A Visualization Tool of the Spatial and Temporal Variability of Wind Resources

Alban Puech

November 2022

1 Introduction

As the world is facing an unprecedented energy crisis, more and more investments are going into intermittent renewable energy sources (IRES). IRENA analysis forecasts wind power to represent 40 per cent of the EU-28 electricity needs in 2050 [1]. Our energy production depending more and more on weather conditions, there is a call for tools that could help in better understanding and dealing with the risks that this can present. In this paper, we present a dashboard-like visualization interface that aims at making the spatial and temporal variability of wind power more accessible to decision-makers. We also hope to bridge the gap between the available climate data and the energy industry looking for more intelligible analysis tools.

After describing the data that we used, we will briefly cover some of the technical considerations related to the development and deployment of the app. We will finally describe the different visualizations, and defend the design and technical choices.

2 Dataset and motivations

Dataset. The dataset used here is an ERA5-derived time series of European country-aggregate electricity wind power generation [2]. The authors use the ERA5 reanalysis data [3] to compute the nationally aggregated hourly capacity factor of 28 European countries. The capacity factor of a wind turbine is the ratio of the electrical energy produced by the turbine for the period of time considered to the total amount of electrical energy that could have been produced at full power operation during the same period. Reanalysis combines past weather observations and current atmospheric models to generate climate and weather historic data. It allows getting a complete weather record from sparse - both in space and time - past data. In addition to the wind speed data, the authors use the wind farm spatial distribution of 2017, taken from `https://thewindpower.net`. However, it is worth mentioning that, because the absolute wind power capacity is not used to compute the capacity factors, only the relative spatial distribution of wind turbines is assumed to be constant. The capacity factor of each country is estimated by aggregating the capacity factor computed for each grid box, weighted by its estimated installed capacity. The capacity factor in each grid box is derived using the 100m wind speeds and the power curve of the type of wind turbine maximizing the energy produced during the entire period (1979-2019), as indicated in [4].

Motivations. This visualization tool originates from a larger (research) project, which scope is to study the spatial and temporal distribution of low wind power (LWP) events in Europe. By getting a better understanding of how countries correlate in terms of LWP events with their neighbours, the goal is to later explore the potential of cross-border cooperation in balancing Europe's renewable energy production. This has already been studied in [5] or [6], but at smaller scales.

3 Development and Deployment

Software framework. Our visualization platform is built using Dash, a python library built on top of React.js that allows building data applications with perfect integration of Plotly.js. Dash was made popular for its use in bioinformatics database visualization [7]. The choice of this particular framework was mainly motivated by its compatibility with the Pandas and Plotly libraries, that we were already using in the context of our larger project. Dash uses callbacks to modify or add new HTML elements as a result of the user's action. We estimate the amount of time spent exploring the data set and producing the visualizations to a total of 20 hours. The web app development took 100 hours, and the deployment, debugging phase and report writing required an additional 25 hours.

Offline pre-processing. Our data set contains 359402 rows, each corresponding to an hour between 1979-01-01 01:00:00 and 2019-12-31 23:00:00. In addition to the timestamp column, the 28 other columns show the hourly mean capacity factor of each of the 28 selected European countries. We changed the dimension of our dataset to have one entry per timestamp and per country, multiplying the number of rows by 28 (the number of countries), and keeping only two columns: the timestamp and the capacity factor. We also pre-computed the day, month and year features out of the timestamp column to accelerate the access of the web app to those features. We finally compressed the dataset.

Online data processing. The compressed dataset is read during the loading of the web application, and is saved in the RAM. The numerical features are further down-casted to save memory and computational power. When the application is loaded, only a minimum amount of data processing is executed to allow the app to quickly reload if it has to restart after a server reboots. Only 5 datasets are pre-computed. For instance, one of them contains the mean capacity factor for each hour of the day, computed across all days of the year, for each year and each country. Those pre-computed datasets are necessary to reduce the size of the data frames on which are executed the operations, and keep the rest of the processing - executed on-demand when a specific plot is requested - short. Only the most used data frames are pre-computed to make sure that the overhead will be amortized in the span of an average user session. The rendering time of most visualizations was gradually improved throughout an iterative process, and by running multiple benchmarking experiments.

