Simulating wind characteristics through direct optimization procedures: illustration with three Russian sites

Aleksei Kangash[a], Muhammad Shakeel Virk[b], Pavel Maryandyshev[a], **Alain Brillard**[1a,c]

[a] *Northern (Arctic) Federal University, naberezhnaya Severnoy Dviny 17, 1632002, Arkhangelsk, Russia*

[b] *UiT The Arctic University of Norway, Lodve Langes Gate 2, 8515 Narvik, Norway*

[c] *University of Haute-Alsace, Laboratoire Gestion des Risques et Environnement, 3bis rue Alfred Werner, 68093 Mulhouse, France*

*Email: p.marjyandishev@narfu.ru

Abstract

Wind energy assessment of a territory where a wind park is planned to be built is important. This can be performed through an appropriate evaluation of the wind characteristics in this territory. To simulate the wind speeds, a Weibull function is recommended whose parameters are classically determined either applying logarithms or using one of the formulas proposed in the literature. In the present study, direct optimization procedures are applied, which consist to minimize the squared difference between the experimental and simulated densities or probabilities. These procedures are applied on the wind characteristics collected from the ERA5 website during forty-one years at three Russian sites close to Arkhangelsk. These direct optimization procedures are proved to give lower errors than the classical one or the formulas of the literature. They also lead to lower values of the estimated Annual Energy Production for a Vestas V90-2.0 wind turbine. Direct optimization

[1] Corresponding author. Alain Brillard
Email: alain.brillard@uha.fr
https://orcid.org/0000-0003-0615-8880
University of Mulhouse, Lab. Gestion des Risques et Environnement, 3bis rue Alfred Werner, 68093 Mulhouse, France
Northern (Arctic) Federal University, naberezhnaya Severnoy Dviny 17, Arkhangelsk, Russia

procedures are also applied to determine the optimal parameters associated with a unique or a superposition of two von Mises distribution functions to simulate the wind directions in these three Russian sites.

Keywords. Wind characteristics; Weibull distribution function; von Mises distribution function; Direct optimization procedure; Optimal parameters

1. Introduction

The continuous development of the world economies leads to an increasing energy consumption. Fossil energy sources still account for the major part in the energy balance of many countries throughout the world. This highly affects the climate and environmental changes observed and described by scientists around the world. The use of fossil fuels for energy production indeed causes the formation and release into the atmosphere of hazardous substances such as sulfur, nitrogen and carbon oxides, as well as particulate matter. Carbon dioxide is one of the main gases that contribute to the greenhouse effect. Compounds of substances that are released into the atmosphere by combustion of fossil fuels can have a serious impact on human health and wildlife. In northern countries, the difficulty of delivering fossil fuel here increases its cost several times. There is a further risk of spills during transportation. In addition, fossil fuel power plants produce large amounts of pollutants as well as noise emissions. Considering these disadvantages, the use of large quantities of fossil fuels is unsustainable in remote northern territories. These issues make the use of renewable energy sources a priority for many countries throughout the world. The use of environmentally friendly natural renewable energy sources can provide effective, sustainable and safe energy production. Renewable energy sources have almost no significant impact on the environment, when compared to fossil sources and are available in large amounts in many regions of the world, even if they are non-permanent sources.

Wind energy is one of the main renewable sources. Wind energy industry has strengthened its position in the electricity production throughout the world in the recent years. In some countries, wind power plays an important role in the balance of power generation. More and more new wind parks are being built in the world every year. However, at the moment it is difficult to replace fossil fuels with wind power everywhere. Therefore, scientific research and new technical developments that can increase the efficiency of wind energy usage are needed.

Estimating the wind energy potential of the territory where a wind park is planned to be built is very important. A preliminary estimation based on available data concerning the wind characteristics and a forecast of these characteristics in a reasonable future can give an idea of whether the location of the wind park in this area will be efficient and how much energy can be generated by its installation. An important parameter of a wind park project is its Annual Energy Production (AEP). This parameter can be increased with a correct location of wind turbines. An inaccurate assessment of the wind energy resources of the area, as well as an incorrect location of the planned wind turbines can cause inefficient production from the wind farm. The energy production will be less than planned and this situation may lead to significant economic losses. Wind resource assessment of the territory is very important for areas where no wind parks were constructed before. Therefore, each territory requires appropriate studies before installation of a wind park. Wind resource assessment can contribute to the successful implementation of a wind park project and prevent from fatal mistakes at the design stage. Northern areas have high wind energy potential [1–3].

In the present study, a Vestas V90-2.0 wind turbine is intended to be located in three Russian sites: Dolgoshchelye (lat. 66.3, long. 43.3), hereafter called in short Dolgo, Mezen (lat. 62.4, long. 38.5) and Solovetsky Islands (lat. 65.7, long. 35.4), hereafter called Solov. The wind directions and speeds were collected over forty-one years at the altitude of 100 m from ERA5 site [4]. Simulations of the wind speeds using a Weibull function were performed through different formulas or procedures already proposed in the literature, [5,6], for example. Direct optimization procedures were also applied which allow determining the optimal parameters of a Weibull distribution function. These direct optimization procedures avoid the use of the logarithm which overrides the variations of the data to be considered. The simulations returned by these direct optimization procedures are compared with that returned using a classical procedure or formulas of the literature. To validate the values of the parameters, different error formulas are used. The differences between the values of the parameters have a significant impact on the wind potential of a site, as evaluated for example through the Annual Energy Potential. Direct optimization procedures are also applied to simulate the wind directions, considering a unique or the superposition of two von Mises distribution functions.

2. Material and Methods

2.1. Characteristics of the Russian sites and of the Vestas V90 wind turbine

The behavior of a Vestas V90-2.0 wind turbine to be located in three Russian sites is investigated. The three sites are located in the Arkhangelsk region (oblast), northwest of the Russian Federation, see Fig. 1 a).

a)

b) c)

Fig. 1. Position of the three Russian sites in the Arkhangelsk region a), and photos of the Mezen b) and Solovetsky Islands c).

The sites where the wind turbine is intended to be placed are almost flat and do not present irregular obstacles, see Fig. 1 b) and c). The Dolgoshchelye site is located around a small village in the tundra. For the Dolgoshchelye site, the maximal and minimal wind speeds are respectively equal to 14.61 and 0.91 m/s, respectively. For the Mezen site, the maximal and minimal wind speeds are respectively equal to 14.16 and 0.85 m/s, respectively, quite comparable to that of the Dolgoshchelye site. For the Solovetsky Islands site, the maximal and minimal wind speeds are respectively equal to 20.76 and 0.84 m/s, respectively, the maximal wind of this last site being much higher than the maximal values for the two other sites. In the three cases, the wind directions cover almost the whole range 0-360°. More detailed statistical analyses of the wind speeds and directions in the three sites are given in section 3.1.

The Vestas V90-2.0 wind turbine to be installed in the three sites has a rated power equal to 2 kW [7]. Its hub height is 80 m. More details concerning this wind turbine are given in section 3.2.3.

2.2. Wind characteristics

Wind characteristics were collected from the ERA5 website [4], at the altitude of 100 m and each hour during forty-one years (1979-2020). The Windographer software was used to collect the data. Because of the huge number of data (368,184), daily means were first computed to reduce the number of data to 15,341.

2.3. Analysis of wind data and determination of the main parameters

2.3.1. Simulation of the wind speeds distribution through a Weibull distribution function using the classical procedure

Concerning the wind speed, a two-parameters Weibull distribution function is recommended by the standards IEC 61400-12-1:2017 [8], to assess the values of the wind speed according to their frequency. The density of the Weibull function is written as:

$$f_W(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} \exp\left(-\left(\frac{v}{c}\right)^{k}\right), \qquad (1)$$

where k is the dimensionless shape parameter and c (m/s) is the scale parameter. The cumulative distribution function of this Weibull function is:

$$W(v) = 1 - \exp\left(-\left(\frac{v}{c}\right)^{k}\right). \qquad (2)$$

To determine the two parameters in the present context, the daily mean wind speeds are first ordered in an increasing way. Then a frequency is associated with each daily mean speed, according to the two following possibilities:

- Either a constant frequency equal to $1/(n_d + 1)$ to each daily speed, where $n_d = 15{,}341$ is the number of observed daily mean speeds in the present study,
- Or the daily mean speeds are assembled in n_c classes, for example according to Brook-Carruthers' formula:

$$n_c = 5 \times log_{10}(n_d), \qquad (3)$$

a constant frequency equal to $\#j/(n_c+1)$ is here associated with each daily mean speed in the class j, where $\#j$ is the number of daily mean speeds in the class j.

In the present study, only the second method will be used.

Cumulative probabilities are then deduced from these frequencies.

