated σ_{lnMF} increases as time scale decreases (Fig. 10). Notably, a clear time dependence is visible for 5-days average and, more evidently, for daily averages, showing lower variability of the MF during summer

months, and higher variability during winter months, which suggests a corresponding behaviour of the variability of the WS (confirmed by observations, Fig. 2).

Fig. 11 Error on the coefficient of variation (CV) computed subtracting the CV of the observed WS to the CV of the corrected GCMs data. Compared to the uncorrected GCMs data the error on the CV is lower for almost all stations

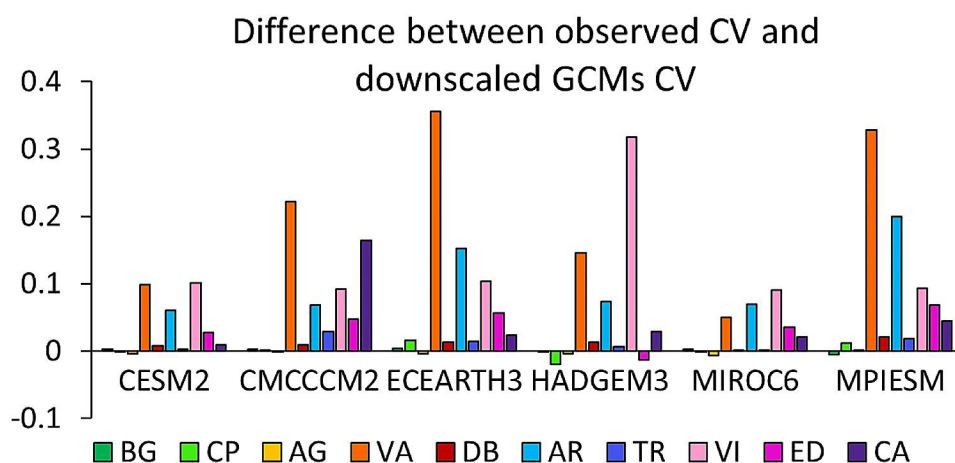


Table 2 Results of the trend tests (Linear Regression, LR, Mann-Kendall, MK and Moving-Window, MW) on the annual average of WS for the two scenarios (SSP1-2.6 and SSP5-8.5). For every model, the number of stations passing the test is indicated. The coefficient of the linear regression is indicated when the corresponding test is passed

| WS | SSP1-2.6 | | | | SSP5-8.5 | | | |
|----------|--|--------|--------------|--------------|--|--------|--------------|--------------|
| | LR [$\text{ms}^{-1} \cdot 100\text{y}^{-1}$] | | MK | MW | LR [$\text{ms}^{-1} \cdot 100\text{y}^{-1}$] | | MK | MW |
| | No. stations | Coeff. | No. stations | No. stations | No. stations | Coeff. | No. stations | No. stations |
| CESM2 | 0 | - | 0 | 0 | 0 | - | 0 | 0 |
| CMCCCM2 | 0 | - | 0 | 5 | 0 | - | 0 | 5 |
| ECEARTH3 | 0 | - | 0 | 0 | 1 | 0.05 | 0 | 5 |
| HADGEM3 | 0 | - | 0 | 2 | 0 | - | 0 | 4 |
| MIROC6 | 0 | - | 0 | 1 | 0 | - | 0 | 7 |
| MPIESM | 0 | - | 0 | 5 | 0 | - | 0 | 1 |

3.3 Downscaling output

The adjusted WS time series were compared with the observed data, revealing a significantly improved alignment in contrast to the unadjusted ones. Specifically, three key observations can be made:

- i) Discrepancies in the total average values become negligible. The use of a correction factor based upon average values results in deviations from observed averages not exceeding 10%.
- ii) Enhanced representativity is evident when considering the coefficient of variation (CV) of WS (Fig. 11). Deviations from the observed CV values do not exceed 20% for 8 out of 10 stations. Conversely, station VA exhibits a pronounced deviation from the observed CV, meaning likely that the downscaled GCMs data display greater variability, compared against the observed data. This outcome could be caused by little adherence of MF in this station to a lognormal distribution (especially on the right tail, see Figs. 8 and 9).
- iii) Enhanced fit is observed in the autocorrelation patterns between the observed and the downscaled GCMs data (Fig. 5). This finding is particularly relevant within the

scope of WS analysis, as no downscaling methods that we know of deals with for this facet. We suggest that this circumstance is noteworthy when carrying out WS analysis (e.g. proper representation of autocorrelation in time may be key for short term forecast when using the STRC, e.g. Bocchiola and Rosso 2006).

