ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Gaussian copula modeling of extreme cold and weak-wind events over Europe conditioned on winter weather regimes

To cite this article: Paulina Tedesco *et al* 2023 *Environ. Res. Lett.* **18** 034008

View the article online for updates and enhancements.

You may also like

- Comparison of natural and artificial forcing to study the dynamic behaviour of bell towers in low wind context by means of ground-based radar interferometry: the case of the Leaning Tower in Pisa
Mario Marchisio, Luca Piroddi, Gaetano Ranieri et al.
- Solar Origin of Bare Ion Anomalies in the Solar Wind and Interplanetary Coronal Mass Ejections
Yeimy J. Rivera, Susan T. Lepri, John C. Raymond et al.
- GLOBAL X-RAY PROPERTIES OF THE O AND B STARS IN CARINA
Y. Nazé, P. S. Broos, L. Oskinova et al.

This content was downloaded from IP address 158.36.47.138 on 05/10/2023 at 09:45

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Gaussian copula modeling of extreme cold and weak-wind events over Europe conditioned on winter weather regimes

Paulina Tedesco[1,2,*], **Alex Lenkoski**[3], **Hannah C Bloomfield**[4,5] and **Jana Sillmann**[6,7]

[1] Department of Physics, University of Oslo, PO box 1048, Blindern, 0316 Oslo, Norway
[2] Information Technology Department, Norwegian Meteorological Institute, Oslo, Norway
[3] Statistical Analysis and Machine Learning, Norwegian Computing Center, Oslo, Norway
[4] School of Geographical Sciences, University of Bristol, Bristol, United Kingdom
[5] Department of Meteorology, University of Reading, Reading, United Kingdom
[6] Center for Earth System Research and Sustainability, Research Unit for Sustainability and Climate Risks, Universität Hamburg, Hamburg, Germany
[7] CICERO Center for International Climate Research, Oslo, Norway
[*] Author to whom any correspondence should be addressed.

E-mail: paulina.tedesco@gmail.com

Keywords: Gaussian copula, extreme compound event, logistic regression, weather regimes, renewable energy, wind power, subseasonal variability

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
31 August 2022

REVISED
25 January 2023

ACCEPTED FOR PUBLICATION
27 January 2023

PUBLISHED
14 February 2023

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Abstract

A transition to renewable energy is needed to mitigate climate change. In Europe, this transition has been led by wind energy, which is one of the fastest growing energy sources. However, energy demand and production are sensitive to meteorological conditions and atmospheric variability at multiple time scales. To accomplish the required balance between these two variables, critical conditions of high demand and low wind energy supply must be considered in the design of energy systems. We describe a methodology for modeling joint distributions of meteorological variables without making any assumptions about their marginal distributions. In this context, Gaussian copulas are used to model the correlated nature of cold and weak-wind events. The marginal distributions are modeled with logistic regressions defining two sets of binary variables as predictors: four large-scale weather regimes (WRs) and the months of the extended winter season. By applying this framework to ERA5 data, we can compute the joint probabilities of co-occurrence of cold and weak-wind events on a high-resolution grid (0.2°).Our results show that (a) WRs must be considered when modeling cold and weak-wind events, (b) it is essential to account for the correlations between these events when modeling their joint distribution, (c) we need to analyze each month separately, and (d) the highest estimated number of days with compound events are associated with the negative phase of the North Atlantic Oscillation (3 days on average over Finland, Ireland, and Lithuania in January, and France and Luxembourg in February) and the Scandinavian blocking pattern (3 days on average over Ireland in January and Denmark in February). This information could be relevant for application in sub-seasonal to seasonal forecasts of such events.

1. Introduction

Affordable and clean energy is one of the UN Sustainable Development Goals (SDGs). Energy is also crucial for achieving many of the other SDGs. The energy sector currently accounts for more than two-thirds of the global greenhouse gas emissions [1]. Consequently, a decarbonization of the energy sector is required to meet the SDGs and the Paris agreement [2], by increasing the share of renewable power generation [3]. Europe is leading this transition, although it is still one of the world's biggest energy consumers and greenhouse gas emitters [4].

The production of clean energy is highly weather-dependent. Very cold or warm temperatures increase the demand due to heating and cooling, respectively [5]. Extreme weather conditions, such as European blocking in wintertime, can lead to high electricity

© 2023 The Author(s). Published by IOP Publishing Ltd

demand and low renewable power production (energy shortfall). How to deal with periods of low renewable production is a big challenge in the design of renewable energy systems [6]. Better subseasonal and seasonal forecasts help improve decision-making and planning [e.g. 7]. They are of value for power producers to better prepare for extreme meteorological events, and could be an important planning tool for traders, plant operators, and investors for managing climate variability related risk [8], as they provide relevant information for price forecasting [9]. Although studies have traditionally focused on single drivers [10], more recent research is considering the interactions between them [e.g. 11–16].

Given that electricity production and demand depend on the weather, they vary on multiple timescales [e.g. 17–19]. Hence, it is important to understand how large-scale circulation systems influence peaks of demand and energy production to identify periods of over- and under-supply. Over time, the impact of large-scale atmospheric patterns on extreme winter temperatures mostly does not change, despite slow fluctuations of these patterns [20]. Meteorological conditions preceding energy shortfall are described as anomalous high pressure systems combined with below normal temperatures [11, 21, 22]. Moreover, moderate energy droughts (with duration and severity exceeding the 75th percentile), are expected every half a year [16]. Further research is still needed to understand the impact of atmospheric variability on surface variables that combined lead to low supply. As far as we know, no other research has proposed a meteorological based methodology to model the bivariate probabilities of cold spells and weak-winds.

