

Statistical analysis of the electric energy production from photovoltaic conversion using mobile and fixed constructions

Artur Bugała^{1,*}, Karol Bednarek¹, Leszek Kasprzyk¹, and Andrzej Tomczewski¹

¹Institute of Electrical Engineering and Electronics, Poznan University of Technology, Piotrowo 3a Street, 60-965 Poznań, Poland

Abstract. The paper presents the most representative – from the three-year measurement time period – characteristics of daily and monthly electricity production from a photovoltaic conversion using modules installed in a fixed and 2-axis tracking construction. Results are presented for selected summer, autumn, spring and winter days. Analyzed measuring stand is located on the roof of the Faculty of Electrical Engineering Poznan University of Technology building. The basic parameters of the statistical analysis like mean value, standard deviation, skewness, kurtosis, median, range, or coefficient of variation were used. It was found that the asymmetry factor can be useful in the analysis of the daily electricity production from a photovoltaic conversion. In order to determine the repeatability of monthly electricity production, occurring between the summer, and summer and winter months, a non-parametric Mann-Whitney U test was used as a statistical solution. In order to analyze the repeatability of daily peak hours, describing the largest value of the hourly electricity production, a non-parametric Kruskal-Wallis test was applied as an extension of the Mann-Whitney U test. Based on the analysis of the electric energy distribution from a prepared monitoring system it was found that traditional forecasting methods of the electricity production from a photovoltaic conversion, like multiple regression models, should not be the preferred methods of the analysis.

1 Introduction

Human beings obtain energy in various ways to subsequently use it to meet their specific functional purposes. The most operationally convenient form of energy is electric energy as it can be transported, stored, and processed. Currently a lot of attention is paid to all the solutions connected with using energy from ecological sources, such as photovoltaic systems, wind turbines, biomass, tidal energy, etc. [1-15].

During the processes of design and execution of power supply systems from renewable sources of energy various types of needs and requirements have to be considered. As the energy is generated and supplied to the recipients – that is currents and voltages occur in the elements of the system – the occurrence and impact of electromagnetic fields should be taken into account. This is connected both with the correctness of operation of all the elements of the system (associated with generation, processing and transport of energy) as well as the impact on the environment (ecological reasons), but also rationalization of energy management. It is important to consider these impacts on various levels: both high current systems (connected with energy generation and supply) and microstructures (photovoltaic conversion of energy or operation of electronic elements) [16-27].

All human activity is associated with optimization behaviours and considerations. People try to achieve their goals in a way most beneficial for them. This is associated with a great variety of optimization criteria, methods of creating criterial functions and the used methods, which depend on a given task [4, 5, 22, 28-36].

The study discusses optimum energy yield achieved from the photovoltaic conversion for the PV modules installed in a fixed and two-axis tracking construction. Due to strongly stochastic character of obtaining energy (sunlight depending on seasons of the year, weather conditions, etc.), the results of the conducted tests were subject to statistical processing, which constitutes the main part of this article [4, 5, 37-43].

2 Parameters of Statistical Analysis

The study presents the exemplary courses of daily variation of electrical energy production with the use of photovoltaic modules in the fixed and two-axis tracking construction for the selected day of the summer, autumn, spring, and winter month. On the basis of the analysis of daily distribution of values of the registered electrical energy, it is possible to indicate the days with almost symmetrical shape of the course of this value in relation to the afternoon hours and the days characterized by strong irregularities. The skewness coefficient described by the following relation was used as a measure of asymmetry [43]:

$$A_s = \frac{\bar{x} - M_o}{s} \quad (1)$$

where:

\bar{x} – average value,
 M_o – median,
 s – standard deviation.

* Corresponding author: artur.bugala@put.poznan.pl

The higher value of the skewness coefficient, the bigger asymmetry in relation to the average value. The asymmetry factor equalling zero indicates the symmetry of the variable's distribution, the positive value indicates right side asymmetry, while the negative value left side asymmetry.

The coefficient of variation vis the quotient of variation of a given value around the average value from the population (standard deviation of the population) to the designated average value. It is assumed that if the determined coefficient of variation does not exceed 10%, the features indicate statistically insignificant variation [42]. Big values of standard deviation in relation to the average value may lead to the limitation of the quality of the forecasting model.