3.4 Trend tests results

A comprehensive assessment of the time series downscaled data was conducted for two distinct scenarios, SSP1-2.6 and SSP5-8.5. The results of the trend analyses show as follows:

- i) Absence of statistically significant trends in the annual average wind speeds (Table 2). However, it is noteworthy that a considerable portion of the models exhibited non-stationary behaviour in certain weather stations. This phenomenon was more pronounced in the SSP5-8.5 scenario, maybe indicating an amplification of non-stationary behaviour of wind patterns in a higher emissions scenario.
- ii) Substantial positive trends were identified in the 95th quantile of wind speed across all models (Table 3).

Table 3 Results of the trend tests (Linear Regression, LR, Mann-Kendall, MK and Moving-Window, MW) on the 95th quantile of WS for the two scenarios (SSP1-2.6 and SSP5-8.5). For every model, the number of stations passing the test is indicated. The coefficient of the linear regression is indicated when the corresponding test is passed

| 95th | SSP1-2.6 | | | SSP5-8.5 | | | | |
|----------|--|--------|--------------|--------------|--|--------|--------------|--------------|
| | LR [$\text{ms}^{-1} \cdot 100\text{y}^{-1}$] | | MK | MW | LR [$\text{ms}^{-1} \cdot 100\text{y}^{-1}$] | | MK | MW |
| | No. stations | Coeff. | No. stations | No. stations | No. stations | Coeff. | No. stations | No. stations |
| CESM2 | 9 | 0.32 | 9 | 10 | 8 | 0.26 | 7 | 9 |
| CMCCCM2 | 1 | -0.14 | 1 | 10 | 2 | 0.15 | 2 | 8 |
| ECEARTH3 | 7 | 0.27 | 4 | 8 | 2 | 0.22 | 0 | 8 |
| HADGEM3 | 4 | 0.25 | 4 | 9 | 9 | 0.24 | 8 | 10 |
| MIROC6 | 8 | 0.24 | 9 | 9 | 1 | 0.46 | 2 | 9 |
| MPIESM | 10 | 0.39 | 10 | 10 | 6 | 0.30 | 7 | 10 |

Table 4 Results of the trend tests (Linear Regression, LR, Mann-Kendall, MK and Moving-Window, MW) on the annual number of calm wind for the two scenarios (SSP1-2.6 and SSP5-8.5). For every model, the number of stations passing the test is indicated. The coefficient of the linear regression is indicated when the corresponding test is passed

| NCW | SSP1-2.6 | | | SSP5-8.5 | | | | |
|----------|--|--------|--------------|--------------|--|--------|--------------|--------------|
| | LR [$\text{n}_{\text{days}}^{-1} \cdot 100 \text{y}^{-1}$] | | MK | MW | LR [$\text{n}_{\text{days}}^{-1} \cdot 100 \text{y}^{-1}$] | | MK | MW |
| | No. stations | Coeff. | No. stations | No. stations | No. stations | Coeff. | No. stations | No. stations |
| CESM2 | 9 | -22.4 | 8 | 10 | 5 | -15.9 | 4 | 9 |
| CMCCCM2 | 3 | 17.2 | 3 | 9 | 5 | 20.1 | 5 | 10 |
| ECEARTH3 | 3 | -14.8 | 3 | 8 | 1 | 13.4 | 1 | 8 |
| HADGEM3 | 1 | -14.2 | 3 | 10 | 7 | -21.3 | 7 | 9 |
| MIROC6 | 9 | -20.7 | 9 | 9 | 0 | - | 0 | 8 |
| MPIESM | 10 | -31.8 | 10 | 10 | 3 | -18.8 | 4 | 10 |

Specifically, these trends indicated an average increase along the century, ranging from +0.24 to +0.39 m/s (from +10% to +17%), for the SSP1-2.6 scenario. Remarkably, one model displays a negative trend under SSP1-2.6. For the SSP5-8.5 scenario, the trends in the 95th quantile varied between +0.15 and +0.46 m/s per 100 years (from +6% to +20%), with all models displaying positive trends.

