**Krzysztof SIWEK[1], Stanisław OSOWSKI[1,2], Bartosz ŚWIDERSKI[3]**

Warsaw University of Technology (1), Military University of Technology (2), *University of Life Sciences* (3)

# Trend elimination of time series of 24-hour load demand in the power system and its application in power forecasting

***Abstract.*** *The paper is concerned with the elimination of different trends existing in the time series representing the hourly power consumption in the power system of small size. Analysing the hourly need for the power in such system we can observe significant trends associated with the season of the year, type of the day, as well as the particular hour of the day. At prediction task the variability of the time series is of great importance. The lower is this variability the better is the accuracy of prediction. In the paper we will present the method of reducing this variability by removing such trends. The elimination of trends is performed in few phases. The first step is to determine the index corresponding to the regular jth day of the weak for j = 1, 2, …, 7 and also type of such day. After this first step of detrending of the time series we follow the second and third aiming at removing the trend corresponding to the particular hour h of the day (h=1, 2, …, 24) and then the seasonality trend, characterizing the succeeding day of the year (d=1, 2, …, 365). All detrending operations are done by using the appropriately defined indexes. After application of all these steps we get the final detrended time series corresponding to all days under consideration (d=1, 2, …, p). The detrended time series is of much lower variance than the original one. This means the significant simplification of the forecasting problem and increase of probability of achieving better accuracy of forecasting results. The experiments of prediction of such detrended time series for small power region of Łódz performed using two types of neural predictors (MLP and SVM) have proved the superiority of such approach.*

***Streszczenie.*** *Praca dotyczy usuwania różnego rodzaju trendów występujących w szeregu czasowym odpowiadającym obciążeniom godzinnym w systemie elektroenergetycznym. Zwykły szereg czasowy charakteryzujący pobór mocy, zwłaszcza w małym systemie elektroenergetycznym, charakteryzuje się występowaniem trendu związanego z charakterystycznymi cechami danego dnia tygodnia, sezonem oraz godziną doby. Wielkości te występują w miarę regularnie i istnieje możliwość ich znacznego złagodzenia. Usunięcie tych trendów powoduje istotne zmniejszenie różnic obciążeń występujących z godziny na godzinę. Oznacza to zmniejszenie zmienności analizowanego szeregu i w efekcie zwiększenie dokładności jego predykcji w procesie prognozowania obciążeń na nowy dzień. Praca dotyczy aspektu usuwania trendów różnego rodzaju poprzez wprowadzenie tzw. indeksów normalizacyjnych. Pokażemy, że zastosowane podejście pozwala na istotne zmniejszenie odchylenia standardowego odniesionego do wartości średniej obciążenia. Dzięki temu możliwe jest uzyskanie lepszej dokładności predykcji szeregu czasowego. Eksperymenty numeryczne przeprowadzone na danych z Łódzkiego systemu elektroenergetycznego potwierdziły, że istotnej redukcji ulega przede wszystkim błąd maksymalny. (**Usuwanie różnego rodzaju trendów występujących w szeregu czasowym odpowiadającym obciążeniom godzinnym w systemie elektroenergetycznym**)*

**Keywords:** time series, power system, trend removal, forecasting, neural networks
**Słowa kluczowe:** szereg czasowy, system elektroenergetyczny, usuwanie trendu, prognozowanie, sieci neuronowe

## Introduction

Time series forecasting in the power system plays an important role in reducing the cost of delivering the electrical power to customers. The accurate forecast of the 24-hour load for the next day on the basis of its past history is strictly associated with the variability of the considered time series. It is especially true for small power systems representing the small regions of the country. The smaller is this variability measured by the standard deviation of the system, the easier prediction task and higher probability of achieving better accuracy of prediction. By removing different types of trends of considered time series we can reduce this variability in a significant way.

