RESEARCH & DEVELOPMENT IN POWER ENGINEERING, 2019

Estimation of maximum loads of residential electricity users

Sławomir Bielecki.[1]

[1] Warsaw University of Technology, Faculty of Power and Aeronautical Engineering – Institute of Heat Engineering, Nowowiejska Str. 21/25 00-665 Warsaw, Poland, e-mail: slawomir.bielecki@pw.edu.pl

Abstract. The subject matter of coefficient factors is not a frequent object of analysis in the scientific literature. However, because of the use of this term in engineering tasks and because of the constantly changing manner of energy use (resulting from the dissemination of new appliances), there is a need for current analysis and possible revision of the approach based on coefficient factors. The aim of the paper is to compare currently recommended values of this coefficient in the task of estimating the peak output of installations on the example of standards applied in Poland. On the basis of smart metering data and statistical methods, coefficient factors for residential consumers supplied from the municipal distribution network were determined. The results were compared with the currently applied values in accordance with the industry standard issued by the Association of Polish Electrical Engineers (standard N SEP-E-002).

1 Introduction

Due to the non-monotonous and usually characterised by more than one extreme of the individual users (consumers load), the peak (maximum) of load profile of the feeder does not have to occur at the time of peak at any of supplied users. There is a non-linear or even entangled superposition of individual peaks. A simplified model of these dependencies is the coincidence factor which takes different values for different numbers and types of loads forming the analysed structure.

The basic definition of *the coincidence factor* is the ratio of the peak power of the feeder to the sum of the peak power of the loads:

$$k_j = \frac{P_{j-group\,peak}}{\sum_{i=1}^{n} P_{i-peak}} \qquad (1)$$

where: $P_{j\text{-group peak}}$ – peak power in the feeder which the *n* loads; $P_{i\text{-peak}}$ – peak load of the *i*-th load.

It takes values from the range:

$$k_j \in \left[\frac{1}{n}; 1\right] \qquad (2)$$

where *n* is the number of loads supplied from the feeder.

The concept of the *kj* factor is based on the finding that the behavioural patterns of different types of consumers overlap. It is unlikely that all consumers will be used at full capacity at the same time and, therefore, smaller load capacity sections of installations or networks can be designed than the simple sum of the expected peak loads. The value of the *kj* factor depends on the number of loads (consumers) and decreases with the increase in this number.

In literature there are reversals of coincidence factor, called a *diversity factor*. The simplicity of the idea of the coincidence factor as a parameter, characterising the load of the network or installation, forces significant limitations of its accuracy. In particular, the definition does not specify the relationship between the moments of the highest loads and does not directly specify the conditions, ranges and certainty of its occurrence (values may change different time horizons – replacement of equipment, change of habits and use of energy among tenants).

The *kj* coefficient binds the peaks of loads of single loads supplied from a feeder. If a too low value is adopted during designing, it may lead to a reduction of capacity reserve or even overload during peak periods, while a too high value in turn leads to unnecessary oversizing of the installation and the increase in investment costs.

2 Literature review

The idea and method of the coincidence factor has been known since the beginning of the 20th century [1], [2], [3]. Its values are determined primarily for household consumers [4], [5].

Assuming that the power demand in the adopted time intervals of consumers has a normal distribution and using the statistical approach, the *kj* coefficients can be determined on the basis of formulas that directly depend their value on the number of consumers and parameters describing the distribution. In fact, the use of active power and energy by users does not have to create normal distributions [6]. This means that formulas based on this assumption (e.g. Rusck, Velander, Strand-Axelsson) are currently being questioned [7].

Assumptions simplifying and unifying loads adopted to determine the coefficients characterising the use of electricity such as *kj* were needed in the conditions of limited availability of measurement data. With the development of AMI technology, the possibilities of using measurement data in load modelling are increasing. In [8], the calculated values of diversity factors with their

dispersion (presenting the average and maximum values and the percentiles: 99th, 95th, 5th and 1st) for different numbers of consumers on the basis of the measurement data of 1898 consumers were presented.

In [9] the concept of a perceptron-type multilayer neural network for the estimation of coincidence factors values was presented. The network requires information about 10 issues influencing the factor value, concerning the load (monthly peak demand, difference between peak and valley demand), environmental (weather) conditions, users and urbanization parameters. The network requires a data-learning process.

