Zbornik radova 56. Konferencije za ETRAN, Zlatibor, 11-14. juna 2012.
Proc. 56th ETRAN Conference, Zlatibor, June 11-14, 2012

ONE STEP AHEAD DAILY ELECTRICITY PEAK LOAD PREDICTION BASED ON ANN

Jelena Milojković, Faculty of Electronic Engineering, University of Niš, Niš, Serbia, jelena@venus.elfak.ni.ac.rs,
Vančo Litovski, Faculty of Electronic Engineering, University of Niš, Niš, Serbia, vanco@elfak.ni.ac.rs

Abstract *— One step ahead prediction of peak electricity loads is presented based on ANN. Two architectures of ANN were implemented to produce predictions that were used to generate the final value as an average. The time instants when daily peak loads occur are produced simultaneously.*

1. INTRODUCTION

The necessity of load forecasting is nowadays broadly recognized. Precise load forecasting helps the electric utility to make unit commitment decisions, reduce spinning reserve capacity and schedule device maintenance plan properly. In addition, corrective actions may be prepared, such as to avoid load shedding, planning power purchases and bringing peaking units on line. Especially, as noted in [1] accurate short-term forecasts are needed by both generators and consumers of electricity particularly during periods of abnormal peak load demand.

The electricity forecasting period may span from several tenths of minutes to several years so very short term (at tenths of minutes level), short term (hourly), daily, weekly, monthly and yearly load forecasts may be encountered. The proceedings presented here are based on our previous results related to short term prediction [2,3,4]. Here we implement similar methods to generate the forecast of a daily peak value for a given load. In addition, the time of the peak will be predicted. Data were extracted from the 1999 UNITE competition [5].

Our method is based on several hypotheses. First we claim that main influence to the future value may have the most recent observables. These contain most recent information on trend, season and weather. Second, we believe that if quality forecast is to be obtained, one that may be used for action, one is not supposed to try many time steps in advance. One or two time steps are the best one can afford. That is why the time series is presented here as deterministic and one-step-ahead prediction is planned. To help the prediction, however, in an appropriate way, we introduce past values e.g. loads for the same day but in previous weeks. That is in accordance with existing experience claiming that every day in the week has its own general consumption profile [6].

In many load forecasting procedures weather data are used as basic input together with the load time series. We here have a specific opinion about the use of weather data. First, as can be seen from experiments [7] it is not easy to establish a significant correlation between the weather parameters and the peak load value. Second, for the prediction instant no weather data are available. These are to be generated by prediction with equal uncertainty as the main prediction. Finally, the known load values already contain information on the weather if any correlation exists.

As an example of the uncertainty of long term weather prediction based on abundant amount of data let us consider the day of December 19, 2011 which is celebrated as the St. Nicolas by the Orthodox Church exercising the Julian calendar. It is the most celebrated Serbian “Slava” and while “half of the people celebrate the other half is visiting the families that celebrate”. There is always snow at St. Nicolas. There are even proverbs related to the snow at St. Nicolas. On the last day, however, there was no snow and the temperature was above zero all day. Nobody could predict that state one month earlier at November 19, 2011, while everyone could do that at December 18, 2011.

The problem of daily peak load forecasting was considered many times in the literature [8,9,10]. As can be seen statistical methods are used and no time of the maximum was considered.

One of the approaches to load forecast is implementation of artificial neural networks (ANN) [11,12]. The main advantage of the method is related to the property of the ANN to be an universal approximator meaning the main problem of regression: the choice of the approximating function, is solved in advance. A common feature, however, of the existing application is that they ask for a relatively long time series to become effective. Typically it should be not shorter then 50 data points [11].

Following these considerations new forecasting architectures were developed [2,3,4]. Namely, prediction is an activity that is always related to uncertainty. One is supposed to have at least two solutions for them to support each other. 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 [2] as well as in other areas including hourly [3,4] and yearly [8] load prediction.

Here we present extensions of the TCR and FFAP ANNs that allow for implementation in daily peak load prediction together with the prediction of the time of maximum occurance.

The structure of the paper is as follows. After general definitions and statement of the problem we will give a short description of the two solutions. After presenting the experimental results a short discussion of the results and consideration related to future work will be given.