Because of the structure (2) of the Weibull function, a classical attempt to determine the shape and scale parameters of the Weibull function consists to apply twice the natural logarithm to (2):

$$ln\big(-ln\big(1-W(v)\big)\big) = kln(v) - kln(c). \quad (4)$$

A linear regression method is then used to determine the values of k and $kln(c)$, whence of c. This procedure will be called the classical method. This procedure uses the logarithm which is known to override the variations of the data. This quite simple procedure does not require any numerical software.

2.3.2. Simulation of the wind speeds distribution through a Weibull distribution function using direct optimization procedures

The shape and scale parameters of a Weibull function simulating the density or probability of wind speed in some territory can be derived minimizing the square root of the squared differences between the observed $f_{i,obs}$ and simulated frequencies:

$$l_f^2 = \left(\sum_{i=1}^{n_d}\left(f_{i,obs} - \frac{k}{c}\left(\frac{v_i}{c}\right)^{k-1}\exp\left(-\left(\frac{v_i}{c}\right)^k\right)\right)^2\right)^{1/2}, \qquad (5)$$

or between the observed $W_{i,obs}$ and simulated cumulative probabilities:

$$l_p^2 = \left(\sum_{i=1}^{n_d}\left(W_{i,obs} - 1 + \exp\left(-\left(\frac{v_i}{c}\right)^k\right)\right)^2\right)^{1/2}, \qquad (6)$$

In the present study, the Scilab software and especially its routine ‘datafit’ is used to determine the optimal values of the shape and scale parameters. This Scilab ‘datafit’ routine uses the quasi-Newton algorithm.

When considering the objective function defined in (5) based on the observed and simulated frequencies, the optimization procedure will be called Directf. When considering the objective function defined in (6) based on the observed and simulated probabilities, the optimization procedure will be called Directp.

2.3.3. Simulation of the wind speeds distribution through a Weibull distribution function using other models or formulas

Other methods or formulas have already been proposed in the literature to derive the shape and scale parameters of a Weibull function which simulates the wind speeds in a territory, see for example [9] and the references therein. For example, direct computations may be applied such as:

- *Modified maximum likehood*

  The shape parameter k is determined through the resolution of the nonlinear equation:

$$k = \left(\frac{\sum_{i=1}^{n_d} v_i^k \ln(v_i)\, f_{i,obs}}{\sum_{i=1}^{n_d} v_i^k f_{i,obs}} - \frac{\sum_{i=1}^{n_d} \ln(v_i)}{n_d}\right)^{-1}, \qquad (7)$$

  which is solved in the present study using the Scilab routine ‘fsolve’. The scale parameter is then computed as:

$$c = \left(\sum_{i=1}^{n_d} v_i^k f_{i,obs}\right)^{1/k}. \qquad (8)$$

- *Lysen*

  The shape and scale parameters are computed through:

$$k = \left(\frac{\sigma}{\bar{v}}\right)^{-1.086}; c = \bar{v}\left(0.58 + \frac{0.433}{k}\right)^{-1/k}, \qquad (9)$$

  where σ is the standard deviations of the wind speeds and $\bar{v} = \sum_{i=1}^{n_d} v_i / n_d$ their mean.

- *Energy pattern factor or Justus*

The shape and scale parameters are computed through:

$$k = 1 + \frac{3.69}{\left(E_{pfm}\right)^2}; c = \frac{v_{av}}{\Gamma\left(1+\frac{1}{k}\right)}, \qquad (10)$$

where Γ is the Gamma function and:

$$E_{pfm} = \frac{\sum_{i=1}^{n_d} v_i^3}{n_d(\bar{v})^3}. \qquad (11)$$

- *Moments*

The shape and scale parameters are computed through:

$$k = \left(0.9874 \times \frac{\bar{v}}{\sigma}\right)^{1.0983}; c = \frac{\bar{v}}{\Gamma\left(1+\frac{1}{k}\right)}. \qquad (12)$$

The values of the shape and scale parameters returned by these methods or formulas will be compared for the three Russian sites, considering the properties of a Vestas V90-2.0 wind turbine.

To validate the values of the shape and scale parameters determined through these different procedures or formulas, differences between the observed and simulated frequencies or probabilities are computed through the l_f^2-norm (5) involving the frequencies or the l_p^2-norm (6) involving the probabilities. Some authors also introduce a root mean square error (*RMSE*) which is equal to the l^2-norm, divided by $1/\sqrt{n_d}$. A χ^2 formula is usually computed but which is the square of $RMSE \times \sqrt{n_d/(n_d-2)}$, as there are two parameters to be determined (shape and scale parameters). A R^2 formula is also computed which involves the frequencies but which is equal to $1.0 - \left(l_f^2\right)^2 / \sum_{i=1}^{n_d}\left(f_{i,obs} - \overline{f_{i,obs}}\right)^2$, where $\overline{f_{i,obs}}$ is the mean value of the observed frequencies. The maximal differences between the observed and simulated frequencies or probabilities hereafter respectively named l_f^∞- and l_p^∞-norms, can also be computed. In the present study, only the l_f^2, and l_p^2, l_f^∞, and l_p^∞ values will be computed to evaluate the simulations of the wind speeds, obtained through the different procedures or formulas. Inverting Weibull function, the differences between the observed and simulated wind speeds will also be evaluated, when considering the different sets of parameters.

2.3.4. Determination of the power density, of the Annual Energy Production, of the capacity factor, and of the power output of a wind turbine

The wind power density (W/m^2) is an important parameter to characterize the wind potential of a site. It is defined as:

$$P_d = \frac{1}{2}\rho c^3 \Gamma\left(1 + \frac{3}{k}\right), \qquad (13)$$

where ρ is the air density which will be taken equal to 1.225 kg/m^3, value corresponding to sea level (1 atm) and under a temperature of 15 °C. Some authors analyzed the variations of the air density and the impact of this physical parameter on the wind energy assessment, see [10] for example.

Values of this power density may be taken per year in the present context. However, a unique value will here be considered for the forty-one years of observations.

The theoretical Annual Energy Production (MWh) of the Vestas V90-2.0 wind turbine is computed through:

$$AEP = 365 \times 24 \times \int_{v_{ci}}^{v_{co}} P(v) f(v) dv, \qquad (14)$$

where v_{ci} is the cut-in wind speed (m/s), v_{co} is the cut-off wind speed (m/s), P (kW) is the power output of the wind turbine under consideration, given by the manufacturer [7], P being a function of the wind speed, and f is the Weibull density representing the frequencies of the wind speeds and defined in (1).

From this theoretical Annual Energy Production and the rated power P_r of the wind turbine, which is indicated by the manufacturer, it is possible to determine a theoretical capacity factor through:

$$C_{f,t} = \frac{AEP}{P_r \times 365 \times 24}. \qquad (15)$$

The power density P_d, the Annual Energy Production and the capacity factor $C_{f,t}$, respectively defined through (13), (14), and (15), depend on the shape and scale parameters involved in the Weibull distribution function W (2) and its density f (1) which represent the probabilities and densities of the wind speeds observed in the site under consideration.

The power output P of the wind turbine depends on the wind speed and takes different expressions according to the value of this wind speed with respect to the cut-in wind speed v_{ci}, the rated wind speed v_r, and the cut-off wind speed v_{co}:

$$P(v) = \begin{cases} 0 & if\ v < v_{ci} \\ P_f(v) & if\ v_{ci} \leq v < v_r \\ P_r & if\ v_r \leq v < v_{co} \\ 0 & if\ v > v_{co} \end{cases} \qquad (16)$$

these cut-in, rated, and cut-off wind speeds being defined by the manufacturer website [7].

The capacity factor defined in (15) may be expressed in terms of the power output as, [11]:

$$C_{f,t} = \frac{\bar{P}}{P_r};\ \bar{P} = \int_0^{+\infty} P(v) f(v) dv = \int_{v_{ci}}^{v_r} P_f(v) f(v) dv + P_r \int_{v_r}^{v_{co}} f(v) dv. \qquad (17)$$

In [12], the authors propose different formulas to estimate the capacity factor of a wind turbine.

A sensitivity analysis and a parametric study of the Annual Energy Production will be performed for the Dolgoshchelye site, with respect to the power output or the density energy.