In the paper [10] it was pointed out that engineering estimation of maximum power demand in low voltage networks, based on peak demand and corrected by coincidence factors of simultaneity, does not take into account the stochastic character of the phenomenon and does not contribute to obtaining a proper outlook of aspects related to energy quality. Therefore, the authors proposed a model of the Monte Carlo simulation of energy demand by the consumer, using data from the gamma distribution. The result of the simulation is corrected, among others, by temperature data.

In the paper [11], presenting the results of calculations of average values of diversity factors, obtained on the basis of the method of bootstrap for household consumers (in the number from 1 to 60), it was shown that these values for a selected number of consumers (with 1000 bootstrapping samples) do not form normal distributions. The authors suggest that a more appropriate distribution for these values is the gamma distribution. Moreover, comparing the obtained results with [5], the authors indicate that these factors for each assumed number of household appliances, obtained as a result of calculations according to the bootstrap method, are independent values. In the paper [12] the methods of calculation of losses based on fuzzy grouping of monthly energy consumption by consumers, without determining the peak demand, are presented.

The paper [13] attempts to introduce and determine coincidence factors for charging electric vehicles from the network (mutual coincidence factor), defined as the ratio of the peak active power in the system after connecting the vehicle to the network to the sum of the peak charging power of the vehicle and the peak power of other loads in the network. On the basis of data from the urban network area of Beijing and the Monte Carlo method, the ranges of these coefficients were estimated, depending on the charging power. A normal coincidence factor for charging electric vehicles, defined as the ratio of the peak charging power of a vehicle to the product of the number of vehicles charged and the charging power of a single vehicle, has also been proposed. These values, therefore, depend on the number of vehicles and the charging power. The concept of coincidence factors in the analysis of the impact of V2G services on the distribution network was also used in the paper [14].

In the Polish papers [15], [16] a proposal of determination of coefficients of simultaneity based on, among others, the estimated degree of filling of the load

diagram and correlations between the consumers was presented. The calculations of coincidence factor or peak power made in [15], [16] on the examples turned out to be even greater in comparison to the values given in the SEP (the Association of Polish Electrical Engineers) standard [17].

3 Methodology

Having at its disposal the measurement data of active energy consumed by 7671 single residential consumers during one year in 1-hour intervals, it was possible to estimate the peak power demand by consumer. The 1-hour values of the single consumer's active energy input in kWh are numerically equal to the values of the active power load in kW averaged over 1-hour intervals.

Based on a random selection it was possible to create sets simulating the system of consumers supplied from one installation and to determine the relation between the peak load of a hypothetical feeder supplying such a system of consumers and the sum of peaks of individual consumers. This relation is represented by the value of kj factor. The trend of kj value was searched for as a function of the number of residential consumers in the system.

Table 1. Results of statistical tests on the values of active power loads of consumers.

Test	Jarque-Bera	Kolmogorv-Smirnov	Lilliefors
Significance level	0,05	0,05	0,05
Maximum of probability-value	0,001	0	0,001
Minimum of difference between test statistic and critical value *)	136,19	0,486	0,0435
Percent of cases where the null hypothesis**) can be rejected	100%	100%	100%

*) In case of normally distributed values the difference should be negative.
**) The null hypothesis: measured values are normally distributed with unspecified mean and standard deviation.

Distributions of 1-hour loads with the active power of the consumers do not have to form normal distributions. Kurtosis as a dimensionless statistical coefficient for normal distributions assumes the value of 0. In the analysed data of the consumers' loads the lowest value of kurtosis was 1,089 (Fig. 1). Statistical tests of the measurement data provided reasons to reject the hypothesis of the normal distribution of values (Table 1).

Therefore, the application of formulas for calculating coincidence factor based on the assumption of equality of standard deviation of standard loads of consumers will not be appropriate. A better approach is to determine the values of these coefficients directly from their definition in a sufficiently large number of random samples of data sets and to determine statistical measures describing them.