Kurtosis is a measure of the flatness or peakedness of the obtained distribution in relation to the normal distribution. Its value equals zero for the theoretical distribution.

Repeatability of the monthly electrical energy production from photovoltaic conversion was defined with the use of the non-parametric Mann-Whitney U test. The measure of central tendency of the test is median. Zero and alternative hypothesis may be described as follows:

- H_0 : distribution of average ranks of observations in the analysed groups does not differ significantly – the samples come from one population (the medians of the tested variable in both groups do not significantly differ between one another);
- H_1 : distribution of average ranks of observations in the analysed groups differs significantly – the samples come from different populations (the medians of the tested variable differ between one another).

Depending on the population's size, it is possible to calculate the value of the test statistics and on its basis the probability value p , which is compared with the significance level α from the distribution tables, using the following relation [43]:

- for small size of the population:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad (2)$$

where:

n_1, n_2 – respective sizes of the first and second population;
 R_1, R_2 – sums of ranks of the elements from the first and second population;

- for big size of the population:

$$Z = \frac{U' - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 \sum (t^3 - t)}{12(n_1 + n_2)(n_1 + n_2 - 1)}}} \quad (3)$$

$$U' = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \quad (4)$$

where:

t – number of cases included in the tied rank.

The choice of the right hypothesis can be presented with the following relation:

$$p \leq \alpha \Rightarrow H_1 \cup p > \alpha \Rightarrow H_0 \quad (5)$$

Repeatability of the daily peak hours in which the highest value of hourly electrical energy production from photovoltaic conversion is observed was described with the use of the non-parametric Kruskal-Wallis test. This test is used to verify the hypothesis about lack of significance of the differences between the medians of the tested variable in a few populations ($k > 2$). It should be assumed that distributions of the measured value are close to each other.

The zero and alternative hypotheses can be recorded as follows:

- H_0 : distribution of average ranks of observations in the analysed k groups does not differ significantly – the samples come from one population (the medians of the tested variable in all groups do not significantly differ between one another);
- H_1 : distribution of average ranks of observations in the analysed k groups differs significantly (the medians of the tested variable differ between one another in individual groups).

The value of test statistics H can be determined using the following relation [41]:

$$H = \left(\frac{12}{N(N+1)} \sum_{j=1}^k \left(\frac{\sum_{i=1}^{n_j} R_{i,j}}{n_j} \right)^2 \right) - 3(N+1) \quad (6)$$

$$N = \sum_{j=1}^k n_j \quad (7)$$

where:

n_j – size of the populations for ($j=1,2,\dots,k$),
 $R_{i,j}$ – ranks assigned to the variable value for ($i=1,2,\dots,n_j$),
 $(j=1,2,\dots,k)$,
 t – number of cases included in the tied rank.

The probability value p determined on the basis of the test statistics H should be compared with the assumed significance level α .

The choice of the appropriate hypothesis can be presented with the following relation:

$$p \leq \alpha \Rightarrow H_1 \cup p > \alpha \Rightarrow H_0 \quad (8)$$

3 Results of Statistical Analysis

The selected results of statistical analysis of daily distribution of electrical energy production for the recommended days of the summer, autumn, spring, and winter months, including determination of the average value, standard deviation, skewness and coefficient of variation for the photovoltaic module installed in the fixed (us) and two-axis tracking construction (un) were presented in Table 1. The method of determination of the recommended days of the year with their calculation for the typical meteorological year of Poznan was presented in [38].

Table 1. Results of statistical analysis of distribution of the daily electrical energy production for the selected days of the months: July, December, April, and January.

