

# Non-Negative K-SVD as an element of the forecasting electricity demand system

*Marcin Drechny*<sup>1,\*</sup>

<sup>1</sup>UTP University of Science and Technology, Faculty of Telecommunications, Computer Science and Electrical Engineering, Prof. S. Kaliskiego 7 Ave, 85-796 Bydgoszcz, Poland

**Abstract.** The article describes the NN-K-SVD method based on the use of sparse coding and the singular value decomposition to specific values. An example of using the method is the compression of load profiles. The experiment of compression of 125022 power load profiles has been carried out with the use of registered profiles in households and small offices. Two matrices: patterns (atoms) and scaling factors are the result of the discussed algorithm. Features of the created matrices, which can be used in the creation of fast power demand forecasting systems, have been characterized.

## 1 Introduction

The development of low and medium voltage power network infrastructure, especially in suburban and rural areas, is connected, among others, with the connection of new small and medium-sized generation sources and prosumer installations [1,2]. The connection of the above-mentioned generation sources to the power system, in particular of local operation, causes that part or all of the electricity generated by these sources is consumed on a regular basis, can be stored or sent to the transmission network. In this situation, the energy flow changes significantly, which affects the power demand in the power station (power demand profile) [1]. In addition, electricity consumers from households (as well as companies) often change their habits, so as to use electricity generated locally (e.g. from renewable energy sources) as much as possible, which can significantly reduce their purchase cost. This situation requires a different approach to forecasting the demand for the power of individual consumers (households or companies) but also the demand for power in a given area of the power station operation. This requires continuous monitoring of energy generated by sources, stored energy, demand for the power of recipients as well as power demand in a given area, which translates into the collection and processing of a significant amount of data [3].

There are and are being developed all the time wired and wireless systems of the Smart Metering type with local and remote registration, which collect huge amounts of measurement data, including load profiles of electricity consumers [4,5]. The amount of data collected is successively increasing, which often prevents quick and effective analysis,

---

\* Corresponding author: [marcin.drechny@utp.edu.pl](mailto:marcin.drechny@utp.edu.pl)

e.g. for prognostic purposes or to offer recipients dynamic tariffs prepared on the basis of their archival profiles.

One of the solutions that can be used to quickly analyze the large amounts of data collected, e.g. power load profiles, is their compression. In the literature [6] you can find suggestions for the use of various types of compression, but most of the proposals require time-consuming decompression of the load profile which prevents quick and effective data analysis and processing. Among the various solutions, one can find a proposal for the use of the sparse coding method using the SVD (Singular Value Decomposition) method, i.e. the distribution of the data into singular values [7,8,9]. The proposed version of this algorithm is the NN-K-SVD (Non Negative K-SVD) version, which only applies to non-negative values. In this method, estimating the shape of the load profile does not require time-consuming decompression, but only summing up the scaled up several model components called atoms.

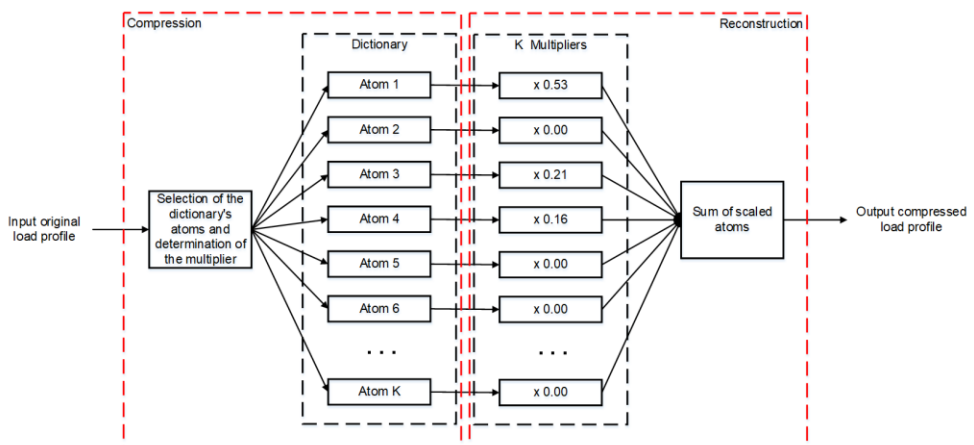
This paper discusses the use of the NN-K-SVD compression method for power load profiles, presents the results of the experiment, including exemplary reference atoms and reconstructed profiles, discusses problems related to profile compression and characterizes the use of the created atom dictionary and matrix of scaling factors in the demand forecasting system for power.

