



**HAL**  
open science

# Assessing the Probability of Extremely Low Wind Energy Production in Europe at Sub-seasonal to Seasonal Time Scales

Bastien Cozian, Corentin Herbert, Freddy Bouchet

► **To cite this version:**

Bastien Cozian, Corentin Herbert, Freddy Bouchet. Assessing the Probability of Extremely Low Wind Energy Production in Europe at Sub-seasonal to Seasonal Time Scales. 2023. hal-04305658

**HAL Id: hal-04305658**

**<https://hal.science/hal-04305658v1>**

Preprint submitted on 24 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Assessing the Probability of Extremely Low Wind Energy Production in Europe at Sub-seasonal to Seasonal Time Scales

Bastien Cozian<sup>1</sup>, Corentin Herbert<sup>2</sup> and Freddy Bouchet<sup>2,3</sup>

1. ENS de Lyon, UCBL, CNRS, Laboratoire de physique, F-69342 Lyon, France

2. ENS de Lyon, CNRS, Laboratoire de physique, F-69342 Lyon, France.

3. LMD/IPSL, ENS, Université PSL, École Polytechnique, Institut Polytechnique de Paris, CNRS Paris, France

E-mail: [bastien.cozian@ens-lyon.fr](mailto:bastien.cozian@ens-lyon.fr)

**Abstract.** The European energy system will undergo major transformations in the coming decades to implement mitigation measures and comply with the Paris Agreement. In particular, the share of weather-dependent wind generation will increase significantly in the European energy mix. The most extreme fluctuations of the production at all time scales need to be taken into account in the design of the power system. In particular, extreme long-lasting low wind energy production events constitute a specific challenge, as most flexibility solutions do not apply at time scales beyond a few days. However, the probability and amplitude of such events has to a large extent eluded quantitative study so far due to lack of sufficiently long data. In this letter, using a 1000-year climate simulation, we study rare events of wind energy production that last from a few weeks to a few months over the January-February period, at the scale of a continent (Europe) and a country (France). The results show that the fluctuations of the capacity factor over Europe exhibit nearly Gaussian statistics at all time scales. A similar result holds over France for events longer than about two weeks and return times up to a few decades. In that case, the return time curves follow a universal curve. Furthermore, a simple Gaussian process with the same covariance structure as the data gives good estimates of the amplitude of the most extreme events. This method allows to estimate return times for rare events from shorter but more accurate data sources. We demonstrate this possibility with reanalysis data.

*Keywords:* rare events, wind energy production, energy meteorology

## 1. Introduction

Over the past decades, the share of renewable energy in the energy mix of many European countries has been rapidly increasing. This trend should only intensify in the coming years, as rapid deployment of renewable energy is an integral part of the global efforts to contain global warming within the bounds of the Paris agreement [1, 2]. As a consequence, electricity production is becoming increasingly sensitive to weather conditions [3, 4]. One of the main challenges of the energy transition is to maintain the balance between electricity production and demand in spite of this increased variability. This challenge is actually a multi-facet one, as the energy production fluctuations occur on a wide spectrum of timescales. For instance, photovoltaic production might fluctuate on a timescale on the order of minutes due to local variations of cloud coverage and surface solar irradiance [5, 6, 7], and wind energy also exhibits variability in this time range [8, 9, 10]. But these energy sources, and in particular wind energy, also fluctuate on longer timescales, due to the prevalence of different meteorological conditions on timescales of days, weeks, months, or even years. In this letter, one of our main goals is to estimate the probability, or *return time*, and amplitude of events where wind energy production remains anomalously low for periods ranging from a few days to the whole winter. In particular, we focus on longer events lasting several weeks, because they impose specific constraints on the design of the energy system. Indeed, many standard flexibility solutions (e.g. demand-side management, battery storage or even pump storage power plants) become unavailable on timescales exceeding a few days, let alone a whole season. Hence, even if they should be expected to be lower than for shorter events, quantifying precisely the needs for flexibility and backup on such long timescales is crucial to ensure the security and robustness of future energy systems, and to assess the corresponding economic cost.

Variability at the sub-seasonal to seasonal timescales has attracted growing interest over the past years. For instance, low wind speed events lasting up to 20 days have been studied over Ireland [11] and over the North Sea [12], using 30-year records from weather stations and a 10-year high-resolution simulation from a regional model, respectively. Similar studies, relying on reanalysis data, have also been conducted for

low wind power events in Germany [13, 14], and Great Britain [15]. However, a fundamental challenge remains, as estimating the characteristics of such events in a statistically robust way requires much more data than is available from observations. Indeed, extreme events play a key role due to their major socioeconomic impact, but require time series much longer than their return time. Estimating the probability and characteristics of rare, long-lasting low anomalies of energy production thus requires new methodological progress.

In this letter, we address this methodological question, and quantify the return time and amplitude of low extremes of wind energy production. We show that return times for events lasting from a few days to the whole winter can be obtained from long simulations with climate models, here 1000 years of data, using a simple energy model to relate energy production to climate variables similar to the one used by van der Wiel et al. [16, 17] to study the meteorological conditions associated with low production and high demand events over Europe. The first main novelty of this work is the computation of return levels associated with return times up to several centuries, and the discussion of their dependence on different scenarios of installed capacity and on various time and space aggregation scales. In particular, we show that for long aggregation times, the capacity factor fluctuations follow a universal curve, independent of the scenario. We further discuss the impact of the duration of the event and the effect of spatial aggregation. For instance, for 20-year events, relative fluctuations of capacity factor for short events (on the order of days) are close to 100% at the scale of France, and 80% at the scale of Europe, while they reduce to about 30% for France and 20% for Europe for two-month events. Such fluctuations still represent a wind power shortfall on the order of 10GW for France and 60GW for Europe for the whole two-month period, which might prove challenging in terms of flexibility and backup. The second main contribution of this work is to show that return times for low wind energy production can be estimated from a simple Gaussian process, whose parameters are estimated from available data. We show that this method provides a good approximation for large spatial scales (here Europe) at any time scale, or at smaller spatial scale (France) for events longer than about two weeks and return times up to a few decades. As an application, we estimate return times from reanalysis data for events which are much more

rare than would be possible using only the raw data.

## 2. Data and methods

### 2.1. Climate and wind energy model

Following the same strategy as other studies based on reanalysis [18, 19] or climate models [16], we construct time series of wind energy production from climate variables, taken from a 1000-year statistically stationary simulation with 3-hourly output using the CESM model [20]. Our setup uses atmosphere and land components only. The sea surface temperature, sea ice cover and greenhouse gases concentration are prescribed at values representative of the 2000s. The same dataset was used recently to study heat waves over Europe [21]. From the 10m-wind speed in the climate model output, we compute the wind speed at turbine height  $WS$ , for each grid point  $(x, y)$  and time  $t$ . The wind speed is then converted to a local capacity factor (ratio of the generated power to its nominal value) through a power curve  $CF(x, y, t) = PC[WS(x, y, t)]$ . The power curve we use is a simple 3-parameter model based on the features of industrial wind turbine power curves [22, 23] (Appendix A).

As we are interested in persistent extreme events, we look at the effect of time aggregation by computing the running average

$$CF_T(t) = \frac{1}{T} \int_t^{t+T} du CF(u), \quad (1)$$

where  $T$  is the integration time. We consider the time series  $CF(t)$  and  $CF_T(t)$  as realizations of a periodic stochastic process (with 1-year period), and we denote  $\langle \cdot \rangle$  the average over realizations, estimated from the empirical average over the 1000 years of data. We shall also study the statistics of the yearly lowest capacity factor event aggregated over some time period  $T$  (Appendix C).

### 2.2. Scenarios of installed capacity

The total capacity factor over any region  $\mathcal{D}$  is computed as

$$CF(t) = \int_{\mathcal{D}} dx dy \frac{IC(x, y)}{IC_{\mathcal{D}}} CF(x, y, t), \quad (2)$$

where  $IC(x, y)$  (in  $\text{W}\cdot\text{m}^{-2}$ ) and  $IC_{\mathcal{D}}$  (in  $\text{W}$ ) are the local and total installed capacity, respectively. We use the installed capacity of five scenarios developed by the e-Highway2050 project<sup>‡</sup> [24]. This project provides distributions of onshore and offshore wind installed capacity based on current and future plans as of 2015. More recent scenarios in line with the Paris Agreement

<sup>‡</sup> Documentation available at <https://docs.entsoe.eu/baltic-conf/bites/www.e-highway2050.eu/results>.

**Table 1.** Total installed capacity of 34+1 countries in Europe (i.e. onshore capacity of 34 countries and offshore capacity in the North Sea), onshore capacity in France, and offshore capacity in the North Sea for the 5 e-Highway2050 scenarios. The share of offshore capacity, exclusively located in the North Sea, is shown in parenthesis.

Scenario	Europe	France	North Sea
X5	813 GW	84 GW	104 GW (13%)
X7	875 GW	124 GW	115 GW (13%)
X10	512 GW	58 GW	76 GW (15%)
X13	304 GW	22 GW	42 GW (14%)
X16	388 GW	54 GW	15 GW (4%)

and consistent with the European Green Deal have since been published [25, 26, 27, 28]. However, the total installed capacity and spatial distribution are more contrasted in the e-Highway2050 scenarios. This allows us to explore the influence of a wider range of choices on extreme events, and to draw more robust conclusions likely to apply to other scenarios as well.

e-Highway2050 provides installed capacities at the scale of “clusters” (Fig. A2), which we aggregate over two domains to study the effect of space aggregation on extremes:  $\mathcal{D} = \text{Europe}$ , or  $\mathcal{D} = \text{France}$ . Because e-Highway2050 assumes that all offshore capacity is located in the North Sea, our Europe region is defined as the onshore capacity for 34 country plus the offshore capacity in the North sea, while our France region contains only onshore capacity. Table 1 shows the total installed capacity for each scenario in each region. The total installed capacity varies from 304 GW to 875 GW. The share of offshore capacity is around 15% in all scenarios, except for scenario X16 where it is significantly lower (4%).

