Comparative analysis of the error prediction of electricity consumption on a monthly and weekly level

Jelena Milojković and Vančo Litovski

Abstract - One month and one week ahead predictions of suburban average electricity load are presented. Although we have a lot of data available for our work, only the most recent of it may be of importance. Consequently, we managed with limited amount of data, and we proposed two independent mutually supporting solutions of artificial neural networks (ANN). ANN have been proven as very reliable in real time system such is electricity consumption. Prediction with ANN is the topic of our previous work where we obtained small prediction errors. In this paper it will be shown a comparative analysis of the prediction error in the cases of monthly and weekly forecasting of the electricity load. In this way, we will try to emphasize importance to undertake these predictions in order to reduce the cost of production, transmission, consumption and other, of electricity load.

***Keynotes*— electric load forecast, error prediction, electricity, artificial neural networks**

I. Introduction

Electric load foerecasting has always been important for planning power generation and operational decision carried out by the manufacturer and utility companies. It is vital in many aspects such as providing price effective generation, system security, scheduling fuel purchases [1]. The load forecast error produces high extra costs: if the load is underestimated one has extra costs caused by the damages due to lack of energy or by overloading system elements; if the load is overestimated, the network investment costs overtake the real needs, and the fuel stocks are overvalued, locking up capital investment. Consequently, the quality of load forecasts, especialy short-term load forecasting, which we describe later, has greatly importance for energy supliers, utility companies, financial institutions, and other participants in electric energy generation, transmission, distribution, and martkets [2]. Accurate forecasting is very important as there are significant financial implications.

The power load value is determined by several environmental and social factors among which the seasonal and daily profiles are the most influential. Temperature and air humidity are the primary parameters determining the energy consumption generally and especially in urban residential areas. Working times, holidays, and seasonal behaviour influence the load-time function.

Jelena Milojković, ICAT, Niš, Bulevar Nikole Tesle 61, lokal 5, 18000 Niš, (email: jelena.milojkovic@icnt.rs)

Vančo Litovski, Independent researcher, Vojvode Mišića 60/18 Niš, Serbia, (email: vanco@elfak.ni.ac.rs)

All together, the load curve is a nonlinear function of many variables that map themselves into it in an unknown way.

In this paper [3] we found the inspiration for our work. Prof. Mendel claims: "Prediction of short time series is a topical problem. Cases where the sample length N is too small for generating statistically reliable variants of prediction are encountered every so often. This form is characteristic of many applied problems of prediction in marketing, politology, investment planning, and other fields." Further he claims: "Statistical analysis suggests that in order to take carefully into account all components the prediction base period should contain several hundreds of units. For periods of several tens of units, satisfactory predictions can be constructed only for the time series representable as the sum of the trend, seasonal, and random components. What is more, these models must have a very limited number of parameters. Series made up by the sum of the trend and the random component sometimes may be predicted for even a smaller base period. Finally, for a prediction base period smaller than some calculated value $N_{\min}$, a more or less satisfactory prediction on the basis of observations is impossible at all, and additional data are required".

Among the fields not mentioned in [3], dealing with really small set of data or "prediction base period", we will discuss here weekly and monthly short-term prediction of electricity loads at suburban level or on the level of a low voltage transformer station. In fact, the amount of data available in this case is large enough to apply any other forecasting method [4,5,6] but looking to the load diagram i.e. weekly (and monthly) load-value curves, we easily recognize that past values of the consumption are not very helpful when prediction is considered. That stands even for data from the previous week (month) and for data from the same week (month) in the previous month (year). Accordingly, we propose the problem of prediction of the load value in the next week (month) to be performed as a deterministic prediction based on very short time series. To help the prediction, however, in an appropriate way, we introduce past values e.g. load for the same week (month) but in previous month (year). That is in accordance with existing experience claiming that every month (week) in the year (month) has its own general consumption profile [7].

