ARTICLE IN PRESS

Energy xxx (2016) 1–13

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

# Deep belief network based electricity load forecasting: An analysis of Macedonian case

Aleksandra Dedinec [a, *], Sonja Filiposka [a], Aleksandar Dedinec [b], Ljupco Kocarev [a, c, d]

[a] Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje, Macedonia
[b] Research Center for Energy and Sustainable Development, Macedonian Academy of Sciences and Arts, Skopje, Macedonia
[c] Macedonian Academy of Sciences and Arts, Skopje, Macedonia
[d] Bio Circuit Institute, University of California San Diego, USA

ARTICLE INFO

*Article history:*
Received 30 November 2015
Received in revised form
14 July 2016
Accepted 18 July 2016
Available online xxx

*Keywords:*
Electricity load forecasting
Neural networks
Deep belief networks
Restricted Boltzmann machines
Short-term forecasting
Hourly electricity load

ABSTRACT

A number of recent studies use deep belief networks (DBN) with a great success in various applications such as image classification and speech recognition. In this paper, a DBN made up from multiple layers of restricted Boltzmann machines is used for electricity load forecasting. The layer-by-layer unsupervised training procedure is followed by fine-tuning of the parameters by using a supervised back-propagation training method. Our DBN model was applied to short-term electricity load forecasting based on the Macedonian hourly electricity consumption data in the period 2008–2014. The obtained results are not only compared with the latest actual data, but furthermore, they are compared with the predicted data obtained from a typical feed-forward multi-layer perceptron neural network and with the forecasted data provided by the Macedonian system operator (MEPSO). The comparisons show that the applied model is not only suitable for hourly electricity load forecasting of the Macedonian electric power system, it actually provides superior results than the ones obtained using traditional methods. The mean absolute percentage error (MAPE) is reduced by up to 8.6% when using DBN, compared to the MEPSO data for the 24-h ahead forecasting, and the MAPE for daily peak forecasting is reduced by up to 21%.



## 1. Introduction

The electric power system plays a key role in the economy sector of one country, making its reliability and continuous improvement a high priority on a long term but also on a day-to-day basis. Thus, load forecasting is of crucial importance for proper operation, maintenance and planning of the electric power system. With respect to the time period, there are four categories of load forecasting [1]. The first one is long-term, wherein 1–50 years electricity consumption is forecasted. If the forecast ranges from one month to one year, then it is considered as mid-term forecasting. Short-term forecasting refers to hour, day or week ahead predictions. Finally, usually considered as a separate category, we have the very short-term forecasting which includes few minutes to an hour ahead prediction of the electricity consumption.

Both, long and mid-term forecasts are of great importance for strategic planning of the development of the electric power systems. This includes scheduling of construction of new generation or transmission capacities, maintenance scheduling, as well as long-term demand-side measurement and management planning [1]. These long and mid-term forecasts of the electricity consumption are usually used as an input to more complex models for electric power system planning, such as MARKAL [2–4], EnergyPlan [5,6] and LEAP [7]. When modelling the forecasting of the electricity consumption, the main difference between long-term and short-term predictions is in the input variables that define or influence the consumption. Long-term forecasting may use only historical data for the consumption as an input variable, provided on a larger scale, such as yearly consumption data [8–10]. Gross domestic production (GDP), GDP per capita and population may be used as additional inputs for this forecasting as it is given in Ref. [11]. In Ref. [12] the input variables for energy import and energy export are being used as well. A slightly different approach for selecting

* Corresponding author.
*E-mail addresses:* aleksandra.kanevche@finki.ukim.mk (A. Dedinec), sonja.filiposka@finki.ukim.mk (S. Filiposka), dedinec@manu.edu.mk (A. Dedinec), lkocarev@ucsd.edu (L. Kocarev).

http://dx.doi.org/10.1016/j.energy.2016.07.090
0360-5442/

the input variables is used in Ref. [13] where the set of parameters includes the gross electricity generation, installed capacity, total subscribership and population.

On the other hand, the short-term electricity consumption forecasting is of fundamental importance for unit commitment or optimal generation capacity scheduling, as well as fuel purchase plans, all in order to minimize the financial risk, improve the reliability of the system as well as plan short-term maintenance. Moreover, today with the introduction of smart grids and the rapidly increasing share of renewable energy sources [14] in the electric power systems, the need for electricity consumption forecasting at hourly level becomes even more important, especially from the aspect of demand side management, dynamic integration of renewable energy sources and planning for storage needs.