As our dataset is quite large (over 500Mb and 10 million rows), every visualization generation would be computationally heavy if done on the original dataset. However, our web app being highly interactive and the plots being all configurable by the user, we couldn't store all the pre-processed data frames as the number of possible visualizations rendered by the app is infinite.

Deployment specifics. We used `eu.pythonanywhere.com`, a hosting service compatible with WSGI-based application framework like Dash to deploy our app. Deploying our tool required us to extensively work on the efficiency of our code, as described previously, as RAM, storage and CPU resources were highly limited. We also had to write a periodic task to ensure that the web app is always preloaded before a user tries to access it. The app is available at the address: `http://albanpuech.eu.pythonanywhere.com/`

4 Interactive Two-Card-Layout

The layout of the web application is based on two cards, placed side-by-side, as shown in Fig. 1 Each card groups together elements that interact with each other (e.g. a dropdown that what data is plotted on a graph). Hence, each of the two cards is independent, with a unique exception: The choropleth map of the left card is used to filter the countries plotted on the right card.

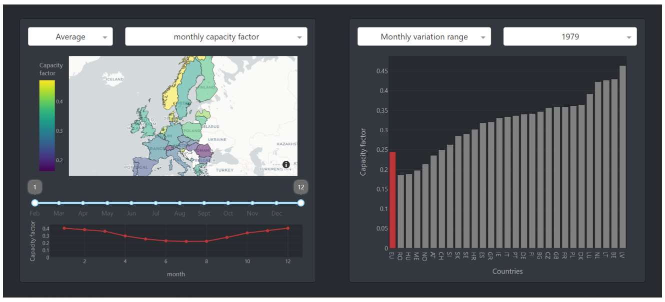

Figure 1: Layout of the web app. The left card highlights the spatial and temporal variability of wind. The right one offers more advanced visualizations.

4.1 Left card: Choropleth map and line plot

The left card displays two main visualizations: A choropleth map and a line plot.

4.1.1 Capacity factor spatial variation (Choropleth map)

European countries are not equal in terms of wind resources [8]. The goal of this map is to highlight the spatial distribution of wind power. It aims at providing an easy way to compare the average and the standard deviation of the capacity factor of European countries, for different time resolutions and over different time periods.

Choice of the visualization type. The choropleth map shows the capacity factor of each European country. Because choropleth maps use a colour scale to represent the value of the considered feature, they can't be used to represent precise values, but are good at showing how different geographic entities compare, which is what we are looking for. Moreover, choropleth maps are used to plot relative data, that doesn't depend on the size of the population of the geographic entities, which is the case of the wind capacity factor.

Data filtering and visualization setting. Two dropdowns positioned above the map can be used to control what data is displayed. They are shown in Fig. 2 The right one is used to choose the resolution of the data: yearly, monthly, or hourly capacity factor. The left one allows choosing whether to plot the average or the standard deviation of the (yearly, monthly or hourly) capacity factor, over the entire period (1979-2019). A range slider, positioned right below the map, is used to filter the years, the months or the hours (according to whether the yearly, monthly or hourly resolution is selected) that are used in

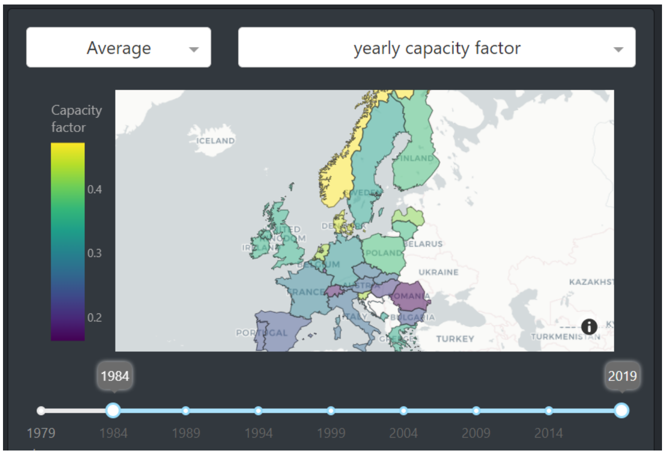

Figure 2: Choropleth map of the average yearly capacity factor over the period 1984-2019. The range slider and the two dropdowns are used to parametrize the visualization.