Several indices exist that describe European climate variability through daily synoptic-scale weather patterns. A method that has been proven to be useful in weather forecasting and climate change applications is the computation of weather regimes (WRs) with the k-means algorithm [e.g. 23–27]. Understanding the surface impact of these WRs is relevant for subseasonal to seasonal energy applications [11, 12, 28, 29]. Furthermore, the response of power systems to these patterns across Europe has recently been studied, with a focus on the North Atlantic Oscillation (NAO) [e.g. 22, 30–33].

Modeling the dependence between demand and energy production is of key importance to understand the occurrence of energy shortfall and prevent it by redesigning the energy systems. However, modeling multivariate distributions can be a challenge. Copulas offer a powerful and flexible tool that returns the joint probability of events as a function of the marginal probabilities of each event. This makes copulas attractive, as the univariate marginal behavior of random variables can be modeled separately from their dependence. The semi-parametric Bayesian Gaussian copula method used here estimates multivariate relationships between variables with univariate marginal distributions that cannot be well approximated with a simple parametric model [34], making it a promising candidate for modeling the joint probabilities of cold and weak-wind events. This method has been used in the post-processing of a multivariate mesoscale weather forecast [35].

Our goal is to develop a methodology for diagnosing the probabilities of meteorological compound events associated with high energy shortfall in the winter, when electric heating is highest [21]. In this study, we focus only on wind power and do not assess solar or hydro. We apply the methodology to cold and weak-wind events over Europe and analyze the results obtained over the continent, where local onshore wind power generation can meet a fraction of the demand. First, we show that the WRs improve the estimation of marginal probabilities of low temperature and weak wind events and that these events are correlated. Then, with Gaussian copulas, we model the dependency between cold and weak-wind events as a function of their marginals, separately from their dependence, and we show where the models are significantly improved when accounting for the correlations between the events. Furthermore, the strong spatial and intraseasonal variability indicates that we need to analyze every month separately on a grid level.

The paper proceeds as follows. Section 2 describes the data and the methodology developed to compute the probabilities of cold and weak-wind events. Section 3 presents the univariate and bivariate probabilities, and the number of expected compound events. Finally, discussions and conclusions are provided in section 4. Additional figures and their respective analysis can be found in the supplementary material.

2. Data and methods

This section describes the data and methods used to estimate joint probabilities of cold and weak-wind events conditioned on the WRs using a Gaussian copula framework.

2.1. Reanalysis data

Temperature is well established as the main weather driver of electricity demand. We use hourly 2 m temperature and 10 m wind data from the ERA5 reanalysis [36], produced by the European Centre for Medium-Range Weather Forecasts. The data are provided on a regular latitude-longitude grid at $0.25° \times 0.25°$, and the period covered in this work is from 1979–2021. We compute daily minimum temperatures from these hourly data. Wind speeds are calculated from the eastward (u) and northward (v) components at an hourly frequency and then aggregated to daily maximum values. Daily minimum wind

speeds are usually close to zero, but low daily maximum values represent well conditions for poor wind power generation. Despite the good coverage and long records, and the fact that ERA5 outperforms other global reanalyses frequently used in the literature, it is discouraged to use global reanalyses to estimate mean winds because of the uncertainty in these data [37].

2.2. Weather regimes

Atmospheric circulation is well known for its variability at multiple time scales being reflected in weather patterns and circulation systems. These WRs are quasi-stationary large-scale circulation patterns [38]. They typically persist for 6–10 days, are spatially well defined, and limited in number. The understanding of the causes of their recurrence, persistence, and transition is crucial for medium-range, subseasonal, and seasonal-to-interannual climate prediction [27, 29], as they influence the weather at the surface and, subsequently, the renewable power generation and electricity demand [12, 31, 39].

WRs are obtained here by applying the k-means algorithm with four centroids on geopotential height at 500 hPa (Z500) data [26, 27]. The algorithm leads to four winter regimes in the Euro-Atlantic area (27°N–81°N, 85.5°W–45°E), and has been adopted by other authors [e.g. 11, 12]. The Z500 fields of the obtained regimes are similar, although not identical, to the teleconnection patterns defined by the National Oceanic and Atmospheric Administration, and can be interpreted as the negative and positive phases of the NAO, the Atlantic Ridge (AR), and the Scandinavian blocking (SCAND) (see figure 1 in the supplementary material). We will therefore use these names to refer to the four WRs. Moreover, it has been shown that temporal sub-sampling [27] and the use of different reanalysis data [11, 12] do not change the spatial structure of the regimes nor the optimal partition ($k = 4$).

The classification method consists of two steps. First, a cosine weight as a function of latitude is applied to the Z500 data and the first fourteen empirical orthogonal functions are computed. The associated principal components (PCs) were used as coordinates of a reduced phase space. Then, the PCs are clustered into four groups with the K-means algorithm, choosing 30 random starts and a maximum of 100 iterations. Every daily map is assigned to a centroid based on its closest Euclidean distance. The number of regimes, $k = 4$, corresponds to the most robust regime partition during winter months [26].

2.3. Logistic regression

In this study, we use logistic regressions to classify cold and weak-wind events. The predictors consist of two sets of indicator variables, one for the five winter months and one for the four WRs. The variables in each group are dichotomous and take only the values 1 or 0. The univariate distributions will later be used in the copula framework to estimate the joint probabilities.