The results of the hourly electricity consumption forecasting may be further used as an input for electricity price forecasting. Especially now, when the electric power industry is becoming liberalized, these predictions may play an important role in decision making for both the power system operators and the market participants [15]. Hence, in Ref. [16] the day-ahead electricity prices are forecasted using data at hourly level and in Ref. [17] the hourly demand is one of the input variables that are used for day-ahead electricity spot prices.

Electricity consumption forecasting may be applied to data at a country level, such as in Refs. [18,19], or it may be applied to smaller scale such as local areas in a country [20], a region [21,22], a city [23], a campus [24,25], in microgrid environments [26] or a household [27,28], which is one of the fundamental topics in smart grids.

When forecasting time series variables, there are two methods for selection of the input variables: the first one is to use only the previous values of the variable in order to do forecasting, which is referred as time series forecasting, and the second is to also include some other variables that are correlated with the variable that is being forecasted. When forecasting short-term electricity consumption, beside the historical electricity consumption data [22], the meteorological and time label data may also be used. Electricity consumption forecasting by using the time series forecasting approach is implemented in Refs. [29] and [30]. However, in Ref. [31] it is concluded that the temperature input variable greatly influences the electricity consumption in a city in Serbia. So, in Ref. [32] temperature and historical data are used for electricity forecasting. In Ref. [25] alongside the temperature and historic load inputs, exact date with hour and type (workday, weekend or holiday) of day are also used. Similar input variables on a coarser scale are used in Ref. [26] such as: month, day of the week and historical load. In Ref. [19] input variables for wind speed and solar irradiance are also considered.

Electricity load forecasting has proven to be a complex problem, which is non-linear and usually cannot be solved with a simple analytical formulation. Thus, there are many models used to solve this problem. As indicated in [33], the models that are used for load forecasting may be categorized into two groups: statistical and artificial intelligence based models. The statistical based models group includes approaches based on multiple regression, auto regressive (AR), moving average (MA), auto regressive moving average (ARMA), and auto regressive integrated moving average (ARIMA) [34,35]. On the other hand, the artificial intelligence based models are based on expert systems [36], grey systems [8–10], artificial neural networks (ANN), support vector machines [29] and fuzzy logic [25,32]. Currently, the most widely used methods are those based on artificial neural networks [37,38] applied in this area since the 1980s. Neural networks have proven to be suitable for power consumption forecasting since they can be non-linear and can approximate any complex function, provided that a sufficient number of hidden layers and a sufficient number of nodes in the hidden layers are being used, and they have proven to present superior results compared to the other previously mentioned methods [33]. However, there are certain problems such as initialization of the parameters, slow convergence, getting stuck in a bad local minimum and the scalability of the neural networks that are still present and the researchers actively work to overcome these problems.

As a solution to these general problems in Ref. [39] a deep belief network (DBN) is proposed. In this approach, a layer-by-layer unsupervised learning method is used to pre-train the initial values of the weights in the network, after which fine tuning is applied using a standard supervised method. The results from this DBN based approach show that the initial parameters of the network are closer to the optimal solution, compared to the random initialization. These results have paved the way for the successful application of the deep belief networks model in many diverse areas [40]. Today, the DBN model is used mostly for image classification [41,42] and as such it is also used in medicine for automatic detection of breast cancer [43] and selecting risk factors and prediction of osteoporosis [44]. Nonetheless, it is also used in mechanical engineering classification of defects in compressor valves [45] and in finance forecasting of the exchange rate [46]. Also, deep belief networks are used for load forecasting in the smart gas and water grids [47].

The general deep belief network model for time-series prediction presented in Ref. [48] may be used as a good starting point for electricity consumption forecasting. However, very few studies have approached the subject of using deep belief networks in electricity consumption forecasting. To this problem, in Ref. [29] it is approached as a time series forecasting problem, where only 1 h ahead is forecasted, using only the consumption of the previous 24 h. Another example is given in Ref. [49] where, although the model is based on deep belief networks, the lack of detailed input data combined with a very unbalanced training vs. testing period and scarce representation of the obtained results, are unable to adequately represent the high potential of using DBN for load forecasting.

The main goal of this paper is to explore the possibilities of using deep belief networks in the problem of 24 ahead electricity consumption forecasting. A modelling method that includes detailed and systematic approach is presented. This paper goes one step beyond the time series forecasting approach using deep belief networks, as special attention is paid to the selection of the input variables, which is one of the most important steps. To this end, there is a detailed analysis of the data from a case study based on which the modelling of the neural network, and thus of the deep belief network is made. A novel input variable is introduced – the cheap tariff indicator, which according to the statistical analyses of the data has great influence on the population behaviour and therefore to the electricity consumption patterns.