II. Related works

The prediction of a time series is synonymous with modelling the underlying physical or social process responsible for its generation. This is the reason of the task

Proceedings of 3rd International Conference on Electrical, Electronic and Computing Engineering
IcETRAN 2016, Zlatibor, Serbia, June 13 – 16, 2016, ISBN 978-86-7466-618-0

difficulty. There were many attempts during the past few decades to propose a solution to the short term load prediction. Among the most comprehensive overviews of the subject we find [4] and [1]. The methods applied may be categorized based on several aspects. By one categorization we see methods that use the weather information such as temperature and/or humidity as controlling variables or not [8]. On the other side a categorization exists based on the underlying mathematical algorithm used for modelling. From that point of view we first come to statistical methods (like auto-regression and time-series) predicting average values and deviations. Among them, the best known are the simple moving average (SMA) and the exponential moving average (EMA) method for prediction of trend [1,9]. That category includes the autoregressive integrated moving average (ARIMA) method [10] and similar as well. Although these statistical techniques are reliable, they fail to give accurate results when quick weather changes occur which form a nonlinear relationship with daily load [11]. Hence results of statistical methods in presence of such events are not satisfactory as desired. Therefore the emphasis has shifted to the application of various deterministic methods. Among the deterministic methods, one can find a two-fold categorization: parametric based method [5,12] and, much frequently encountered the artificial intelligence method that is often represented by implementation of ANNs [13].

The idea in our implementation is reminiscent to the substitution of the simple moving average (SMA) by the exponential moving average (EMA) method for prediction of trend [14,15]. The simple moving average is extremely popular among traders, but one argues that the usefulness of the SMA is limited because each point in the data series is equally weighted, regardless of its position in the sequence. It is common opinion that *most recent data is more significant than the older* and should have a greater influence on the final result. That led us into the subject of prediction based on short time series. Our idea is at the same time inspired by the classical deterministic method known as the *k*-nearest-neighbour [12], in which the data series is searched for situations similar to the current one each time a forecast needs to be made. This method asks for periodicity to be exploited that, in our case, may be helpful but not decisively.

Having all that in mind we undertook a project of developing an artificial neural network (ANN) based method that will be convenient for systematic implementation in stationary time series prediction with reduced set of data. Our first results were applied to prediction of environmental as well as technological data and published in [7,16,17]. Analysis as to why neural networks are implemented for prediction may be found in [7]. The main idea implemented was the following: If one wants to create neural network that may be used for forecasting one should properly accommodate its structure.

Following these considerations we already developed new forecasting architectures. The structures developed were named Time Controlled Recurrent (TCR) and Feed Forward Accommodated for Prediction (FFAP). Both were implemented successfully for prediction in modern developments in micro electronics [17] as well as in other areas including load prediction on yearly basis [18]. Encouraged by the excellent results, both networks were further improved and applied to the prediction of electricity consumption on weekly and monthly basis.

The main goal of this paper is making analisys of the results which we are obtained by using short term forecasting based ANN. We implemented proposed method for weekly and monthly prediction. In [21] we proposed an averaging method that will use both predictions in order to smooth the prediction error so making the final result as dependable as possible. Finally, we proposed a method for finding the proper number of hidden neurons in both networks.

The structure of the paper is as follows. After short review of the method and the new architectures of ANN application to forecasting we will go through the data and algorithm for finding optimal number of hidden neurons. After that we will present prediction error for both ANN, for monthly and weekly prediction. Finally, short discussion of the results and consideration related to future work will be given.

III. SHORT REVIEW OF THE METHOD AND SOLUTION

A time series is a number of observations that are taken consecutively in time. A time series that can be predicted precisely is called deterministic, while a time series that has future elements which can be partly determined using previous values, while the exact values cannot be predicted, is said to be stochastic. We are here addressing only deterministic type of time series.

Consider a scalar time series denoted by y_i, i=1, 2, … m. It represents a set of observables of an unknown function, taken at equidistant time instants separated by the interval Δt i.e. $t_{i+1}=t_i+\Delta t$. One step ahead forecasting means to find such a function $\hat{y}=\hat{f}(t)$, that will perform the mapping

$$y_{m+1} = f(t_{m+1}) = \hat{y}_{m+1} + \varepsilon , \qquad (1)$$

where $\hat{y}_{m+1}$ is the desired response, with an acceptable error ε.

The prediction of a time series is synonymous with modeling of the underlying physical or social process responsible for its generation. This is the reason of the difficulty of the task. There have been many attempts to find solution to the problem. Among the classical deterministic methods we may mention the *k*-nearest-neighbor [22], in which the data series is searched for situations similar to the current one each time a forecast needs to be made. This method asks for periodicity to be exploited that, as already discussed, here is not of much a help.

