Statistical Data Processing during Wind Generators Operation

Maksym Zaliskyi and Yuliia Petrova
National Aviation University, Aviation Radioelectronic Complexes Department, Kyiv, Ukraine
Email: maximus2812@ukr.net; panijulia.p@gmail.com

Marlen Asanov and Eskender Bekirov
V.I. Vernadsky Crimean Federal University, Electrical Energy and Electrical Engineering Department, Simferopol, Ukraine
Email: {asanov.m.m; bekirov.e.a}@cfuv.ru

***Abstract*—The article deals with the problem of data processing in wind turbines operation systems. The current status of wind power was analyzed, particularly the prospect of using wind turbines both onshore and offshore for generating electricity in the Black Sea region. Main attention was paid to actual statistical data processing received during the operation of wind turbines. Statistical model of times between failures and times to repairs was analyzed. This analysis made it possible to calculate the probability density function of availability. In addition, mathematical models that specify statistical data on q-q plot were made, which can be used for solving the forecasting problems.**

***Index Terms*—operation system, statistical data processing, mean time between failures, q-q plot**

I. Introduction

Wind power is being developed actively in the last ten years. According to the statistics the maximum net generating capacity of wind power plants increased from 93 553 MW in 2007 to 466 505 MW in 2016 worldwide [1].

The Black Sea region, especially its west part, also has the wind energy usage potential [2], [3]. Wind patterns in the Black Sea basin were analyzed on the basis of continuous observations from 1999 to 2012. Authors [2] processed data collected by 11 meteorological stations and also satellite information. Two reanalysis wind models were used for more detailed wind climate investigation.

As for 2016 the maximum net generating capacity of wind power plants in Turkey was 5 376 MW [1] and all the wind turbines are onshore. In the same time wind energy potential in Turkey according to various estimates is 120 billion kWh [4]. Author [4] noted that Turkey's growing electricity demands can be easily supplied by wind stations. The current status of wind power in Turkey was described. The potentially promising places for installing new wind turbines were considered.

Among the most suitable regions for wind power development in Turkey its west part was marked out [5], [6]. Author [5] analyzed state of the art of wind power on Turkey's territory, provided maps of yearly mean power density for 50 m high and distribution of wind velocity in 30 m high for this region. Methods for computing the wind parameters were considered. Pointed out that Weibull and Rayleigh functions are used most often for modeling.

Offshore wind turbine project on Turkey's territory also shows great promise [7], [8].

Meteorological data of 54 coastal regions was analyzed [7]. As a result authors picked out 5 regions that are most favorable for offshore wind turbine installation and gave own advices for mounting them.

Wind parameters analysis above Black Sea region showed that maximum potential of wind power density is at the North-West part of Black and Azov Seas [8]. Based on the wind parameters data analysis, accumulated for more than 37 years, maps of these parameters for 50 m of altitude were made.

Along with wind power plants quantity increasing the problem of estimating their reliability also grows. Turbines operation efficiency, amount and quality of electricity generated depend directly on steps intended for maintenance of required reliability characteristics.

Wind turbine operational mode mainly depends on environmental conditions, which can change unpredictably and rapidly. Because of the weather changes separate wind turbine components can be under the influence of increased load, which leads to failure and consequently to reliability decrease. Electricity generation by wind turbine is not constant. The operating characteristics of two identical turbines are not equal even if they are located next to each other or especially if they are situated in different places. Wind turbine consists of many components mechanically or electrically connected to each other, which sometimes work independently and not simultaneously. Described factors make wind turbine reliability estimation very difficult and time-consuming process.

Based on the analysis of numerous failure reports collected for more than 15 years authors concluded that wind turbine annual failure rate considerably depends on

Manuscript received January 25, 2018; revised March 10, 2018; accepted June 18, 2018.
Corresponding author: Maksym Zaliskyi (email: maximus2812@ ukr.net)

doi: 10.18178/ijeetc.8.1.33-38

their power: the higher turbine power, the more frequent failures occur [9], [10]. The greatest number of failures observed in electrical system (23%) and plant control system (18%). In the same time, downtime is biggest during failures of generator, gearbox and drive train.

Similar results were obtained in [11]. After processing data on wind turbines received in Denmark, Germany and Sweden authors picked out subassemblies with the highest failure frequencies, which are electrical system, rotor (i.e. blades & hub), converter (i.e. electrical control, electronics, inverter). The reliability characteristics of turbines of different power with several types of used generators were considered.