parameter	12.07.14		03.07.14		26.07.14	
	un	us	un	us	un	us
average/day	84.63	49.92	74.29	45.75	70.67	44.58
average/sunny hours	135.40	85.57	111.4	78.43	113.07	76.43
standard deviation/day	76.91	62.19	69.06	57.93	69.08	54.54
standard deviation/sunny hours	50.89	59.82	54.90	56.48	53.30	51.64
skewness/day	0.00	0.84	0.09	0.93	0.19	0.79
coefficient of variation/day	0.91	1.25	0.93	1.27	0.98	1.22
coefficient of variation/sunny hours	0.38	0.70	0.49	0.72	0.47	0.68
	18.12.14		11.12.14		29.12.14	
average/day	0.79	1.25	9.29	6.83	6.83	4.79
average/sunny hours	4.75	6.00	37.17	27.33	23.43	16.43
standard deviation/day	1.98	2.70	22.71	15.21	14.07	9.70
standard deviation/sunny hours	2.17	2.53	32.05	19.10	17.04	11.45
skewness/day	2.72	2.28	2.93	2.70	2.33	2.23
coefficient of variation/day	2.50	2.16	2.44	2.23	2.06	2.02
coefficient of variation/sunny hours	0.46	0.42	0.86	0.70	0.73	0.70
	11.04.14		25.04.14		02.04.14	
average/day	37.29	23.29	55.42	40.21	55.54	41.58
average/sunny hours	63.93	46.58	95.00	74.23	111.08	83.17
standard deviation/day	49.01	30.95	69.14	55.76	69.60	59.00
standard deviation/sunny hours	49.13	28.81	66.59	56.69	59.31	59.19
skewness/day	1.34	0.99	0.74	1.04	0.85	1.17
coefficient of variation/day	1.31	1.33	1.25	1.39	1.25	1.42
coefficient of variation/sunny hours	0.77	0.62	0.70	0.76	0.53	0.71

Table 1 (cont.)

	18.01.15		08.01.15		15.01.15	
average/day	1.29	1.71	1.42	1.75	11.71	8.46
average/sunny hours	6.20	8.20	8.50	7.00	46.83	29.00
standard deviation/day	3.21	3.52	3.58	3.92	22.58	17.06
standard deviation/sunny hours	4.35	2.48	4.09	4.97	19.84	20.05
skewness/day	3.06	1.94	2.88	2.59	1.71	2.12
coefficient of variation/day	2.48	2.06	2.53	2.24	1.93	2.02
coefficient of variation/sunny hours	0.70	0.30	0.48	0.71	0.42	0.69

In the analysed case the value A_s for the groups of days of the summer month (03.07.2014, 12.07.2014, 26.07.2014) is close to zero, indicating almost symmetrical distribution. For winter months, the increase in the asymmetry factor both for the value of electrical energy produced by the photovoltaic module in the fixed and two-axis tracking construction can be observed. The highest values of this parameter, exceeding 2, were observed for the winter months (08.01.2015, 15.01.2015, 18.01.2015) with strongly random distribution of the surface density of the solar power radiation. The symmetry of distribution of the daily production of electrical energy determines its bigger predictability, which makes it easier to make a decision on the method of installation of photovoltaic modules in the photovoltaic system and to determine the possibility of covering the daily energy demand of the powered objects during morning to evening hours.

In order to determine the repeatability of the courses of the monthly electrical energy production from photovoltaic conversion occurring between summer months and summer and winter months, a non-parametric Mann-Whitney U test was used, which consisted in comparing two independent groups. In this case the normal distribution of quantitative variables and the equal size of the groups is not required. It is difficult to meet the last condition in the analysed cases due to various numbers of days in individual months of the year. Also, the used non-parametric method is relevant for populations of small size, where variables are measured in a quantitative, ordinal or dichotomous scale. The starting point was formulating test hypotheses: zero H_0 and alternative H_1 :

- H_0 : distribution of average ranks of observations in the analysed groups (months of the year) does not differ significantly;
- H_1 : there are significant differences in the distributions of variables in both groups (months of the year).

The observations with equal value in the created ordered series were assigned ranks – also tied ranks. The detailed results of the Mann-Whitney U test obtained in *Statistica* software for June, July, August, December, and January are presented in Table 2.

Table 2. Results of Mann-Whitney U test for June, July, August, December, January 2014 and 2015.

period	month: July 2014, December 2014								
variable	sum of rang A	sum of rang B	U	Z	p	Z with correction	p	N valid A	N valid B
energy	1239.00	714.00	218.00	3.688586	0.000226	3.702927	0.000213	31	31
period	month: June 2014, January 2015								
variable	sum of rang A	sum of rang B	U	Z	p	Z with correction	p	N valid A	N valid B
energy	1367.00	524.00	28.00	6.297056	0.000000	6.298888	0.000000	30	31
period	month: June 2014, July 2014								
variable	sum of rang A	sum of rang B	U	Z	p	Z with correction	p	N valid A	N valid B
energy	950.00	941.00	454.00	-0.151476	0.879601	-0.151494	0.879586	31	30
period	month: June 2014, August 2014								
variable	sum of rang A	sum of rang B	U	Z	p	Z with correction	p	N valid A	N valid B
energy	1008.50	882.50	386.50	1.125247	0.260485	1.125322	0.260454	30	31
period	month: December 2014, January 2015								
variable	sum of rang A	sum of rang B	U	Z	p	Z with correction	p	N valid A	N valid B
energy	860.50	1092.50	364.50	-1.626080	0.103935	-1.654200	0.098087	31	31