## 3. Capacity factor statistics in Europe and France

### 3.1. Annual mean capacity factor

The annual mean capacity factor  $\langle CF_{T=1y} \rangle$  in our model ranges from 25.5% for the X16 scenario to 28.9% for the X13 scenario. Van der Wiel et al. [16] found significantly smaller values: using 2000 years of data representative of present-day climate, for two different climate models (EC-Earth and HadGEM2-ES), the average daily wind energy production they report (2.1 TWh/day and 1.3 TWh/day, respectively) correspond to annual mean capacity factors of 18.3% and 11.3%, respectively. This discrepancy can be explained by several factors: a smaller set of 15 European countries, and different parameters for the wind energy and climate models. Using their wind energy model parameters, the same set of 15 countries and the same scenario (X5), we find an annual mean

capacity factor of 19.3% with our climate model. The rest of the difference can be attributed to the climate data. With the same installed capacity but using ERA interim data leads to an annual mean capacity factor of 22.6% [16]. Similarly, the Renewables.Ninja dataset of capacity factor timeseries calculated with MERRA-2 reanalysis leads to an annual mean capacity factor of 24.2% with the actual 2017 installed capacity of 24 European countries [30]. This same dataset yields an annual mean capacity factor of 26.3% with the X5 scenario over the same countries.

Hence, while the annual mean capacity factor given by such energy models depends on many parameters which make precise comparison difficult, our values are broadly compatible with the existing literature. Here we will focus on fluctuations of the capacity factor rather than its absolute value.

### 3.2. Seasonal cycle of capacity factor

Wind power has a strong seasonal cycle. Figure 1 shows the statistics of the daily capacity factor  $CF_{T=1\text{ day}}$  in Europe for the five scenarios. We find that it is about 2.9 times larger in winter (40.2% in DJF, averaged over the 5 scenarios) than in summer (13.9% in JJA). This is consistent with [16] who find a winter-summer ratio of 2.5 and 3.5 for their two climate models, but larger than the value they report for ERA-Interim, which is 2.2.

We will focus here on the winter season, when extremely low wind generation events can be expected to have the largest impact. Indeed, it is the time of the year where both the energy demand and fluctuations of wind energy capacity factor are largest (figure 1). Specifically, we will consider events occurring in January and February, because this period is short enough to consider the time series of capacity factor to be approximately stationary (the mean and variance of the daily capacity factor in figure 1 are approximately constant) and still long enough to provide meaningful statistics (e.g. time correlations). Hence, all the winter-averages in the sequel correspond to JF averages.

As Fig. 1 shows, the seasonal capacity factor statistics vary from one scenario to the next. In particular, a higher share of offshore capacity decreases the variability of wind production in our model, in agreement with previous studies [31] (see Appendix A for more details).

### 3.3. Effect of space and time aggregation on capacity factor statistics

The smoothing effect on the variability of aggregating wind energy production from distant areas has been documented in a number of studies based on

observational data [32, 33, 34, 35, 36, 37, e.g.]. Here, we also observe this effect: while the PDF of winter capacity factor at the scale of France is strongly skewed, it becomes close to Gaussian at the scale of Europe, and the variance is reduced by a factor 2.

Similarly, fluctuations on different timescales, characterized by different solutions in terms of backup and flexibility, have different statistical properties. At the elementary level, one can note that aggregating wind energy production over longer and longer timescales makes it more and more Gaussian and reduces its variance, similarly to aggregation in space.

Both these properties are illustrated in Appendix B. Now, we turn to the main point of our paper, which is to discuss quantitatively the smoothing effect due to space and time aggregation on rare fluctuations of the capacity factor, rather than typical ones.

## 4. Return time curves

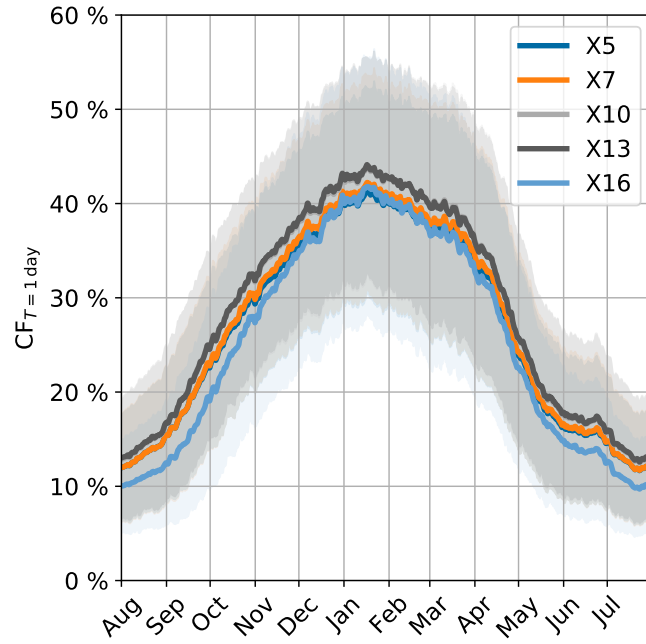
### 4.1. Return times of winter-average capacity factor

In this section, we assess the probability of low winter wind energy production events by computing return time curves [38, 39] (see Appendix C for details).

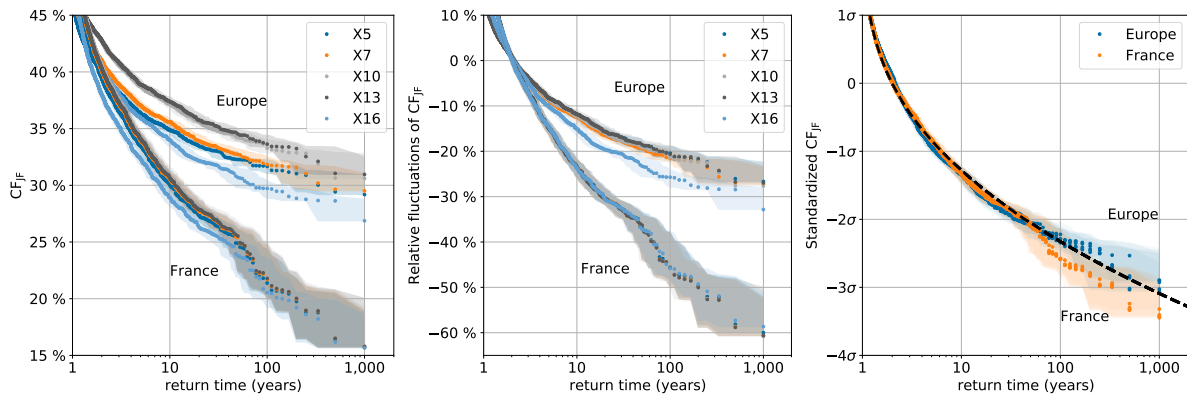
First, we consider return time curves for the capacity factor averaged over JF,  $CF_{JF}$  (figure 2, left panel), for all scenarios, for both France and Europe. Such curves can be interpreted either as the frequency of occurrence of an event of a given amplitude, or the reverse. For instance, at the scale of Europe for scenario X16, one may say that a JF capacity factor of 30% has a return time of 100 years (i.e. a 1 in 100 chance of occurring in any given year), or equivalently that the return level at 100 years is 30%. All return time curves have a convex shape. This means that a simple extrapolation (i.e. a linear fit in figure 2) of return levels for small return times would over-estimate the amplitude of the rarest events.

The return time curves shown in figure 2 depend on the geographical area (France or Europe) and scenario considered (less so at the scale of France, where the distribution of installed capacity varies less from one scenario to the other). For instance, the European JF capacity factor drops below 34% on average once in a decade for scenario X16, but only once in a century for X10 and X13. However, we show below that these dependencies only come from basic statistical properties (mean and variance of capacity factor) and that events with return times between 10 and 100 years obey universal properties.

First, we show in figure 2 (middle panel) return time curves for relative fluctuations of capacity factor  $CF'_{JF} = (CF_{JF} - \langle CF_{JF} \rangle) / \langle CF_{JF} \rangle$ , where  $\langle CF_{JF} \rangle$  is the climatology. Unlike absolute capacity factors, return times for fluctuations of capacity factor around the



**Figure 1.** Daily capacity factor  $CF_{T=1 \text{ day}}$  in Europe: average (solid lines) and standard deviation (shaded area) over 1000 sample years, for 5 e-Highway2050 scenarios of installed capacity. The curves for scenarios X10 and X13, and to a lesser extent X5 and X7, are indistinguishable from each other.