Author analyzed quite a number of databases in [12]. As a result, the system with the most frequent failures occurs is power electronics or power module. The gearbox has biggest downtime.

Wind turbines reliability estimation aimed at organizing its optimal operation and maintenance plays an important role in scientific researches [13]. Authors [13] conditionally divided these researches into three types: works that are devoted to reliability analysis of separately taken turbine, whole wind power plant or series of one-type turbines and, finally, overall power facilities. Wind turbines maintenance since they are operating and wearing out becomes more and more costly. Turbine downtime also increases due to repair works. To the greater extent it may concern offshore wind power plants. Reliability analysis methods that can be used for assess reliability of a wind turbine are considered in [13]. Simulation studies for wind power plant operations are presented. Using the standard reliability estimation methods will not always be effective because a large number of factors affect wind turbine reliability.

To keep the wind turbines reliability at a high level their operation system is used. The processes implemented in the operation system are intended use, maintenance, repair, resource extension etc. Statistical data processing procedures are usually used for effective operation control. This kind of process is made at all stages of device life time. The main tasks to perform data processing are making statistical models of device operation, reliability parameters estimation during certain observation interval, device technical state predicting, resource identifying etc. One of the statistical data source for processing is times between failures of the certain device.

II. Literature Analysis and Problem Statement

Literature analysis as regards to technical systems operation showed that great attention is paid to operation systems designing and updating [14]-[16]. At the same time statistical data processing procedures are being developed.

Issues related to technical systems operation namely reliability indices calculation are considered in [14], [15]. Author [16] suggested a new method for operation system designing based on four principles: adaptability, aggregation, system and process approaches.

In general, literature analysis shows the topicality of statistical data processing for technical systems operation. Timely made managerial decisions based on statistical data processing results favor the increasing of operation system efficiency.

Therefore this work will consider models analysis of statistical data collected during wind turbine operation.

III. Statistical Data Analysis in Operation System of Wind Generators

A. *Statistical Data Model Analysis*

The following reliability characteristics of technical devices are used during statistical data processing in these devices operation:

– Mean time between failures,
– Mean time to repair,
– Steady-state availability,
– Time-dependent availability,
– Gamma-percentile operating reliability measure, etc.

Various distribution laws of times between failures and time to repair can be used for describing the technical devices reliability characteristics. These laws include Weibull distribution, exponential distribution, inverse Gaussian distribution, Birnbaum-Saunders distribution etc. [17]. However, exponential distribution is widely used to simplify calculations in engineering applications.

The actual operation statistical data allows providing device operation process management more appropriate that in turn increases the operation system efficiency.

Let us find the model of actual statistical data collected during wind turbine operation. In this research authors consider one of 177 USW 56-100 type wind turbines included in the Sakskaya wind plant that is situated on the Crimea peninsula. We observed 36 failures (i=1, 2, ⋯, 36) of the wind turbine #112 during operation. Table I shows the total time x_i between failures of the turbine.

TABLE I. Total Time between Failures of Wind Generator #112

Failure i	Total time x_i (hour)	Failure i	Total time x_i (hour)	Failure i	Total time x_i (hour)
1	2445	13	1885	25	4737
2	330	14	3473	26	397
3	73	15	3032	27	1012
4	992	16	10825	28	1257
5	3736	17	3321	29	1788
6	96	18	5007	30	1884
7	1163	19	11331	31	11
8	3750	20	14493	32	1875
9	589	21	643	33	1739
10	44	22	5566	34	352
11	5986	23	1734	35	1885
12	6223	24	1872	36	3473

From Table I we can find the mean time T_0 between failures:

$$T_0 = \frac{1}{36}\sum_{1}^{36} x_i \cong 3028.3\,\text{hours}$$

Standard chi-squared test is used to check the hypothesis about exponential model of times between failures. Initial failure data is a source for plotting the histogram shown in Fig. 1.

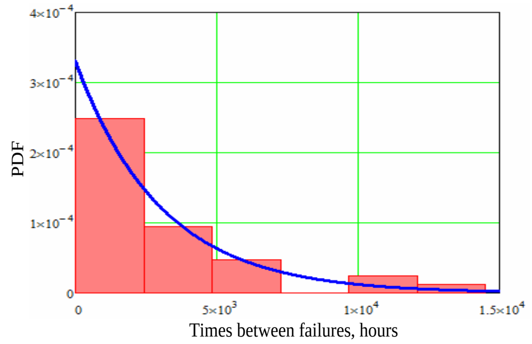

Fig. 1. PDF of times between failures.