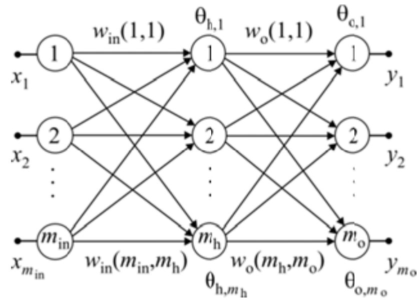

Fig. 1. Fully connected feed-forward neural network with one hidden layer and multiple outputs

In previous work [8, 19, 20, 21] we presented our ideas about method and new architectures of ANN. In the next, we will first briefly introduce the feed-forward neural networks that will be used as a basic structure for prediction throughout this paper.

In prediction of time series, in our case, a set of observables (samples) is given, meaning that only one input signal is available being the discretized time [23]. To get the average monthly/weekly consumption we averaged the data for every month/week of the year/month. According to (1) we are predicting one quantity at a time meaning one output is needed, too. The values of the output are numbers (average power for a period of one month, week). To make the forecasting problem numerically feasible we performed transformation in both the time variable and the response. The time was reduced by t_0 so that

$$t=t^{*}-t_0. \tag{2}$$

Having in mind that t^* stands for the time (in month, weeks), this reduction gives the value of 0 to the time (t_0) related to the first sample. The samples are normalized in the following way

$$y=y^{*}-M \tag{3}$$

where y^* stands for the current value of the target function, M is a constant (for example: M=596.859/595.19, being the average monthly/weekly consumption for a year/month).

If the architecture depicted in Fig. 1 was to be implemented (with one input and one output terminal) the following series would be learned: $(t_i, f(t_i))$, i=1,2,....

Starting with the basic structure of Fig. 1, in [16] possible solutions were investigated and two new architectures were suggested to be the most convenient for the solution of the forecasting problem based on short prediction base period. Here, however, having in mind the availability of data related to previous year/month, these architectures will be properly accommodated.

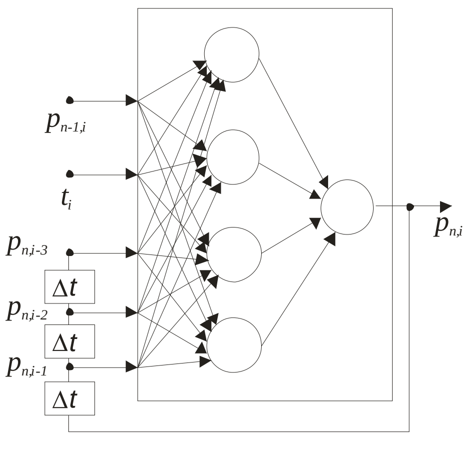

Fig. 2. ETCR. Extended time controlled recurrent ANN derivated from TCR [16]

We presented the first one, named Extended Time Controlled Recurrent (ETCR) architecture. It is made from the TCR ANN on the way described in [16]. This structure is depicted in Fig. 2.

Here in fact, the network is learning a set in which the output value representing the average power consumption for a given month/week in a given year/month is controlled by the present time and by its own previous instances:

$$p_{n,i} = f(t_i, p_{n,i-1}, p_{n,i-2}, p_{n,i-3}, p_{n-1,i}) \quad i=1,2,3... \tag{4}$$

Here n stand for the number of the month/week in the year/month. In that way the values indexed with n are from the actual year/month, while the value indexed n-1 is from the previous year/month. i stands for the i-th sample in the year/month selected. The actual value $p_{n,i}$ is unknown and should be predicted. Incrementing i, in fact, means moving the prediction window one step ahead.

The second structure was named Extended Feed Forward Accommodated for Prediction (EFFAP) and depicted in Fig. 3. We extended the FFAP architecture exactly in the same way as we did with the TCR [16]. Our idea was here to force the neural network to learn the same mapping several times simultaneously but shifted in time. In that way, we suppose, the previous responses of the function will have larger influence on the $f(t)$ mapping.