2. PROBLEM FORMULATION

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 $\hat{y} = \hat{f}(t)$, 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 f that will perform the mapping

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

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

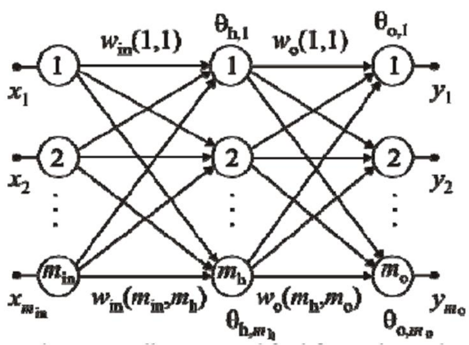

Figure 1. Fully connected feed-forward artificial neural network with one hidden layer and multiple outputs

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.

The network is depicted in Fig. 1. It has only one hidden layer, which has been proven sufficient for this kind of problem [14]. Indices: "in", "h", and "o", in this figure, stand for input, hidden, and output, respectively. For the set of weights, $w(k, l)$, connecting the input and the hidden layer we have: k=1,2,..., m_{in}, l=1,2,..., m_h, while for the set connecting the hidden and output layer we have: k=1,2, ...m_h, l=1,2,..., m_o. The thresholds are here denoted as $\theta_{x,r}$, r=1,2, …, m_h or m_o, with x standing for "h" or "o", depending on the layer. The neurons in the input layer are simply distributing the signals, while those in the hidden layer are activated by a sigmoidal (logistic) function. Finally, the neurons in the output layer are activated by a linear function. The learning algorithm used for training is a version of the steepest-descent minimization algorithm [15]. The number of hidden neurons, m_h, is of main concern. To get it we applied a procedure that is based on proceedings given in [16].

In prediction of time series, in our case, a set of observables (samples) is extracted (one peak value per day) from the UNITE 1997 file. According to (1) we are predicting one quantity at a time. To make the forecasting problem numerically feasible we performed a small transformation in the response. Namely the samples are reduced in the following way

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

where y^* stands for the measured value of the target function, M is a constant (here M=600 MW).

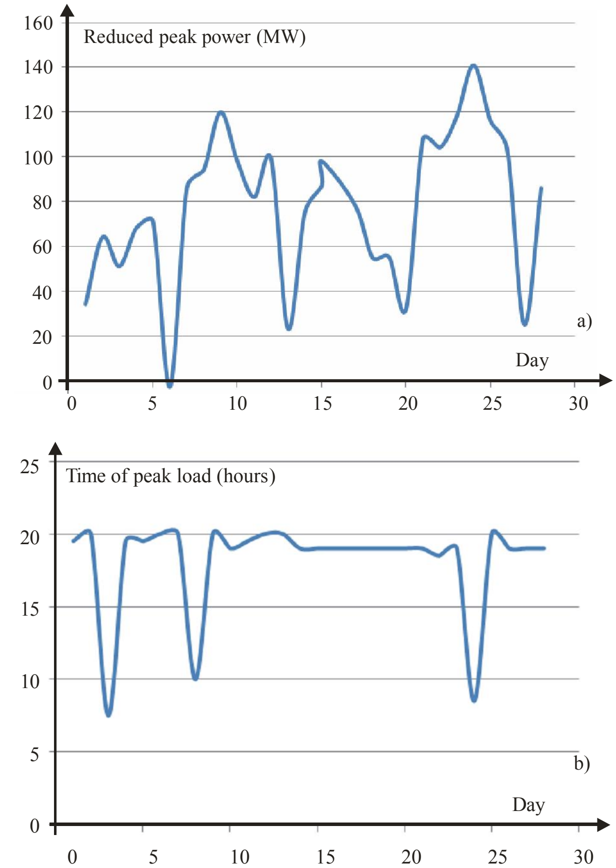

Figure 2. a) Load peak values for 28 days (reduced by 600 MW), and b) time of occurrence of the peak values

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,....

The observables are depicted in Fig. 2. The top part represents the daily peak values in the period September 25, to October 22, 1997. The bottom one represents the corresponding times at which the peak values occur.

3. THE ETCR SOLUTION

Starting with the basic structure of Fig. 1, in [2,3,4] 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.