In addition in Fig. 1 probability density function (PDF) of theoretical exponential distribution is given as:

$$f_0(t) = 3.304 \cdot 10^{-4} e^{-3.30410^{-4} t} h(t)$$

where $h(t)$ is a Heaviside function.

Calculated chi-square index value is equal to:

$$\chi^2_{\text{calc}} = 3.874$$

which is less than threshold value $\chi^2_{\text{th}} = 9.5$, so the hypothesis about exponential distribution law of times between failures is accepted with a significance level equal to 0.05.

We recorded the repair time of 36 failures (i=1, 2, ⋯, 36) of the wind turbine #112 during operation. Table II shows the total time r_i to repair of the turbine.

TABLE II. TOTAL TIMES TO REPAIRS OF WIND GENERATOR #112

Failure i	Repair time r_i (hour)	Failure i	Repair time r_i (hour)	Failure i	Repair time r_i (hour)
1	3	13	3	25	2.5
2	25	14	3	26	120.5
3	24	15	0.5	27	201
4	4	16	2	28	226
5	6	17	2	29	481
6	99	18	10	30	102
7	3.5	19	9	31	2
8	178	20	2.5	32	303.5
9	111	21	214	33	164
10	20	22	11	34	3.5
11	72	23	21	35	15
12	9	24	1.5	36	354

From Table II we can find the mean time T_R to repair:

$$T_R = \frac{1}{36}\sum_{1}^{36} r_i \cong 78.028 \text{ hours}$$

While checking the hypothesis about exponential distribution law of time to repair by chi-square test, five intervals of histogram were taken and following value was calculated:

$$\chi^2_{\text{calc}} = 5.762$$

which is less than threshold value $\chi^2_{\text{th}} = 7.8$, so the hypothesis about exponential distribution law of equipment time to repair is accepted with a significance level equal to 0.05.

So, probability density function of time to repair is the following:

$$f_1(t) = 0.013 e^{-0.013t} h(t)$$

Let's find the statistical model of availability.

In the probabilistic sense, availability is equal to the ratio of mean time between failures to the sum of mean time between failures and mean time to repair. In the statistical sense, availability factor is defined as:

$$A = \frac{\hat{T}_0}{\hat{T}_0 + \hat{T}_R}$$

where $\hat{T}_0$ and $\hat{T}_R$ are estimates of mean time between failures and mean time to repair.

For the exponential models mean distribution value obeys the chi-square distribution. That is, for considered statistical data distributions of mean time between failures and mean time to repair can be written as:

$$f_{\hat{T}_0}(t) = 5 \cdot 10^{-110} t^{35} e^{-0.011895t} h(t)$$

$$f_{\hat{T}_R}(t) = 7.793 \cdot 10^{-53} t^{35} e^{-0.4614t} h(t)$$

After using the standard functional transformation method of system of two random values in general form function will look as following:

$$f(A) = \frac{1}{(1-A)^2} \int_0^\infty t f_{\hat{T}_0}\left(\frac{At}{(1-A)}\right) f_{\hat{T}_R}(t) dt$$

The expression for probability density function of availability is found by solving this integral and is as following:

$$f(A) = \frac{5.117 \cdot 10^{-36} A^{35} (1-A)^{35}}{(1-0.9742A)^{72}}, \quad 0 \le A \le 1$$

Probability density function of availability is shown in Fig. 2.

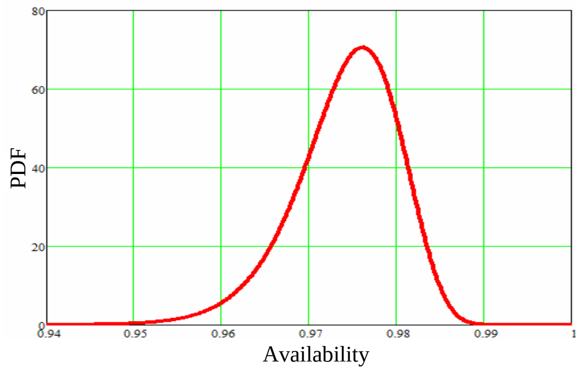

Fig. 2. PDF of availability.

Calculated value of availability according to the wind turbine operation data is equal to 0.975. As from the Fig. 2 this value is within essential values of calculated probability density function.

