Journal of Artificial Intelligence in Electrical Engineering, Vol. 3, No. 9, June 2014

Efficient Short-Term Electricity Load Forecasting Using Recurrent Neural Networks

Vahid Mansouri [1], Mohammad E. Akbari [2]
[1]Department of Electrical Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran
Email: vahidmansouri2010@gmail.com (Corresponding author)

[2] Department of Electrical Engineering, Ahar Branch, Islamic Azad University, Ahar, Ira.

ABSTRACT

Short term load forecasting (STLF) plays an important role in the economic and reliable operation of power systems. Electric load demand has a complex profile with many multivariable and nonlinear dependencies. In this study, recurrent neural network (RNN) architecture is presented for STLF. The proposed model is capable of forecasting next 24-hour load profile. The main feature in this network is internal feedback to highlight the effect of past load data for efficient load forecasting results. Testing results on the three year demand profile shows higher performance with respect to common feed forward back propagation architecture.

KEYWORDS: Short term load forecasting (STLF) , Recurrent neural network (RNN) , hourly load forecast, Load Data normalization.

1. INTRODUCTION

Load forecasting is necessary for the proper scheduling activities and corrects operation of electric utilities. Based on the interval of time period, load forecasting can be separated into three categories: short-term load forecasting (STLF) which is usually from one hour to one week, medium term load forecasting (MTLF) , typically from one week to one year, and long-term load forecasting (LTLF) which is longer than one year [1,2]. Each type has an important role to play in the financial and reliable operation of electric utilities. STLF is essential for tasks like real time generation control, security analysis, dis-patch and short-term or spot trading and energy operation planning [2].

Different models have been hired in power systems for realizing forecasting accuracy and these include: regression, statistical and state space methods [3-6]. Artificial intelligence based approaches have been explained based on expert systems, evolutionary programming, fuzzy systems, artificial neural networks and various arrangements of these. The widely used approach in previous works has been based on designing a mathematical model of the power system and then to achieve simulations to control required values [7-9]. The main task with this has been to make perfect non-linear mathematical models. Neural networks are an important tool in modeling nonlinear systems. Also, comprehensive system data is not always readily available and the great calculating effort necessary [10]. Complexity of the electrical network and existence of many varying parameters with nonlinear inter-relations make it difficult to solve the problem, optimally through conventional methods.

Recently, ANNs (Artificial Neural Networks) have been applied to a range of problems in the field of power system control including load forecasting [11], [12]. The theory of ANNs is mainly motivated by the establishment of simple formal models of biological neurons and their interconnections called ANNs. In the power engineering domain, the application of already simplified tools of ANNs to technical problems is the main objective. The ANNs are brain-inspired computers which may solve similar tasks as the biological brain. A number of STLF tools have been recently developed using ANNs modeling methods. A class of Radial Basis Function Networks (RBFN) used for STLF [13]. D.Specht used probabilistic neural network by estimating probability density functions (p.d.f) for load forecasting [14]. Weather variables such as temperature used in [15] are either averaged information of forecasted values, there is always an opportunity that the uncertainty could disturb output of load forecasting.

A number of factors have been evaluated to decide about the usefulness of these issues as inputs data for short term assessment. These include weather variables and past demand. As for the past loads, data examination shows that trends are not maintained steadily for all hours of a day, or even for dissimilar week days. Presented in this paper is a method employing artificial neural networks (ANN) to extract trends from past data of the similar parameters and use that to predict probable upcoming values. Since previous load data affect future load prediction, we have used a feedback included in neural networks. The proposed system consists of an artificial intelligence load forecaster according to recurrent neural network which is adopted to consider the past demand data profile using internal feedback connections. The feedback creates dependency on previous data to predicted values inside network structure. In the next section, we describe our proposed recurrent neural networks (RNN) architecture. Section 3 explains load data pre-processing before feeding them as the input of RNN. Section 4 illustrates results of implementing RNN load forecaster and compares it with back-propagation neural networks (BPNN) and finally, section 5 concludes the paper.