2.3.5. Simulation of the wind distributions through von Mises distribution functions

Usually, simulations of the wind directions are performed considering a unique or the superposition of two von Mises functions, [11]. The use of a von Mises function first requires the conversion of the wind distributions in radians. The density of a von Mises function is defined as:

$$f_{M,\mu,\kappa}(\theta) = \frac{\exp\big(\kappa\cos(\theta - \mu)\big)}{2\pi I_0(\kappa)} = \frac{I_0(\kappa) + 2\sum_{p=1}^{\infty} I_p(\kappa)\cos(p(\theta - \mu))}{2\pi I_0(\kappa)}, \qquad (18)$$

where θ is the wind distribution (Rad), μ is the shape or a location parameter (Rad), κ is the scale or concentration parameter (no unit), and I_j is the modified Bessel function of order j. The associated cumulative probability is given as:

$$F_{M,\mu,\kappa}(\theta) = \frac{\theta I_0(\kappa) + 2\sum_{p=1}^{\infty} I_p(\kappa)\big(\sin\big(p(\theta - \mu)\big) + \sin\big(p(\mu)\big)\big)/p}{2\pi I_0(\kappa)}. \qquad (19)$$

In the case of a superposition of two von Mises distributions, a weight $w \in (0,1)$ is further introduced and the density of this superposition of two von Mises functions is the weighted sum of the previously defined densities:

$$f_{M,\mu_1,\kappa_1,\mu_2,\kappa_2}(\theta)$$

$$= w\frac{I_0(\kappa_1) + 2\sum_{p=1}^{\infty} I_p(\kappa_1)\cos(p(\theta-\mu_1))}{2\pi I_0(\kappa_1)}$$

$$+ (1-w)\frac{I_0(\kappa_2) + 2\sum_{p=1}^{\infty} I_p(\kappa_2)\cos(p(\theta-\mu_2))}{2\pi I_0(\kappa_2)}. \qquad (20)$$

The cumulative probability associated with this superposition of two von Mises distributions is equal to:

$$F_{M,\mu_1,\kappa_1,\mu_2,\kappa_2}(\theta)$$

$$= w\frac{\theta I_0(\kappa_1) + 2\sum_{p=1}^{\infty} I_p(\kappa_1)\big(\sin\big(p(\theta-\mu_1)\big) + \sin\big(p(\mu_1)\big)\big)/p}{2\pi I_0(\kappa_1)}$$

$$+ (1-w)\frac{\theta I_0(\kappa_2) + 2\sum_{p=1}^{\infty} I_p(\kappa_2)\big(\sin\big(p(\theta-\mu_1)\big) + \sin\big(p(\mu_1)\big)\big)/p}{2\pi I_0(\kappa_2)}. \qquad (21)$$

In the present study, the wind distributions are first ordered in an increasing way and classified in 20 classes, according to Brook-Carruthers' formula (3). In the present study, the optimal parameters of a unique or of a superposition of two von Mises function are again determined using the Scilab software and direct optimization procedures dealing with either the density or the probability. These procedures are respectively denoted as Directf1, Directp1, Directf2 and Directp2.

To validate the simulations performed with these procedures, the l_f^2- and l_p^2-norms, as derived from (5)-(6), and the maximal differences (l_f^∞- and l_p^∞-norms) between the observed and simulated cumulative probabilities will be computed and compared.

3. Results and discussion

3.1. Wind characteristics in the three Russian sites

The mean and standard deviations of the daily mean wind speeds collected from the ERAS5 website during forty-one years for the three Russian sites are gathered in Table 1.

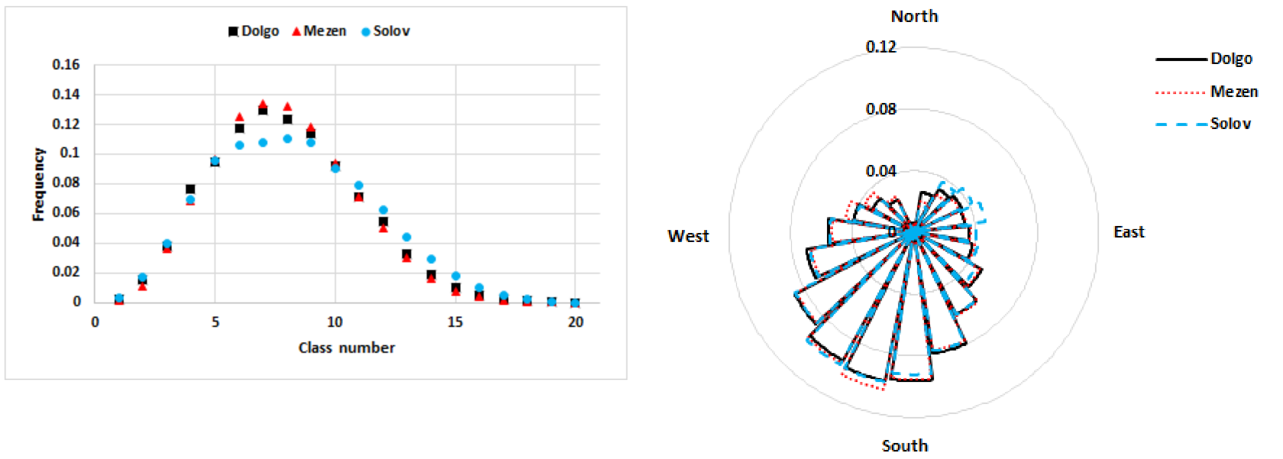
Table 1. Mean and standard deviations of the daily mean speeds and directions for the three Russian sites.

Site	Wind speed		Wind direction	
	Mean (m/s)	Standard deviations (m/s)	Mean (rad)	Standard deviations (rad)
Dolgo	6.018	2.043	3.340	1.355
Mezen	5.760	1.886	3.393	1.337
Solov	8.561	3.272	3.276	1.337

The mean of the daily mean speeds is much higher in the Solovetsky Islands than in the two other sites. Its standard deviations is also much higher, which means that the wind speeds here present larger variations around a higher mean, than in the two other sites. The wind speeds in Mezen present the lowest mean and standard deviations.

The wind directions present quite similar mean and standard deviations for the three sites.

According to Brook-Carruthers' formula (3), the collected wind speeds and directions were assembled into 20 classes. For the wind directions, the first sector is centered around the north. Wind roses and wind speed distributions are shown in Fig. 2, for the three Russian sites.



a) b)

Fig. 2. Wind speed frequencies a) and wind direction roses b) for the Russian site Dolgoshchelye (black squares or solid line), Mezen (red triangle or dotted line) and Solovetsky Islands (blue points or hyphened line).

The wind speed frequencies exhibit small differences between the three sites: the Mezen site presents slightly higher frequencies around the peak and slightly lower frequencies for the last classes. The Solovetsky site presents slightly lower frequencies around the peak which occurs here for higher classes and slightly higher frequencies for the last classes.

The wind roses also present small differences between the three sites, in the major or minor directions: the Solovetsky Islands site presents higher frequencies in the Northeast direction than the two other sites; the Mezen site presents higher frequencies in the Northwest direction than the two other sites. For the three sites, the wind mainly blows from the Northwest direction, which confirms the values of Table 1.

3.2. Simulations of the wind characteristics through Weibull functions

3.2.1. Determination of the shape and scale parameters of the Weibull distribution function through the classical procedure

The parameters of a Weibull function which simulates the wind speeds are first determined through the classical procedure, according to the procedure described in section 2.3.1. The values are gathered in Table 2 for the three sites, together with the four selected errors.

Table 2. Values of the shape and scale parameters, according to the model described in section 2.3.1, l_f^2 and l_p^2 errors as defined in (5)-(6), and maximal differences between observed and simulated frequencies l_f^∞ and probabilities l_p^∞.

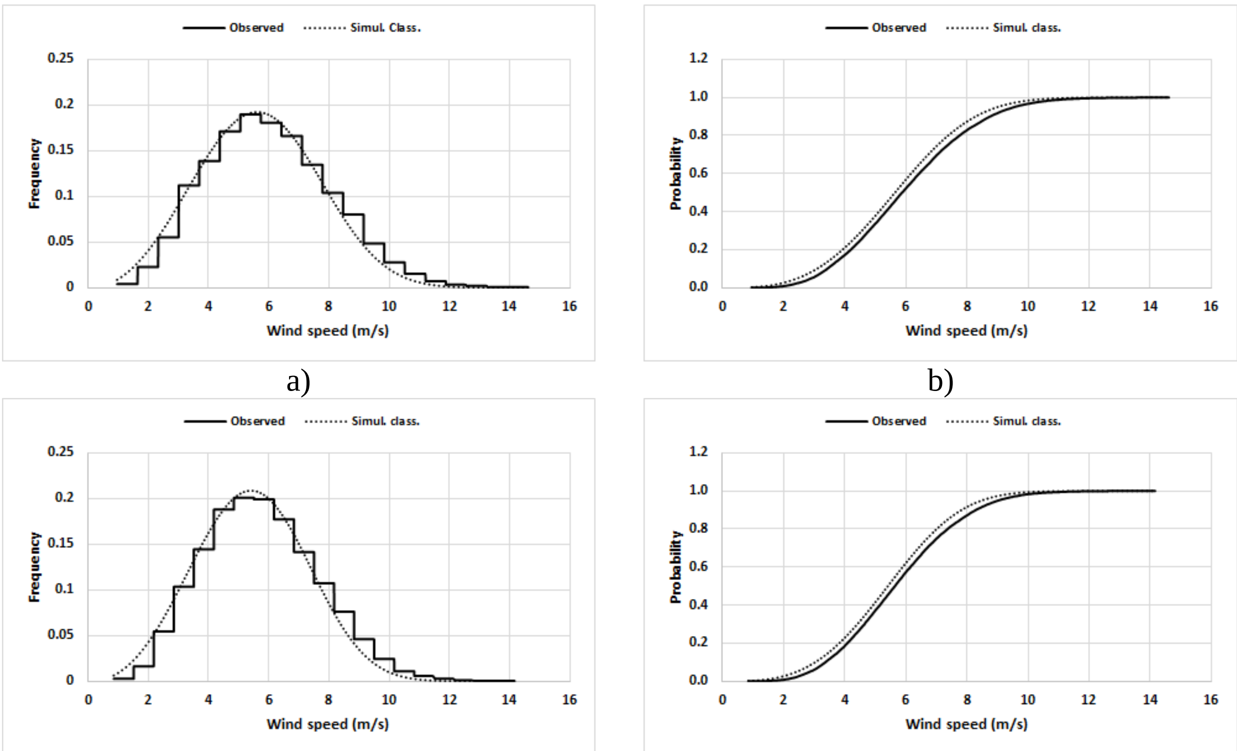
	k	c (m/s)	l_f^2	l_p^2	l_f^∞	l_p^∞
Dolgo	3.133	6.346	1.416	5.173	0.036	0.053