**Figure 2.** Return time curves for low average JF capacity factor  $CF_{JF}$  (left panel), relative fluctuations of the JF capacity factor  $CF'_{JF} = (CF_{JF} - \langle CF_{JF} \rangle) / \langle CF_{JF} \rangle$  (middle panel), and standardized JF capacity factor (right panel), see main text for details. In the first two panels, the 5 scenarios are shown in different colors. In the last panel, all scenarios for each region are shown in the same color. The return time curve for a standard normal distribution is shown as a black dashed line. 95% confidence intervals (shaded) are computed with a bootstrap method (Appendix C).

mean depend little on the scenario. Note, however, that a gap remains for large return times between scenario X16 and the others. This is due to the larger variance of the capacity factor in this scenario, presumably due to its small amount of offshore wind energy, as shown in the return time curves for standardized capacity factor (i.e.  $CF'_{JF}/\sigma$  with  $\sigma = \sqrt{\langle (CF'_{JF})^2 \rangle}$ ), where all scenarios collapse (Fig. 2, right panel). In addition, the return time curves of standardized capacity factor for France and Europe also collapse. Without this normalization (figure 2,

left and middle panels), the convexity of the curve is stronger for Europe than for France, which shows that although the return level of typical events (e.g. for 2-year return time, corresponding to the median of the distribution) is similar for both cases, space aggregation decreases significantly the amplitude of rare events. For instance, on average, the capacity factor in France drops by about 46% once a century, but only by about 20% at the European scale. This mitigation effect of spatial aggregation on rare events is in fact only due to the reduced variance (Appendix B), as figure 2 shows (right panel). The return time

curves for standardized capacity factor match well the standard normal distribution (black dashed line), at least below 100-year return times. We can therefore conclude that the statistics of extreme low JF capacity factor events with return times below 100 years can be assumed to be Gaussian, and all the information about scenarios and spatial scale are encoded in the average and variance. For rarer events, with return times above 100 years, the uncertainty is larger, and although it seems that events are more extreme than Gaussian at the scale of France and less extreme at the scale of Europe, the available data does not allow to draw definitive conclusions.

#### 4.2. Effect of time aggregation on return time curves

Above we considered extreme events for the JF capacity factor, corresponding to an integration time  $T = 59$  days (C.1). We now study return times of capacity factor fluctuations for shorter events, down to  $T = 1$  day. To do so, we can either fix the return time and plot the return level as a function of the aggregation time  $T$  (figure 3, left panel; we chose a 20-year return time), or conversely fix the return level and plot the return time as a function of the aggregation time  $T$  (figure 3, right panel, -20% return level).

In both panels we note again that the different scenarios lead to very similar statistics, except X16 which yields slightly more extreme events at the scale of Europe, and that spatial aggregation leads to smaller amplitude or less frequent events for Europe compared to France. We observe a smoothing effect of time aggregation: the amplitude of capacity factor fluctuations, for fixed return time (figure 3, left panel) decreases with increasing  $T$ . At the European scale, the curve is markedly concave, which means that flexibility measures allowing to smooth energy generation over several days or a few weeks are very efficient (for instance the amplitude of fluctuations is reduced by half when integrating the capacity factor over 20 days), while, by contrast, over longer time scales the amplitude of fluctuations depends less on the aggregation time (integrating from 40 days to 59 days only reduces the fluctuations by a few percents). A similar behavior has been reported for 10-year wind speed events as a function of their duration [11]. At the scale of France, these two regimes are less pronounced and the reduction of the amplitude of the fluctuations with the aggregation time is more linear. This stronger concavity at the European scale means that spatial aggregation is more efficient at reducing the amplitude of rare  $CF_T$  fluctuations with increasing  $T$ : it is 20% smaller at the European scale than at the French scale for daily events, while it is about 50% smaller for 1-month events.

Equivalently, the smoothing effect is seen on the

return time curve of capacity factor fluctuations with fixed amplitude (figure 3, right panel), which increases approximately exponentially with  $T$ .

## 5. Estimating return times with Gaussian processes

The analysis carried out in section 4 suggests that extreme capacity factor events behave essentially in a Gaussian manner for long enough aggregation time. It should be expected, however, that it ceases to be the case for events on shorter time and/or spatial scales. To investigate the limits of validity of this approximation, we compare the statistics based on capacity factor time series with a simple Gaussian stochastic process, with the same covariance structure.

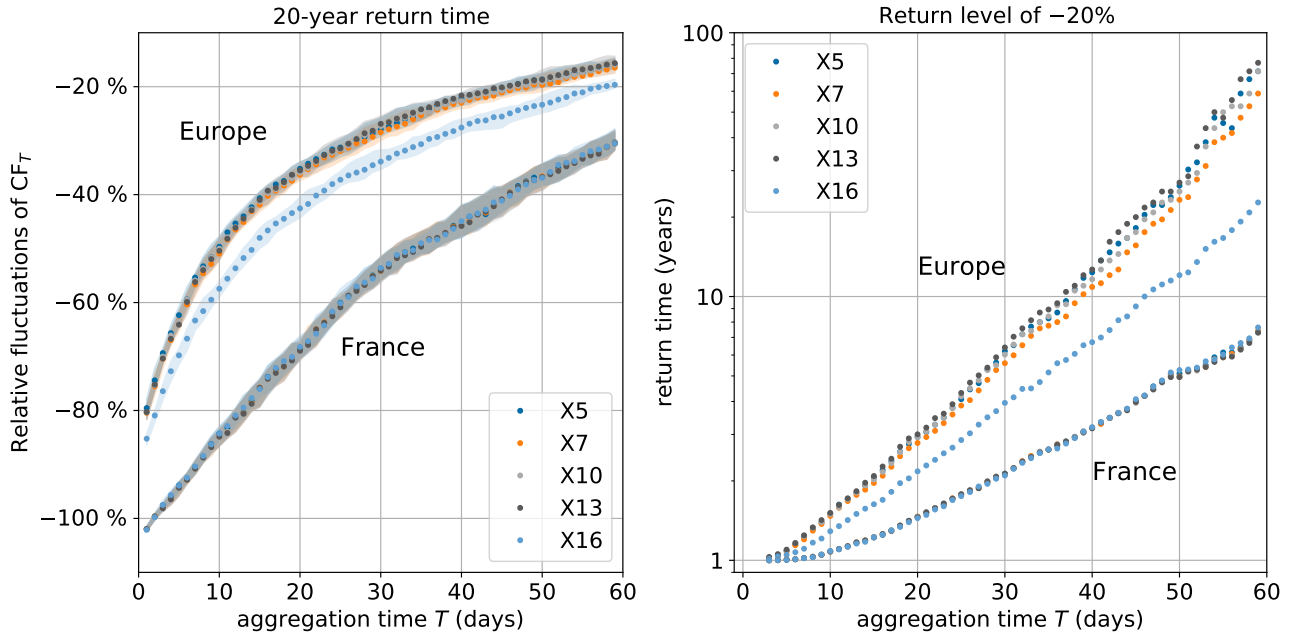
### 5.1. The Gaussian process

First we compute the autocorrelation function of the capacity factor fluctuations time series  $CF'_{T=1\text{ day}}(t)$ , daily-averaged to remove the diurnal cycle (figure 4). Because of the seasonality of wind energy production, the autocorrelation function should depend on the calendar time  $t_{\text{cal}}$  as well as the time lag  $\tau$ :  $R(t_{\text{cal}}, \tau) = \langle CF'(t_{\text{cal}})CF'(t_{\text{cal}} + \tau) \rangle / \langle CF'(t_{\text{cal}})^2 \rangle$ . As the wind energy production statistics vary slowly over the January-February period (section 3.2), we assume that the dependence on the calendar date can be neglected over these two months:  $R(t_{\text{cal}}, \tau) \approx R(0, \tau)$ . We observe that the autocorrelation function (figure 4) decays approximately as the sum of two exponentials with characteristic timescales  $\tau_1 \approx 2$  days and  $\tau_2 \approx 15$  days for Europe,  $\tau_1 \approx 1.5$  days and  $\tau_2 \approx 6$  days for France. Similar results are obtained for all scenarios.

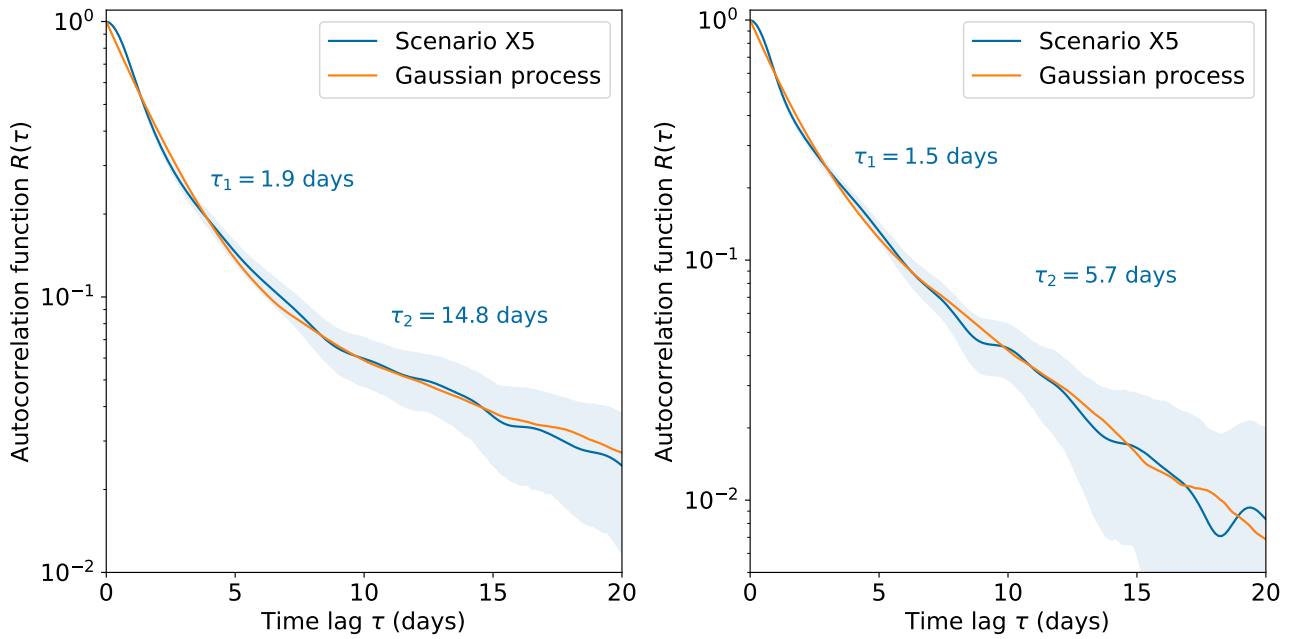
We therefore construct a simple model for capacity factor fluctuations, as the sum  $x_G(t) = x_1(t) + x_2(t)$  of two *Ornstein-Uhlenbeck processes* [40, e.g.] corresponding to these two time scales:

$$dx_i(t) = -\frac{1}{\tau_i}x_i(t)dt + \sqrt{2D_i}dW_i(t), \quad i \in \{1, 2\}, \quad (3)$$

where  $W_1$  and  $W_2$  are two independent *Wiener processes*. This process is Gaussian as the sum of two Gaussian processes. Its autocorrelation function is a sum of two exponentials which matches well the one obtained from the climate data (figure 4). The noise amplitudes  $\{D_i\}_i$  are chosen such that the variance  $\tau_1 D_1 + \tau_2 D_2$  is equal to the variance of the daily capacity factor, and such that their relative weight  $\frac{\tau_2 D_2}{\tau_1 D_1}$  matches the relative weight fitted on the autocorrelation function. A  $59 \times 10^4$ -day long realization of this stochastic process is computed with the `StochRare` Python module [41], using a Euler-Maruyama method.

