Time Series vs Statistical Approaches in Estimating Wind Turbines Energy Yield

Raphael Martins de Figueiredo Carvalho

Master Dissertation

FEUP Supervisor: Prof. Álvaro Henrique Rodrigues

U. PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Master in Mechanical Engineering

June/2023

Resumo

A utilização do vento como fonte renovável de energia tem sido uma constante nos últimos anos, razão essa associada às mudanças climáticas provocadas principalmente pela utilização de combustíveis fósseis. Essa fonte primária é capaz de ser aproveitada por turbinas eólicas com o objetivo de gerar eletricidade, pelo que a sua produção de energia considera o regime de vento do local de operação. Esse regime é caracterizado por meio de uma série temporal de vento, no qual contem dados de velocidade, direção, temperatura e pressão desse recurso ao longo de um período determinado de medição.

Para caracterizar a produção de energia e considerando um modelo definido de turbina, uma das abordagens mais tradicionais empregadas nessa avaliação é o de produção de energia anual (AEP). Para isso, as séries de dados de vento são aproximadas por parâmetros estatísticos e ajustamentos funcionais.

Entretanto, a utilização desse fator pode nem sempre ser o mais confiável para estimar a produção de energia da turbina. Isso deve ao fato de que os parâmetros atmosféricos e muitos outros são reduzidos a um valor médio anual, o que pode causar incertezas e variações bruscas, quando comparado com a produção de energia obtida diretamente das séries de vento.

Por isso, o objetivo deste trabalho é realizar um estudo da bondade do ajustamento de parâmetros estatísticos a séries a partir de dados das características do vento em dois casos de estudo, representando locais com características climáticas diferentes. Além disso, também será estudado possibilidades de análise de energia no período não anual, como por exemplo semestral ou noite/dia, a fim de descobrir se os ajustamentos nesses casos são melhores que no modelo anual.

Essa análise será conduzida em dois casos de estudo, a região da Polónia e a região Nordeste do Brasil (Rio Grande do Norte), para essas diferentes discretizações de tempo.

Time Series vs Statistical Approaches in Estimating Wind Turbines Energy Yield

Abstract

The use of wind as a renewable source of energy has been a constant in recent years, a reason associated with climate changes caused mainly by the use of fossil fuels. This primary source is capable of being harnessed by wind turbines to generate electricity, so its energy production considers the wind regime at the place of operation. This regime is characterized by means of a wind time series (WTS), which contains data on speed, direction, temperature, and pressure of this resource over a given period of measurement.

To characterize the energy production and considering a defined turbine model, one of the most traditional approaches used in this evaluation is the annual energy production (AEP). For this, the wind data series are approximated by statistical parameters and functional adjustments.

However, using this factor may not always be the most reliable way to estimate turbine energy production. This is due to the fact that the atmospheric parameters and many others are reduced to an annual average value, which can cause uncertainties and sudden variations, when compared with the energy production obtained directly from the wind series.

Therefore, the objective of this work is to carry out a study of the goodness of adjustment of statistical parameters to series from data on wind characteristics in two case studies, representing locations with different climatic characteristics. In addition, possibilities of energy analysis in the non-annual period, such as half-yearly or night/day, will also be studied in order to find out if the adjustments in these cases are better than in the annual model.

This analysis will be conducted in two case studies, the region of Poland and the northeastern region of Brazil (Rio Grande do Norte), for these different time discretizations.

Acknowledgements

First, I would like to express my sincere thanks to Professor Álvaro Rodrigues for his enthusiasm in the area of Wind Energy, managing to transmit his knowledge in a solid way, in addition to his availability, support and guidance throughout the work period.

I wish to express my appreciation to engineer José Carlos Matos and engineer Nuno Cardoso from Novenergia, who provided great support from the beginning, contributing with their vast knowledge in the area and making time available for doubts that appeared.

I would like to extend my thanks to my family, who were always with me at all times, in addition to the friends I made in Coimbra, for all the wonderful moments and also for being an extension of my family here.

I would also like to thank my girlfriend Iara who was my safe haven throughout the whole time, motivating me and giving me all the necessary support. Your altruism and love were essential in this journey.

I express my deep gratitude to all those who, in one way or another, contributed to the success of this project. Your dedication and support are greatly appreciated.

Contents

List of Figures

List of Tables

Nomenclature and abbreviations

AEP – Annual Energy Production

EPF – Energy Pattern Factor

EPFM – Energy Pattern Factor Method

EML – Empirical Method of Lysen

EYA – Energy Yield Assessment

EYTS – Energy Yield Time Series

MLM – Maximum Likelihood Method

MOM – Method of Moments

PPA – Power Purchase Agreement

WAsP - Wind Atlas Analysis and Application Program

WRA – Wind Resource Assessment

σ – Standard deviation

ρ – Density

k – Shape factor

A – Scale factor

1. Introduction

1.1. Background/Framework

The electricity represents one of the most current and important demands on the world stage, as it enables the implementation of projects and the development of the industry as a whole, in addition to being widely used in other sectors such as residential, transport in the case of electric vehicles, powering appliances for agricultural purposes and many other equipment with different purposes for people's daily situations. Throughout history, numerous energy sources have been used to obtain electricity, such as oil, gas, coal, water in the case of hydroelectric plants, among others. However, with climate changes occurring every year and their harmful consequences, it is necessary to resort to alternative energy sources that do not contribute to atmospheric pollution, namely the large-scale production of carbon dioxide in the atmosphere.

From this, it was possible to perceive that in the last decades, there was a movement of research and development of new sources of energy, resources that could be captured in a simple way and with capacity of electricity production that supplied the demands of the community, either it, local, national, or worldwide. Thus, new technologies for converting renewable energy sources into power, such as wind and solar energy, were developed. The first example operates in capturing wind by means of wind turbines, transforming kinetic energy into electricity and the second operates using photovoltaic panels converting sunlight energy into power.

An example of how renewable energy sources can replace fossil fuels is by obtaining hydrogen gas (green hydrogen), Figure 1, which can later be converted into electricity. Despite this, the technology for using hydrogen gas for electricity production purposes, for example, has not reached technological maturity, that is, whether a reliable and well-structured energy conversion system already exists. For this reason and as an alternative to fossil fuels, one of the most used forms of energy to generate power is wind, through the simple use of wind turbines. Although there have been prototypes of wind technology since the 19th century, it has only been 30 years since it reached maturity.

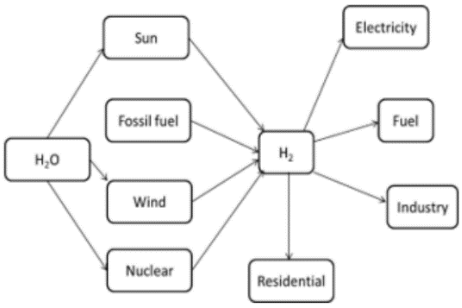

Figure 1: Primary energy sources for the production of hydrogen and its uses. (Shanmugaratnam et al., 2021)

However, for the implementation of wind technology, there are some factors that must be considered, such as the annual energy production (AEP) that is expected in a given location for a certain model of wind turbine. One of the analysis methods for this factor is based on a wind time series (WTS) at the location where a wind farm is to be installed, with information on speed, wind direction, air temperature, atmospheric pressure, and other parameters at time intervals of generally 10 minutes.

Besides the definition of the project conditions, such as the locations where the turbines will be positioned and their model, the expected total period of operation, the other technologies to be used and the size of the project in general, another process is the calculation of the AEP, the production estimate that can be revised in the face of different methodological approaches or new design information. As the operation of a wind farm involves intensive capital investment, the AEP calculation can serve as a basis for forecasting revenues and expected amortization, enabling wind farm investors to estimate the financial return on the investment.

Therefore, the calculation of the AEP is conducted from the wind series, in which the collected data are approximated by statistical adjustment models and their parameters. The best-known fits in this application are the Rayleigh distribution and the most used, the Weibull fitting, which uses shape and scale parameters. There are other parameters that are equally necessary to obtain production estimates, such as the shear factor and air density, which are treated on the basis of annual average values.

However, the statistical adjustments perform estimation predictions by simplifying parameters on an annual basis. So, the major concern of it is how close the estimated value of annual energy production by these models approximates and reproduces the AEP values if the calculation were done using the wind series directly. In this case, the values of the shear factor and density parameters would be considered every 10 minutes and not an annual approximation.

The analysis of energy production on an annual basis, though, is not always the most required by electricity consumer companies, that is, there are applications in which this annual period time is not always sufficient for customers who buy energy. When there is a Power Purchase Agreement (PPA), which is an energy purchase agreement between a power generation company and a consumer company, the second party does not always want an annual profile, but a monthly or daily discretization, as in figure 2. In this case, those who buy energy from the wind farm to sell electricity to consumers usually request production studies on a short time scale, in order to predict the quantity and quality of service that will be provided, in addition to making it easier to predict risks in cases of low production rather than relying on overestimates.

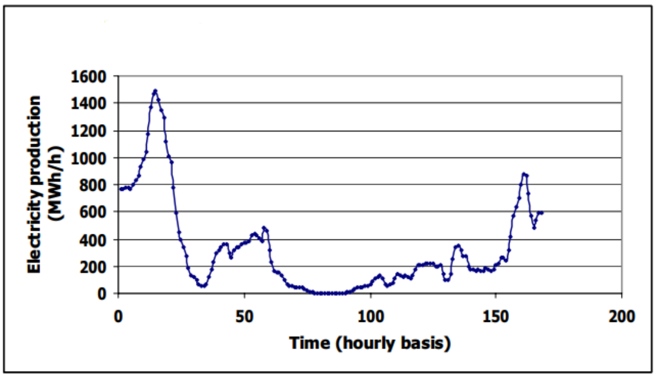

Figure 2: Example of Energy production profile during a 1-week period (JRC et al., 2006)

As production is continuous over time, there is also the case that the majority of production in AEP occurs outside of peak energy demand times, which is not advantageous for companies since electricity sales prices will not be the highest, decreasing their profitability. In SPOT markets, for example, the purchase and sale of electricity occurs in real time, with prices determined at the time of the transaction. Therefore, the demand for electricity is hourly, so the AEP calculation is neither considered nor accurate for this discretization, the previously mentioned statistical distribution hardly produces a good fit for this scenario.

An equally important case that is related to the optimization of energy systems that needs a different time discretization use is the hybridization of technologies, such as, for example, wind and solar together. In this, the energy forecast/profile is hourly to check the overlap of wind and sun in the simultaneous production of electricity, or on each season of the year (Figure 3) as there might be variations depending on the place. Furthermore, this discretization is useful in checking the potential need for energy storage when a technology does not produce efficiently, such as on days without sun or days without wind.

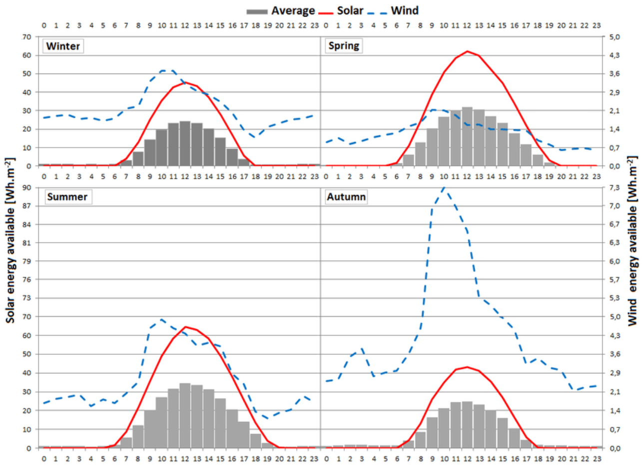

Figure 3: Daily averages of solar and wind energy availabilities between seasons (Siqueira et al., n.d.)

Therefore, before implementing a wind farm, it is necessary to carry out a wind resource analysis (WRA) in order to characterize potential energy production in a defined location. This analysis is done using wind data at a given location that was obtained through measurements of speed, direction and other parameters over a given time. The estimation of energy production, therefore, can be done using wind climate data series or using other approaches such as statistical models to perform the forecast.

For statistical approximations, therefore, one must consider whether the adjustment produced can obtain better results if used for shorter periods, either Winter/Summer, every semester or night and day. In addition, it is necessary to analyze whether depending on the method of obtaining the Weibull parameters, the level of accuracy of the energy production estimation calculations increases or decreases.

1.2. Objectives

This work will address the study of the goodness of adjustment of statistical parameters to series from data on wind characteristics in two case studies, representing locations with different climatic characteristics.

So, the objectives of this study are:

- Understand to what extent these annual adjustments allow reproducing the AEP values that would have been obtained directly using the wind series and the other necessary parameters.
- Analyze whether the data series with a discretization of 1 or more years is relevant for obtaining good results with the adjustments.
- Study the different discretizations in the WTS in order to consider seasonality phenomena and their impacts on the AEP estimate
- Understand if the adjustment of shear factor parameters in a greater discretization of the wind series can contribute to better results.
- Understand if the variation of the scale factors of the Weibull adjustment influences the quality of the results.

1.3. Structure

Chapter 2 gives an overview of how the wind resource is traditionally analyzed. Section 2.1 highlights the process of measuring a wind regime in a given location and the characteristics of the equipment used to do so. Section 2.2 presents the concept of Wind Time Series (WTS) and how the measured wind data is translated into a series. Section 2.3 already details the common process of characterizing this wind regime through a statistical adjustment. Section 2.4 defines other important parameters for the study of energy production.

Chapter 3 focuses on presenting the adjustment methods that are usually used to estimate energy production, in addition to criteria that allow qualifying the goodness of adjustment of such methods. These methods and criteria will be further employed in chapter 5.

Chapter 4 defines the case studies present in this work, as well as the design conditions of the reference turbine to be used. In it, the objectives are referenced again and also the methodology to be used to conduct the study of energy production in different discretizations. Section 4.2 defines the study environment of each case, the local characteristics, and base values of some important parameters for energy calculation.

Chapter 5 already refers to the results obtained in each case study. It will contain measured values, calculation of some parameters and analysis of each time approach used. For each case study, there will be a separation of the analysis according to the defined time discretizations.

Chapter 6 contains the conclusions and the main findings of the project. In addition, it proposes recommendations on future work related to this work.

2. Wind resource evaluation

The activities required to implement a wind farm must have the best cost-benefit ratio because wind technology is capital intensive, that means, it requires a significant investment in planning and construction before the operation of the wind farm. A wind energy project must be carried out with the least amount of uncertainty feasible at every stage, including the estimate of energy production, as mistakes made during the project implementation phase are hardly corrected, or even not corrected at all.

With the evolution of technologies that manage to take advantage of the wind energy source for conversion into other forms of energy, society's demands for electricity, for example, are increasingly being met. However, it is necessary to study the viability of energy production in a given location to analyze whether the wind farm can supply energy in different periods of time with a good cost benefit. The factors that directly impact wind energy and its price volatility are multiple and more complex (Hosius et al., 2023).

In order to have a standard approach in the assessment and evaluation of the wind resource to subsequently carry out estimates of energy production in a given location, some steps are conducted, as in figure 4.

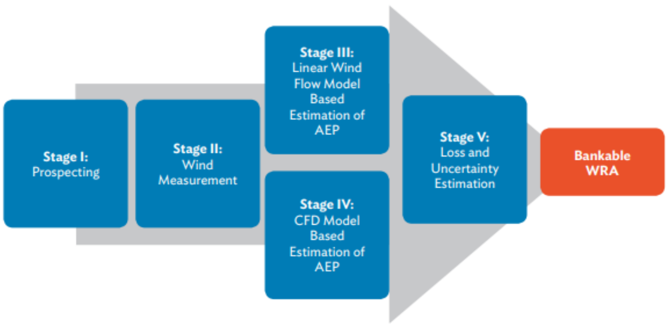

Figure 4: Wind resource assessment in five stages (ADB, 2014)

In this case and having the objective of this work presented, steps 2 and 3 will be analyzed. The AEP (annual energy production) is represented as an approximation approach to production statistics and for that, some considerations are made. After collecting wind data, this information is given in the form of a wind time series and traditionally used in a statistical distribution whose parameters are obtained through approximations.

Therefore, before carrying out any calculation of energy production, it is necessary to consider how the wind data are obtained, how they are related during the measured time period and also how the characterization of these data in a distribution is done.

2.1.Field measurements

The use of wind energy in a given location for the production of electricity requires a study on the viability in terms of production, that is, an analysis to understand whether in the chosen area of implementation of this technology the expected production is met. For this, it is essential to have wind data in place and the most effective method to capture this information is through a measurement campaign.

The measurement campaign can be done with some technologies, but it is usually carried out by installing anemometric towers that measure the speed and direction of the wind at different heights above the ground, as well as other relevant meteorological features such as air temperature, relative humidity, and atmospheric pressure. Installing a mast in a specific location to obtain wind information is one of the most traditional and reliable ways of measuring data (Shende et al., 2023).

As measurement campaigns aim to collect enough wind data at a given location in order to provide a characterization of the wind resource, the duration of measurement campaigns is usually over 1 year. This allows the dataset to be able to integrate the effects of seasonality into the wind regime.

The location of installation of equipment for the measurement campaign, traditionally masts, is based on ensuring that the position of the tower has the least possible interference from other obstacles that may influence the measured wind parameters. In addition, the tower must be located in a region not too far from the location where the wind turbines are to be positioned, so that the collected data will reliably portray the wind regime at the installation location of the turbines.

Height is an important variable to consider as wind speed varies with height above the ground and therefore the measuring station consists of anemometers for three components of wind direction, in addition to wind sensors positioned at different elevations. The type of anemometer to be chosen is related to the operation requirements during the measurement and also the cost associated with each one. The sonic anemometer (Figure 5a), for example, does not need recalibration or frequent maintenance but has a high cost and high energy consumption. The cup anemometer (Figure 5b), in turn, is widely used in measuring stations due to its low cost and its simplicity of use and handling.

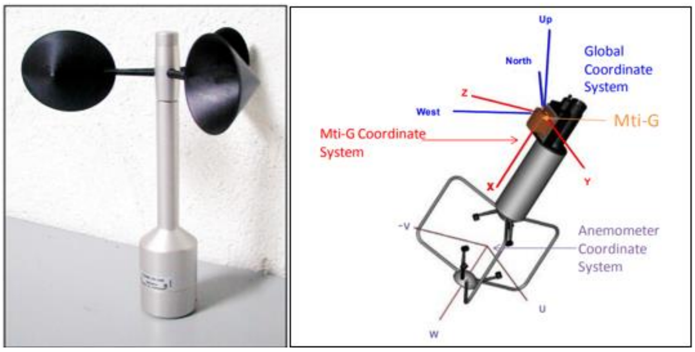

Figure 5: Thies Clima 4.3350 cup anemometer (Roibas-Millan et al., 2017) (a) and Sonic Anemometer/MTi-G schematic (Stevens et al., 2013) (b)

The definition of the height points at which the anemometers will be installed on the mast are based on the wind turbine model previously chosen to be installed in the wind farm. Ideally, the wind measurement heights should be close to the height of the turbine rotor axis, in which some literature indicates that measurements should be taken at least 2/3 of the height of the rotor axis. Measurement heights close to the planned hub height will reduce the vertical extrapolation uncertainties of the wind conditions (Measnet_SiteAssessment_V2.0.Pdf, n.d.).

In addition to wind speeds, another parameter collected by the measuring tower is the direction that is obtained by a piece of equipment called a wind vane. It is composed of a rotating structure, usually in the form of arrows or propellers, which aligns with the direction of the wind flow. It is noted that sonic anemometers can also provide wind direction and even flow inclination.

In the measuring towers there are also temperature and atmospheric pressure sensors. An example of the measuring mast and its components can be seen in figure 6.

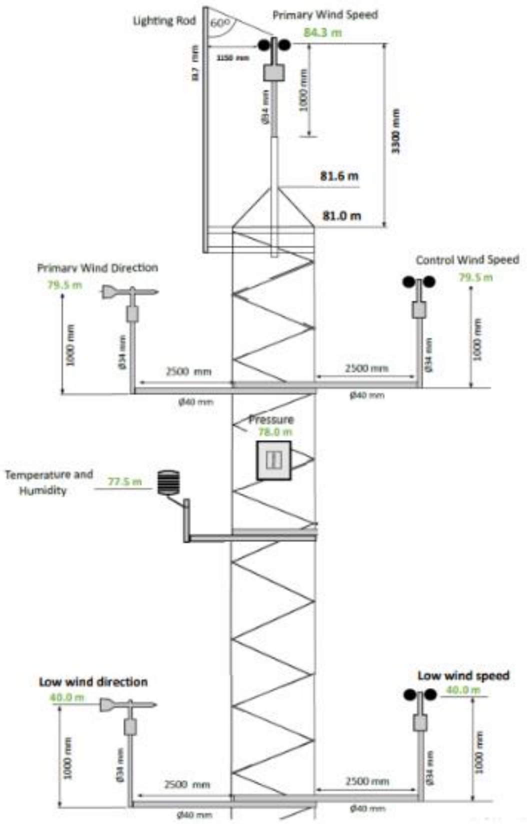

Figure 6: Example of a mast scheme (GENERG SGPS)

Despite this, turbines are getting bigger, both in terms of the diameter of the blades and the height of the rotor axis, so the anemometers and the towers must also be bigger to capture the wind values in a more representative way at a given height of the wind turbine. rotor shaft. This can mean higher costs for building more extensive towers and, consequently, an increase in project risk. The need for more flexible methods of monitoring wind is therefore clear (Lang & Mckeogh, 2011)

Therefore, new forms of technologies were designed for these cases that also have easy mobility on the ground, when compared to measuring towers, LIDAR (Light Detection and Ranging) and SODAR (Sound Detection and Ranging).

The first method consists of a remote sensing technology equipment that, through monochromatic laser rays with a certain frequency, reach the air particles and through backscattering, the reflection of light has a shifted frequency, which allows the estimation of the speed of the wind (Figure 7).

Figure 7: Lidar system in an offshore situation

The second method consists of a principle similar to LIDAR but instead of electromagnetic waves, it uses sound waves that propagate through the air and find variations in wind speed and direction at different altitudes to then be reflected towards the SODAR equipment. (Figure 8)

Figure 8: Sodar system (Triton)

One of the advantages of a remote measurement technology is the low environmental impact, as the machine used does not require large buildings, in addition to the good behavior of the sensors in conditions of intense cold, in which conventional anemometers can freeze and cause data loss, in information gaps. Operation difficulties related to large amount of data or power supply may, however, occur and in foggy days for example, LIDAR measurements are usually poor.

Despite the existence of these different remote measurement technologies, for the work carried out and considering a medium to long-term measurement campaign, the mast measurement tower is the most suitable.

It is important to mention that for the development of exploitation projects, methods have emerged to try to create a wind behavior database in order to be used before installing any measuring tower. An example of this that is widely used is the global wind ATLAS, which represents a tool that provides wind resource data and helps to understand the wind resource potential of each location, as in the example in figure 9.

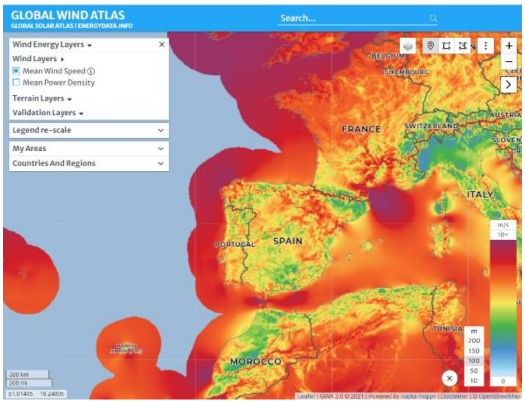

Figure 9: Global Wind Atlas interface

2.2. The wind Time Series

For the implementation of a wind farm, it is not enough just to know the place where it is intended to generate electricity, it is necessary to conduct studies to understand if there is a financial, logistical, and energetic viability. The estimation of energy production is not simple, it is necessary to characterize it in a given time, considering factors of orography, land cover, the presented obstacles and, above all, the characterization of velocities throughout time.

After carrying out the measurement campaigns in a given period, the collected wind data are stored in a format, in which they are preserved without any value modification, that is, no correction is made to the data considered with suspected error. The set of these data collected and transposed in a temporal scale represents a wind time series. A time series analysis draws on regular observations over a defined time interval for a particular phenomenon (Glass et al., 2009)

Regarding the analysis of wind time series, it is conducted from a database of wind speeds every 10 mins interval, in which the value defined in this range is actually an average of 300 values. So, every 2 seconds a value is registered by the data logger. Therefore, the use of time series resorts to these data in such a way as to predict the energy production of a wind farm.

The variables collected and presented in the wind time series are:

- Average, maximum, minimum, and standard deviation of wind speed, [m/s], for each anemometer.

- Mean and standard deviation of wind direction relative to true (geographic) north

- Average air temperature, [ºC].

- Mean atmospheric pressure, [hPa].

From this, the calculation of energy production estimates, in a specific location, uses a fundamental tool for analyzing the feasibility of implementing a wind farm, the wind time series (WTS). The series, thus are a collection of wind parameter data measured at a specific location and at a specific time interval via masts, as previously described. The majority of wind measurements are performed through the use of simple mechanical devices, as the traditional cup anemometer (Mortensen, 1994). Based on this information, engineers can predict how much power can be generated in a given area. An example of a wind time series distribution can be seen in Figure 10 below.

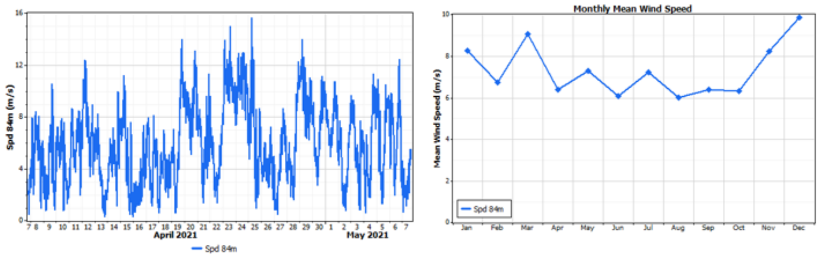

Figure 10: Example of Wind Time Series

In addition to temporal discretization, wind energy time series can also have variable durations. While some series may last only a year, others may go on for many years, for example 5 years or a decade. The amount of time deemed pertinent for the analysis in question determines how long the time series will last.

The advantages of using the time series are based on the fact that it considers the dynamics of time during the entire measurement period and calculates how the wind speed over time can be standardized to make accurate estimates of energy yield. In addition, this approach is not restricted to one type of time series, it is generally used in numerous regions with energy potential that have different types of climates, which also allows considering the seasonality and variation of the data.