Mezen	3.260	6.058	1.554	5.406	0.035	0.053
Solov	2.710	9.065	1.015	5.149	0.022	0.055

The values of the shape and scale parameters present high differences between the three sites. The Solovetsky Islands site has the lowest shape value and the much higher scale value. On the contrary, the Mezen site has the higher shape value and the lowest scale value. The Dolgoshchelye site presents intermediate values of the shape and scale parameters.

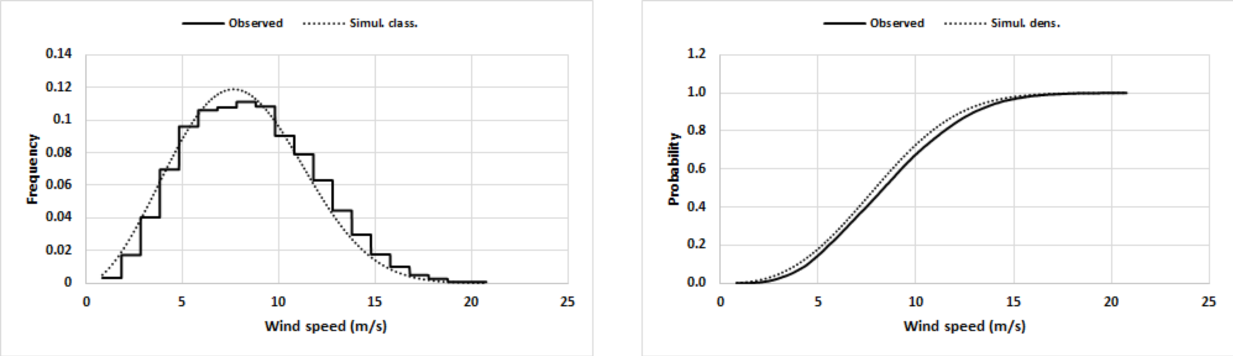
The l_f^2- and l_p^2-norms (5)-(6) are lower for the Solovetsky Islands site and higher for the Mezen site. The l_f^∞-norm is lower for the Solovetsky Islands site and higher for the Dolgoshchelye site. The l_p^∞- is slightly higher for the Solovetsky Islands than for the two other sites.

The observed and simulated (through the classical procedure) frequencies and probabilities of the mean wind speeds are gathered in Fig. 3 for the three Russian sites.



a) b)

c) d)



e) f)

Fig. 3. Observed (solid line) and simulated (dotted line), with the classical method, densities and cumulative probabilities of the wind speeds according to a constant frequency in each class of the 20 wind speeds for the Dolgoshchelye (a) and b)), Mezen (c) and d)) and Solovetsky Islands (e) and f)) sites.

The density curves present slightly different shapes, while the cumulative probability curves look similar. Whatever the site, the simulated frequency curve starts above and ends below the observed one. The probability curve is above the observed one, whatever the site.

Inverting the Weibull function (2) with the parameters determined through the classical method, it is possible to build the wind speeds which correspond to probabilities equal to $(i-1)/(n_d-1)$, $i = 1, \dots, n_d$, and to measure the differences between the observed and simulated wind speeds. Fig. 4 presents the curves of the observed and simulated wind speeds for the Dolgoshchelye site.

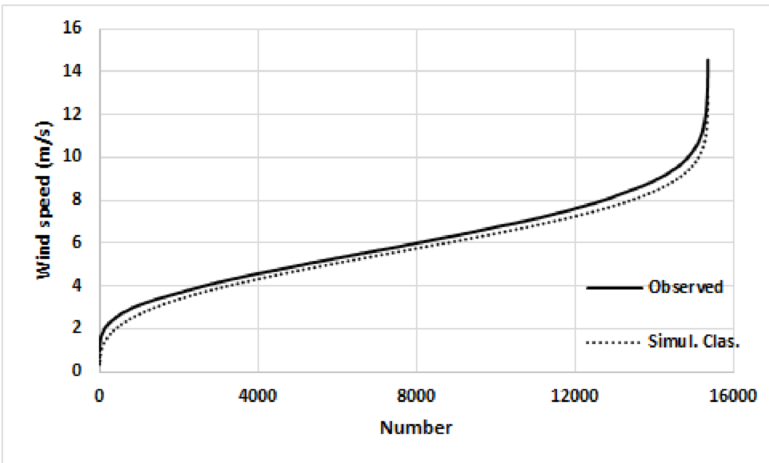


Fig. 4. Observed (solid line) and simulated (dotted line) wind speeds for the Dolgoshchelye site.

The maximal difference between the observed and simulated wind speeds is equal to 1.51 m/s. The l^2 norm of the differences is equal to 45.27 m/s. Similar results can be obtained for the two other sites.

3.2.2. Determination of the shape and scale parameters of the Weibull distribution function through the direct optimization procedures

When considering the direct optimization procedure applied to the Weibull function, the objective function defined in (6) and to be minimized has the shape presented in Fig. 5, for *k* and *c* both varying in the interval [1,10]. Here only the Dolgoshchelye site is considered, the other Russian sites leading to quite similar results.

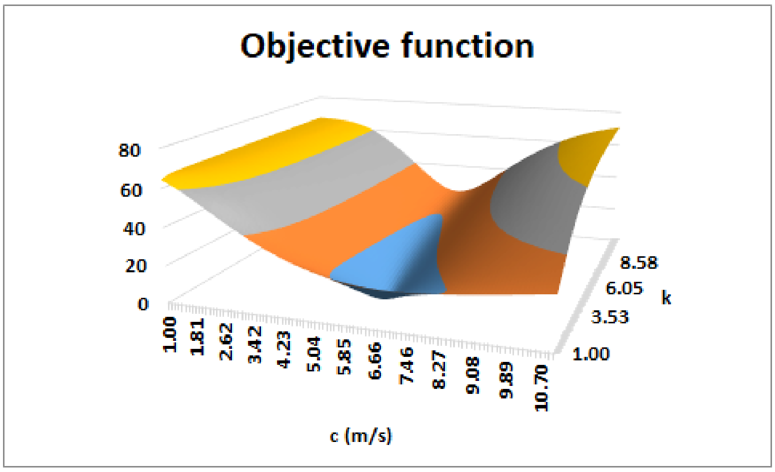


Fig. 5. Values of the objective function defined in (6) for *k* and *c* both varying in the interval [1,10] and for the Dolgoshchelye site.

This objective function presents a minimum for values of *k* close to 2 and of *c* close to 6. This minimizer looks unique. Quite similar observations can be brought for the two other sites (not shown here).

The values of the shape k and scale c parameters defined through the different models indicated in the expressions (7)-(12) of section 2.3.3 are also gathered in Table 3.