2. RECURRENT NEURAL NETWORKS (RNN) STRUCTURE

Since previous load data affects forthcoming values, to guess any parameter one requires having some previous information on aspects that affect that parameter, or trends that refer to the parameter of interest. The architecture or structure of RNN predictor underpins its capacity to represent the dynamic properties of a statistically non stationary discrete time input signal and hence its ability to predict or forecast some future value. The basic building blocks of all discrete time predictors are adders, delayers, multipliers and for nonlinear case zero-memory nonlinear elements. The manner in which these elements are interconnected, describes the architecture of the predictor. Depending on particular task and available data using recurrent neural networks (RNNs) is the appropriate method of forecasting for this case study. RNNs are fundamentally different from feed forward architectures in the sense that they work, in addition to an input space, on an internal state space representing what already has been processed by the network [16]. State space enables illustration of temporally and/or sequentially extended dependencies over time intervals. It is because of using local and global feedback together with delay elements inside their structures. General structure of the network is depicted in Figure 1.

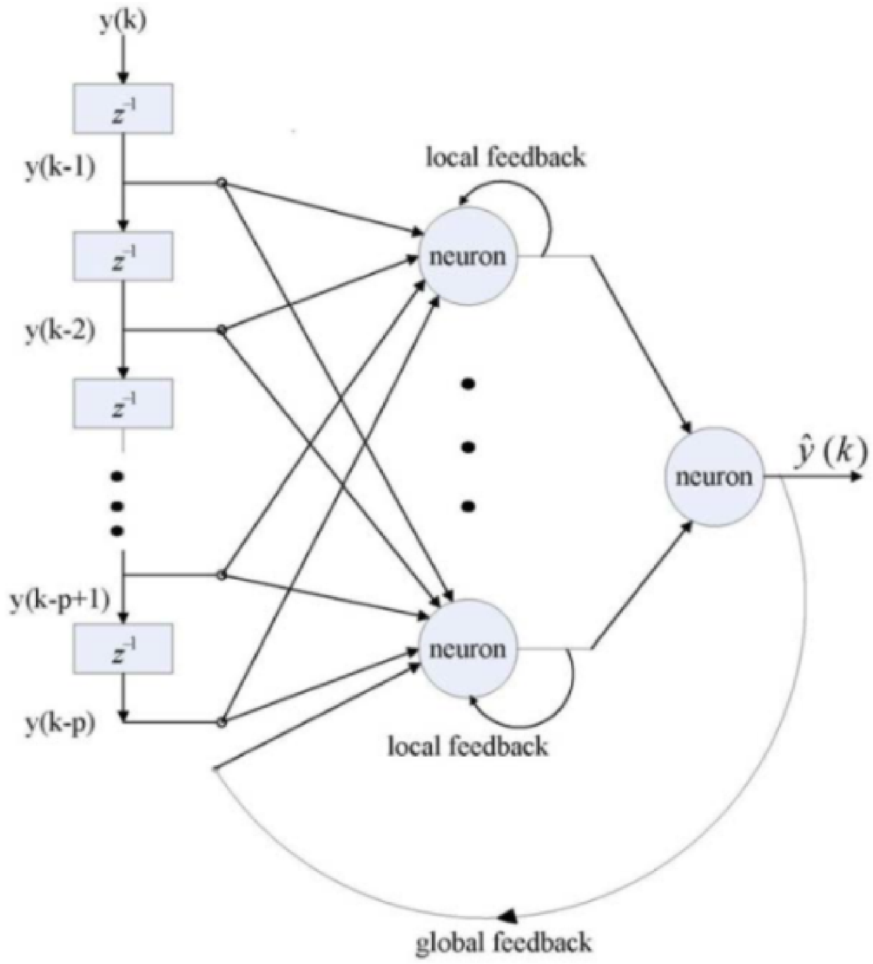

Fig. 1. Structure of a recurrent neural network (RNN) with local and global feedback and delay elements.