## 2 Compression with the use of NN-K-SVD method

### 2.1 General idea of compression with the use of sparse coding

The principle of compression using the sparse coding method used by the NN-K-SVD algorithm will be presented on the example of compression of power load profiles.

The general principle of compression and decompression using the sparse coding method consists of two stages [7,9]. In the first stage, a base of patterns, called atoms, is created. Then the compressed profiles are distributed over the sum of the patterns. Pattern numbers and values of scaling factors of these patterns are stored. Reconstructing a single compressed profile is a simple linear combination of selected profiles and their scaling values. The principle of using the profile compression and decompression algorithm is shown in Figure 1.



**Fig. 1.** General principle of compression and decompression method with sparse coding [9].

For several or even several thousand profiles it is required to store information about the atoms used for compression and their scaling factors. This is done in properly prepared matrices. The reconstruction of a single compressed profile is a linear combination of appropriate atoms and scaling coefficients as represented by the expression (1):

$$x_i \approx \sum_{k=1}^K d_k \cdot a_{i,k} \quad (1)$$

where,  $x_i$  – restored  $i$ -th profile,  $d_k$  –  $k$ -th atom of the pattern,  $a_{i,k}$  – scaling coefficients for  $i$ -th reconstructed profile and  $k$ -th atom,  $K$  - number of atoms.

In Figure 1, the compressed profile is represented by three atoms: first, third and fourth with scaling factors: 0.53, 0.21 and 0.16 respectively. Generally, in the matrix notation, the expression describing the reproduction of many profiles placed in the matrix is shown in the following expression:

$$\mathbf{X} \approx \mathbf{D} \cdot \mathbf{A} \quad (2)$$

where,  $\mathbf{X}$  – matrix of restored load profiles,  $\mathbf{D}$  – matrix of dictionary atoms,  $\mathbf{A}$  – matrix of scaling factors.

What is gained using such a compression method? If one daily load profile is averaged every 15 minutes, the 96 values must be stored (registered). Assuming that this profile will be archived using only 3 patterns with the use of only 3 scaling factors, it can be easily stated that 96 values are represented by only 6 values. This means a compression ratio of 16:1. In some cases, it is required to register the load profile with minute averaging which will increase the number of recorded values to 1440. In this case, three atoms and their scaling factors give a compression ratio of 240:1.

The problem is the fidelity of the profile reconstruction. When increasing the compression rate, i.e. reducing the number of atoms used, the reconstructed profile can be heavily distorted. However, if the compression ratio is reduced, the profile will be reconstructed correctly but the data number will decrease by a small amount, and then profile compression is not indicated. The most important in this case is the compromise between the degree of compression and the accuracy of the reconstruction of the compressed profiles. These issues are discussed in detail in [9].

## 2.2 Example of use NN-K-SVD algorithm to load profile compression

The method of using the NN-K-SVD method for compression of power load profiles with the use of sparse coding is described in detail in [9]. It consists of two stages. The first is named Sparse Coding stage, the second is Dictionary update.

In the first stage, in order to determine the matrix of scaling factors  $\mathbf{A}$  from the expression 2 assuming a given number of scaling factors, the sparse coding method is used. Electric power load profiles take positive values, therefore methods are used that set coefficients only as non-negative [10]. The reproduction error, as the target function, should be minimized assuming the unchanging shape of atoms.

The second stage depends on correcting, by the SVD method, the shape of atoms and matching them to the shapes of archived profiles. The first and second stages are performed with the predetermined number  $J$  times or until the assumed minimum error value.

In order to present the operation of the method, an experiment was carried out during which 127750 electric power load profiles, hourly averaged, registered mainly in households and small offices, were subjected to compression. During the initial data processing, 2728 profiles with all zero values and incomplete profiles in which all values were not recorded were removed. The next profiles have been normalized.

The following compression parameters have been assumed:

- size of the atom dictionary K: 60 and 100,
- number of compressing / reconstructing atoms s: 1 and 5,
- number of iterations of the J: 5 algorithm.