Logistic regressions are widely used linear models for binary classification. Let Y be the binary outcome variable indicating failure or success with $0, 1$. Then, p stands for the probability of a positive event, i.e. $p = P(Y = 1)$. The mathematical expression of the logit function is:

$$logit(p) = log\left(\frac{p}{1-p}\right). \tag{1}$$

It is assumed that the logit transformation of the outcome variable has a linear relationship with the predictor variables: $x_1, x_2, \ldots, x_k$. Then $\beta_0, \beta_1, \ldots, \beta_k$ are the parameters estimated via the maximum likelihood method when performing a logistic regression of Y on $x_1, x_2, \ldots, x_k$:

$$logit(p) = log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k. \tag{2}$$

We are usually interested in predicting the probability that a particular sample belongs to a particular class. The formula of the probability $P(Y = 1)$ is:

$$p = \frac{e^{\beta_0+\beta_1 x_1+\ldots+\beta_k x_k}}{1+e^{\beta_0+\beta_1 x_1+\ldots,\beta_k x_k}} = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\ldots+\beta_k x_k)}}. \tag{3}$$

2.4. Copulas

Our goal is to establish a methodology for modeling the probabilities of co-occurrence of cold temperatures and weak winds. However, estimating joint densities is not an easy task since only a few non-Gaussian families are defined, and non-parametric estimation is demanding. Nonetheless, density estimation in one variable is relatively easy, given that many convenient families exist (e.g. logistic, exponential, uniform) and that the non-parametric approach is efficient and accurate. The copulas framework for modeling multivariate distributions provides a flexible representation and separates the univariates from their true nature of dependence.

In the field of probability theory and statistics, a copula function $C : [0,1]^n \to [0,1]$ is defined as a multivariate distribution:

$$C(u_1, u_2, \ldots, u_n) = P(U_1 \leqslant u_1, U_2 \leqslant u_2, \ldots U_n \leqslant u_n), \tag{4}$$

such that marginalizing gives $U_i \sim uniform(0,1)$. Copulas are useful because we can transform any arbitrary random variable into a uniform and back. The function that transforms uniforms to any other univariate distribution is the inverse of the cumulative distribution function (CDF). To do the opposite transformation, from an arbitrary distribution to

the uniform(0, 1), we apply the inverse of the inverse CDF, the CDF.

2.4.1. Gaussian copulas

Copula theory ensures that, for every joint multivariate distribution, there exists a unique copula. In the case of the Gaussian copula function, finding its parameters is limited to finding the correlation matrix of the random variables we want to study.

A Gaussian copula is given by:

$$C(u_1, u_2, \ldots, u_n) = \Phi_\Sigma \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right), \tag{5}$$

where Φ_Σ represents the CDF of a multivariate normal with covariance Σ and mean 0, and Φ^{-1} is the inverse CDF for the standard normal.

Given a multivariate distribution:

$$\begin{aligned} F_X(X) &= P(X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_n \leqslant x_n) \\ &= \Phi_\Sigma(x_1, x_2, \ldots, x_n), \end{aligned} \tag{6}$$

we can extract its Gaussian copula:

$$\begin{aligned} F_X(X) &= \Phi_\Sigma \left(F_1^{-1}(F_1(X)), F_2^{-1}(F_2(X)), \ldots, F_n^{-1}(F_n(X))\right) \\ &= \Phi_\Sigma \left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n)\right) \\ &= \Phi_\Sigma \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right) \\ &= C\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right), \end{aligned} \tag{7}$$

and plug in any marginal into the copula function.

The inverse CDF transforms the uniforms into normal distributions. Then, the multivariate normal's CDF squashes the distribution so that the marginals are uniform and with Gaussian correlations. Thus, the Gaussian Copula is a distribution over the unit hypercube $[0, 1]^n$ with uniform marginals.

2.4.2. Semiparametric copula estimation

It is often the case that the marginal distributions do not belong to standard families. In such cases, it might be appropriate to use a semi-parametric strategy that involves representing the associations among variables with a simple parametric approach and estimating the marginals nonparametrically.

We use an extended rank likelihood method of semiparametric inference for copula, which is a function of the association parameters only [34]. It can be applied without any assumptions of the marginal distributions, making it appropriate for the joint analysis of continuous and ordinal discrete data. The *R* package 'sbgcop' [40] provides a tool for estimation and inference for the Gaussian copula parameters.

2.5. Model performance

We use permutation tests to evaluate the significance of the results at a 10% level in terms of the BSS. In the logistic regression framework, the BSS is used to assess the performance of models against a seasonal baseline model that does not include any information about the WRs (the only predictors are the indicator variables for the months). The BSS is also used to evaluate the copula model's performance against an independent model with no correlation.

The most common measure for probabilistic forecasts is the Brier score (BS) [41, 42]. It is assumed that the events only can occur in one class on each of the n occasions. For dichotomous events, the forecast probabilities are f_i for an occasion i and the BS is:

$$BS = \frac{1}{n}\sum_{i=1}^{n}(f_i - E_i)^2, \tag{8}$$

where E_i, takes the value 1 if the event occurred and 0 otherwise. The score is negatively oriented, meaning that a perfect forecast has a $BS = 0$. Less accurate forecasts exhibit higher scores, but since individual forecasts and observations are both bounded by zero and one, the range of possible values for the BS is $0 \leqslant BS \leqslant 1$.

The BSS is often used and, since $BS_{perfect} = 0$, it takes the form:

$$BSS = \frac{BS - BS_{ref}}{0 - BS_{ref}} = 1 - \frac{BS}{BS_{ref}}. \tag{9}$$

Negative values mean that the forecast is less accurate than the reference forecast; when the forecast presents no skill compared to the reference $BSS = 0$; and a perfect skill compared to the reference forecast reflects in a skill score equal to 1.