The values of coincidence factor kj differ depending on the selected consumers making up the set. In order to determine the representative size of the factor sought, the bootstrapping approach was used.
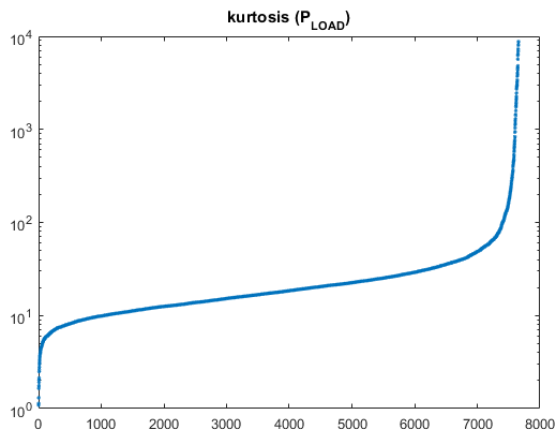
Fig. 1. Diagram of the sorted value of kurtosis of active power loads for each of the 7671 consumers.

The bootstrap is a simulation method of statistical inference developed in 1979 [18] and can be applied to both independent and dependent variables, also with unequal distribution. The method is predestined for situations when the distribution of statistics or a random variable describing a feature is unknown. Bootstrapping may be a better method of estimating the distribution of an estimated parameter than classical methods based on a central limit statement [19]. In a non-parametric approach, this method is free from model assumptions. The method allows to estimate the parameter characterising a given population on the basis of a randomly selected output sample. The approximation of statistics values on the basis of bootstrap samples is based on the Monte Carlo method. It consists in drawing independent samples of no more than the size of the original sample [20], calculating their statistics, where their average value is an estimation of the sought-after parameter describing the community. The method using bootstrapping to calculation the *kj* coefficients was applied in the paper [11].

The possessed measurement data of loads of consumers were considered as the primary sample. Bootstraps are sets of data on loads drawn from set of users. The key issue is to select a sufficient number of independent secondary samples. There are many rules for selecting the number of draws. A correct approach, although not very practical (a large number of sets), is to select the number of draws as equal to all possible variants of the bootstrap set elements of the assumed number [21]. Apart from complicated algorithms of optimal number of samples (e.g. [22]), often to determine the bootstrap distribution estimator, fixed values are taken as the element from the number of the output sample [19] or a simple recommendation that at least 1000 samples of the bootstrap type should be used [23].

In the task of determining the value of *kj* factors, the own algorithm of random selection of a set of consumers was developed. The built algorithm allowed for one-time use of data of each consumer in a number of sets. A single set concerned the selected number of consumers forming the system of loads supplied from one feeder. Therefore, the number of secondary samples was determined by the number of consumers *n*. The creation of sets was repeated 250 times, obtaining an integer number from $250 \times \left[\frac{L}{n}\right]$ samples for each number *n*, where *L=7671* is the number of all the measured consumers. Thus, numerous samples from 958 750 (at *n=2*) to 9 500 (at *n=200*) were obtained.

4 Results

Fig. 2 presents the statistical measurements of population of coincidence factors values calculated on the basis of the algorithm as a function of the number of residential consumers. For comparison, the values recommended by the SEP standard [17] were imposed in two extreme cases, namely for older modernised installations (*SEP1*) and installations with electric heating of water (SEP2), as well as the theoretically minimum value (*min*) determined according to formula (2). The average value and the median *kj* were approximated by the power function. The graph of standard deviations in relation to the average of *kj* value depending on the number of flats presents Fig. 3.

Fig. 4. presents a comparison of peak power estimates in relation to the number of consumers (flats) obtained in three ways, namely on the basis of:
- *kj* coefficients according to the SEP standard,
- *kj* coefficients calculated in this paper,
- direct determination of peak power on the basis of analysed measurement data.

Peak capacities according to the SEP standard (*SEP1* and *SEP2*) are the products of the number of flats and given in this standard: relevant *kj* coefficients and the power demanded by one flat (7 kVA for modernised installations or 30 kVA for flats with electric heating of usable water). The mean value of the apparent peak power of 7671 consumers analysed in this paper was 3.4 kVA and 2.3 kVA median and 99th percentile was 5.2 kVA.