Although the wind series contain important data on the characterization of the wind regime at a given place, it cannot always be admitted that this database is completely reliable and representative, as there is intermittency and possible random variations of air masses, temperature, and pressure values throughout time. In order to get a more precise assessment of the climatic conditions in a given area, wind reanalysis series are also utilized, which corresponds to a modeling technique that integrates observational data, such wind series, with atmospheric modeling data. The atmospheric reanalysis series are a temporal and geographic description of the climate, produced through models with observations, containing estimates of quantities for several meteorological parameters (Matos, 2022).

Reanalysis series are used to extend the temporal representativeness of short-period series. In this case, methods such as MCP (Measure-Correlate-Predict) are used to estimate energy production in a location where a short measurement campaign has been carried out. Otherwise, if there was already a mast station measuring wind data for 10 years, the reanalysis series would not be needed at all.

Data accounting before carrying out any AEP calculation process must consider the existence of information failure in some intervals due to sensor freezing or any other interference in the measuring station. In longer periods of measurement of wind data, it is more likely that there is large intervals with error of measured speeds, that is, the existence of blocks of data without information can be found in the series more frequently.

According to general literature, the availability or recovery rate of the series can be represented as the proportion between the effectively obtained data that contains non-error information in relation to the total amount of data, including any type of information generated, including errors. Some common error messages appearing in wind time series may have characters like "999", “NaN” or blank spaces.

As the objective of the entire implementation of a wind farm is based on reducing calculation inaccuracies, mainly energy, a common activity to deal with the existence of gaps in values in the series is to perform data collation. However, the collation is only justified if there is indeed a large number of data failure blocks and if the energy production calculation is not done in a statistical approach, in which small failures are not relevant.

One of the approaches used to fill gaps in data blocks is to perform a pattern analysis of velocity information and conduct a linear interpolation between the values. For example, if there is a block of data failure for an entire day, the population of this data will be based on the values of the previous day and the next day.

If the missing data is not filled in and also for the calculation of the AEP done directly from the series, there is a correction of the energy production, based on the present availability. Equation [1] is used to obtain the availability-corrected AEP.

$$\boldsymbol{AEP_{corrected} = AEP_{previous}/availability} \qquad [1]$$

In which the availability can be calculated from the equation [2].

$$\boldsymbol{availability = \frac{Number\ of\ useful\ data\ points}{Number\ of\ total\ data\ points}} \qquad [2]$$

Therefore, the determination of wind series represents a first step and a fundamental role in the analysis of energy production. From that moment on, new steps will be taken to better forecast the site's annual energy production (AEP). The common methodology of these steps can be represented in Figure 11 below, in which after obtaining wind measurements over time, extrapolations of the data will be carried out at a given turbine height to then combine the distribution of wind speeds with a power curve of a turbine model and forecast the energy produced on the site.

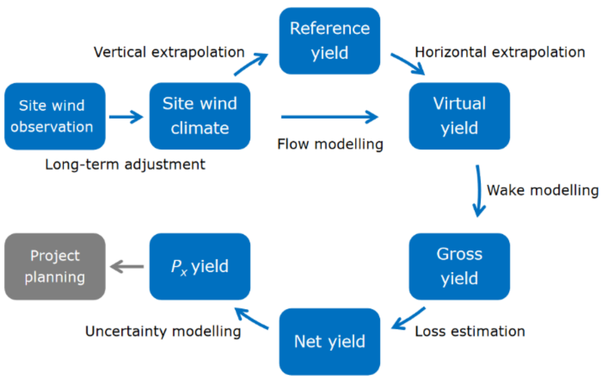

Figure 11: Energy Yield prediction process

2.3. Characterization of wind resource

Bearing in mind that the wind resource is intermittent, that is, its availability varies according to local meteorological and climatic conditions, it is necessary to understand the distribution of wind in a given time in order to estimate the energy production closest to the reality. To understand this availability of wind energy over time, it is necessary to analyze the distribution of speeds, at daily, weekly, and monthly intervals (Jani et al., 2023).

One of the characterizations of the wind regime defined by the measurement data in a given is the wind occurrence rose. This representation is made from the count of occurrences of wind in a given range of direction and its subsequent normalization to make the polar graph. There is usually a division of the 360º into 16 equal intervals of 22.5º, but this is not always done in this way, it might depend on the project requirements. The wind rose allows, for example, to analyze whether the wind turbine operating region has a more prevalent direction for the angle of attack.

From the wind data contained in the time series, the set of speed information is generally characterized by histogram, in order to characterize the regime through a representation that allows describing the behavior of the measured velocities, respectively, regarding their central tendency, shape and dispersion.

So, the construction of this representation is done by counting the number of occurrences of velocities that are in the same bin, that is, each velocity range will have the same value "length", such as 1 m/s, 0.5 m/s or another and the "height" of the columns in the histogram will be the frequency of the velocities that are within each range. An example of the rose of occurrences and a histogram is shown in figure 12.

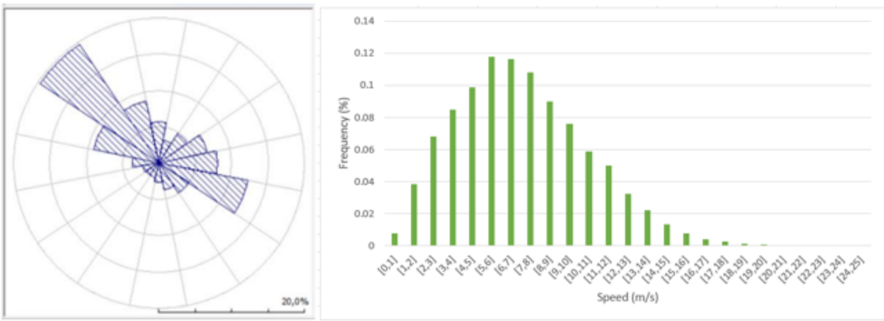

Figure 12: Rose of occurrences and histogram

The estimate of energy production is then directly associated with the discretization of the speed bins, that is, its occurrence distribution profile. Regarding the approaches for characterizing the frequency distribution and considering the objective of this study, two options can be reported, the first being to use only the time series and obtain the frequencies manually, without any statistical interference. The second approach is to use statistical methods to try to predict the speed occurrence frequencies from a measurement database.

Wind is a random variable and is subject to temporal and spatial fluctuations. The use of a statistical distribution allows modeling these fluctuations and providing an adequate mathematical representation of the wind behavior, as from a probability density function.

According to the literature, one of the distributions most used by researchers is the Weibull distribution, for the purpose of optimizing the algorithms that determine the parameters of the statistical model (Al-Quraan et al., 2023). Therefore, it is important to analyze how close the

Weibull distribution is to the real velocity distribution, that is, to the time series that does not have any statistical approach. Each time the Weibull method becomes closer to the real distribution, that is, more adjusted, the estimate of energy production is made with greater precision.

As a Weibull distribution fitting is therefore the statistical method most commonly used in wind energy studies to obtain a smooth distribution (Dayal et al., 2021), it is necessary to understand its characterization. Its distribution is the generalization of the Rayleigh model, which contains a scale factor, *A*, and a shape factor, *k*, and therefore the Weibull method is used when the Rayleigh distribution fails to obtain a sufficiently adjusted approximation to reality. The Weibull distribution function [2] and the cumulative distribution function (CDF) [3] are expressed according to these two parameters.

$$\begin{cases} \boldsymbol{f(u) = \frac{k}{A}\left(\frac{u}{A}\right)^{k-1} e^{-(u/A)^k}} \\ \boldsymbol{F(u) = CDF(u) = 1 - e^{-(u/A)^k}} \end{cases} \qquad [2,3]$$

The cumulative distribution function of the velocity u gives the fraction of time (or probability) that the wind velocity is equal or lower than u, while in equation 1, f(u) represents the frequency of occurrence of velocity u. An example of a Weibull distribution adjustment can be seen in Figure 13, in which, from an example of series of wind data at a location in Portugal, and with the WAsP program, it was possible to calculate the model parameters (*A*,*k*) and obtain a fit to the distribution of wind speeds.

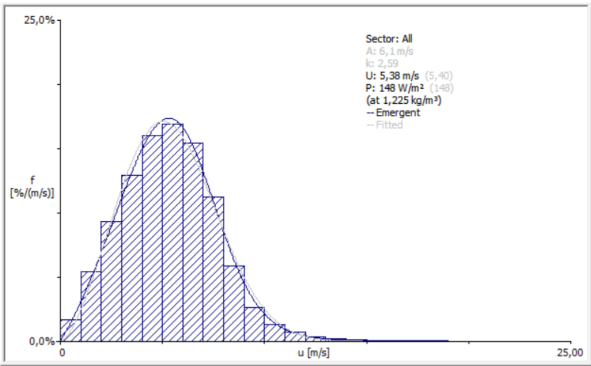

Figure 13: Example of Weibull adjustment

The WAsP program is widely used in the field of wind energy because, based on meteorological and topographic data from a given region, together with the combination of turbine models distributed in that area, it is possible to estimate the annual energy production. In this case above, a histogram is represented with the current distribution of occurrence of velocities and a Weibull model adjustment made in the software. As the approximation, in this case, follows quite well the frequency variations of the velocity intervals, it can be said that it has a good fit.

According to Hennessey (1977) and Justus et al. (1978), the advantages of using of the Weibull distribution to adjust the measured wind data are:

- It is a function with only two parameters to be determined, thus having a simple application
- Knowing the parameters, *A* and *k* at a given height, one can directly extrapolate the distribution to other heights

- Provides a good representation of the asymmetry of the wind distribution.

The values of parameters *A* and *k* are related to the measurement height in relation to the surface and therefore, from different heights, the distribution can widen or not, resulting in lower or higher maximum probability values. If there are obstacles or the terrain has a slope, it is possible to see the presence of the speed-up effect in which these barrier elements cause the wind air masses to be compressed and have a velocity acceleration.

The scale parameter (*A*) indicates the wind speed at which the probability of occurrence of it is greatest, expressed in meters per second (m/s), and the shape parameter indicates the dispersion in the probability distribution. As the *k* parameter increases, the Weibull fit is also modified, indicating that there is more occurrence of stronger and weaker winds, as in figure 14.

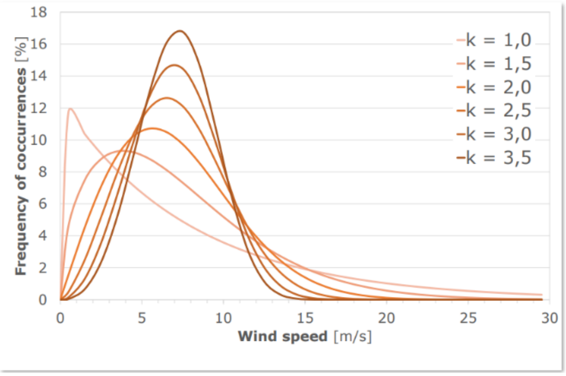

Figure 14: Example of variation of k and different fittings (Rodrigues, 2022)

When *k* is equal to 2, the function assumes a Rayleigh distribution, while when it assumes a value of 3.5, the function assumes a normal distribution, close to a Gaussian distribution. For the case of the Rayleigh function, the distribution is represented by equation [4].

$$\boldsymbol{f_r} = \frac{\boldsymbol{\pi}}{\mathbf{2}}\frac{\boldsymbol{u}}{\overline{\boldsymbol{V}}^{\mathbf{2}}}\boldsymbol{e}^{-\frac{\pi}{4}\left(\frac{u}{\overline{V}}\right)^2} \qquad [4]$$

In which $\overline{V}$ is the average speed of the series.

However, very high or very low Weibull *k* values do not always indicate a good fit, as this factor depends on the climate of the location where the measurement campaign was conducted. In monsoon weather, for example, a single Weibull fit is not enough to represent the wind regime, so a combination of adjustments is needed as in figure 15.

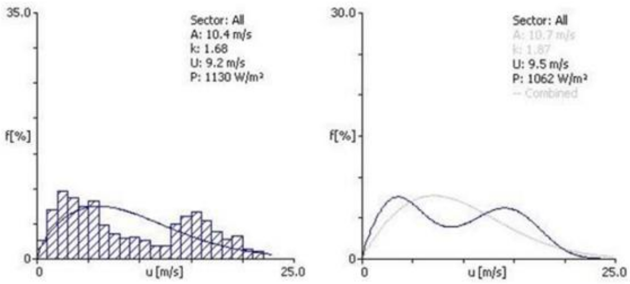

Figure 15: Before and after Combined Weibull fitting (Rodrigues, 2022)

Another parameter that can be obtained using the Weibull parameters is the average speed of the data series, according to equation [5].

$$\bar{v} = A\Gamma\left(1 + \frac{1}{k}\right) \quad [5]$$

where Γ is the gamma function.

Therefore, whenever the frequency of occurrence for each speed *u* is obtained and by relating it to the power curve distribution of a certain wind turbine, the gross annual energy production (AEP) can be calculated in a more general way using the following equation [6]:

$$AEP_{gross} = \sum[f(u) * P(u)] * 8760 \quad [6]$$

Where f(u) is the frequency of occurrence of wind speed [%] and *P*(u) is the power produced by the wind turbine at wind speed [kW]. The 8760 factor represents the number of hours the wind blows at that speed *u* in a year (365 days x 24 hours a day). The multiplication of *f(u)* with 8760 can be denominated as a parameter H(u) that represents the total number of hours per year in which the wind blows at such speed.

It is important to reiterate that for this equation, considering the use of statistical approximations, traditionally the Weibull adjustment, the portrayed velocity is not the specific value, but the representative value of each velocity interval.

Statistical approximations, when used to calculate energy production, adjust the wind regime, and are associated with the power curve of the wind turbine model chosen by the project.

The power curve of a wind turbine, in turn, represents the relationship between wind speed and the electrical power generated by the turbine. It represents a fundamental step in the calculation of energy production, as it is integrated with the wind speed values obtained from the wind time series.

The operation of a wind turbine can be characterized by some parameters, as in the example in Figure 16. The turbine starts its operation at a cut-in speed, which is the minimum wind speed necessary for the wind turbine to start generating power and thereafter, the power output increases with the wind speed in a curve of third power, cubic. The curve grows until reaching a point of rated speed and rated power, in which the turbine is operating at its maximum efficiency point. The rotor is stopped when the wind speed exceeds the cut-out speed, in order to prevent damage (Dupont et al., 2017).

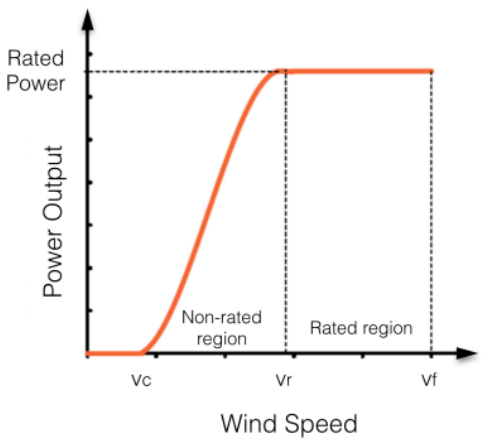

Figure 16: Typical wind turbine power curve (Dupont et al., 2017)

The characterization of the wind regime, therefore, by statistical approximations, namely the most common being the Weibull adjustment, is the first and most traditional step to perform the calculation of AEP. These and other parameters are adjusted in an annual representation, that is, when calculating the energy production by using the statistical adjustment, an annual average of the density and the shear factor is made.

The energy market does not establish a single way to analyze the production of energy and later the sale of electricity in final consumption may not require energy values in the annual period. An example of this is the case of the wind farm having a Power Purchase Agreement (PPA) between the energy producer and the buyer, which establishes terms and conditions for the sale of renewable energy over a specific period of time. PPAs can be seen as a hedging tool by the market, as they offer an opportunity for energy buyers to achieve price certainty beyond 3–5 years, and at the same time meet their sustainability objectives (Jiménez et al., 2023)

This exposes some questions, such as whether extending the series for a period of 1 year provides the presentation of good adjustment results and, furthermore, if changing the discretization, the adjustment is more reliable, produces better results.

An example of the different requirements of an analysis of energy production, that is, of an energy representation made for a non-annual period is in the SPOT markets. This type of financial market trades assets instantly, with service delivery, for example, electricity almost immediately. In this case, it makes no sense to present an energy analysis conducted for a year, as the energy buyer has the demand in a shorter period of time, such as month, week or even daily, as in the example of figure 17.

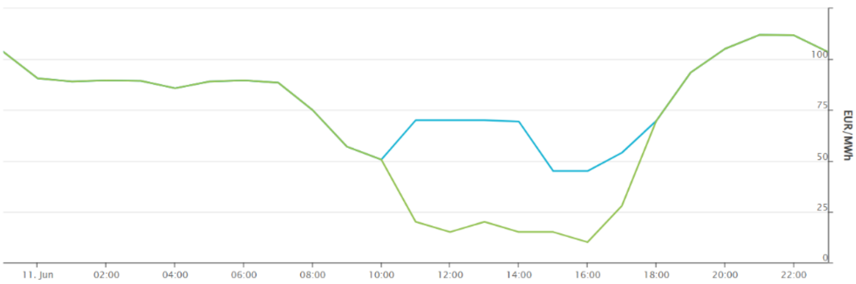

Figure 17: Example of daily market price -SPOT market (*SPOT Hoje | OMIE*, n.d.)

Another example in which the AEP calculation is not required is seen in the hybridization models of technologies that take advantage of solar and wind energy, for example. In this case, there is the operation of wind turbines with photovoltaic panels and, as the energy sources are different, there may be periods of low production in one or the other, or in both. Therefore, an energy analysis is necessary for daily and weekly demands, as in the case of figure 18.

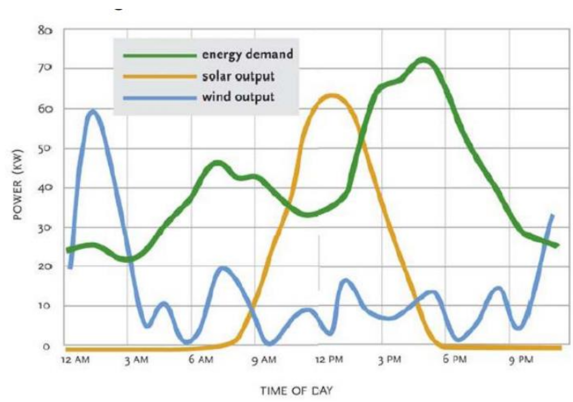

Figure 18: Wind power and solar energy generation curves compared with power demand of grid (Edrisian et al., 2013)

2.4. Other relevant parameters for estimating production

Although the series of wind measurements already provide parameters such as wind speed, direction, temperature, and pressure at certain time intervals, it is necessary to analyze some more parameters to estimate electricity production. These parameters, therefore, are not obtained directly from the data series but from relations between the parameters already obtained in them.

Since current wind turbines have high rotor shaft heights and many times greater than the maximum height of the measuring station, the speed values need to be extrapolated in order to conduct the AEP estimate in a more precise way. Therefore, a new parameter called the shear factor is obtained to be able to relate the velocities at different heights.

A second parameter of interest in the study of energy production refers to the fact that for each time interval and a given speed, there is a corresponding air density value, so the variation of it must be considered for obtaining a specific power at each value of this parameter.

In addition, it is also necessary to consider that during the period of operation of wind turbines, they are subject to turbulence variations, together with extreme winds.

2.4.1. Shear factor

The choice of a turbine model requires that the speed conditions obtained in a Wind Time Series (WTS) be recalculated for the height of the hub rotor shaft and with that, the distribution of speed occurrences is also modified, along with the estimate of energy production, either with methods statistical or not. Vertical variation of wind speed is an important parameter for wind turbine design, especially for those with large diameter rotors (Lopez-Villalobos et al., 2022).

According to Custódio (2007), it is possible to establish the wind speed at different heights through a logarithmic relationship, defined by equation [7].

$$\frac{v_1}{v_2} = \frac{ln\left(\frac{h_1}{z_0}\right)}{ln\left(\frac{h_2}{z_0}\right)} \quad [7]$$

In which V_{h1} and V_{h2} are the velocities of the respective heights h_1 and h_2, z_0, the roughness length in place.

In order to be able to perform a speed extrapolation at a height different from the station's measurement elevations, at a power law model, a parameter called the shear factor (α) is used. This parameter is determined by the relationship between the wind speed at different heights, and is generally expressed as an exponential function, expressed in equation [8].

$$\alpha = \frac{\ln\left(V_{h1}/V_{h2}\right)}{\ln\left(h_1/h_2\right)} \quad [8]$$

The velocity profile according to the height change can be represented by the power law as in figure 19.

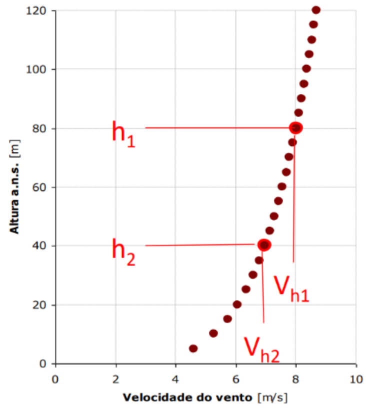

Figure 19: Example of velocity profile (Rodrigues, 2022)

It is important to consider that the calculation of the shear factor is not calculated at a random time, without any reference. The IEC 61400 standard requires that the shear factor exponent be calculated over the rotor -swept area, from lower blade tip height, and for each turbine position (Figure 20).

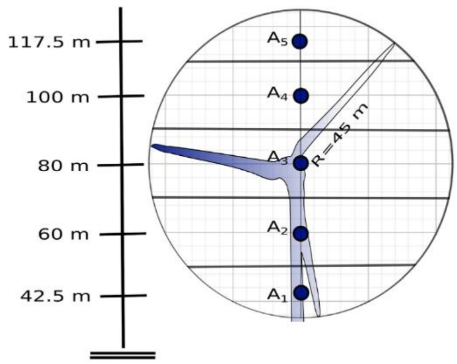

Figure 20: Illustration of determination of equivalent rotor wind speed over the rotor swept area divided into four segments corresponding to the four heights measurements. (Lopez-Villalobos et al., 2022)

Wind shear factor also depends on atmospheric stability, which could change, e.g., as the surface warms, and so the power law with a constant shear exponent becomes even less accurate for determining turbine-height winds (Badger et al., 2016).

An example of the change in shear factor values according to weather conditions can be seen in figure 21, in which for different months, the average values change.

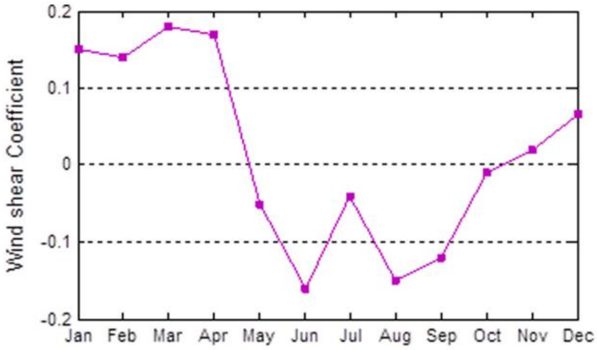

Figure 21: Monthly variation of the wind shear coefficient (Abbes & Belhadj, 2014)

In addition to this difference in a longer period, in months, the shear factor variation can also be seen in a daily period, as in figure 22, in which extrapolations were made from 65 m to 120 m.

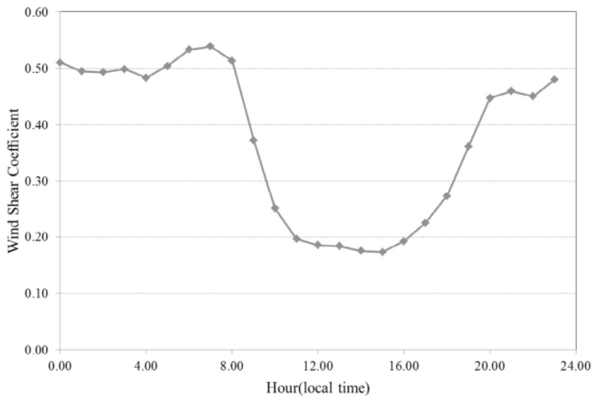

Figure 22:Diurnal variation of wind shear coefficient for the elevations 65 and 120 m agl at the Phangan station (Werapun et al., 2017)

2.4.2. Air Density

The estimation of energy production also considers another parameter that is not obtained directly from the measuring stations but whose value is obtained from the time series and which, together with the power curve of the wind turbine, indicates the respective power value at a certain speed. This parameter is the air density.

Air density can vary significantly from factors such as air temperature, atmospheric pressure, and relative humidity, and therefore, its correct estimation ensures that errors in production estimation are also minimized. Note that air density is likewise a function of temperature, humidity, and especially atmospheric pressure.(Guerrero-Villar et al., 2019) This density relationship with these other parameters that are obtained in the time series is described according to perfect gas equation [9].

$$\boldsymbol{\rho = \frac{p}{RT}} \qquad [9]$$

In which, p is the air pressure in the given period of time, R is the universal gas constant of 287 J/Kg. K and T, the absolute temperature.

Due to the fact that atmospheric conditions determine the value of air density, it is possible to perceive the variation of this parameter over time and, above all, to try to analyze whether there is a pattern of variation between years. Figure 23 shows an example of density variation on a monthly basis, in a 20-year database.

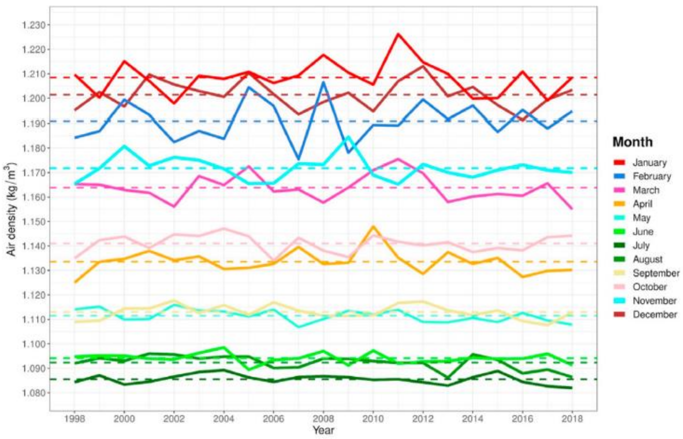

Figure 23: Monthly variation of air density based on the observational data for 1998-2018.(Liang et al., 2022)

Air density affects the power available in the wind in a linear way, as the power available in the wind that will pass through the wind turbine is defined by equation [10]. This power is the amount of energy in the air that can be generated from wind. This power is determined by the air density, the swept area A of the turbine blades, and the wind speed V.

$$\boldsymbol{P = \frac{1}{2}\rho A V^3} \qquad [10]$$

It should be noted that this is a simplified formula and that it ignores a number of important elements, like the height of the wind turbine, the wind profile, and losses from aerodynamic resistance, among others. More intricate models that take these variables into account are employed in actual projects to estimate wind power more precisely.