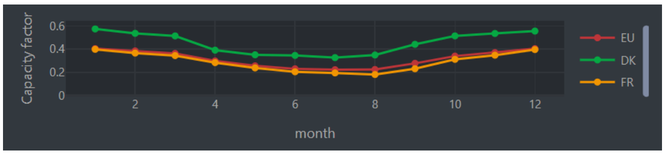

Figure 3: Line plot showing the average monthly capacity factor of France, Denmark and of the entire region (France+Denmark) for the period 1979-2019. The scale starts at 0 to allow a better comparison between countries.

the calculation of the average or the standard deviation of the capacity factor displayed on the map.

Scale. The colour scale range adapts to the range of the data shown on the choropleth map, and its limits can then change when other parameters are used for the plot. This is because the choropleth map is used to compare countries with each other, but is not used to compare the capacity factor across different time periods: the line plot can be used for that. Another reason for this design choice is that keeping the same scale would make the comparison of the capacity factor of the countries impossible when the range of the data is very small (e.g. when plotting the standard deviation of the monthly factor over July and August, that have very similar capacity factors.)

4.1.2 Capacity factor temporal variation (line plot)

We display a line plot of the capacity factor below the choropleth map. The role of the line plot is to show the intraday, intrayear and yearly capacity factor variations of European countries, depending on what resolution the user chose. It also allows comparison of the capacity factor between countries, over different time windows.

Data filtering and visualization settings. The users can select one or multiple countries by holding shift when clicking on the choropleth map. When no country is selected (initial state), the average capacity factor over all 28 European countries available is displayed, and the line is labelled as "28C". When one or more countries are selected, new lines, each corresponding to the selected countries, are added to the plot. This allows the user to compare a country with the Europe-aggregated data, or compare multiple countries, as shown in Fig. 3. The range slider described earlier acts both on the choropleth map and on the line plot, and can then be used to filter the time period that is displayed.

Scale. The scale starts at 0 so that the user is not misled by small changes in the capacity factor, which would appear as a large line-height variation if the axis was truncated. Another reason is that this chart is also used to compare countries, and we thus want the relative height of the lines to accurately represent the difference in capacity factor between the two countries.

4.2 Right card: Advanced visualizations

The left map displays "raw" data, with little preprocessing involved. It provides a simple yet clear description of the spatial and temporal distribution of the capacity factor. The right card provides more elaborated visualizations, that require more pre-processing. Those visualizations are to be interpreted with the data shown on the left card. The layout was designed to allow the two plots to be side by side, so that the user does not need to switch between them.

4.2.1 Intrayear (Intraday) variation range of the monthly (hourly) capacity factor

Motivations. Understanding the Wind generation seasonal patterns is of great interest to better integrate wind energy into the grid [5]. For example, we may want to quantify the capacity factor gap between winter and summer seasons and compare this gap across countries. The "Intrayear variation range of the monthly capacity factor" bar plot shows the variation range of the monthly capacity factor of each country. The same plot is provided for the intraday variation range of the hourly capacity factor. When no country is selected on the left card, all European countries are plotted, in ascending order of values, as shown in Fig. 4. This is because the goal of this visualization is to provide a ranking of the countries based on their capacity factor variability. Moreover, the bar corresponding to the variability of the 28-countries-aggregated capacity factor is displayed in red and highlights the po-

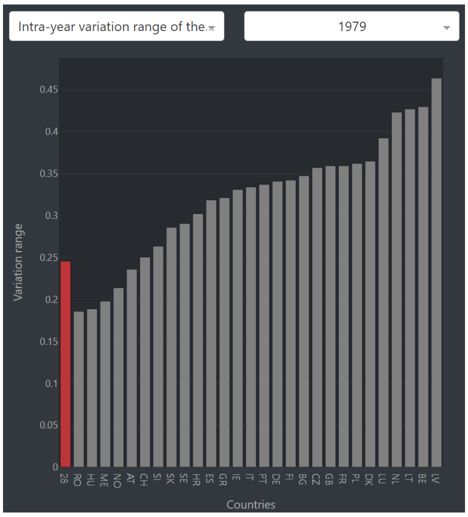

Figure 4: Bar chart of the intrayear variation range of the monthly capacity factor of the 28 countries in 1979. The bar corresponding to the 28 countries-aggregated capacity factor range is shown in red.