Below, in Figures 2-5, the results of the NN-K-SVD method are graphically presented in the form of selected best-recovered profiles. The black solid line shows the original load profile, while the red dotted line shows the restored profile. The profile recovery error was determined by calculating the RMS Mean Square Error (RMSE) error from the following expression:

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (x_i - \hat{x}_i)^2}, \quad (3)$$

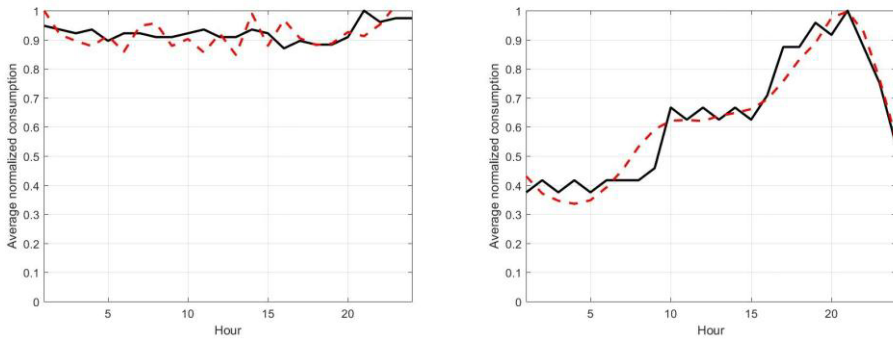
where, M – numbers of profiles,  $x_i$  – i-th original load profile,  $\hat{x}_i$  - i-th reconstructed load profile.

Table 1 shows the comparison of RMSE errors of minimum, maximum and average profile reconstruction for different compression variants.

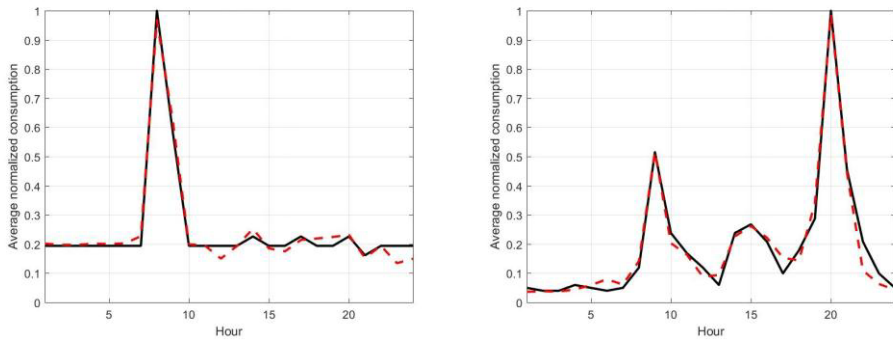
**Table 1.** RMSE error values after reconstruction of load profiles.

| RMSE | Compression parameters: K / s |        |         |         |
|------|-------------------------------|--------|---------|---------|
|      | 60 / 1                        | 60 / 5 | 100 / 1 | 100 / 5 |
| min  | 0,0097                        | 0,0054 | 0,0081  | 0,0055  |
| max  | 0,1679                        | 0,1019 | 0,1087  | 0,1019  |
| mean | 0,0435                        | 0,0284 | 0,0433  | 0,0285  |

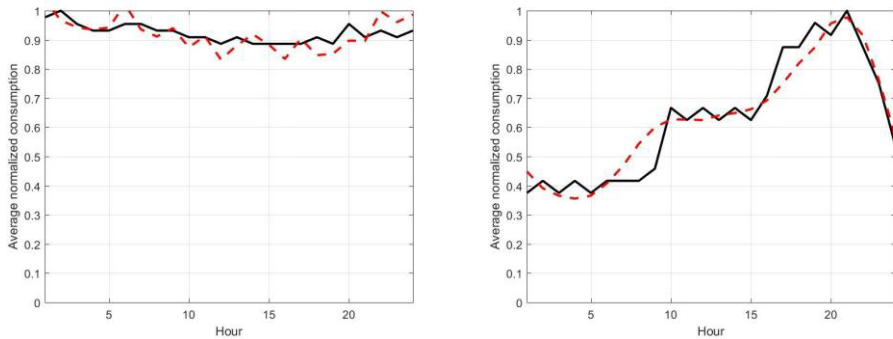
As might be expected, increasing the number of atoms from which a compressed profile is reconstructed results in less errors. It is important that increasing the number of dictionary atoms does not significantly improve the quality of the profile being reconstructed. It is important when optimizing the method aimed at increasing the compression rate without reducing the accuracy of the reconstruction using a smaller number of standard atoms.