2.4.3. Turbulence intensity

The operation of every wind turbine requires a direct interaction with the layers of air in the atmosphere that meet the turbine blades, causing the rotation of the rotors to produce energy, that is, a conversion of kinetic energy into mechanical power and then electrical power.

However, the movement of the air layers does not present a homogeneous model of characterization, that is, there is turbulence and complex and chaotic wind variations over time, either by a change in temperature and pressure or even by wake effects from the turbines. The wake effect refers to the phenomenon caused by the passage of air in a wind turbine, creating an altered wind pattern in its wake, that is, in the region immediately behind the turbine, which may result in turbulence. These effects cause downstream turbines in wind farms to extract less energy from the wind and also can cause structural issues (Öztürk et al., 2023). Figure 24 shows a schematic of the regions before and after the wake effect.

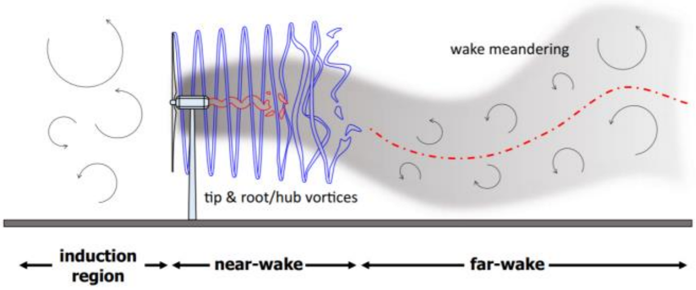

Figure 24: Interaction of a wind turbine with the air flow (Porté-Agel et al., 2020)

However, it is important to reiterate that there is no significant conversion of energy contained in the turbulence scales, as they are of small scales and therefore absorbed by the structure of the blades, and not very useful from the point of view of energy conversion.

The turbulence regime presupposes the existence of a time and space scale, in which the velocity parameter is defined by the sum of two components, average value plus a fluctuation. Thus, the turbulence intensity, $Iturb$, can be seen as a parameter that describes the variability of wind speed over time caused by turbulent fluctuations and can be obtained according to equation [11]. This relationship describes this parameter when used for industrial purposes, but in the broadest meaning, the turbulence intensity is the square root of the mean value of the velocity fluctuation component.

$$\boldsymbol{I_{turb}} = \frac{\boldsymbol{\sigma}}{\overline{\overline{\boldsymbol{V}}}} \qquad [11]$$

In which σ is the standard deviation of the wind speed and $\bar{V}$, the average speed measured over a 10-minute interval in the series.

This parameter is considered, for example, when choosing wind turbines, in order to ensure that the equipment, especially the blades, are sized to withstand fluctuations in speed and the respective load requests. The derived turbulence intensity is calculated for the measurement positions at measurement height and for the wind turbine positions at hub height. (*Measnet_SiteAssessment_V2.0.Pdf*, n.d.) Therefore, in places with a large variation in speed or with a high value of turbulence intensity, it may be necessary to implement more resistant turbines in order to guarantee a good useful life.

2.4.4. Extreme Winds

Another parameter that is also related to the variability of speeds caused by different pairs of temperature and pressure in the atmosphere is the extreme wind, which can be described as maximum velocity value (10 minutes average), expected in 50 years return period.

Extreme wind conditions are essential for evaluating the structural robustness of wind turbines. It is crucial to build wind turbines that can withstand these challenging conditions and maintain their structural integrity because during extreme wind events, wind turbines are subject to increased loads and mechanical efforts.

However, the calculation of extreme wind based on a wind time series is not standardized, as it is a prediction of a maximum value of a magnitude for a long period of time. The wind data are represented by a period of short duration when compared to the 50 years of wind analysis, so it is necessary to consider a model adapted to the characteristics of wind speed.

For the calculation of this parameter, it was decided to use the Gumbel model, in which this distribution is characterized by two parameters, the location (α) and the scale (β), and it is possible to estimate the probability of occurrence of extreme winds above a certain limit. The Gumbel distribution has been used as an extreme value distribution in many research areas, including wind energy (D. Kang et al., 2015). The probability function that a velocity X exceeds x is given by equation [12].

$$\boldsymbol{P(X > x) = 1 - exp\left\{- \mathbf{exp}\left[-\left(\frac{x-\alpha}{\beta}\right)\right]\right\}} \qquad [12]$$

In which x is the maximum velocity in 50 years. The location parameter and scale parameter values can be obtained using relations [13-14].

$$\boldsymbol{\beta = \frac{\sigma\sqrt{6}}{\pi}} \qquad [13]$$

$$\boldsymbol{\alpha = V_{max} - 0,577 * \beta} \qquad [14]$$

In which σ is the standard deviation and V_{max}, the maximum velocity obtained in the WTS.

3. Statistical Approaches for Wind Regime Analysis

3.1. Methods for estimating Weibull parameters

The Weibull distribution represents a model with good flexibility to accommodate a wide variety of distribution shapes via the scale factor and shape factor parameter pair. Technical evaluations usually use the Weibull distribution to identify the optimal wind turbine technology in a given location, in addition to being able to obtain data such as wind power density, annual energy output, and capacity factor (Alrashidi, 2023).

In addition to this, the distribution is easy to interpret because the shape factor, for example, indicates the probability of a certain wind speed occurring, which can be centered in a range of speeds or shifted depending on the local climate. The scale factor, on the other hand, indicates the wind speed at which the probability of occurrence of it is greatest, which already allows forecasting a possible range of more frequent speeds during the operation of a wind turbine.

The definition of the Weibull parameters allows a discretization of the frequency and later, with the power curve, an estimation of energy production, being therefore necessary to obtain the best estimate of the pair of parameters and to reduce the errors in the estimation.

The advantages of using a statistical approach to predict the energy yield of a wind farm are based on the fact that there is no need to have a very large and complex amount of data, because in software such as WAsP and others, the AEP calculation using directly the series may take some time. Thus, the statistical models can adapt to the amount of information available, that is, the calculation of parameters will be done independently and will consider the entire data set.

However, a major disadvantage of using statistical approaches is the simplification of wind data, which can influence the calculation of the AEP, that is, depending on the behavior of the data, the Weibull distribution may not be well adjusted to the wind histogram as in figure 25. This can lead to poorly made estimates of the AEP, that is, cause a greatly altered overestimate or underestimation, far from the closest value to the real that is the AEP with the wind time series.

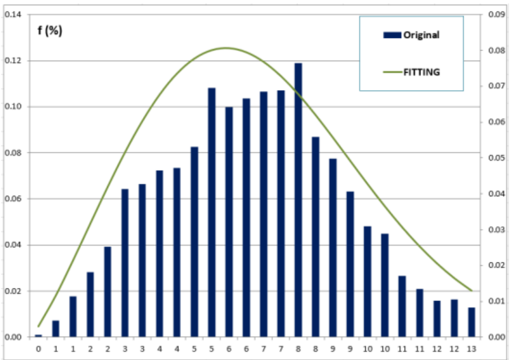

Figure 25: Example of a poor Weibull fitting

In this case, it seems there is a clear case of poor fit, but it is necessary to have more robust criteria than simple appearance in order to support these conclusions, criteria addressed in the next section 3.2.

Therefore, for this work analysis, it was decided to employ five models for obtaining the Weibull parameters, Maximum Likelihood Method (MLM), Energy Pattern Factor Method (EPFM), the WAsP algorithm, Method of Moments (MOM) and the Empirical Method of Lysen (EML). Weibull parameters estimate the characteristics for a specific location and therefore need to be calculated and compared with different methods than the adoption of a single statistical model (Yaniktepe et al., 2023).

It is important to point out that as the amount of data extends over time, the calculation of the parameters becomes more and more accurate. The ideal is to have a relatively long period that allows to make an overview of all the data and from that, there is also the possibility of discretizing the time series and calculating the parameters on an annual basis or during defined periods such as Winter and Summer.

The calculations of these parameters for estimating energy production must be done after extrapolating the results to the height of the chosen wind turbine, otherwise the accuracy is lower.

3.1.1. Maximum Likelihood Method (MLM)

The statistical model of the Maximum Likelihood Method uses time-series wind data for calculating Weibull parameters and consists of estimating those values that best fit the time-series data, based on the likelihood of information. Thus, the distribution will be more adjusted to reality so that later it can combine the information with the power curve of a wind turbine.

The statistical approach of MLM is based on carrying out iterative processes in which an initial value for the shape factor *k* is first assumed and after finding a convergence of values for this parameter, the next step is to determine the value of the scale factor *A* . For the first stage, the relation, proposed by *Harter and Moore (1965a, 1965b)*, used to define *k* is given by equation [15], while for the second part, the relation in [16] is used. The aim is to find those values that maximize the likelihood function of the data.

$$\boldsymbol{k = \left(\frac{\sum_{i=1}^{N} U_i^k \ln(U_i)}{\sum_{i=1}^{N} U_i^k} - \frac{\sum_{i=1}^{N} \ln(U_i)}{N}\right)^{-1}} \qquad [15]$$

$$\boldsymbol{A = \left(\frac{\sum_{i=1}^{N} U_i^k}{N}\right)^{\frac{1}{k}}} \qquad [16]$$

Where N is the number of non-zero data points obtained from the time series and Ui is the average velocity for each 10-minute time interval in the series.

The difference between shape factor values obtained when intending to use the method for different time intervals (Annual or Winter/Summer) can be seen in figure 26. In this example, a series from a Portugal mountain region was used that had data for 5 years, from 2015 to 2019. If the MLM approach were based on each whole year, the values of *k* obtained would be described in the annual series (gray), while if it was chosen to have a Winter (blue) and Summer (orange) approach, the values already have modifications.

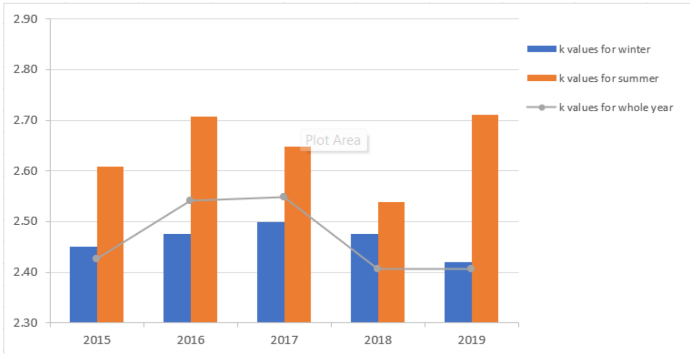

Figure 26: Example of *k* values for different discretizations-MLM

Although some values of *k* practically coincide using one approach or another, it is possible to notice that there are differences in some years, which can be explained by the more detailed discretization of the Winter/Summer series and possibly by the weather variation.

In the case of increasing the discretization, as was done in the previous example, the comparison of energy production values is usually done by first calculating the AEP and then doing the calculation with the half-yearly productions, to see how they can relate. It is interesting to try to understand how the Weibull parameters can influence the calculation of production, both for an annual period and for larger discretizations.

Analogously to the shape factor, the scale factor also undergoes changes depending on the time discretization performed, as can be seen in figure 27, where the same comparison of *A* values was made for each whole year and values obtained by the approach of Winter and Summer.

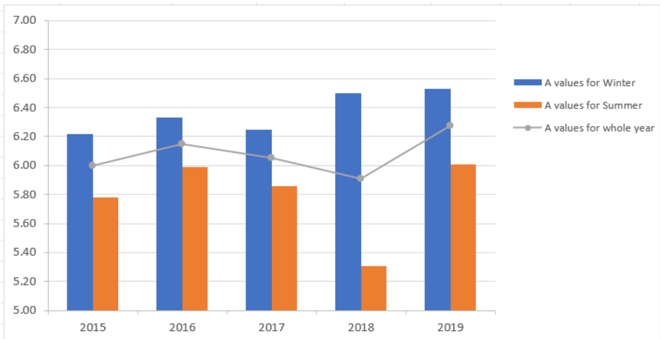

Figure 27: Example of A values for different discretizations -MLM

3.1.2. Energy Pattern Factor Method (EPFM)

The second method approached EPFM, also called Power Density Method, uses the wind speed profile obtained through in situ measurements or numerical modeling. The objective of this method is to adjust the occurrence distribution in such a way that the wind power available in the wind is equal to the equivalent power.

This statistical approach is based on the principle that the estimate of energy production can be made according to the amount of air mass passing through the rotor. Energy pattern factor is a sign of wind speed variability (Akdağ & Güler, 2015). From this, a parameter called energy pattern factor is defined, which is the ratio between the total power available in the wind and the power corresponding to the cube of the mean wind speed [17].

$$E_{pf} = \frac{\boldsymbol{Total\ amount\ of\ power\ available\ in\ the\ wind}}{\boldsymbol{Power\ calculated\ from\ cube\ of\ the\ mean\ speed}} \qquad [17]$$

The wind power density depends directly on the third moment of the Weibull distribution (Gugliani et al., 2018).Therefore, the energy pattern factor can also be defined by the velocity ratio of the mean of the wind speed cube with cube of the mean wind speed, according to relation [18], where n is the total number of data obtained in the time series.

$$E_{pf} = \frac{\frac{1}{n}\sum_{i=1}^{n} v_i^3}{\left(\frac{1}{n}\sum_{i=1}^{n} v_i\right)^3} \qquad [18]$$

Another way to express the EPF is through the relation [19].

$$E_{pf} = \frac{\Gamma\left(1+\frac{3}{k}\right)}{\Gamma^3\left(1+\frac{1}{k}\right)} \qquad [19]$$

Once this factor is obtained, the parameter *k* is approximated according to Akdag and Dinler (2009) to relation [20].

$$k = 3.957 E_{pf}^{-0.898} \qquad [20]$$

According to the literature survey conducted for this study shows that *Epf* is between 1.45 and 4.4 for most wind distribution in the world and therefore, the constants were established after performing power densities tests (Akdağ & Dinler, 2009).

The Energy Factor pattern can be related to the shape parameter as shown in figure 28.

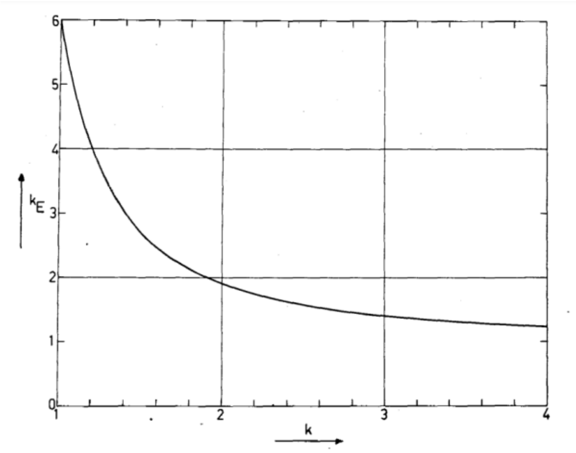

Figure 28: The energy pattern factor of a Weibull shape factor (Lysen, 1983)

And therefore, when the shape factor is already known, the value of the scale parameter, *A*, can also be obtained by using the relation [5].

In terms of discretization of the measurement period, the variation of the *k* and *A* parameters can also be seen according to the choice of an analysis of whole years or with the Winter/Summer combination for each year, figure 29. In this figure, there is the Portugal mountain region series as an example, from the years 2015 to 2019.

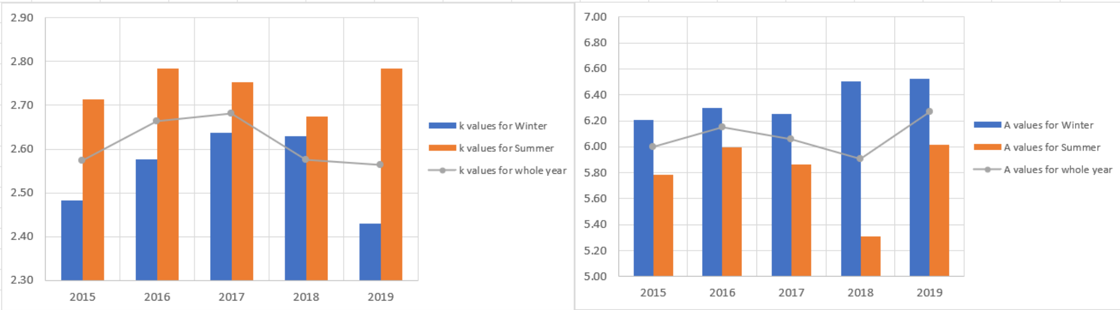

Figure 29: Example of *k* and *A* values for different discretizations -EPFM

3.1.3. WAsP Method

The WAsP algorithm method is the one used in the software of the same name, which is used for estimating energy production. For the sake of curiosity, in this software, the program is fed with the temporal data series of the measurement campaign and for the calculation of the AEP, it is necessary some factors such as topographic information, wind maps in high resolution and the roughness of the soil, in order to assign a study of the estimation with a given number and model of wind turbines.

However, for the case study, the focus is on the statistical methodology that WAsP uses to obtain the Weibull parameters and produce an adjustment with the wind time series. The approach taken is based on iterative processes that starts with two definitions of wind power density [21] and [22].

$$\boldsymbol{WPD = \frac{1}{2}\rho A^3 \Gamma\left(\frac{3}{k}+1\right)} \qquad [21]$$

$$\boldsymbol{WPD = \frac{1}{2N}\rho \sum_{i=1}^{N} U_i^3} \qquad [22]$$

Wind power density can be characterized as a measure of harnessing wind energy in a specific location and considers both wind speed and air density.

From these two relations [21-22], it is possible to obtain another one to find the scale parameter *A* [23], depending on the number of data points from the wind time series (N), the respective speeds for each 10-minute time interval (U_i) and the shape parameter *k*, which will start with an initial value stipulated to start the iterative process.

$$\boldsymbol{A = \sqrt[3]{\frac{\frac{1}{N}\sum_{i=1}^{N} U_i^3}{\Gamma\left(\frac{3}{k}+1\right)}}} \qquad [23]$$

After finding parameter A, obtaining the shape factor *k* is done according to relation [24] by an iteration process of values until convergence is reached. In it, a new element is introduced, an X parameter that is related to the cumulative distribution for an average velocity of the series ($\overline{U}$), according to equation [25].

$$\left(\frac{\frac{1}{N}\sum_{i=1}^{N} U_i^3}{A}\right)^k = -\ln(X) \qquad [24]$$

$$X = exp\left[-\left(\frac{\overline{U}}{A}\right)^k\right] \qquad [25]$$

Analogously to the other methods, the temporal discretization also directly impacts the values of the Weibull parameters. By using the same example series from Portugal mountain area in the Weibull statistical approximation approach, it is also possible to notice the difference in values when the velocity profile is characterized for whole years with a profile that characterizes the Winter/Summer period of a year and then considers the average of values (Figure 30).

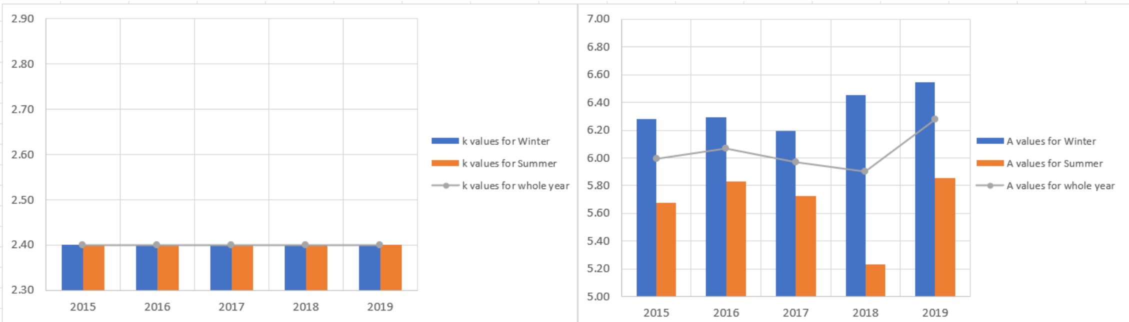

Figure 30: Example of *k* and *A* values for different discretizations - WAsP

In this specific case, we already see a difference from other methods in the convergence of *k* values that have the same values for Winter/Summer and whole year discretizations. However, this convergence does not strictly mean that the method of this algorithm is the most recommended, it is necessary to have methods for comparing the statistical models to check their respective reliability.

3.1.4. Method of Moments (MOM)

The Method of Moments is based on the numerical iteration of the mean wind speed and the standard deviation of the wind speeds used to estimate the parameters of a probability distribution based on the sample moments.

The Weibull parameters, in this case, are defined in terms of the wind speed population mean [26] and [27] and standard deviation.

$$\boldsymbol{v_m = \left(\frac{1}{N}\right)\sum_{i-1}^{N} v_i} \quad [26]$$

$$\boldsymbol{\sigma = \left\{\left(\frac{1}{N-1}\right)\sum_{i=1}^{N}(v_i - v_m)^2\right\}^{1/2}} \quad [27]$$

Where v_m is the average wind speed, v_i is the speed for each interval, N is the total number of observations and σ as the standard deviation of wind speeds.

Furthermore, the standard deviation can be represented by the gamma function and depends on the parameters A and k, as in the case of the average velocity equation [5]. This alternative relationship can be seen in equation [28].

$$\boldsymbol{\sigma = A\left[\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)\right]^{1/2}} \quad [28]$$

Equation [5] and [28] can then be related in the following way [29] and thus obtain the shape parameter with equation [30].

$$\boldsymbol{\frac{\sigma}{\bar{\bar{v}}} = \sqrt{\frac{\Gamma\left(1+\frac{2}{k}\right)}{\left[\Gamma\left(1+\frac{1}{k}\right)\right]^2} - 1}} \quad [29]$$

$$\boldsymbol{k = \left(\frac{0.9874}{\frac{\sigma}{\bar{\bar{v}}}}\right)^{1.0983}} \quad [30]$$

Obtaining the scale parameter is naturally found through equation [5].

3.1.5. Empirical Method of Lysen (EML)

As its name already suggests, this method calculates the Weibull parameters from empirical results of wind speeds at different locations. Therefore, after conducting analysis of the parameters in different circumstances of wind regime, models were established for the calculation.

The shape parameter is calculated based on a study previously carried out by Justus et al. (1977), in which the values of the mean wind speed and the standard deviation are needed. Equation [31] represents the relationship to obtain *k*.

$$\boldsymbol{k = \left(\frac{\sigma}{v_m}\right)^{-1.086}} \qquad [31]$$

Where v_m is the average wind speed and σ, the standard deviation of wind speeds.

The scale parameter, in turn, is obtained through equation [32] (Elie Bertrand et al., 2020), also related through empirical factors.

$$\boldsymbol{A = v_m\left(0,568 + \frac{0,433}{k}\right)^{-1/k}} \qquad [32]$$

3.2. Assessment of method accuracy

Obtaining the Weibull parameters for estimating production through the methods may differ from each other, due to the estimation model employed, that is, for the same wind time series, with the same discretization, it is possible to obtain different distribution adjustments and consequently production estimates. Therefore, it is necessary to establish accuracy criteria to compare these methods and understand which one has the lowest variance values.

These accuracy criteria can provide information about the level of accuracy that each statistical method can have in predicting power production. Evaluation of goodness of fit is an important step in the process of choosing the best distribution function for a certain region.(Chong & Ragai Henry Rigit, 2021)

A precise adjustment of the Weibull fitting makes it possible to more reliably predict the amount of energy that can be generated in a given area. Considering the implementation of a wind farm and that the study on its feasibility should be carried out in terms of the efficiency of wind turbines with the wind resource and also on the expected financial return, it is recommended that there be the least possible uncertainty in the estimation of production, avoiding project estimates that exceed or are underestimated in relation to the available project resources.

Therefore, in this subchapter, methods and efficiency tests that can be used for each statistical method for obtaining the Weibull parameters will be discussed. These include statistical indicators such as the Root Mean Square Error (RMSE), Chi-Square error (χ^2), Correlation coefficient (R) and the Determination Coefficient (R^2).

3.2.1. Root mean square error (RMSE)

The root mean square error method consists of analyzing the errors of a large number of data by means of the average deviation of the observed values in relation to the values predicted by the Weibull model. The RMSE is often employed to explore the differences between real-world and model data; the differences are termed residuals. (S. Kang et al., 2021)

The focus of this criterion is based on calculating the square root of the mean squares of the residual errors, while the lower the RMSE value, the better the statistical model used. This criterion demands a large amount of data such as the wind time series. The root mean square error (RMSE) has been used as a standard statistical metric to measure model performance in meteorology, air quality, and climate research studies. This method is defined according to relation [33].

$$\boldsymbol{RMSE} = \left[\frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \boldsymbol{x_i})^{\boldsymbol{2}}\right]^{\boldsymbol{1/2}} \qquad [33]$$

Where n is the number of wind speed dataset (bins), y_i is the actual wind distribution data, from the measured data without using Weibull, and x_i is the predicted wind distribution from the Weibull.

A particularity of this method is that the measured errors (y_i-x_i) undergo a procedure of being squared before being averaged and therefore, the sum of errors has different weights assigned to it. The assumption of this method is that the errors are unbiased and follow a normal distribution (Chai & Draxler, 2014).

From this, for randomly generated pseudo-errors with zero mean and unitary variance Gaussian distribution, it is noticed that as the number of samples increases, the error distribution using RMSE becomes more reliable, as seen in figure 31.

n	RMSEs
4	0.92, 0.65, 1.48, 1.02, 0.79
10	0.81, 1.10, 0.83, 0.95, 1.01
100	1.05, 1.03, 1.03, 1.00, 1.04
1000	1.04, 0.98, 1.01, 1.00, 1.00
10 000	1.00, 0.98, 1.01, 1.00, 1.00
100 000	1.00, 1.00, 1.00, 1.00, 1.00
1000 000	1.00, 1.00, 1.00, 1.00, 1.00

Figure 31: RMSE distribution according to sample size (Chai & Draxler, 2014)

In this example, you can see that as the sample size reaches a relatively high number of data, it is possible to reconstruct the error distribution close to its "true" or "exact solution".

For the example of the case of Portugal mountain area of 7 years duration for the measure campaign, and for example, with the first three different methods used for estimating the Weibull parameters, it was possible to obtain different RMSE results, as shown in Table 1 below.