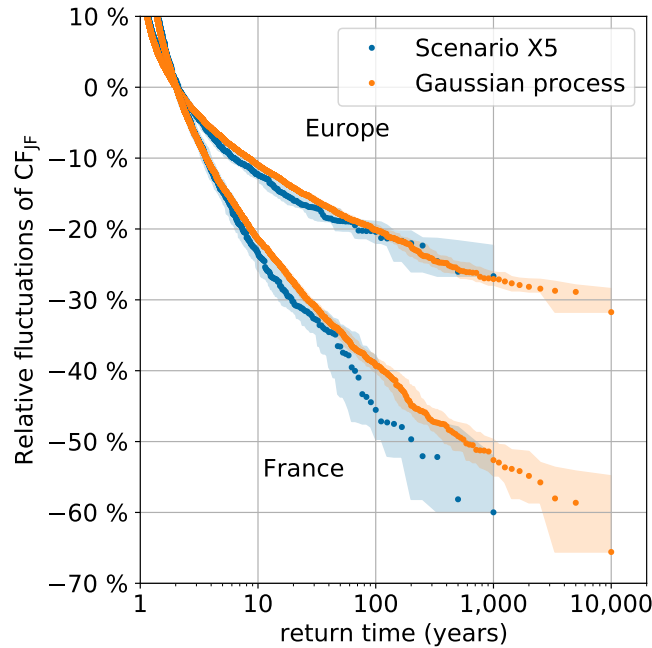


**Figure 3.** Return level  $CF'_T$  of a 1-in-20-year low fluctuation (left panel) and return time of a  $CF'_T = -20\%$  event (right panel), as a function of the aggregation time  $T$ , for 5 scenarios and for both Europe and France. For the left panel, the confidence interval is the 95% confidence interval of the return time curve for a 20 year return time, which is computed as in figure 2 for every value of  $T$ .



**Figure 4.** Autocorrelation functions of daily capacity factor fluctuations  $CF'_{T=1 \text{ day}}(t)$  for Europe (left) and France (right) for scenario X5 (blue) and the Gaussian process (orange). Note the logarithmic scale in the y-axis. The 95% confidence interval (shaded area) is computed with a bootstrap method. The time scales  $\tau_1$  and  $\tau_2$  are fitted with a mean square error method.





**Figure 5.** Return time curve for low average JF capacity factor fluctuations  $CF'_{JF}$  for scenario X5 (blue), and for a Gaussian process with the same covariance structure (orange), for Europe and for France. The 95% confidence intervals (shaded areas) are computed with a bootstrap method.

### 5.2. Return time estimates

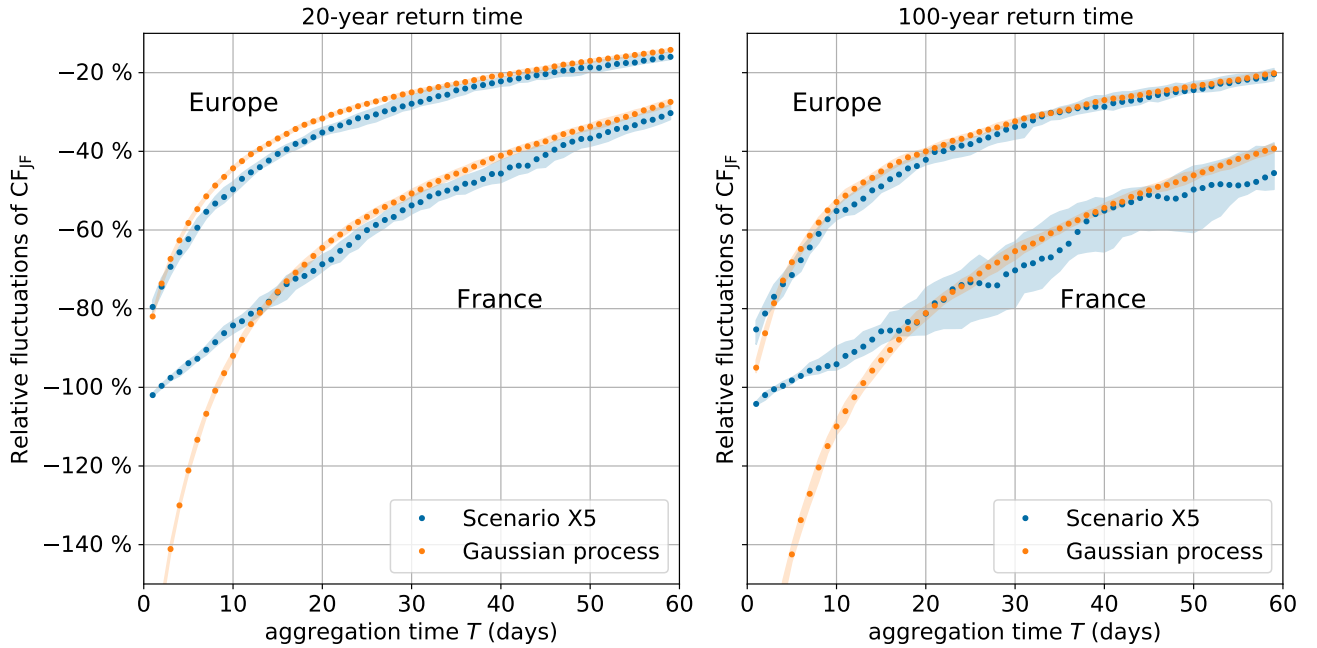
Figure 5 compares return time curves for JF capacity factor fluctuations  $CF'_{JF}$  based on climate data and for the Gaussian process, for Europe and for France. It shows that the approximation performs well for small return times, below 5 years. On the other hand, for intermediate return times between 6 and 50 years the Gaussian process lies outside the confidence interval estimated from climate data: the fluctuations are systematically more extreme than the Gaussian process at fixed return time, or more frequent at fixed return level. For instance, a relative fluctuation of  $-16\%$  for Europe ( $-30\%$  for France) has a return time of approximately 20 years for climate data and 30 years for the Gaussian process. Conversely, for Europe, a 1-in-20-year event corresponds to a capacity factor fluctuation of  $-16\%$  for climate data and  $-14\%$  for the Gaussian process. For the most extreme events, with return times above 100 years, the Gaussian process is a relatively good approximation at the continental scale but not at the scale of France, where it underestimates the severity or frequency of the events even more than for intermediate return times. Note however that the uncertainty is larger in this regime.

In figure 6, we compare again the Gaussian process with climate data, but this time, varying the aggregation time, for given return times of 20 years (left panel) and 100 years (right panel). The above conclusions still hold: overall, the Gaussian process is a better approximation at the scale of Europe than at

the scale of France. In particular, it fails completely for short events (below 15 days) over France. In the case of Europe, the approximation performs slightly better for rarer events. In all cases, except 100-year events over France (for which the error bars are much larger), the discrepancy is larger for events lasting between two weeks and a month, and the amplitude of the event is systematically larger than the Gaussian process estimate. However, relative errors remain relatively small. For instance, for a two-week event, the 20-year return level is  $-42\%$  for climate data and  $-38\%$  for the Gaussian process.

### 5.3. Application to reanalysis data

We have shown above that a simple model based on a Gaussian process provides a good approximation of return time curves at the scale of Europe. One of the main practical interests is that it allows to compute return times for events which are not observed in the original dataset. Since most datasets are quite short (at most a few decades), this is a very interesting feature to study rare events. Here we apply this method to the Renewables.Ninja dataset [30], which provides hourly wind capacity factors for European countries, computed with the MERRA-2 reanalysis product over the period 1980-2016. It addresses some of the limitations of our climate data, such as an increased spatial resolution and the use of bias-correction with respect to available data from national



**Figure 6.** Same as figure 5, but looking at the return level of 1-in-20-year (left panel) and 1-in-100-year events (right panel) for different integration time  $T$ . The two figures share the same y-axis.