The determined value of test probability p indicates that while comparing the summer and winter months (June – January, July – December), the assumed significance level α has bigger values. Due to that there are grounds to reject the zero hypothesis and to assume the alternative hypothesis according to which there are significant differences in the monthly distributions of electrical energy production from photovoltaic conversion for both periods. This can also be confirmed in the results of monthly distribution of electrical energy production for the installed fixed and tracking photovoltaic modules for the indicated time periods. Figure 1 and Figure 2 present the distribution of electrical energy obtained from photovoltaic modules installed in a fixed and 2 – axis tracking configuration.

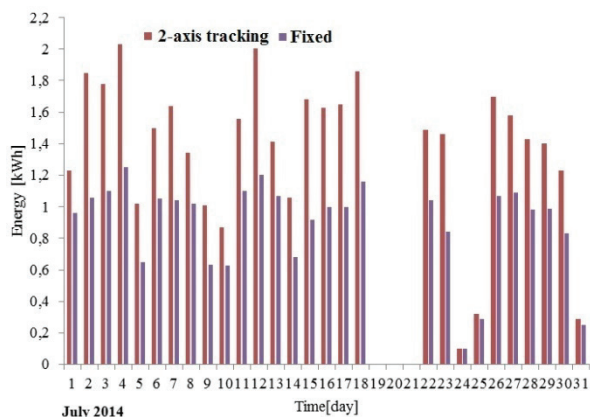


Fig.1. Monthly production of electrical energy by the fixed and tracking photovoltaic modules, July 2014.

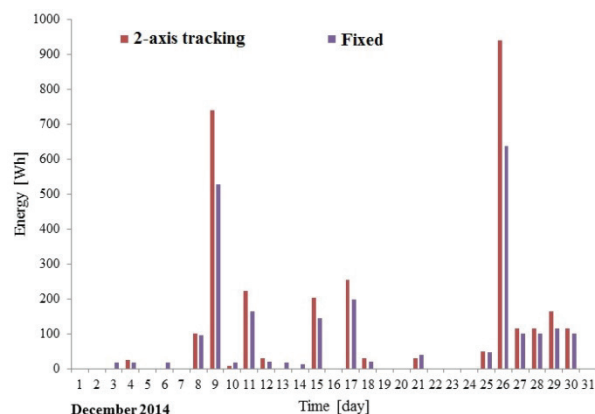


Fig.2. Monthly production of electrical energy by the fixed and tracking photovoltaic modules, December 2014.

In the case of summer months (June – July, June – August) no grounds to reject the zero hypothesis were identified. The monthly distributions of electrical energy do not significantly differ between one another. Also, no statistically significant differences were demonstrated for the two winter months: December 2014 and January 2015.

In order to analyse the repeatability of the daily peak hours in which the highest value of hourly electrical energy production is observed, respective 60-minutes periods of the day were indicated for each day of the analysed month: May, June, July and August 2014. Comparison of four independent groups was conducted using the non-parametric Kruskal-Wallis test which is the extension of the Mann-Whitney test. Both tests do not require to meet many assumptions characteristic for parametric tests. The statistical significance of the Kruskal-Wallis test indicates

* Corresponding author: artur.bugala@put.poznan.pl

differences between the tested groups. The starting point was formulation of test hypotheses: zero H_0 and alternative H_1 :

- H_0 : distribution of average ranks in the analysed groups (months of the year) does not differ significantly;
- H_1 : there are significant differences between the groups in relation to the repeatability of peak hours.

The value of parameter $p = 0.3279$ determined as a part of *Statistica* software is higher than the assumed significance level $\alpha = 0,05$. Hence, there are no grounds to reject the zero hypothesis assuming lack of significant differences in the periods of peak hours for the analysed months of the year. Therefore the occurring differences are not statistically significant. The additionally conducted Mann-Whitney U test for two groups confirmed the occurrence of significant differences between the peak hours (average ranks of observation) for the summer month (June 2014) and winter month (January 2015) with the assumed significance level $\alpha = 0.05$. The results of the tests are presented in Table 3.

