DE GRUYTER OPEN

Journal of ELECTRICAL ENGINEERING, VOL 68 (2017), NO3, 212–215

COMMUNICATIONS

Benchmarking the reliability of medium-voltage lines

Michal Kolcun, [*] **Mirosław Kornatka,** [**] **Anna Gawlak,** [**] **Zsolt Čonka** [*]

This paper discusses the problem of reliability of the Polish medium voltage power lines. It presents the benchmarking of power lines and values of SAIDI. The probability density distribution of the medium voltage lines length as well as selected indicators of the MV power lines reliability, being operated by the distribution companies under test, are also introduced. The analysis is based on the actual MV lines and uses the nonparametric method of kernel density estimation.

K e y w o r d s: radial distribution network, reliability of MV line, kernel density estimation, SAIDI

1 Introduction

An electric distribution network must continually supply energy to the consumers connected to the grid. Reliability is the key notion in the estimation of the supply continuity and service quality.

Evaluation of the reliability of MV networks is one of the main points considered in the analysis of the distribution power system. In the case of failure, both the Distribution Network Operators (DNOs) and its customers can suffer severe economic losses.

The DNOs are required to make public the following reliability indices: SAIDI, SAIFI (for long-term planned and unplanned interruptions) and MAIFI. These three are the main indices in the distribution network of various voltage levels [4] and are defined as follows:

SAIDI – System Average Interruption Duration Index gives the average amount of time per year while the energy supply to a customer is interrupted. It is expressed in minutes per customers per year and calculated by means of the following expression [4]

$$\mathrm{SAIDI} = \frac{\sum r_i\, N_i}{N_T} \tag{1}$$

where the summation is taken over all incidents, either at all voltage levels or only at selected voltage levels; r_i is the restoration time for each incident; N_i is the number of customers interrupted by each incident; N_T is the total number of customers in the system for which the index is calculated.

SAIFI – System Average Interruption Frequency Index corresponds to the average number of times per year while the supply of a customer is interrupted. It is expressed in interruptions per customer per year and calculated by means of the following expression [4]

$$\mathrm{SAIFI} = \frac{\sum N_i}{N_T}\,. \tag{2}$$

The expression for calculating MAIFI is the same as the one for calculating SAIFI (2), where the summation is taken over all incidents resulting in short interruptions.

The paper provides an assessment of the Polish MV lines reliability, based on the available statistical data from 2014 for eight distribution companies. The probability distribution function of the length of MV line is determined. The probability distribution functions of failure rates and of duration of power outages are also determined. The fundamental value of SAIDI is presented for the tested MV voltage lines. Calculations are performed in R (the language and environment for statistical computing) using the feature nonparametric methods for data analysis.

2 Methods of analysis of reliability data

Very often few reliability data are available, which are additionally averaged by the DSOs. Therefore, we should take this into account in the analysis. We can utilize nonparametric kernel density estimation (KDE) techniques for building these statistical representations of the averaged data [7, 8].

These techniques estimate the probability distribution function (PDF) directly from the data without any assumptions about the underlying distributions. This is in contrast with most parametric methods in elementary statistics which assume the data is quantitative, for example, that the population has a normal distribution and the sample size is sufficiently large. KDE is a nonparametric technique for density estimation, *ie* the estimation of PDF, which is one of the fundamental issues in statistics. It can be viewed as histogram density estimation with improved statistical properties. In a histogram, the choice of the number and width of bins can heavily influence the appearance.

Technical University of Košice Department of Electric Power Engineering, Košice, Slovakia, Technical University of Czstochowa, Institute of Electrical Power Engineering, Czstochowa, Poland

DOI: 10.1515/jee-2017–0031, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X

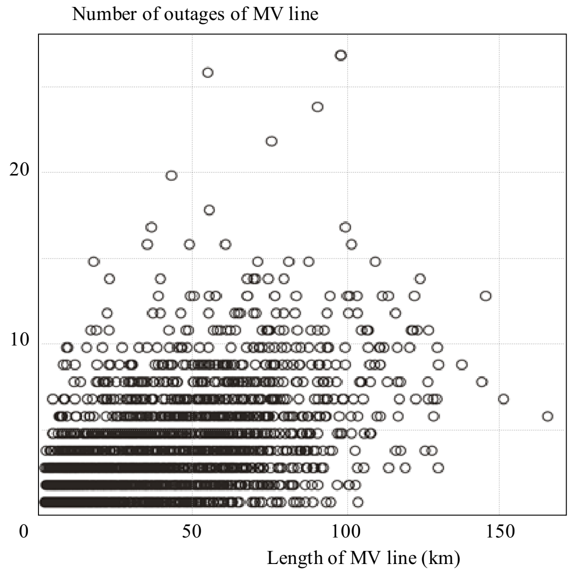

Fig. 1. The scatter plot of the number of recorded failures as a function of the length of the MV lines

The KDE is a useful method for the density of data samples. It can help an analyst determine how to model the data and answer the question whether the KDE looks like a normal curve or there are outliers in the data. It is a finite mixture distribution, being a sum of small density kernels that are centered at each data point. The shapes of the kernel are determined by the choice of a kernel function. The width of the kernel is determined by the bandwidth.