The first one, named *extended time controlled reccurent* (ETCR) was inspired by the time delayed recurrent ANN. It is a recurrent architecture with the time as input variable so controlling the predicted value. Our intention was to benefit from both: the generalization property of the ANNs and the success of the recurrent architecture. Its structure is depicted in Fig. 3. In this figure i stands for the sample counter and in fact represents the time variable i.e. the day. t_i stands for the daily peak value time while y_i is the daily peak value. Here in fact, the network is learning two sets of variables. The first is the output value representing the daily peak power consumption for the next day is controlled by the present time (variable i) and by its own previous instances. The second is the daily peak value time which is controlled by the same data:

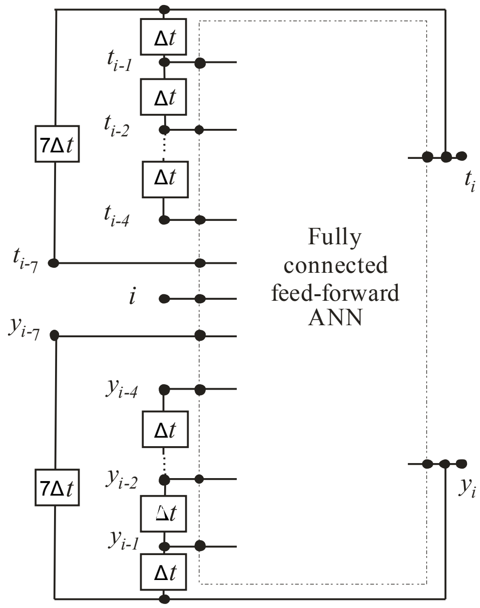

Figure 3. ETCR. Extended time controlled recurrent ANN

Table 1. Training data for the ETCR ANN

i	y_{i-1}	y_{i-2}	y_{i-3}	y_{i-4}	y_{i-7}	t_{i-1}	t_{i-2}	t_{i-3}	t_{i-4}	t_{i-7}	y_i	t_i
8	86	-3	71	68	34	40	40	39	39	39	94	20
9	94	86	-3	71	64	20	40	40	39	40	120	40
10	120	94	86	-3	51	40	20	40	40	15	98	38
11	98	120	94	86	68	38	40	20	40	39	82	39
12	82	98	120	94	71	39	38	40	20	39	99	40
13	99	82	98	120	-3	40	39	38	40	40	23	40
14	23	99	82	98	86	40	40	39	38	40	75	38
15	75	23	99	82	94	38	40	40	39	20	87	38
16	87	75	23	99	120	38	38	40	40	40	98	38
17	98	87	75	23	98	38	38	38	40	38	78	38
18	78	98	87	75	82	38	38	38	38	39	55	38
19	55	78	98	87	99	38	38	38	38	40	55	38
20	55	55	78	98	23	38	38	38	38	40	32	38
21	32	55	55	78	75	38	38	38	38	38	108	38
22	108	32	55	55	87	38	38	38	38	38	104	37
23	104	108	32	55	98	37	38	38	38	38	118	38
24	118	104	108	32	78	38	37	38	38	38	141	17
25	141	118	104	108	55	17	38	37	38	38	116	40
26	116	141	118	104	55	40	17	38	37	38	103	38
27	103	116	141	118	32	38	40	17	38	38	25	38
28	25	103	116	141	108	38	38	40	17	38	86	38

$$y_i = f_1(i, y_{i-1}, y_{i-2}, y_{i-3}, y_{i-4}, y_{i-7}, t_{i-1}, t_{i-2}, t_{i-3}, t_{i-4}, t_{i-7}) \quad , i=7, 8,... \qquad (2a)$$

$$t_i = f_2(i, y_{i-1}, y_{i-2}, y_{i-3}, y_{i-4}, y_{i-7}, t_{i-1}, t_{i-2}, t_{i-3}, t_{i-4}, t_{i-7}) \quad , i=7, 8,... \qquad (2b)$$

In these first proceedings we chose four recent samples and one one-week old to control the output. That choice was confirmed by the results obtained so no new attempts were maid to complicate the training set of the ANN.