Look for example at the hourly consumption of the electrical energy of the small power region of Lodz in Poland (Fig. 1) in one year. We can easily see the trends concerning these dependencies. The upper figure represents the power consumption for the period of 1 year. We can see quite clearly the yearly seasonal changes: smaller consumption of energy in the summer months and much higher consumption for winter days. To observe the other (weekly and hourly) trends we present in the bottom figure the period limited to one week only. We can see the daily changes associated with different days of the week (from Monday to Sunday) as well as different hours of the day (small consumption at night hours and much higher consumption at the day.

The most straightforward approach to reduce the variation of the power time series is to remove the trends due to the daily, weekly and seasonal cycles of the power demand [1,2,3]. The paper develops this direction of the time series processing. We propose the indexation of the time series aimed at the elimination of the trends due to the type of the day, daily hour cycle and the seasons of the year. The transformed time series after removing these trends are of much lower variability. The profit of trend removal will be measured in this work by comparing the ratio of standard deviation to mean for the original data and the transformed data. We will show that our approach to the problem can greatly reduce this measure.

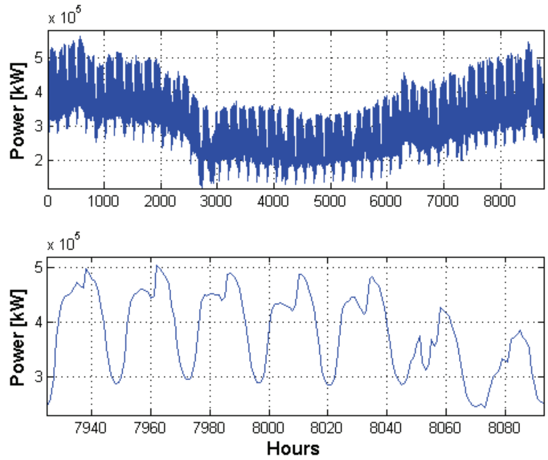


Fig. 1 The example of hourly consumption of the electrical energy in a small power region. The upper curve presents the data of the whole year and the bottom one of the chosen week

## Elimination of trends

Because of different trends existing in the time series we have to perform their elimination in few phases. First we eliminate the trend associated with the type of the day. Then this detrended time series is subject to elimination of the trend characteristic for particular hour of the day. At last step we remove the seasonal trends associated with the succeeding days of the year. At each step of detrending processes the next operations are performed on the already detrended time series. Removing trends at each step is done by indexing the time series.

***Daily trend***

In the first step of trend elimination we remove the trend associated with $j$th day of the weak ($j$ = 1, 2, …, 7). First we define the appropriate index, calculating the mean value of the load for each $j$th day of the year, dividing it by the mean of the year and averaging over all years taking part in calculations. Let us denote this index by $\alpha_{dw}$. Then we use the following formula

$$\alpha_{dw}(j) = \underset{years}{mean}\left(\frac{P_m(y,j)}{P_m(y)}\right) \tag{1}$$

In this expression $P_m(y,j)$ denotes the average power consumption of $j$th day of the week corresponding to $y$th year, and $P_m(y)$ is the mean power for all days of $y$th year. After its calculation we remove the daily trend by dividing the real time series $\mathbf{P}(w,j)=[P_1(w,j), P_2(w,j),…, P_{24}(w,j)]$ of each $j$th day of $w$th week by the value of daily index $\alpha_{dw}(j)$, i.e.,

$$\mathbf{P}_1(w,j) = \frac{\mathbf{P}(w,j)}{\alpha_{dw}(j)} \tag{2}$$

The trend associated with the particular day of the week is responsible for the normal patterns of the days within the week. However there are special situations when the regular working days are located in different ways with respect to non-working days. To consider it we have done additional indexation associated with the load patterns corresponding to workdays and holidays. Additionally we have defined 5 special types of the days:

- the workday just before non-working day (holiday or weekend)
- the workday just after non-working day (holiday or weekend)
- the workday between two non-working days
- all other working days
- the non-working days.