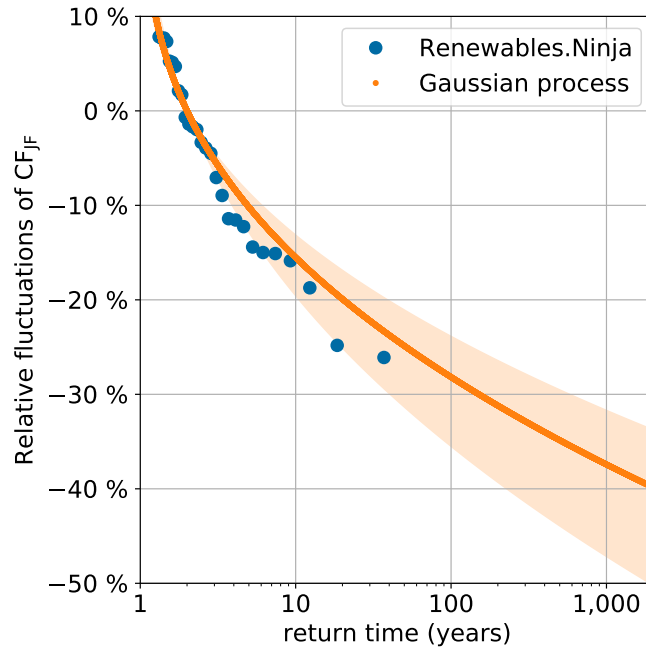
Transmission System Operators. We use it here as a proof of concept for the method. Figure 7 shows the return time curve for JF capacity factor fluctuations over Europe estimated directly from the dataset, and its extension to arbitrarily rare events using a Gaussian process. Because the dataset is short, there is some uncertainty in the estimate of the second correlation time  $\tau_2$ , which propagates to the return time estimates with the Gaussian process (Appendix D). Again, the Gaussian process estimate matches well the direct return time curve estimate when data is available. While the installed capacity in this dataset (as of 2017 for version v1.1) is different from the scenarios considered above, the return time curves are very similar, in agreement with our finding in section 4. This approach allows us to extrapolate return times for rare events: for instance, for a 100-year event, we expect relative fluctuations of  $-28\%$ . For this scenario where the total installed capacity is  $IC_{\mathcal{D}} = 110$  GW and the average JF capacity factor is  $\langle CF_{JF} \rangle = 31\%$ , a 100-year return time event represents an average loss of 9.5 GW over 2 months, or a loss of 13.5 TWh.

## 6. Conclusion

The question we have addressed here is the estimation of the probability and amplitude of events where wind energy production remains low for periods ranging from a few days to the entire January-February period. Using a very long climate simulation, we have computed return time curves for low capacity factor

events over Europe or France using different scenarios of installed capacity up to 1000-year return times. A first result is that the return times of extreme winter capacity factor depend on the geographic area and installed capacity scenario only through the average and standard deviation of the capacity factor; after standardization they follow a universal curve. This simplifies the discussion of projections of extreme events of wind energy production, as the properties of the various scenarios can be obtained by simple rescaling. We further show that the return time curves can be estimated using a simple Gaussian process with the same covariance as the capacity factor time series generated from the climate model. This approximation performs well at the scale of Europe, although events tend to be slightly more extreme than Gaussian, especially at intermediate return times (on the order of 10 years) and durations (from a week to a month). The same remark applies in a magnified way at smaller spatial scale (France), where the quality of the approximation degrades significantly for events with return times above 100 years or durations below two weeks. Nevertheless, for long events at the scale of Europe, this provides a method to estimate return times even when observational or model data is scarce. We illustrate this by constructing a return time curve based on a capacity factor dataset constructed from more precise data, using reanalyses, but much shorter than our climate data. This return time curve can be extended to arbitrarily rare events.

Finally, we have discussed the effect of space



**Figure 7.** Return time curve for low JF capacity factor fluctuations  $CF'_{JF}$  for the Renewables.Ninja dataset estimated directly (blue) and using a Gaussian process with the same covariance structure (orange). The 95% confidence interval is computed from the uncertainty in the estimate of the correlation time  $\tau_2$ , due to the short length of the data (see Appendix D).

and time aggregation for rare events. Our results show that spatial aggregation significantly reduces the amplitude of extreme events: for daily events, the capacity factor fluctuations in Europe is about 20% lower than in France, and for 1-month events it is about 50% lower, at 20-year return times. Similarly, we have shown that longer events are much less extreme than short ones: less than 20% loss compared to the average capacity factor for two-month events on a European scale, 40% for two weeks and 80% for one day. At the scale of France, these figures become 30% for two-month events, 70% for two weeks and 100% for one day. Nevertheless, in absolute terms the wind energy shortfall for two-month events is on the order of 20-59 GW for Europe and 3-15 GW for France, depending on the scenario installed capacity, for 20-year return times, and about 20% more for 100-year return times. Our results, obtained with a simple energy model, therefore suggest that such long events might put important constraints on future energy systems, given the fact that flexibility means become scarcer on such long timescales. This provides an incentive for studying rare, long-lasting events in more comprehensive energy models.

### Acknowledgments

This work was granted access to the HPC/AI resources of CINES under the allocation 2018-A0050110575, 2019-A0070110575 and 2020-A0090110575 made by

GENCI. Data analysis was carried out using the resources of the *Centre Blaise Pascal* at ENS de Lyon. We are grateful to Emmanuel Quemener for his help with the platform. This research has been partially funded by the *Institut des Mathématiques pour la Planète Terre (IMPT)*. We would like to thank Laurent Dubus for his very valuable comments on a preliminary version of this manuscript.

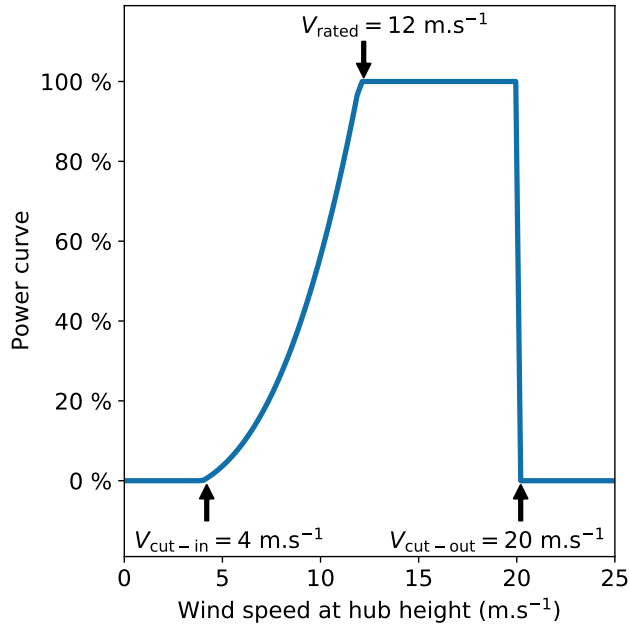
### Appendix A. Additional information on the wind energy model

The  $0.9^\circ \times 1.2^\circ$  grid of CESM is remapped with a bilinear interpolation method to the  $0.5^\circ \times 0.5^\circ$  grid§ used by the ECEM project [23]. This allows a better resolution of e-Highway2050 “cluster” regions (section 2.2). Similarly to studies using reanalysis data [15, 14], no bias correction is applied to the wind speed, in order to preserve the self-consistency of the climate model output. Furthermore, such bias correction methods have been validated for average wind speeds but not for extreme values, and we can expect some degree of error smoothing due to the large scales considered here.

To compute the wind speed at turbine height WS, we choose a hub height of  $h_{\text{hub}} = 100$  m for both onshore and offshore turbines, and assume a power law profile with exponent  $\alpha = \frac{1}{7} \approx 0.143$  [42, 43]:

§ Available on the Copernicus Climate Data Store <https://cds.climate.copernicus.eu> [accessed in October 2020]

$WS = WS_{10m} \left( \frac{h_{hub}}{h_{ref}} \right)^\alpha$ , where  $WS_{10m}$  is the wind speed at the reference height  $h_{ref} = 10$  m.



**Figure A1.** Power curve used in the energy model.

To relate the generated power to the wind speed, we use a simple power curve (figure A1) with a cubic increase of the generated power after some cut-in value, until the rated power is reached. The turbine stops after a cut-off wind speed for security reasons. The equation for the power curve is:

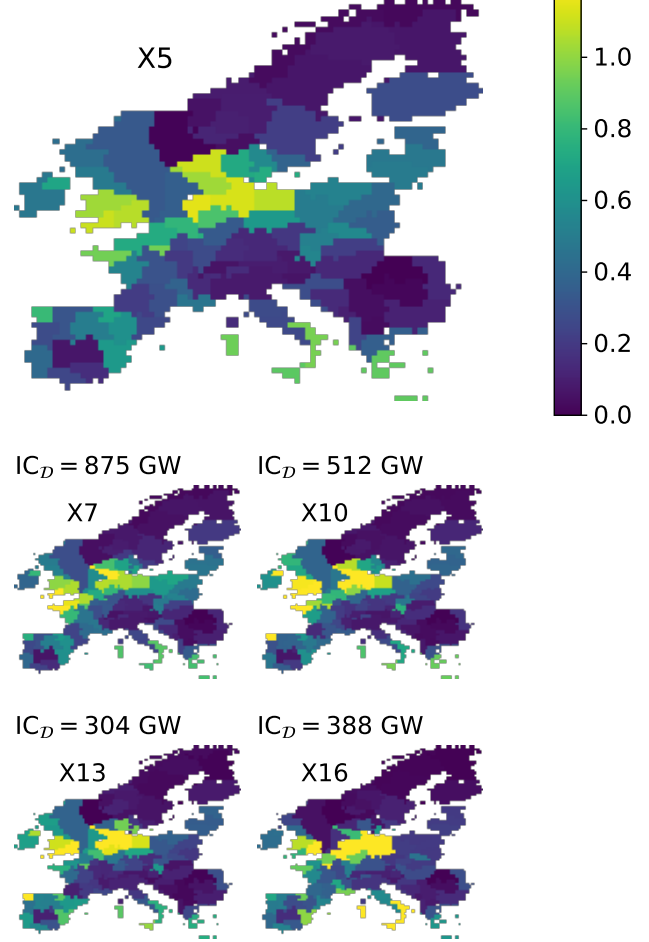
$$PC = \begin{cases} 0 & \text{if } WS < V_{ci} \\ \frac{WS^3 - V_{ci}^3}{V_r^3 - V_{ci}^3} & \text{if } V_{ci} \leq WS < V_r \\ 1 & \text{if } V_r \leq WS < V_{co} \\ 0 & \text{if } V_{co} \leq WS \end{cases} \quad (\text{A.1})$$

The parameters ( $V_{ci} = 4$  m.s<sup>-1</sup>,  $V_r = 12$  m.s<sup>-1</sup>,  $V_{co} = 20$  m.s<sup>-1</sup>) are chosen to correspond to a VESTAS V110-2.0MW turbine<sup>||</sup> as done in the European Climatic Energy Mixes (ECEM) project [23].