For precautionary purposes, the 99th percentile (5.2 kVA) was used in the peak power calculations on the basis of the previously determined *kj* coefficients. The estimated peak capacity is, therefore, the product of the number of consumers, the 5.2 kVA value and the relevant *kj* value calculated at the previous stage of this analysis. Two variants of estimations were presented – on the basis of appropriate average values of *kj* (*S calc kj mean*) and 95th percentile (*S calc kj 95-perc*).

In the second stage of the bootstrapping analyses, peak loads in absolute units for one feeder with residential consumers were estimated directly. The algorithm looked similar to the previous one, but this time it did not calculate the value of *kj*. The graph (Fig. 4). shows the average values of the samples obtained from the peak values depending on the number of consumers (*Pmean*) and the 95th percentile of peak power of the feeder (*P 95-perc*). According to the SEP standard, it is assumed that in residential buildings the power factor cosφ≈1, so these two values in kW would be numerically equal to the values in kVA. The average value *(Pmean)* was converted into apparent power (*Pmean/cos(phi)*) on the basis of the determined median power factor of the analysed consumers cosφ, which in this case was 0.3.
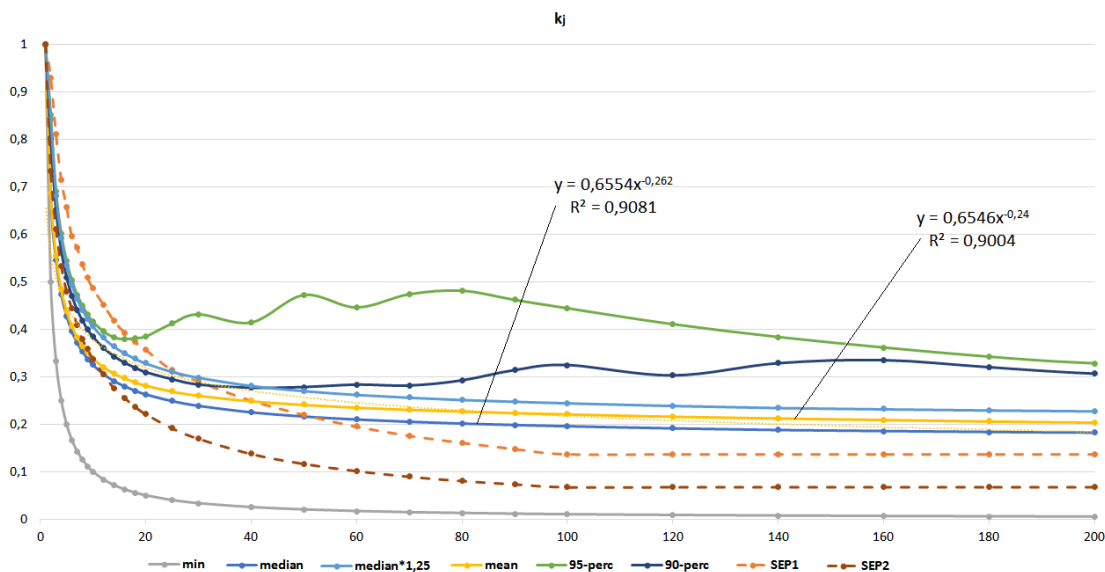
Fig. 2. The coincidence factor values obtained from calculations based on sets of randomly selected consumers as a function of their number in the installation – mean, median, 95th and 90th percentile, median increased by 25% (*median*1.25*), coefficients according to SEP standard (*SEP1, SEP2*) and theoretically minimal value of the coefficient (*min*) according to formula (2).
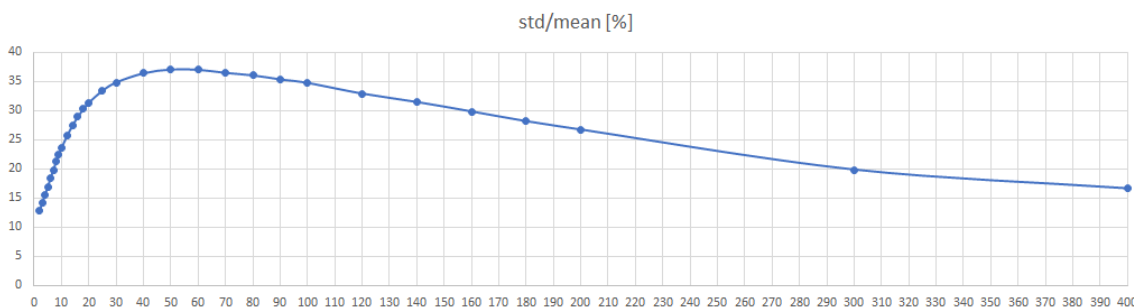

