An Analysis of Wind Energy Potential in the North Sea

Mohammad Fadel Berakdar [*1] **Gwendolyn Neitzel** [*2] **David Voigt** [*3] **Alireza Yahyanejad** [*4]

Abstract

In the face of climate change, it is widely agreed that energy production has to rely on more sustainable and renewable forms of harnessing energy. Offshore wind turbine parks play a crucial role in increasing the share of green energy. This paper explores probabilistic methods of assessing the wind energy potential and potential trends of data collected on Helgoland by considering wind speeds as a Weibull distributed random variable. Further, for forecasting, the monthly expected wind power density is extrapolated using a Gaussian process regression model.

1. Introduction

With the rising need for clean and sustainable energy due to climate change, offshore wind energy is playing an important role in the future energy mix of many European countries (European Commission, 2020). In the context of Germany, the North Sea is considered a prime location for offshore wind parks, with 41 farms already established and many more in planning (Chirosca et al., 2022). For assessing the viability of offshore wind parks, an often used metric is the wind power density (Wang & Liu, 2021; Miao et al., 2020; Mohammadi et al., 2016). In this paper, we follow Mohammadi et al. (2016) by considering the wind speed as a Weibull distributed random variable (Section 3.1). Further, we conduct a trend analysis (Section 3.1) on the change of the parameters of the yearly wind speed distributions and look for evidence for the phenomenon of global terrestrial stilling, that is, a decrease in the global mean annual surface wind speed starting in the 1980s (Tian et al., 2019). As proposed by Mohammadi et al. (2016), we compute the wind power density probabilistically as the expectation of the monthly and yearly Weibull distributions (Section 3.2). We then employ a Gaussian process regression model for the prediction of future wind power densities (Section 3.3). In Section 4, we present the results of applying the aforementioned methods to real-world data.[1]

2. Data

The wind speed data used in this paper was collected by the DWD on Helgoland, spanning the years from 1996 to 2022. The wind speed is recorded at a temporal resolution of ten-minutes[2], with the anemometer being $4.38\mathrm{m}$ above ground. Averaged over the whole measurement period, the data has a yearly mean completeness of about 91%. Notably, the year 2019 is missing 88% of its measurements. We consider this, and a noticeable drop in the yearly wind speed mean and variance of the following years, as evidence that the measurement process was changed, and therefore discard them. Notably, the years 1996 to 1998 have a completeness of about $3\%, 85\%$ and 84%, respectively. For preventing a distorting effect on the Weibull distribution estimations, we also did not consider these years for our analysis, thereby improving the yearly mean completeness to 99%. For handling outliers, we dropped measurements exhibiting a z-score > 8.

3. Methods

3.1. Fitting of the Weibull Distribution and Trend Analysis

Based on the observed wind data, we model the wind speed probability distribution both on a monthly and yearly basis. Whilst there are several other models, like the three-parameter Weibull distribution or the normal distribution for fitting wind speed frequencies, the most prominent one is the two-parameter Weibull distribution (Shi et al., 2021; Akpinar & Akpinar, 2004; Mohammadi et al., 2016). For the scale parameter $\lambda > 0$ and shape parameter $\beta > 0$, the probability density function (PDF) and cumulative density

[*]Equal contribution [1]Matrikelnummer 6117917, mohammad-fadel.berakdar@student.uni-tuebingen.de, MSc Bioinformatics [2]Matrikelnummer 5425507, gwendolyn.neitzel@student.uni-tuebingen.de, MSc Mathematik [3]Matrikelnummer 5416770, david.voigt@student.uni-tuebingen.de, MSc Computer Science [4]Matrikelnummer 6645496, alireza.yahyanejad@student.uni-tuebingen.de, MSc Computer Science.

Project report for the "Data Literacy" course at the University of Tübingen, Winter 2023/24 (Module ML4201). Style template based on the ICML style files 2023.

[1]The complete source of this paper can be found in our repositoriy.