The total capacity factor  $CF(t)$  over a region  $\mathcal{D}$  in equation (2) depends on the scenario only through its spatial distribution  $\frac{dx dy IC(x,y)}{IC_{\mathcal{D}}}$ . Installed capacities in the e-Highway2050 project are defined at the scale of about a hundred “clusters” built from European NUTS-3 regions (Fig. A2). This allows for a greater spatial granularity than scenarios defined at the scale of countries such as ENTSO-E’s Ten-Year Network Development Plan scenarios [26] or European Commission’s reference scenarios [25]. Figure A2 shows the spatial distribution of installed capacity at this cluster scale for all 5 e-Highway2050 scenarios. Contrarily to onshore clusters, e-Highway2050 does

<sup>||</sup> From [https://thewindpower.net/turbine\\_en\\_590\\_vestas\\_v110-2000.php](https://thewindpower.net/turbine_en_590_vestas_v110-2000.php) [accessed in October 2020]

$IC_{\mathcal{D}} = 813$  GW



**Figure A2.** Spatial distribution of installed capacity  $\frac{dx dy IC(x,y)}{IC_{\mathcal{D}}}$  (in %) for all 5 e-Highway2050 scenarios. The total installed capacity  $IC_{\mathcal{D}}$  is shown at the top left of each panel.

not provide a geographical description of equivalent “cluster” regions in the North Sea. Therefore, we use a dataset constructed for offshore energy [29] to manually assign maritime regions to offshore wind capacities.

In the main text we focus on properties which are common to all scenarios, but the distribution of installed capacity has a clear effect both on the annual (as reported in section 3.1, it varies from 25.5% to 28.9%) and seasonal mean capacity factor (Fig. 1). Indeed, the mean winter capacity factor varies from one scenario to the other between 39% and 41.7%. This difference is much smaller than the winter-average of the daily standard deviation (from 10.9% for X5 to 13.6% for X16, see figure 1), but of the same order of magnitude as the inter-annual variability, computed as the standard deviation of the mean winter capacity factor (from 2.9% for X5 to 3.6% for X16, not shown). Small differences in the mean winter capacity factor leads to significant differences in term of total

energy production when they are integrated over the whole winter: for instance a 37% mean capacity factor corresponds to a total winter energy production 7.5% lower than a 40% mean capacity factor.

Some scenarios have specific characteristics: X5 and X7 have a higher capacity factor in summer than X16, but a similar capacity factor in winter, while X10 and X13 always have a capacity factor larger than all other scenarios. In scenario X16, where the share of offshore is by far the lowest, the daily standard deviation is at least 2 percentage points higher than in all other scenarios, and the inter-annual standard deviation is at least 0.6 percentage points higher than in all other scenarios.

### Appendix B. Illustration of the effect of aggregation on capacity factor statistics

In this section, computations are made for scenario X5, but the conclusions also hold for the other scenarios.

On figure B1, we compare two time series from our simulations of the 3-hourly capacity factor in winter and its probability density function (PDF) aggregated at the national scale (taking the example of France) and at the continental scale (Europe). At the scale of France, the capacity factor undergoes strong fluctuations, with daily ramp-up and ramp-down events of about 80%. The PDF of the capacity factor is strongly skewed: low-production events are much more frequent than high-production events. A long lasting low production event is observed from late-January to early-February in the sample time series shown in figure B1. Hence, at this spatial scale the smoothing effect of space aggregation appears to be limited by the correlations of the wind field. On the other hand, at the European scale, the PDF of winter capacity factor is nearly Gaussian. While the average capacity factor is almost the same in both cases, the standard deviation (13%) is about half the one at the scale of France (27%), illustrating the strong smoothing effect of space aggregation at the continental scale. The persistent low production event in France is no longer visible at the European scale.

To characterize the impact of the timescale on the statistics of the fluctuations, we show in figure B2 the PDF of the moving average of the capacity factor  $CF_T$  (1) over France and Europe, for three integration times:  $T = 1, 7$  and 30 days. At the scale of France, the main effect of time aggregation is that it makes the PDF of the time-averaged capacity factor more and more Gaussian as the length of the averaging window increases. This effect is less spectacular at the European scale, as the capacity factor statistics are already closer to Gaussian at small integration times, due to the effect of spatial aggregation. In both cases,

the variance decreases significantly: it reduces by a factor of 2.5 for France and 2.3 for Europe when the integration time goes from  $T = 1$  day to  $T = 30$  days.

### Appendix C. Return time estimates

We compute return times for seasonal minima of capacity factor aggregated over a time interval  $T$ . More precisely, for a stochastic process  $\{X(t)\}$  in a time interval  $[T_1, T_2]$ , we define the random variable

$$a_T = \min_{t \in [T_1, T_2 - T]} \left\{ \frac{1}{T} \int_t^{t+T} du X(u) \right\}, \quad (C.1)$$

where  $T$  is the integration (or aggregation) time.

Here,  $X$  is a periodic stochastic process, with period  $T_p = 1$  year,  $T_1$  is the first time step of January and  $T_2$  is the last time step of February. We assume that the correlation time of the process is much smaller than the 1-year periodicity of  $\{X(t)\}$ , such that the samples of the random variable  $a_T$  obtained from all the winters in the climate simulation can be considered statistically independent. We thus have  $N = 1000$  independent samples, i.e. one event per climate year. This independence assumption is reasonable with the climate model used here, but the approximation would degrade if long-term memory effects, for instance due to the ocean, were included in the model.

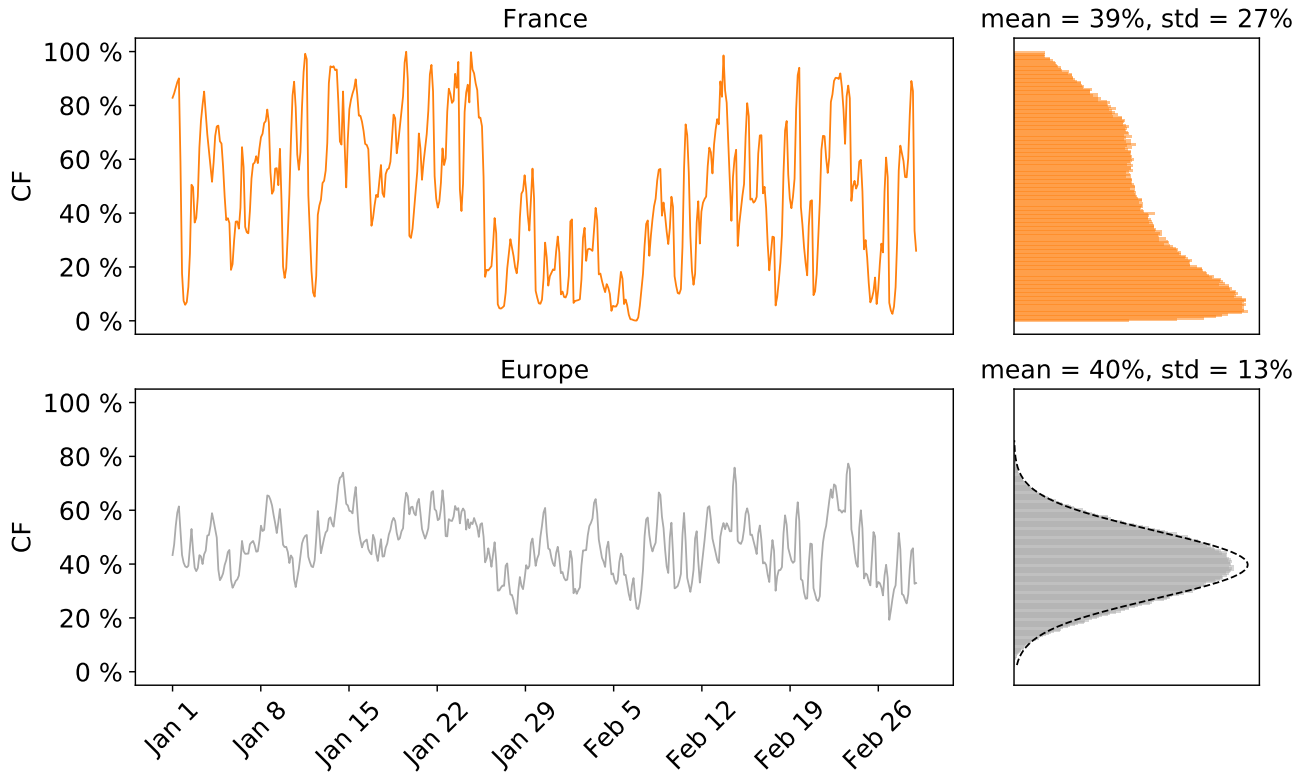
In section 4, the stochastic process  $\{X(t)\}$  alternatively stands for the capacity factor  $CF_T$  itself, its relative fluctuations  $CF'_T = (CF_T - \langle CF_T \rangle) / \langle CF_T \rangle$ , and the standardized fluctuations  $CF'_T / \sigma$  with  $\sigma = \sqrt{\langle CF'^2_T \rangle}$  the standard deviation.