Table 3. Results of the Mann-Whitney U tests confirming the significance of differences between the peak hours.

period	month: June 2014, January 2015			
sum of rang A	sum of rang F	p	N valid A	N valid F
354.00	207.00	0.030572	24	9

4 Conclusions

The used non-parametric methods are particularly relevant for small populations. Both the Mann-Whitney U test and non-parametric Kruskal-Wallis test demonstrated that there are significant differences between the values of the registered daily electrical energy and instantaneous power from the photovoltaic conversion for summer and winter, summer and autumn and summer and spring months. However, no statistically significant differences between the variable measured for months of the same season of the year were identified with the assumed significance level.

The analysis of variation and repeatability of monthly electrical energy production for almost the whole measurement year was also completed to include determination of, among other things, monthly average value, standard deviation, median, range and coefficient of variation.

Both the quantitative and the qualitative comparative analysis of the monthly electrical energy production from photovoltaic conversion indicates lower variation of courses for the summer months. Basic parameter in the presented analysis is the coefficient of variation in the meaning of ratio of standard deviation to arithmetic mean. The coefficient of variation value changes within the range from 0.36 – 0.51 for the months from June to August 2014. This is almost fivefold lower value compared to the results obtained for December the same year and over threefold lower compared to January the following

year. The value of coefficient of variation exceeding 0.60 indicates that a significant percentage of the average value is its standard deviation. Hence, the distribution of the measured value is not homogeneous and consequently the arithmetic mean should not be the main statistic measure. Using only the value of the standard deviation in the comparative analysis may turn out to be insufficient due to considerable differences in the average value of the monthly production of electrical energy for the extreme months of the measurement year.

The determined median value with considerably higher values for the summer months indicates that half of all the observations is found respectively below or over its value in the ordered series. The range has comparable values for the months from the same period (season of the year). However, this is not the measure resistant to the outliers which can disturb the variation of the measured value, at the same time not resulting in the change of the range's value.

The conducted statistical analysis may be useful for evaluation of energy potential of central Poland represented by the city of Poznan, with regard to planning new investments in the sector of renewable energy resources, mainly using technologies based on conversion of solar radiation energy to electrical energy or heat.

References

1. K. Abdelhamid, R. Ibtouen, *Solar Energy*, **103**, 171-182 (2014)
2. M. Bortolini, M. Gamberi, A. Graziani, *Energy Conv. Manag.* **86**, 81-92 (2014)
3. K. Branker, M. J.M. Pathak, J.M. Pearce, *Renew. Sust. Energ. Rev.* **15**, 4470-448 (2011)
4. A. Bugała, G. Frydrychowicz-Jastrzębska, D. Janczak, J. Dach, P. Boniecki, *International Multidisciplinary Scientific Geoconference (SGEM)*, 337-344, (2016)
5. A. Bugała, G. Frydrychowicz-Jastrzębska, D. Janczak, W. Czekala, J. Dach, M. Zaborowicz, *International Conference on Energy and Environment*, (ICEE, 721-727, 2015)
6. D. Burzyński, L. Kasprzyk, *E3S Web of Conferences*, **14**, 01041 (2017)
7. N. Faghil, J. Zarinchang, *Ekspluat. Niezawodn.*, **14**, 234-248 (2012)
8. D. Głuchy, D. Kurz, G. Trzmiel, *Prz. Elektrotechniczny*, **90/4**, 78-80 (2014)
9. D. Janczak, J. Dach, A. Bugała, P. Sołowiej, M. Neugebauer, *International Multidisciplinary Scientific Geoconference (SGEM)*, 243-250, (2016)
10. A. Kaabeche, M. Belhamel, R. Ibtouen, *Solar Energy*, **85**, 2407-2420 (2011)
11. M. Kolhe, *IEEE Trans. Energy Convers.* **24**, 511-519 (2009)