[2]Averaged over the last ten-minute interval.

function (CDF) are given by

$$p(v) = \begin{cases} \frac{\beta}{\lambda}\left(\frac{v}{\beta}\right)^{\beta-1} \exp\left(-\left(\frac{v}{\lambda}\right)^{\beta}\right) & \text{for } v > 0 \\ 0 & \text{else} \end{cases} \quad (1)$$

$$P(v) = \begin{cases} 1 - \exp\left(-\left(\frac{v}{\lambda}\right)^{\beta}\right) & \text{for } v > 0 \\ 0 & \text{else} \end{cases}. \quad (2)$$

Via the gamma function, $\Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}dt$, the k-th raw moment and the variance of a random variable V with Weibull distribution can be expressed as follows:

$$\mathrm{E}[V^k] = \lambda^k \Gamma\left(1 + \frac{k}{\beta}\right) \quad (3)$$

$$\mathrm{Var}[V] = \lambda^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right]. \quad (4)$$

We determine the parameters λ and β for every month and year using the maximum likelihood method applied to (1) (Mohammadi et al., 2016). Due to the seasonal variability of wind speed, the long-term trend analysis is conducted on the yearly (rather than monthly) estimated wind speed PDFs. For this, we employ a least-squares linear regression and assess the statistical significance through a two-sided permutation test at the 5% level. Since the two Weibull parameters fully characterize the distribution, we are then able to use this analysis to infer trends in other quantities of interest, like mean and standard deviation.

3.2. Wind Power Density

The wind power density is often used as an approximation of the physical upper limit of how much energy a wind turbine is able to harness (Wang & Liu, 2021; Mohammadi et al., 2016; Miao et al., 2020). The power density is given by

$$\mathrm{P} = \frac{1}{2} \cdot \rho \cdot \mathrm{v}^3 \quad (5)$$

$$[\mathrm{P}] = [\rho] \cdot [\mathrm{v}^3] = \frac{\mathrm{kg}}{\mathrm{m}^3} \cdot \frac{\mathrm{m}^3}{\mathrm{s}^3} = \frac{\mathrm{W}}{\mathrm{m}^2} \quad (6)$$

where ρ is the air density, and v is the wind speed. Notice that the wind power density is independent of the area covered by the wind turbine's rotors. Using equation (5) the expected power density can be found by calculating the expected value under the given distribution $p(\mathrm{v})$ as follows:

$$E[P] = \int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}} \frac{1}{2} \cdot \rho \cdot \mathrm{v}^3 \cdot p(\mathrm{v})\,\mathrm{dv} \quad (7)$$

$$= \frac{1}{2} \cdot \rho \cdot \int_{\mathrm{v_{min}}}^{\mathrm{v_{max}}} \mathrm{v}^3 \cdot p(\mathrm{v})\,\mathrm{dv}. \quad (8)$$

For a range of $v_{\mathrm{min}} = 0\,\frac{\mathrm{m}}{\mathrm{s}}$ and $v_{\mathrm{max}} = \infty\,\frac{\mathrm{m}}{\mathrm{s}}$, the solution to this integral is given by the third raw-moment as shown in (3). However, in using these limits, we would assume that wind turbines are able to convert energy during arbitrarily high or low wind speeds. Thus, a more sensible assumption is to use the so-called cut-in $\mathrm{v_{in}}$ and cut-out wind speed $\mathrm{v_{out}}$ of wind turbines. For a generic wind turbine, a common assumption is $\mathrm{v_{in}} \in [2, 4]\,\frac{\mathrm{m}}{\mathrm{s}}$ and $\mathrm{v_{out}} \in [20, 23]\,\frac{\mathrm{m}}{\mathrm{s}}$ (Elise Dupont, 2017; Wang & Liu, 2021).

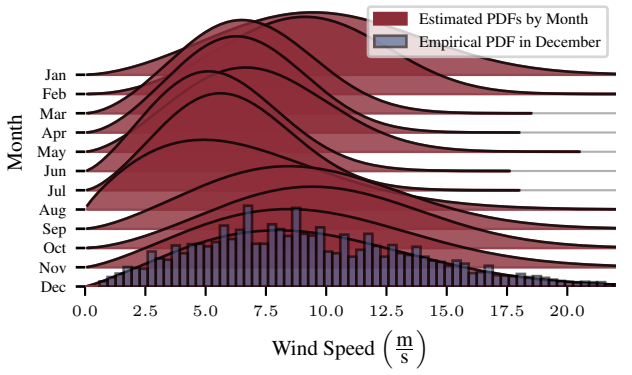

Figure 1. The monthly estimated Weibull distributions in the year 2010. The parameter λ decreases in summer, whilst β decreases, resulting in a shift to the left and lower standard deviation.