Table 1: RMSE values for different methods for Weibull parameters estimation

Method	k	A	RMSE
MLM	2.47	6.08	0.0045
EPFM	2.61	6.08	0.0034
WAsP	2.40	6.04	0.0054

Thus, the model that offers the best adjustment is the one that contains the lowest RMSE value, which in this example of case study is the EPFM.

3.2.2. Chi-square

The chi-square method is also used for the evaluation of the accuracy of the predicted wind Weibull distribution in relation to real wind distribution and its criterion is to determine whether there is a significant difference between the two distributions. Therefore, his method (χ^2) is calculated by summing the squared differences between the observed frequencies and the expected frequencies, divided by the expected frequencies, as in relation [34].

$$\boldsymbol{x^2 = \sum_{i=1}^{n} \frac{(y_i - x_i)^2}{x_i}} \qquad [34]$$

Where n is the number of wind speed dataset (bins), y_i is the actual wind distribution data, from the measured data without using Weibull, and x_i is the predicted wind distribution from the Weibull.

For the example given, chi-square data were also obtained for the same three different methods, as can be seen in Table 2 below. In it, the method that offers the best distribution is the one with the smallest value of χ^2 because it indicates a smaller difference between the observed distribution and the expected theoretical distribution, in this case the WAsP algorithm method.

Table 2:Chi square values for different methods for Weibull parameters estimation

Method	k	A	Chi-square
MLM	2.47	6.08	0.076
EPFM	2.61	6.08	0.049
WAsP	2.40	6.04	0.021

3.2.3. Correlation coefficient (R)

The criteria of the correlation coefficient R already introduces another relationship between the data predicted by the Weibull distribution and the data observed in the wind time series. The approach in this criterion consists of defining a linear relationship between the two datasets, in which the parameter defining a good approximation is found in the interval [-1;1], in which a perfect correlation is one that approaches the extremes, while a lack of correlation between the data is seen when the R value approaches zero (Mukaka, 2012).

Thus, the closer the R value is to the extremes, the better the fit of the Weibull model. The definition of the coefficient of R in relation to the predicted and observed wind data can be seen in equation [35].

$$\boldsymbol{R} = \frac{\boldsymbol{n}\sum_{i=1}^{n}(\boldsymbol{y_i x_i}) - \sum_{i=1}^{n}\boldsymbol{y_i}\cdot\sum_{i=1}^{n}\boldsymbol{x_i}}{\sqrt{\boldsymbol{n}\sum_{i=1}^{n}\boldsymbol{y_i}^2 - (\sum_{i=1}^{n}\boldsymbol{y_i})^2}\cdot\sqrt{\boldsymbol{n}\sum_{i=1}^{n}\boldsymbol{x_i^2} - (\sum_{i=1}^{n}\boldsymbol{x_i})^2}} \quad [35]$$

In the case of the example given, the data of the correlation coefficient R obtained for the three Weibull parameter estimation methods can be seen in Table 3, in which the EPFM method also presents a better value than the others, even though the other two also present good correlation values, very close to 1.

Table 3: R values for different methods for Weibull parameters estimation

Method	k	A	R
MLM	2.47	6.08	0.989
EPFM	2.61	6.08	0.999
WAsP	2.40	6.04	0.998

3.2.4. The Determination Coeffiecient (R2)

The last criterion used in the case studies is the use of an R^2 determination coefficient that seeks to measure the variability of the observed wind speed data, in addition to being able to assess the quality of the Weibull adjustment made to the observed data. According to this accuracy criterion, R^2 values range from 0 to 1, where values close to 1 indicate that the fit model is adequate to accurately predict energy production.

Therefore, this model, according to (Kavak Akpinar & Akpinar, 2005), is able to measure the strength of the linear ratio of the distribution of estimated and actual frequencies through the relation [36].

$$\boldsymbol{R^2 = \frac{\sum_{i=1}^{n}(y_{i-}z_i)^2 - \sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(y_{i-}z_i)^2}} \qquad [36]$$

For the example case of Portugal mountain region, data on the coefficient of determination were also obtained for each method, over the 7 years of measured wind regime. According to table 4, there are the data obtained for the coefficient values, in which the method that guarantees greater precision for the measured period was also the EPFM.

Table 4: R^2values for different methods for Weibull parameters estimation

Method	k	A	$\boldsymbol{R^2}$
MLM	2.47	6.08	0.985
EPFM	2.61	6.08	0.991
WAsP	2.40	6.04	0.974

4. Case studies

The study presented in this document aims to analyze the goodness of the adjustment of statistical parameters to a wind series, in order to conduct the calculation of energy production estimates for wind turbines. The wind series, as mentioned in Chapter 2, are first obtained by a measurement campaign and subsequently the collected wind data is translated into time series.

In order to study the effectiveness of a WTS adjustment, two cases with a wind potential for the implementation of wind turbines will be analyzed, the first in Poland and the second in Brazil. Before any individual introduction about each case, the importance of this analysis can already be mentioned, due to the fact that the two regions are located in different zones of the planet and, consequently, have different climates.

For the calculation of the annual energy production (AEP), identical steps will be taken in each case study. Thus, the approach of standardizing the analysis steps allows the same study to be reproduced by other parties, which increases the reliability of the results.

As mentioned in section 2.4 , although for most purposes the analysis of energy production is commonly done in an annual period, it is important to report that this time discretization is not always ideal for energy buyers as they seek to understand how the production profile in short periods of time will affect the availability of selling energy. Thus, the calculation of energy production in different periods of time does not require statistical adjustments.

The AEP calculations in the two cases will not be limited to profiles in whole years, which means that for each case study there will be 4 energy analysis options. The strategy of using these discretizations is used to analyze whether it is possible to achieve better quality Weibull adjustments, but always bearing in mind the annual production estimates. The discretizations adopted for the study are:

- AEP based on data of a long measurement period
- AEP every entire year
- AEP using a Winter/Summer discretization
- AEP using a Night/Day discretization

According to the objective of this work, in each case study and in each previously mentioned discretization, the calculation of the AEP will be done with two approaches, the first of which is based on the wind time series and the second, using the five mentioned statistical approaches (MLM, EPFM, WAsP, MOM and EML) to produce a Weibull fitting to velocity distribution of each wind time series (WTS).

In addition, this approach is based exclusively on data obtained by a met station, which implies that the information translated into a series is the result of measurements made of wind speeds in all types of conditions, considering the adverse ones in very cold temperatures, there may be data failures due to freezing of sensors, as referred in section 2.2. Although there are already backup systems in case a sensor fails, the possibility of joint failure of all of them must be considered, which causes time intervals without measurements or with errors.

Failures derived from interference in the sensors, of any kind, are translated into the data series and often there is bridging through linear interpolation or other gap-filling approaches. After this step, equation [1] for the AEP correction is no longer used, as the data availability becomes 100%.

However, this filling only makes sense if there are large gaps in the data and, moreover, if the calculation of the AEP is done directly by the wind time series. Otherwise, using statistical approximations, the correction does not make sense, since the adjustment method fits any time frame. Therefore, for each case study, it is interesting to analyze the respective data availability.

After this first step of filling in or not the data gaps, it is necessary to proceed with the choice of wind turbines to be used in the wind farm. As the purpose of this work is based on analyzing the comparison between energy yield forecast using time series or through statistical approaches, it assumes that the region of implementation of the wind turbines, as well as the study of optimization of the locations and amount of equipment, have already been defined. So, these factors will not be analyzed.

4.1. The reference wind turbine

The next step, therefore, is the choice of a turbine model to be used for the calculation of AEP, which in turn considers the characteristic power curve of the wind turbine. The power curve represents a graph as seen previously in figure 17 that relates the different wind speeds with the respective generated powers. The AEP is then calculated from the sum of the energies obtained, which are derived from multiplying the power generated in each time interval by the duration of that interval.

Knowing that each wind turbine model has a specific power curve, the choice of the turbine used in this work was based on the international standard IEC 61400-1, which categorizes such equipment into classes and subclasses. The class is determined by the extreme winds, i.e., the maximum velocity value (10 minutes average), expected in 50 years return period, as already mentioned. The subclasses are defined by the calculated turbulence intensity.

After these parameters are calculated, the choice is made based on the following table [5].

Table 5: Wind turbine assessment guidelines from IEC 61400-1

Class	I			II			III		
Vmax,50 (m/s)	50			42.5			37.5		
Subclass	A	B	C	A	B	C	A	B	C
Intensity of turbulence [%]	16	14	12	16	14	12	16	14	12

Extreme wind estimates for class determination are made based on the Gumbel distribution, seen in section 2.4, for each series of case studies.

So, a preliminary analysis was carried out to obtain the turbulence intensity and the extreme wind of each one, resulting in the values in table 6.

Table 6: Values of extreme winds and turbulence intensity on the 2 case studies

	$V_{max,50}$ (m/s)	I_{turb} (%)
Poland	31.86	10.19
Brazil	30.32	10.23

From these tables, it is possible to see that both the case studies of Poland and Brazil indicate the choice of a class III and subclass C wind turbine.

From this, the choice of a single turbine was made based on the definition of the class with the subclass and from market research on the wind turbine models that are within these conditions. Of the analyzed models, there are some that fulfill both requirements, but the standard curve chosen can be seen in figure 32. This curve represents a type of wind turbine that can be applied to both case studies.

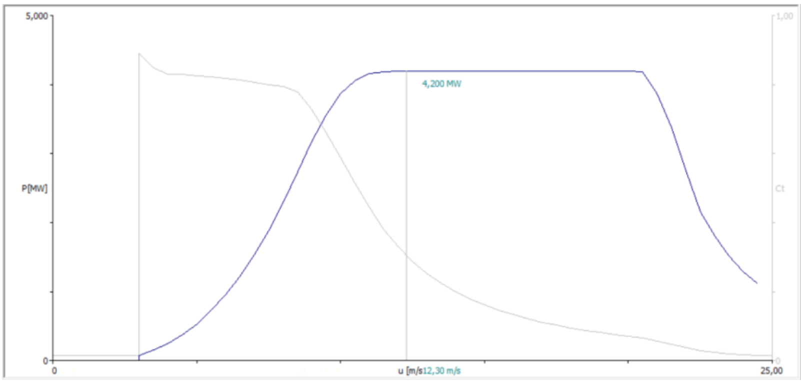

Figure 32: Power curve of the chosen wind turbine, class III

The chosen turbine has a rated power of 4.2 MW and, according to the curve in figure 34, a cut-in speed of 3 m/s and a cut-out speed of 24.5 m/s. In addition, the height of the rotor axis of this turbine model is 105 m.

In this curve, besides the representation of the power values relative to the speeds, there is also another important parameter to be considered in the study of the implementation of a wind farm, which is the axial thrust coefficient (Ct). This parameter represents the fraction of kinetic energy produced from the wind that is converted into useful work by the turbine.

The Ct parameter also quantifies the wake phenomenon produced by the turbines because it affects the amount of energy that the turbine extracts from the wind. If the wind turbine works with a high value for this parameter, there will be a more pronounced wake with a more turbulent velocity distribution.

The curves in Figure 34 then represent two fundamental parameters that are generally listed by the wind turbine manufacturer. However, the Ct coefficient will not be used in the scope of this work as the estimates will consider a single wind turbine.

The power generated by a wind turbine (P_{wt}) can be modeled by the following relationship [37]:

$$\boldsymbol{P_{wt}} = \begin{cases} \boldsymbol{0} & \boldsymbol{V(t) < V_{in}\ or\ V(t) > V_{out}} \\ \boldsymbol{P_{rated}} & \boldsymbol{V_{rated} < V(t) < V_{out}} \\ \boldsymbol{q * P_{rated}} & \boldsymbol{V_{in} < V(t) < V_{rated}} \end{cases} \qquad [37]$$

Where P_{rated} is the power defined by the wind turbine manufacturer, $V_{in}$is the cut-in speed, V_{rated} is the rated wind speed, V_{out} is the cut-off speed and v(t) is the wind speed at desired height.

The "q" factor is defined by the relation [38].

$$\boldsymbol{q = \frac{(V(t) - V_{in})}{(V_{rated} - V_{in})}} \qquad [38]$$

However, in the scope of this study, the relations [37] and [38] are not used because the values in the turbine power curves are already established by the manufacturer.

4.2. Scenarios

The two case studies have a particular climate and vegetation, with no similarity between them. This study was conducted to try to find the influence of time series procedures and statistical models to characterize the wind regime, in an attempt to understand if it is indeed possible to calculate and predict the AEP with the lowest possible uncertainties, on each approach.

Before any discretization of the Time Series, it is important to point out that in the study it will be called the base case, since its AEP is calculated without any static interference or temporal discretization.

The first chosen discretization is done in such a way as not to divide the measurement period, to characterize the whole time period and the AEP calculation as homogeneous. That means, if for a given case the measurement period is 5 years, estimate of the annual energy production will be the same for all years.

The second discretization is based on establishing an AEP estimate with the variation of the wind regime throughout the years, the so called inter annual variability. Therefore, it is expected that there are variations between years in production, as in figure 33, which shows an example of the distribution of speeds throughout the years.

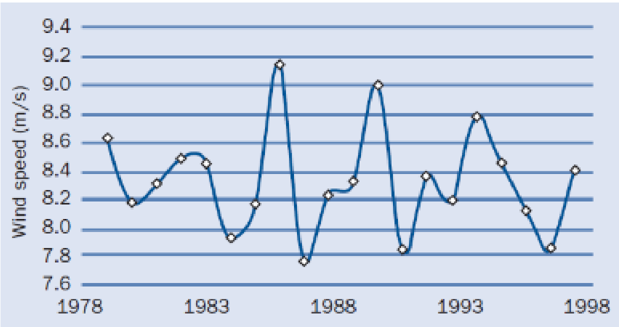

Figure 33: Inter annual variability (*The Annual Variability of Wind Speed*, Wind Energy - The Facts.)

The third discretization chosen is to separate the Winter and Summer periods, calculating the energy production obtained from each season and then adding them to obtain the AEP. In fact, this discretization does not only consider the three Winter and Summer months , for the 6 coldest months (January, February, March, October, November, and December), the total energy is seen as in the "Winter" category, while the other six warmer months (April, May, June, July, August, and September) are included in the "Summer" category. This discretization aims to describe seasonality for each, as in the example of figure 34.

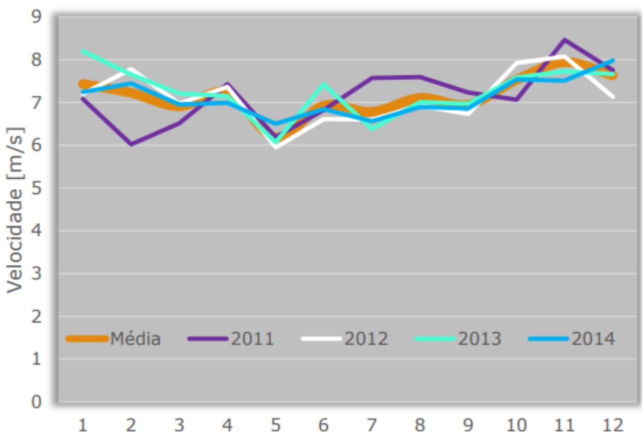

Figure 34: Example of seasonality (Rodrigues, 2022)

However, this example is a poor representation of the seasonality of the wind, as it is possible to see that the regime is practically constant throughout the year.

The last and fourth discretization is done in an attempt to further reduce the period of analysis and considering, for example, market demands to have short-scale production profiles. Therefore, the discretization made is Night and Day, in which each series of case studies are separated into periods of the night that comprise from 7:00PM to 6:50AM, and periods of the day that are between 7:00AM to 6:50PM.

The choice of this last Night/Day discretization has the objective of considering different values for the shear factor, instead of an average annual value, given the dependence of this factor on atmospheric stability. Atmospheric stability tends to be greater at night, which leads to more open wind profiles, with higher shear factors.

Therefore, for the two case study scenarios, steps of a standardized analysis were defined below in the calculation of the AEP, according to each discretization previously mentioned:

1) Interpolate, if necessary, for missing values and failed blocks of data

2) Calculate shear factor values for each time interval from velocities measured at two heights above the ground and extrapolate, for the height of the turbine rotor shaft,105 m above ground

3) Calculate density values for each time interval, for the height of the turbine rotor shaft

4) Calculate AEP from the time series (without using Weibull fitting)

5) Get Weibull parameters in each statistical method

6) Calculate AEP from the Weibull adjustment parameters

7) Calculate the accuracy criteria parameters for each statistical method

The first step is intended to fill the data gaps, trying to predict the information related to the speeds not captured using a linear interpolation and considering the speed patterns on the previous and the following day.

The second step uses equation [8] to calculate the shear factor values in each velocity interval, that is, after interpolation of missing data, the series will undergo an extrapolation in order to predict the velocities obtained at 105 m, at the height of the turbine rotor shaft. This was only possible because in all two series, the velocity data were measured at heights of 40 and 80 m, therefore, one more reason to standardize this step in all case studies.

For the third step, that is, to determine the density values for each time interval at a height of 105 m , it was first necessary to find the density values at a height of 80 m, using the relation [9]. After that, an air density extrapolation was made from a height of 80 m to 105 m, in all the two cases using the following equation [39]:

$$\boldsymbol{\rho_2 = \rho_1 * \left(\frac{T_2}{T_1}\right)^{\left(-\frac{g}{R(T_{2}-T_{1})}\right)}} \quad [39]$$

Where *ρ2* is the air density at a height of 105 m, *ρ1* is the air density at a height of 80 m, *T2* and *T1* are the absolute temperatures corresponding to heights of 105 m and 80 m, respectively, *g* is the gravitational acceleration (approximately 9.81 m/s^2), *R* is the ideal gas constant (approximately 287.1 J/(kg·K)).

However, the temperature in each time interval at height 105 m is not given, so it is necessary to use equation [40] to obtain an approximation of T2.

$$\boldsymbol{T_2 = T_1 + 0,16} \quad [40]$$

This statistic is an average based on observations and climatological studies, showing an average rate of temperature decline with increasing altitude in the standard atmosphere.

The fourth step already represents the first instance of AEP calculation, however, using only the time series. The procedures carried out in this activity are represented in the flowchart, in figure 35.

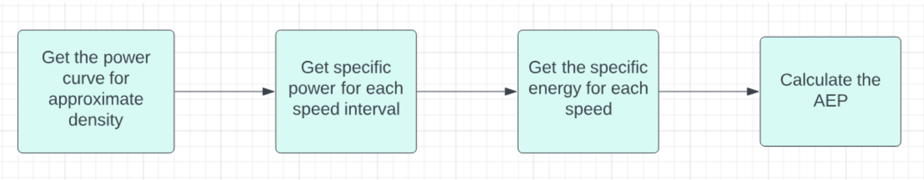

Figure 35: Flowchart of AEP calculation without Weibull fitting

All calculations performed in this study were performed using the EXCEL software, both for the statistical approximation methods and for the AEP calculations (with and without Weibull parameters). Furthermore, in the spreadsheets, the power curve present in appendix A was available as an auxiliary database.

Therefore, the choice of the representative density for each wind interval was obtained using the "VLOOKUP" formula in order to find the values closest to the list of wind turbine densities.

To find the power relative to each velocity interval, it was necessary to perform a linear interpolation, in which, with the density found previously, there will no longer be a matrix of columns as in figure 36 where it was necessary to carry out two simultaneous interpolations. Thus, the interpolation made to find the power relative to each data point is given by equation [41].

	ρ1	ρ2
V1	P11	P12
V2	P21	P22

Figure 36: Matrix scheme of power values for each density and speed value

$$\frac{(\boldsymbol{P_{i+1}} - \boldsymbol{P_{i-1}})}{(\boldsymbol{V_{i+1}} - \boldsymbol{V_{i-1}})} = \frac{(\boldsymbol{P_{*}} - \boldsymbol{P_{i-1}})}{(\boldsymbol{V_{*}} - \boldsymbol{V_{i-1}})} \qquad [41]$$

Where P_* is the extrapolated velocity at 105 m of each series data and V_*, the power obtained according to the power curve.

The specific energy for each speed value is obtained by the product of the power found after the interpolation made in the previous step with the factor (1/6). This factor represents the amount of time in hours that the given average speed was present, captured by the turbine, that is, the 10 minutes. Thus, the AEP obtained in the time series, is the sum of the energy found for each wind speed data.

After calculating the AEP, the fifth step of the study is to calculate the Weibull parameters for each series and also for each discretization, that is, whether whole years, Winter/Summer, or Night/Day. There will be the calculation of the scale and shape parameters according to the MLM, EPFM and WAsP methods.

The sixth step, in turn, already includes the calculation of AEP by statistical approaches, which means that any discretization of the series will be reduced to the conditions of the power curve. For example, in time series analysis, all velocity data was computed in the calculation and now it is not. Because the turbine only operates above 3 m/s and below 24.5 m/s, speed values outside this range are not considered. Thus, the activities performed in this sixth step are represented in the flowchart in Figure 37.

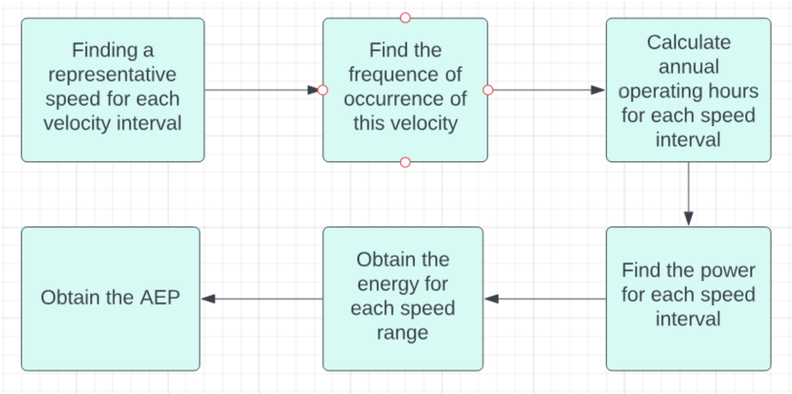

Figure 37: Flowchart of AEP calculation using Weibull fitting

The representative speed is chosen by trying to find the average speed for each interval. For this, the series of wind data is divided into the speed intervals provided by the power curve of the wind turbine and then this representative speed is calculated between each interval. Among the intervals of each velocity, for example, [3; 3.5] and [24; 24.5], the average of the velocities within each interval is calculated, obtaining the representative velocity.

The frequency of occurrence of each velocity is obtained using equations 1 or 2 that describe the Weibull model, after obtaining the shape and scale parameters by statistical methods.

Step 3, of this approach with Weibull, represents the annual operating hours for each representative velocity, of each velocity interval of the power curve, that is, the number of hours that the wind blows for each velocity or *H(u)* as referred to in section 2.3. This is obtained by multiplying the frequency of each speed interval by constant part 8760, which represents the total hours in a year (365 days x 24 hours a day). However, this multiplication is only effective if the AEP analysis is done for whole years, that is, for the other 2 discretizations of Winter/Summer and Night/Day, the H(u) is multiplied by 4380 which is half.

Obtaining the power relative to each interval is done using equation [40], similarly to the AEP process without Weibull and the specific energy at each interval is not obtained by multiplying the power by the factor (1/6) but by the factor "*H(u)* ". However, instead of calculating only with the time series that considered a density value for each interval, in the approach with statistical methods, only an annual average density or average densities are considered for each discretization.

Therefore, the AEP is found by summing the energies relative to each velocity interval.

Step 7 of this work is the calculation of accuracy parameters of the methods. As each statistical method has its own model for Weibull adjustment, it is important to understand whether their use is reliable for the series presented in each case study. For that, there are 4 reliability analysis criteria to be conducted (RMSE, Chi-square, correlation coefficient and coefficient of determination), as established in section 3.2. These four criteria will be used for the four established time discretizations and also for the two cases.

4.1.1. Poland

The first case study is located in Poland, more specifically in a region called Gluchow, according to the location map in figure 38. This region, as it is located in central Europe, is characterized by a temperate continental climate with characteristics of low temperatures, rain, and the occurrence of snow in Winter and high temperatures in Summer.

Figure 38: Poland map (Windhunter, 2013)

According to the company "Meteoblue", the distribution of wind in the last 30 years in this region can be represented by figure 39, in which in Winter there are indeed winds with higher speeds than in Summer.

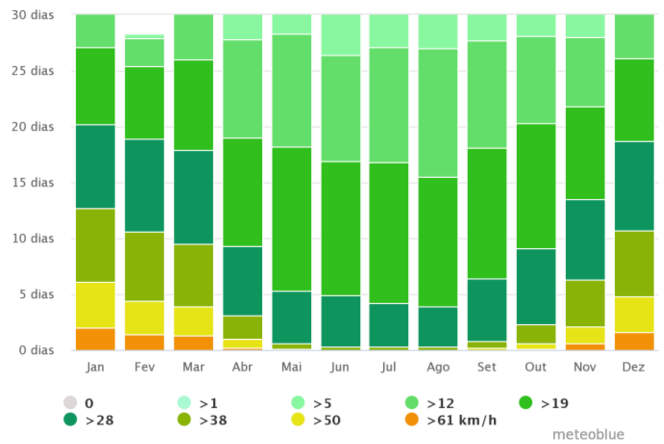

Figure 39: Wind average month velocities - Gluchow (Meteoblue)

According to the measuring station arranged in Gluchow, the wind time series was obtained for a period of 7 whole years, that is, data between January 1st and December 31st of the years 2015 until 2021.

From the data presented in the wind time series, the total amount of speed information at each 10 min interval, at a height of 80 m was 368208. This number represents the total amount of records but the amount of data available, that is, with speed information that doesn't have errors was 362431. Therefore, using equation [2], data availability was 98.43%. In spite of the great availability in this particular case, missing data will be replaced by interpolation according to the technic described in section 2.

From this series it is possible to obtain the histogram and the wind rose (see figure 40) according to the WAsP software, in which the series is fed into the program with data on speed, temperature, pressure and wind direction at different intervals. One can notice that the predominant wind azimuth is West and that the predominant speed is in the range of bin [6;7] m/s.

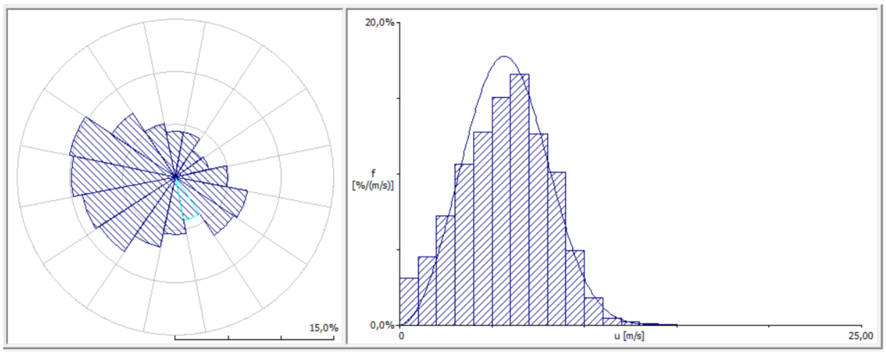

Figure 40: Wind rose of occurrences and histogram (Poland)

It is important to report that this series of data obtained information for the heights of 40 and 80 m, and in figure 41, the information refers to the height of 80 m.