- iii) Consistent non-stationary behaviour of NCW values across all models and nearly all weather-stations (Table 4). Furthermore, the application of the Mann-Kendall test and the linear regression test identified significant trends in NCW for specific stations. However, the direction of these trends remains ambiguous: it is not clear whether the NCW is projected to either increase (as suggested by 5 models for the scenario SSP1-2.6 and 2 models for the SSP5-8.5), or decrease (as suggested by one model for the SSP1-2.6 and 2 models for the SSP5-8.5).

3.5 Wind scenarios in Mediterranean area

Overall, our trend analysis results highlight the considerable uncertainty, and variability of projected wind speed scenarios in the Mediterranean area, where Lombardy region is laid. Trend's magnitude, direction and location are dependent upon the area, and influenced by the chosen GCM and

downscale method, often with opposite results. Notably, the main source of uncertainty typically arises from variance between GCMs (Koletsis et al. 2016).

Future scenarios in the Mediterranean region show a predominantly negative trend in WS. Donat et al. (2011), using Climate Models prior to 2007, found loss of wind storms (98th quantile) for the Mediterranean region. Similarly, Tobin et al. (2015), and Hueging et al. (2013) found a negative trend for wind energy production in the Mediterranean area. In another recent study (Bonanno et al., 2023), specific for Italy, a negative trend of wind power production was found, under the sole RCP 8.5 scenario. More specifically, the trend was detected in coastal regions, where wind power plants are located, while in inner areas, and particularly in the North, the trend is negligible. Some other scholars (Bloom et al. 2008; Kjellström et al. 2011) reported future negative trends, limited to the winter season, whereas in summer positive trends are found. This may be explained by alteration in large scale circulation patterns, e.g. resulting from a rise in anticyclonic impact, and a fall in cyclonic impact, while summer positives trend can be attributed to growth of air temperature.

However, numerous studies have reported highly uncertain results, also in regions close to case study area. Medugorac et al. (2021), when studying with RCMs storm surges from the Adriatic Sea, the main source of air recirculation in Po Valley, found small probability of change in wind fields. Often times, significant differences are found between

GCMs, making difficult to detect a consistent trend (e.g. Rockel and Woth 2007; Outten and Esau 2013). Pryor et al. (2006) studied the 90th percentile trends in future scenarios, finding fluctuations of about $\pm 15\%$ at the end of 2100. However, the magnitude of these trends was no bigger than the scatter between GCMs. The divergence between GCMs in the same study resulted in an inability to predict future storminess.

4 Conclusions and outlook

In this study we applied a statistical method to correct the bias of WS data from GCMs. We kept the monthly signal of the GCMs, and through a downscaling technique we obtained daily WS data, resembling statistical properties of the observations (i.e. average, variability and autocorrelation). The downscaled timeseries demonstrated highly accurate replication of the observations, with no significant variations observed among the GCMs, but with some variations noted among the weather stations.

The downscaling algorithm was used to project future WS scenarios (namely SSP1-2.6 and SSP5-8.5) during 2015–2100. To detect potential implications on wind circulation resulting from climate change, trend tests were pursued using the projected daily WS data. Even though average annual wind speed exhibits non-stationary dynamics, particularly in the context of the highest emissions scenario (SSP5-8.5), no significant trends were evidenced. The 95th percentile of wind speed exhibited consistent, and significantly positive trends in most models for both scenarios, maybe calling for a future rise in extreme wind speeds. NCW values exhibited non-stationary behaviour and significant trends in select stations. Nevertheless, future NCW indicator remains uncertain.

The findings of this study suggest that climate change is expected to affect the circulation of wind in Lombardia region. However, it remains challenging to predict the evolution of WS quantitatively, by the end of the 21st century. Due to the contrasting behaviour of the variables, it appears that the probability distribution of WS is likely to be altered in the future. A more comprehensive investigation, incorporating a greater number of GCMs and weather stations, may result in less uncertain findings, and may be pursued henceforth. Furthermore, estimation of the STRC algorithm parameters may be tested for robustness/dependence against other variables affecting wind circulation (e.g. pressure gradient, vorticity). Finally, the examination of seasonal or monthly data separately should be undertaken, because WS clearly displays strong seasonal dependence.

Acknowledgements The authors would like to express their gratitude to Dr. Maria Ranci and Dr. Marta Salvati from Arpa Lombardia for

supplying the wind speed data collected by the weather stations and for their contribution in the development of the study.

Author contributions Conceptualization: L.F., D.B., Methodology: L.F., L.S., Formal Analysis: L.F., Writing-Draft: L.F., L.S., Writing-Review and Editing: D.B., Supervision: D.B.