The model was applied to an available historical real data describing the electricity consumption in the Republic of Macedonia. More specifically, three data sets are used. The first data set includes the electricity consumption data of the consumers that are directly connected to the electric power transmission network, most of them big industrial companies. The second data set represents the consumers that are connected to the distribution network, where the highest share of the consumption is from the household sector. And the third data set includes the total electricity consumption in the country (a sum of the first two data sets). The data for the electricity consumption for the period from 2008 to 2014 at hourly level is used, as well as the hourly temperature data for the same time period. The testing phase and accordingly the obtained results are for the period of the last two years: 2013 and 2014.

Please cite this article in press as: Dedinec A, et al., Deep belief network based electricity load forecasting: An analysis of Macedonian case, Energy (2016), http://dx.doi.org/10.1016/j.energy.2016.07.090

The results of this model are compared with the latest actual data, with a typical feed-forward multi-layer neural network and with the results obtained from a traditional neural network model used by the Macedonian system operator for providing 24-h ahead forecasted data. The comparisons show that the deep belief network model can be successfully applied for 24-h ahead load forecasting, providing results that surpass the traditional methods.

The outline of the paper is as follows. In Section 2, a description of the neural networks based models is given. Additionally, the concept of the deep belief networks is described, as well as the way this model is integrated into the traditional neural networks model. In Section 3 the case study that is analyzed in this paper is presented by providing a short overview of the Macedonian electric power system, followed by the analyses of the historical electricity consumption. The key results of short-term electricity forecasting in Macedonia obtained using deep belief networks accompanied by appropriate discussion are presented in Section 4. Section 5 concludes the paper.

## 2. Method

As it was discussed in the introduction, traditional neural networks are a common basic method used for electricity load forecasting. After defining the structure of the neural network, in this paper, deep belief network method is integrated into the traditional neural network model.

### 2.1. Neural networks

The main advantage of the artificial neural networks is that they include a process of learning, by which the relationship between the input and the output variables is determined [19]. They focus on the variables that are very significant for the output and ignore the information that has little impact on the output.

The neural network used in this paper has multi-layer feed-forward perceptron (MLP) structure, where the network is represented as a directed acyclic graph, whose structure has at least three layers – an input layer, one or more hidden layers and an output layer. The input layer gathers the model's inputs vector $x$, while the output layer gives the model's output vector $y$. The neurons in the hidden layer are activated by the hyperbolic tangent sigmoid transfer function.

The non-linear mapping between the input $x$ and the output $y$, when there are three layers in the network is given by the following equation:

$$y = \sum_{j=0}^{h}\left[w_j f\left(\sum_{i=0}^{d} w_{ji} x_i\right)\right] \tag{1}$$

where $d$ represents the number of input variables, $h$ represents the number of hidden layer neurons and the variable $y$ is a single linear output. The parameters $w_j$ and $w_{ji}$ represent the weights and biases that connect the layers.

During the training (learning) phase the weights and biases are adjusted in order to minimize the cost function, which in our case is the mean-squared error, given by:

$$E_D = \frac{1}{2}\sum_{i=1}^{N}\{y_i - t_i\}^2 = \frac{1}{2}\sum_{i=1}^{N} e_i^2 \tag{2}$$

where $y_i$ is the actual data and $t_i$ is the forecasted data. For the optimization problem of updating the weights and biases, error back propagation algorithm is used.

For the generalization (testing) phase we use the most common metric in the electricity load forecasting: MAPE (mean absolute percent error) and MAE (mean absolute error), given by the following equations:

$$MAE = |e_i| = |y_i - t_i| \tag{3}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{|e_i|}{t_i} \times 100\right) \tag{4}$$

### 2.2. Deep belief networks

As mentioned in Section 2.1, a back propagation algorithm is used during the training phase of the neural networks. Back propagation is a supervised learning method which uses pairs of input and desired output data in order to modify the weights of the connections of the neurons on different layers. However, the initial values of the weights affect the learning process and usually a random initialization of the parameters is used. This may lead to slower convergence and also the training process may get trapped at bad local optima [48,49]. Additionally, as the weights are randomly initialized, the results of the forecast are different for each process of network training. Another drawback of the back propagation algorithm is that it is not scalable. In fact, as the network gets larger, the performance of the network is not improved (or it may even be worse), which is also accompanied by time issue problems. As a solution to these problems, in Ref. [39] a deep belief network (DBN) is proposed. The idea is to use, a layer-by-layer unsupervised learning method in order to pre-train the initial values of the weights in the network. A layer-by-layer unsupervised training procedure implies that each layer captures the features of the previous one, starting from capturing the features of the lowest layer i.e. the training set. After the unsupervised training, fine tuning is applied using the standard supervised method (back propagation). In this way the initial parameters of the network are closer to the optimal solution, compared to the random initialization.