Fig. 3. Percentage of standard deviations to the mean value of the kj coefficients of the samples as a function of the number of flats supplied from one feeder.
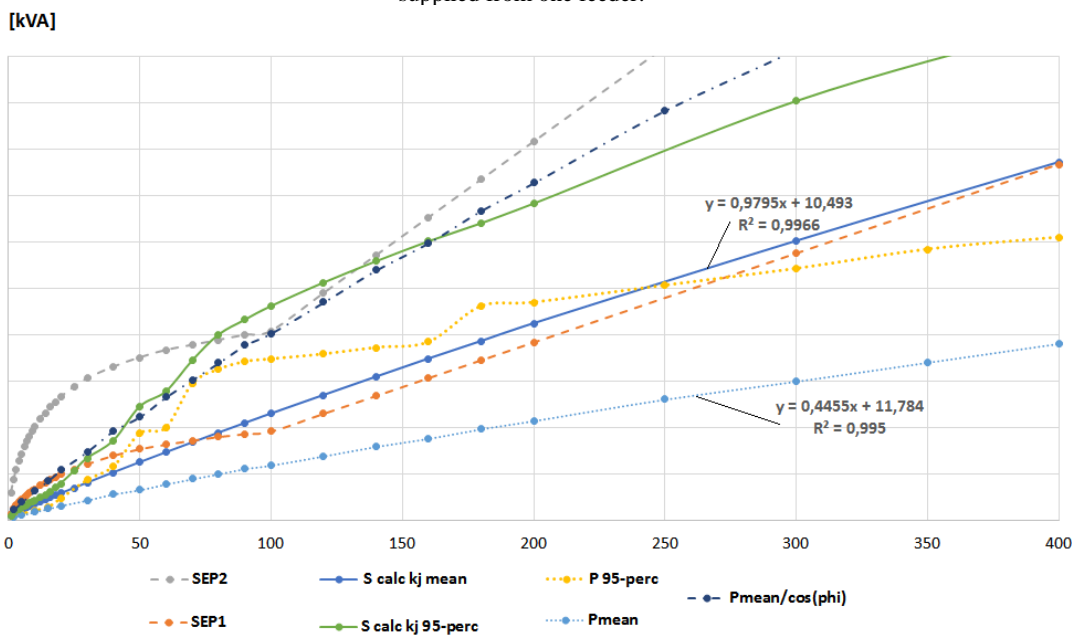

Fig. 4. Comparison of the determined peak capacities based on: SEP standards (*SEP1* and *SEP2*); determined coincidence factors kj - mean (*S calc kj mean*) and 95th percentile (*S calc kj 95-perc*); directly determined peak capacities in absolute values - mean (*P mean*), 95th percentile (*P 95-perc*) and mean taking into account the average value of tangent power of consumers. Values are presented as a function of the number of flats.

5 Discussion

The values of statistical measures, apart from the higher percentile, are monotonously decreasing in relation to the increasing number of flats supplied from the common feeder. The average value for more than 20 consumers is located in the middle between the median values: calculated and increased by 25%. The highest dynamics of changes in kj value is observed for small numbers of consumers (below 20) supplied from one feeder (Fig. 2). The highest values of standard deviations from the samples in relation to their mean value reach 40% (Fig. 3).

The determined kj coefficients in comparison to the SEP standard take smaller values for smaller numbers of consumers. For a number of flats below 10, the smallest of the kj values (according to the SEP standard with electric water heating) are between the average value and 90th percentile of the obtained kj values from random samples. The highest values of kj coefficients (for modernised installations) according to SEP are higher than 125% of obtained median of samples for the number of flats below 30. For larger numbers of flats the SEP coefficients turn out to be lower than those obtained as a result of analyses (Fig. 2). The values of median and average are similar, the biggest differences do not exceed 12% (maximum 0.025 in absolute values).