3. Results

We first illustrate the marginal probabilities. Then, we analyze the joint probabilities, which can be interpreted as the percentage of days we expect compound events for a given month and regime.

3.1. Marginal probabilities of cold and weak-wind events

Marginal probabilities of cold events and weak-wind events are modeled for each pair of regime and month using logistic regressions.

Figure 1 illustrates the regions where the BSS computed for cold events is significant, i.e. where the WRs provide relevant information to the logistic model. It shows, in addition, the probabilities of observing temperatures lower than the 10th in these regions. What stands out from figure 1 is that regions with significant BSS cover extended areas. Even when probabilities are close to zero, like in November or

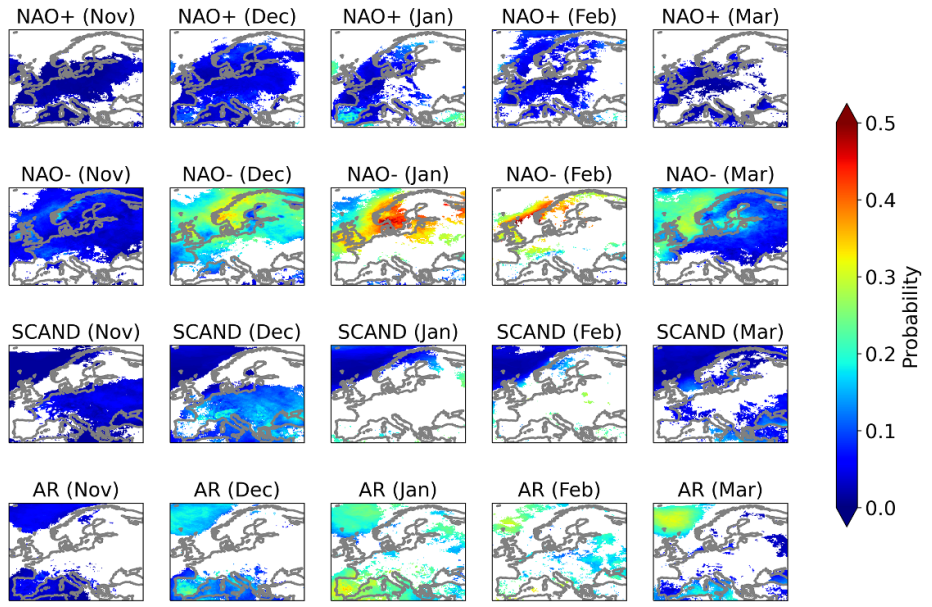

Figure 1. Marginal probabilities of cold events where the BSS is significant. Colors show the probabilities of cold events for each month and WR estimated with logistic regressions. Cold events are defined for each grid box as days with daily minimum temperature below 10th percentile computed for the entire winter season (NDJFM). Probability values where the BSS is negative or not significant at a 10% level are masked out. The BSS was computed with respect to a climatology model. Area of study: 35°N–72°N, 11°W–40°E. Figure based on ERA5 data (1979–2021).

over continental areas in March, considering the WRs improves the performance of the logistic regressions used to model cold events. If we focus on continental areas where the BSS is significant, we observe probability values higher than 0.3 over Scandinavia for NAO− and over the Iberian Peninsula for AR. Furthermore, although the areas with significant BSS are similar throughout the season for a given regime, the strong intraseasonal variability of the probabilities indicates that it is convenient to analyze the temperature events for each month separately rather than for the whole season.

A strong negative NAO index is usually associated with cold winter temperatures across Northern Europe and Northern Asia, which agrees with the pattern observed for the NAO− cluster in figure 1. Provided that the 10th percentile was computed at each grid box for the entire season (NDJFM), the highest probabilities are observed during the coldest months. Probabilities are higher than 0.2 from December to February over land, in smaller areas of Northern Europe and Scandinavia, and from December to March over the sea, which we speculate is due to a slower response of the water.

The AR pattern is characterized by an anticyclonic circulation over the Atlantic with a weak surface temperature anomaly over Western Europe (see figure 2 in the supplementary material). Despite the weak anomaly pattern, the AR is the regime with the second highest significant probabilities of cold events, reaching local probability values higher than 0.3 across the Iberian Peninsula in January. Furthermore, the pattern shows a significant signal in the North Atlantic, with probabilities increasing from November to March. Still, we are mostly concerned about strong signals over land, which is the case of the Iberian Peninsula in December and January.

The other two regimes, SCAND and NAO+, exhibit probabilities close to zero when the BSS is significant and positive. Thus, for NAO+ and SCAND, the model is only improved by the WRs in regions where cold events are less likely to occur.

Figure 2 is analogous to figure 1 but has been computed for daily maximum wind speeds lower than the 10th percentile. This variable presents persistent zonal bands of significant probabilities at different latitudes when NAO+, NAO−, or SCAND are dominant. However, significant probabilities of weak-wind events are lower than 0.3 for all the regimes. The WRs improve the results mainly over the North Atlantic and Scandinavia during NAO−, and in Northern and Central Europe during SCAND. On the other hand, NAO+ shows two zonal bands of significant probabilities, one with probabilities under 0.2 over the Mediterranean and one with probabilities close to zero over Northern Europe.

