APPROACHES TO STOCHASTIC MODELING OF WIND TURBINES

Migran N. Gevorkyan,
Anastasiya V. Demidova
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
Email: gevorkyan_mn@rudn.university,
demidova_av@rudn.university

Ivan S. Zaryadov
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
and Institute of Informatics Problems,
FRC CSC RAS, IPI FRC CSC RAS,
44-2 Vavilova str., Moscow, 119333, Russia
Email: zaryadov_is@rudn.university

Robert A. Sobolewski
Department of Power Engineering,
Fotonics and Lighting Technology,
Bialystok University of Technology,
45D Wiejska str., 15-351 Bialystok, Poland
Email: r.sobolewski@pb.edu.pl

Anna V. Korolkova
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
Email: korolkova_av@rudn.university

Dmitry S. Kulyabov
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
and Laboratory of Information Technologies
Joint Institute for Nuclear Research
6 Joliot-Curie, Dubna,
Moscow region, 141980, Russia
Email: kulyabov_ds@rudn.university

Leonid A. Sevastianov
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
and Bogoliubov Laboratory of Theoretical Physics
Joint Institute for Nuclear Research
6 Joliot-Curie, Dubna,
Moscow region, 141980, Russia
Email: sevastianov_la@rudn.university

KEYWORDS

Weibull distribution, approximation, lognormal distribution, gamma distribution, beta distribution, wind speed, statistics

ABSTRACT

Background. This paper study statistical data gathered from wind turbines located on the territory of the Republic of Poland. The research is aimed to construct the stochastic model that predicts the change of wind speed with time. **Purpose.** The purpose of this work is to find the optimal distribution for the approximation of available statistical data on wind speed. **Methods.** We consider four distributions of a random variable: Log-Normal, Weibull, Gamma and Beta. In order to evaluate the parameters of distributions we use method of maximum likelihood. To assess the the results of approximation we use a quantile-quantile plot. **Results.** All the considered distributions properly approximate the available data. The Weibull distribution shows the best results for the extreme values of the wind speed. **Conclusions.** The results of the analysis are consistent with the common practice of using the Weibull distribution for wind speed modeling. In the future we plan to compare the results obtained with a much larger data set as well as to build a stochastic model of the evolution of the wind speed depending on time.

INTRODUCTION

This work is devoted to the problem of stochastic modeling of speed of wind, which is used to generate electrical power in wind plants located on the territory of the Republic of Poland. As a first step several distributions for accuracy of the wind speed approximation will be examined. For this purpose Log-normal, Weibull, Gamma and Beta are chosen. All these distributions have shape-location-scale parametrisation. For statistical data processing the authors used Python 3 with `numpy`, `scipy.stats` (see Jones et al. (2001)) and `matplotlib` (see Droettboom et al. (2017)) libraries and also `Jupyter` (see *Project Jupyter home* (2017))—an interactive shell. We used books (see Norman L. Johnson (1994, 1995); Nelson (1982)) as reference materials for distributions properties. Articles

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

(see Frchet (1927); Weibull (1951)) are the primary sources in which the Weibull distribution is presented for the first time. Articles (see Lun and Lam (2000); Seguro and Lambert (2000); Bowden et al. (1983); Yeh and Wang (2008); Islam et al. (2011); Garcia et al. (1998)) describe the use of the Weibull distribution in the modeling of wind turbines and wind speed.

THE DESCRIPTION OF THE STATISTICAL DATA STRUCTURE

The set of statistical data is stored in the file `csv` consisting of the following columns:

1) T — time of fixation of wind speed and direction by sensors installed on the wind power turbine (hh:mm format);
2) X_1 — output power of wind turbine [kW] (the negative values mean the power is consumed rather then generated);
3) X_2 — wind speed [m/s] (measured by anemometer installed at the top of wind turbine nacelle);
4) X_3 — wind direction [deg] (measured by anemometer installed at the top of wind turbine nacelle; measured clockwise, the value 0 to the N);
5) X_4 — wind speed 10 m [m/s] obtained at 10 m above the ground m;
6) X_5 — wind direction 10 m [deg] (obtained at 10 m above the ground; measured clockwise, the value from 0 to the N);
7) X_6 — wind speed 50 m [m/s] (obtained at 50 m above the ground);
8) X_7 — wind direction 50 m [deg] (obtained at 50 m above the ground; measured clockwise, the value ftom 0 to the N).