In general availability models for different probability density function of times between failures and times to repairs are given in [18].

B. q-q Plot Usage for Determining Extremal Values of Distribution

Sometimes need to find the extremum values of operating time (minimum and maximum values) to calculate, for example, maintenance rate or wind turbine service life.

During mathematical models construction and testing hypothesis about statistical data probability distribution the so called p-p plot and q-q plot are used in mathematical statistics.

Let's consider the procedures for making the mathematical model of actual statistical data (Table I) using q-q plot. Well known that q-q plot shows the dependence of sample values (or their logarithms) that are sorted in increasing order on quantile of the given distribution type. Normal distribution $N(0;\ 1)$ and exponential distribution with parameter equal to $3.304 \cdot 10^{-4}$ will be taken as a verifiable distribution type.

Empirical probabilities for each point are found by the following formula:

$$P_i = \frac{i}{n+1}$$

where i is a number of random value in order statistic ($i \in [1;n]$), $n = 36$ is a sample size.

Quantiles of distribution are found for each empirical probability according to the following expression:

$$V(P_i) = \inf\left(s \in \mathrm{R} : P_i \le F(s)\right)$$

where $F(s)$ is a given distribution function.

q-q plot constructing reduces to drawing the ranked logarithms of failures data $\ln(t_i)$ on the quantiles grid.

After that, obtained dependence is approximated by linear and piecewise-linear functions according to the least-squares method. If the linear approximation $y(x) = a_0 + a_1 x$ is built, then its coefficients a_0 and a_1 are found from the following system of equations:

$$\begin{cases} n a_0 + a_1 \sum_i x_i = \sum_i y_i, \\ a_0 \sum_i x_i + a_1 \sum_i (x_i)^2 = \sum_i x_i y_i. \end{cases}$$

If the approximation is made by piecewise-linear polygonal regression:

$$y(x) = a_0 + a_1 x + a_2 x^+,$$
$$x^+ = (x - x_{sw}) h(x - x_{sw}),$$

where x_{sw} is a switching point, then coefficients a_0, a_1 and a_2 are found by solving the following system of equations:

$$\begin{cases} n a_0 + a_1 \sum_i x_i + a_2 \sum_i x_i^+ = \sum_i y_i, \\ a_0 \sum_i x_i + a_1 \sum_i (x_i)^2 + a_2 \sum_i x_i x_i^+ = \sum_i x_i y_i, \\ a_0 \sum_i x_i^+ + a_1 \sum_i x_i x_i^+ + a_2 \sum_i \left(x_i^+\right)^2 = \sum_i x_i^+ y_i. \end{cases}$$

Using the polygonal regression implies the evaluation of transition point from one line to another (switching point). In the case of theoretical normal distribution was considered that switching point corresponds to median of the distribution ($V(P_i) = 0$).

The following analytical relations for the respective approximation functions can be written after calculating the coefficients of linear and piecewise-linear approximation for the considered statistical data:

$$y(x) = 7.264 + 1.614x, \quad y(x) = 7.858 + 2.409x - 1.589x^+$$

Data for the theoretical normal distribution and its approximation by linear (LNM) and piecewise-linear (PWNM) function is on q-q plot that is shown in Fig. 3.

The statistical data test for linearity by method suggested in [19] showed that this data is nonlinear. Therefore piecewise-linear approximation of statistical data is more adequate for finding the extremum values of distribution.

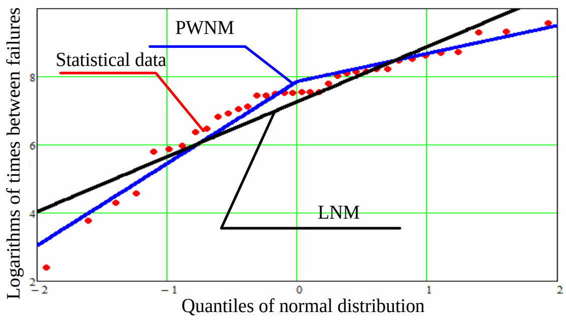

Fig. 3. Q-q plot in case of normal theoretical distribution.