The fundamental feature of a Recurrent Neural Network (RNN) is that the network contains at least one feed-back connection, so the activations can flow round in a loop. That permits the networks to do temporal processing and learn sequences, e.g. achieve sequence recognition/reproduction or temporal association/prediction. Recurrent neural network architectures can have many different arrangements, including Elman [17] and Jordan [18] networks. One common type used in this study consists of a standard Multi-Layer Perceptron (MLP) plus added loops. These can exploit the powerful non-linear mapping abilities of the MLP, and also have some form of memory. For simple architectures and deterministic activation functions, learning can be achieved using similar gradient descent procedures to those leading to the back-propagation algorithm for feed-forward networks. The simplest form of fully recurrent neural network is an MLP with the previous set of hidden unit activations feeding back into the network along with the inputs:

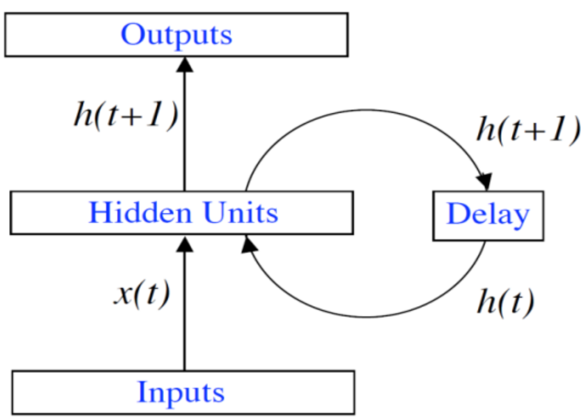

Fig. 2. Structure of the proposed fully recurrent neural network.

The time ***t*** has to be discretized, with the activations updated at each time step. The time scale might correspond to the operation of real neurons, or for artificial systems any time step size appropriate for the given problem can be used. A delay unit needs to be introduced to hold activations until they are processed at the next time step. If the neural network inputs and outputs are the vectors $x(t)$ and $y(t)$, the three connection weight matrices are W_{IH}, W_{HH} and W_{HO}, and the hidden and output unit activation functions are f_H and f_O, the behavior of the recurrent network can be described as a dynamical system by the pair of non-linear matrix equations (1) and (2):

$$h(t+1) = f_H\,(W_{IH}\;x(t)+W_{HH}\,h(t)) \tag{1}$$

$$y(t+1) = f_O\,(W_{HO}(\,h(t+1)) \tag{2}$$

In general, the state of a dynamic system is a set of values that summarizes all the information about the past behavior of the system that is necessary to provide a unique explanation of its future performance, apart from the effect of any external factors. In this case the state is defined by the set of hidden unit activations $h(t)$.

Thus, in addition to the input and output spaces, there is also a state space. The order of the dynamical system is the dimensionality of the state space, the number of hidden units. The recurrent network can be converted into a feed-forward network by unfolding over time as shown in Fig3. This means all the previous theory about feed-forward network learning follows through.

The Back-Propagation Through Time (BPTT) learning algorithm is the natural extension of standard back-propagation that performs gradient descent on a complete unfolded network. If a network training sequence starts at time t_0 and ends at time t_1, the total cost function is simply the sum over time of the standard error function $E_{sse/ce}(t)$ at each time-step(see eq.3):

$$E_{total}(t_0,t_1)=\sum_{t=t_0}^{t_1}E_{sse/ce}(t) \tag{3}$$

and the gradient descent weight updates have contributions from each time-step(see eq.4):

$$\Delta w_{ij}=-\eta\frac{\partial E_{total}(t_0,t_1)}{\partial w_{ij}}=-\eta\sum_{t=t_0}^{t_1}\frac{\partial E_{sse/ce}(t)}{\partial w_{ij}} \tag{4}$$