The indicators of wind speed and direction were read out from the sensors every 10 minutes for about 9 months. In total, the table contains 39606 entries.

To make an initial choice of distributions that may be suitable for wind speed approximation, the histograms of wind speed are drawn. Visual assessment of these histograms suggest that the adequate choice will be a "heavy-tailed" distribution. But for the wind direction approximation these distributions are not suitable, as can be seen from the figure 1.

To read out the data we used the function `genfromtxt` from `numpy` (see Jones et al. (2001)) lib:

```
ws1, ws2, ws3 =
    np.genfromtxt('data.csv',
    delimiter=';', skip_header=True,
    usecols=(2, 4, 6), unpack=True)
```

where `'data.csv'` is data file, `delimiter=';'` is columns separator, `skip_header = True` specifies ignoring of the first line as the names of the columns, `usecols=(2, 4, 6)` makes function to use only 2, 4, 6 columns (numbering begins with zero) and `unpack=True` — contents of each column should be written in separate arrays `ws1`, `ws2` and `ws3` for further analysis of the data separately.

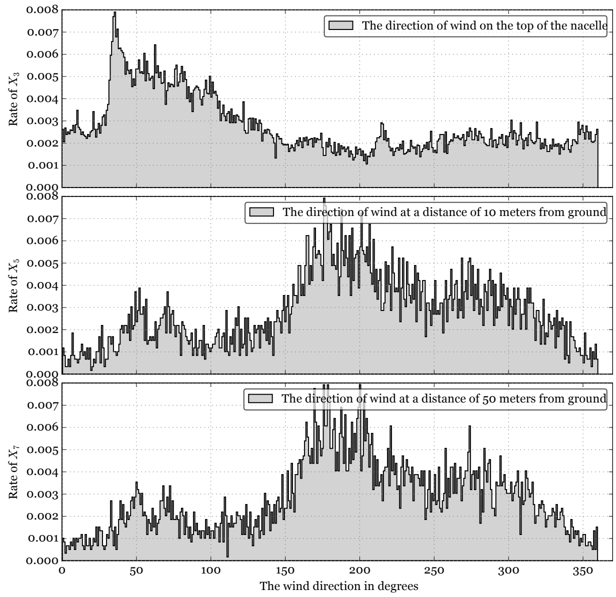

Fig. 1. Histogram of wind direction at three levels of height

PROBABILITY DISTRIBUTIONS

Each of distributions is parameterized by three parameters: α — shape factor, l — location factor and s — scale factor. In the case of the beta distribution the second scale factor is added, denoted by β-letter. All distributions parameters are positive real numbers: $\alpha, \beta, s, l \in \mathbb{R}$, $\alpha, \beta, s > 0$, $l \geqslant 0$.

The probability density function (PDF) of a Log-Normal random variable X is:

$$f_{LN}(x;\alpha,l,s) = \\ = \begin{cases} \dfrac{1}{(x-l)\alpha\sqrt{2\pi}}\cdot \\ \quad\cdot\exp\left(-\dfrac{1}{2}\left(\dfrac{\ln(x-l)-\ln s}{\alpha}\right)^2\right), & x \geqslant l. \\ 0, & x < l. \end{cases}$$

The probability density function of a Weibull (see Frchet (1927); Weibull (1951)) random variable X is:

$$f_W(x;\alpha,l,s) = \\ = \begin{cases} \dfrac{\alpha}{s}\left(\dfrac{x-l}{s}\right)^{\alpha-1}\exp\left[-\left(\dfrac{x-l}{s}\right)^{\alpha}\right], & x \geqslant l, \\ 0, & x < l. \end{cases}$$

The probability density function of a Gamma random variable X is:

$$f_\Gamma(x;\alpha,l,s) = \begin{cases} \dfrac{(x-l)^{\alpha-1}\exp\left(-\frac{(x-l)}{s}\right)}{s^\alpha\Gamma(\alpha)}, & x \geqslant l, \\ 0, & x < l. \end{cases}$$

where $\Gamma(\alpha)$ is gamma-function.