For the theoretical exponential distribution the same calculations were made and the following analytical relations for respective approximation functions were obtained:

$$y(x) = -572.7 + 1.249x$$

$$y(x) = -72.868 + 0.968x + 0.726(x - 5039)^+$$

But in this case there are two features. Firstly, in q-q plot not logarithms are marked, but actual values of times between failures. Secondly, during approximation by piecewise-linear function switching point x_{swopt} optimization was made according to the following rule:

$$x_{swopt} = \inf\left(s \in \mathrm{R} : \sigma(s) \le \sigma(x_{swi})\right)$$

where $\sigma(x_{swi})$ is a standard deviation of statistical data from piecewise-linear approximation function values for all possible switching points values x_{swi}.

Data for the theoretical exponential distribution and its approximation by linear (LEM) and piecewise-linear (PWEM) function is on q-q plot that is shown in Fig. 4.

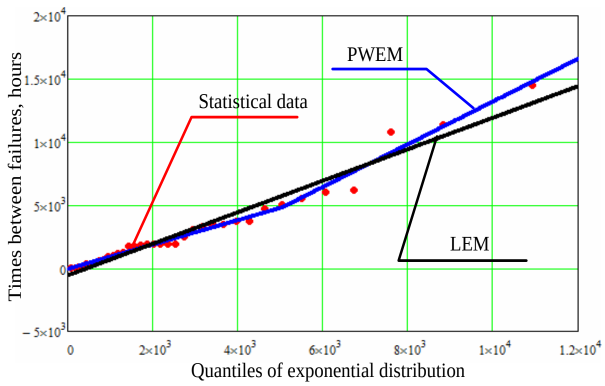

Fig. 4. q-q plot in case of exponential theoretical distribution.

The statistical data test in this q-q plot by method suggested in [19] also showed that this data is nonlinear. Therefore piecewise-linear approximation is also more adequate for finding the extremum values.

Table III contains extremum values of empirical probabilities of failure P_i and corresponding values for time to failure (in hours) calculated according to LNM, PWNM, LEM and PWEM models.

TABLE III. RESULTS OF FORECASTING

P_i	0.9	0.95	0.99	0.995	0.999
LNM	11300	20310	61020	91270	209400
PWNM	7392	9955	17400	21350	32540
LEM	8132	10750	16840	19460	25540
PWEM	8072	11630	19880	23430	31680

IV. CONCLUSION

Rapid development of wind power in the last ten years is indicative of its great potential. Along with the increase of wind turbines amount there is a growing need for their proper and timely maintenance. Wind turbine, in turn, is an expensive and complex device that complicates its repair procedures and raises the significance of evaluating its reliability characteristics.

As practice shows, technical devices operation system specifies the efficiency of their intended use. In particular, this efficiency is determined by available procedures of statistical data processing and making corresponding decisions. So, topical scientific and engineering task of processing the statistical data concerning operating and repair time was considered. The distribution law of this data was checked and mathematical model that describes it was built using piecewise-linear regression on q-q plot.

Obtained probability density function of availability for wind turbine operational data has mathematical expectation that is approximately equal to point estimate of availability calculated by classical formula and is 0.975. Four models (LNM, PWNM, LEM, PWEM) for statistical data description calculated in this work can be used for solving forecasting problems. Calculation results of distribution extremum values verified the advantages of polygonal models that are more adequately specify statistical data (the least standard deviation that is equal to 434.75 has PWEM model).

Obtained results can be used during solving the design problems or updating the wind generators operation system.

REFERENCES

[1] Renewable Capacity Statistics 2017. Abu Dhabi: International Renewable Energy Agency (IRENA), 2017.

[2] F. Onea and E. Rusu, “Wind energy assessments along the Black Sea basin,” *Meteorol. Appl.*, vol. 21, no. 2, pp. 316–329, 2014.

[3] F. Onea and E. Rusu, “Efficiency assessments for some state of the art wind turbines in the coastal environments of the Black and the Caspian seas,” *Energ. Explor. Exploit.*, vol. 34, no. 2, pp. 217–234, 2016.

[4] C. Ilkilic, “Wind energy and assessment of wind energy potential in Turkey,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 2, pp. 1165–1173, 2012.

[5] M. Gokcek, “Developing wind energy in Turkey,” in *Paths to Sustainable Energy*, Artie Ng, Ed. InTech, 2010, pp. 315–330.

[6] M. E. Basoglu and B. Cakir, “Promising energy sources within the framework of 2023 energy targets in Turkey: Solar and wind,” *International Journal of Electronics and Electrical Engineering*, vol. 5, no. 2, pp. 183–188, 2017.

[7] M. Argın and V. Yerci, “The assessment of offshore wind power potential of Turkey,” in *Proc. 9th Int. Conf. on Electrical and Electronics Engineerin*, Nov. 2015, pp. 966-970.