In order to understand the interaction between velocity and height, that is, to characterize the vertical profile of velocities in this case study, the shear factor was used according to the relation [8]. Thus, the overall shear factor of this series, performing an average of the value of this index in each time interval was 0.266. Therefore, with the average speed values for each height (40, 80 and 105) m represented in table 7, the vertical wind speed profile of Poland is shown in figure 41.

Table 7: Average velocity values for each height (Poland)

Height (m)	40	80	105
Velocity (m/s)	5.34	6.38	6.89

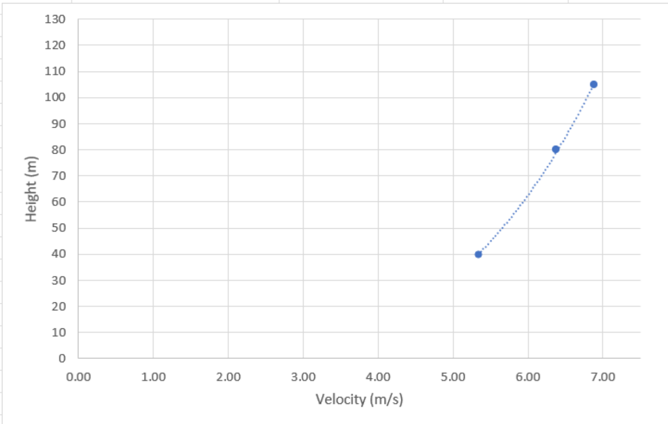

Figure 41: Vertical velocity profile (Poland)

The average temperature found in the series is 9.84°C and the average density, calculated by relation [9] is 1.22 kg/m^3.

4.1.2. Brazil

The second case study is located in the northeast of Brazil, more specifically in the coastal region of the state of Rio Grande do Norte, as in figure 42, which represents the northeast region (a) of the country and the state in which the series was obtained (b).

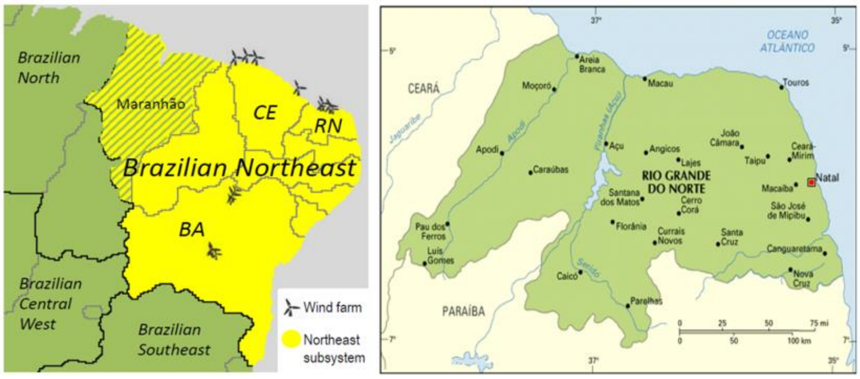

Figure 42: (a) Northeast region of Brazil (Jong, n.d.) and (b) State of Rio Grande do Norte

This area is located in the tropical region with a climate characterized as tropical humid, in which temperatures have an annual average of 20°C, and a relatively high rainfall.

Regarding wind data according to the company "Meteoblue", for 30 years, the profile of annual average wind speeds can be seen in figure 43, in which the highest average wind speeds are also seen in the winter of Southern hemisphere, in this case, July and August.

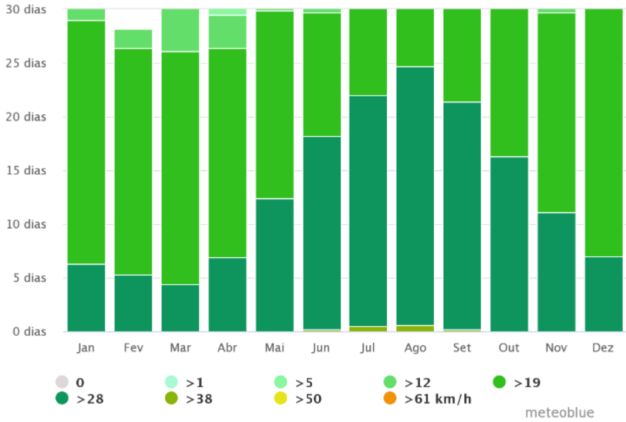

Figure 43: Wind average month velocities - Rio Grande do Norte, Brazil (Meteoblue)

The wind time series was obtained for a period of 3 years, from 20/06/2009 to 19/06/2012.

From the data presented in the wind time series, the total amount of wind speed information, at a height of 80 m is 157824 data records but the amount of data available, that is, with speed information that doesn't have errors was 156423. Therefore, using equation [2], the data availability was around 99.11%. In this case, availability is also high, so there is no need to collate data. In this case, the availability is so high that missing data was not replaced.

Using the WAsP software and feeding it with the series obtained in with data on speed, temperature, pressure, and wind direction at different intervals it was possible to obtain the wind rose and the speed histogram, seen in figure 44. In it, one can notice that the most predominant wind direction is Southeast and that the most predominant speed is in the range of bin [8;9] m/s. The change from a smaller velocity interval to a larger one, when comparing case 1 with 2, can already be justified by the fact that Northeast Brazil is located close to the

Intertropical Convergence Zone (ITCZ), an area where air masses meet that contributes to the formation of low-pressure systems and the intensification of winds.

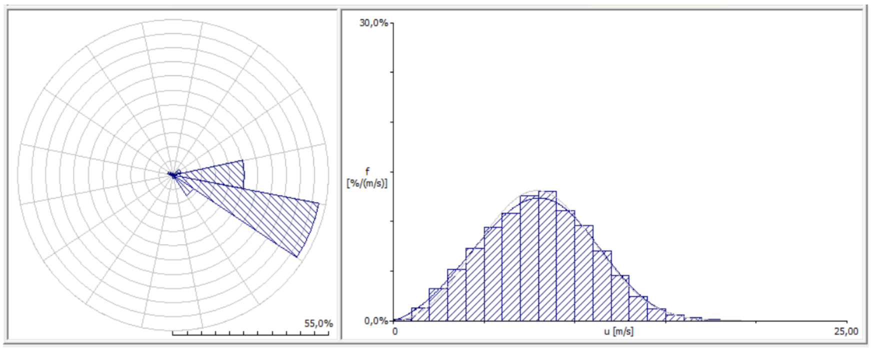

Figure 44: Wind rose of occurrences and histogram (Brazil)

The wind time series was also obtained with speeds at heights of 40 and 80 m and in figure 44, the information refers to the height of 80 m.

To characterize the vertical velocity profile, the overall average shear factor of this series was 0.07. Therefore, with the average speed values for each height (40, 80 and 105) m represented in table 8, the vertical speed profile of Brazil is shown in figure 45.

Table 8: Average velocity values for each height (Brazil)

Height (m)	40	80	105
Velocity (m/s)	7.66	8.00	8.15

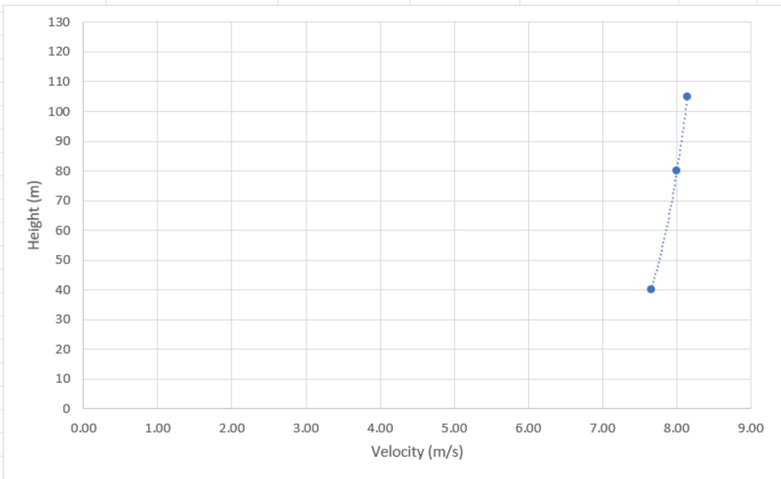

Figure 45: Vertical velocity profile (Brazil)

Besides that, the average temperature found in the series is 26.8°C and the average density, calculated by relation [9] is 1.16 kg/m^3.

5. Results

5.1. Poland

5.1.1. Long measurement period (7 years)

The Polish series, as mentioned in section 4.2, has a 7-year wind time series, in the case of 01/01/2015 to 31/12/2021.

As the chosen wind turbine model has a hub axis height of 105 m, some parameter data provided by WTS were extrapolated to this height. The main quantities that went through this process were velocity and air density, according to equations [8] and [39], respectively.

Considering the power curve of Appendix, A, in which the rated power is 4.2 MW, and defining H(u) as 10 minutes, the energy obtained directly from the 7-year series was 112011 MWh. Therefore, considering the discretization in 7 years, the AEP was obtained by dividing this total value by the time interval of 7 years, finding a value of 16002 MWh.

The next step, after calculating the AEP directly from the 7-year WTS, was to estimate the Weibull parameters for each of the five methods (MLM, EPFM, WAsP, EML, and MOM). The results of the shape and scale parameters for each method can be found in table 9.

Table 9: Weibull parameters (7 years)

Weibull parameters		
Method	k	A
MLM	2.53	7.76
EPFM	2.65	7.75
WAsP	2.40	7.66
EML	2.55	7.76
MOM	2.54	7.76

This table shows that the values of the Weibull parameters, both in terms of shape and scale, were estimated with little variability between the methods. In fact, it is possible to see in figure 46 how much the parameters of each method vary in relation to the average value of these results. The mean value of the methods for the shape parameter was 2.53, while the scale parameter was 7.74 m/s.

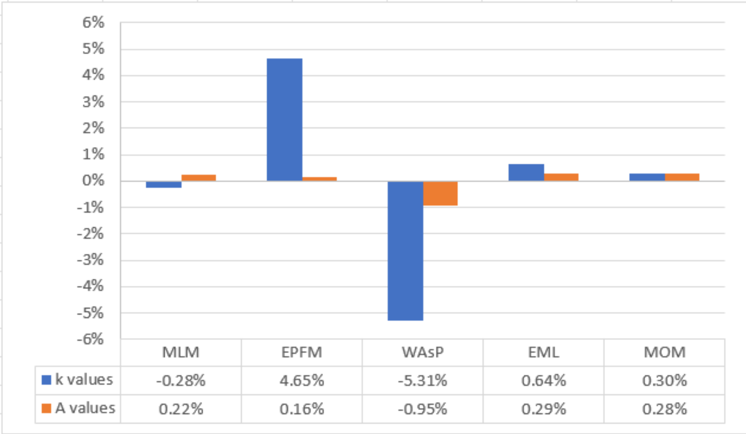

Figure 46: Weibull parameters variation against their mean value

In this case, the EPFM and WAsP methods were the ones that had the greatest variability in the *k* parameter in relation to the mean value. However, the existence of a greater variability of the parameter of a method, in relation to the average value of the methods, does not necessarily indicate a bad adjustment. It is necessary to see how the adjustment of each method is made in relation to the histogram of the wind series and conduct then the evaluation of the reliability criteria.

Before that, it is also important to analyze the AEP estimated for each method and obtain a comparison of the value found in relation to the AEP obtained directly from the time series. In Table 10, it is possible to see the value found by each method and how close it is to the Time Series estimate.

Table 10: AEP for each method and their variation against AEP from Time Series (Poland)

Method	AEP (MWh)	AEP Variation (%)
MLM	16339	2.10%
EPFM	16375	2.33%
WASP	15897	-0.66%
EML	16351	2.18%
MOM	16345	2.14%

According to this table, the method that lead to the AEP estimate closest to the value found from the time series was the WAsP Method, with an underestimation of only 0.66%. This concretizes the idea of not taking hasty measures in choosing the best adjustment method by only analyzing the variability of method parameters in relation to the mean value.

The adjustment made by each method to the 7-year series can be seen in figure 47. In it, all methods have very similar adjustments, so the next step is the analysis of accuracy criteria.

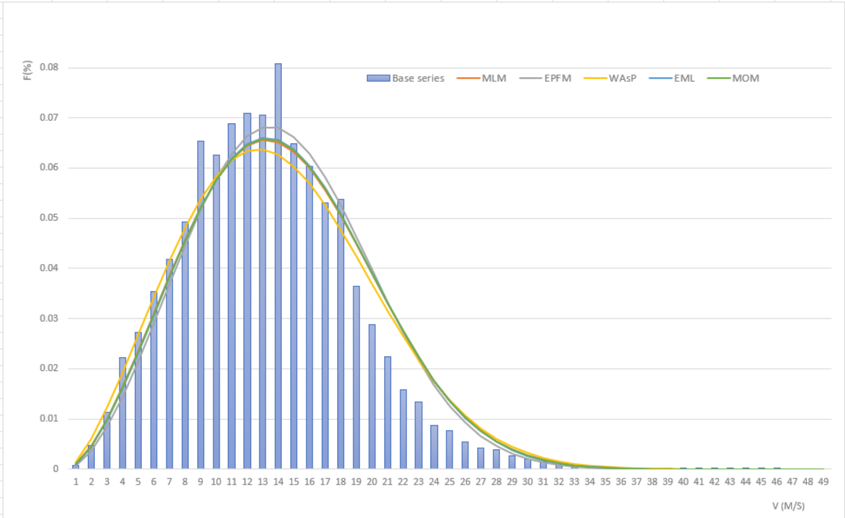

Figure 47: Weibull adjustment for each method related to the Time Series (7 years)

After performing the analysis of the adjustment criteria, presented in section 3.2, values were obtained for each method. Table 11 shows the goodness of fit results of each statistical method in relation to the wind time series.

Table 11 : Accuracy test values for each method (7 years)

Statistical methods	The Weibull parameters		Accuracy test efficiency			
	k	A	RMSE	X^2	R	R^2
MLM	2.53	7.76	0.0040	0.0208	0.9966	0.9872
EPFM	2.65	7.75	0.0032	0.0474	0.9976	0.9919
WAsP	2.40	7.66	0.0051	0.0254	0.9928	0.9794
EML	2.55	7.76	0.0038	0.0218	0.9968	0.9881
MOM	2.54	7.76	0.0039	0.0212	0.9967	0.9877

From this table, it is possible to see that for three reliability criteria, the one that offers the best adjustment model is the EPFM, both for higher values of R and R^2, and for the lowest value in the RMSE. In the Chi-Square criterion, the model that offers the best fit with the lowest value is the MOM.

Besides that, all methods look good in terms of goodness of fit. The EPFM method is the one that offers the best for three reliability criteria, the RMSE, R and R^2. The MLM is the one that offers the best reliability criteria value with Chi-Square.

If the work stopped here and without the need for further discretization, it would be possible to state that the models in general have good adjustments to the WTS. Because the EPFM method is the one with the best results in the reliability criteria, it would be a strong candidate to serve as a statistical model for estimating the AEP for the Time Series . However, the difference in many of these criteria is from the fourth decimal place, which can be explained by truncation and rounding errors. In the end, the best method is the one capable of reproducing the AEP value calculated directly by the Time Series, which in this case was the WAsP.

For the calculation of the AEP, a parameter used, in addition to the speeds, was the air density but with the annual average value. So, probably, the differences in the AEP and accuracy results are due to the influence of the turbine power curve. According to section 4.1, the turbine disregards power values for speeds below 3 and above 24.5 m/s, something that the calculation with the Time Series does not do. If there was a good adjustment at the nominal power level, for example, there would be no significant impact, since for speeds from 12 to 20 m/s the turbine produces the same.

Therefore, the next step to take, after this calculation, is to try to carry out the same study as the AEP, with the adjustment of the data series using Weibull parameters but with greater time discretizations.

5.1.2. Yearly discretization

In this second approach, the AEP calculation will not be done by simply finding the total energy produced over the years (7 years in Poland) and dividing by 7 years, that means, finding the average. In this case, the energy calculation will be done each year, that is, there will be seven AEP values, one for each year. The 7-year series was separated into 7 different databases, discretizing an entire year.

So, for each year, the respective AEP will be calculated and the final value to be considered is their average. This allows annual energy differences to be seen more clearly between years. In figure 48 there are the results of the AEP measured each year (blue) compared with the value obtained by the discretization of 7 years (orange), both measured directly from the Time Series, without the intervention of a statistical method yet.

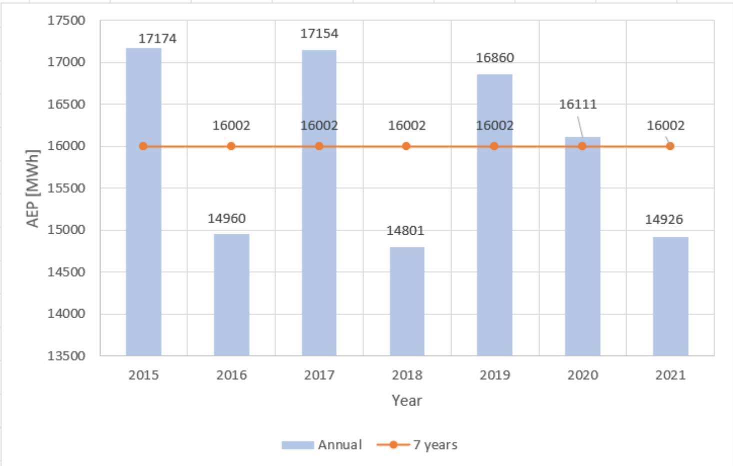

Figure 48: AEP values for each year *versus* AEP for 7 years calculation (Both from Time Series)

This figure demonstrates that the AEP values with the annual discretization oscillate in relation to the mean of the 7-year discretization (orange) and that the average AEP between them (blue) is 15998 MWh, that is, having an underestimation of 0.02%, which is not significant.

Table 12, in turn, already shows the values of the shear factor, density for each year and the respective AEP.

Table 12: Shear factor, Density and AEP calculated for each year (Poland)

Year	Shear factor	Density	AEP (MWh)
2015	0.269	1.222	17174
2016	0.272	1.223	14960
2017	0.253	1.225	17154
2018	0.272	1.222	14801
2019	0.263	1.218	16860
2020	0.269	1.220	16111
2021	0.263	1.227	14926

Only through these values, it is not possible to obtain a concrete relationship on the variation of shear factor and density with the AEP. Especially because the values of these parameters are very close, both for the shear factor, which has an average of 0.266, and for the density, with an average of 1.222 kg/m3. Thus, it is necessary to perform calculations with larger discretizations, as will be discussed in the next subchapters.

The calculation of the Weibull parameters for each method was also done for each year of the total 7 years available in the wind time series. The results of the shape and scale parameter values can be seen in table 13.

Table 13: Weibull parameters for each year and method (Poland)

	MLM		EPFM		WAsP		EML		MOM	
Year	k	A	k	A	k	A	k	A	k	A
2015	2.41	8.04	2.56	8.04	2.40	8.04	2.44	8.05	2.44	8.04
2016	2.44	7.47	2.57	7.46	2.40	7.46	2.47	7.47	2.46	7.47
2017	2.49	7.97	2.62	7.97	2.40	7.97	2.52	7.98	2.51	7.97
2018	2.49	7.37	2.62	7.36	2.40	7.36	2.52	7.37	2.51	7.37
2019	2.63	7.91	2.72	7.91	2.40	7.91	2.67	7.92	2.66	7.92
2020	2.61	7.80	2.70	7.79	2.40	7.79	2.65	7.81	2.64	7.81
2021	2.55	7.50	2.66	7.50	2.40	7.50	2.58	7.51	2.57	7.51

Considering that for each method, their respective Weibull parameters have already been calculated, as shown in the table above, the next step was to obtain the AEP for each year, according to these estimates. In table 14 these results are demonstrated with the average AEP for each method.

Table 14: AEP for each method and year (Poland)

	AEP (MWh)				
	MLM	EPFM	WASP	EML	MOM
2015	17241	17385	17265	17305	17270
2016	15216	15220	15128	15226	15223
2017	17070	17181	16782	17132	17088
2018	14853	14842	14609	14860	14858
2019	16964	17032	16253	17032	17024
2020	16535	16556	15890	16600	16593
2021	15374	15412	14988	15423	15420
Average	16179	16233	15845	16225	16211

This average AEP value of each method obtained by this second approach (yearly discretization) can be compared with the AEP value obtained by the first approach (7 years), as shown in Table 15.

Table 15: AEP variation between 7 years and yearly approach

	AEP (MWh)		
	7 years	Average for each year	AEP Variation (%)
MLM	16339	16179	-0.98%
EPFM	16375	16233	-0.87%
WASP	15897	15845	-0.32%
EML	16351	16225	-0.77%
MOM	16345	16211	-0.82%

The obtained values in table 14 can be compared with the AEP values calculated directly by the time series. The annual energy production variations of each method in relation to the direct approach, without statistical method can be seen in figure 49.

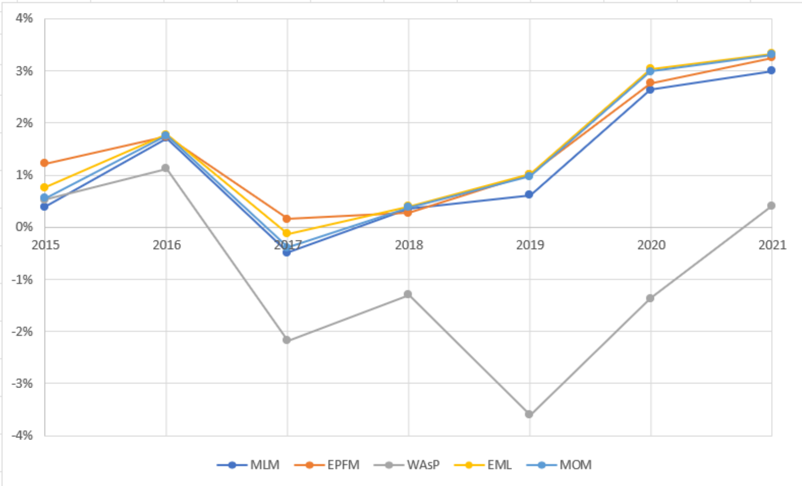

Figure 49: AEP variation comparing each statistical method with Time Series (Poland)

This figure demonstrates the importance of calculating energy for different approaches and, in this case for Poland, not considering the calculation period of 7 whole years as the most reliable. This is because in the 7-year calculation, the WASP method was the one that presented the lowest energy variation compared to the time series value, and therefore it was a candidate for the method that best fits the time series. However, this figure portrays that WAsP is the only one that starts to have different results from the other methods, from 2017 and not always having the smallest variation of AEP with the time series.

When analyzing the figure again, it is noticed that although the WAsP method does not beat the other methods in terms of lowest variation in all years, it continues to beat the other methods in 3 years (2016, 2020 and 2021). The MLM method has the best ability to reproduce the time series AEP in two years (2015 and 2019), while the EPFM and EML methods are better in one year (2018 and 2017, respectively). The MOM method does not have the lowest variability in any year.

The adjustments carried out by each of the methods in each year can be consulted in Appendix B, figures (B1.1-B1.4).

In terms of adjustment criteria, in this case where the discretization is yearly performed, it was decided to use the determination coefficient (R^2) criterion, as it is more commonly used in the industry. Therefore, the results obtained from this test for each method are represented in table 16.

Table 16: Determination Coefficient values for each statistical method (yearly discretization - Poland)

	R^2 Values				
	MLM	EPFM	WAsP	EML	MOM
2015	0.975	0.984	0.975	0.977	0.977
2016	0.988	0.987	0.986	0.989	0.988
2017	0.979	0.985	0.974	0.981	0.981
2018	0.985	0.989	0.980	0.986	0.986
2019	0.983	0.987	0.977	0.985	0.984
2020	0.989	0.993	0.982	0.991	0.990
2021	0.985	0.989	0.976	0.987	0.986

For a better visualization of the values and to understand which method establishes the best adjustments, figure 50 represents the values of the accuracy criteria of each method over the years.

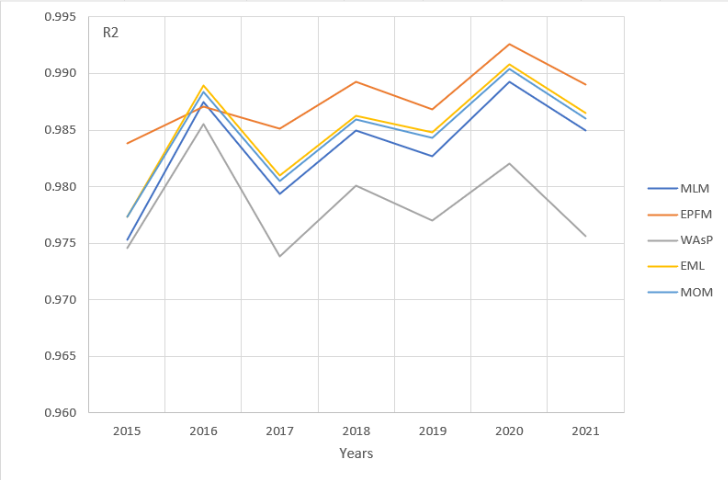

Figure 50: Annual Determination Coeffiecient values for each method

Through this figure, it can be seen that there is a method with better accuracy values in 6 years, which is the EPFM (orange). The WAsP, which in the AEP variation analysis, in relation to the time series energy calculation, obtained the best results, in this case it was the opposite, it obtained the worst accuracy results.

5.1.3. Winter/Summer discretization

The Winter/Summer discretization, as mentioned earlier, has the objective of studying whether the separation into smaller periods, in the case of half-yearly production, the calculation of the AEP manages to meet the seasonality phenomena. Therefore, the analysis conducted in this third discretization separates the year into two semesters, the first called Winter because it considers the six coldest months of the year (January, February, March, October, November, and December). The second semester, called Summer, considers the six hottest months of the year (April, May, June, July, August, and September).

The first analysis obtained is relative to the monthly average speeds in each of the years from 2015 to 2021, as can be seen in figure 51.