In that way for the approximation function we may write the following

$$\{p_{n,i+1}, p_{n,i}, p_{n,i-1}, p_{n,i-2}\} = f(t_i, p_{n-1,i}\} \quad i=1.2.3... \tag{5}$$

The new network is approximating the future (unknown) value $p_{n,i+1}$, based on the actual time t_i, the actual consumption $p_{n,i}$, the past consumption values for the given year/month ($p_{n,i-k}$, k=1,2,3), and the past consumption values for the same month/week at the actual time of the previous year/month ($p_{n-1,i}$).

In the next the procedure of implementation of ETCR and EFFAP network will be described.

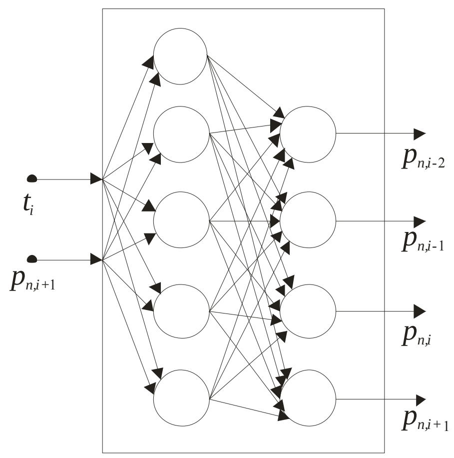

Fig. 3. EFFAP. Extended feed forward accommodated for prediction ANN derivated from FFAP [16] according to (5)

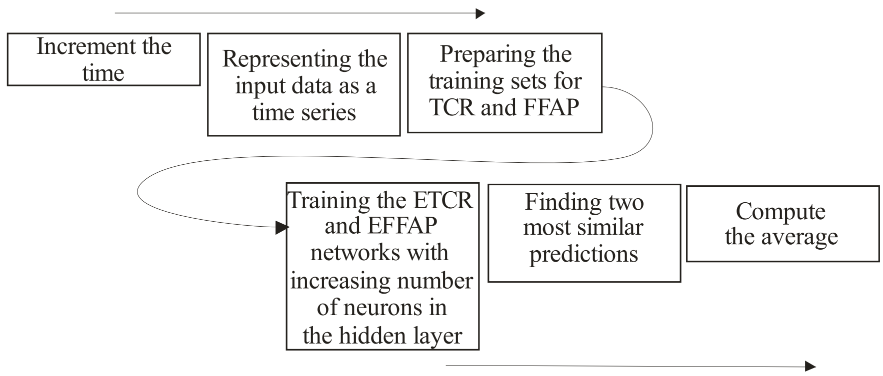

Fig. 4. Steps in obtaining a prediction

The new procedure we are promoting in [21] is depicted in Fig. 4. We start with a time series obtained from [23]. Then we arrange the training sets in two ways appropriate to the two ANN structures we use.

For both EFFAP and ETCR we make eight predictions with eight networks with rising number of hidden neurons starting with 3 and ending with 10. In that way we obtain two vectors of predictions; one for the EFFAP and the other for the ETCR ANN.

In the next step we search the two vectors for the most similar prediction that is for predictions which support each other. These are picked from their vectors as final ETCR and final EFFAP prediction. The process ends by adopting the final prediction obtained as an average of the above two. Namely, if the two predictions are supporting each other they are of equal importance while none may be qualified as the better one. So, the average is the best representative.

IV. Data and Errors

In our work we used data from the UNITE competition [23]. Since there are data for two years only, we created 24 instances for monthly and 101 instances for weekly consumption. Having in mind, however that our method asks for a value of the load for the same month in the previous year, the first 12 instances are to be reserved. Furthermore, to start the prediction we need some values of the previous months. For these reasons we started the prediction with the fourth part of the data i.e. from the 19th month. The weekly prediction started at the end of the first year (last week of December) which, as will be discovered later is of importance for the prediction results.

TABLE I
THE MOST SIMILAR ETCR AND EFFAP SOLUTIONS ON RESTORED ORIGINAL INPUT DATA FOR MONTHLY PREDICTION

t_n	ECTR		EFFAP		Average (p)	Expected
	m_h	p	m_h	p		
19	3	501,358	4	512,897	507,128	510,2098
20	9	502,551	3	493,109	497,830	496,3975
21	7	511,576	7	563,501	537,539	536,7299
22	4	598,954	7	614,244	606,600	617,8031
23	4	684,131	7	681,784	682,957	680,6868
24	4	697,091	3	719,561	708,326	717,1586

In order to get even better insight into the results, the prediction error was calculated and depicted in Table II. As can be seen the error of the average value compared with the expected one is less than 2% in all six cases. A graphical representation of Table I is given in Fig. 5.