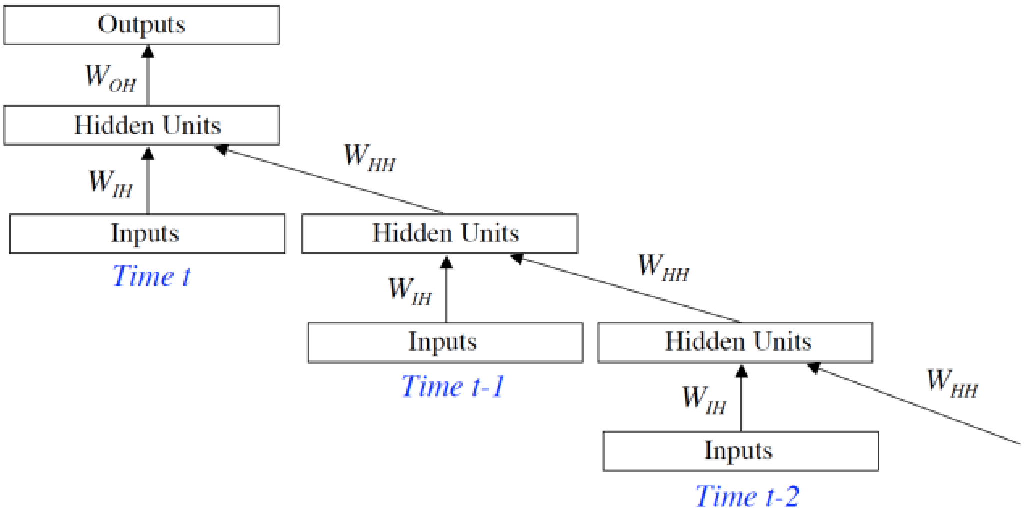

Fig. 3. The recurrent network converted into a feed-forward network by unfolding over time.

The basic partial derivatives $\partial E_{sse/ce}/\partial w_{ij}$ now have contributions from the multiple instances of each weight $w_{ij} \in \{W_{IH}, W_{HH}\}$ and depend on the inputs and hidden unit activations at previous time steps. The errors now have to be back-propagated through time as well as through the network.

2.1 Hourly Load Forecast

This problem involves the prediction of the load for the following hour. The main elements that affect the hourly load prediction are the hour of the day, t, the temperature at that hour, T(t), the load at that hour, L, and the temperature at the following hour, T(t+1). Fig.4. shows the inputs and output for the hourly prediction problem.

First, in order to maintain uniformity, the training data are split up into four categories, as mentioned earlier, based on seasonal and daily variations. In each case, depending on hourly problem forecasting, the inputs and outputs are identified and suitable networks selected. In the case of hourly prediction, the neural network chosen had four inputs and one output. The training and testing process was implemented on a windows seven operating system using MATLAB software on a Laptop l computer with 8GB of RAM and Core2 processor. The proposed RNN is illustrated in Fig.4.

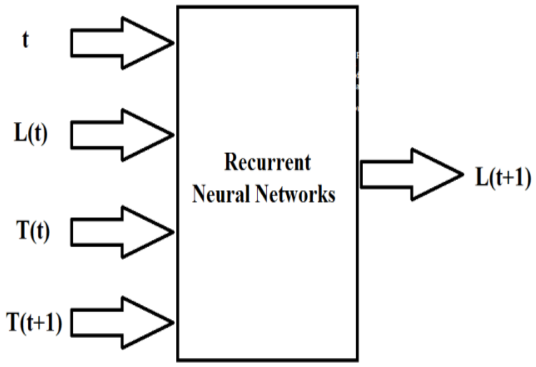

Fig. 4. Hourly load forecasting RNN architecture.

3. Load data pre-processing

The objective of data pre-processing is to offer a valid data set in which all variables are normalized, all corrupted data is either rebuilt or at least recognized and all data is ordered. The output of this pre-processing block is a data matrix in which each row represents the information available from one day, see Fig. 5. Actual load data extracted from [19]. Input data are imported into MATLAB software. The data are then rearranged so that each row contains the hourly load from 6 a.m. to 5 a.m. on the next day. This consideration is taken due to the fact that the load on the period from midnight to 6 a.m. is actually more influenced by the events on the previous day than by those to happen on its actual calendar date. Filtering is intended to eliminate corrupted data. This is done by running a low-pass filter that eliminates sudden changes on the load (approximately more than 50% change on a 2-hour period).