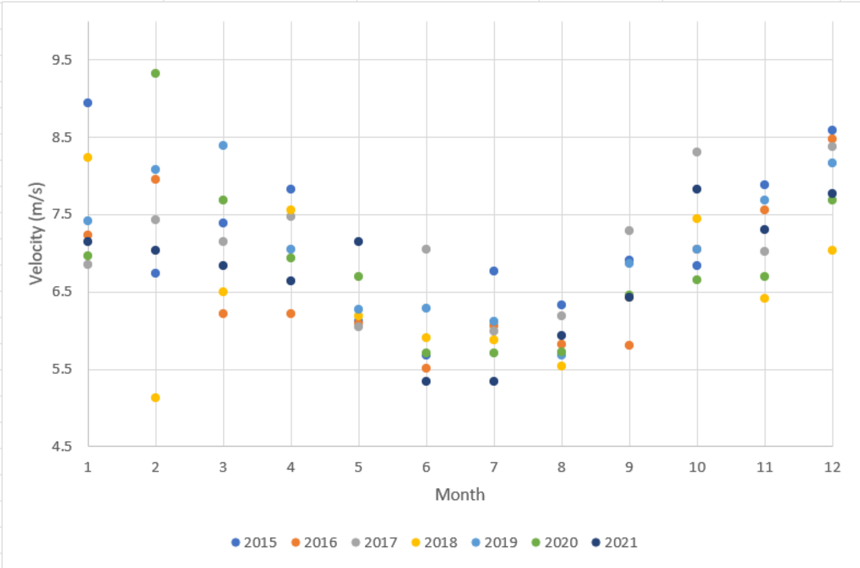

Figure 51: Monthly average velocity in each year for 105 m above ground level

In this case, seasonality is not very significant for some years but in others like 2018 and 2020 it is already possible to see its effects on the average monthly speed.

In relation to the other parameters obtained by the wind series, such as shear factor and density, it is expected that there is greater variability between the Winter and Summer values, for both magnitudes. In figure 52 it is possible to visualize the values obtained from these parameters for the two periods.

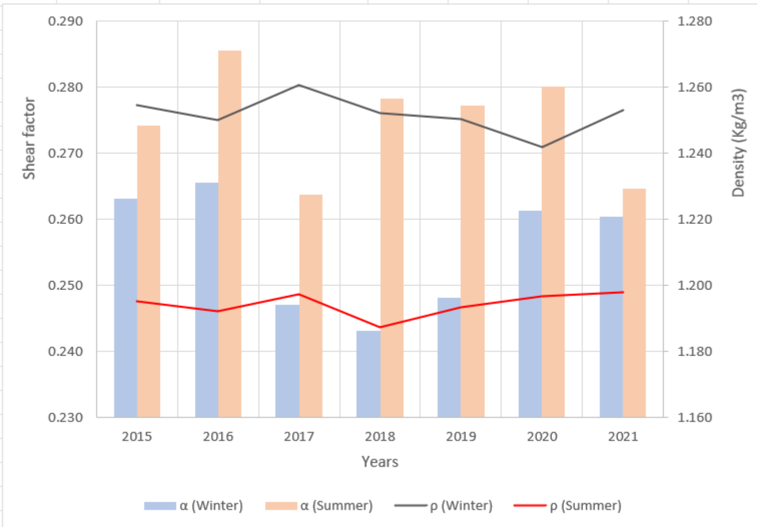

Figure 52: Shear factor and density values for Winter and Summer in each year

The lower density values in Summer than in Winter can be explained by the increase in temperature. An example to demonstrate this is the average temperature in the Winter of 2019 which was 4.5ºC with a density of 1.25 kg/m3, while the average temperature in the Summer of the same year was 16.3ºC with a density of 1.19 kg/m3. Even so, the average temperature in the hottest semester is still relatively low, compared to other countries further south in Europe.

In terms of AEP, the energy results for each semester and the respective annual values calculated directly from the time series are represented in table 17 below.

Table 17: Season energy values and AEP values for each year (Poland)

	Semester Energy (MWh)		**AEP (MWh)**
Year	Winter	Summer	
2015	9675	7499	17174
2016	9038	5922	14960
2017	9449	7703	17151
2018	8128	6672	14801
2019	9928	6932	16860
2020	9535	6575	16111
2021	8427	6499	14926

The years 2016, 2018 and 2021 had lower energy productions compared to the others.

The next step, related to the calculation of the Weibull parameters, was based on this half-yearly discretization. Therefore, for each method, two parameters were obtained for Winter and another pair for Summer. In table 18 there is an example of these results for the MLM method. The remaining results of the other methods can be consulted in the tables (C1.1-C1.4), of Appendix C.

Table 18: Weibull parameters calculated through MLM (Winter/Summer - Poland)

MLM parameters				
	Winter		Summer	
Years	k	A	k	A
2015	2.45	8.68	2.53	7.43
2016	2.64	8.36	2.43	6.67
2017	2.45	8.34	2.67	7.49
2018	2.54	8.23	2.51	7.03
2019	2.65	8.5	2.56	7.17
2020	3.14	8.99	2.59	6.98
2021	2.76	7.84	2.52	6.91

From this, it was possible to obtain the semiannual energy production values for each method, for each year. These values are in table 19.

Table 19: Winter and Summer energy values for each method and year (Poland)

	Semester energy (MWh)									
	MLM		EPFM		WAsP		EML		MOM	
Year	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer
2015	9687	7541	9775	7539	9614	7355	9703	7542	9703	7541
2016	9294	6064	9324	6007	8862	6033	9307	6059	9301	6062
2017	9143	7683	9256	7696	9045	7321	9173	7683	9166	7683
2018	9010	6763	9082	6734	8761	6617	9050	6762	9045	6763
2019	9543	7039	9596	7039	9615	6849	9704	7059	9704	7059
2020	10750	6649	10682	6621	9460	6430	10775	6648	10767	6649
2021	8396	6525	8409	6509	7849	6386	8396	6544	8412	6525

When performing the sum of the energy productions of the Winter months with the Summer months, the AEP is obtained in these discretizations of less than one year. Analogously to the other discretizations, it is possible to establish the variation of the AEP estimated from the statistical methods in relation to the AEP obtained directly from the Time Series. This variation can be seen in figure 53.

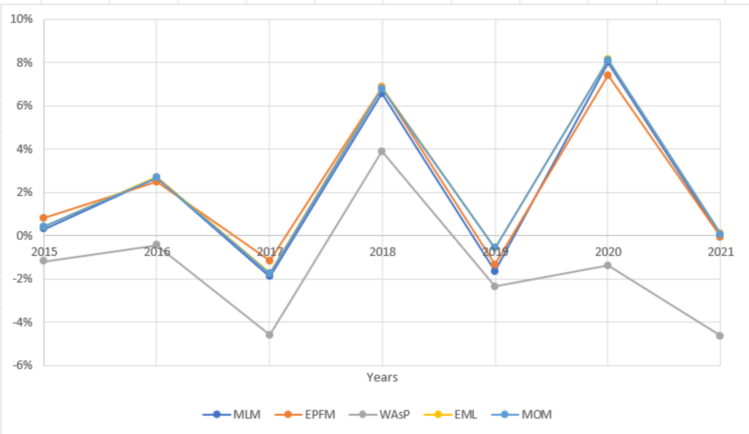

Figure 53: AEP variation between statistical methods with the Time Series, for each year (Winter/Summer discretization -Poland)

Through this graph, it can be inferred that in this discretization made, the WAsP method in three years is the one that has the lowest variability of AEP in relation to the calculation made directly by the Time series. MLM is the one that has the lowest variability in two years, while the rest of the methods only win, in terms of the lowest variability, for one year. This data can be found in table 20.

Table 20: AEP variation between statistical methods with the Time Series, for each year (Winter/Summer discretization-Poland)

	ΔAEP (Winter/summer)				
	MLM	EPFM	WASP	EML	MOM
2015	0.31%	0.81%	-1.19%	0.42%	0.41%
2016	2.67%	2.49%	-0.43%	2.72%	2.69%
2017	-1.89%	-1.16%	-4.58%	-1.72%	-1.76%
2018	6.57%	6.86%	3.90%	6.83%	6.80%
2019	-1.65%	-1.34%	-2.35%	-0.58%	-0.58%
2020	7.99%	7.40%	-1.37%	8.14%	8.10%
2021	-0.03%	-0.05%	-4.63%	0.09%	0.07%

Regarding the reliability criteria, only the determination coefficient will be used, as in the case of the previous discretization. For this, the years 2017 and 2018 were chosen as those to be analyzed in terms of R^2 values for Winter and Summer. Table 21 shows the results for these two years.

Table 21: Determination Coefficient values for each statistical method (Winter/Summer-Poland)

	R^2 Values			
	2017		2018	
	Winter	Summer	Winter	Summer
MLM	0.9777	0.9774	0.9840	0.9736
EPFM	0.9835	0.9576	0.9902	0.9729
WASP	0.9752	0.9576	0.9790	0.9729
EML	0.9786	0.9775	0.9852	0.9811
MOM	0.9713	0.9741	0.9849	0.9737

In this case, it is possible to notice that the reliability values are higher in the winter period than in the summer. In addition, the method that has the best fit in 2017 and 2018 in winter is the EPFM, while in the summer of the same years it is the EML. However, as the main analysis factor is the ability to reproduce the AEP from the WTS, the method that for 3 years has the lowest variability of energy production is WAsP.

5.1.4. Night/Day discretization

The last discretization is done in order to consider shear factor values and densities that are not annual averages, as well as the third discretization seen previously. In this case, the study approach as defined in section 4.2 is to characterize the wind series in periods of night and day.

The average annual speeds in the night and in the day, during the years 2015 to 2021 can be seen in Figure 54.

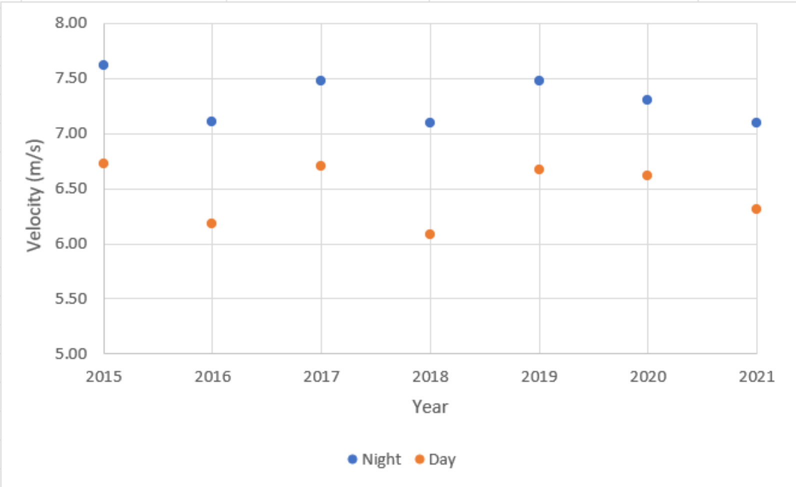

Figure 54: Average annual speed for night and day

Regarding the parameters of shear factor and air density, a difference in the values of these two magnitudes was expected due to the atmospheric stability present at night and because this period has lower temperatures. The result of values of this parameter can be seen in figure 55.

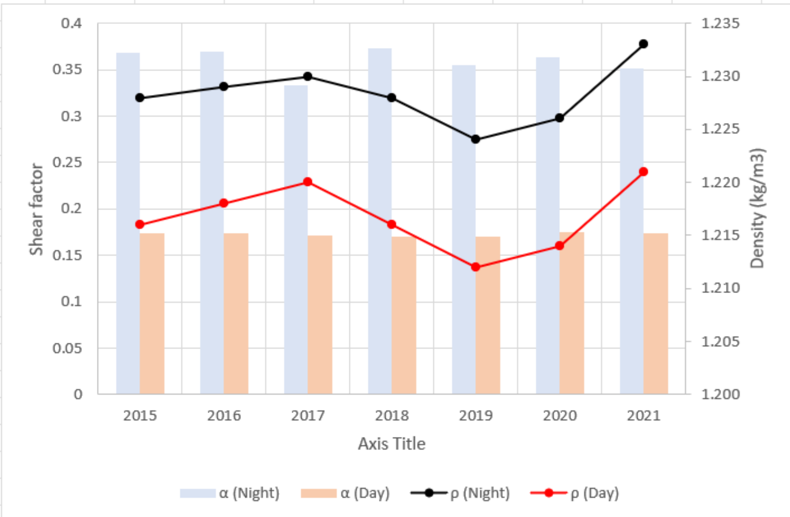

Figure 55: Shear factor and density values for Night and Day in each year

The averages of these parameters are shown in Table 22.

Table 22: Night and Day Annual average values for shear factor and air density

Shear factor		Density (kg/m3)	
Night	Day	Night	Day
0.359	0.173	1.228	1.217

Confirming the hypotheses, the mean value of the shear factor at night is greater than that of the day, about twice as much, due to the greater atmospheric stability in that period. The density is also higher at night, which justifies the presence of higher temperatures during the morning and, consequently, lower density.

When carrying out this discretization, the calculation of the AEP is given by the sum of the annual portions of energy from the night with the portions of energy from the day. The AEP data obtained directly from the Wind Time Series can be seen in Table 23.

Table 23: Night and Day energy values and AEP values for each year

	Energy (MWh)		
Year	Night	Day	AEP
2015	9765	7410	17174
2016	8559	6400	14959
2017	9492	7658	17151
2018	8590	6211	14800
2019	9477	7383	16860
2020	8913	7197	16110
2021	8393	6533	14925

The data obtained from AEP coincide with the values from the annual discretization and Winter/Summer. The next step is to obtain the estimates of the Weibull parameters for the night and day periods, for each statistical method. The result for MLM can be seen in Table 24. The remaining results of the other methods can be consulted in the tables (C1.5-C1.8), of Appendix C

Table 24: Weibull parameters calculated from MLM (Night and Day)

	MLM parameters			
	Night		Day	
	A	k	A	k
2015	8.21	2.58	7.68	2.14
2016	7.67	2.6	7.07	2.17
2017	8.19	2.68	7.7	2.29
2018	7.67	2.76	7.00	2.25
2019	8.04	2.78	7.60	2.34
2020	7.84	2.63	7.51	2.35
2021	7.42	2.42	7.00	2.13

From this, it was possible to obtain the semiannual energy production values for each method, for each year. These values are in table 25.

Table 25: Night and Day energy values for each method and year

	Night /Day energy (MWh)									
	MLM		EPFM		WAsP		EML		MOM	
Year	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day
2015	8996	7894	9266	8124	8675	8436	9054	7895	9049	7875
2016	8015	6857	8244	6980	7725	7263	8025	6837	8025	6837
2017	9013	7979	9114	8059	8556	8173	9048	7968	9043	7968
2018	8055	6725	8129	6728	7578	6980	8089	6723	8089	6706
2019	8781	7818	9034	8009	8191	7942	8839	7816	8835	7813
2020	8345	7660	8654	7907	7978	7743	8350	7639	8366	7619
2021	7507	7391	8202	7197	7412	7112	7482	6649	7481	6649

From these data, the AEP is obtained by adding the portions of the night with those of the day. The AEP variation obtained by statistical methods compared to the AEP data obtained directly by the Wind Time Series is shown in figure 56.

Figure 56: AEP variation between statistical methods with the Time Series, for each year (Night and Day discretization)

Through this graph, it is possible to see that with this type of night and day discretization, the variations in AEP estimated by statistical models no longer follow a pattern of behavior very well, as happened in other previous cases. Now, the WAsP method is not the only one that has different results from the others, but also the EPFM which is the only one that overestimates the production in all years.

In terms of lower AEP variability, that is, the ability to reproduce the AEP obtained directly from the wind series, the MLM wins in two years (2020 and 2021), the EPFM wins in the years 2017 and 2019, the WAsP model wins in 2015 and 2016 and MOM wins in 2018. This data can be seen in Table 26.

Table 26: AEP variation between statistical methods with the Time Series, for each year (Night/Day discretization)

ΔAEP (Night/Day)					
Year	MLM	EPFM	WASP	EML	MOM
2015	-1.66%	1.26%	-0.37%	-1.31%	-1.46%
2016	-0.59%	1.77%	0.19%	-0.65%	-0.65%
2017	-0.93%	0.13%	-2.46%	-0.78%	-0.81%
2018	-0.14%	0.39%	-1.64%	0.08%	-0.04%
2019	-1.54%	1.09%	-4.31%	-1.22%	-1.26%
2020	-0.65%	2.80%	-2.42%	-0.76%	-0.78%
2021	-0.18%	3.17%	-2.69%	-5.32%	-5.33%

For the reliability test, the one used for this discretization will also be the R^2 for the years 2019 and 2020 in all methods. Table 27 shows the results for these two years.

Table 27: Determination Coefficient values for each statistical method (Night/Day discretization)

	R^2 Values			
	2019		2020	
	Night	Day	Night	Day
MLM	0.9929	0.9867	0.9881	0.9893
EPFM	0.9911	0.9850	0.9944	0.9866
WASP	0.9726	0.9725	0.9735	0.9712
EML	0.9950	0.9707	0.9928	0.9702
MOM	0.9927	0.9699	0.9927	0.9662

In this case, it is possible to notice that the reliability values are higher in the night period than in the day. Although there are some methods with higher values for this criterion, the decisive factor as mentioned earlier in the other discretizations is the reproducibility of the Wind Series AEP. In this case, there is not a single method that varies less, but three, MLM, EPFM and WAsP.

Therefore, after these four discretizations previously analyzed, it is possible to make a comparison between them in terms of better adjustment results. In all these approaches, the only method that manages to beat the others in terms of more times with lower rates of change is the WAsP. Therefore, this method is the best choice of the Wind Time Series.

In addition, it is interesting to understand in which discretization of the four discussed, it manages to make the best estimates of energy production. For this, the AEP variations of the WAsP method were compared with the AEP of the WTS, in each discretization (see table 28).

Table 28: AEP variation for each discretization - Poland

	ΔAEP for each discretization		
	Annual	Winter/Summer	Night/Day
2015	0.53%	-1.19%	-0.37%
2016	1.13%	-0.43%	0.19%
2017	-2.17%	-4.58%	-2.46%
2018	-1.30%	3.90%	-1.64%
2019	-3.60%	-2.35%	-4.31%
2020	-1.37%	-1.37%	-2.42%
2021	0.41%	-4.63%	-2.69%

Through this table, it is possible to see that despite the Summer/Winter and Night/Day approach do not consider annual mean values for parameters such as shear factor, air density and Weibull parameters, these discretizations weren’t the most reliable. The annual approach was the most reliable because in seven years, it had the lowest AEP variation rates in four years.

5.2. Brazil

5.2.1. Long Measurement period (3 years)

The case study of Brazil, as mentioned in section 4.2, has a 3-year wind time series, from 20/06/2009 to 19/06/2012.

The turbine to be used will be the same as the case study in Poland, with a rated power of 4.2 MW (Appendix A) at a hub height of 105 m and with air velocities and densities extrapolated for that height.

Analogously to the case of Poland, the total energy of these 3 years was calculated, directly from the Wind Time Series, which is 62759 MWh. Therefore, the AEP represented each year was obtained by dividing this total value by the time interval of 3 years, finding a value of 20920 MWh.

The next step, after obtaining the AEP by the Time Series, is to estimate the Weibull parameters of each method (MLM, EPFM, WAsP, EML, and MOM) in order to perform the adjustment to the wind time series. The results of the shape and scale parameters for the five methods can be seen in Table 29.

Table 29: Weibull parameters (3 years)

	Weibull parameters	
Method	k	A
MLM	2.95	9.13
EPFM	2.89	9.14
WAsP	2.95	9.13
EML	2.94	9.13
MOM	2.93	9.13

As in the case of Poland, in this first study approach, the estimated values of the Weibull parameters show little variation between the methods. In figure 57, it is possible to see the variation of the parameters for each method in relation to their average value, which in the case of the shape parameter was 2.93 and the scale parameter, 9.13.

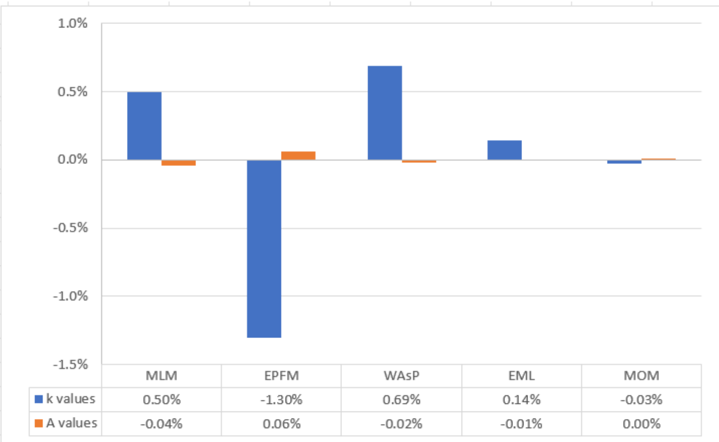

Figure 57: Weibull parameters variation against their mean value (3 years)

In this case, the methods with the largest shape parameter variations were EPFM and WAsP but it is not possible to draw any goodness of fit conclusion based on this analysis. The adjustment made by each method to the 3-year series can be seen in figure 58.

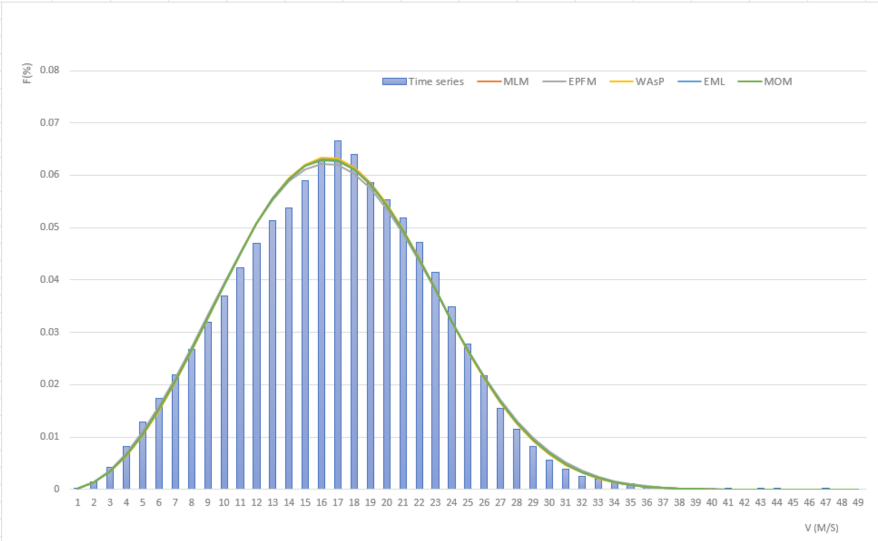

Figure 58: Weibull adjustment for each method related to the Time Series (3 years)

With the Weibull parameters already estimated, the next step was to calculate the AEP for each method. However, the value found is not the final one yet, as it is necessary to make the correction according to the availability of 99.11%. This correction is made according to equation 1 and only for the time series because for statistical methods it does not make sense to use this concept. The AEP's obtained by statistical methods and their variation with the AEP of the time series, after correction, are shown in table 30.

Table 30: AEP for each method and their variation against AEP from Time Series (3 years)

Method	AEP (MWh)	AEP Variation (%)
MLM	21036	-0.34%
EPFM	20973	-0.64%
WASP	21036	-0.34%
EML	21020	-0.41%
MOM	21004	-0.49%

According to this table, all methods perform underestimations of AEP compared to the value obtained by the time series and those with less variation are MLM and WAsP, with only 0.34% in both.

Regarding the goodness of fit criteria, the four mentioned in section 3 were used, with the results available in table 31.

Table 31: Accuracy test values for each method (3 years)

Statistical methods	The Weibull parameters		Accuracy Test Efficiency			
	k	A	RMSE	X^2	R	R^2
MLM	2.95	9.13	0.00231	0.00602	0.99695	0.99619
EPFM	2.89	9.14	0.00263	0.00762	0.99640	0.99550
WAsP	2.95	9.13	0.00231	0.00602	0.99695	0.99619
EML	2.94	9.13	0.00236	0.00620	0.99686	0.99612
MOM	2.93	9.13	0.00242	0.00641	0.99676	0.99603

This table shows that by presenting the same estimates of Weibull parameters, the MLM and WAsP methods have the same values in the goodness of fit criteria. These two methods are those with the best values, in the four criteria. Despite this, all statistical methods have good values, especially in the determination coefficient and in the correlation coefficient, since they are all close to 1.

The choice of the best method should consider the ability to reproduce the AEP obtained from the wind series, which characterizes the wind regime at the site. Contrary to what happens in the case of Poland, both for this reproducibility criterion and for the values obtained in the goodness of fit tests, the MLM and WAsP methods are the ones that produce the best fit for the WTS.

Now, it is necessary to check whether this pattern occurs for the other discretizations, mainly for the Winter/Summer and Night/Day approach.

5.2.2. Yearly discretization

For the second approach, the calculation of the AEP will be done for each year and not consider the sum of the energies of the three years and then divide by three to have the same AEP in all years. Therefore, there will be 3 different databases, discretizing an entire year.

Figure 59 represents the results of the AEP obtained directly from the Wind Time Series, in each year (blue) with the value obtained by the first energy analysis (section 5.2.1), in orange, also without statistical intervention. It is important to remember that the availability was obtained for each year (see table 32) and after that the AEP comparison was made.

Table 32: Availability for each year

Year	Availability	AEP_{before} (MWh)	$AEP_{corrected}$ (MWh)
2009/2010	97.89%	18944	19353
2010/2011	99.45%	21135	21251
2011/2012	99.99%	22679	22681

Figure 59: AEP values for each year versus AEP for 3 years calculation (Both from Time Series)

Through this figure, the AEP over the years has an increase around the AEP value of the first discretization, that is, having a more linear than oscillatory behavior around 20920 MWh. The average AEP obtained each year is also 20920 MWh, with a negligible error (15th decimal place).

In addition, one of the analysis factors is the shear factor, air density and the respective AEP directly from the Time Series, already mentioned above. Table 33 shows these values.

Table 33: Shear factor, Density and AEP calculated for each year (Brazil)

Year	Shear factor	Air density	AEP (MWh)
2009/10	0.074	1.155	19353
2010/11	0.068	1.157	21251
2011/12	0.073	1.157	22681

Only through these data it is not possible to obtain a direct relationship between the variation of these parameters with the values of AEP of the Time Series, being necessary to perform calculations with larger discretizations.

After that, Weibull parameters were estimated each year by all five methods also studied in the case of Poland. Table 34 shows the results of these estimates of the pair (k, A).