12. M. Maleki, F. Pourfayaz, *Solar Energy*, **115**, 471-483 (2015)
13. M. B. Shadmand, M. Pasupuleti, R. Balog, *IEEE Energy Conversion Congress and Exposition*, 3626-3632, (2011)
14. K. Skowronek, G. Trzmiel, *Prz. Elektrotechniczny*, **83/11**, 108-110, (2007)
15. J. Szymenderski; D. Typanska, *Computational Problems of Electrical Engineering* (IEEE Xplore, 2016)
16. K. Bednarek, R. Nawrowski, A. Tomczewski, *Electromagnetic Fields in Electrical Engineering, Studies in Applied Electromagnetics and Mechanics*, vol. **22**, IOS Press, Amsterdam, Berlin, Oxford, Tokyo, Washington, 363-368 (2002)
17. K. Bednarek, R. Nawrowski, A. Tomczewski, *Prz. Elektrotechniczny*, **84/1**, 62-64 (2008)
18. K. Budnik, W. Machczynski, J. Szymenderski, *Archives of Electrical Engineering*, **64**(2), 227 (2015)
19. K. Ciosk, *Prz. Elektrotechniczny*, **88/12B**, 47-49 (2012)
20. A. Dobrzycki, S. Mikulski, W. Opydo, *International Conference on High Voltage Engineering and Application* (ICHVE, 1-4, 2014)
21. G. Dombek, Z. Nadolny, *Ekspluat. Niezawodn.*, **18**, 413-417 (2016)
22. L. Kasprzyk, A. Tomczewski, K. Bednarek, *Prz. Elektrotechniczny*, **87/12b**, 82-85 (2011)
23. M. Lebioda, J. Rymaszewski, E. Korzeniewska, *International Conference on Microtechnology and Thermal Problems in Electronics* (MicroTherm, 2014)
24. L. Nowak, J. Mikołajewicz, W. Pietrowski, *Prz. Elektrotechniczny*, **85/6**, 63-66 (2009)
25. W. Opydo, A. Dobrzycki, *Electrical Engineering*, **94/1**, 37-48 (2012)
26. R. Pawlak, E. Korzeniewska, Z. Stempień, *Selected Issues of Electrical Engineering and Electronics* (WZEE, 2016)
27. Ł. Putz, M. Kurzawa, *Computational Problems of Electrical Engineering* (IEEE Xplore, 2016)
28. K. Bednarek, *Prz. Elektrotechniczny*, **84/12**, 138-141 (2008)
29. K. Bednarek, J. Jajczyk, *Prz. Elektrotechniczny*, **85/12**, 29-32 (2009)
30. J. Jajczyk, L. Kasprzyk, *International Conference on Heat Transfer, Thermal Engineering and Environment* (WSEAS, 226-231, 2008)
31. J. Jajczyk, *Computational Problems of Electrical Engineering* (IEEE Xplore, 1-3, 2016)
32. L., Kasprzyk, *Prz. Elektrotechniczny*, **88/7B**, 131-133 (2012)
33. Ł. Knypiński, L. Nowak, C. Jędrzycka, *COMPEL*, **34/3**, 882-892(2015)
34. Ł. Knypiński, L. Nowak, A. Demenko, *COMPEL*, **34/2**, 448-455 (2015)
35. A. Lewicki, A. Bugała, H. Piekarska-Boniecka, A. Lewicka, P. Boniecki, M. Cieślak, *International Conference on Chemical and Biological Sciences* (ICCBS 2016)
36. R. Nawrowski, K. Bednarek, A. Tomczewski, *International Conference on Parallel Computing in Electrical Engineering* (PARELEC, 218-222, 2000)
37. A. Bugała, G. Frydrychowicz-Jastrzębska, Z. Zbytek, J. Dach, D. Janczak, *International Conference on Frontiers of Sensors Technologies* (ICFST, 2016), *MATEC Web of Conferences*, **59** (2016)
38. G. Frydrychowicz-Jastrzębska, A. Bugała, *Energies* **8/2**, 1025-1041 (2015)
39. D. Głuchy, D. Kurz, G. Trzmiel, *Prz. Elektrotechniczny*, **89/6**, 281-283 (2013)
40. G.S. Maddala, *Ekonometria*, (PWN, Warszawa 2014)
41. M. Sobczyk, *Statystyka* (PWN, Warszawa, 2016)
42. A. Zeliaś, *Metody statystyczne*, (Polskie Wydawnictwo Ekonomiczne, Warszawa, 2000)
43. <http://www.statsoft.pl> [access 20.01.2017]