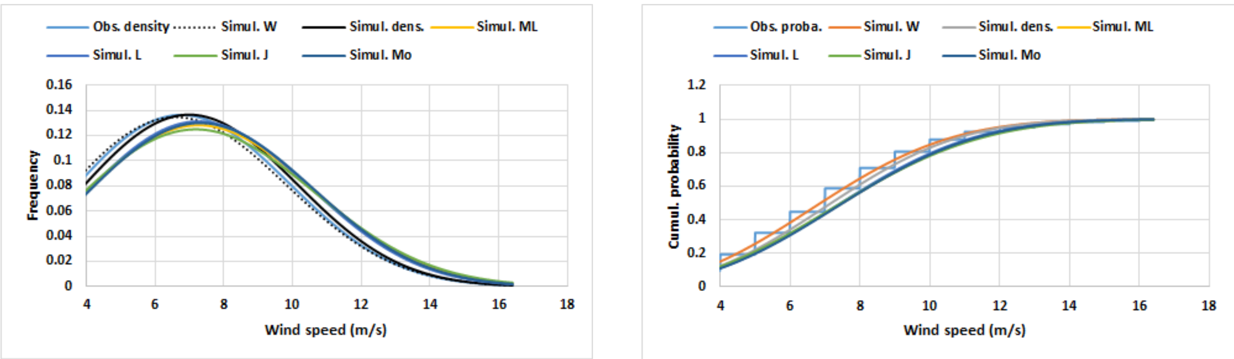
Table 3. Values of the shape and scale parameters, according to the direct optimization methods described in sections 2.3.2 and to formulas presented in section 2.3.3, l_f^2- and l_p^2-errors as defined in (5)-(6), and maximal differences between the observed and simulated frequencies (l_f^∞) and probabilities (l_p^∞).

	Directf	Directp	Max. like.	Lysen	Energy pattern factor (Justus)	Moments
Dolgo						
k	3.089	2.973	3.186	3.232	2.995	3.230
c (m/s)	6.499	6.307	6.723	6.679	6.740	6.740
l_f^2	1.540	1.174	1.608	1.557	1.669	1.714
l_p^2	3.615	4.961	8.080	7.652	7.709	8.469
l_f^∞	0.043	0.032	0.036	0.036	0.031	0.038
l_p^∞	0.069	0.104	0.144	0.139	0.136	0.148
Mezen						
k	3.091	3.219	3.307	3.362	3.075	3.361
c (m/s)	6.016	6.204	6.422	6.381	6.444	6.444
l_f^2	1.721	1.276	1.771	1.712	1.941	1.915
l_p^2	3.818	5.283	8.534	8.124	8.166	9.049
l_f^∞	0.043	0.033	0.038	0.036	0.039	0.039
l_p^∞	0.072	0.111	0.153	0.148	0.144	0.159
Solov						
k	2.574	2.671	2.837	2.842	2.738	2.836
c (m/s)	9.065	9.437	9.621	9.552	9.622	9.622
l_f^2	1.012	0.731	0.944	0.924	0.821	0.943
l_p^2	3.220	5.045	7.106	6.583	6.706	7.111
l_f^∞	0.019	0.026	0.024	0.023	0.022	0.024
l_p^∞	0.058	0.096	0.125	0.121	0.118	0.125

Whatever the site, the values of the shape parameter k and of the scale parameter c highly depend of the method or formula. The Energy pattern factor method returns the lowest values of the shape parameter k. The Lysen formula returns the highest values of the shape parameter k and scale parameter c. The direct optimization method Directp (based on the probability) returns the lowest l_f^2-error, which is the smallest error among the six methods or formulas, and the highest l_p^2-error. The direct optimization method Directf (based on the density) returns the lowest l_p^2-error among the six methods. The moments formula returns the highest l_f^2- and l_p^2-errors.

For the Dologoshchelye site, the mean value $\bar{k}$ and the standard deviations σ_k of the k parameters returned by the different models or formulas are equal to 3.1232 and 0.1151, respectively. The mean value $\bar{c}$ and the standard deviations σ_c of the c parameters returned by the different models or formula are equal to 6.6147 and 0.1609 m/s, respectively.

The density and cumulative Weibull distributions associated with these six sets of shape and scale parameters are gathered in Fig. 6.



a) b)

Fig. 6. Observed and simulated density a) and probability b) curves for the Dolgoshchelye site. The observed curves are in light blue, the simulated curves obtained with the direct optimization with respect to the probability in red, with the direct optimization with respect to the density in grey, with maximum likehood in yellow, Lysen in dark blue, Justus in green, and moments in violet.

It is difficult to identify the different curves even if they present quite significant differences, see also Table 3. Figure 7 focuses on the observed and simulated density and probability curves for the three Russian sites, the simulated curves being obtained with the two direct optimization procedures described in section 2.3.2.

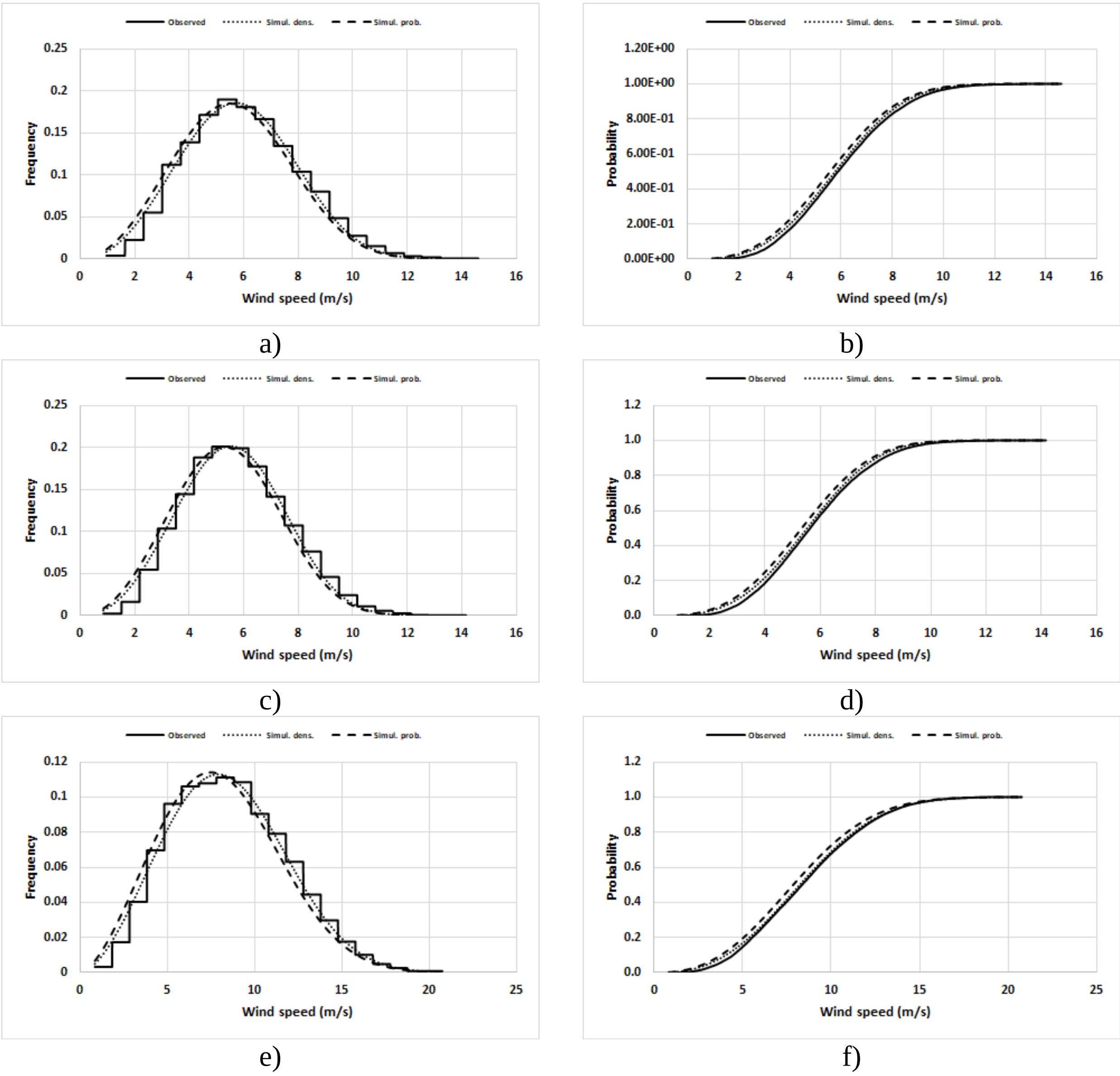


Fig. 7. Observed and simulated density and probability curves for the Dolgoshchelye (a) and b)), Mezen (c) and d)) and Solovetsky Islands sites (e) and f)). The observed curves are the solid lines, the simulated curves obtained with the direct optimization with respect to the density are the dotted lines, with the direct optimization with respect to the density are the hyphened lines.

For the three sites, the probability curves seem to be better simulated when considering the direct optimization curves involving the densities. Whatever the site, the density curves indeed go more between the steps of the observed density curves when considering the direct optimization curves involving the densities, especially after the peak.

The simulations of the wind speeds presented in this section were based on observations performed at 100 m height, although the hub height of the Vestas V90-2.0 is 85 m. Corrections should be brought to these observed wind speeds, according to a formula given in the literature when considering a quite flat territory, see [12,13], for example.