Table 34: Weibull parameters for each year and method (Brazil)

	MLM		EPFM		WAsP		EML		MOM	
Year	k	A	k	A	k	A	k	A	k	A
2009/2010	2.93	8.69	2.89	8.70	2.94	8.69	2.92	8.69	2.91	8.70
2010/2011	2.82	9.20	2.82	9.20	2.82	9.20	2.80	9.20	2.79	9.20
2011/2012	3.15	9.47	2.99	9.50	3.17	9.49	3.16	9.48	3.15	9.48

Based on these estimated parameters, the next step was to obtain the AEP for each year using these statistical methods. Table 35 presents these results.

Table 35: AEP for each method and year (Brazil)

	AEP (MWh)				
Year	MLM	EPFM	WASP	EML	MOM
2009/2010	19473	19457	19486	19486	19483
2010/2011	21048	21048	21048	21048	20996
2011/2012	22485	22295	22550	22535	22518

With the AEP data obtained by the statistical methods, it is now possible to see their ability to reproduce the annual energy production from the Time Series. This is done by analyzing the energy variation between the two approaches and then, represented in figure 60.

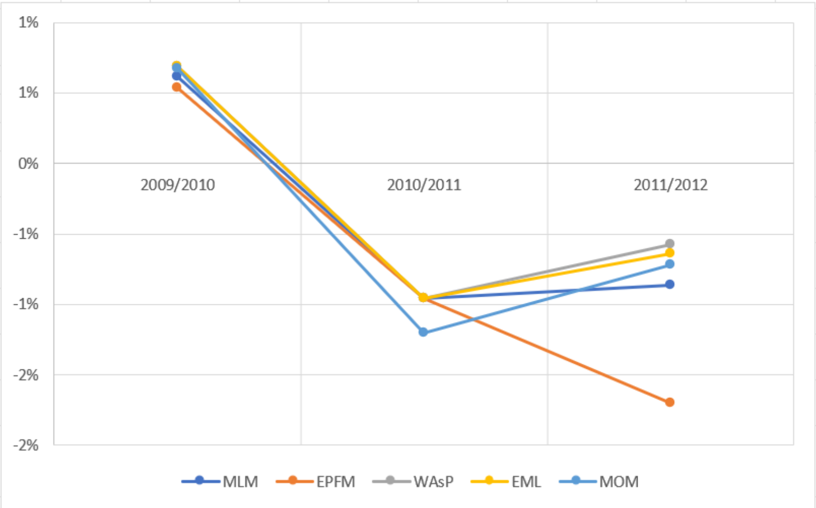

Figure 60: AEP variation comparing each statistical method with Time Series (Brazil)

These variations over the years can also be seen in Table 36.

Table 36: AEP variation for each method against AEP from Time Series

	ΔAEP (Annual)				
Year	MLM	EPFM	WASP	EML	MOM
2009/2010	0.62%	0.53%	0.68%	0.68%	0.67%
2010/2011	-0.96%	-0.96%	-0.96%	-0.96%	-1.20%
2011/2012	-0.86%	-1.70%	-0.58%	-0.64%	-0.72%

Through the graph and the table, it can be seen that the energy variations of the methods compared to the Wind Time Series are relatively low, which already indicates a good reproducibility of the AEP of the base series. Despite the variations between the methods being very low, it is already possible to analyze that the EPFM method and the WAsP are the ones that come closest to the AEP of the WTS in two years. The MLM and EML methods have the best reproducibility together with EPFM and WAsP, in the year of 2010/2011 and MOM, in no year manages to beat the other methods.

An example of adjustment can be seen in figure 61 for the year 2009/2010, with the fittings of each method made to the base case (WTS). Adjustments for other years can be seen in figure B2.1. (Appendix B).

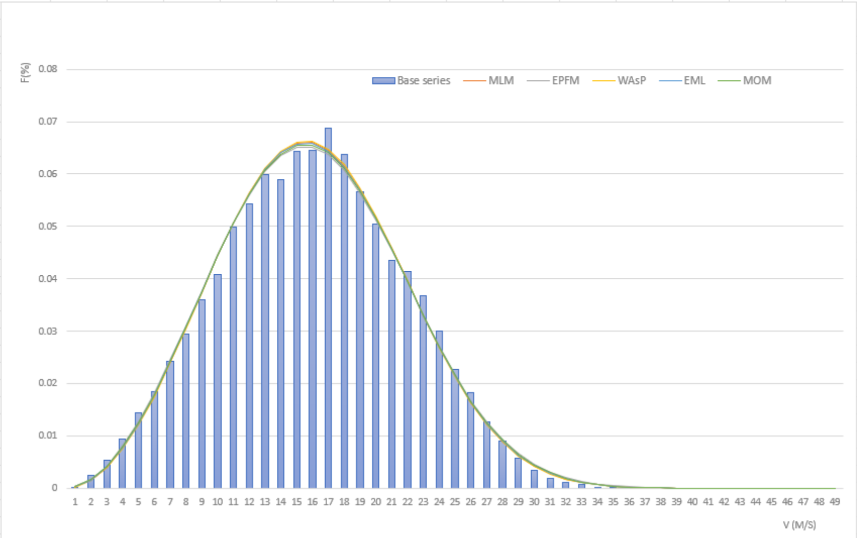

Figure 61: Weibull adjustment for each method related to the Time Series (Brazil – 2009/2010)

The next step was to perform goodness of fit criteria and understand whether the methods have good values. In terms of adjustment criteria, in this case where the discretization is yearly performed, it was decided to use the determination coefficient (R^2) criterion, as it is more commonly used in the industry. Table 37 shows the values found for each method and in each year.

Table 37: Determination Coefficient values for each statistical method (Yearly discretization- Brazil)

	R^2 Values				
Years	MLM	EPFM	WAsP	EML	MOM
2009/2010	0.9958	0.9952	0.9960	0.9957	0.9956
2010/2011	0.9911	0.9911	0.9911	0.9909	0.9907
2011/2012	0.9929	0.9860	0.9935	0.9932	0.9929

Through this table, it can be seen that the statistical methods adjust very well to the Wind series, since the values of this goodness criterion are all very close to 1. The non-significant differences between them and therefore, with this criterion, it is not possible to define a better adjustment method.

The choice is made based on the ability to play the Time Series AEP, which as mentioned earlier, the MLM and WAsP methods come closest.

5.2.3. Winter/Summer discretization

This discretization, as in the case of Poland, will separate the years of the series into Winter and Summer, in which the first refers to the months of April, May, June, July, August, and September. The second semester, called Summer, considers the six hottest months of the year (January, February, March, October, November, and December).

The first analysis obtained is relative to the monthly average speeds in each of the years, as can be seen in figure 62.

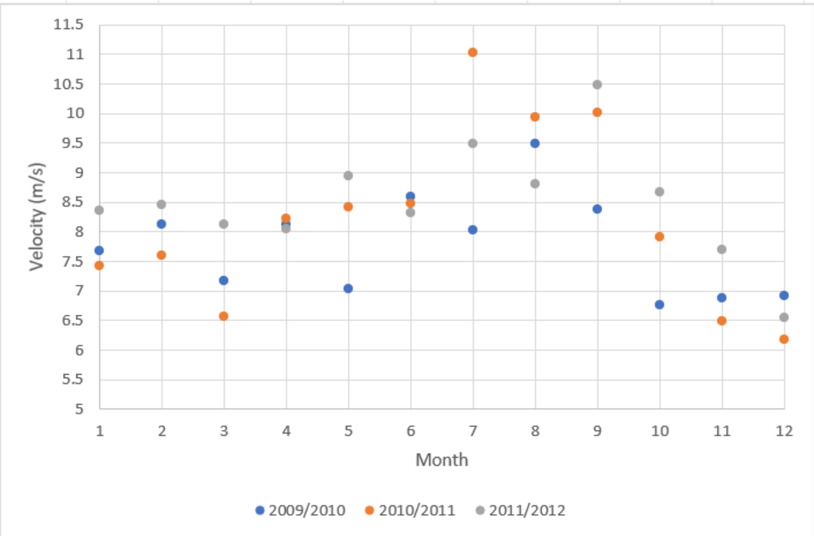

Figure 62: Monthly average velocity in each year for 105 m above ground level (Brazil)

Because Brazil is located in the southern hemisphere, the coldest period in the study region is in the months of July, August, and September, represented in the graph as the maximum speed points in all years. Seasonality in 2010/2011 is more noticeable than in other years.

Regarding the shear factor and density parameters, these values can be consulted in figure 63. As expected, the density values do not differ much when comparing winter and summer, due to the fact that the study site in Brazil does not have well-defined seasons. An example of this is that in 2011 the average temperature of the winter months was 25.6ºC, while in the summer it was 27ºC.

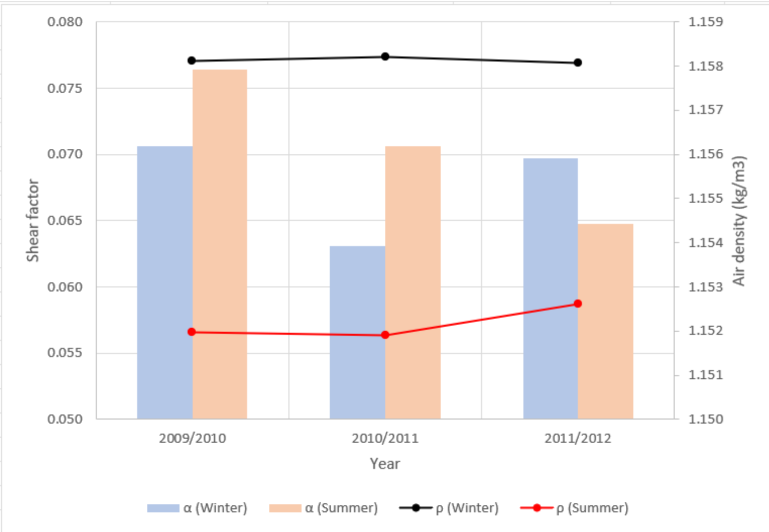

Figure 63: Shear factor and density values for Winter and Summer in each year (Brazil)

In terms of AEP obtained directly by the Time Series, table 38 shows the energy results of the winter and summer periods and the total (AEP). This table already accounts for the energy correction made by the availability of each year and in each discretization. The availability can be seen in table 39.

Table 38: Season energy values and AEP values for each year (Brazil)

	Semester Energy [MWh]		**AEP [MWh]**
Year	Winter	Summer	
2009/2010	10784	8603	19388
2010/2011	12971	8254	21225
2011/2012	12065	10616	22681

Table 39: Season availability (Brazil)

	Availability	
Year	Winter	Summer
2009/2010	96.25%	99.55%
2010/2011	100.00%	98.90%
2011/2012	99.99%	100.00%

The next step is to obtain the Weibull parameters through statistical methods. Therefore, for each method, two parameters were obtained for Winter and another pair for Summer. In table 40 there is an example of these results for the MLM method. The remaining results of the other methods can be consulted in the tables (C2.1-C2.4), of Appendix C.

Table 40: Weibull parameters calculated from MLM (Winter/Summer)

	MLM parameters			
	Winter		Summer	
Years	k	A	k	A
2009/2010	3.24	9.21	2.74	8.16
2010/2011	3.35	10.33	2.68	7.96
2011/2012	3.20	9.88	3.17	9.06

From this, it was possible to obtain the semiannual energy production values for each method, for each year. These values are in table 41.

Table 41: Winter and Summer energy values for each method and year (Brazil)

	Semester energy (MWh)									
	MLM		EPFM		WAsP		EML		MOM	
Year	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer
2009/2010	10874	8653	10746	8673	10866	8658	10851	8667	10868	8667
2010/2011	12721	8278	12482	8282	12735	8278	12683	8278	12697	8275
2011/2012	11925	10561	11771	10486	11934	10569	11944	10547	11934	10565

Therefore, the AEP is obtained from the sum of the winter and summer portions. The difference in the AEP obtained by statistical methods compared to the AEP resulting from the Time Series can be seen in figure 64 and table 42.

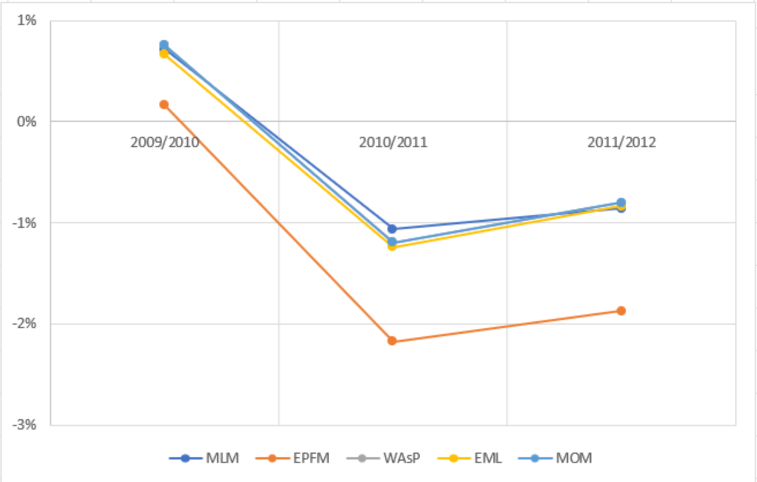

Figure 64: AEP variation between statistical methods with the Time Series, for each year (Winter/Summer discretization - Brazil)

Table 42:AEP variation between statistical methods with the Time Series, for each year (Winter/Summer discretization - Brazil)

	∆AEP (Winter/summer)				
	MLM	EPFM	WASP	EML	MOM
2009/2010	0.72%	0.16%	0.75%	0.67%	0.75%
2010/2011	-1.06%	-2.17%	-1.19%	-1.24%	-1.19%
2011/2012	-0.86%	-1.87%	-0.80%	-0.84%	-0.80%

Through this graph and this table, it is possible to see that the methods, in general, make good adjustments, and in some years, the methods have an error of less than 1% of overestimation or underestimation. In terms of comparison, all methods are best-fitting choices in at least one of the three years, except EML, which does not win in any of the years.

An interesting point is that in the first year, the EPFM method is the one that best fits the WTS. However, in the next two years, this method completely deviates from the other methods, having the greatest variations between them. This is due to the order of magnitude of the Weibull parameters estimated by him, since both in winter and in summer, of the second and third year, the values of the other four parameters are similar and that of the EPFM already produces other results.

Regarding the reliability criteria, only the determination coefficient will be used, as in the case of the previous discretization. Table 43 shows the results for these the three years, regarding Winter and Summer values.

Table 43: Determination Coefficient values for each statistical method (Winter/Summer - Brazil)

	R^2 Values					
	2009/2010		2010/2011		2011/2012	
	Winter	Summer	Winter	Summer	Winter	Summer
MLM	0.9957	0.9930	0.9886	0.9918	0.9901	0.9933
EPFM	0.9914	0.9936	0.9771	0.9921	0.9814	0.9870
WASP	0.9957	0.9930	0.9886	0.9918	0.9900	0.9933
EML	0.9955	0.9925	0.9872	0.9918	0.9904	0.9925
MOM	0.9956	0.9925	0.9873	0.9916	0.9900	0.9926

From this table, it can be seen that all methods have good adjustment values, and an interesting fact is the slight reduction in the value of the EPFM method in the year 2010/2011 more than in the other methods. As seen previously, from that year onwards, the method begins to increase the variation of AEP with the Time Series, making this change in value in this criterion plausible.

5.2.4. Night/Day discretization

As in the other case study, this last discretization will be done in order to consider shear factor values and densities that are not annual averages.

The average annual speeds in the night and in the day, during the three years can be seen in Figure 65.

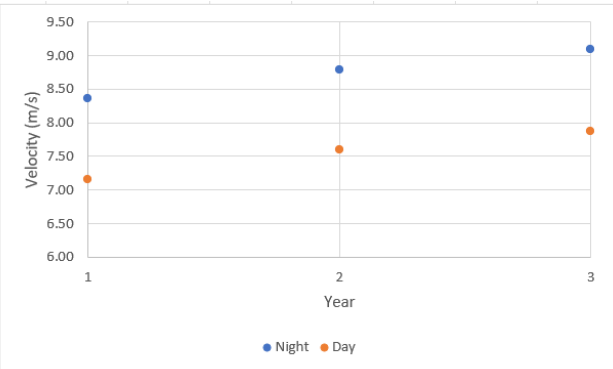

Figure 65: Average annual speed for night and day (Brazil)

Regarding the shear factor and air density parameters, the values obtained for this discretization approach are shown in figure 66.

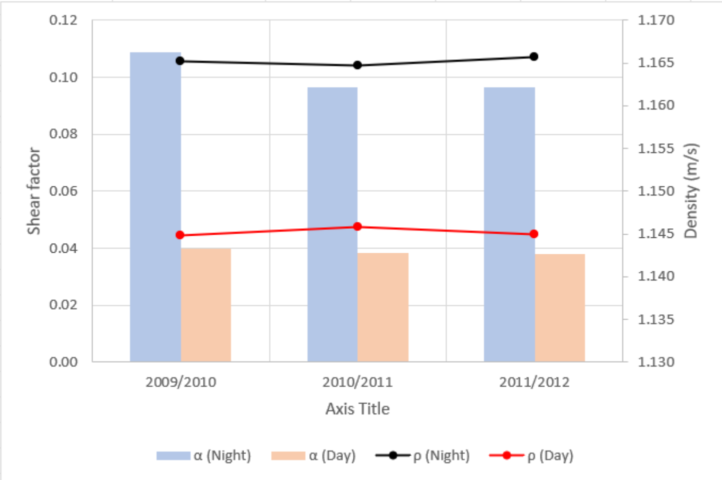

Figure 66: Shear factor and density values for Night and Day in each year (Brazil)

The averages of these parameters are shown in Table 44.

Table 44: Night and Day Annual average values for shear factor and air density (Brazil)

Shear factor		**Density (kg/m3)**	
Night	Day	Night	Day
0.101	0.039	1.165	1.145

In this case, in addition to the variation in the shear factor values, it is possible to see that there is no significant difference in the air density values when comparing the period of night with day. The climatic factor and the average high temperatures present in this location justify this scenario.

Before performing any AEP calculations directly from the Time Series, it is necessary to consider the availability of each day and night period, as shown in Table 45.

Table 45: Night and day availability (Brazil)

	Availability	
	Night	Day
2009/2010	97.80%	97.96%
2010/2011	99.45%	99.45%
2011/2012	100.00%	99.99%

Therefore, the energy values of each part (night/day), calculated directly from the WTS and their respective AEP are represented in table 46.

Table 46: Night and Day energy values and AEP values for each year (Brazil)

Energy (MWh)			
Year	Night	Day	AEP
2009/2010	11050	8307	19356
2010/2011	11994	9257	21251
2011/2012	12753	9928	22681

The next step is to obtain the estimates of the Weibull parameters for the night and day periods, for each statistical method. The result for MLM can be seen in Table 47. The remaining results of the other methods can be consulted in the tables (C2.5-C2.8), of Appendix C.

Table 47: Weibull parameters calculated from MLM (Night and Day- Brazil)

	MLM parameters			
	Night		Day	
	k	A	k	A
2009/2010	3.34	9.33	2.72	8.04
2010/2011	3.23	9.80	2.55	8.56
2011/2012	3.55	10.09	2.90	8.82

From this, it was possible to obtain the semiannual energy production values for each method, for each year. These values are in table 48.

Table 48: Night and Day energy values for each method and year (Brazil)

	Night /Day energy (MWh)									
	MLM		EPFM		WAsP		EML		MOM	
Year	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day
2009/2010	11265	8542	11076	8542	11216	8542	11192	8542	11192	8539
2010/2011	11930	9371	11782	9426	11945	9365	11907	9371	11897	9365
2011/2012	12666	10052	12384	10048	12718	10069	12708	10052	12708	10052

From these data, the AEP is obtained by adding the portions of the night with those of the day. The AEP variation obtained by statistical methods compared to the AEP data obtained directly by the Wind Time Series is shown in figure 67.

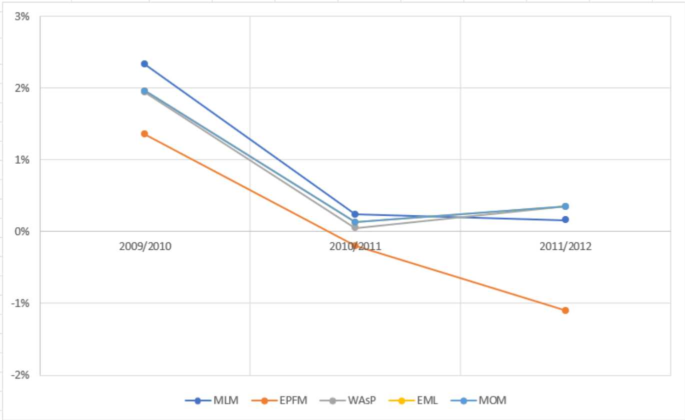

Figure 67: AEP variation between statistical methods with the Time Series, for each year (Night and Day discretization - Brazil)

According to this AEP variation analysis of the statistical methods in relation to the Time Series, the same of the previous discretization occurs in this one. The methods perform good adjustments in terms of being able to reproduce the WTS AEP, as the error found is a maximum of 2% of estimation.

Another point is that in this discretization, the MLM, EPFM and WAsP methods outperform the other methods in certain specific years. EML and MOM, despite good adjustments, are not the best among the five. Considering the analysis made in the other discretizations, except the first one, the EPFM and WAsP methods are the ones that have the greatest consistency in the reproduction of the Time Series AEP.

The variation described in figure 67 can be found in table 49.

Table 49: AEP variation between statistical methods with the Time Series, for each year (Night and Day discretization - Brazil)

	ΔAEP (Night/Day)				
	MLM	EPFM	WASP	EML	MOM
2009/2010	2.33%	1.35%	1.94%	1.96%	1.96%
2010/2011	0.24%	-0.20%	0.05%	0.13%	0.13%
2011/2012	0.16%	-1.10%	0.35%	0.35%	0.35%

Regarding the criterion used in the last discretizations, of the determination coefficient, the values obtained for Night and day, in each year are in table 50.

Table 50: Determination Coefficient values for each statistical method (Night/Day - Brazil)

	R^2 Values					
	2009/2010		2010/2011		2011/2012	
	Winter	Summer	Winter	Summer	Winter	Summer
MLM	0.9987	0.9924	0.9880	0.9890	0.9915	0.9956
EPFM	0.9920	0.9902	0.9766	0.9847	0.9824	0.9877
WASP	0.9933	0.9965	0.9869	0.9910	0.9961	0.9933
EML	0.9957	0.9914	0.9822	0.9855	0.9916	0.9934
MOM	0.9956	0.9925	0.9871	0.9870	0.9904	0.9883

Contrary to what happened in the first case study, the four different discretizations for Brazil failed to converge to a more reliable method. Among the options, those that had smaller variations, that is, that best managed to reproduce Wind Time's AEP were MLM, EPFM and WAsP.

For these three methods, the discretization that produced smaller AEP rates of change with WTS was Night/Day.

6. Conclusion and further works

In the course of this dissertation, several statistical methods were approached to obtain the Weibull parameters used in the calculation of AEP. In fact, the main factor that was present in the case studies and in all the discretizations was the AEP, which served as a parameter for comparison with the energy production coming directly from the Wind Time Series.

As explained in this work, the calculation of AEP to estimate energy production does not always condition the good adjustment of statistical methods to WTS, simply because parameters such as shear factor and air density are considered on an annual average basis. This can lead to estimation values with significant errors in the energy production of a turbine, in addition to not considering seasonality phenomena over the course of a year.

In general terms, both in the case of Poland and Brazil, the statistical methods obtained a good adjustment, and the estimated Weibull parameters, for all discretizations, were higher in Brazil than in Poland. In addition, the AEP variation interval of each method in relation to the AEP obtained directly from the wind series was greater in the first case than in the second.

For the Polish case study, the statistical methods were generally able to reproduce the energy obtained directly from the WTS with few errors, in which this reproducibility was between -3 and 3%. Considering that the first discretization does not consider any seasonality phenomenon, in addition to assigning the same AEP to all years, the results obtained from it were not considered relevant for the study, only to serve as a comparison.

Throughout the discretizations, there were methods that had good WTS energy reproducibility rates such as MLM and WAsP. However, in terms of lower AEP variation rates when compared to the energy obtained directly from the Time Series, the WAsP method was the most suitable. In addition, contrary to the hypothesis that perhaps a greater discretization of the wind series can contribute to better results, the best discretization for the WAsP method was the annual one.

For the WAsP method and the MLM, the best discretization in terms of AEP reproduction of the wind series was the annual one, while with other methods, the one with the best results was Night/Day and Winter/Summer. This may be related to the different combinations of shear factors and density for each year, as for the Winter/Summer discretization there is a large difference in values between the average density of each period due to the average temperatures, but there is no significant difference of shear factor. In the case of the Night/Day discretization, there is a big difference between the shear factors of the night in relation to the day, but not for the densities.

For the case study of Brazil, the statistical methods had less variability of results than that of Poland. A possible explanation for this may be related to the local climate, which is more stable in terms of density variation and shear factor than in Poland. Thus, the Weibull parameters estimated by each method are better adjusted to the wind series.

Another factor that may explain the better adjustments in the case of Brazil is that no single method was obtained as the most reliable to reproduce the WTS energy production, in which AEP's reproducibility were even smaller than Poland's, with -2.5 to 2.5%. In the end, three methods were more suitable, namely MLM, EPFM and also WAsP. Regarding the discretization, the option that best adjusted to WTS in terms of energy production reproducibility was night/day. This contributes to the idea that the adjustment of different shear factors to a wind series can contribute to better energy production results.

A method present in both case studies as the one with the best ability to reproduce the AEP obtained directly from the wind series is the WAsP. This can be explained by the fact that this method is the only one that presents a double validation of data with an iterative process of convergence of Weibull parameters.

The results obtained were important to understand the influence of climate and the atmospheric parameters derived from it for estimating the energy production of a wind turbine. Of course, statistical methods also play a role in influencing the adjustment, depending on your mathematical model, but it is essential to consider the sensitivity of the shear factor variability and air density to estimate the AEP.

This study paves the way for future research that may explore aspects not yet addressed, such as assessing the uncertainties with both time series and statistical approaches in estimating wind turbine energy yield, related to orography, soil roughness and associated obstacles. Another interesting topic that can be derived from the study done is applying time series and statistical models to estimate energy yield for various types of wind turbines.

References

Abbes, M., & Belhadj, J. (2014). Development of a methodology for wind energy estimation and wind park design. *Journal of Renewable and Sustainable Energy*, *6*. https://doi.org/10.1063/1.4895919

ADB, A. D. B. (Ed.). (2014). *Guidelines for Wind Resource Assessment: Best Practices for Countries Initiating Wind Development*.