To remove such trend we define the day type index $\alpha_{dt}$ in the way

$$\alpha_{dt}(t) = \underset{years}{mean}\left(\frac{P_{1m}(y,t)}{P_{1m}(y)}\right) \tag{3}$$

where $P_{1m}(y,t)$ denotes the average power consumption $\mathbf{P}_1$ of the days of the same type $t$ (workdays or holidays) corresponding to $y$th year, and $P_m(y)$ is the mean power for all days of $y$th year. The final value of this index is averaged over all years under consideration. Removing this trend is equivalent to division of the real time series $\mathbf{P}_1$ by the value of $\alpha_{dt}(t)$, i.e.,

$$\mathbf{P}_2(w,t) = \frac{\mathbf{P}_1(w,t)}{\alpha_{dt}(t)} \tag{4}$$

Fig. 2 shows the distribution of values of the daily indexes $\alpha_{dw}(j)$ and the day type indexes $\alpha_{dt}(t)$ for the time series presented in Fig. 1.

To compare the time series before and after daily trend removal we have calculated the measure of variability

$$k_{sm} = \frac{std(P(d,h))}{mean(P(d,h))} \tag{5}$$

for the original time series and the time series after elimination of trend. For original time series the value of this measure was equal $k_{sm}$= 0.289. After removing daily trend this value was reduced to $k_{sm}$= 0.267.

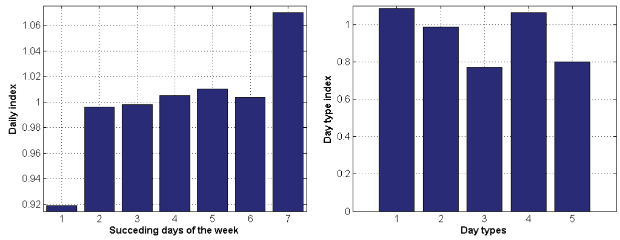


Fig. 2 The values of indexes $\alpha_{dw}(j)$ and $\alpha_{dt}(t)$ responsible for daily trend removal.

***Hourly trend***

The next step is removing the trend corresponding to the particular hour $h$ of the day ($h$=1, 2, …, 24). The procedure is identical to the previous ones. The hourly index $\alpha_h(h)$, is defined as

$$\alpha_h(h) = \underset{years}{mean}\left(\frac{P_{2m}(y,h)}{P_{2m}(y)}\right) \tag{6}$$

Removing the trend corresponding to the particular hour of each $d$th day is equivalent to the following operation

$$P_3(d,h) = \frac{P_2(d,h)}{\alpha_h(h)} \tag{7}$$

Fig. 3 presents the change of the index $\alpha_h(h)$ for 24 hours of the days. For each day of the year this value depends only on the hour.

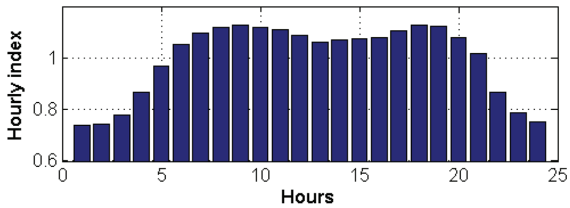


Fig. 3 The value of index $\alpha_h(h)$ responsible for hourly trend removal

After removing trend corresponding to the particular hour we have observed the reduction of the measure of variability from $k_{sm}$=0.267 of the starting time series to $k_{sm}$=0.231 of the time series after hourly trend removal.

***Seasonal trend***

The last operation is to remove the seasonality trend characterizing the succeeding days of the year ($d$=1, 2, …, 365). Let us denote the seasonality index by $\alpha_s$. We define this index in the following form

$$\alpha_s(d) = \underset{years}{mean}\left(\frac{P_{3m}(y,d)}{P_{3m}(y)}\right) \tag{8}$$

where $P_{3m}(y,d)$ denotes the average power consumption $\mathbf{P}_3$ of all 24 hours of the day $d$ corresponding to $y$th year, and $P_{3m}(y)$ is the mean power for all days of $y$th year. Fig. 4 presents the change of the seasonality index $\alpha_s(d)$ for all

365 days of the year. Each day of the year is described by the specific value of the index, characteristic for this day.