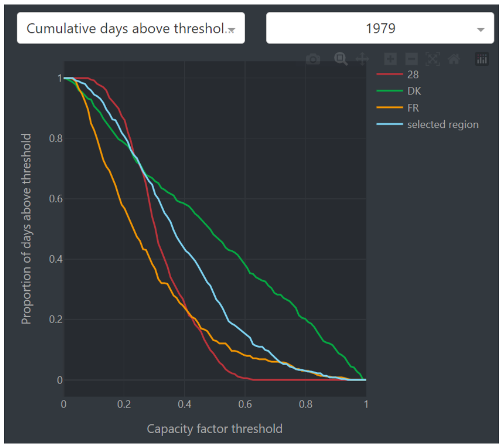

Figure 5: Line plot of the cumulative days above thresholds for the year 1979. We display the data for France and Denmark, and the data aggregated over the region (France+Denmark) in blue.

tential of European cross-border cooperation and grid interconnections in mitigating the temporal variability of wind energy.

Visualization settings. The scale starts at 0 for the difference in bar heights to accurately represent the difference in capacity factor variation range. Selecting one or more countries on the left card allows filtering the data displayed on the plot. The colour used for each country is consistent across all visualizations of the web app, allowing the user to quickly spot the selected countries on each plot. Finally, the user can select the year using a dropdown located on the top part of the card.

4.2.2 Monthly and hourly capacity factor box plots.

While the previous plots gave information about how variable the monthly and hourly capacity factors are for each given country, they do not give any information on the raw capacity factor values. The user might want to have both pieces of information at the same time, which is why we also propose a scatter plot of the "Min, Max, Avg intraday hourly capacity factor", shown in Fig. 4

Visualization settings. The scatter plot displays the mean with a cross and the min and max values with error bars. The countries are ordered by mean values. The rest of the visualization settings are similar to the one described in the previous subsection.

4.2.3 Cumulative time plot

Motivations. One way to compare the capacity factor of different regions is to look at their proportion of days that had a capacity factor higher than some threshold. This is what can be done using the "Cumulative days above threshold" plot.

Visualization settings. When no country is selected, the data for the 28 countries-aggregated capacity factor is displayed, and the lines corresponding to each European country are displayed in light grey. This allows the user to quickly see how each country compares to the other 27 European countries by hovering over its corresponding line. When one or more countries are selected, their lines are displayed on the graph. We also add a line that corresponds to the data aggregated over the entire selected region, as in Fig. 5. Again, this allows the user to consider the selected countries as a unique region in the context of perfect grid interconnection.

4.2.4 YoY (year-over-year) monthly capacity factor comparison

Motivations. So far, most of the visualizations were focusing on the spatial distribution of wind power, and on the comparison of capacity factor-derived features across countries. However, the increasing investments into wind energy have pushed for more research on the year-over-year evolution of wind energy resources, and on the long-term evolution of wind resources [9]. In the "YoY (year-over-year) monthly capacity factor comparison" plot, we display the intra-year evolution of the capacity factor for the selected country and the selected year.

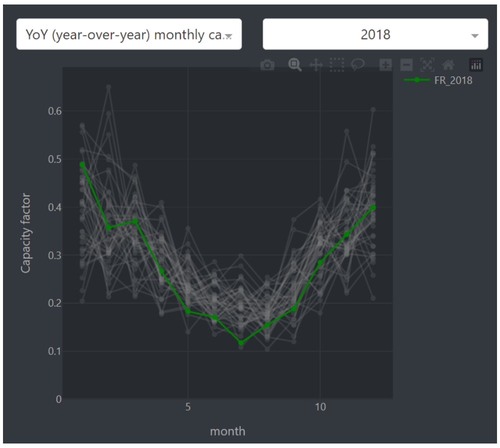

Figure 6: Year-over-Year monthly capacity factor comparison for France, The highlighted year is 2018. The grey lines correspond to the other years of the period (1979-2019). The lines can be hovered to know the year they correspond to.