A deep belief network is a probabilistic, generative model that can learn to probabilistically reconstruct its inputs and is composed of multiple simple learning modules. In this paper, each pair of layers of the neural network is pre-trained by using restricted Boltzmann machine (RBM). Restricted Boltzmann machines are a special type of generative energy based models that can learn a probability distribution over its set of inputs [50]. An RBM has a single layer of hidden units that are not connected to each other and have undirected, symmetrical connections to a layer of visible units. The restriction is that their neurons must form a bipartite graph. The main advantage of the RBMs is that the hidden units are conditionally independent given the visible states, as there are no connections between hidden units.

The standard type of RBM has binary valued hidden and visible units. A joint configuration ($\mathbf{v}$,$\mathbf{h}$) of the visible and hidden units has the following energy:

$$\boldsymbol{E}(\mathbf{v},\mathbf{h}) = -\sum_{\boldsymbol{i\in visible}} \boldsymbol{a_i v_i} - \sum_{\boldsymbol{j\in hidden}} \boldsymbol{b_j h_j} - \sum_{\boldsymbol{i,j}} \boldsymbol{v_i h_j w_{ij}} \tag{5}$$

where $v_i$,$h_j$ are the binary states of visible unit $i$ and hidden unit $j$, $a_i$,$b_j$ are the biases and $w_{ij}$ is the weight between them. A lower energy indicates that the network is in a more desirable state. This

Please cite this article in press as: Dedinec A, et al., Deep belief network based electricity load forecasting: An analysis of Macedonian case, Energy (2016), http://dx.doi.org/10.1016/j.energy.2016.07.090

energy function is used to calculate the probability that is assigned to every possible pair of a visible and a hidden vector:

$$\boldsymbol{p}(\mathbf{v},\mathbf{h}) = \frac{1}{\boldsymbol{Z}}\boldsymbol{e}^{-\boldsymbol{E}(\mathbf{v},\mathbf{h})} \tag{6}$$

where Z is a partition function, which is a sum of $e^{-E(\mathbf{v},\mathbf{h})}$ over all possible configurations, and is used for normalizing:

$$\boldsymbol{Z} = \sum_{\mathbf{v},\mathbf{h}} \boldsymbol{e}^{-\boldsymbol{E}(\mathbf{v},\mathbf{h})} \tag{7}$$

The probability that the network assigns to a visible vector, **v**, is given by:

$$\boldsymbol{p}(\mathbf{v}) = \frac{1}{\boldsymbol{Z}}\sum_{\mathbf{h}} \boldsymbol{e}^{-\boldsymbol{E}(\mathbf{v},\mathbf{h})} \tag{8}$$

The gradient or the derivative of the log probability of a training vector with respect to a weight has a simple form:

$$\frac{\partial \log(\boldsymbol{p}(\boldsymbol{v}))}{\partial \boldsymbol{w_{ij}}} = \langle \boldsymbol{v_i h_j}\rangle_{data} - \langle \boldsymbol{v_i h_j}\rangle_{model} \tag{9}$$

where $<\ldots>_p$ represents averages with respect to distribution $p$. This means that the learning rule for stochastic steepest ascent in the log probability of the training data is:

$$\Delta \boldsymbol{w_{ij}} = \varepsilon(\langle \mathbf{v_i h_j}\rangle_{\mathbf{data}} - \langle \mathbf{v_i h_j}\rangle_{\mathrm{model}}) \tag{10}$$

where $\varepsilon$ is a learning rate.