From the Weibull functions whose parameters are determined through the direct optimization methods presented in section 2.3.2, it is possible to build the wind speeds which correspond to probabilities equal to $(i-1)/(n_d-1)$, $i = 1, \dots, n_d$, and to measure the differences between the observed and simulated wind speeds. The maximal differences between the observed and simulated wind speeds are equal to 1.14 and 1.09 m/s, when considering the frequencies or probabilities, respectively. These maximal differences are slightly lower than that (1.51 m/s) obtained when considering the Weibull function whose parameters are determined through the classical method. The l^2-norms of the differences between the observed and simulated wind speeds are equal to 49.56 and 29.47 m/s, respectively. The l^2-norm of the differences between the observed and simulated wind speeds obtained when considering the Weibull function with parameters deduced from the classical method was computed at 45.27 m/s, between the two l^2-norms.

3.2.3. Determination of the power density, AEP and capacity factor

For a Vestas V90-2.0 wind turbine, the cut-in wind speed v_{ci} is equal to 3.0 m/s, the rated wind speed v_r is equal to 12.5 m/s, and the cut-off wind speed v_{co} is equal to 20.5 m/s. The rated power P_r is equal to 2 MW. In the interval $[v_{ci}, v_r]$, the power output P_f (MW) of the wind turbine may be approximated through the polynomial function of degree 6:

$$P_f(v) \sim 2.4925 \times 10^{-5} \times v^6 - 1.1420 \times 10^{-3} \times v^5 + 2.0165 \times 10^{-2} \times v^4 - 0.1762 \times v^3 + 0.8298 \times v^2 - 1.9213 \times v + 1.6817. \qquad (22)$$

The determination coefficient of this polynomial function of degree 6 is higher than 0.999.

As deduced from the formulas (13)-(15) in section 2.3.4 and of the values of the scale and shape parameters determined in the previous section, the values of the power density, of the annual energy production and of the capacity factor are gathered in Table 4.

Table 4. Values of the power density P_d, of the Annual Energy Production AEP and of the capacity factor C_f, as deduced from the Weibull functions indicated in Table 3, through the expressions (13)-(15).

	Directf	Directp	Max. like.	Lysen	Energy pattern factor	Moments
Dolgo						
P_d (W/m^2)	137.2	150.3	166.7	163.5	167.4	168.0
AEP (MWh)	3695.2	3983.5	4352.4	4262.2	4436.8	4371.6
C_f (%)	0.211	0.227	0.248	0.243	0.253	0.250
Mezen						
P_d (W/m^2)	119.2	131.1	145.5	142.9	146.5	147.1
AEP (MWh)	3179.8	3447.6	3790.9	3707.7	3892.8	3814.7
C_f (%)	0.181	0.197	0.216	0.212	0.222	0.218
Solov						
P_d (W/m^2)	405.1	457.6	485.9	475.6	485.4	486.0
AEP (MWh)	8454.5	9060.4	9439.3	9340.8	9373.1	9440.3
C_f (%)	0.483	0.517	0.539	0.533	0.535	0.539

The power density, Annual Energy Production and capacity factor highly depend on the model. The direct optimization method based on the density returns the lowest values of the power density, of the Annual Energy Production and of the capacity factor, whatever the site. The Energy pattern factor and moments methods return the highest values of these parameters. Concerning the Annual Energy Production, the relative increase between the lowest and highest values is equal to 20.0% for Dolgoshchelye, to 22.4% for Mezen and to 11.7% for Solovetsky Islands sites, which are very high percentages.

3.2.4. Sensitivity analysis and parametric study concerning the Annual Energy Production in the case of the Dolgoshchlye site

The formula (14) giving the Annual Energy Production involves the power output, the cut-in v_{ci} and cut-off v_{cr} rates of the wind turbine, as indicated by the manufacturer. The cut-in and cut-off rates are fixed by the manufacturer, as they deal with efficiency or security reasons. Consequently, the uncertainty of the Annual Energy Production comes from that of the power out curve or of the Weibull function. A sensitivity analysis is performed on these two terms, but in an additive way, the case of a multiplicative uncertainty being indeed trivial. In the case of additive uncertainties on the power $\delta P(v)$ and on the Weibull density $\delta f(v)$, the uncertainty on the Annual Energy Production is computed through:

$$AEP + \delta AEP = 365 \times 24 \int_{v_{ci}}^{v_{co}} (P(v) + \delta P(v))(f(v) + \delta f(v))dv$$

$$= 365 \times 24 \int_{v_{ci}}^{v_{co}} (P(v)f(v) + \delta P(v)f(v) + P(v)\delta f(v) + \delta P(v)\delta f(v))dv. \quad (23)$$

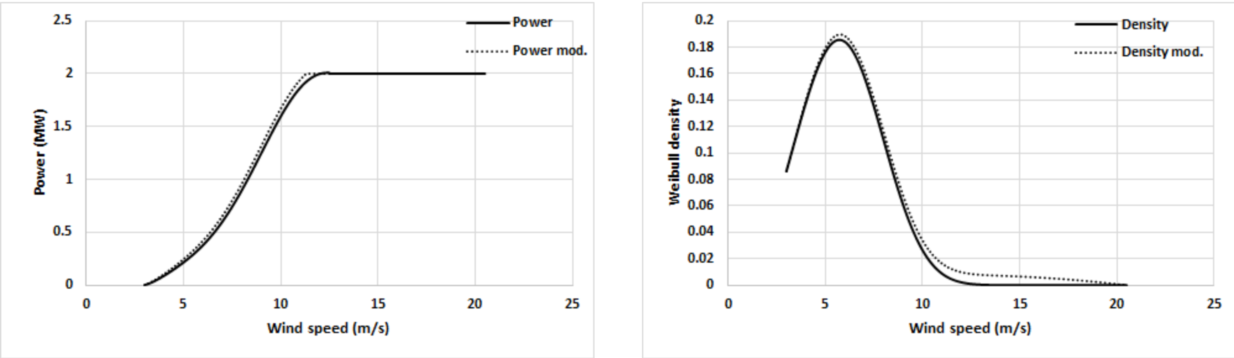
The power curve given in (17) was slightly modified through the formula:

$$P_{fm}(v) = max\left(min\left(P_f(v) - \frac{(v - v_{ci})(v - v_{co})}{1000.0}, P_f(v_{co})\right), 0.0\right). \quad (24)$$

The Weibull density given in (1) and determined through the direct optimization method with the probabilities, leading to the values $k = 2.973, c = 6.307$ m/s, see Table 3, was modified according to the formula:

$$f_{Wm}(v) = f_W(v) - \frac{(v - v_{ci})(v - v_{co})}{10000.0}. \quad (25)$$

The curves of the original and modified power outputs and densities are gathered in Fig. 8.



a) b)

Fig. 8. Original (solid line) and modified (dotted line) power output and density in the interval $[v_{ci}, v_{co}]$.

The integral of the absolute difference between the power curves is equal to 0.44 MW, to be compared to the integral of the power output, which is equal to 24.90 MW, thus representing a relative variation of 1.7%. The integral of the absolute difference between the density curves is equal to 0.089, to be compared to the integral of the density which is equal to 1.0, thus representing a relative variation of 8.9%.

Only modifying the power output according to (24), the Annual Energy Production increases from 3983.5 to 4317.1 MW, which represents a relative variation of 8.4%. Only modifying the density according to (25), the Annual Energy Production decreases from 3983.5 to 3943.6 MW, which represents a relative variation of 1%. Modifying the power output and density expressions, the Annual Energy Production increases from 3983.5 to 5557.9 MW, which represents a relative variation of 39.5%! Even if the relative variation of the power output is lower than that of the density, its impact on the variation of the Annual Energy Production is much higher, due to the expression of the integral leading to the Annual Energy Production.

Considering the Dolgoshchelye site, a parametric analysis with respect to the *k* and *c* parameters of a Weibull function can be performed on the Annual Energy Production, taking the five values $(\bar{k}, \bar{c})$

and $\left(\bar{k} \pm \sigma_k, \bar{c} \pm \sigma_c\right)$. The values of the Annual Energy Production obtained for these four last values are gathered in Table 5.

Table 5. Values of the Annual Energy Production for the five values $\left(\bar{k}, \bar{c}\right)$ and $\left(\bar{k} \pm \sigma_k, \bar{c} \pm \sigma_c\right)$, and relative variations with respect to the value of the Annual Energy Production obtained when taking $\left(\bar{k}, \bar{c}\right)$ and equal to 4176.8 MWh, third line), and also with respect to the mean value of the Annual Energy Productions determined through the different methods (Table 4) and equal to 4183.6 MWh, fourth line).