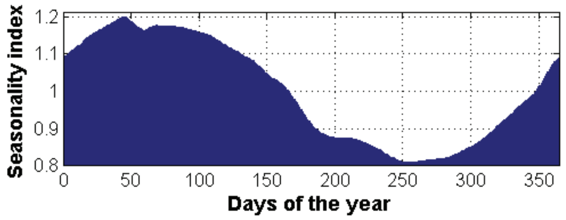


Fig. 4 The distribution of values of index $\alpha_s(d)$ responsible for seasonal trend removal

After this step of detrending we can define the transformed load pattern $P_4(d,h)$ of each hour of the days under consideration in the final form

(9) $$P_4(d,h) = \frac{P_3(d,h)}{\alpha_s(d)}$$

This final form of the time series $\mathbf{P}_4(d)=[P_4(d,1), P_4(d,2),\ldots, P_4(d,24)]$ represents the detrended time series of the power consumption corresponding to all days under consideration ($d$=1, 2, …, $p$).

Removing the trends existing in the time series reduces the variation of the considered time series, making it more balanced of reduced standard deviation. The last step of transformation has reduced the measure of variability from the starting value $k_{sm}$= 0.231 to the final one equal $k_{sm}$=0.128.

Fig. 5 presents the detrended time series of the power demand corresponding to the original one, presented in Fig.1. The upper figure depicts the hourly data corresponding to the whole period under consideration. The original power consumption was varied within the range (110-550)MW. The range of power changes in the detrended time series was reduced now to (200-470)MW. This is significant reduction which eases the prediction process of this time series. It means the simplification of the forecasting problem and increases of probability of achieving better accuracy of results.

The bottom curve represents the time series limited to one week chosen from the above curve. It is evident that the this time series is much more balanced than the original one (presented in Fig. 1b) and is deprived of the day type trends. The load pattern representing different days of the week is much more alike for different days.

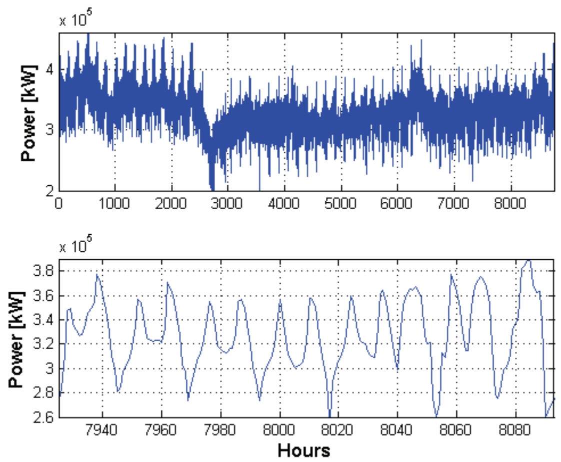


Fig. 5 The detrended time series of the power demand corresponding to Fig. 1: a) the whole set of data, b) the chosen one week data

To compare the final result of detrending we have gathered in Table 1 the ratios of the standard deviation to the mean value of the power consumption for all 24 hours of the day calculated for the whole data of consideration. They represent the original data $P(d,h)$ and the detrended time series $P_4(d,h)$.

Table 1 The ratio of the standard deviation to mean load of the hourly power demand of the original $P(d,h)$ and detrended ($P_4(d,h)$) time series