The probability density function of a Beta random variable X is:

$$f_{\mathcal{B}}(x;\alpha,\beta,l,s) = \\ = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{s\Gamma(\alpha)\Gamma(\beta)}\left(\dfrac{x-l}{s}\right)^{\alpha-1}\left(1-\dfrac{x-l}{s}\right)^{\beta-1}, & x \geqslant l, \\ 0, & x < l. \end{cases}$$

If in PDF formulas of Log-Normal, Weibull and Gamma distributions let $l = 0$, and for Beta distribution let $s = 1$, we get the formulas of distributions most frequently used in Norman L. Johnson (1994); Nelson (1982).

DETERMINATION OF DISTRIBUTIONS PARAMETERS

In `scipy.stats` (see Jones et al. (2001)) following objects are defined: `lognorm`, `weibull_min`, `gamma` and `beta`. These objects implement distributions we work with. Every one of these objects has PDF function `pdf(x, a, [b,] loc, scale)` and CDF (cumulative distribution) function `cdf(x, a, [b,] loc, scale)`, where `x` — function argument, `a`, `b` — shape parameters α, (and β for Beta-distribution), `loc` and `scale` are location and scale parameters.

For parameters estimation of our distributions the library `scipy.stats` provides the function `fit(data)`, which calculates the parameters of distributions by maximum likelihood method and the empirical data. We used this function to calculate parameters of the considered distributions. Then we used `pdf` and `cdf` functions to compute values of the probability density function and cumulative distribution function.

There is the example of the code for the case of Log-Normal distribution:

```
s, loc, scale =
    scipy.stats.lognorm.fit(ws1)
xs = np.linspace(np.min(ws1),
    np.max(ws1), 1000)
logN_PDF =
    scipy.stats.lognorm.pdf(xs, s,
    loc, scale)
logN_CDF =
    scipy.stats.lognorm.cdf(xs, s,
    loc, scale)
```

The results are presented graphically on figures 2–9.

The figures were plotted for theoretical distributions, the parameters of which have been determined on the basis of the entire dataset. From the analysis of the quantile-quantile plots (Q-Q plots) we can conclude that the Weibull distribution is best suited for approximation of available data (although only slightly), outmatching them only in the approximation of extreme values of a random variable.

We also performed computations with the considered distributions parameterized by only two parameters

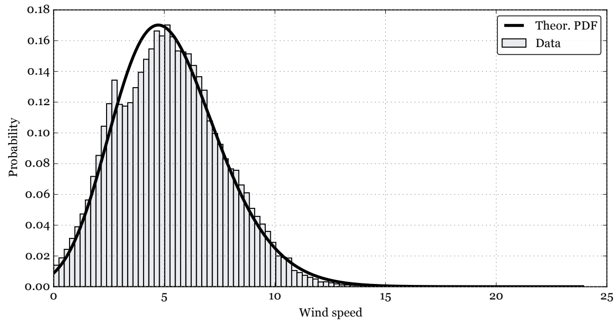

Fig. 2. PDF of **Log-Normal** distribution compared with data histogram

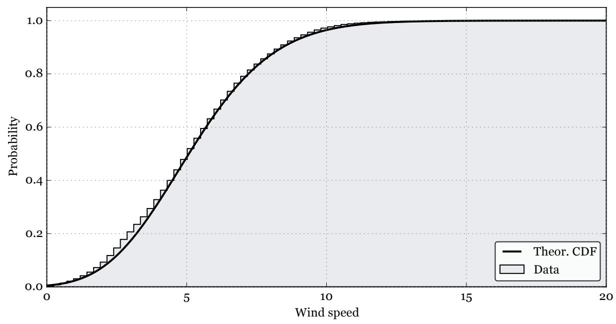

Fig. 3. CDF of **Log-Normal** distribution compared with empirical distribution function