3.3. Gaussian Process Regression

For forecasting, we employ a Gaussian process regression (GPR) model to perform a time-series regression on the monthly estimated power density and extrapolate future values. Thereby, we assume that the wind speed data $Y \in \mathbb{R}^N$ is from a latent function $f : \mathbb{R}^N \to \mathbb{R}^N$ with some added measurement noise: $y = f(x) + \epsilon$. The GPR framework consists of a Gaussian process as prior, multivariate Gaussian likelihood and Gaussian process posterior conditioned on the data (Rasmussen & Williams, 2005):

$$p(f) = \mathcal{GP}(f; 0, k), \quad (9)$$

$$p(y \mid f) = \mathcal{N}(Y; f, \sigma^2 I), \quad (10)$$

$$p(f \mid Y) = \mathcal{GP}(f; m', k') \propto p(Y \mid f) \cdot p(f). \quad (11)$$

The resulting posterior (and prior) distribution collapses into a multivariate normal distribution when evaluated at some new locations $X' \in \mathbb{R}^M$.

4. Results

4.1. Weibull Parameter Estimation and Trend Analysis

For the available years (1999-2018) we estimated the parameters of the Weibull distribution on a monthly and yearly basis. Figure 1 depicts the resulting PDFs for the months of 2010 as well as the empirical histogram in December. When applying the square root rule to the number of bins, this fit has an expected relative error per bin of 18%. Averaged over

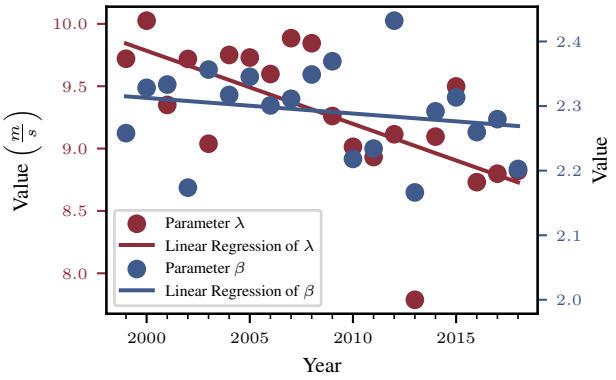

Figure 2. The linear regression of the yearly Weibull parameters shows a noticeable decrease in λ, but less so in β.

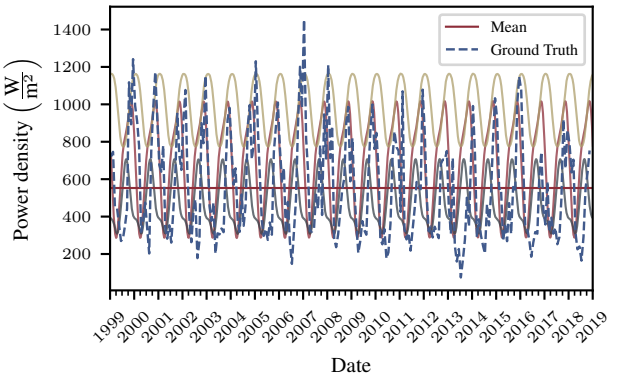

Figure 3. The ground truth data plotted together with three samples $f_i \sim \mathcal{GP}(f; 0, k)$ drawn from the Gaussian Process prior distribution with the added mean $\overline{Y}$ of the data Y.

the monthly and yearly distributions, the expected relative error per bin is 23% and 22%, respectively.