In general, we observe that the marginal probabilities obtained for the NAO−, NAO+, and SCAND patterns support earlier results [11, 12]. However, the monthly probabilities calculated for AR conditions

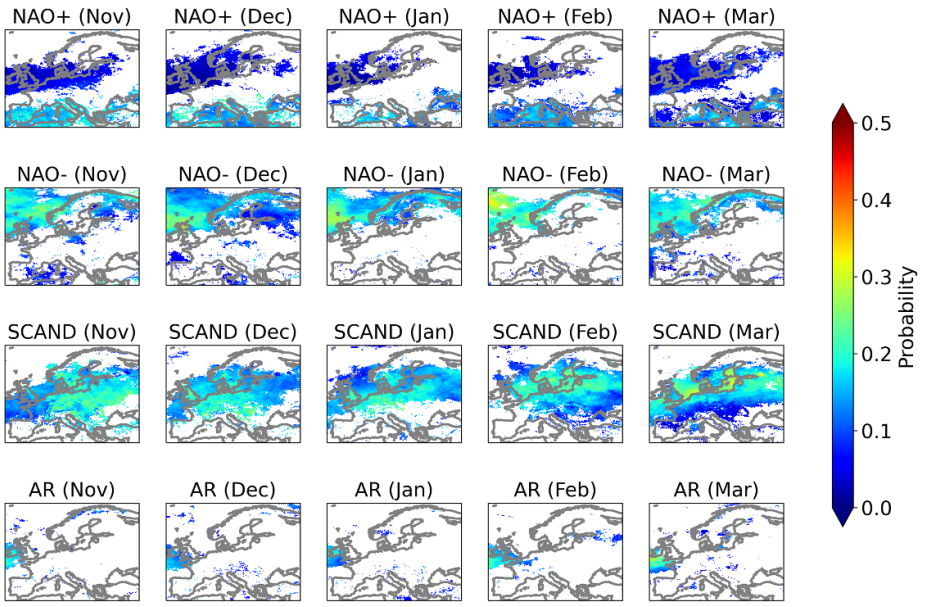

Figure 2. Marginal probabilities of weak-wind events where the BSS is significant. Colors show the probabilities of weak-wind events for each month and WR estimated with logistic regressions. Weak-wind events are defined for each grid box as days with daily maximum wind speeds below 10th percentile computed for the entire winter season (NDJFM). Probability values where the BSS is negative or not significant at a 10% level are masked out. The BSS was computed with respect to a climatology model. Area of study: 35°N–72°N, 11°W–40°E. Figure based on ERA5 data (1979–2021).

indicate a stronger response of the temperatures to large-scale upper-level circulation than the seasonal aggregated results reported in previous studies [11, 12]. Given that November is the warmest month of the winter season and fewer temperature events are registered, the results for this month excluded from the analysis in the following. We have also excluded the results for NAO+ for the same reason. In addition, figures of unmasked marginal probabilities are provided in the supplementary material for both variables.

3.2. Joint probabilities of cold and weak-wind events

The Gaussian copula framework makes it possible to generate a joint distribution from arbitrary marginal probabilities taking as an argument the covariance matrix (illustrated in the supplementary material). In this case, the BSS was computed with respect to a model with no correlations.

Significant joint probabilities of cold and weak-wind events are shown in figure 3. Here, we mean by significant that the correlations estimated with the rank likelihood method [40] improve the results. This occurs in smaller regions of Northern and Central Europe when NAO−, SCAND, or AR conditions are dominant, in particular, in January and February. However, probabilities higher than 0.1, where the BSS is significant, occur only for NAO− in January and SCAND in February at specific locations. For NAO+ (not shown), the correlations do not significantly improve the results.

3.3. Estimated number of days with compound cold and weak-wind events

By considering the number of days in each month, we can translate the joint probabilities into the number of days we expect compound events.

Figure 4 shows that the number of events varies with the regime, month, and region considered. Two regimes stand out during the coldest months: NAO− and SCAND. In December, at a country level, compound events are expected to occur 0–1 on average in most countries. For NAO−, this number rises to 2 days in France, Belgium, Switzerland, and the UK and 3 days in Ireland. In January, during NAO− and SCAND, several countries in Northern Europe and Scandinavia experience compound events on average 2–3 days. We also see that these two patterns have a strong north-south gradient. Contrarily, when AR dominates, there is no evident gradient, and the number of compound events is 1–2 in the Iberian Peninsula, Ireland, the UK, and Scandinavia. The highest number of average days per country is estimated for France, Luxembourg, and Belgium during NAO− conditions in February, and local values of 5–6 days are observed in France. Also, in February, but during SCAND, an average of 2 days is expected in the Baltic countries and Poland. Finally, in March, the average number of expected compound events per country drops again to 0–1, except in Estonia, where the average is 2 days. We highlight the spatial and temporal variability of the results. Although the maximum average number of days at a country level is 3, and only for a few combinations of WR and month, figure 4 shows smaller regions where we