Funding Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

Data availability Data will be made available upon reasonable request.

Declarations

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arpa (2023) Meteorological analysis of July 2023 storms. Available in Italian at: https://www.arpalombardia.it/media/p5qjndr/rapporto_evento_arpa_temporali_luglio_2023_web.pdf. Last accessed on: 02/11/2023
- Baccini M, Biggeri A, Grillo P, Consonni D, Bertazzi PA (2011) Health impact assessment of fine particle pollution at the regional level. *Am J Epidemiol* 174(12):1396–1405. <https://doi.org/10.1093/aje/kwr256>
- Baran S, Lerch S (2015) Log-normal distribution based Ensemble Model output statistics models for probabilistic wind-speed forecasting. *Q J R Meteorol Soc* 141(691):2289–2299. <https://doi.org/10.1002/qj.2521>
- Barbariol F, Pezzutto P, Davison S, Bertotti L, Cavaleri L, Papa A, Favaro M, Sambo E, Benetazzo A (2022) Wind-wave forecasting in enclosed basins using statistically downscaled global wind forcing. *Front Mar Sci* 9:1002786. <https://doi.org/10.3389/fmars.2022.1002786>
- Bloom A, Kotroni V, Lagouvardos K (2008) Climate change impact of wind energy availability in the Eastern Mediterranean using the regional climate model PRECIS. *Nat Hazards Earth Syst Sci* 8(6):1249–1257. <https://doi.org/10.5194/nhess-8-1249-2008>
- Bocchiola D (2007) Use of scale recursive estimation for assimilation of precipitation data from TRMM (PR and TMI) and NEXRAD.