We define the return time  $r(a)$  of the return level  $a$  as the average time to wait before seeing a year with  $a_T \leq a$ . It can be easily estimated as follow. Let us define the random variable  $Y(a)$  by  $Y(a) = 1$  if  $a_T < a$  and  $Y(a) = 0$  otherwise. By construction,  $Y(a)$  has a binomial distribution with parameter  $F(a) = \mathbb{E}[Y(a)]$ , the cumulative distribution function of  $a$ :

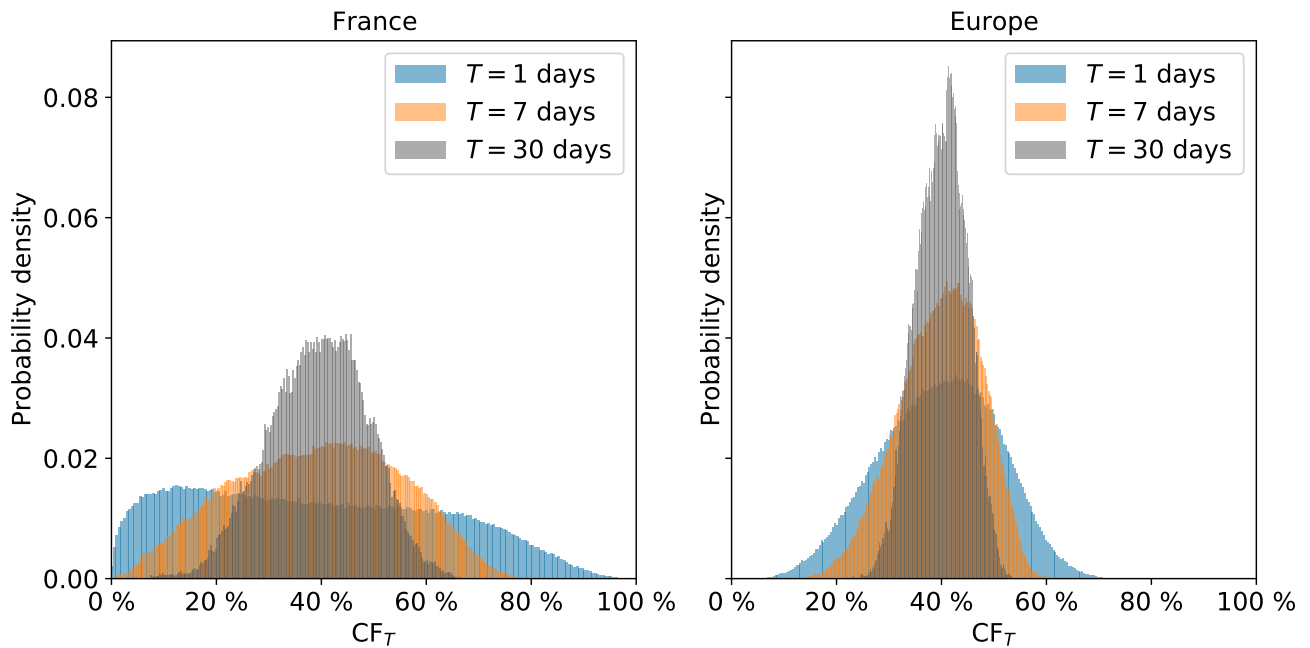
$$F(a) = \mathbb{P}[a_T \leq a] = \int_{-\infty}^a \mathbb{P}[a_T = a'] da', \quad (C.2)$$

which can be estimated with  $\hat{F}(a) = \hat{n}(a)/N$ , where  $\hat{n}(a)$  is the number of samples of  $a_T$  lower than the threshold  $a$ .

We note  $\tau(a)$  the random variable which is the number of years to wait before seeing an event with  $a_T < a$ . The probability of  $\tau(a) = nT_p$  is related to the negative binomial distribution for one failure. Indeed, the probability  $\mathbb{P}[\tau(a) = T_p]$  that the event  $a_T < a$  occurs within one year is  $F(a) = \mathbb{P}[a_T \leq a]$ , and the probability to see this event at year  $n$  is  $\mathbb{P}[\tau(a) = nT_p] = (1 - F(a))^{n-1} F(a)$ . The return time



**Figure B1.** Time series of the 3-hourly winter capacity factor for year 989 of the climate simulation (left), and probability density function (PDF) of the 3-hourly winter capacity factor computed on the 1000 years of data (right), at the scale of France (top row) and Europe (bottom row), for scenario X5. The mean and the standard deviation of the winter capacity factor are indicated above each PDF. A Gaussian fit of the PDF is shown in black dashed line.



**Figure B2.** Probability density function of the moving average of the capacity factor  $CF_T(t)$  in winter, for  $T = 1$  day,  $T = 7$  days and  $T = 30$  days, at the scale of France (left) and Europe (right), in scenario X5.

is thus  $r(a) = \mathbb{E}[\tau(a)] = \frac{T_p}{F(a)}$ . A direct estimator is therefore

$$\hat{r}(a) = \frac{NT_p}{\hat{n}(a)}. \quad (\text{C.3})$$

This choice of extreme events gives a natural definition of the return time. For example, an event  $a_{2y}$  with a return time of two years ( $r(a_{2y}) = 2T_p$ ) is defined such that  $F(a_{2y}) = \frac{1}{2}$ , which means  $a_{2y}$  is the median of the distribution: one half of the events are above  $a_{2y}$  and the other half is below  $a_{2y}$ . Return times are defined in a similar way by [11], who computed them for wind speed at several sites in Ireland using short observational records.

In practice, to compute a return time curve, we first sort the  $N$  samples in increasing order  $\{a_T^k\}_{k \in \{1, \dots, N\}}$  such that the number of times  $a_T$  is lower than  $a_T^k$  is its rank in the sorted sample:  $\hat{n}(a_T^k) = k$ . Then, we simply plot  $\{a_T^k\}_{k \in \{1, \dots, N\}}$  on the  $y$ -axis and  $\left\{\frac{NT_p}{k}\right\}_{k \in \{1, \dots, N\}}$  on the  $x$ -axis, as done in [39].

We calculate the 95% confidence interval by applying a non-parametric bootstrap method [44, 45].

#### Appendix D. Estimating return times with short datasets

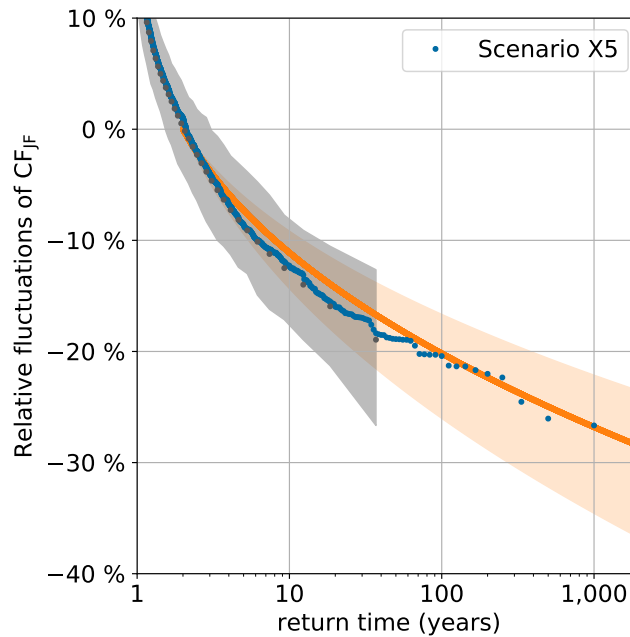
Using a single time scale  $\tau_1$  gives a poor approximation of the autocorrelation function and thus a poor estimate of the return time curve (not shown). Therefore it is necessary to correctly estimate the second time scale  $\tau_2$  if one wants to make prediction on the return time curve of extreme low JF capacity factor fluctuations  $CF'_{JF}$ .

There are two types of errors associated with a limited amount of data: the error on the return time curve directly computed with the data available (the “direct error”), and the error on the fit of  $\tau_2$  which propagates when making an estimation of the return time curve with a Gaussian process (the “ $\tau_2$ -error”).

Both errors are computed with a bootstrap method by drawing randomly a large number (typically  $10^3 - 10^4$ ) of 37-year-long bootstrap samples. Each bootstrap sample is used to estimate directly a return time curve (for the direct error), or to estimate  $\tau_2$  and compute the return time curve of the associated Gaussian process (for the  $\tau_2$ -error). In figure D1, the 95% confidence interval is shown in grey for the direct error and orange for the  $\tau_2$ -error. The method based on a Gaussian process yields smaller errors than a direct estimate.

In section 5.3, we use the  $\tau_2$ -error found in our climate simulation to estimate the 95% confidence interval for the 37-years-long Renewables.Ninja dataset.