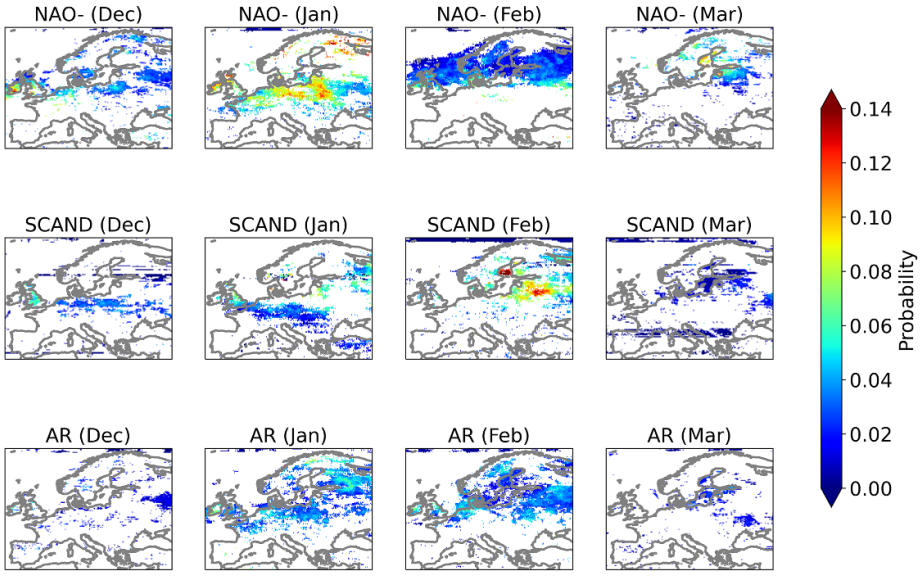

Figure 3. Joint probabilities of cold and weak-wind events where the BSS is significant. Colors show the probabilities of co-occurrence of cold and weak-wind events for each WR and month. The events were computed for daily minimum temperatures and maximum wind speeds below the 10th percentile of the data in the extended winter season (NDJFM). Probability values where the BSS is not significant at a 10% level are masked out. The BSS was computed with respect to a model with no correlation between cold and weak-wind events. Area of study: 35°N–72°N, 11°W–40°E. Figure based on ERA5 data (1979–2021).

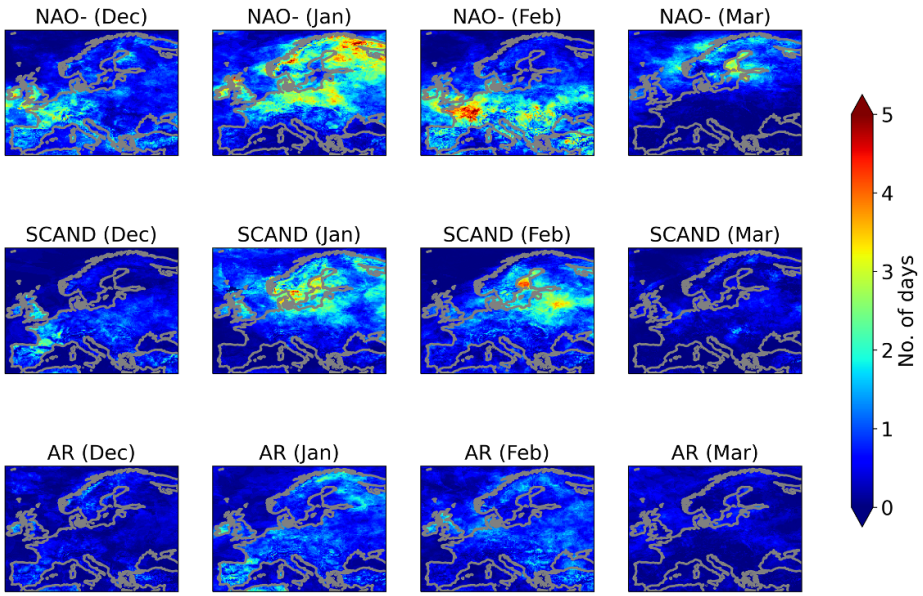

Figure 4. Number of compound cold and weak wind events. The events were computed as the daily minimum temperatures and maximum wind speeds below the 10th percentile of the data in the extended winter season (NDJFM). Area of study: 35°N–72°N, 11°W–40°E. Figure based on ERA5 data (1979–2021).

expect compound events 5 days a month in January and February associated with blocking conditions.

Overall, figure 4 supports what previous studies have found: (a) that large-scale circulation patterns influence the energy demand [22, 30, 31] and the wind power generation over Europe [22, 32, 33, 43], and (b) that blocking patterns are associated with anomalous cold and low-wind speed conditions over Northern and Central Europe, which, in turn, increase the demand and reduce the wind power generation [22, 30, 33, 43–45].

4. Discussion and conclusions

Low temperatures are associated with high electricity demand during the winter, while wind speed is directly related to electricity production. The position of anomalous pressure systems and planetary waves

disturb the zonal flow at 500 hPa, which, in turn, influences the progression of WRs that impact surface variables, such as temperature and wind speed. Here, we develop a Gaussian copula-based framework for modeling monthly bivariate distributions of cold weak-wind events conditioned on large-scale WRs. This flexible methodology can be adapted to different univariate distribution models.

Blocking conditions usually cause above-average demand and below-average wind and solar generation in Central and Northern Europe [11, 39]). Composites of temperature and wind speed indicate that the impacts of NAO− and SCAND on near-surface variables are associated with cold and weak-wind conditions [11, 12]. Thus, more extreme events are expected for these two clusters. We confirmed this hypothesis by analyzing the number of past events divided by the possible outcomes, which provides insight into how many events we can expect given the defined threshold instead of relying on the departures from average conditions, as most studies do.

First, we estimated the correlations with the 'sbgcop' package and modeled the marginals probabilities with logistic regressions using two sets of binary predictors, one for the months in the winter season and one for the WRs. Permutation tests were used to assess the performance of the logistic models in terms of the BSS computed with respect to a seasonality model that only considers the set of monthly predictors. The results indicate that incorporating the WR information into the models improves the performance over large geographical areas. This means that large-scale systems impact the occurrence of extreme temperature and wind speed events and, consequently, the peaks of demand and production of energy.

Then, we computed the joint probabilities with Gaussian copulas using the estimated correlations and marginal probabilities. The BSS for joint probabilities reveals an improved skill in smaller regions when accounting for the associations between the event variables, mainly in latitudes higher than 45°N. It follows from the calculation of the joint probabilities that the co-occurrence of cold and weak-wind events is restricted to specific geographical areas, depending on the month and the predominant WR. These probabilities have been converted to the number of days we expect compound events for a given month and WR. Overall, the highest number of days is expected during NAO− and SCAND from December to February, in particular, in coastal and inland regions in Northern Europe with high wind farm density. The results also expose the need to consider each month separately when analyzing compound events of cold and weak winds. We conclude that the gridded results exhibit a spatial variability of extreme events that is not captured by the national aggregates presented in other studies [e.g. 28].