- Adv Water Resour 30(11):2354–2372. <https://doi.org/10.1016/j.advwatres.2007.05.012>
- Bocchiola D, Rosso R (2006) The use of scale recursive estimation for short term quantitative precipitation forecast. *Phys Chem Earth* 31:1228–1239. <https://doi.org/10.1016/j.pce.2006.03.019>
- Bonanno R, Viterbo F, Maurizio RG (2023) Climate change impacts on wind power generation for the Italian peninsula. *Reg Environ Change* 23:15. <https://doi.org/10.1007/s10113-022-02007-w>
- Collier SJ, Elliott R, Lehtonen TK (2021) Climate change and insurance. *Econ Soc* 50(2):158–172. <https://doi.org/10.1080/03085147.2021.1903771>
- De Michele C, Montanari A, Rosso R (1998) The effects of non-stationarity on the evaluation of critical design storms. *Water Sci Technol* 37(11):187–193. [https://doi.org/10.1016/S0273-1223\(98\)00332-1](https://doi.org/10.1016/S0273-1223(98)00332-1)
- Deidda R (2000) Rainfall downscaling in a space-time multifractal framework. *Water Resour Res* 36(7):1779–1794. <https://doi.org/10.1029/2000WR900038>
- Donat MG, Leckebusch GC, Wild S, Ulbrich U (2011) Future changes in European winter storm losses and extreme wind speeds inferred from GCM and RCM multi-model simulations. *Nat Hazards Earth Syst Sci* 11(5):1351–1370. <https://doi.org/10.5194/nhess-11-1351-2011>
- Gonzalez-Aparicio I, Monforti F, Volker P, Zucker A, Careri F, Huld T, Badger J (2017) Simulating European wind power generation applying statistical downscaling to reanalysis data. *Appl Energy* 199:155–168. <https://doi.org/10.1016/j.apenergy.2017.04.066>
- Groppelli B, Bocchiola D, Rosso R (2011) Spatial downscaling of precipitation from GCMs for climate change projections using random cascades: a case study in Italy. *Water Resour Res* 47(3). <https://doi.org/10.1029/2010WR009437>
- Gross P (2022) Attack on Pearl Harbour: How the weather helped the Japanese. Click on Detroit. Available at: <https://www.clickondetroit.com/features/2022/12/06/attack-on-pearl-harbor-how-the-weather-helped-the-japanese/>. Last accessed: 29/02/2024
- Gupta VK, Waymire EC (1993) A statistical analysis of mesoscale rainfall as a random cascade. *J Appl Meteorol Climatol* 32(2):251–267. [https://doi.org/10.1175/1520-0450\(1993\)032<0251:ASAO MR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<0251:ASAO MR>2.0.CO;2)
- Hubbard KG (1994) Spatial variability of daily weather variables in the high plains of the USA. *Agric for Meteorol* 68(1–2):29–41. [https://doi.org/10.1016/0168-1923\(94\)90067-1](https://doi.org/10.1016/0168-1923(94)90067-1)
- Hueging H, Haas R, Born K, Jacob D, Pinto JG (2013) Regional changes in wind energy potential over Europe using regional climate model ensemble projections. *J Appl Meteorol Climatol* 52(4):903–917. <https://doi.org/10.1175/JAMC-D-12-086.1>
- ISTAT (2024) <https://www.istat.it/it/files//2015/04/Classificazioni-statiche-Anni-2022-2024.zip>. Available at the online archive of ISTAT: <https://www.istat.it/it/archivio/156224>. Last accessed: 29/02/2024.
- Keller AA, Garner KL, Rao N, Knipping E, Thomas J (2022) Downscaling approaches of climate change projections for watershed modeling: review of theoretical and practical considerations. *PLOS Water* 1(9):e0000046. <https://doi.org/10.1371/journal.pwat.0000046>
- Kendall MG (1975) Rank correlation methods. Oxford Univ, New York
- Kjellström E, Nikulin G, Hansson ULF, Strandberg G, Ullerstig A (2011) 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations. *Tellus A: Dynamic Meteorol Oceanogr* 63(1):24–40. <https://doi.org/10.1111/j.1600-0870.2010.00475.x>
- Koletsis I, Kotroni V, Lagouvardos K, Soukissian T (2016) Assessment of offshore wind speed and power potential over the Mediterranean and the Black seas under future climate changes. *Renew Sustain Energy Rev* 60:234–245. <https://doi.org/10.1016/j.rser.2016.01.080>
- Lowe BJ (2000) Polybius 10.10.12 and the existence of salt-flats at Carthago Nova. *Phoenix* 39–52. <https://doi.org/10.2307/1089089>
- Maranzano P (2022) Air quality in Lombardy, Italy: an overview of the environmental monitoring system of ARPA Lombardia. *Earth* 3(1):172–203. <https://doi.org/10.3390/earth3010013>
- Marden JI (2004) Positions and QQ plots. *Stat Sci*, 606–614
- Marsan D, Schertzer D, Lovejoy S (1996) Causal space-time multifractal processes: predictability and forecasting of rain fields. *J Geophys Res: Atmos* 101(D21):26333–26346. <https://doi.org/10.1029/96JD01840>
- Medugorac I, Pasarić M, Güttler I (2021) Will the wind associated with the Adriatic storm surges change in future climate? *Theor Appl Climatol* 143:1–18. <https://doi.org/10.1007/s00704-020-03379-x>
- Miglietta MM, Matsangouras IT (2018) An updated climatology of tornadoes and waterspouts in Italy. *Int J Climatol* 38(9):3667–3683. <https://doi.org/10.1002/joc.5526>
- Nolan P, Lynch P, McGrath R, Semmler T, Wang S (2012) Simulating climate change and its effects on the wind energy resource of Ireland. *Wind Energy* 15(4):593–608. <https://doi.org/10.1002/we.489>
- Outten SD, Esau I (2013) Extreme winds over Europe in the ENSEMBLES regional climate models. *Atmos Chem Phys* 13(10):5163–5172. <https://doi.org/10.5194/acp-13-5163-2013>
- Over TM, Gupta VK (1994) Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large-scale forcing. *J Appl Meteorol Climatol* 33(12):1526–1542. [https://doi.org/10.1175/1520-0450\(1994\)033<1526:SAOMRD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<1526:SAOMRD>2.0.CO;2)
- Ozay C, Celiktaş MS (2016) Statistical analysis of wind speed using two-parameter Weibull distribution in Alaçatı region. *Energy Conv Manag* 121:49–54. <https://doi.org/10.1016/j.enconman.2016.05.026>
- Pant S, Cha EJ (2019) Potential changes in hurricane risk profile across the United States coastal regions under climate change scenarios. *Struct Saf* 80:56–65. <https://doi.org/10.1016/j.rser.2014.10.028>
- Peng X, Deng D, Wen J, Xiong L, Feng S, Wang B (2016) A very short term wind power forecasting approach based on numerical weather prediction and error correction method. *China International Conference on Electricity Distribution (CICED)*, Xi'an, China, 2016, pp. 1–4. <https://doi.org/10.1109/CICED.2016.7576362>
- Pishgar-Komleh SH, Keyhani A, Sefeepari P (2015) Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooh county of Iran). *Renew Sustain Energy Rev* 42:313–322. <https://doi.org/10.1016/j.rser.2014.10.028>
- Prats X, de la Torre D, Delgado L (2022) In-flight cost index optimisation upon weather forecast updates. *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)*, Portsmouth, VA, USA, pp. 1–10. <https://doi.org/10.1109/DASC55683.2022.9925796>
- Pryor SC, Schoof JT, Barthelmie RJ (2006) Winds of change? Projections of near-surface winds under climate change scenarios. *Geophys Res Lett* 33(11). <https://doi.org/10.1029/2006GL026000>
- Reyers M, Pinto JG, Moemken J (2015) Statistical–dynamical downscaling for wind energy potentials: evaluation and applications to decadal hindcasts and climate change projections. *Int J Climatol* 35(2):229–244. <https://doi.org/10.1002/joc.3975>
- Rockel B, Woth K (2007) Extremes of near-surface wind speed over Europe and their future changes as estimated from an ensemble of RCM simulations. *Clim Change* 81(Suppl 1):267–280. <https://doi.org/10.1007/s10584-006-9227-y>
- Schertzer D, Lovejoy S (1987) Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J Geophys Res: Atmos* 92(D8):9693–9714. <https://doi.org/10.1029/JD092iD08p09693>