**Fig. 2.** Example profiles before compression and after reconstruction ( $K = 60, s = 1$ ).

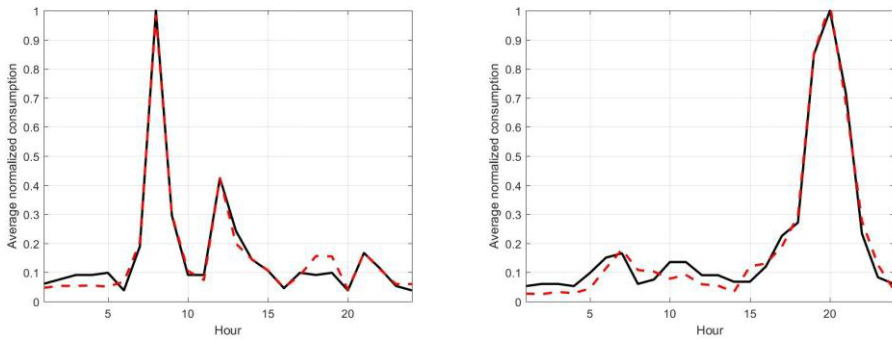


**Fig. 3.** Example profiles before compression and after reconstruction ( $K = 60, s = 5$ ).



**Fig. 4.** Example profiles before compression and after reconstruction ( $K = 100, s = 1$ ).

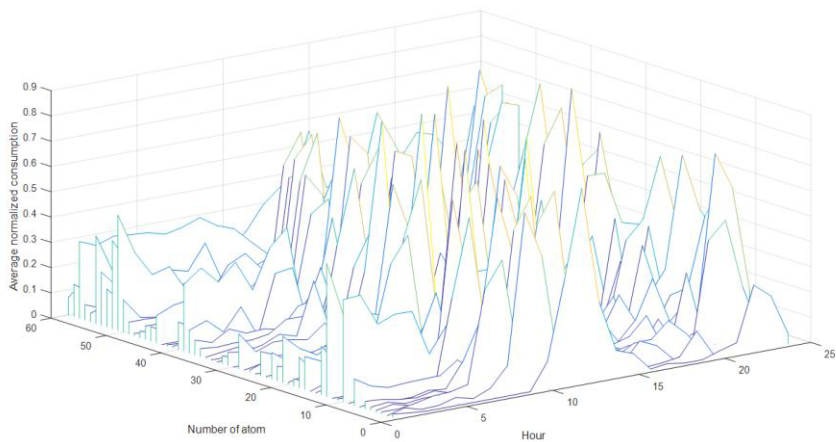
Both for profiles composed of 1 and 5 atoms, the shape of the pattern can be estimated. However, for compression using 5 atoms, more complex profiles are well reconstructed, as shown in Figures 3 and 5. Figures 2 and 4 show that the trend of the profiles being reconstructed is not quick-changing, that is, the slow-changing profiles reproduced accurately. Therefore, the most important aspect of using this method are formed atoms, because their shape has the greatest impact on the accuracy of profile reconstruction.



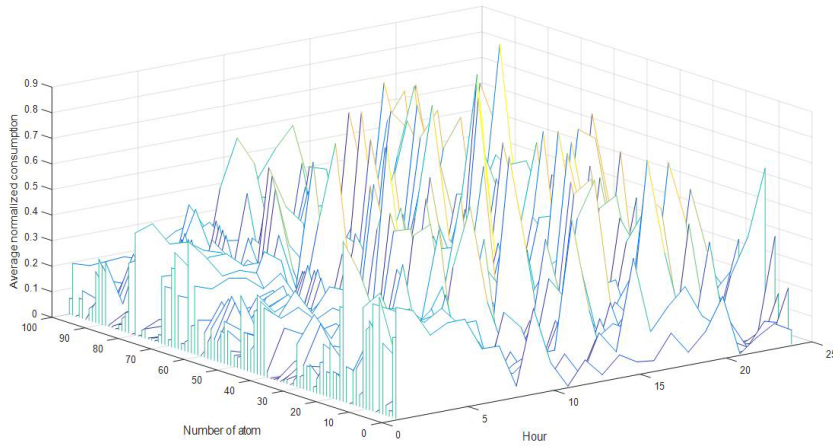
**Fig. 5.** Example profiles before compression and after reconstruction ( $K = 100$ ,  $s = 5$ ).