	$\left(\bar{k}, \bar{c}\right)$	$\left(\bar{k} + \sigma_k, \bar{c}\right)$	$\left(\bar{k} - \sigma_k, \bar{c}\right)$	$\left(\bar{k}, \bar{c}+\sigma_c\right)$	$\left(\bar{k}, \bar{c}-\sigma_c\right)$
AEP (MWh)	4176.8	4143.1	4208.8	4464.3	3895.8
		0.81%	0.76%	6.88%	6.73%
	0.16%	0.97%	0.60%	6.71%	6.88%

Variations of the *k* parameter approximately equal to 3.7% do not highly modify so much the Annual Energy Production (relative variations less than 1%). On the contrary, variations of the *c* parameter approximately equal to 2.6% significantly modify the Annual Energy Production (variations greater than 6%).

From the values of the parameters presented in Table 4, it is possible to perform an economic analysis of the wind turbine, like in [14], for example.

3.2.5. Simulations of the wind directions with a superposition of two von Mises functions

The wind distributions obtained from the ERA5 website [4] during forty-one years, converted in radians and gathered in 20 classes lead to the observed density and probability curves presented in Fig. 9.

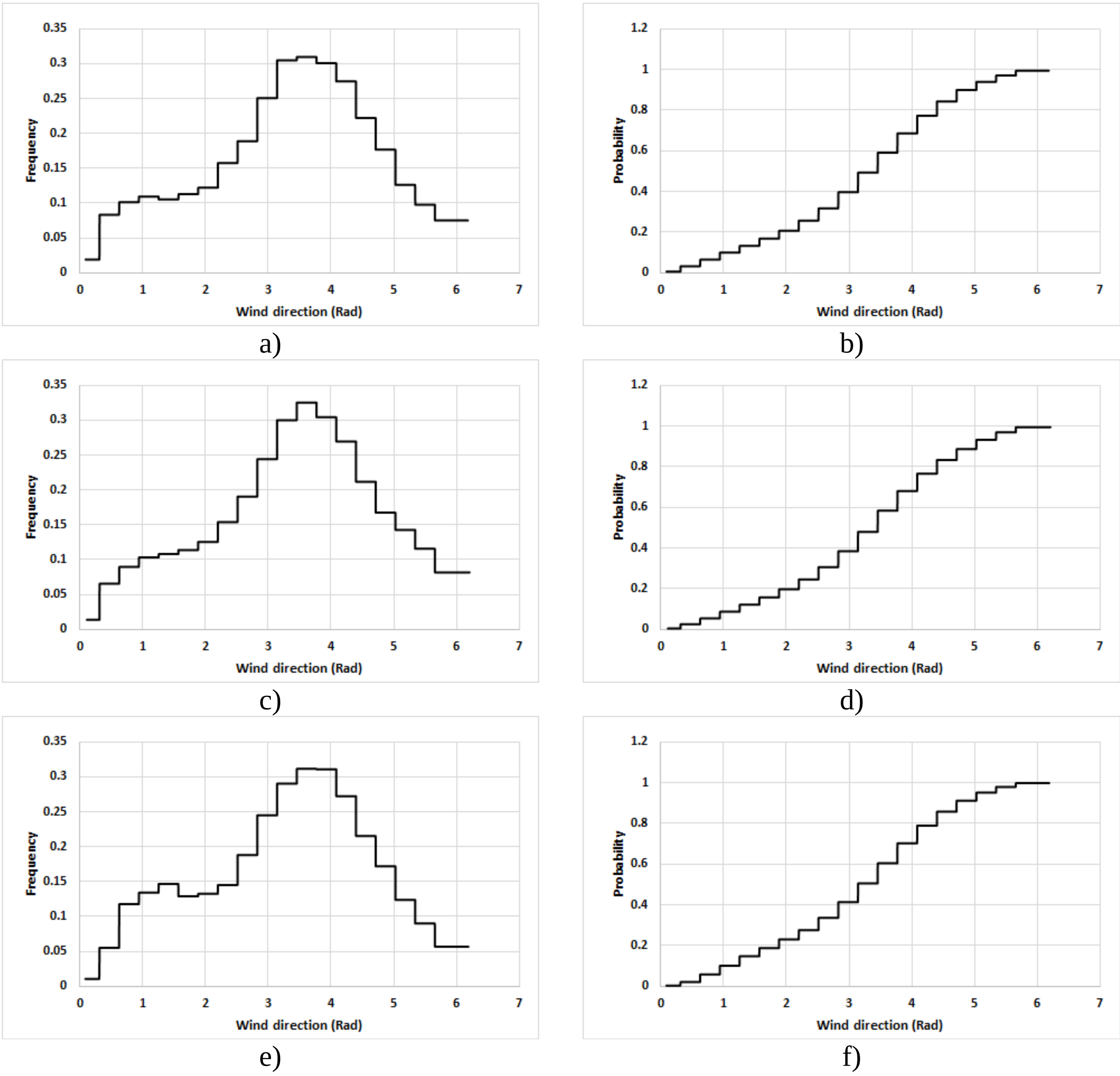


Fig. 9. Observed density and cumulative probability curves of the wind directions for the Dolgoshchelye (a) and b)), Mezen (c) and d)) and Solovetsky (e) and f)) sites.

Clearly, two peaks appear on the density curve Fig. 9 a), c) and e), which correspond to the wind rose given in Fig. 2 a), dotted line, the second one being more important and corresponding to winds oriented Southwest.

The optimal values of the parameters involved in the unique or in the superposition of two von Mises functions which intend to simulate the wind densities and probabilities directions, according to the direct optimization procedures described in section 2.3.5 are gathered in Table 4. The different errors between the observed and simulated density and probability curves are also gathered in Table 6.

Table 6. Values of the parameters of a unique or of a superposition of two von Mises to simulate the wind directions, according to the model described in section 2.3.5, l_f^2- and l_p^2-errors as adapted in (5)-(6), and maximal differences between the frequencies l_f^∞ and probabilities l_p^∞.

	Directf1	Directp1	Directf2	Directp2
Dolgo				
μ (Rad)	3.661	3.661	-	-
κ (Rad)	0.813	0.813	-	-
μ_1 (Rad)	-	-	1.064	1.021
μ_2 (Rad)	-	-	3.683	3.527
κ_1 (Rad)	-	-	4.205	4.284
κ_1 (Rad)	-	-	0.974	1.107
w	-	-	0.071	0.103
l_f^2	2.259	2.259	1.512	2.259
l_p^2	7.602	7.602	5.345	2.594
l_f^∞	0.046	0.046	0.045	0.046
l_p^∞	0.121	0.121	0.102	0.052
Mezen				
μ (Rad)	3.666	3.666	-	-
κ (Rad)	0.825	0.825	-	-
μ_1 (Rad)	-	-	1.187	1.111
μ_2 (Rad)	-	-	3.690	3.537
κ_1 (Rad)	-	-	4.335	4.740
κ_1 (Rad)	-	-	0.967	1.061
w	-	-	0.063	0.082
l_f^2	2.208	2.208	1.654	2.208
l_p^2	6.732	6.732	4.858	2.618
l_f^∞	0.050	0.050	0.044	0.050
l_p^∞	0.116	0.116	0.099	0.054
Solov				
μ (Rad)	3.648	3.648	-	-
κ (Rad)	0.778	0.778	-	-
μ_1 (Rad)	-	-	1.244	1.170
μ_2 (Rad)	-	-	3.705	3.550
κ_1 (Rad)	-	-	4.289	4.022
κ_1 (Rad)	-	-	1.082	1.232
w	-	-	0.121	0.151
l_f^2	3.392	3.392	1.506	3.392
l_p^2	8.902	8.902	5.337	2.590
l_f^∞	0.069	0.069	0.044	0.069
l_p^∞	0.133	0.133	0.102	0.052

Whatever the site and considering a unique von Mises function, the optimal values of the parameters determined through the two direct optimization procedures are exactly the same. In the case of the superposition of two von Mises functions, slight differences appear between the two sets of optimal values. In both cases, the l_f^2 and l_f^∞ errors are smaller when considering the direct optimization procedure involving the densities. The l_p^2 and l_p^∞ errors are also smaller when considering the direct optimization procedure involving the probabilities.

Figure 10 gathers the observed and simulated density and probability curves for the wind directions in the three sites.