TABLE II
PREDICTION ERROR FOR MONTHLY PREDICTION

t_n	Error (%) ECTR	Error (%) EFFAP	Error (%) Average
19	1,735	-0,5267	0,604
20	-1,240	0,6625	-0,289
21	4,687	-4,988	-0,151
22	3,051	0,576	1,813
23	-0,506	-0,161	-0,334
24	2,798	-0,335	1,232

It is interesting to note that the prediction errors of the ETCR and the EFFAP ANNs are much larger (less than 6%). That means that the worst prediction would never exceed that value. By good luck, however, in this case, cancellation occurred during the computation of the average which led to an extraordinary good result.

TABLE III
THE MOST SIMILAR ETCR AND EFFAP SOLUTIONS ON RESTORED ORIGINAL INPUT DATA FOR WEEKLY PREDICTION

t_n	ECTR		EFFAP		Average p	Expected p
	m_h	p	m_h	p		
51	5	746.759	5	736.506	741.633	615.027
52	7	662.406	8	663.523	662.964	647.869
53	3	579.127	9	706.465	642.796	661.6578
54	9	740.493	8	635.385	687.939	683.78
55	10	675.972	5	668.981	672.477	696.83
56	5	697.742	8	698.717	698.23	726.75
57	9	761.235	10	762.086	761.66	726.583
58	6	716.076	6	719.692	717.884	690.366
59	6	670.976	4	687.522	679.249	668.848
60	4	662.313	6	663.963	663.138	649.366

Here, we made a similar table as for monthly prediction, now, for weekly prediction and we follow the algorithm. So, for example, for the 54th week the prediction of the ETCR ANN built by nine hidden neurons and the EFFAP ANN built by eight neurons were the most similar ones. These two were chosen and the average calculated.

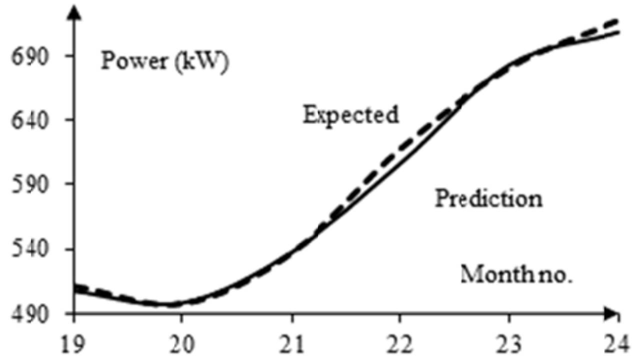

Fig. 5. Visualization of the last two columns of Table I

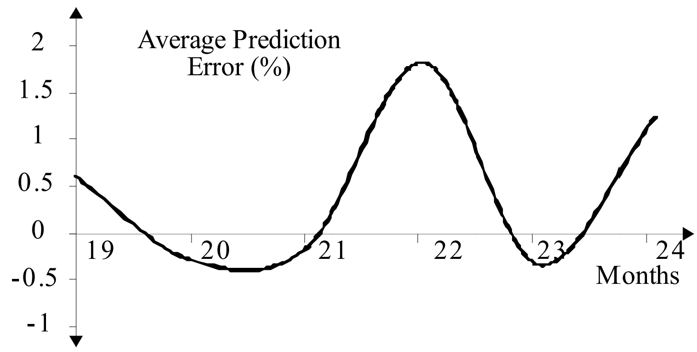

Fig. 6. Monthly prediction error (in %) of the averaged solution

TABLE IV
PREDICTION ERROR FOR WEEKLY PREDICTION

t_i	Error (%) ECTR	Error (%) EFFAP	Error (%) Average
51	-21.4	-19.8	-20.6
52	-2.24	-2.42	-2.33
53	12.5	-6.77	2.85
54	-8.3	7.08	-0.608
55	3.0	4.0	3.5
56	4.0	3.86	3.92
57	-4.77	-4.89	-4.83
58	-3.72	-4.25	-3.99
59	-7.06	-7.60	-7.33
60	-1.99	-2.25	-2.12

Note, to complete the prediction the values produced by (3) were to be restored. That practicaly meant that all entries of Table I and Table III were obtained by incrementation M, earlier depicted. Fig. 5 depicts the two last columns of Table I. Namely the expected and the predicted values are drawn together. Fig. 7 depicts the two last columns of Table III.