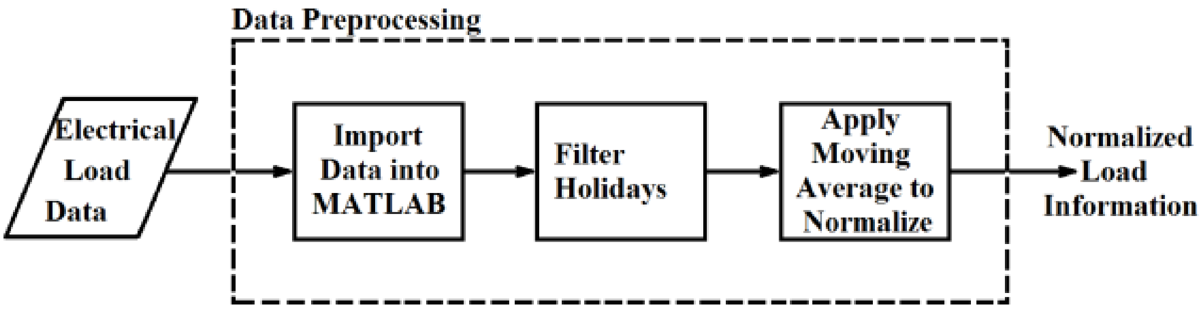

Fig. 5. Block diagram of electrical load data preprocessing.

4. Results and discussion

The hourly load prediction problem is considered in this paper. Fig. 6 shows the hourly load prediction results versus the actual load for the proposed RNN method compared with typical Back-Propagation (BP) neural network in a 24 hour period. Fig.7. illustrates percentage error calculation of forecasted load in Fig.6. using our proposed RNN method and BP architecture in a one day time period.

The load forecast results indicate that the RNN load forecast is more accurate than the BP load forecast, during 24 hours of a day. The average percentage errors for RNN and BP forecasting are 2.38% and 3.18%, respectively. Clearly these errors depend on the inputs, training and test data. The average errors can be minimized further by confirming the following: (a) use of more homogeneous training data; (b) better selection of inputs; (c) larger choice of training set.