For an m-element set of available empirical data $(x_1 - x_m)$ the univariate kernel density estimator may be defined as [7]

$$\widehat{f}(x) = \frac{1}{mh} \sum_{i=1}^{m} K\left(\frac{x - x_i}{h}\right) \tag{3}$$

where $K\left(\frac{x-x_i}{h}\right)$ is the kernel of the estimator, in this case standard normal density rescaled by the bandwidth.

Kernel is a non-negative function that integrates to one and has a mean equal to zero.

The h (bandwidth) is the smoothing parameter. The bandwidth is chosen in such a way that it decreases as the size of the sample increases and the value of mean integrated squared error (MISE) decreases down to zero. When the value of h is too small, it causes too many local extreme values of the estimator $\widehat{f}(x)$. When its value is too large, the estimator becomes excessively smoothed, which does not reflect the properties of a real empirical sample. In practice there exist a number of useful procedures for determining the value of the bandwidth. A complete analysis of the issue can be found in [7]. In this paper the bandwidth is calculated by means of the standard R method [9].

Because the estimated value reliability parameters, *eg* the failure rate, have only positive values, it is necessary to modify the form of kernel estimator so that the random variable could take only positive values. Whether we use the limited Epanechnikov kernel or unlimited normal part of kernel located at the edge of the range, it may extend beyond the permitted area. The essence of limiting media kernel consists in a symmetrical reflection of that part of each kernel, which lies outside the permitted range. The KDE takes the form [7]

$$\widehat{f}(x) = \frac{1}{mh} \sum_{i=1}^{m} \chi_{[0,\infty]} \left[K\left(\frac{x - x_i}{h}\right) + K\left(\frac{x + x_i}{h}\right)\right] \tag{4}$$

where $\chi_{[0,\infty]} = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}$

In KDE, the choice of bandwidth can heavily influence its appearance.

The Tukey box plot, which is another method for presenting statistical information, is free from this disadvantage. The box plot is a useful nonparametric, quick way of examining of data graphically. Box plots may seem more primitive than a histogram or kernel density estimate but they do have some advantages [9]. Box plots display variation in samples of a statistical population without making any assumptions on the underlying statistical distribution. The spacings between the different parts of the box indicate the degree of dispersion and skewness in the data as well show outliers. Box plots allow to represent various estimators visually, notably the interquartile range (IQR).

3 Experiments

The paper provides an assessment of the Polish MV lines reliability, based on the available statistical data from 2014 for the eight distribution companies. These distribution companies supply more than 5 million energy customers and the total length of MV lines operated by them is more than 100 thousand km. The share of cable lines in the area under study is low at about 16 %. The total number of failures occurred in this period in cable and overhead MV lines were 9 184.

In order to analyze the condition of the MV line all failures were accurately recorded. For each disconnection the MV line recorded its duration and the number of customers without power supply. The number of MV lines considered in the study was 2 893 and their total length was 90 315 km. Statistical data available can determine the sequences of the highest failure rates, *ie* locate the lines that need to improve their reliability.

Figure 1 shows the dispersion of the number of failures occurring in the individual MV lines as a function of their length.

No significant correlation was found between the line length and the number of failures. It should be noted that the longest lines don't have the largest number of faults. The greatest number of failures recorded in the MV overhead lines occurred in lines less than 100 km

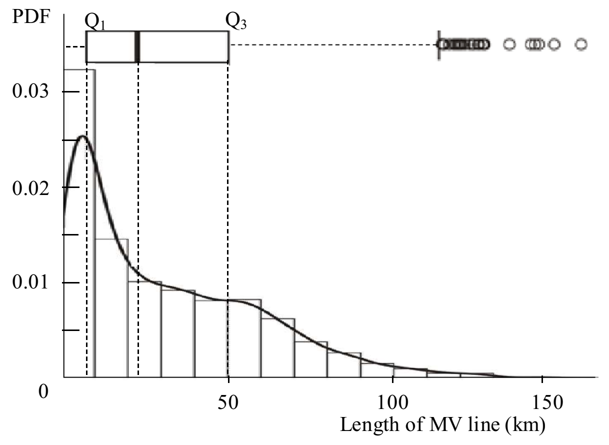

Fig. 2. Histogram, box plot and PDF (kernel — Gaussian, bandwidth — 5 145) of the cable and overhead MV lines