### 2.3 Base atoms

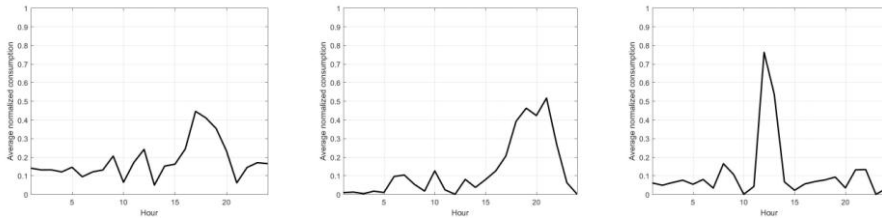
As mentioned earlier, one of the key elements of the presented compression algorithm is the created atom base. The accuracy of the reconstruction of compressed profiles depends from the amount and shape of atoms. In the example presented in chapter 2.2, the created base of atoms consists of 60 and 100 elements. The shapes of all the atoms that were created during the tests are shown in Figures 6-7. However, selected profiles from the 100 element dictionary are shown in Figure 8.



**Fig. 6.** All atoms of the 60-element dictionary.



**Fig. 7.** All atoms of the 100-element dictionary.



**Fig. 8.** Selected example atoms of the 100-element dictionary.

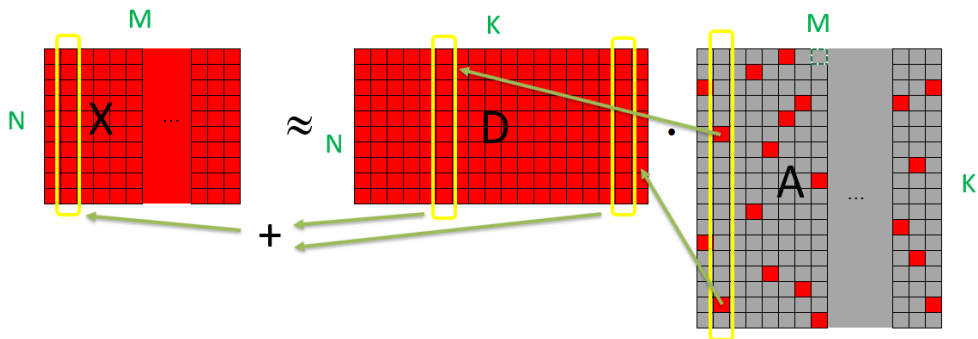
Based on figures 6-8, it can be concluded that using the shape of the atom used to compress the original load profile, as well as the scaling factor, you can estimate the potential behaviour of the electricity consumer in the future. In this case, it is not necessary to store all several or even several hundred thousands archival profiles, but only a few values referring to the atom number and the scaling factor value.

### 3 The possibility of using atoms and scaling coefficients in forecasting system

Prediction algorithms, in particular ultra-short-term power demand, are based on various types of information, including network configuration, meteorological data, date and time, and archival data recorded in the form of e.g. load profiles [11]. As mentioned in chapter 1 and 2, the dynamic development of power grids and new services, e.g. offering dynamic tariffs to consumers, forces a new approach to management of these networks, in particular the collection of a significant amount of data and their quick processing. Measurements and registration of power or electricity performed every hour, 30 or even 15 minutes may not allow quick response to dynamic changes in the power demand. On the market there are more and more electricity meters with minute averaging which solves the problem of the accuracy of profile registration but at the same time forces the storage and processing of more and more data. In this situation, it is advisable to use data compression methods, including those that allow quick access to them without time-consuming decompression.

An example of such compression is the NN-K-SVD method described above. The advantage of this method is the quick access to information based on the analysis of the location of non-zero scaling factors in the matrix  $A$  and the analysis of the values of these coefficients. In publication [12], this feature of the NN-K-SVD method has been implemented to identify the behaviour of electricity consumers.

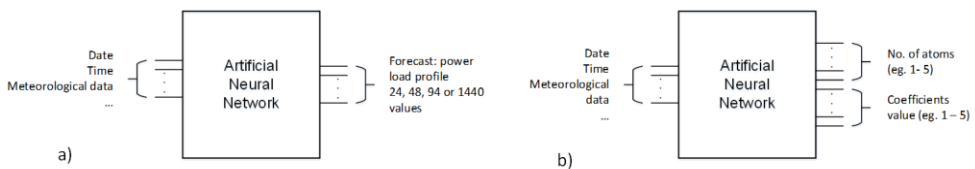
Based on expression 2, the graphic way of storing information in individual matrices is shown in Figure 8. The green colour of letters describes the dimensions of individual matrices. Yellow colour boxes and the green arrows show the process of reconstructing the second load profile located in the  $X$  matrix. This profile consists of two patterns - atoms, numbered 6 and 17 with the 18 element dictionary atoms  $D$ . The scaling factors from matrix  $A$  are multiplied by the corresponding atom from the matrix  $D$ . Next, the scaled atoms are summed up by reconstructing the compressed load profile.