Since a RBM is represented by a bipartite graph, it is easy to get an unbiased sample of $<v_i h_j>_{data}$. The hidden unit activations are mutually independent given the visible unit activations (and vice versa):

$$\boldsymbol{P}(\mathbf{v}|\mathbf{h}) = \prod_{\boldsymbol{i}=1}^{\boldsymbol{m}} \boldsymbol{P}(\boldsymbol{v_i}|\mathbf{h}) \tag{11}$$

And the individual activation probabilities, i.e. the state of a visible node, given a hidden vector, are represented by:

$$\boldsymbol{P}(\boldsymbol{v_i}=1|\mathbf{h}) = \boldsymbol{\sigma}\left(\boldsymbol{a_i} + \sum_{\boldsymbol{j}} \boldsymbol{h_j w_{ij}}\right) \tag{12}$$

where $\sigma(x) = 1/(1 + exp(-x))$ is the logistic sigmoid function. Respectively, for randomly selected training input **v**, the binary state $h_j$ of each hidden unit $j$ is set to 1 with probability:

$$\boldsymbol{P}(\boldsymbol{h_i}=1|\mathbf{v}) = \boldsymbol{\sigma}\left(\boldsymbol{b_i} + \sum_{\boldsymbol{i}} \boldsymbol{v_i w_{ij}}\right) \tag{13}$$

For calculating the $<v_i h_j>_{model}$ part of eq. (10), the following training process is used [50]. First, the visible units $v_i$ are set to be equal to the training sample. Afterwards, the hidden states $h_j$ are calculated according to equation eq. (13). One step “reconstruction” of the visible $v_i'$ and hidden $h_j'$ units is produced after repeating the process once more, using equations (12) and (13). Therefore, the weights are updated according to:

$$\Delta \boldsymbol{w_{ij}} = \varepsilon\left(\langle \mathbf{v_i h_j}\rangle_{\mathbf{data}} - \langle \boldsymbol{v_i' h_j'}\rangle_{recon}\right) \tag{14}$$

As a typical RBM uses binary logistic units for visible nodes, in this paper for the input data that are not binary, a conversion to continuous-valued inputs is used as described in Ref. [51]. Actually, those input variables are scaled to the range of (0,1) and then they are treated as probability for binary random variable to take value 1.

The procedure of the layer-by-layer unsupervised learning method, which is used in this paper, is illustrated in Fig. 1 for electricity load forecasting when using three layers of the neural network. RBM1 is composed of the input and the hidden layer of the neural network, and RBM2 is composed of the hidden and the output layer of the neural network. The values of the units in the hidden layer obtained from RBM1 are used as values for the input layer in RBM2. Certainly, the model can respectively be expanded for larger networks with more hidden layers.

### 2.3. Selection of the input variables of the neural network

One of the most important steps in order to obtain good results when modelling neural networks and deep belief networks for forecasting purposes is the selection of the input variables. When solving regression problems, and specifically when forecasting time series variables, there are two approaches (as it is indicated in the Introduction). The first one is to use only the previous x values of the variable in order to do forecasting (time series forecasting), and the other method is to use selected input variables, which may include selected points of the historical values and also other variables that are correlated to the one that is being forecasted.

## 3. Case study: the Macedonian electric power system

As a case study, the Macedonian electric power system is used in this paper focusing on the aim to provide a short-term load forecast of the electricity consumption in Macedonia. Therefore, a short overview of the electric power system and analyses of the historical electricity consumption data in Macedonia are presented in this section.

### 3.1. Overview

The electric power system of the Republic of Macedonia is composed of:

- Electricity generation capacities of around 2000 MW;
- Electricity transmission system, operated by the state owned company AD MEPSO – Skopje and
- Electricity distribution system, operated by EVN Macedonia AD.

The Macedonian transmission network operator – MEPSO provides hourly data for the electricity consumption, as well as the 24-h ahead forecasted electricity consumption, both accessible at its official site [52]. This data is divided into two groups. The first group is represented by the consumers that are directly connected to the electric power transmission network, which mainly involves the big industrial companies. The second group represents the consumers that are connected to the distribution network, where the highest share of the consumption is from the household sector. The data for the 24-h ahead forecasting of the first group of consumers, the big industrial companies, is as reported from the companies themselves. The forecasting of the second group is as calculated according to the distribution system operator EVN. Their model for forecasting of the electricity consumption is mainly based on traditional neural networks.