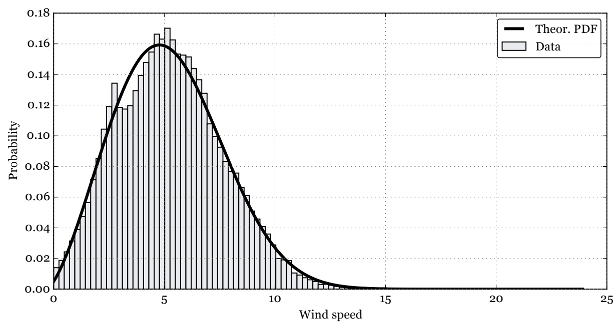

Fig. 4. PDF of **Weibull** distribution compared with data histogram

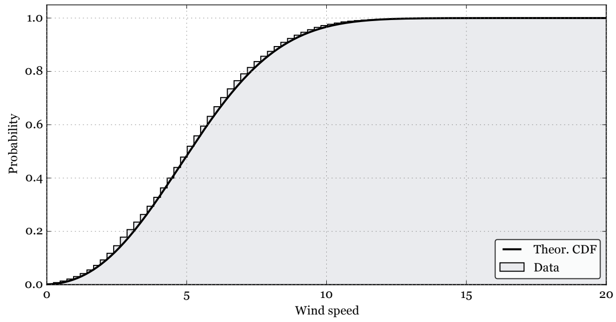

Fig. 5. CDF of **Weibull** distribution compared with empirical distribution function

(let $l = 0$, and for Beta distribution an addition let $s = 1$). After plotting the results of calculations we found out that the two-parameter Weibull distribution has superiority over other two-parameters distributions (Log-Normal, Gamma and Beta), which is not true for

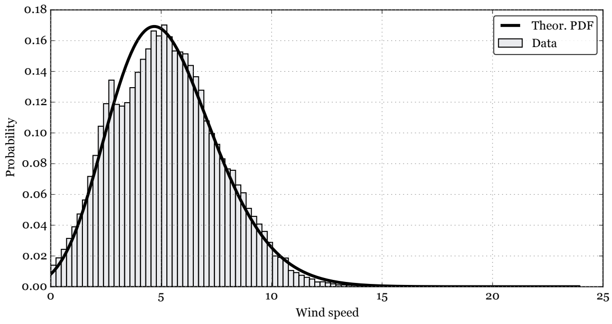

Fig. 6. PDF of **Gamma** distribution compared with data histogram

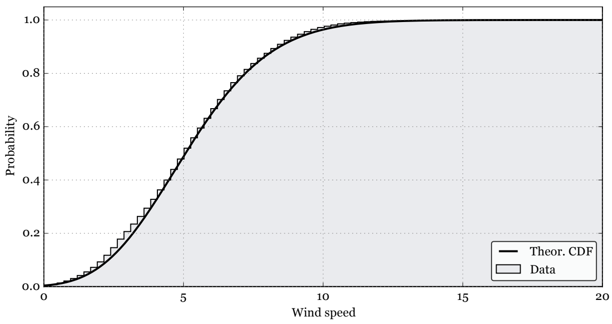

Fig. 7. CDF of **Gamma** distribution compared with empirical distribution function