**Fig. 8.** Graphical representation of the expression 2.

Each reconstructed profile consists of several atoms which scaling coefficients are placed in the matrix  $A$ . Forecasting the demand for the power of a given customer or area, you can analyze the shape of atoms from which the profile and the scaling coefficient of a given atom are composed. The advantage of this method is the reduction of the amount of stored data and quick access to archival information.

The proposal to applying the results of the NN-K-SVD method in the prognostic system based on artificial neural networks can be used for teaching and then predicting the matrixes  $A$  and  $D$ . An exemplary idea of such a system is presented in Figure 9.



**Fig. 9.** The idea of a prognostic system using artificial neural networks a) a classical approach, b) using atom numbers and scaling factors.

The input data to the forecasting system are the same in both cases. The output of this system is, in the case of the classical approach (a), load profile values, which can be as high as 1,440. In the proposed solution (b), the output of the forecasting system is several values corresponding to profile numbers and their scaling factors. Profit in the form of reducing the number of physical outputs of the artificial neural network, translates into simplifying



its construction (mainly reducing the number of neurons) will allow to increase the speed of its operation, in particular, speed up the time-consuming and often repeated learning process. Appropriate preparation of the atom dictionary so that you can use one atom along with the scaling factor while maintaining an acceptable level of error, will significantly simplify and accelerate the operation of the forecasting system.

## 4 Summary

The NN-K-SVD method in literature is mainly proposed for the compression of various types of signals, including power load profiles. Its disadvantage is the complex and time-consuming process of compression, i.e. determining the matrix of atoms and scaling values. However, this method has more advantages than disadvantages. Its main advantage is the uncomplicated and fast reproduction of the signal. Interesting features of this method are also the properties of the matrix of atoms and matrix coefficients. Appropriate interpretation of atomic shapes along with the scaling values assigned to them can be used not only for the purposes of profile reconstruction but also for the behaviour of electricity recipients as well as in power demand forecasting systems, especially where a quick decision or forecast is required.

This work is an analysis of the applicability of this method in the forecasting system and requires the development of a model of such a system and its practical verification, which is included in the plans for further research.

## References

1. A. Keyhani, *Design of smart power grid renewable energy systems*, 2nd edition Willey, (2017)
2. E. M. B. Cavalheiro, A. H. B. Vergílio, Ch. Lyra, *Computers & Operations Research*, Elsevier, **96**, 272-280 (2018)
3. P. D. Diamantoulakisb, V. M. Kapinasb, G. K. Karagiannidisab, *Big Data Research*, Elsevier, **2** (3), 94-101 (2015)
4. N. Uribe-Pérez, L. Hernández, D. de la Vega, I. Angulo, *Appl. Sci.* **6** (3), 68 (2016)
5. B. Dubalski, P. Kiedrowski, *Rynek Energii*, **5**, 48-53 (2010)
6. M. P. Tcheou, L. Lovisololo, M. V. Ribeiro, E. A. B. da Silva, M. A. M. Rodrigues, J. M. T. Romano, P. S. R. Diniz, *IEEE Transactions on Smart Grid*, **5** (1), 291-302 (2014)
7. M. Aharon, M. Elad, A. Bruckstein, *IEEE Transactions on Signal Processing*, **5** (11), 4311-4322 (2006)
8. J. C. Stacchini de Souza, T. M. L. Assis, B. Ch. Pal, *IEEE Transactions on Smart Grid*, **8** (1), 275-284 (2017)
9. Y. Wang, Q. Chen, Ch. Kang, Q. Xia, M. Luo, *IEEE Transactions on Power Systems*, **32** (3), 2142-2151 (2017)
10. P. O. Hoyer, *Proc. of the 12th IEEE Workshop on Neural Networks for Signal Processing*, 557-565 (2002)
11. A. Grzyb, P. Piotrowski, *Przegląd Elektrotechniczny*, **1** (93), 316-319 (2017)
12. M. Drechny, *Proc. of Innovative Materials and Technologies in Electrical Engineering (i-MITEL)*, 1-4 (2018)