[8] B. Aydogan, “Offshore wind power atlas of the Black Sea Region,” *J. Renew. Sustain. Ener.*, vol. 9, Feb. 2017.

[9] B. Hahn, M. Durstewitz, and K. Rohrig, “Reliability of wind turbines - Experiences of 15 years with 1500 WTs,” Institut für Solare Energieversorgungstechnik (ISET), Verein an der Universität Kassel, Kassel, Germany, 2006.

[10] D. Topic, D. Sljivac, and M. Stojkov, “Availability of different wind power plant configurations based on components performance statistics,” *Int. Rev. Electr. Eng-I*, vol. 10, no. 3, pp. 414–420, 2015.

[11] P. Tavner, F. Spinato, G. van Bussel, and E. Koutoulakos, “Reliability of different wind turbine concepts with relevance to offshore application,” in *Proc. European Wind Energy Conf. & Exhibition*, March 2008, pp. 166–170.

[12] S. Sheng, “Report on wind turbine subsystem reliability—A survey of various databases,” Nat. Renew. Energy Lab., Washington, DC, USA, Tech. Rep. REL/PR-5000-59111, Jun. 2013.

[13] E. Byon, L. Ntaimo, C. Singh, and Y. Ding, “Wind energy facility reliability and maintenance,” in *Handbook of Wind Power Systems: Optimization, Modeling, Simulation and Economic Aspects*, Pardalos, Rebennack, Pereira, Iliadis and Pappu, Ed., Springer, 2013, pp. 639–672.

[14] A. Hoyland and M. Rausand, *System Reliability Theory*, New York: John Wiley & Sons, Inc., 1994.

[15] T. Nakagawa, *Maintenance Theory of Reliability*, London: Springer-Verlag, 2005.

[16] O. V. Solomentsev, V. H. Melkumyan, M. Y. Zaliskyi, and M. M. Asanov, “UAV operation system designing,” in *Proc. of IEEE 3rd Int. Conf. on Actual Problems of Unmanned Air Vehicles Developments*, October 2015, pp. 95–98.

[17] Horst Rinne. (2014). The Hazard Rate: Theory and Inference (with supplementary MATLAB–Programs). [Online]. Available: https://core.ac.uk/display/56346083

[18] O. Solomentsev, M. Zaliskyi, and O. Zuiev, “Estimation of quality parameters in the radio flight support operational system,” *Aviation*, vol. 20, no. 3, pp. 123–128, 2016.

[19] V. Kuzmin, O. Solomentsev, and M. Zaliskyi, “Methodical approach for statistical data testing on linearity,” *Problems of Informatization and Control*, vol. 52, no. 4, pp. 63–67, 2015. (in Ukrainian).

Maksym Zaliskyi, Assoc. prof., Ph.D. (Eng). Date of birth: 1984. Education: National Aviation University, Kyiv, Ukraine (2007). Affiliations and functions: associate professor in the Department of Aviation Radioelectronic Complexes, National Aviation University, Kyiv, Ukraine. Research interests: statistical data processing, reliability, radioelectronic devices operation systems. Publications: 70 scientific papers.

Marlen Asanov, Assoc. prof., Ph.D. (Phys-Math). Date of birth: 1984. Education: National Aviation University, Kyiv, Ukraine (2007). Affiliations and functions: associate professor in the Electrical Energy and Electrical Engineering Department, V.I. Vernadsky Crimean Federal University, Simferopol, Ukraine. Research interests: Renewable energy sources, statistical data processing, reliability analysis, mathematical modeling. Publications: 55 scientific papers.

Eskender Bekirov, Prof., Dr (Eng). Date of birth: 1951. Education: Kazakh Polytechnic Institute, Almaty, Kazakhstan (1974). Affiliations and functions: professor, head of the Electrical Energy and Electrical Engineering Department, V.I. Vernadsky Crimean Federal University, Simferopol, Ukraine. Research interests: Renewable energy sources, statistical data processing. Publications: 210 scientific papers.

Yuliia Petrova, Assoc. prof., Ph.D. (Eng). Date of birth: 1977. Education: National Aviation University, Kyiv, Ukraine (2001). Affiliations and functions: associate professor in the Department of Aviation Radioelectronic Complexes, National Aviation University, Kyiv, Ukraine. Research interests: statistical data processing, radioelectronic devices operation systems. Publications: 25 scientific papers.