There is an unavoidable degree of uncertainty affecting reanalysis datasets; the most significant disagreements in DJF are encountered within continental areas [37]. Representativeness can also be a problem because wind farms are often located in places where the wind is stronger than its surroundings, and thus the mean value of the grid box might be inaccurate. Moreover, one of the challenges of working with wind speed data from reanalyses is that they are available at 10 meters, whereas typical hub heights are 80–120 meters. Therefore, an extrapolation method is typically used to estimate hub height winds from surface winds. The power law [e.g. 21] was tested for the computation of marginal weak-wind probabilities (not shown) but discarded because it adds even more uncertainty to the already biased winds.

The scope of this study was limited to the analysis of meteorological variables associated with poor conditions for onshore wind power generation during periods of high demand. It is well known that the relationship between wind speed and wind power is nonlinear. Still, wind power is largely determined by wind speed, and our results are, therefore, useful for understanding problems of undersupply. Furthermore, we believe our work will contribute to designing more robust energy systems for periods of co-occurrence of low wind energy generation and high demand.

The skill could be improved by modeling the marginal probabilities with more complex algorithms and considering more variables. Future studies should also explore how solar and hydro can complement wind power generation in situations with high probabilities of co-occurrence of cold and weak-wind events. We also recommend examining the applicability of this methodology to energy demand and energy shortfall variables commonly used in the literature [e.g. 12, 16, 21]. These variables have been studied in a copula-based framework [16], but with a focus on seasonal values aggregated at a country level. The pronounced subseasonal and spatial variability observed in the present study suggests that decision-makers would further benefit from monthly results on a grid when analyzing energy demand and shortfall variables.

In summary, the results confirm that the dependence between the wind speed and temperature event variables plays an essential role in modeling compound meteorological events and that low-frequency circulation patterns control parts of the distribution of extreme events. We also provide evidence of strong intraseasonal and spatial variability of compound events.

Data availability statement

No new data were created or analysed in this study.

Acknowledgment

The authors have received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 776787 (S2S4E). This work has also been supported by the Research Council of Norway through grants 303411 (Machine Ocean) and 309562 (Climate Futures).

This research was conducted using computational resources provided by the Norwegian Meteorological Institute and the University of Oslo.

The authors would also like to thank Frode Stordal for interesting discussions and Jean Rabault for encouraging the publication of this research.

ORCID iDs

Paulina Tedesco https://orcid.org/0000-0002-0941-4766
Alex Lenkoski https://orcid.org/0000-0002-6664-0292
Hannah C Bloomfield https://orcid.org/0000-0002-5616-1503
Jana Sillmann https://orcid.org/0000-0002-0219-5345