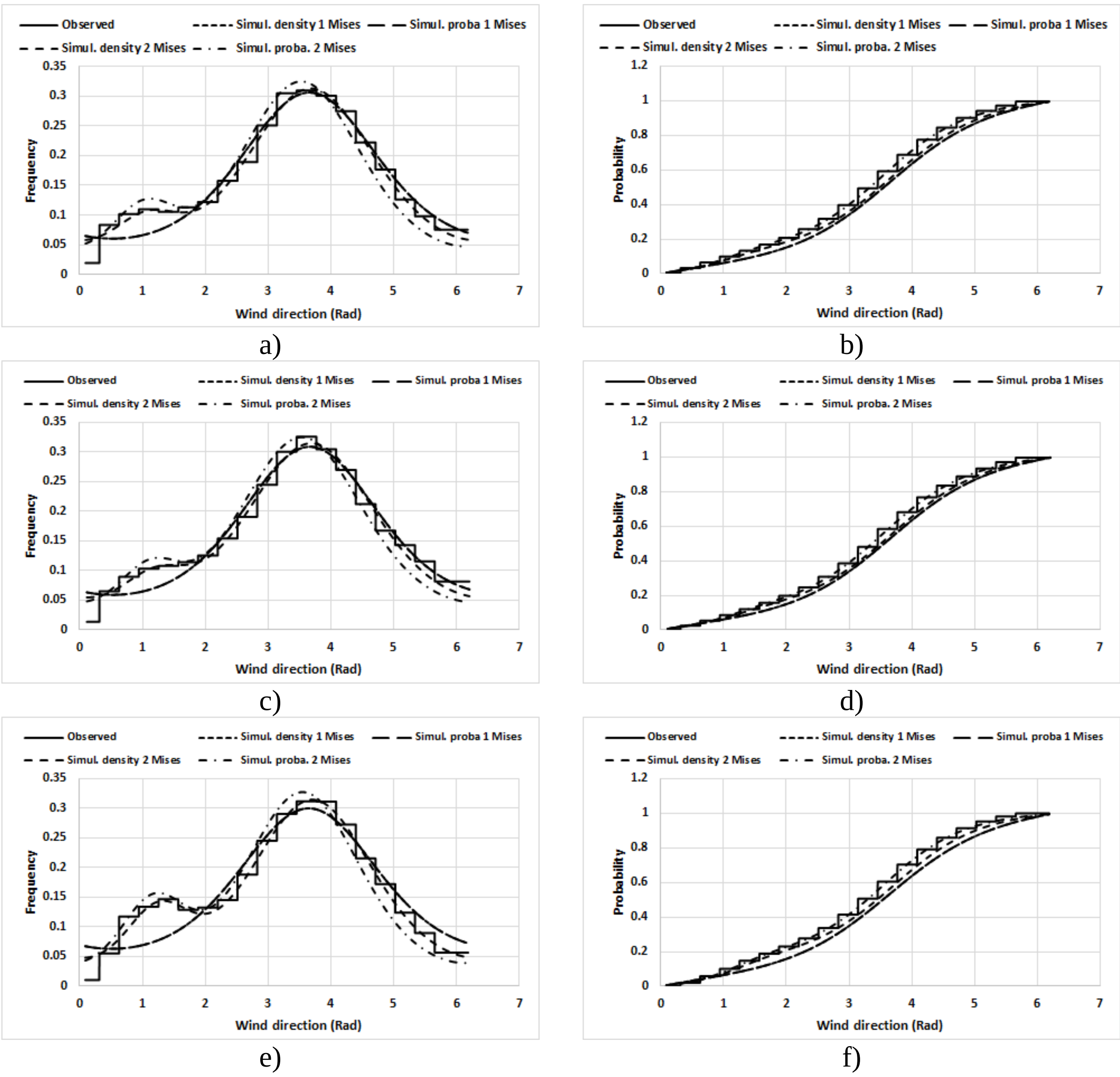


Fig. 10. Observed (solid line) and simulated with a unique von Mises (dotted line), or with a superposition of two von Mises (hyphened line) distributions, for the density and cumulative probability of wind directions and for the Dolgoshchelye (a) and b)), Mezen (c) and d)) and Solovetsky (e) and f)) sites.

The simulations with a unique von Mises function are much poorer than that with a superposition of two von Mises functions. This is the consequence of two peaks, even if the first is much smaller than the second one. Clearly, the simulations with the superposition of the two von Mises distributions with the weight w better reproduce the two peaks and the shape of the observed cumulative probability.

Considering the superposition of two von Mises functions, the l_f^2 and l_p^2 errors are quite large, whatever the direct optimization procedure. This is surely the consequence of the shape of the observed density and probability curves. Quite important jumps indeed appear for the lowest values of the wind direction.

4. Conclusion

Predicting the power delivered by a wind turbine is important, at least from an economic point of view. Such prediction can be realized analyzing the wind characteristics in the chosen site, simulating them with appropriate tools and taking into account the chosen wind turbine. In the present study, the characteristics of wind data collected from the ERA5 website during forty-one years and concerning three Russian sites in the Arkhangelsk region were analyzed. Concerning the wind speeds, different mathematical methods or formulas, among which two direct optimization ones, were used to determine the values of the shape and scale parameters of a Weibull distribution function representing the wind speeds organized in 20 classes. The direct optimization methods consist to minimize an objective function which involves the squared differences between the observed and simulated frequencies or probabilities, with respect to the two parameters of a Weibull function. The errors between the observed and simulated wind speeds were lower using the direct optimization methods. The Annual Energy Potential deduced from the values of the parameters of the Weibull function simulating the wind speeds was significantly lower when considering the direct optimization methods. A sensitivity analysis and a parametric study were conducted on this Annual Energy Production with respect to the power output of the turbine and to the density of a Weibull function representing the wind speed probabilities. For the wind directions, a superposition of two von Mises distributions was applied with good agreement for each site. Here again, direct optimization methods were applied to derive the parameters involved in this superposition of two von Mises distributions. As soon as wind turbines will be installed in the three chosen sites, their behavior will be compared with the simulations performed in the present study. Working in an Arctic region, icing and de-icing phenomena surely will occur and should be taken into account, as they penalize the optimal behavior of the wind turbines.

Funding. The authors did not receive support from any organization for the submitted work.

Conflict of interest. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Weis, T.M., Ilinca, A.: Assessing the potential for a wind power incentive for remote villages in Canada. Energy Policy. 38, 5504–5511 (2010). https://doi.org/10.1016/j.enpol.2010.04.039
2. Souba, F., Mendelson, P.B.: Chaninik Wind Group: Lessons learned beyond wind integration for remote Alaska. The Electricity Journal. 31, 40–47 (2018). https://doi.org/10.1016/j.tej.2018.06.008
3. Ghani, R., Kangash, A., Virk, M.S., Maryandyshev, P., Mustafa, M.: Wind energy at remote islands in arctic region—A case study of Solovetsky islands. Journal of Renewable and Sustainable Energy. 11, 053304 (2019). https://doi.org/10.1063/1.5110756
4. ERA5 hourly data on single levels from 1979 to present. Accessed October 12th, 2021., https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form
5. Freitas de Andrade, C., Ferreira dos Santos, L., Silveira Macedo, M.V., Costa Rocha, P.A., Ferreira Gomes, F.: Four heuristic optimization algorithms applied to wind energy: determination of Weibull curve parameters for three Brazilian sites. Int J Energy Environ Eng. 10, 1–12 (2019). https://doi.org/10.1007/s40095-018-0285-5

6. Kollu, R., Rayapudi, S., Narasimham, S., Pakkurthi, K.: Mixture probability distribution functions to model wind speed distributions. Int J Energy Environ Eng. 3, 27 (2012). https://doi.org/10.1186/2251-6832-3-27
7. Vestas: Vestas V90 wind turbine. https://www.vestas.com/en/products/2%20mw%20platform/v90%202_0_mw#!. Accessed October 5th, 2021, https://www.vestas.com/en/products/2%20mw%20platform/v90%202_0_mw#!
8. International Electrotechnical Commission: Wind energy generation systems. Part 12-1. (2017)
9. Khalid Saeed, M., Salam, A., Rehman, A.U., Abid Saeed, M.: Comparison of six different methods of Weibull distribution for wind power assessment: A case study for a site in the Northern region of Pakistan. Sustainable Energy Technologies and Assessments. 36, 100541 (2019). https://doi.org/10.1016/j.seta.2019.100541
10. Jung, C., Schindler, D.: The role of air density in wind energy assessment – A case study from Germany. Energy. 171, 385–392 (2019). https://doi.org/10.1016/j.energy.2019.01.041
11. Masseran, N., Razali, A.M., Ibrahim, K., Latif, M.T.: Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia. Energy Conversion and Management. 72, 94–102 (2013). https://doi.org/10.1016/j.enconman.2012.11.025
12. Diyoke, C.: A new approximate capacity factor method for matching wind turbines to a site: case study of Humber region, UK. Int J Energy Environ Eng. 10, 451–462 (2019). https://doi.org/10.1007/s40095-019-00320-5
13. Sunderland, K., Woolmington, T., Blackledge, J., Conlon, M.: Small wind turbines in turbulent (urban) environments: A consideration of normal and Weibull distributions for power prediction. Journal of Wind Engineering and Industrial Aerodynamics. 121, 70–81 (2013). https://doi.org/10.1016/j.jweia.2013.08.001
14. Quan, P., Leephakpreeda, T.: Assessment of wind energy potential for selecting wind turbines: An application to Thailand. Sustainable Energy Technologies and Assessments. 11, 17–26 (2015). https://doi.org/10.1016/j.seta.2015.05.002