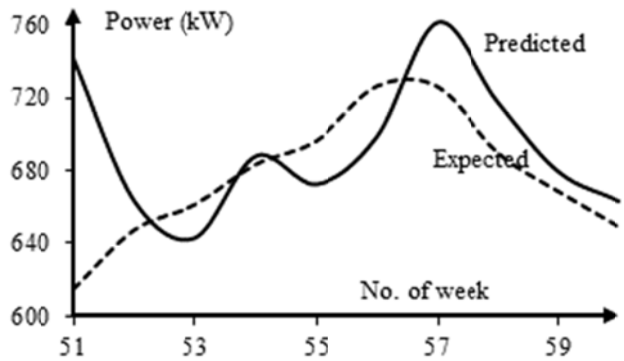

Fig. 7. Visualization of the last two columns of Table III

Finally, in order to get even better insight into the results, the weekly/monhly prediction error was calculated and depicted in Table II/TableIV. A graphical representation of TableII/Table IV (averaged prediction error for monthly and weekly is given in Fig. 6 and Fig.8, respctively).

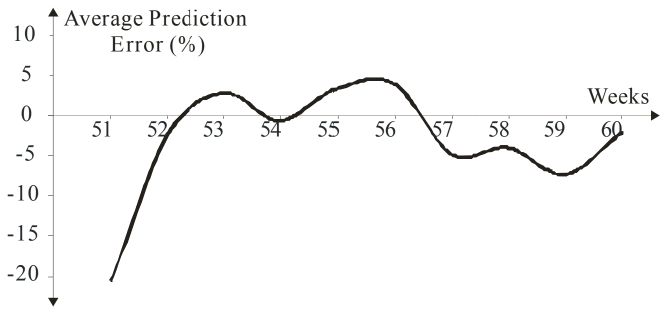

Fig. 8. Weekly prediction error (in %) of the averaged solution

It is easy to recognize that after escaping from the “fatal” last week of the year, the prediction goes smoothly with prediction error no larger than 8%.

IV. CONCLUSION

Results obtained by a new method related to one week and month ahead prediction of suburban average electricity load, based on short time series, were presented and compared. It was shown first that for the subject of short term prediction of electricity load, even though a large amount of data may be available, only the most recent of it may be of importance. That gives rise to prediction based on limited amount of data. We here proposed implementation of some instances of architectures of artificial neural networks as potential systematic solution of that problem as opposed to heuristics that are in use. To further rise the dependability of the predicted data averaging of two independent predictions was proposed. A specific approach to the choice of the number of hidden neurons was implemented. Example was given related to monthly and weekly forecasting of the electricity load at suburban level. Prediction was carried out on real data taken the literature. Acceptable prediction errors were obtained.

ACKNOWLEDGMENT

This research was partially supported by the Ministry of Education Science and technological Development of Serbia within the project TR32004.

REFERENCES

[1] S. Tzafestas, "Computational Intelligence Techniques for Short-Term Electric Load Forecasting", J. of Intelligent and Robotic Systems, Vol. 31, No. 1-3, 2001, pp.7-68.

[2] A. R., Sarode, M. R., Salodkarm, "Short-Term Load Forecasting Using ANN Technique", International J. of Eng. Sciences & Emerging Technologies, Vol. 5, Issue 10, 2015, pp. 08-12.

[3] A. S. Mandel', "Method of Analogs in Prediction of Short Time Series: An Expert-statistical Approach", *Automation and Remote Control*, Vol. 65, No. 4, 2004, pp. 634-641.

[4] P. Murto, "Neural Network Models for Short -Term Load Fore-casting", MS Thesis, Helsinki University of Technology, 1998.

[5] F., Cavallaro, "Electric load analysis using an artificial neural network", *Int. J. of Energy Research*, Vo l. 29, 2005, pp. 377–392.