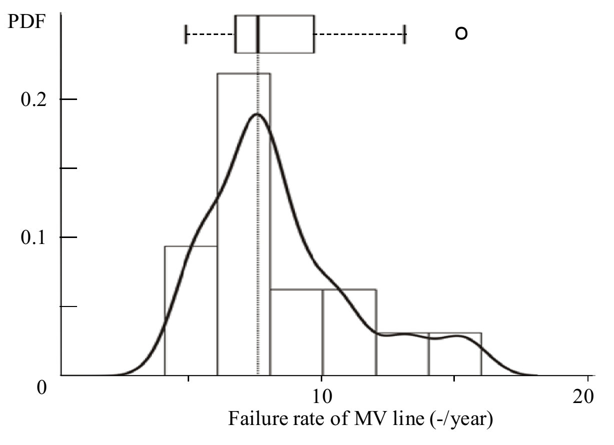

Fig. 3. Histogram, box plot and approximation of pdf (kernel - Gaussian, bandwidth - 0 5188) of the interruption duration of MV line

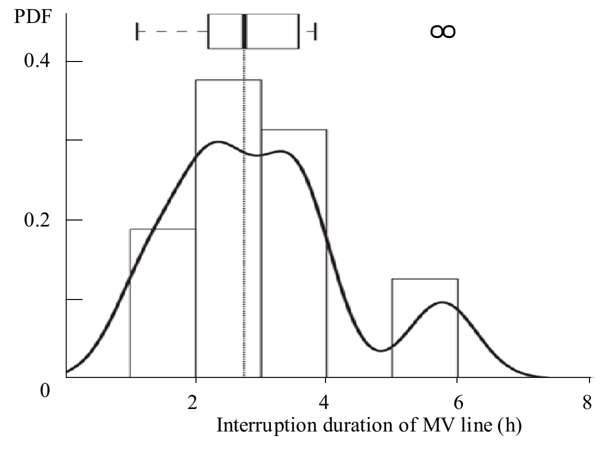

Fig. 4. Histogram, box plot and pdf (kernel – Gaussian, bandwidth – 0.9306) of failure rate of MV line

long. The most extreme case was that of a 96 km long line, which was off 27 times.

Analyzing the number of failures only does not reflect the picture of the situation. There is a need of a more in-depth analysis of the problem of the MV network failure. For this purpose the PDF, histogram and box plot for the length of the MV line were determined and failure rate as well as the interruption duration in the supply of electricity due to damage were obtained. Graphical representation of the same data on different types of charts helps extract the useful information and allows one to observe the facts which cannot be seen on the basis of the calculated basic statistical measures (mean value and standard deviation).

Figure 2 shows the statistical analysis of the distribution of the length of the cable and overhead MV lines.

With the histogram and KDE it is possible to determine that the modal value of the distribution equals 7 km. The distribution curve of the line length has a very long tail. The diagram at upper part (box plot) shows that the median value of the distribution is about 22.66 km. We can also read the 1st quartile (Q1) – 7.09 km and 3rd quartile (Q3) – 50.13 km. On the box plot, MV lines with a length of more than 114 km are marked as outliers of the distribution (beyond the value of one and a half of the span interquartile (IQR), therefore, these lines are marked by circles).

Figure 3 presents the distribution of the failure rate for individual distribution companies. The failure rate is calculated as the average value for the entire distribution company in a given year. Determining the PDF of the failure rate provides a tool for the assessment of the network condition and for benchmarking distribution companies.

The minimum value of the failure rate was 4.83 failures per 100 km of MV during the year. One of the divisions had a failure rate – 15.3 failures/year/ 100 km – which is an outlier distribution. The median and modal values of distribution are similar – 7.5 failures/year/ 100 km.

According to (1) System Average Interruption Duration Index determines the average duration of interruptions supply per customer. The distribution of the interruption duration of MV lines in individual distribution companies is presented in Fig. 4.

Although the data presented come from areas with similar climatic conditions and no natural disaster occurred in 2014, particular divisions operate the MV network of different character. Two distribution companies operate the MV network with its large share located in rural areas. On each type of graphs this can be seen as outlier values. The median distribution is 2.7 h and the value of IQR is 1.3 h.

The main issue of the study was to the answer the question which lines are responsible for the excessive value of the index SAIDI? For this purpose, the scattered values of SAIDI index of the single MV line were determined as a function of their length, as shown in Fig. 5.

The unreliability of the MV line is caused by a small portion of the line. 50 % of the SAIDI value is contributed by failures in 383 lines, the length of which constitutes only 24 % of the total line length (the grey rectangle in the scatter plot).