| Hour $h$ | $\frac{std(P(d,h))}{mean(P(d,h))}$ | $\frac{std(P_4(d,h))}{mean(P_4(d,h))}$ |
|---|---|---|
| 1 | 0.208 | 0.153 |
| 2 | 0.225 | 0.152 |
| 3 | 0.237 | 0.150 |
| 4 | 0.244 | 0.148 |
| 5 | 0.264 | 0.153 |
| 6 | 0.285 | 0.152 |
| 7 | 0.300 | 0.148 |
| 8 | 0.293 | 0.140 |
| 9 | 0.266 | 0.119 |
| 10 | 0.244 | 0.099 |
| 11 | 0.231 | 0.089 |
| 12 | 0.226 | 0.086 |
| 13 | 0.224 | 0.086 |
| 14 | 0.231 | 0.088 |
| 15 | 0.235 | 0.091 |
| 16 | 0.245 | 0.098 |
| 17 | 0.283 | 0.133 |
| 18 | 0.298 | 0.151 |
| 19 | 0.295 | 0.156 |
| 20 | 0.282 | 0.152 |
| 21 | 0.243 | 0.126 |
| 22 | 0.195 | 0.101 |
| 23 | 0.177 | 0.095 |
| 24 | 0.189 | 0.138 |
| whole data set | 0.289 | 0.128 |

We can observe the significant change of results. The ratio std/mean in all cases has been reduced twice and even more, after introducing indexation. The average improvement of this ratio calculated over all 24 hours is 2.25.

Observe, that prediction task has been moved now to the time series represented by $P_4(d,h)$ of much smaller variability. Smaller variance of this time series means easier prediction task for the forecasting systems and higher probability of achieving better accuracy.

After predicting the time series $P_4(d,h)$ we can return to the original values. Taking into account the accumulation of all indexing operations the real load pattern corresponding to $h$th hour of $d$th day can be presented as follows

(10) $$P(d,h) = P_4(d,h)\alpha_{dw}(j)\alpha_{dt}(t)\alpha_h(h)\alpha_s(d)$$

The index $j$ corresponds here to proper day of the week ($j$ = 1, 2, …,7) and $t$ denotes the actual type of the working or nonworking day. For each day $d$ and hour $h$ the proper values of indices should be applied. Knowing them in advance the prediction task is simplified to prediction of the detrended values $P_4(d,h)$. This task can be done by applying the neural network.

**Neural prediction of detrended time series**

The prediction of the detrended time series $P_4(d,h)$ will be done in this work by applying two types of supervised neural networks: the multilayer perceptron (MLP) and Support Vector Machine in regression mode (SVR) [3,7]. The predictive model is built for each hour independently. In building this model for $d$th day and $h$th hour we assume that all previous values of $P_4(i,j)$ are available for $i= d-1, d-2, \dots$ and $j= 24, 23, \dots 1$. The assumed model of prediction takes into account the predicted value of minimal $T_{min}(d)$ and maximal $T_{max}(d)$ temperatures of the next $d$th day, the values of the load in the last 2 hours of two previous day, as well as the load of $h$th and ($h$-1) hours of the same type of day (the same day a week ago). This general model may be written in the following form

$$\hat{P}_4(d,h) = f(P_4(d-1,h), P_4(d-1,h-1), P_4(d-2,h), P_4(d-2,h-1), P_4(d-7,h), P_4(d-7,h-1), T_{\min}(d), T_{\max}(d)) \tag{11}$$

The values denoted by hat represent the predicted and without hat – the real power consumption. The expression $f()$ is the approximation function implemented by the neural network, either MLP or SVR.

The expression (11) defines explicitly the input signals applied to the neural predictors. They are the variables appearing in the brackets on the right side of the equation. Irrespective of the applied neural network they are composed of 8 signals: six correspond to the previous (detrended) load and two to the predicted minimal (night) and maximal (day) temperatures of the day.

At application of MLP as a predictor we have to assume at the beginning the number of hidden sigmoidal neurons. We have solved this problem in experimental form using the validation data set. 24 MLP networks were trained, each for the particular hour of the day.

In the case of SVR we have applied the Gaussian kernel function of the hyperparameters $\sigma$, the regularization constant $C$ end the tolerance parameter $\varepsilon$ guaranteeing the best results on the testing data not taking part in learning. The process of optimizing the values of $C$, $\varepsilon$ and $\sigma$ was done simulataneously, for each hour of the day. Many different values of $C$, $\varepsilon$ and $\sigma$ combined together in the learning process have been used in the learning process and their optimal values are those for which the classification error on the validation data set was the smallest one. For each hour the separate SVR network was trained (24 networks together) using the modified sequential programming algorithm [8,9] implemented in Matlab [6].