- [1] Rogelj J, Luderer G, Pietzcker R C, Kriegler E, Schaeffer M, Krey V and Riahi K 2015 *Nature Climate Change* **5** 519–527
- [2] Van Vuuren D P, Stehfest E, Gernaat D E, Van Den Berg M, Bijl D L, De Boer H S, Daioglou V, Doelman J C, Edelenbosch O Y, Harmsen M *et al.* 2018 *Nature climate change* **8** 391–397
- [3] Bloomfield H C, Brayshaw D J, Shaffrey L C, Coker P J and Thornton H 2016 *Environmental Research Letters* **11** 124025
- [4] Bloomfield H, Brayshaw D J, Shaffrey L, Coker P J and Thornton H E 2018 *Environmental Research Letters* **13** 054028
- [5] Tomson T and Tamm G 2006 *Solar Energy* **80** 600–606
- [6] Lave M, Kleissl J and Arias-Castro E 2012 *Solar Energy* **86** 2190–2199
- [7] Watanabe T, Takamatsu T and Nakajima T Y 2016 *Journal of Applied Meteorology and Climatology* **55** 2165–2180
- [8] Holttinen H 2005 *Wind Energy* **8** 173–195
- [9] Apt J 2007 *Journal of Power Sources* **169** 369–374
- [10] Anvari M, Lohmann G, Wächter M, Milan P, Lorenz E, Heinemann D, Tabar M R R and Peinke J 2016 *New Journal of Physics* **18** 063027
- [11] Leahy P G and McKeogh E J 2013 *Wind Energy* **16** 575–586
- [12] Patlakas P, Galanis G, Diamantis D and Kallos G 2017 *Wind Energy* **20** 1033–1047
- [13] Jung C and Schindler D 2018 *Applied Energy* **230** 845–854
- [14] Ohlendorf N and Schill W P 2020 *Environmental Research Letters* **15** 084045
- [15] Cannon D J, Brayshaw D J, Methven J, Coker P J and Lenaghan D 2015 *Renewable Energy* **75** 767–778
- [16] van der Wiel K, Stoop L P, van Zuijlen B R H, Blackport R, van den Broek M A and Selten F M 2019 *Renewable and Sustainable Energy Reviews* **111** 261–275 ISSN 1364-0321 URL <https://www.sciencedirect.com/science/article/pii/S1364032119302862>
- [17] van der Wiel K, Bloomfield H C, Lee R W, Stoop L P, Blackport R, Screen J A and Selten F M 2019 *Environmental Research Letters* **14** 094010
- [18] MacLeod D, Torralba V, Davis M and Doblas-Reyes F 2018 *Meteorological Applications* **25** 1–10 ISSN 13504827 URL <https://onlinelibrary.wiley.com/doi/10.1002/met.1660>
- [19] Bloomfield H C, Brayshaw D J, Gonzalez P L M and Charlton-Perez A 2021 *Earth System Science Data* **13** 2259–2274 ISSN 1866-3516 URL <https://essd.copernicus.org/articles/13/2259/2021/>
- [20] Hurrell J W, Holland M M, Gent P R, Ghan S, Kay J E, Kushner P J, Lamarque J F, Large W G, Lawrence D, Lindsay K, Lipscomb W H, Long M C, Mahowald N, Marsh D R, Neale R B, Rasch P, Vavrus S, Versteinsten M, Bader D, Collins W D, Hack J J, Kiehl J and Marshall S 2013 *Bulletin of the American Meteorological Society* **94** 1339–1360 publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society URL <https://journals.ametsoc.org/view/journals/bams/94/9/bams-d-12-00121.1.xml>
- [21] Ragone F and Bouchet F 2021 *Geophysical Research Letters* **48** e2020GL091197 ISSN 1944-8007 eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091197> URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091197>
- [22] Mathew S 2006 Performance of wind energy conversion systems *Wind Energy: Fundamentals, Resource Analysis and Economics* ed Mathew S (Berlin, Heidelberg: Springer) pp 145–178 ISBN 978-3-540-30906-2 URL [https://doi.org/10.1007/3-540-30906-3\\_5](https://doi.org/10.1007/3-540-30906-3_5)
- [23] Troccoli A, Goodess C, Jones P, Penny L, Dorling S,



**Figure D1.** Errors on the return time curve estimate due to lack of data: (grey) error due to the direct effect of data size (“direct error”, see the main text) and (orange) error on the sampling of  $\tau_2$  (“ $\tau_2$ -error”, see the main text). For both errors, 95% confidence intervals are shown. The same return time curve for the JF capacity factor fluctuations  $CF'_{JF}$  as in figure 5 is shown in blue.

- Harpham C, Dubus L, Parey S, Claudel S, Khong D H, Bett P E, Thornton H, Ranchin T, Wald L, Saint-Drenan Y M, De Felice M, Brayshaw D, Suckling E, Percy B and Blower J 2018 *Advances in Science and Research* **15** 191–205 ISSN 1992-0636 URL <https://asr.copernicus.org/articles/15/191/2018/>
- [24] e Highway 2050 2015 Europe’s future secure and sustainable electricity infrastructure URL <https://docs.entsoe.eu/baltic-conf/bites/www.e-highway2050.eu/results>
- [25] for Climate Action (European Commission) D G, for Energy (European Commission) D G, for Mobility and Transport (European Commission) D G, De Vita A, Capros P, Paroussos L, Fragkiadakis K, Karkatsoulis P, Höglund-Isaksson L, Winiwarter W, Purohit P, Gómez-Sanabria A, Rafaj P, Warnecke L, Deppermann A, Gusti M, Frank S, Lauri P, di Fulvio F, Florou A, Kannavou M, Forsell N, Fotiou T, Siskos P, Havlík P, Tsiropoulos I, Evangelopoulou S, Witzke P, Kesting M, Katoufa N, Mitsios I, Asimakopoulou G and Kalokyris T 2021 *EU reference scenario 2020: energy, transport and GHG emissions : trends to 2050* (LU: Publications Office of the European Union) ISBN 978-92-76-39356-6 URL <https://data.europa.eu/doi/10.2833/35750>
- [26] ENTSO-E and ENTSOG 2022 TYNDP 2022 Scenario Report
- [27] for Climate Action (European Commission) D G, for Energy (European Commission) D G, for Mobility and Transport (European Commission) D G, Zampara M, Obersteiner M, Evangelopoulou S, De Vita A, Winiwarter W, Witzke H P, Tsani S, Kesting M, Paroussos L, Höglund-Isaksson L, Papadopoulos D, Capros P, Kannavou M, Siskos P, Tasios N, Kouvaritakis N, Karkatsoulis P, Fragkos P, Gomez-Sanabria A, Petropoulos A, Havlík P, Frank S, Purohit P, Fragiadakis K, Forsell N, Gusti M and Nakos C 2016 *EU reference scenario 2016: energy, transport and GHG emissions : trends to 2050* (LU: Publications Office of the European Union) ISBN 978-92-79-52374-8 URL <https://data.europa.eu/doi/10.2833/001137>
- [28] ENTSO-E and ENTSOG 2020 TYNDP 2020 Scenario Report
- [29] Saint-Drenan Y M, Troccoli A and Dubus L 2020 European maritime region definition type: dataset URL <https://zenodo.org/record/4009400>
- [30] Staffell I and Pfenninger S 2016 *Energy* **114** 1224–1239 ISSN 0360-5442 URL <https://www.sciencedirect.com/science/article/pii/S0360544216311811>
- [31] Drew D R, Cannon D J, Brayshaw D J, Barlow J F and Coker P J 2015 *Resources* **4** 155–171 ISSN 2079-9276 number: 1 Publisher: Multidisciplinary Digital Publishing Institute URL <https://www.mdpi.com/2079-9276/4/1/155>
- [32] Archer C L and Jacobson M Z 2007 *Journal of applied meteorology and climatology* **46** 1701–1717
- [33] Kempton W, Pimenta F M, Veron D E and Colle B A 2010 *Proceedings of the National Academy of Sciences* **107** 7240–7245
- [34] Hasche B 2010 *Wind Energy* **13** 773–784
- [35] Fertig E, Apt J, Jaramillo P and Katzenstein W 2012 *Environmental research letters* **7** 034017
- [36] Louie H 2014 *Wind Energy* **17** 793–810
- [37] Handschy M A, Rose S and Apt J 2017 *Renewable Energy* **101** 1124–1130
- [38] Gumbel E J 1941 *The Annals of Mathematical Statistics* **12** 163–190 ISSN 0003-4851, 2168-8990 publisher: Institute of Mathematical Statistics URL <https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-12/issue-2/The-Return-Period-of-Flood-Flows/10.1214/aoms/1177731747.full>
- [39] Lestang T, Ragone F, Bréhier C E, Herbert C and Bouchet F 2018 *Journal of Statistical Mechanics: Theory and Experiment* **2018** 043213 ISSN 1742-5468 publisher: IOP Publishing URL <https://doi.org/10.1088/1742-5468/aab856>



- [40] Gardiner C W 2009 *Handbook of Stochastic Methods for physics, chemistry, and the natural sciences* 4th ed (Springer, Berlin)
- [41] Herbert C and Bouchet F 2017 *Physical Review E* **96** 030201 publisher: American Physical Society URL <https://link.aps.org/doi/10.1103/PhysRevE.96.030201>
- [42] Kármán T V 1921 *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik* **1** 233–252 ISSN 1521-4001 eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/zamm.19210010401> URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/zamm.19210010401>
- [43] Jerez S, Thais F, Tobin I, Wild M, Colette A, Yiou P and Vautard R 2015 *Renewable and Sustainable Energy Reviews* **42** 1–15 ISSN 1364-0321 URL <https://www.sciencedirect.com/science/article/pii/S1364032114008144>
- [44] Wilks D S 2011 Chapter 5 - Frequentist Statistical Inference *International Geophysics (Statistical Methods in the Atmospheric Sciences* vol 100) ed Wilks D S (Academic Press) pp 133–186 URL <https://www.sciencedirect.com/science/article/pii/B9780123850225000051>
- [45] DelSole T and Tippett M 2022 *Statistical Methods for Climate Scientists* 1st ed (Cambridge University Press) ISBN 978-1-108-65905-5 978-1-108-47241-8 URL <https://www.cambridge.org/core/product/identifier/9781108659055/type/book>