The median and mean values as a function of the number of flats are monotonously decreasing in the whole analysed range. To 14 flats the 95th percentile practically coincides with the median value increased by 25%. The 90th percentile does not behave monotonously as a function of the number of flats in the range between 40 and 160, while the 95th percentile in the range 20…90. In these ranges (number of flats 20...160) the highest standard deviation values are observed among the samples of kj in relation to their mean values (30-40%) – Fig. 3. For this number of flats, more samples with higher kj values were observed, despite the fact that the monotonous course of averages and median kj were maintained. In the data set there were series of peak load points of relatively high values, which were equally probable to be drawn at the mentioned numbers of users and increased the extreme kj values.

The characteristic feature is the occurrence of a linear relationship between the number of loads and the average value of directly determined peak power values of the feeder – Fig. 4. The linear relationship is confirmed by the matching of a straight trend line with a high determination factor R^2. The lowest peak power values were obtained for mean values under the assumption $\cos\varphi \approx 1$. These values are significantly lower than according to the SEP standard.

A separate issue is the demand for reactive power of residential consumers. Nowadays, the assumption adopted in accordance with the SEP standard with a significant advantage of resistive devices in residential installations does not seem appropriate. An additional aspect complicating the issue is the controversy around the power theory [24]. Apart from such detailed issues, on the basis of the measurement data from the meters of 7671 residential users, the median power factor was

determined, which reached only $\cos\varphi=0.3$. Taking this value into account for the determined average value of peak power resulted in the fact that for a smaller number of users this value is similar to the estimation made on the basis of 95th percentile of calculated kj. For a large number of flats, this value is between the extreme variant according to the SEP standard and the values determined on the basis of 95th percentile kj. The assumption of such a low $\cos\varphi$ value for all users does not have to be justified. Reactive power is compensable. It is therefore possible to compensate each other for reactive power flows in the installation. On the one hand, it requires an additional flow analysis and, on the other hand, it enables the specific use and optimisation of this capacity within the framework of the extended energy management concept [25].

Peak apparent power values estimated on the basis of average kj values and assumed peak power of a single flat were also linear in relation to the number of residential users. For a smaller number of flats (up to 60), the values are lower than the least demanding variant of the SEP standard. The values of 95th percentiles of the directly estimated active peak power to approximately 250 flats vary between the values corresponding to the peak power estimated on the basis of the average of the determined kj and the directly estimated average peak power, corrected by the power factor value. Above 250 flats, 95th percentile of directly estimated peak power are quasi-linear, below the value based on the SEP standard.

6 Conclusions

The method of peak power estimation based on coincidence factors kj is simple to apply in engineering practice. The key step is to adopt the correct values for these factors. The distributions of the power used values do not have to form normal distributions, so it will not be appropriate to determine them on the basis of the assumption of similar values of standard deviations. Estimating kj directly from the definition requires access to large sets of measurement data, which in the era of smart metering should not be a problem. In this case, a simpler and more effective method may be to determine directly the values in absolute peak power units, without calculating the coincidence factor. Especially since the kj-method additionally requires estimation of the peak power of a single consumer.

The analyses carried out indicate that the peak power may be overestimated by the applicable standards giving the kj-values This overestimation results mainly from assuming a too high value of the peak power of a single user. The calculations of the peak power of the feeder conducted directly for the average values from random samples indicate a linear dependence on the number of residential consumers. The results obtained on the basis of kj averaged from random samples are characterised by a similar dependence. The results obtained on the basis of the kj-value from the standards do not have a general linear trend. The impact of reactive power demand complicates the issue due to the complexity and small recognition of reactive power using at the level of household consumers.