**The numerical results**

The numerical experiments we have done using the data of small power region of Lodz. The calculations have been performed in two phases. The aim of the first phase was the detrending process of the whole data. As a result of this we have determined all four indexing coefficients. As a result we have got the detrended time series $P_4(d,h)$ corresponding to all data set. This data set has been normalized column wise by dividing all entries by the maximum value corresponding to of each hour.

The second phase of data processing is the prediction of the detrended time series. It has been provided by using neural predictors of MLP and SVR structures. The whole set of data has been split into two parts. Two third of it was used for learning the predictor and one third for testing the network. Both predictors used the same structure of input data defined by expression (10). Eight input signals to the neural networks have been applied, as presented in the previous section. In both networks we used single output neuron, responsible for predition of the normalized power of the particular hour. 24 neural predictors corresponding to each of 24 hours of the day were trained.

The number of hidden neurons of MLP was changing in the networks from 5 to 8 neurons (all of sigmoidal nonlinearity). In the case of Gaussian kernel SVR the tolerance $\varepsilon$ was fixed to 0.005, while the optimal values of $C$, and $\sigma$ have been determined by trying the predefined values of them and using the validation data set. As a result of such introductory experiments we have fixed the regularization constant for all 24 SVR networks on $C$=100 and the Gaussian width $\sigma$=10. These hyperparameters have been used in real learning procedure of all SVRs.

After designing the neural predictors they have been tested on the data set not used in learning. As a result of it we got the forecasted values of $\hat{P}_4(d,h)$, on the basis of which we were able to determine the real predicted values $\hat{P}(d,h)$ for the days used in testing (equation 10). The results of prediction have been assessed using different measures of errors, including mean absolute percentage error (MAPE), maximum percentage error (MAXPE) and root mean squared error (RMSE).These quality factors have been defined as follows

$$MAPE = \frac{1}{n}\sum_{h=1}^{n} \frac{\left|P(h) - \hat{P}(h)\right|}{P(h)} \cdot 100\% \tag{12}$$

$$MAXPE = \max\left\{ \frac{\left|P(h) - \hat{P}(h)\right|}{P(h)} \cdot 100\% \right\} \tag{13}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{h=1}^{n} \left|P(h) - \hat{P}(h)\right|^2} \tag{14}$$

In these definitions $n$ is the quantity of hours used in testing, $P(h)$ – the real power consumption of $h$th hour and $\hat{P}(h)$ - the predicted value of $P(h)$.

Table 2 depicts the obtained results of testing the neural predictors in the form of MAPE, MAXPE and RMSE. They have been compared with the results of prediction by using crude data, without detrending procedure (the direct application of SVM and MLP).

Table 2 The results of prediction by using different solutions of forecasting systems

| Prediction method | MAPE [%] | MAXPE [%] | RMS [kW] |
|---|---|---|---|
| Detrending+SVM | 3.4855 | 49.1397 | 1.4531e4 |
| Detrending+MLP | 3.5352 | 50.0892 | 1.5162e4 |
| Direct SVM | 3.5824 | 81.0938 | 1.5782e4 |
| Direct MLP | 3.6559 | 94.7229 | 1.5836e4 |

Analysing these results we can see that direct application of neural predictors (without detrending) is evidently worse, although the difference in MAPE and RMSE is not very significant (3.38% instead of 3.58% for the best solution). Both SVM and MLP predictors were of comparable accuracy, although the SVM was slightly better. However very high improvement was observed for maximum percentage error. Application of detrending procedure has reduced this error in a very significant way. For example at direct application of SVM the MAXPE=81.09%. After detrending the data this error was reduced to 49.14% (39.5% of relative improvement). The explanation for this may be the fact, that the detrending procedure reduces significantly the abrupt changes of the time series components corresponding to the specific types and certain hours of the day. Hence their accurate prediction is much easier.