[6] H., Hahn, S., Meyer-Nieberg, and S., Pickl, "Electric load fore-casting methods: Tools for decision making", *European J. of Operational Research*, Elsevier, Vol. 199, 2009, pp. 902–907.

[7] J. Milojković, V. B. Litovski, "New methods of prediction implemented for sustainable development", Proc. of the 51th Conf. ETRAN, Herceg Novi, Monte Negro, June 2007, Paper no. EL1.8 (in Serbian).

[8] H. A., Malki, N. B., Karayiannis and M., Balasubramanian, "Short-term electric powerload forecasting using feedforward neural networks", *Expert Systems,* Vol. 21, No. 3, 2004, pp. 157-167.

[9] H.M. AI-Hamadi, and S.A. Soliman, "Short-term electric load forecasting based on Kalman filtering algorithm with movingwindow weather and load model", Electric Power Systems Research, Vol. 68, No. 1, 2004, pp. 47-59.

[10] N. Amjady, "Short term hourly load forecasting using time-series modeling withpeak load estimation capability". *IEEE Trans. on Power Systems*, Vol. 16, No. 3, 2001, pp. 498–505.

[11] H. Seetha, and R. Saravanan, "Short Term Electric Load Prediction Using Fuzzy BP", *J. of Computing and Information Technology – CIT,* Vol. 15, No. 3, 2007, pp. 267–282.

[12] E.A. Plummer, "*Time series forecasting with feed-forward neural networks: guidelines and limitations*", M.S. Thesis, University of Wyoming, Laramie, 2000.

[13] R. M. Khan, and A. Abraham, "Short Term Load Forecasting Models in Czech Republic Using Soft Computing Paradigms", *Int. J. of Knowledge-Based Intelligent Engineering Systems,* Vol. 7, No. 4, 2003, pp. 172-179.

[14] J. E. P., Box, and G., Jenkins, "*Time Series Analysis, Forecasting and Control*"**,** Holden-Day, San Francisco, CA, 1990.

[15] D.C., Montgomery, C.L., Jennings, and M., Kulahci, "*Introduction to Time Series Analysis and Forecasting*", Wiley, Hoboken, NJ, 2008.

[16] J. Milojković, and V. B. Litovski, "Comparison of some ANN based forecasting methods implemented on short time series", Proc. of the 9th Symp. NEUREL-2008, Belgrade, ISBN 978-1- 4244-2903-5, Sept. 2008, pp. 175-178.

[17] J. Milojković, a ndV. B. Litovski, "Short term forecasting in Electronics", *Int. J. of Electronics*, Vol. 98, No. 2, 2011, pp. 161-172.

[18] J. Milojković, V. B. Litovski, O., Nieto-Taladriz, and S., Bojanić, "Forecasting Based on Short Time Series Using ANNs and Grey Theory – Some Basic Comparisons"**,** In Proc. of the 11th Int. Work-Conf. on Artificial Neural Networks, IWANN 2011, June 2011, Torremolinos-Málaga (Spain). J. Cabestany, I. Rojas, and G. Joya (Eds.): Part I, LNCS 6691, pp. 183–190, 2011, © Springer-Verlag, Berlin, Heidelberg.

[19] J. Milojković, and V. B. Litovski, „Dynamic Short-Term Fore-casting Of Electricity Load Using Feed-Forward ANNs", *Int. J. of Engineering Intelligent Systems for Electrical Engineering and Communication*, Vol. 17, No. 1, March 2009, pp. 38-48.

[20] J. Milojković, and V. B. Litovski, "Short -term Forecasting of Electricity Load Using Recurrent ANNs", "*Electronics*", Vol. 14, No. 1, June 2010, pp. 45-49.

[21] J. Milojković, and V. B. Litovski, "One month Ahead Prediction of Suburban Average Electricity Load", Proc. of the second Conf. IcETRAN, Srebrno Jezero, Serbia, June 2015, Paper no. EL I2.2.

[22] E.A., Plummer, "Time series forecasting with feed-forward neural networks: guidelines and limitations", M.S. Thesis, University of Wyoming, Laramie, USA, July 2000.

[23] -,World-wide competition within the EUNITE network, http://neuron.tuke.sk/competition.