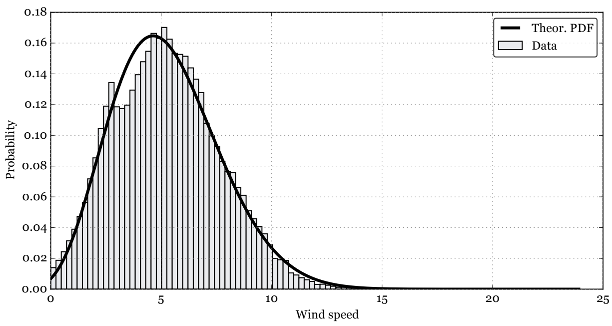

Fig. 8. PDF of **Beta** distribution compared with data histograms

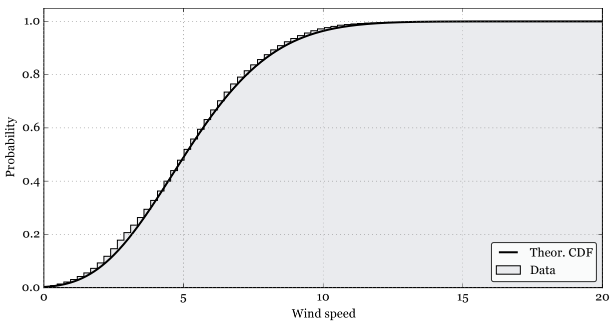

Fig. 9. CDF of **Beta** distribution compared with empirical distribution function

three-parameter case (Fig. 10–13).

CONCLUSIONS

The results of statistical data processing correspond to the results presented in the literature, where Weibull distribution is the most often used distribution for the

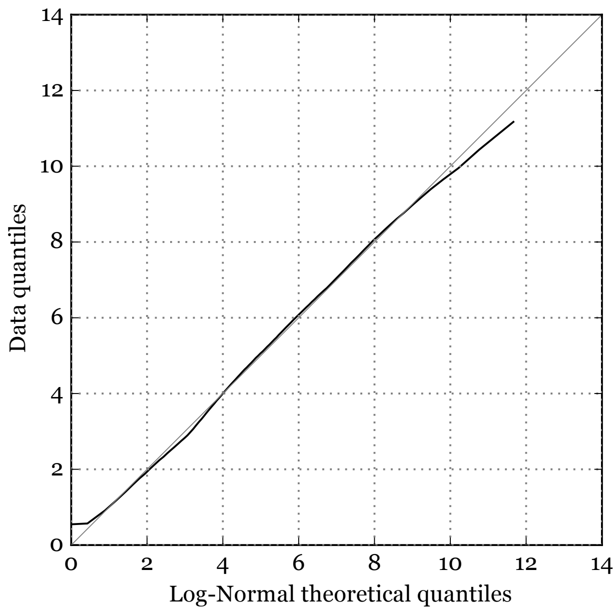

Fig. 10. Q-Q plot for LogNormal distribution

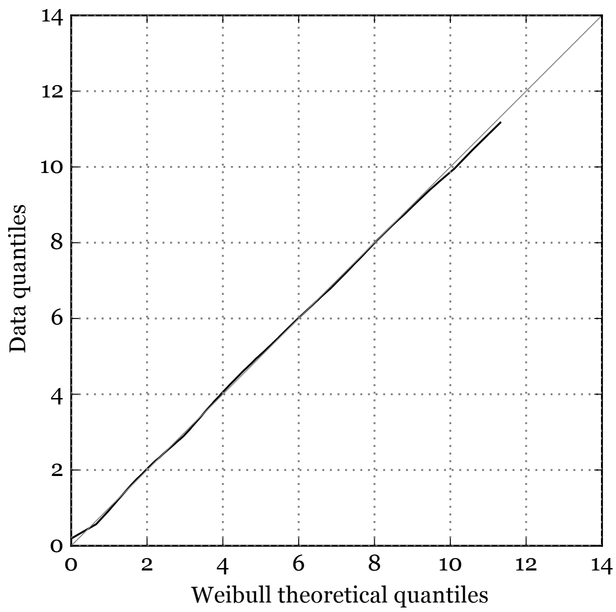

Fig. 11. Q-Q plot for Weibull distribution

wind speed approximation (see Lun and Lam (2000); Seguro and Lambert (2000); Bowden et al. (1983); Yeh and Wang (2008); Islam et al. (2011); Garcia et al. (1998)).