References

1. W. Lackie, The Influence of Load and Diversity Factors on Methods of Charging for Electrical Energy, *Journal of the Institution of Electrical Engineers*, vol. 42, pp. 100-114, (1909).
2. H. Gear, Diversity factor, *the Chicago Section of the American Institute of Electrical Engineers*, (1910).
3. C. Bary, Coincidence-Factor Relationships of Electric-Service-Load Characteristics, *AIEE Transactions,* vol.64, pp. 623-629, (1945).
4. A. Sargent, R. Broadwater, J. Thompson and J. Nazarko, Estimation of Diversity and kWh-to-Peak-kW Factors from Load Research Data, *IEEE Trans. on Power Systems,* vol.9, pp. 1450-1456, (1994).
5. J. Nazarko, R. Broadwater and N. Tawalbeh, Identification of Statistical Properties of Diversity and Conversion Factors from Load Research Data, *Electrotechnical Conference,* vol.1, (1998).
6. K. Billewicz, Tests normally distributed population of the values energy (in Polish), *Przegląd Elektrotechniczny*, no. 4, pp. 78-79, (2008).
7. J. Dickert and P. Schegner, Residential Load Models for Network Planning Purposes, *IEEE 2010 Modern Electric Power Systems*, Wrocław, (2010).
8. J. Peppanen and J. Taylor, Enhanced Load Modeling with Expanded System Monitoring, *IEEE Clemson University Power Systems Conference (PSC)*, (2018).
9. X. Liu, H. Zheng, S. Song and G. Fu, Selection Method of Community Load Coincidence Factor Based on BP Neural Network, *IEEE 3rd International Conference on Computer Science and Network Technology*, (2013).
10. D. H. McQueen, P. Hyland and S. Watson, Monte Carlo Simulation of Residential Electricity Demand for Forecasting Maximum Demand on Distribution Networks, *IEEE Trans. on Power Systems*, vol. 19, no. 3, pp. 1685-1689, (2004).
11. V. Chatlani, D. Tylavsky, D. Montgomery and M. Dyer, Statistical Properties of Diversity Factors for Probabilistic Loading of Distribution Transformers, *IEEE 39th North American Power Symposium,* (2007).
12. S. Kaykahie and S. Kowsari Movahed, A New Approach for Calculating Load and Loss Factor Base on a New Approach for Calculating Load and Loss Factor Base on Consumer Data with Fuzzy Modelling, *CIRED 22nd International Conference on Electricity Distribution*, (2013).
13. X. Tong, C. Guo, X. Yang and C. Chen, Research on Characteristics of Electric Vehicle Charging Load and Distribution Network Supportability, *IEEE PES Asia-Pacific Power and Energy Conference*, (2016).
14. U. C. Chukwu, O. A. Nworgu and D. O. Dike, Impact of V2G penetration on Distribution System Components Using Diversity Factor, *IEEE SOUTHEASTCON*, (2014).
15. S. Ciura and H. Kocot, Rzeczywiste i obliczeniowe obciążenia poszczególnych elementów sieci, zasilających odbiorców bytowych, przyłączonych do sieci niskiego napięcia., *Rynek energii elektrycznej. Aktualne problemy energetyki.*, P. Pijarski, Z. Połecki, Eds., Politechnika Lubelska, pp. 33-42, (2018).
16. S. Ciura and P. Kubek, Wyznaczanie obciążeń maksymalnych w sieciach nn – teoria i praktyka, *Rynek energii elektrycznej. Rozwój i funkcjonowanie rynków energii.*, Z. Połecki, P. Pijarski, Eds., Politechnika Lubelska, pp. 7-18, (2016).
17. N SEP-E-002. Instalacje elektryczne w obiektach budowlanych. Instalacje elektryczne w budynkach mieszkalnych., COSiW SEP, (2003).
18. B. Efron, Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics*, vol.7, pp.1-26, (1979).
19. K. Singh and M. Xie, Bootstrap: A Statistical Method, Rutgers University, (2008).
20. B. Efron, Estimating the Error Rate of a Prediction Rule: Improvement on Cross-validation, *Journal of the American Statistical Association*, vol.78, (1983).
21. J. Fox, Applied Regression Analysis and Generalized Linear Models, *SAGE Publications Inc.*, (2008).
22. R. Davidson and J. MacKinnon, Bootstrap tests: how many bootstraps?, *Econometric Reviews – Taylor&Francis Journals*, vol. 19(1), (2000).
23. B. Efron and R. Tibshirani, An Introdution to the Bootstrap, Chapman&Hall, (1993).
24. L. Czarnecki, Meta-theory of electric powers and present state of power theory of circuits with periodic voltages and currents, *Przegląd Elektrotechniczny* no.6, pp. 26-31, (2013).
25. S. Bielecki and T. Skoczkowski, An enhanced concept of Q-power management, *Energy* no.162, pp. 335-353, (2018).