The least-squares linear regression of the yearly parameters can be seen in Figure 2. It shows a decrease of the parameter λ of $-0.059 \frac{m}{s} \frac{1}{\text{yr}}$ with an RMSE of $0.42 \frac{m}{s}$. For β, the slope is $-0.002 \frac{1}{\text{yr}}$, with an RMSE of $0.07 \frac{m}{s}$. At the 5% level of the two-sided permutation test, the null hypothesis that there is no linear trend in the development of the parameter λ can be rejected (p-value$= 0.001$), so the observation is statistically significant. For the parameter β, however, the p-value is 0.383, so we cannot reject the null hypothesis. Assuming the trend observed in λ and the empiric mean $\bar{\beta} = 2.29$ for β, we derive the following trends in other quantities: As the mean and standard deviation are proportional to λ, we can infer a linear trend of $-0.052 \frac{m}{s} \frac{1}{\text{yr}}$ and $-0.024 \frac{m}{s} \frac{1}{\text{yr}}$, respectively. Further, under these assumptions, the mean power density of the mean wind speed decreases (non-linearly) from 1999 to 2018 by 30%.

4.2. Expected Power Density

By using (7), we numerically computed the expected power density under each of the estimated monthly Weibull distributions with $v_{\min} = 3 \frac{\text{m}}{\text{s}}$ and $v_{\max} = 22 \frac{\text{m}}{\text{s}}$. The resulting values can be found as the ground truth in Figure 3 and 4.

4.3. Gaussian Process Regression

The first step for conducting a GP regression is to decide on a kernel k for the prior (9). By analysing the plotted raw monthly expected power density, we identify a yearly periodic trend and a slight global downward trend. The former can be captured by the Exponential-Sine-Squared kernel (ES) (MacKay, 1998). Supported by the empirical data, we choose a periodicity of 365 days and an amplitude of $450 \frac{\text{W}}{\text{m}^2}$. For the length scale, we settle for a rather small value (< 2 days) in order to capture local patterns in the periodicity more accurately. For allowing the amplitude to slowly decay over long periods of time, we multiply the ES kernel by a very smooth (length scale > 50 years) Squared-Exponential (SE) kernel. For reproducing the global downwards trend observed in section 4.1, we also choose a very smooth Squared-Exponential kernel with an output scale of $100 \frac{\text{W}}{\text{m}^2}$. For accounting for the monthly midterm trend, we add a Rational-Quadratic kernel (Rasmussen & Williams, 2005) with a scale of $10 \frac{\text{W}}{\text{m}^2}$ and a length scale of one month. Finally, the additive kernel, consisting of a White-Noise (WN) and a Matern kernel (MA) (Abramowitz et al., 1968), is responsible for explaining measurement noise in the data and thus also prevents overfitting. Thereby, the kernel has the form

$$k = \underbrace{\text{ES} \cdot \text{SE}}_{\text{yearly trend}} + \underbrace{\text{RQ}}_{\text{monthly trend}} + \underbrace{\text{ES}}_{\text{longterm trend}} + \underbrace{\text{WN} + \text{MA}}_{\text{noise}}.$$

By drawing samples[3] $f_i \sim \mathcal{GP}(f; 0, k)$ from the prior distribution (9), we gauge how accurately our kernel choice fits to the given data. The results are shown in Figure 3. We can observe that the amplitude and periodicity of some of the samples mirrors the data's in an accurate manner, supporting our choice of hyperparameters. After conditioning the Gaussian process prior on the normalised data, we obtain the posterior distribution $\mathcal{GP}(f; m', k')$ (11). The evaluated posterior mean and standard deviation can be found in Figure 4. We can observe that the mean function is following the periodicity of the ground truth without overfitting extreme power densities. Furthermore, the uncertainty of the model (standard deviation) increases gradually the further the prediction is from the data. Notably, the model also captures the global downward trend. When comparing the prediction to the test data in the years 2017 and 2018, we can observe that the prediction fits the ground truth accurately and all data points are well within one standard deviation of the mean.

[3] with the mean of the data added.