The box plot in a right side shows a large number of outliers of SAIDI values (outside the scope of one and

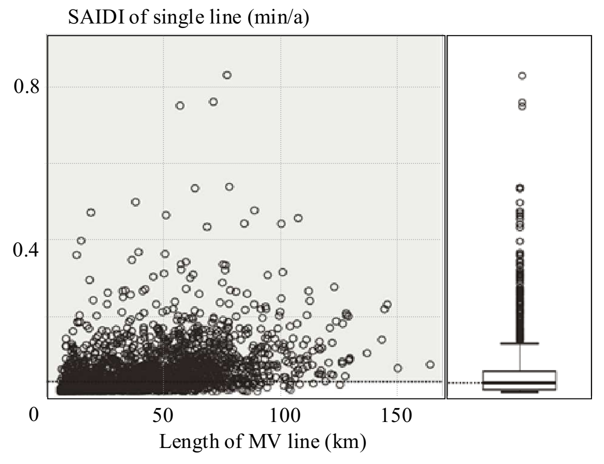

Fig. 5. Values of the SAIDI of single MV line *vs* the length of each line

a half the IQR span). As it can be seen in Fig. 5, up to 75 % of the MV lines has less than 0.05430 values of SAIDI min/a (3rd quartile in box plot). The median value distribution of SAIDI index was 0.01950 min/a.

4 Conclusion

Nonparametric methods are suitable for the evaluation of the tested network failure and for locating excessive failure. The methods presented in the paper of analysing the reliability data allows for a more complete and at the same time easier data mining and graphical representation of the measure of the statistical distributions of a power network.

References

[1] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems* 2nd ed., New York, Plenum Press, (1996).

[2] 6th and CEER, "Benchmarking Report on the Quality of Electricity and Gas Supply", *Ref: C16-EQS-72-03*, CEER, Brussels, published 2016, (also previous editions).

[3] M. Kolcun and P. Frak, "Grid reliability analysis as a tool for effective asset management", *5th International Scientific Symposium ELEKTROENERGETIKA 2009 Stará Lesná*, (2009) pp. 60–62.

[4] "Minister of Economy Order the Matter of Detailed Operation of the Power System", 4th May 2007, $ 41.

[5] "National Report 2016 of the President of Energy Regulators Office", www.ure.gov.pl.

[6] *Manual Movement and Operation of the Distribution Network, TAURON S. A. Kraków, (2015) - Polish.*

[7] B. W. Silverman, *Density Estimation for Statistics and Data Analysis* Chapman & Hall/CRC, London (1986).

[8] M. Kornatka, "The Weighted Kernel Density Estimation Methods for Analysing Reliability of Electricity Supply", *17th International Scientific Conference on Electric Power Engineering (EPE)*, Prague, (2016) pp. 1–4.doi: 10.1109/EPE.2016.7521729.

[9] R Core Team, *R: A Language and Environment for Statistical Computing*, Foundation for Statistical Computing, (2016).

Received 16 December 2016

Michal Kolcun was born in 1954 in Ruska Vola nad Popradom, Slovakia. In 1979 he graduated at the Faculty of Electric Power Engineering of the Moscow Power Engineering Institute. In 1989 he defended his PhD on the same institute in Moscow. In 1993 he habilitated to associated professor at the department of Electric Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. In 2000 he inaugurate to professor, his thesis title was High-tension electrical power engineering. Since 2006 he is honorary professor at Budapest Polytechnics. Since 1979 he is working with the Department of Electric Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. His scientific research is focusing on a power system control and computer application in electric power engineering.

Mirosław Kornatka (PhD) was born in 1964. He graduated at the Faculty of Electrical Engineering at the Czstochowa University of Technology where he also received his Master degree in 1989.He completed his doctor thesis in 2000 at the Faculty of Electrical Engineering at the Wrocaw University of Technology. He is currently adiunkt at the Faculty of Electrical Engineering.

Anna Gawlak is Assistant Professor at the Faculty of Electrical Engineering Technical University of Czstochowa. She received her MSc in 1985 and her PhD in 1994 in Electrotechnics from Technical University of Czstochowa. In 2011 she was awarded her the degree of Habilitation at the Technical University in Ostrava. The main research interests are: the efficiency of power distribution, power losses and energy distribution networks, and network optimization.

Zsolt Čonka is Assistant Professor at the Faculty of Electrical Engineering and Informatics at the Technical University in Košice. He was born in 1988 in Trebišov Slovakia. He graduated at the Faculty of Electrical Engineering and Informatics at Technical University in Košice. In 2016 he defended his PhD on the Technical University of Košice. His scientific research is focusing on a FACTS devices, power system stability, and protection systems.