Interesting observation may be made analyzing the results for succeeding hours of the day. Fig. 6 depicts the values of MAPE and MAXPE errors for individual hours corresponding to the best solution (detrending+MLP).

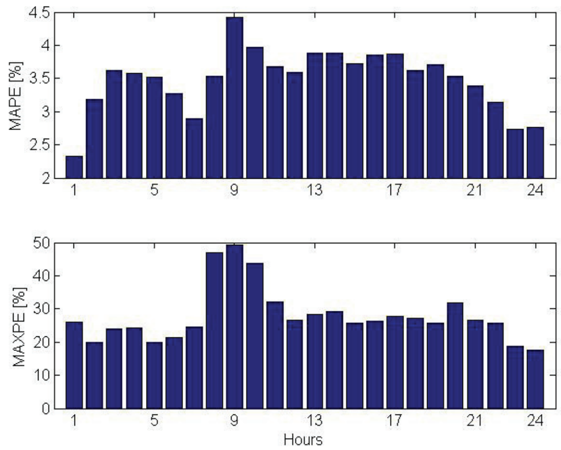


Fig. 6 The distribution of MAPE and MAXPE for succeeding hours of the day for the testing data not taking part in learning

The best accuracy are observed for hours of relatively small need of power (the night hours). The hours of the highest values of errors correspond to the time of rush hours following from the daily activity of industry and individual humans.

**Conclusions**

The paper has presented the novel approach to the elimination of different trends in the power system. The aim of this is to reduce the variability of the time series under prediction. The detrending process has been done in few steps, successively applied to the original time series. It includes the trends associated with different days of the week, hourly trend as well as the seasonal one. The numerical experiments made for the small power region of Poland has proved that elimination of trends has significantly (more than twice) reduced the ratio of std to mean for the time series under consideration. Such great reduction of this variability makes the process of prediction much easier.

Application of neural networks as the predictors for the detrended power series has proved that this method of data preprocessing results in significant improvement of the accuracy of prediction, especially reduction of the maximum error. The developed method is especially recommended for the small power regions, where the variability of the load is very high.

ACKNOWLEDMENTS

This work was supported by The Polish Ministry of Science and Education as a research grant within the years 2010-2012.

REFERENCES

[1] Box G. E,. Jenkins G. M, Time series analysis – forecasting and control, San Francisco, Holden-Day, (1976)
[2] Gonzalez-Romera, E., Jaramillo-Moran, M. A., Carmona-Fernandez D.: Monthly electric energy demand forecasting based on trend extraction. *IEEE Trans. Power Systems*, 21, (2006), 1946-1953
[3] Haykin, S.: Neural networks, a comprehensive foundation. Macmillan, N. Y. (2002)
[4] Koehler A., Snyder R., Ord J. K., Forecasting models and prediction intervals fort he multiplicative Holt-Winters method, *Int. J. Forecasting*, 17 (2001), 269-286
[5] Osowski S., Siwek K., Szupiluk R., Ensemble Neural Network Approach for Accurate Load Forecasting in the Power System, *Applied Mathematics and Computer Science*, 19 (2009), No.2, 303-315
[6] Matlab user manual, MathWorks, Natick, (2008)
[7] Osowski, S., Siwek, K.: Regularization of neural networks for load forecasting in power system. IEE Proc. GTD, 149, (2002), 340-345
[8] Osowski S., Sieci neuronowe SVM w zastosowaniu do problemów regresji, Przegląd Elektrotechniczny, 2002, vol. 78, No 9, ss. 225-229
[9] Schölkopf, B., Smola, A.: Learning with Kernels. Cambridge, MA: MIT Press (2002)

***Authors***: *dr inż. Krzysztof Siwek, Warsaw University of Technology, Email: ksiwek@iem.pw.edu.pl, prof. dr hab. inż. Stanisław Osowski, Warsaw Unversity of Technology and Military University of Technology, E-mail: sto@iem.pw.edu.pl; dr inż. Bartosz Swiderski, University of Life Sciences, Email swidersb@iem.pw.edu.pl,.*