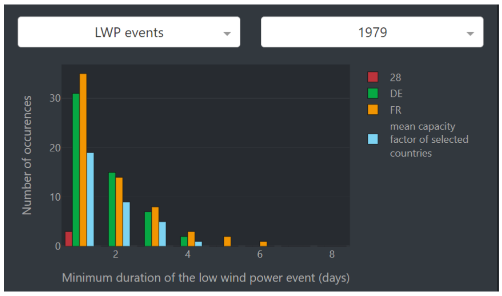

Figure 7: Number of low wind power events in France, Germany, and in the entire (France+Germany) region for different minimum LWP events duration in 1979. The selected region had fewer low wind power events because a country can on some occasions compensate for the low capacity factor of the other, especially when they have different seaboards and are exposed to different wind regimes.

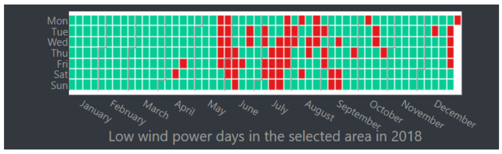

Figure 8: Calendar plot of the low wind power days in France in 2018. Low wind power days are displayed in red.

Visualization settings. The lines corresponding to the other years of the period 1979-2019 are displayed in light grey, allowing the user to quickly figure out how the capacity factor of a given year compares to the other years of the period. This is shown in Fig. 6. As the goal here is really to compare the intra-annual capacity factor over the years for a single region, selecting multiple countries results in showing only one single line corresponding to the data aggregated over the entire selected region.

4.3 Low wind power events

Motivations. compared to solar photovoltaics, wind energy has much more unpredictable variations. In particular, the study of the temporal distribution of low-wind-power (LWP) events has gained more attention in the literature [10, 11]. In our web app, we define LWP events to be consecutive days during which the daily capacity factor is constantly below a threshold of 10 per cent. When the user selects "LWP events" in the dropdown, two plots are displayed. The first one is a bar plot of the number of occurrences of low wind power events for each minimum duration. The second one is a calendar plot that indicates the low wind power days in the selected region.

Visualization settings. When no country is selected, the plot is replaced by a message explaining that at least one country has to be selected. This is because the bar chart would not be readable if we were to display the data of all countries. While we could display the 28 countries-aggregated data on other plots, it is not the case here since most bars in the bar plot would be of size 0, which would probably confuse the user. Indeed, there is no low wind power event of duration longer than 2 days at the scale of the 28 countries-aggregated data. When the user selects multiple countries, the bar chart displays grouped bars corresponding to each country. This allows for comparing the number of occurrences of LWP events of each minimum duration across selected countries. We additionally add the data corresponding to the selected-region-aggregated data. This allows the user to see how grid interconnection mitigates the risk of observing LWP events, as shown in Fig. 7. Indeed, a selected region often has a lower number of LWP events than each of its constitutive countries, since LWP events don't necessarily happen at the same time in all constitutive countries. The calendar plot indicates the low wind power days at the scale of the selected region. This plot gives information on the temporal distribution of those days within the considered year. An example of such plots is shown in Fig. 8

4.4 Choropleth correlation plots

Motivation. The previous plots that we described allow the user to compare countries in terms of their number of LWP events. However, understanding how the capacity factor of neighbours countries are correlated is of major

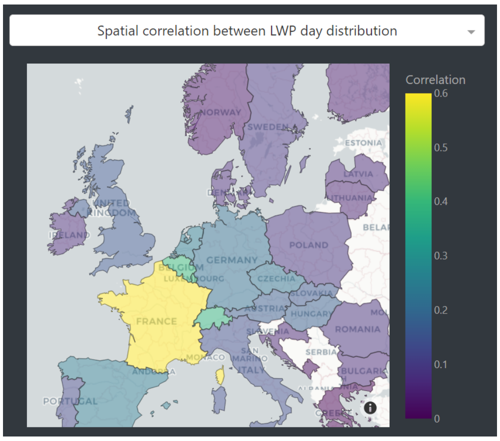

Figure 9: Correlation between the LWP events distribution of France and the ones of the 28 other European countries.

importance to balance the spatial variability of wind energy [12, 13, 14]. For this reason, we propose two different plots. Both of them require one country to be selected on the left card. The first plot shows the Pearson correlation coefficient between the selected country and the other countries in terms of LWP day distribution. The second one shows the same statistics, but for the raw capacity factor value.