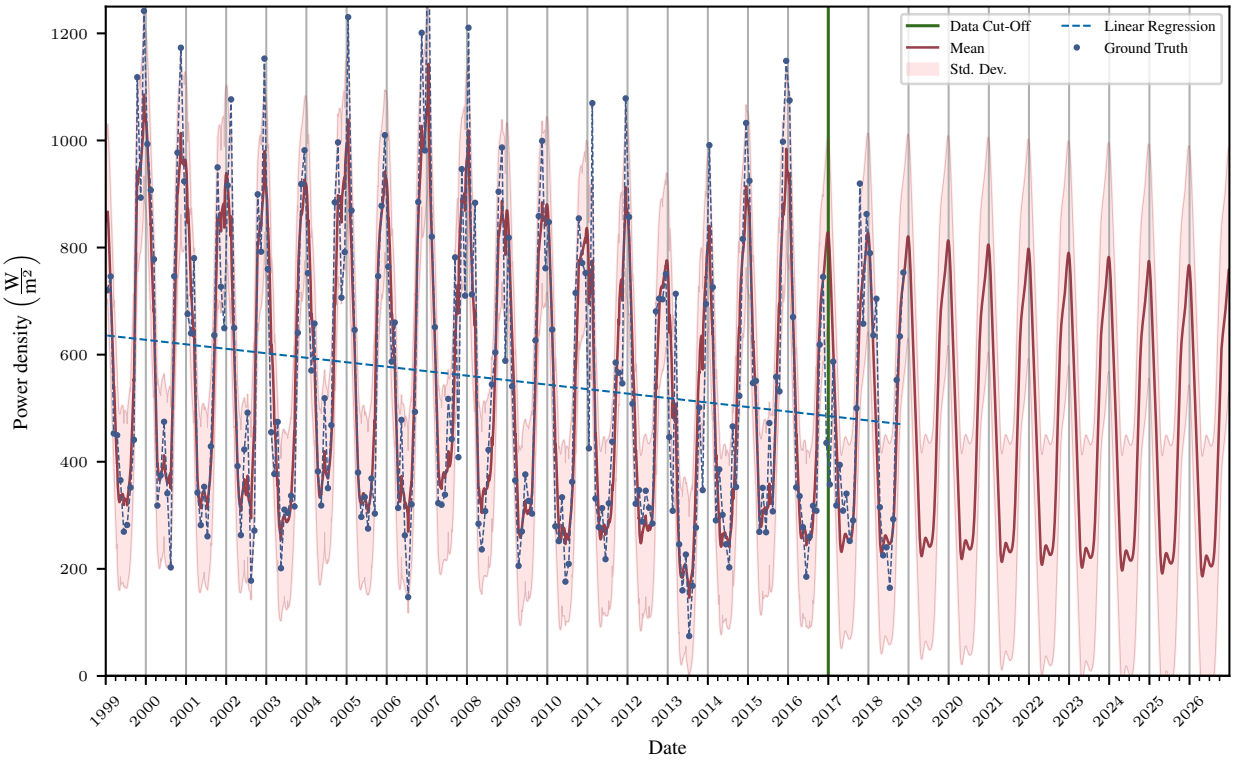

Figure 4. The mean and standard deviation of the posterior distribution $\mathcal{N}(f(x); m', k')$ evaluated at x is plotted together with the linearly interpolated ground truth. Starting at the data cut-off, the model is extrapolating (predicting) the power density into the future.

5. Discussion & Conclusion

In this paper, we analysed the wind energy potential of the island Helgoland in the North Sea. We have found a decrease in the mean annual wind speed of $-0.052 \frac{m}{s} \frac{1}{\mathrm{yr}}$, consistent with the global phenomenon of terrestrial stilling (Tian et al., 2019). In this location, a reversal in the stilling, as was observed globally (Zeng et al., 2019), could not be found. If the trend prevails, diminishing wind energy production is to be expected at this particular location. Further, we have shone light on the monthly variations in expected power density, a valuable insight for evaluating potential contribution to the overall energy supply. Lastly, we explored Gaussian process regression as a tool for extrapolating the expected power density. We have found that it is adequate for this purpose, being able to adapt to a variety of subtrends and cyclical behaviours due to the explicit and flexible choice of the prior kernel function. By comparing the test data with the prediction, we consider the fit to be accurate and the uncertainty quantification realistic.