According to this definition when preparing the training data for the ETCR ANN, sets of vectors were created by extracting data from the original similarly to the time series reconstruction technique that stems from the embedding theorem developed in [17,18]. The ith input training vector would be:

$\mathbf{x}_i=\{i, y_{i-1}, y_{i-2}, y_{i-3}, y_{i-7}, t_{i-1}, t_{i-2}, t_{i-3}, t_{i-7}\}$,

while the corresponding training output vector would be

$\mathbf{z}_i=\{y_i, t_i\}$.

In this proceedings $i\in\{8,28\}$. Namely, 21 training lessons were used. The training data are given in Table 1.

The task was to predict the peak value and its time at October 23.1997. The resulting ANN had 11 input, two output, and 5 hidden neurons. After proper excitation the prediction was z_{29}={665.799, 18.96}. As can be seen from Table 2. both predictions are within the 10% error region.

4. THE EFFAP SOLUTION

The second structure was named *extended feed forward accommodated for prediction* (FFAP) and depicted in Fig. 4. We use the same notation as in Fig. 3. 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. Note that y_{i+1} is learned meaning a set of data shifted by one in time was used in this case.

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

$$\{y_{i+1}, y_i, y_{i-1}, y_{i-2}, y_{i-3}, t_{i+1}, t_i, t_{i-1}, t_{i-2}, t_{i-3}\} = \mathbf{f}(i, y_{i-6}, t_{i-6}) \quad , i=3,4, ... \qquad (3)$$

The new network is approximating the future (unknown) values y_{i+1} and t_{i+1}, based on the actual time i, the actual peak value coordinates (y_i, t_i), three past peak value coordinates (y_{i-k}, t_{i-k} , k=1,2,3), and the past peak value coordinates for the same day of the previous weeks (y_{i-6}, t_{i-6}).

Table 2. Prediction results (NN = number of neurons in the hidden layer, PPV = predicted peak value, EPV expected peak value, PT = predicted time of the peak value, and ET = expected time of the peak value)

Type	NN	PPV	EPV	%	PT	ET	%
ETCR	5	665.799	717.4	-7.193	18.96	20	-5.2
EFFAP	5	771.032	717.4	7.476	16.96	20	-15.2
Average		718.416	717.4	0.14	17.96	20	-10.2

The resulting ANN had 3 input, 10 output, and 5 hidden neurons. After proper excitation the prediction was z_{29}= {771.032, 16.96}. As can be seen from Table 2., now, while the prediction of the peak value is of the same magnitude but with opposite sign compared to the ETCR solution, the prediction of the peak time is slightly worse.

5. SUMMARY AND CONCLUSION

As expected the prediction obtained by application of the two ANN structures differ. These are both necessary, however, in order to mutually support since, in prediction, no other reference is available. The main criterion for acceptance is the mutual similarity of the results produced by different methods. Since none of them may be considered better in advance one is to profit of both by using the average. The averaged prediction values and the corresponding errors are also given in Table 2. Now we finally have z_{29}={718.416, 17.96}. While the peak value is hit almost exactly, the averaged prediction error being 0.14% only, its time of occurrence was missed by approximately 10%. All together an excellent result, above all expectations.

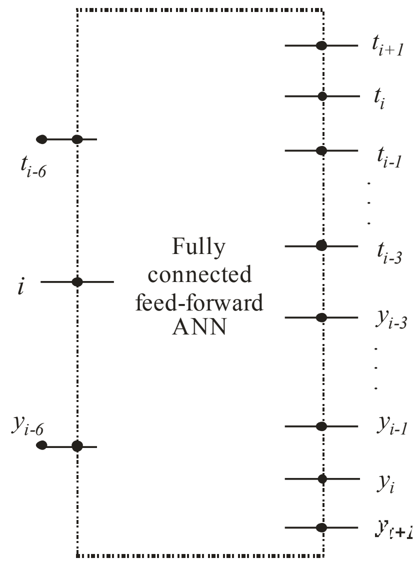

Figure 4. The Extended feed forward accommodated for prediction ANN (EFFAP) according to (3)

Based on the preliminary results reported above we find the method proposed feasible for implementation in short term prediction of daily peak loads with no use of environmental data. Note, to the best of our knowledge, it is the only method reported that predicts the time of occurrence of the peak value. It will be further studied in order to establish as much stability of the prediction as possible. The number of previous days and weeks used for prediction will be considered. Multistep ahead prediction will be searched, too.