Our future work will be aimed at the construction of stochastic models that can approximate the wind speed depending on time (see Miano and Milano (2015)). On the other hand, we expect to verify the results of this work by using more dilated and large data array.

ACKNOWLEDGMENT

The work is partially supported by RFBR grants No's 15-07-08795 and 16-07-00556. Also the publication was supported by the Ministry of Education

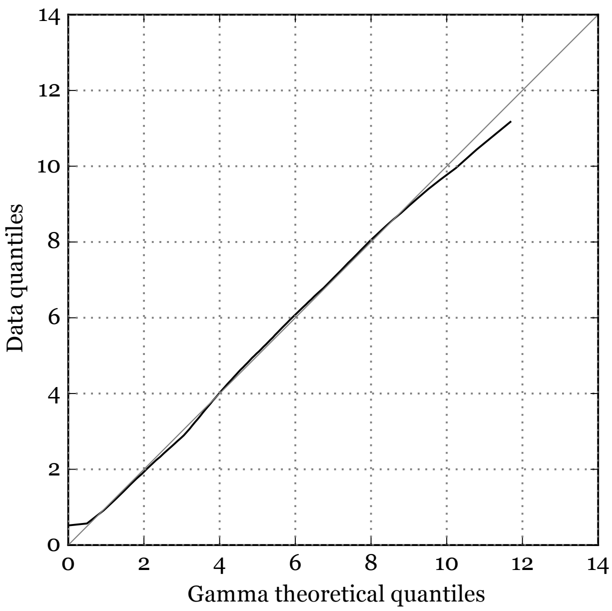

Fig. 12. Q-Q plot for Gamma distribution

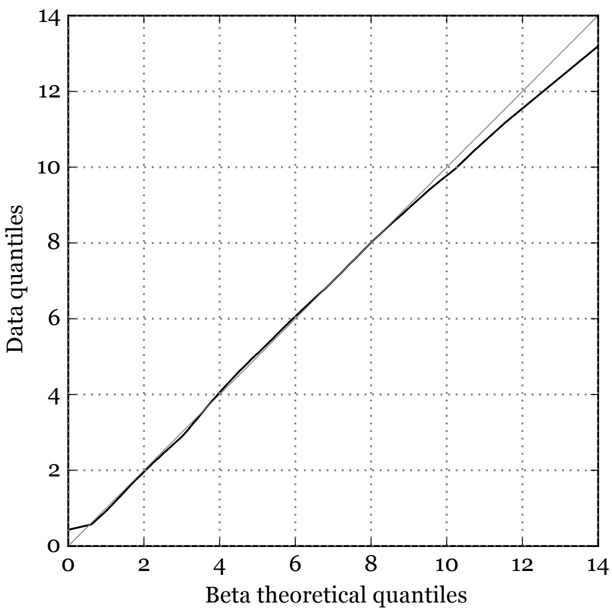

Fig. 13. Q-Q plot for Beta distribution

and Science of the Russian Federation (the Agreement No 02.a03.21.0008). The computations were carried out on the Felix computational cluster (RUDN University, Moscow, Russia) and on the HybriLIT heterogeneous cluster (Multifunctional center for data storage, processing, and analysis at the Joint Institute for Nuclear Research, Dubna, Russia).

REFERENCES

Bowden, G. J., Barker, P. R., Shestopal, V. O. and Twidell, J. W. (1983), The weibull distribution function and wind power statistics, *Wind Engineering* 7, 85–98. Provided by the SAO/NASA Astrophysics Data System.