References

[1] IRENA 2020 Global renewables outlook: energy transformation 2050 *Abu Dhabi*
[2] Sawyer S, Fried L, Shukla S, and Liming Q 2017 Global wind report 2016 - annual market update 04
[3] Rogelj J, Luderer G, Pietzcker R C, Kriegler E, Schaeffer M, Krey V and Riahi K 2015 Energy system transformations for limiting end-of-century warming to below 1.5 °C *Nat. Clim. Change* **5** 519–27
[4] Liobikienė G and Butkus M 2017 The European Union possibilities to achieve targets of Europe 2020 and Paris agreement climate policy *Renew. Energy* **106** 298–309
[5] Taylor J and Buizza R 2003 Using weather ensemble predictions in electricity demand forecasting *Int. J. Forecast.* **19** 57–70
[6] Huber M, Dimkova D and Hamacher T 2014 Integration of wind and solar power in Europe: assessment of flexibility requirements *Energy* **69** 236–46
[7] Orlov A, Sillmann J and Vigo I 2020 Better seasonal forecasts for the renewable energy industry *Nat. Energy* **5** 108–10
[8] Cortesi N, Torralba Vonica, González-Reviriego N, Soret A and Doblas-Reyes F J 2019 Characterization of European wind speed variability using weather regimes *Clim. Dyn.* **53** 4961–76
[9] Pinson P 2013 Wind energy: forecasting challenges for its operational management *Statistical Science* **28** 564–85
[10] Zscheischler J *et al* 2018 Future climate risk from compound events *Nat. Clim. Change* **8** 469–77
[11] van der Wiel K, Bloomfield H C, Lee R W, Stoop L P, Blackport R, Screen J A and Selten F M 2019 The influence of weather regimes on European renewable energy production and demand *Environ. Res. Lett.* **14** 094010
[12] Bloomfield H C, Brayshaw D J and Charlton-Perez A J 2020 Characterizing the winter meteorological drivers of the European electricity system using targeted circulation types *Meteorol. Appl.* **27** e1858
[13] Drücke J, Borsche M, James P, Kaspar F, Pfeifroth U, Ahrens B and Trentmann Jorg 2021 Climatological analysis of solar and wind energy in Germany using the Grosswetterlagen classification *Renew. Energy* **164** 1254–66
[14] Kaspar F, Borsche M, Pfeifroth U, Trentmann J, Drücke J and Becker P 2019 A climatological assessment of balancing effects and shortfall risks of photovoltaics and wind energy in Germany and Europe *Adv. Sci. Res.* **16** 119–28
[15] Turner S W D, Voisin N, Fazio J, Hua D and Jourabchi M 2019 Compound climate events transform electrical power shortfall risk in the Pacific Northwest *Nat. Commun.* **10** 1–8
[16] Felipe N O, Martius O, Allen S, Bloomfield H and Schaefli B 2021 A copula-based assessment of renewable energy droughts across Europe *Renew. Energy* **201** 667–77
[17] Sinden G 2007 Characteristics of the UK wind resource: long-term patterns and relationship to electricity demand *Energy Policy* **35** 112–27
[18] Bessec M and Fouquau J 2008 The non-linear link between electricity consumption and temperature in Europe: a threshold panel approach *Energy Econ.* **30** 2705–21
[19] Bloomfield H C, Brayshaw D J, Shaffrey L C, Coker P J and Thornton H E 2016 Quantifying the increasing sensitivity of power systems to climate variability *Environ. Res. Lett.* **11** 124025
[20] Schuhen N, Schaller N, Bloomfield H C, Brayshaw D J, Lledó Lç, Cionni I and Sillmann J 2022 Predictive skill of teleconnection patterns in twentieth century seasonal hindcasts and their relationship to extreme winter temperatures in Europe *Geophys. Res. Lett.* **49** e2020GL092360
[21] Karin van der W, Stoop L, van Zuijlen B, Blackport R, Broek M and Selten F 2019 Meteorological conditions leading to extreme low variable renewable energy production and extreme high energy shortfall *Renew. Sustain. Energy Rev.* **111** 261–75
[22] Bloomfield H C, Brayshaw D J, Shaffrey L C, Coker P J and Thornton H E 2018 The changing sensitivity of power systems to meteorological drivers: a case study of Great Britain *Environ. Res. Lett.* **13** 054028
[23] Neal R, Fereday D, Crocker R and Comer R E 2016 A flexible approach to defining weather patterns and their application in weather forecasting over Europe *Meteorol. Appl.* **23** 389–400
[24] Ferranti L, Corti S and Janousek M 2015 Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector *Q. J. R. Meteorol. Soc.* **141** 916–24
[25] Matsueda M and Palmer T N 2018 Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts *Q. J. R. Meteorol. Soc.* **144** 1012–27
[26] Michelangeli P-A, Vautard R and Legras B 1995 Weather regimes: recurrence and quasi stationarity *J. Atmos. Sci.* **52** 1237–56
[27] Cassou C 2008 Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation *Nature* **455** 523–7
[28] Bloomfield H C, Brayshaw D J, Gonzalez P L M and Charlton-Perez A 2021 Pattern-based conditioning enhances sub-seasonal prediction skill of European national energy variables *Meteorol. Appl.* **28** e2018
[29] Cassou C 2010 Euro-Atlantic regimes and their teleconnections *ECMWF Annual Seminar 2010*
[30] Ely C R, Brayshaw D J, Methven J, Cox J and Pearce O 2013 Implications of the North Atlantic Oscillation for a UK-Norway renewable power system *Energy Policy* **62** 1420–7
[31] Thornton H, Scaife A, Hoskins B and Brayshaw D 2017 The relationship between wind power, electricity demand and winter weather patterns in Great Britain *Environ. Res. Lett.* **12** 064017
[32] 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 *Renew. Energy* **36** 2087–96

[33] Zubiate L, McDermott F, Sweeney C and O'Malley M 2017 Spatial variability in winter NAO–wind speed relationships in western Europe linked to concomitant states of the east atlantic and Scandinavian patterns *Q. J. R. Meteorol. Soc.* **143** 552–62
[34] Hoff P D 2007 Extending the rank likelihood for semiparametric copula estimation *Ann. Appl. Stat.* **1** 265–83
[35] Möller A, Lenkoski A and Thorarinsdottir T L 2013 Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas *Q. J. R. Meteorol. Soc.* **139** 982–91
[36] Hersbach H, Bell B, Berrisford P, Horányi A J M S, Muñoz Sabater J, Nicolas J, Radu R, Schepers D, Simmons A and Soci C *et al* 2019 Global reanalysis: goodbye ERA-Interim, hello ERA5 *ECMWF Newsl.* **159** 17–24
[37] Ramon J, Lledó Lç, Torralba Vonica, Soret A and Doblas-Reyes F J 2019 What global reanalysis best represents near-surface winds? *Q. J. R. Meteorol. Soc.* **145** 3236–51
[38] Reinhold B B and Pierrehumbert R T 1982 Dynamics of weather regimes: quasi-stationary waves and blocking *Mon. Weather Rev.* **110** 1105–45
[39] Grams C, Beerli R, Pfenninger S, Staffell I and Wernli H 2017 Balancing Europe's wind-power output through spatial deployment informed by weather regimes *Nat. Clim. Change* **7** 557–62
[40] Hoff P 2018 sbgcop: semiparametric Bayesian Gaussian copula estimation and imputation R package version 0.980
[41] Brier G W 1950 Verification of forecasts expressed in terms of probability *Mon. Weather Rev.* **78** 1–3
[42] Wilks D S 2006 *Statistical Methods in the Atmospheric Sciences* (*International Geophysics*) (San Diego, CA: Elsevier Science)
[43] Cradden L C, McDermott F, Zubiate L, Sweeney C and O'Malley M 2017 A 34-year simulation of wind generation potential for Ireland and the impact of large-scale atmospheric pressure patterns *Renew. Energy* **106** 165–76
[44] Ravestein P, Van der Schrier G, Haarsma R, Scheele R and Van den Broek M 2018 Vulnerability of European intermittent renewable energy supply to climate change and climate variability *Renew. Sustain. Energy Rev.* **97** 497–508
[45] 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 *Environ. Res. Lett.* **8** 044035