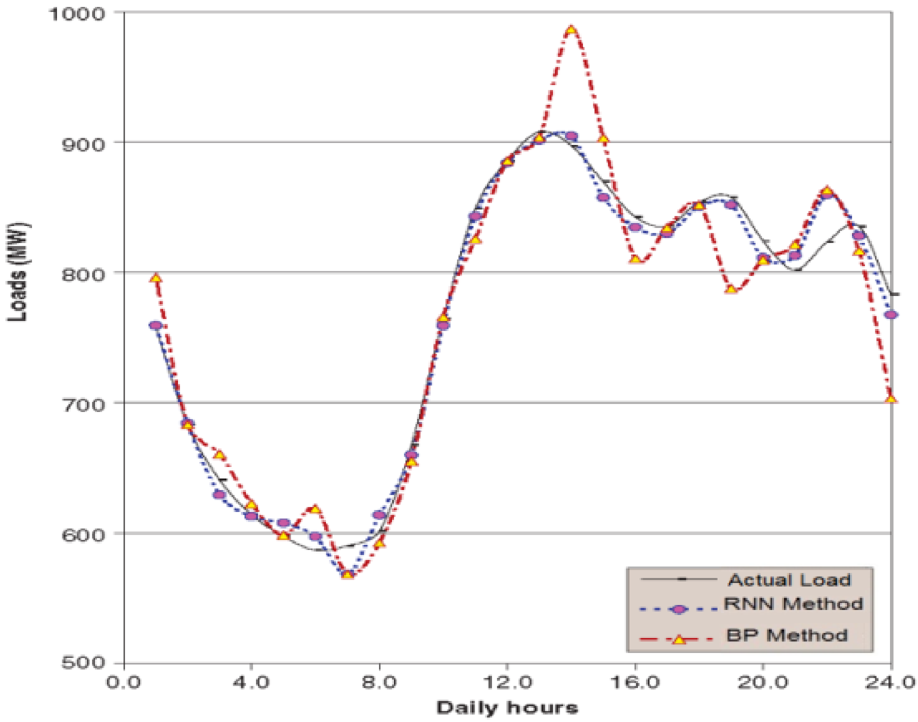

Fig.6. Comparison of hourly load forecasting results using our proposed RNN (Recurrent Neural Network) method and BP(Back Propagation) architecture with actual load in a 24 hour time.

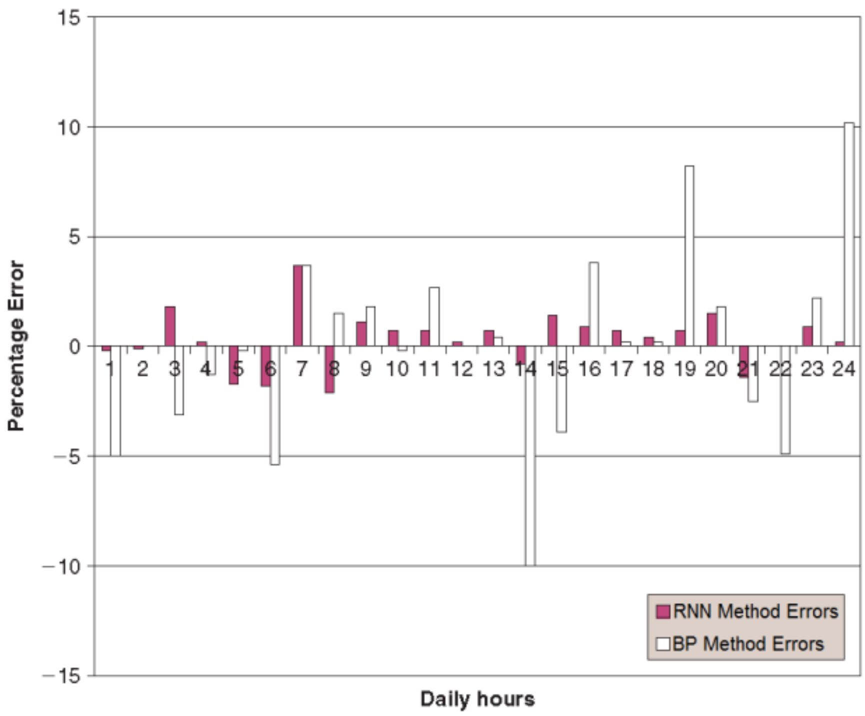

Fig.7. Calculation of hourly percentage error of forecasted load with respect to actual load using two methods, including: our proposed RNN (Recurrent Neural Network) method and BP (Back Propagation)

The results shown here approve that the accuracy of the results depends heavily on the previous load profile and hence proposed RNN architecture generates more accurate and reliable forecasting results by considering the effect of past data inside its feedback structure.

5. CONCLUSIONS

Typical neural networks architecture, as an important tool for non-linear system identification, is already used in short-term load forecasting. In this paper Recurrent Neural Networks (RNN) with internal feedback structure proposed for STLF. Since load profile is highly affected from its past trend profile, RNN used to consider the effect of previous data in the prediction procedure. It is achieved by putting internal feed back inside network structure for hourly load forecast. Also preprocessing is achieved to generate more uniform data after normalization and prevent outliers during training process. Results show higher performance of the proposed architecture with regard to common BP feed forward networks during hourly load forecast.