ACKNOWLEDGEMENT

This research was partially funded by The Ministry of Education and Science of Republic of Serbia under contract No. TR32004

REFERENCES

[1] Taylor, J.W., “An evaluation of methods for very short-term load forecasting using minute-by-minute British data”, *International Journal of Forecasting (Energy Forecasting)*, Vol. 24, No. 4, 2008, pp.645–658.

[2] Milojković, J.B., Litovski, V.B., “Short term forecasting in Electronics”, *International Journal of Electronics*, ISSN: 0020-7217, Vol. 98, No. 2, 2011, pp. 161-172.

[3] Milojković, J, and Litovski, V.B., „Dynamic Short-Term Forecasting of Electricity Load Using Feed-Forward ANNs“, *Int. J. of Engineering Intelligent Systems for Electrical Engineering and Communication*, ISSN 1472-8915, Vol. 17, No. 1, March 2009, pp. 38-48.

[4] Milojković, J., and Litovski, V., „Dynamic One Step Ahead Prediction of Electricity Loads at Suburban Level”, *Proc. of the First IEEE Int. Workshop on Smartgrid Modeling and Simulation – at IEEE SmartGridComm 2011*, SGMS2011, Brussels, October 2011, Proc. on disc, paper no. 25.

[5] World wide competition within the eunite network. (2001). [Online] Available: http://neuron.tuke.sk/competition.

[6] Murto P., ‘*Neural Network Models for Short-Term Load Forecasting*”, MS Thesis, Helsinki University of Technology, 1998.

[7] Hyndman, R.J., and Fan, S., "*Forecasting long-term peak half-hourly electricity demand for South Australia*", Report for Electricity Supply Industry Planning Council (South Australia, Monash University, Australia, 2009.)

[8] Elattar, E.E., Goulermas, J., and Wu, Q. H., “Electric Load Forecasting Based on Locally Weighted Support Vector Regression”, *IEEE Transactions On Systems, Man, and Cybernetics—Part C: Applications and Reviews*, Vol. 40, No. 4, July 2010, pp. 438-447.

[9] As’ad, M., "Finding the Best ARIMA Model to Forecast Daily Peak Electricity Demand", *Applied Statistics education and Research Collaboration (ASEARC) – Conf. Papers,* Uni. of Wollongong, Australia, 2012, Paper 11.

[10] Sigauke, C., and Chikobvu, D., “Prediction of daily peak electricity demand in South Africa using volatility forecasting models”, *Energy Economics*, Elsevier, Vol. 33, No. 5, Sept. 2011, pp. 882-888.

[11] Zhang, B.G., “Forecasting with artificial neural networks: The state of the art”, *Int. J. of Forecasting*, Vol. 14, No. 1, March 1998, pp. 35-62.

[12] Hippert, H. S., Pedreira, C. E., and Souza, R. C., “Neural networks for short-term load forecasting: A review and evaluation,” *IEEE Transactions on Power Systems*, vol. 16, pp. 44–55, February 2001.

[13] Milojković, J., Litovski, V., Nieto-Taladriz, O., and Bojanić, S., “Forecasting Based on Short Time Series Using ANNs and Grey Theory – Some Basic Comparisons”, *In J. Cabestany, I. Rojas, and G. Joya (Eds.): Part I, LNCS 6691*, pp. 183–190, 2011, © Springer-Verlag, Berlin, Heidelberg.

[14] Masters, T., "*Practical Neural Network Recipes in* C++", Academic Press, San Diego, 1993.

[15] Zografski, Z., “A novel machine learning algorithm and its use in modeling and simulation of dynamical systems”, *Proc. of 5th Annual European Computer Con.*, COMPEURO '91, Hamburg, Germany, 1991, pp. 860-864.

[16] Baum, E.B., and Haussler, D., “What size net gives valid generalization”, *Neural Computing*, 1989, Vol. 1, pp. 151-160.

[17] Takens, F., “Detecting strange attractors in turbulence,” *Lecture Notes in Mathematics* (Springer Berlin), Vol. 898, 1981, pp. 366–381.

[18] Sauer, T., Yorke, J. A., and Casdagli, M., “Embedology,” *Journal of Statistical Physics*, Vol. 65, May 1991, pp. 579–616.