Akdağ, S. A., & Dinler, A. (2009). A new method to estimate Weibull parameters for wind energy applications. *Energy Conversion and Management*, *50*(7), 1761–1766. https://doi.org/10.1016/j.enconman.2009.03.020

Akdağ, S. A., & Güler, Ö. (2015). A novel energy pattern factor method for wind speed distribution parameter estimation. *Energy Conversion and Management*, *106*, 1124–1133. https://doi.org/10.1016/j.enconman.2015.10.042

Al-Quraan, A., Al-Mhairat, B., Malkawi, A. M. A., Radaideh, A., & Al-Masri, H. M. K. (2023). Optimal Prediction of Wind Energy Resources Based on WOA—A Case Study in Jordan. *Sustainability (Switzerland)*, *15*(5). Scopus. https://doi.org/10.3390/su15053927

Alrashidi, M. (2023). Estimation of Weibull Distribution Parameters for Wind Speed Characteristics Using Neural Network Algorithm. *Computers, Materials & Continua*, *75*(1), 1073–1088. https://doi.org/10.32604/cmc.2023.036170

Badger, M., Peña, A., Hahmann, A. N., Mouche, A. A., & Hasager, C. B. (2016). Extrapolating Satellite Winds to Turbine Operating Heights. *Journal of Applied Meteorology and Climatology*, *55*(4), 975–991. https://doi.org/10.1175/JAMC-D-15-0197.1

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chong, C. H., & Ragai Henry Rigit, A. (2021). Wind Speed Statistical Analysis and Energy Assessment for Pulau Triso, Sarawak, Malaysia. *2021 International Conference on*

Green Energy, Computing and Sustainable Technology (GECOST), 1–6. https://doi.org/10.1109/GECOST52368.2021.9538687

Dayal, K. K., Cater, J. E., Kingan, M. J., Bellon, G. D., & Sharma, R. N. (2021). Wind resource assessment and energy potential of selected locations in Fiji. *Renewable Energy*, *172*, 219–237. https://doi.org/10.1016/j.renene.2021.03.034

Dupont, E., Koppelaar, R., & Jeanmart, H. (2017). Global available wind energy with physical and energy return on investment constraints. *Applied Energy*, *209*. https://doi.org/10.1016/j.apenergy.2017.09.085

Edrisian, A., Samani, H., Sharifan, A., & Naseh, M. R. (2013). The New Hybrid Model of Compressed Air for Stable Production of Wind Farms. *International Journal of Emerging Technology and Advanced Engineering*, *3*, 37–43.

Guerrero-Villar, F., Dorado-Vicente, R., Fike, M., & Torres-Jiménez, E. (2019). Influence of ambient conditions on wind speed measurement: Impact on the annual energy production assessment. *Energy Conversion and Management*, *195*, 1111–1123. Scopus. https://doi.org/10.1016/j.enconman.2019.05.067

Gugliani, G. K., Sarkar, A., Ley, C., & Mandal, S. (2018). New methods to assess wind resources in terms of wind speed, load, power and direction. *Renewable Energy*, *129*, 168–182. https://doi.org/10.1016/j.renene.2018.05.088

Hosius, E., Seebaß, J. V., Wacker, B., & Schlüter, J. Chr. (2023). The impact of offshore wind energy on Northern European wholesale electricity prices. *Applied Energy*, *341*, 120910. https://doi.org/10.1016/j.apenergy.2023.120910

Jani, H. K., Nagababu, G., & Prasad, K. M. (2023). *Enhancing offshore wind resource assessment with LIDAR-validated reanalysis datasets: A case study in Gujarat, India | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.ijft.2023.100320

Jiménez, J. J., Tzianoumis, A., Stokes, L., Yang, Q., & Livina, V. N. (2023). Long-term wind and solar energy generation forecasts, and optimisation of Power Purchase Agreements. *Energy Reports*, *9*, 292–302. https://doi.org/10.1016/j.egyr.2022.11.175

Jong, P. (n.d.). *Location of the wind farm clusters in the Northeast subsystem...* ResearchGate. Retrieved 4 June 2023, from https://www.researchgate.net/figure/Location-of-the-wind-farm-clusters-in-the-Northeast-subsystem-of-Brazil-Key-BA-Bahia_fig23_319991149

JRC, D., Shaw, S., & Peteves, S. (2006). *Bridging the European Wind Energy Market and a Future Renewable Hydrogen-Inclusive Economy*.

Kang, D., Ko, K., & Huh, J. (2015). Determination of extreme wind values using the Gumbel distribution. *Energy, 86*, 51–58. https://doi.org/10.1016/j.energy.2015.03.126

Kang, S., Khanjari, A., You, S., & Lee, J.-H. (2021). Comparison of different statistical methods used to estimate Weibull parameters for wind speed contribution in nearby an offshore site, Republic of Korea. *Energy Reports*, *7*, 7358–7373. https://doi.org/10.1016/j.egyr.2021.10.078

Kavak Akpinar, E., & Akpinar, S. (2005). A statistical analysis of wind speed data used in installation of wind energy conversion systems. *Energy Conversion and Management*, *46*(4), 515–532. https://doi.org/10.1016/j.enconman.2004.05.002

Lang, S., & Mckeogh, E. (2011). LIDAR and SODAR Measurements of Wind Speed and Direction in Upland Terrain for Wind Energy Purposes. *Remote Sensing*, *3*, 1871–1901. https://doi.org/10.3390/rs3091871

Liang, Y., Wu, C., Ji, X., Zhang, M., Li, Y., He, J., & Qin, Z. (2022). Estimation of the influences of spatiotemporal variations in air density on wind energy assessment in China based on deep neural network. *Energy*, *239*, 122210. https://doi.org/10.1016/j.energy.2021.122210

Lopez-Villalobos, C. A., Martínez-Alvarado, O., Rodriguez-Hernandez, O., & Romero-Centeno, R. (2022). Analysis of the influence of the wind speed profile on wind power production. *Energy Reports*, *8*, 8079–8092. https://doi.org/10.1016/j.egyr.2022.06.046

Lysen, E. H. (1983). *Introduction to wind energy*.

Matos, J. C. (2022). *Perdas de produção de eletricidade em parques eólicos: Expectativas vs. realidade*. INEGI. http://www.inegi.pt/pt/noticias/perdas-de-producao-de-eletricidade-em-parques-eolicos-expectativas-vs-realidade/

Measnet_SiteAssessment_V2.0.pdf. (n.d.). Retrieved 8 June 2023, from https://www.measnet.com/wp-content/uploads/2016/05/Measnet_SiteAssessment_V2.0.pdf

Mortensen, N. G. (1994). *Wind Measurements for Wind Energy Applications—A Review*.

Mukaka, M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal : The Journal of Medical Association of Malawi*, *24*(3), 69–71.

Öztürk, B., Hassanein, A., Akpolat, M. T., Abdulrahim, A., Perçin, M., & Uzol, O. (2023). On the wake characteristics of a model wind turbine and a porous disc: Effects of freestream turbulence intensity. *Renewable Energy*, *212*, 238–250. Scopus. https://doi.org/10.1016/j.renene.2023.05.002

Porté-Agel, F., Bastankhah, M., & Shamsoddin, S. (2020). Wind-Turbine and Wind-Farm Flows: A Review. *Boundary-Layer Meteorology*, *174*. https://doi.org/10.1007/s10546-019-00473-0

Roibas-Millan, E., Cubas, J., & Pindado, S. (2017). Studies on Cup Anemometer Performances Carried out at IDR/UPM Institute. Past and Present Research. *Energies*, *10*(11), Article 11. https://doi.org/10.3390/en10111860

Shanmugaratnam, S., Yogenthiran, E., Koodali, R., Ravirajan, P., Velauthapillai, D., & Shivatharsiny, Y. (2021). Recent Progress and Approaches on Transition Metal Chalcogenides for Hydrogen Production. *Energies*, *14*, 8265. https://doi.org/10.3390/en14248265

Shende, V., Patidar, H., Baredar, P., & Agrawal, M. (2023). Estimation of wind characteristics at different topographical conditions using doppler remote sensing instrument—A comparative study using optimization algorithm. *Environmental Science and Pollution Research*, *30*(16), 48587–48603. https://doi.org/10.1007/s11356-023-25689-z

Siqueira, J., Feiden, A., & Gurgacz, F. (n.d.). *Analysis of the energetic dynamism between solar and wind energy available in the south of Brazil*. Retrieved 2 June 2023, from https://www.researchgate.net/publication/314175018_Analysis_of_the_energetic_dynamism_between_solar_and_wind_energy_available_in_the_south_of_Brazil/link/59a0b90d0f7e9b0fb8994561/download

SPOT hoje | OMIE. (n.d.). Retrieved 10 June 2023, from https://www.omie.es/pt/spot-hoy

Stevens, W., Squier, W., Mitchell, W., Gullett, B., & Pressley, C. (2013). Measurement of motion corrected wind velocity using an aerostat lofted sonic anemometer. *Atmospheric Measurement Techniques Discussions*, *6*, 703–720. https://doi.org/10.5194/amtd-6-703-2013

The annual variability of wind speed. (n.d.). Wind Energy The Facts. Retrieved 16 June 2023, from https://www.wind-energy-the-facts.org/the-annual-variability-of-wind-speed.html

Werapun, W., Tirawanichakul, Y., & Waewsak, J. (2017). Wind Shear Coefficients and their Effect on Energy Production. *Energy Procedia*, *138*, 1061–1066. https://doi.org/10.1016/j.egypro.2017.10.111

Yaniktepe, B., Kara, O., Aladag, I., & Ozturk, C. (2023). Comparison of eight methods of Weibull distribution for determining the best-fit distribution parameters with wind data measured from the met-mast. *Environmental Science and Pollution Research*, *30*(4), 9576–9590. https://doi.org/10.1007/s11356-022-22777-4

7. APPENDIX A: Power curves of the chosen wind turbine

	Air density [kg/m³]													
Wind speed [m/s]	**1.225**	**0.95**	**0.975**	**1.0**	**1.025**	**1.05**	**1.075**	**1.1**	**1.125**	**1.15**	**1.175**	**1.2**	**1.25**	**1.275**
3.0	**81**	51	54	57	60	62	65	68	70	73	76	79	84	87
3.5	**172**	123	127	132	136	141	145	150	154	159	163	168	177	181
4.0	**285**	210	217	224	231	238	244	251	258	265	272	278	292	299
4.5	**424**	318	328	337	347	357	366	376	386	395	405	415	434	444
5.0	**597**	452	465	478	492	505	518	531	544	557	571	584	610	623
5.5	**809**	616	633	651	669	686	704	721	739	757	774	792	827	844
6.0	**1062**	813	835	858	881	904	926	949	972	995	1017	1040	1085	1108
6.5	**1361**	1045	1074	1103	1131	1160	1189	1218	1247	1275	1304	1332	1389	1418
7.0	**1709**	1317	1353	1389	1425	1461	1496	1532	1568	1603	1639	1674	1744	1779
7.5	**2101**	1628	1671	1715	1758	1802	1845	1888	1931	1974	2016	2058	2143	2185
8.0	**2545**	1982	2034	2086	2137	2189	2240	2292	2343	2394	2444	2494	2594	2644
8.5	**3014**	2375	2435	2496	2556	2616	2674	2732	2790	2848	2904	2959	3067	3120
9.0	**3458**	2791	2856	2921	2986	3052	3112	3172	3232	3292	3348	3403	3510	3562
9.5	**3807**	3180	3246	3312	3377	3443	3499	3556	3613	3669	3715	3761	3845	3884
10.0	**4038**	3543	3602	3662	3722	3781	3824	3866	3909	3951	3980	4009	4059	4079
10.5	**4143**	3842	3884	3926	3969	4012	4035	4059	4083	4107	4119	4131	4150	4158
11.0	**4191**	4055	4078	4100	4122	4145	4154	4162	4171	4180	4184	4187	4193	4195
11.5	**4199**	4152	4160	4168	4176	4185	4188	4190	4193	4196	4197	4198	4199	4200
12.0	**4200**	4185	4188	4191	4194	4198	4198	4199	4199	4200	4200	4200	4200	4200
12.5	**4200**	4197	4197	4198	4199	4200	4200	4200	4200	4200	4200	4200	4200	4200
13.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
13.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
14.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
14.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
15.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
15.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
16.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
16.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
17.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
17.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
18.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
18.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
19.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
19.5	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
20.0	**4200**	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200	4200
20.5	**4186**	4186	4186	4186	4186	4186	4186	4186	4186	4186	4186	4186	4186	4186
21.0	**3870**	3870	3870	3870	3870	3870	3870	3870	3870	3870	3870	3870	3870	3870
21.5	**3373**	3373	3373	3373	3373	3373	3373	3373	3373	3373	3373	3373	3373	3373
22.0	**2745**	2745	2745	2745	2745	2745	2745	2745	2745	2745	2745	2745	2745	2744
22.5	**2154**	2154	2154	2154	2154	2154	2154	2154	2154	2154	2154	2154	2154	2154
23.0	**1805**	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805
23.5	**1526**	1526	1526	1526	1526	1526	1526	1526	1526	1526	1526	1526	1526	1526
24.0	**1283**	1283	1283	1283	1283	1283	1283	1283	1283	1283	1283	1283	1283	1283
24.5	**1116**	1116	1116	1116	1116	1116	1116	1116	1116	1116	1116	1116	1116	1116

Figure A1: Power curves of the chosen wind turbine

8. APPENDIX B: Weibull adjustment to the time series

B.1. Poland

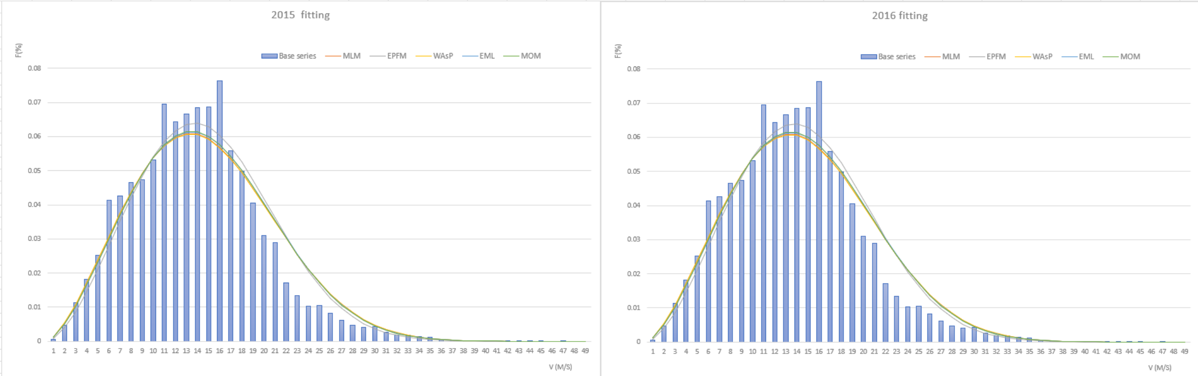

Figure B1.1: (a) 2015 and (b) 2016 adjustments

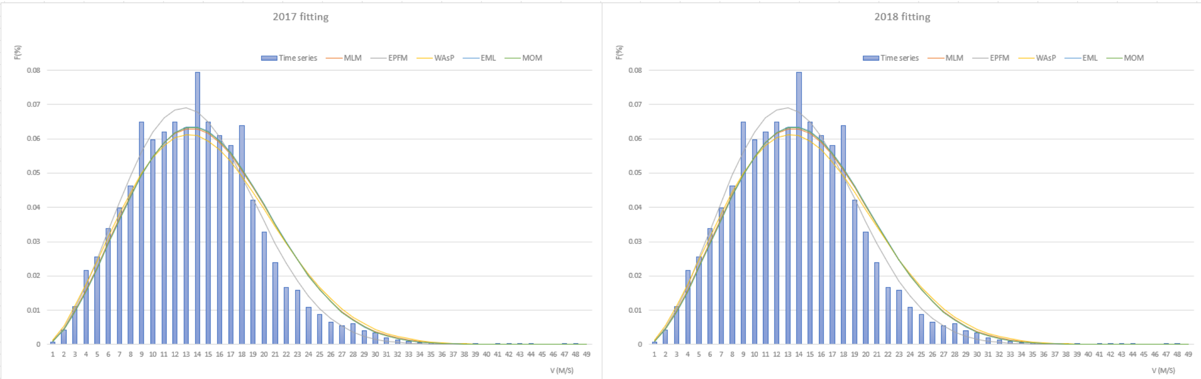

Figure B1.2: (a) 2017 and (b) 2018 adjustments

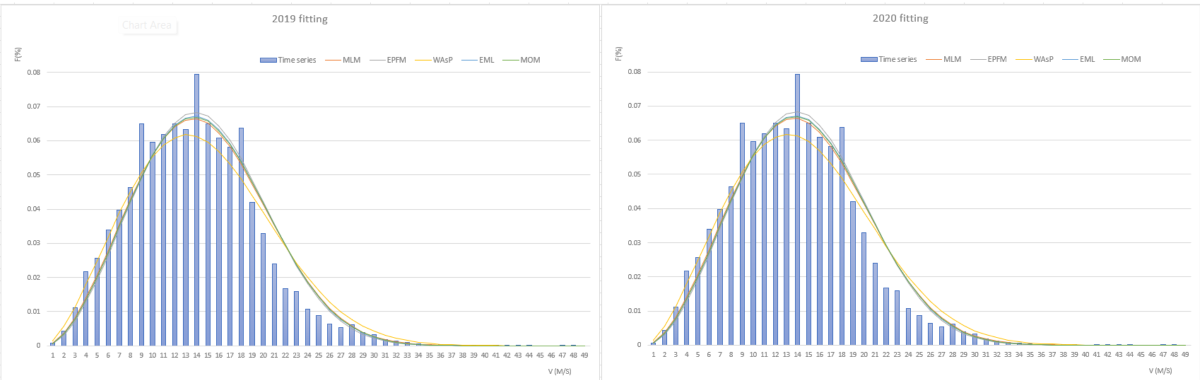

Figure B1.3: (a) 2019 and (b) 2020 adjustments

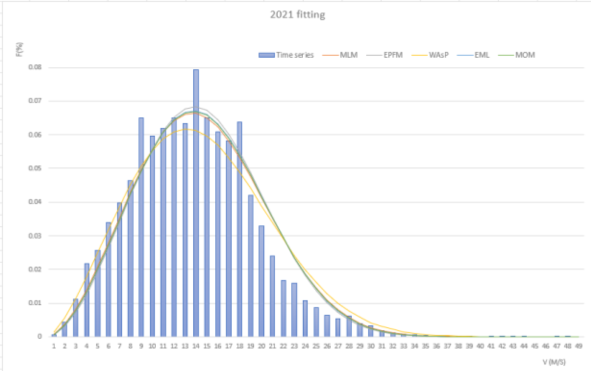

Figure B1.4: 2021 adjustment

B.2. Brazil

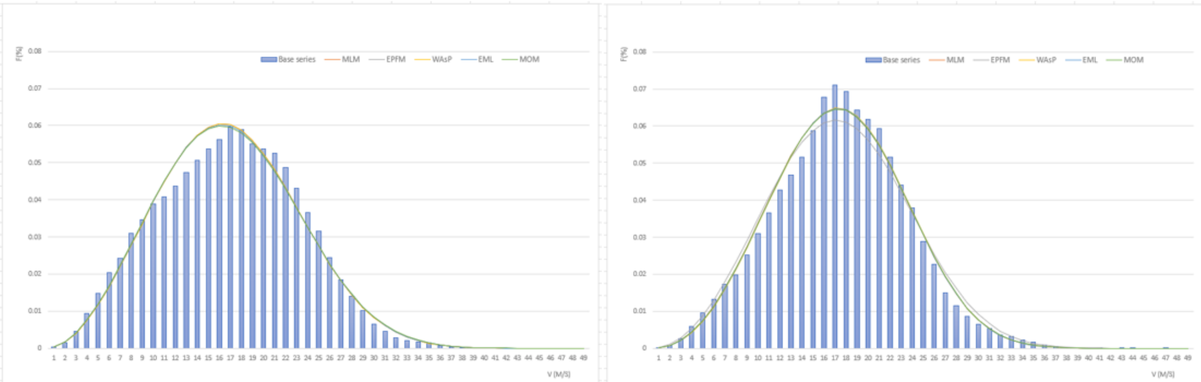

Figure B2.1.: (a)2010/2011 and (b) 2011/12 adjustments

9. APPENDIX C: Weibull parameters for the different discretizations

C.1. Poland

Table C1.1: Weibull parameters calculated from EPFM (Winter/Summer)

	EPFM parameters			
	Winter		Summer	
Years	k	A	k	A
2015	2.58	8.66	2.40	7.34
2016	2.72	8.19	2.40	6.65
2017	2.66	8.30	2.40	7.32
2018	2.67	8.13	2.40	6.95
2019	2.73	8.66	2.40	7.07
2020	2.99	8.57	2.40	6.86
2021	2.80	7.61	2.40	6.83

Table C1.2.: Weibull parameters calculated from WAsP (Winter/Summer)

	WAsP parameters			
	Winter		Summer	
Years	k	A	k	A
2015	2.40	8.66	2.40	7.34
2016	2.40	8.19	2.40	6.65
2017	2.40	8.30	2.40	7.32
2018	2.40	8.13	2.40	6.95
2019	2.40	8.66	2.40	7.07
2020	2.40	8.57	2.40	6.86
2021	2.40	7.61	2.40	6.83

Table C1.3.: Weibull parameters calculated from EML(Winter/Summer)

	EML parameters			
	Winter		Summer	
Years	k	A	k	A
2015	2.47	8.68	2.54	7.43
2016	2.66	8.36	2.45	6.67
2017	2.47	8.35	2.67	7.49
2018	2.58	8.24	2.53	7.03
2019	2.47	8.68	2.59	7.18
2020	3.15	9.00	2.60	6.98
2021	2.76	7.84	2.53	6.92

Table C1.4.: Weibull parameters calculated from MOM (Winter/Summer)

	MOM parameters			
	Winter		Summer	
Years	k	A	k	A
2015	2.47	8.68	2.53	7.43
2016	2.65	8.36	2.44	6.67
2017	2.46	8.35	2.67	7.49
2018	2.57	8.24	2.52	7.03
2019	2.47	8.68	2.58	7.18
2020	3.14	9.00	2.59	6.98
2021	2.75	7.85	2.52	6.91

Table C1.5: Weibull parameters calculated from EPFM (Night/Day)

	EPFM parameters			
	Night		Day	
	A	k	A	k
2015	8.30	2.78	7.76	2.41
2016	7.76	2.80	7.14	2.42
2017	8.22	2.77	7.71	2.49
2018	7.70	2.82	7.01	2.44
2019	8.14	2.92	7.67	2.58
2020	7.96	2.85	7.61	2.63
2021	7.72	2.95	7.25	2.61

Table C1.6: Weibull parameters calculated from WAsP (Night/Day)

	WASP parameters			
	Night		Day	
Year	A	k	A	k
2015	8.08	2.40	7.94	2.40
2016	7.54	2.40	7.29	2.40
2017	8.01	2.40	7.79	2.40
2018	7.46	2.40	7.14	2.40
2019	7.80	2.40	7.66	2.40
2020	7.68	2.40	7.55	2.40
2021	7.37	2.40	7.21	2.40

Table C1.7: Weibull parameters calculated from EML (Night/Day)

	EML parameters			
	Night		Day	
Year	A	k	A	k
2015	8.21	2.69	7.67	2.19
2016	7.66	2.72	7.06	2.21
2017	8.19	2.75	7.69	2.31
2018	7.68	2.82	7.00	2.27
2019	8.04	2.92	7.59	2.40
2020	7.82	2.76	7.49	2.43
2021	7.39	2.68	6.96	2.27

Table C1.8: Weibull parameters calculated from MOM (Night/Day)

	MOM parameters			
	Night		Day	
Year	A	k	A	k
2015	8.21	2.68	7.66	2.18
2016	7.66	2.72	7.06	2.20
2017	8.19	2.74	7.69	2.31
2018	7.68	2.82	6.99	2.26
2019	8.04	2.91	7.59	2.39
2020	7.83	2.75	7.48	2.42
2021	7.39	2.67	6.96	2.27

C.2. Brazil

Table C2.1.: Weibull parameters calculated from EPFM (Winter/Summer)

	EPFM parameters			
	Winter		Summer	
Years	k	A	k	A
2009/2010	3.02	9.24	2.78	8.16
2010/2011	3.07	10.38	2.75	7.95
2011/2012	3.01	9.91	3.00	9.09

Table C2.2.: Weibull parameters calculated from WAsP (Winter/Summer)

	WAsP parameters			
	Winter		Summer	
Years	k	A	k	A
2009/2010	3.23	9.21	2.75	8.16
2010/2011	3.35	10.34	2.68	7.96
2011/2012	3.21	9.88	3.18	9.06

Table C2.3.: Weibull parameters calculated from EML (Winter/Summer)

	EML parameters			
	Winter		Summer	
Years	k	A	k	A
2009/2010	3.21	9.21	2.73	8.17
2010/2011	3.30	10.34	2.68	7.96
2011/2012	3.22	9.88	3.15	9.06

Table C2.4.: Weibull parameters calculated from MOM (Winter/Summer)

	MOM parameters			
	Winter		Summer	
Years	k	A	k	A
2009/2010	3.21	9.22	2.73	8.17
2010/2011	3.30	10.35	2.67	7.96
2011/2012	3.21	9.88	3.15	9.07

Table C2.5.: Weibull parameters calculated from EPFM (night/day)

	EPFM parameters			
	Night		Day	
	k	A	k	A
2009/2010	3.05	9.36	2.77	8.03
2010/2011	3.02	9.84	2.66	8.55
2011/2012	3.15	10.17	2.87	8.83

Table C2.6: Weibull parameters calculated from WAsP (night/day)

	WAsP parameters			
	Night		Day	
	k	A	k	A
2009/2010	3.30	9.32	2.72	8.04
2010/2011	3.23	9.81	2.54	8.56
2011/2012	3.59	10.10	2.90	8.83

Table C2.7: Weibull parameters calculated from EML (night/day)

	EML parameters			
	Night		Day	
	k	A	k	A
2009/2010	3.27	9.32	2.72	8.04
2010/2011	3.19	9.81	2.55	8.56
2011/2012	3.58	10.09	2.90	8.82

Table C2.8: Weibull parameters calculated from MOM (night/day)

	MOM parameters			
	Night		Day	
	k	A	k	A
2009/2010	3.27	9.32	2.71	8.04
2010/2011	3.18	9.81	2.54	8.56
2011/2012	3.58	10.10	2.90	8.82