- Schultz MG, Betancourt C, Gong B, Kleinert F, Langguth M, Leufen LH, Stadler S (2021) Can deep learning beat numerical weather prediction? *Philosophical Trans Royal Soc A* 379(2194):20200097. <https://doi.org/10.1098/rsta.2020.0097>
- Sun W, Tang S, Liu X, Zhou S, Wei J (2022) An Improved Ship Weather Routing Framework for CII reduction accounting for wind-assisted rotors. *J Mar Sci Eng* 10:1979. <https://doi.org/10.3390/jmse10121979>
- Tapley TD, Waylen PR (1990) Spatial variability of annual precipitation and ENSO events in western Peru. *Hydrol Sci J* 35(4):429–446. <https://doi.org/10.1080/02626669009492444>
- Terna (2023) Renewable energy sources in Italy: wind power stations. Available Italian at : <https://www.terna.it/it/sistema-elettrico/dis-pacciamento/fonti-rinnovabili>. Last accessed: 02/22/2023.
- Tobin I, Vautard R, Balog I, Bréon FM, Jerez S, Ruti PM, Yiou P (2015) Assessing climate change impacts on European wind energy from ENSEMBLES high-resolution climate projections. *Clim Change* 128:99–112. <https://doi.org/10.1007/s10584-014-1291-0>
- Tustison B, Foufoula-Georgiou E, Harris D (2002) Scale-recursive estimation for multisensor quantitative precipitation Forecast verification: a preliminary assessment. *J Geophys Research: Atmos* 107(D8):CIP–2. <https://doi.org/10.1029/2001JD001073>
- Vaittinada Ayar P, Vrac M, Bastin S, Carreau J, Déqué M, Gallardo C (2016) Intercomparison of statistical and dynamical downscaling models under the EURO-and MED-CORDEX initiative framework: present climate evaluations. *Clim Dyn* 46:1301–1329. <https://doi.org/10.1007/s00382-015-2647-5>
- Wood Mackenzie (2023) Global wind energy to top 1 TW threshold by the end of 2023. <https://www.woodmac.com/press-releases/global-wind-energy-to-top-1-tw-threshold-by-the-end-of-2023/>, last accessed at: 02/11/2023.
- World Meteorological Organization (2017) Guidelines on the calculation of climate normals (WMO-No. 1203). WMO, Geneva
- Zhang P, Zhang J, Chen M (2017) Economic impacts of climate change on agriculture: the importance of additional climatic variables other than temperature and precipitation. *J Environ Econ Manag* 83:8–31. <https://doi.org/10.1016/j.jeem.2016.12.001>
- Zhang F, Sun YQ, Magnusson L, Buizza R, Lin SJ, Chen JH, Emanuel K (2019) What is the predictability limit of midlatitude weather? *J Atmos Sci* 76(4):1077–1091. <https://doi.org/10.1175/JAS-D-18-0269.1>
- Zhao C, Zhang H, Wang M, Jiang H, Peng J, Wang Y (2021) Impacts of climate change on wind erosion in Southern Africa between 1991 and 2015. *Land Degrad Dev* 32(6):2169–2182. <https://doi.org/10.1002/ldr.3895>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.