Visualization settings. In the first plot, most correlation coefficient values are much smaller than one, so that using a scale range of [0,1] would make most of the countries have the same colour. For this reason, we decided to use the value obtained for the most correlated country (excluding the selected country itself, which has a correlation of 1 with itself) as the upper bound for the range. This is shown in Fig. 9. Another possibility was to use a log scale, but we wanted to keep the scale consistent with the one of the second plot, for which we did not want to use a log scale.

5 Conclusion and Future Work

In this paper, we presented a new web app that offers multiple visualizations of the temporal and spatial variability of wind. The tool proposes configurable plots that allow the user to deeply analyse the ERA5-derived capacity factor dataset [2]. In the future, we will add the solar energy capacity factor data, and new visualizations illustrating the effects of combining different energy sources in mitigating their variability [15].

References

[1] IRENA, "Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects (A Global Energy Transformation paper)," tech. rep., International Renewable Energy Agency, 2019.

[2] H. Bloomfield, D. Brayshaw, and A. Charlton-Perez, "Era5 derived time series of european country-aggregate electricity demand, wind power generation and solar power generation: hourly data from 1979-2019," 2020.

[3] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut, "The era5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020.

[4] H. Bloomfield, D. Brayshaw, D. Brayshaw, P. Gonzalez, P. Gonzalez, and A. Charlton-Perez, "Subseasonal forecasts of demand and wind power and solar power generation for 28 european countries," *Earth System Science Data*, vol. 13, pp. 2259–2274, 05 2021.

[5] F. Kaspar, M. Borsche, U. Pfeifroth, J. Trentmann, J. Drücke, and P. Becker, "A climatological assessment of balancing effects and shortfall risks of photovoltaics and wind energy in germany and europe," *Advances in Science and Research*, vol. 16, pp. 119–128, 07 2019.

[6] M. Child, C. Kemfert, and D. Bogdanov, "Flexible electricity generation, grid exchange and storage for the transition to a 100 per cent renewable energy system in europe," *Renewable Energy*, vol. 139, 02 2019.

[7] Shammamah Hossain, "Visualization of Bioinformatics Data with Dash Bio," in *Proceedings of the 18th Python in Science Conference* (Chris Calloway, David Lippa, Dillon Niederhut, and David Shupe, eds.), pp. 126 – 133, 2019.

[8] P. Enevoldsen, F.-H. Permien, I. Bakhtaoui, A.-K. von Krauland, M. Z. Jacobson, G. Xydis, B. K.

Sovacool, S. V. Valentine, D. Luecht, and G. Oxley, "How much wind power potential does europe have? examining european wind power potential with an enhanced socio-technical atlas," *Energy Policy*, vol. 132, pp. 1092–1100, 2019.

[9] C. Jung and D. Schindler, "Development of onshore wind turbine fleet counteracts climate change-induced reduction in global capacity factor," *Nature Energy*, vol. 7, 07 2022.

[10] Y. Cai and F.-M. Bréon, "Wind power potential and intermittency issues in the context of climate change," *Energy Conversion and Management*, vol. 240, p. 114276, 2021.

[11] N. Ohlendorf and W.-P. Schill, "Frequency and duration of low-wind-power events in germany," *Environmental Research Letters*, vol. 15, 08 2020.

[12] J. Olauson and M. Bergkvist, "Correlation between wind power generation in the european countries," *Energy*, vol. 114, pp. 663–670, 11 2016.

[13] F. Monforti, M. Gaetani, and E. Vignati, "How synchronous is wind energy production among european countries?," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1622–1638, 2016.

[14] A. Malvaldi, S. Weiss, D. Infield, J. Browell, P. Leahy, and A. M. Foley, "A spatial and temporal correlation analysis of aggregate wind power in an ideally interconnected europe," *Wind Energy*, vol. 20, no. 8, pp. 1315–1329.

[15] C. Frank, S. Fiedler, and S. Crewell, "Balancing potential of natural variability and extremes in photovoltaic and wind energy production for european countries," *Renewable Energy*, vol. 163, pp. 674–684, 2021.