Droettboom, M., Caswell, T. A., Hunter, J., Firing, E., Nielsen, J. H., Root, B., Elson, P., Dale, D., Lee, J.-J., Varoquaux, N., Seppnen, J. K., McDougall, D., May, R., Straw, A., de Andrade, E. S., Lee, A., Yu, T. S., Ma, E., Gohlke, C., Silvester, S., Moad, C., Hobson, P., Schulz, J., Wrtz, P., Ariza, F., Cimarron, Hisch, T., Kniazev, N., Vincent, A. F. and Thomas, I. (2017), matplotlib/matplotlib: v2.0.0. https://doi.org/10.5281/zenodo.248351

Frchet, M. R. (1927), Sur la loi de probabilit de l'cart maximum, *Annales de la Socit Polonaise de Mathematique* pp. 93–116.

Garcia, A., Torres, J., Prieto, E. and de Francisco, A. (1998), Fitting wind speed distributions: a case study, *Solar Energy* 62(2), 139–144.

Islam, M., Saidur, R. and Rahim, N. (2011), Assessment of wind energy potentiality at kudat and labuan, malaysia using weibull distribution function, *Energy* 36(2), 985–992.

Jones, E., Oliphant, T., Peterson, P. et al. (2001), SciPy: Open source scientific tools for Python. [Online; accessed 19.01.2017]. http://www.scipy.org/

Lun, I. Y. and Lam, J. C. (2000), A study of weibull parameters using long-term wind observations, *Renewable Energy* 20(2), 145–153.

Miano, R. Z. and Milano, F. (2015), Construction of sde-based wind speed models with exponential autocorrelation.

Nelson, W. B. (1982), *Applied Life Data Analysis (Wiley Series in Probability and Statistics)*.

Norman L. Johnson, Samuel Kotz, N. B. (1994), *Continuous Univariate Distributions*, Vol. 1 of *Wiley Series in Probability and Statistics*, Wiley-Interscience.

Norman L. Johnson, Samuel Kotz, N. B. (1995), *Continuous Univariate Distributions, Vol. 2*, Vol. 2 of *Wiley Series in Probability and Statistics*, Wiley-Interscience.

Project Jupyter home (2017). [Online; accessed 19.01.2017]. https://jupyter.org

Seguro, J. and Lambert, T. (2000), Modern estimation of the parameters of the weibull wind speed distribution for wind energy analysis, *Journal of Wind Engineering and Industrial Aerodynamics* 85(1), 75–84.

Weibull, W. (1951), A statistical distribution function of wide applicability, *Journal of Applied Mechanics* pp. 293–297.

Yeh, T. H. and Wang, L. (2008), A study on generator capacity for wind turbines under various tower heights and rated wind speeds using weibull distribution, *IEEE Transactions on Energy Conversion* 23(2), 592–602.

AUTHOR BIOGRAPHIES

MIGRAN N. GEVORKYAN received his Ph.D. in Mathematics in 2013. Since then, he has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). His current research activity

focuses on mathematical modeling. His email address is gevorkyan_mn@rudn.university.

ANASTASIYA V. DEMIDOVA received her Ph.D. in Mathematics in 2014. Since then, she has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). Her email address is demidova_av@rudn.university.

IVAN S. ZARYADOV received his Ph.D. in Mathematics in 2010. Since then, he has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on probability theory. His email address is zaryadov_is@rudn.university.

ROBERT A. SOBOLEWSKI He has worked in Department of Power Engineering, Fotonics and Lighting Technology in Bialystok University of Technology. His email address is r.sobolewski@pb.edu.pl.

ANNA V. KOROLKOVA received her Ph.D. in Mathematics in 2010. Since then, she has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). Her current research activity focuses on mathematical modeling. Her email address is korolkova_av@rudn.university.

DMITRY S. KULYABOV received his Ph.D. in Physics in 2000. Since then, he has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on mathematical modeling. His email address is kulyabov_ds@rudn.university.

LEONID A. SEVASTIANOV received his D.Sc. in Phys.-Math. in 1999. Since then, he has worked as full professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on mathematical modeling. His email address is sevastianov_la@rudn.university.