However, there are some shortcomings. One of the inherent limitations of the GPR is that it yields functions mapping onto $\mathbb{R}$, while the power density is only limited to $\mathbb{R}_+$. Another important limitation is that our data was recorded at a height of $4.38\mathrm{m}$, whereas the real height of wind turbines can exceed $100\mathrm{m}$. One commonly applied rule of thumb to mitigate this issue is multiplying the wind speeds by the factor $\left(\frac{h}{4.38\mathrm{m}}\right)^{0.14}$ or the power density by its cube, where h is the proposed hub height of a turbine (Akpinar & Akpinar, 2004). However, this is far from exact, as the winds at different heights could be subject to different circulations. Furthermore, we did not take the direction of the wind into account and made the simplification that wind turbines are always perfectly aligned. Also, for equation (5), we assumed a constant air pressure rather than considering the monthly changing pressure. Lastly, with the use of the maximum-likelihood method, we assumed the independence of measurements. While this is generally not given, we were still able to produce adequate estimations.

Contribution Statement

David Voigt conducted the Gaussian process regression. Gwendolyn Neitzel was mainly concerned with the Weibull model fitting and the trend analysis. Mohammad Fadel Berakdar focused on the data preprocessing. Alireza Yahyanejad worked on data loading. All authors were involved in data exploration and jointly wrote the report.

References

Abramowitz, M., Stegun, I. A., et al. *Handbook of mathematical functions*, volume 10. Dover, New York, 1968.

Akpinar, E. K. and Akpinar, S. A statistical analysis of wind speed data used in installation of wind energy conversion systems. *Energy Conversion and Management*, 46, 5 2004. doi: https://doi.org/10.1016/j.enconman.2004.05.002.

Chirosca, A.-M., Rusu, L., and Bleoju, A. Study on wind farms in the north sea area. *Energy Reports*, 8:162–168, 2022. ISSN 2352-4847. doi: https://doi.org/10.1016/j.egyr.2022.10.244.

Elise Dupont, Rembrandt Koppelaar, H. J. Global available wind energy with physical and energy return on investment constraints. *Applied Energy*, 209, 10 2017. doi: https://doi.org/10.1016/j.apenergy.2017.09.085.

European Commission. An EU Strategy to harness the potential of offshore renewable energy for a climate neutral future, 2020. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:741:FIN&qid=1605792629666.

MacKay, D. J. C. Introduction to gaussian processes. 1998.

Miao, H., Dong, D., Huang, G., Hu, K., Tian, Q., and Gong, Y. Evaluation of northern hemisphere surface wind speed and wind power density in multiple reanalysis datasets. *Energy*, 200:117382, 2020. ISSN 0360-5442. doi: https://doi.org/10.1016/j.energy.2020.117382.

Mohammadi, K., Alavi, O., Mostafaeipour, A., Goudarzi, N., and Jalilvand, M. Assessing different parameters estimation methods of weibull distribution to compute wind power density. *Energy Conversion and Management*, 108:322–335, 2016.

Rasmussen, C. E. and Williams, C. K. I. *Gaussian Processes for Machine Learning*. The MIT Press, 11 2005. ISBN 9780262256834. doi: https://doi.org/10.7551/mitpress/3206.001.0001.

Shi, H., Dong, Z., Xiao, N., and Huang, Q. Wind speed distributions used in wind energy assessment: A review. *Frontiers in Energy Research*, 9, 11 2021. doi: https://doi.org/10.3389/fenrg.2021.769920.

Tian, Q., Huang, G., Hu, K., and Niyogi, D. Observed and global climate model based changes in wind power potential over the northern hemisphere during 1979–2016. *Energy*, 167:1224–1235, 2019. ISSN 0360-5442. doi: https://doi.org/10.1016/j.energy.2018.11.027.

Wang, Z. and Liu, W. Wind energy potential assessment based on wind speed, its direction and power data. *Scientific reports*, 11(1):16879, 2021. doi: https://doi.org/10.21203/rs.3.rs-594128/v1.

Zeng, Z., Ziegler, A. D., Searchinger, T., Yang, L., Chen, A., Ju, K., Piao, S., Li, L. Z. X., Philippe, C., Chen, D., Liu, J., Azorin-Molina, C., Chappell, A., Medvigy, D., and Wood, E. F. A reversal in global terrestrial stilling and its implications for wind energy production. *Nature Climate Change*, 2019. doi: https://doi.org/10.1038/s41558-019-0622-6.