REFERENCES

[1] Eugene A. Feinberg and Dora Genethliou, "Chapter 12 Load Forecasting" Weather (2006), Issue: August, Publisher: Springer, pp. 269-285

[2] Milos Bozic, Milos Stojanovic and Zoran Stajic, "Short-term electric load forecasting using least square support vector machines" Facta Universitatis, Series: Automatic Control and Robotics Vol. 9, No 1, pp. 141-150,2010

[3] R. C. Garcia, et al., "GARCH Forecasting Model to Pre-dict Day-ahead Electricity Prices," IEEE Transactions on Power Systems, Vol. 20, No. 2, May 2005, pp. 867-874. doi:10.1109/TPWRS.2005.846044

[4] M. Stevenson, "Filtering and Forecasting Spot Electricity Prices in the Increasingly Deregulated Australian Elec-tricity Market," Quantitative Finance Research Centre, University of Technology, Sydney, 2001.

[5] N. Hubele, et al., "Identification of Seasonal Short-term Load Forecasting Models Using Statistical Decision Functions," IEEE Transactions on Power Systems, Vol. 5, No. 1, 1990, pp. 40-5. doi:10.1109/59.49084

[6] M. El-Hawary, et al, "Short-Term Power System Load Forecasting Using the Iteratively Reweighted Least Squares Algorithm," Electrical Power Systems Research, Vol. 19, 1990, pp. 11-22. doi:10.1016/0378-7796(90)900 03-L

[7] V. S. Kodogiannis and E. M. Anagnostakis, "A Study of Advanced Learning Algorithms for Short-term Load Fo-recasting," Engineering Applications of Artificial Intelli-gence , Vol. 12, 1999, pp. 159-173. doi:10.1016/S0952-1976(98)00064-5

[8] G.-C. Liao and T.-P. Tsao, "Application of Fuzzy Neural Networks and Artificial Intelligence for Short-term load Forecasting," Electrical Power Systems Research, Vol. 70, 2004, pp. 237-244. doi:10.1016/j.epsr. 2003.12.012

[9] H. Yamin, M. Shahidehpour and Z. Li, "Adaptive short-term Price Forecasting

using artificial Neural Net-works in the Restructured Power Markets," Electrical Power and Energy Systems, Vol. 26, 2004, pp. 571-581. doi:10.1016/j.ijepes.2004.04.005

[10] A. K. Topalli, I. Erkmen and I. Topalli, "Intelligent Short-term Load Forecasting in Turkey," Electrical Pow-er and Energy Systems, Vol. 28, 2006, pp. 437-447. doi: 10.1016/j.ijepes.2006.02.004

[11] R.C.Bansal, "Overview and Literature Survey of Artificial Neural Networks Applications to Power Systems (1992-2004)",IE Journal, Vol86, March, 2006.

[12] M.Tarafdar Haque, A.M.Kashtiban, "Application of Neural Networks in Power systems: A Review", Proceedings of world Academy of Science and Technology, Vol 6, June 2006

[13] Zbigniew Gontarand Nikos Hatziargyriou, "Short Term load forecasting using Radial basis neural network", SM, IEEE2001 IEEE Porto Power Tech Conference, September, Porto, Portugal.

[14] J Donald. F. Specht, "Probabilistic neural networks for classification mapping, or associative memory", Proc. IEEE Inf. Conf. Neural Networks, San Diego, CA, VOI. 1, pp. 525-532, July 1988.

[15] P. Mandal, T. Senjyu, N. Urasaki and T. Funabashi, "A Neural Network Based Several-Hour-Ahead Electric Load Forecasting using Similar Days Approach," Elec-trical Power and Energy Systems, Vol. 28, 2006, pp. 367-373. doi:10.1016/j.ijepes.2005.12.007.

[16] Bodn, " A guide to recurrent neural networks and back propagation", Report from NUTEK-supported project AIS-8: Application of data analysis with learning systems, 1999-2001, Holst, A.(ed), SICS Technical Report T2002:3, SICS, Kista, Sweden, 2002.

[17] L. Elman, Finding Structure in time, Cognitive science, 14, pp.179-211, 1990.

[18] Jordan, "Attractor Dynamics and parallelism in a connectionist sequential machine, In Proceedings of the Eighth conference of the Cognitive Science Society, pp.531-546, 1986.

[19] ww.energyseec.com/downloadiranedatabank.asp. Last visited Jan. 2015.