### 3.2. Electricity consumption analyses

The analyses of the electricity consumption play a key role when selecting the input data upon which the forecasting of the consumption depends on. Therefore, in this section an overview of the

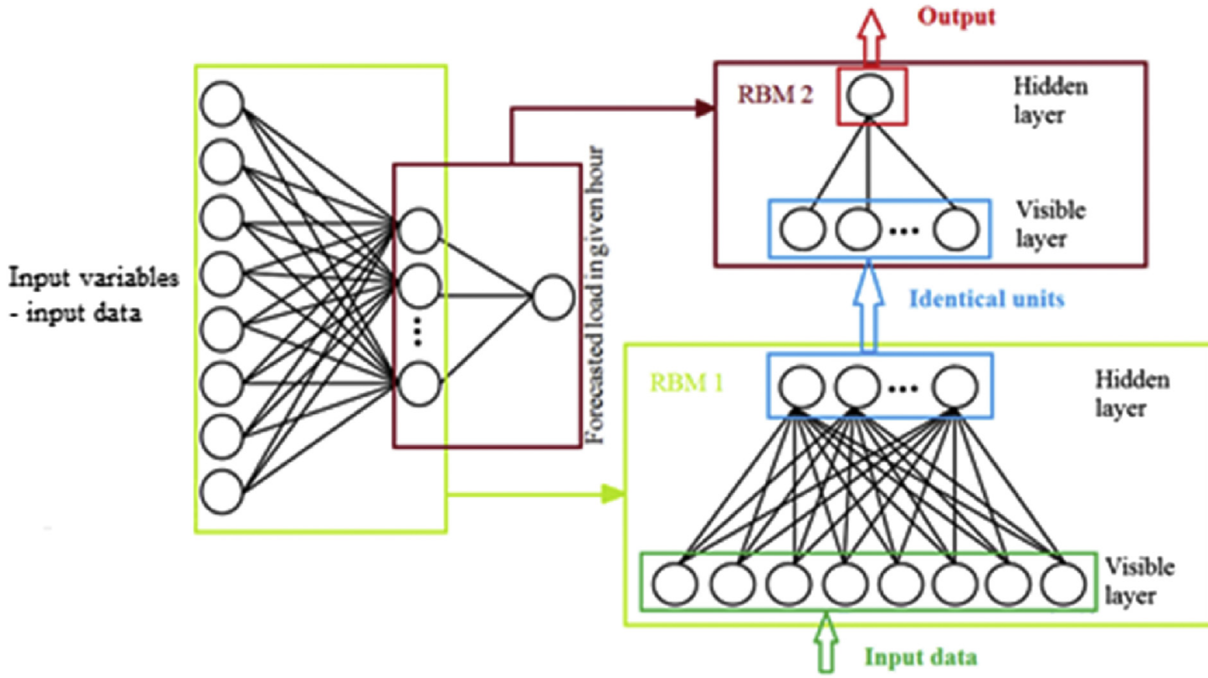



**Fig. 1.** Example of deep belief network with three layers.

electricity consumption in the Republic of Macedonia is provided.

The total electricity load in the Republic of Macedonia on monthly basis for the years 2008–2014 is presented in Fig. 2. It can be noticed that the highest consumption is during the heating season, which means that high share of the electricity consumption is used for heating. Also, the total electricity consumptions per year for the period from 2008 to 2014 are presented in Fig. 3, where the highest consumption is in 2011, followed by 2008. It is noticeable, however, that there is no pattern that is repeated on a yearly basis. In Fig. 3 the total electricity consumption is presented as a summation of the load of the consumers that are directly connected to the transmission network, and the load of the consumers connected to the distribution network. The share of the directly connected consumers to the transmission network ranges from 13% to 23% of the total electricity consumption.

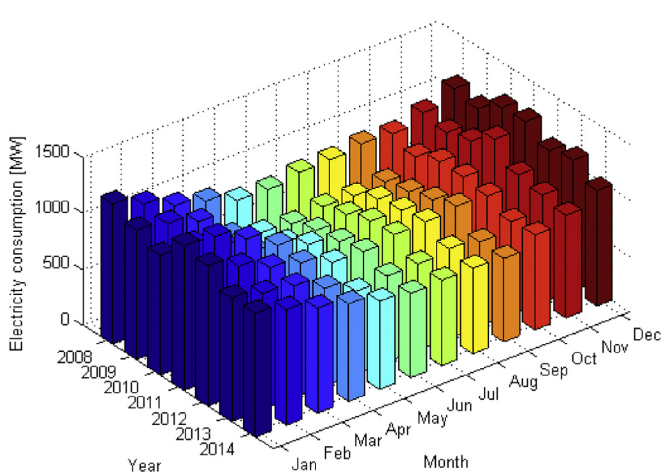



**Fig. 2.** Electricity consumption in the Republic of Macedonia on monthly basis for years 2008–2014.

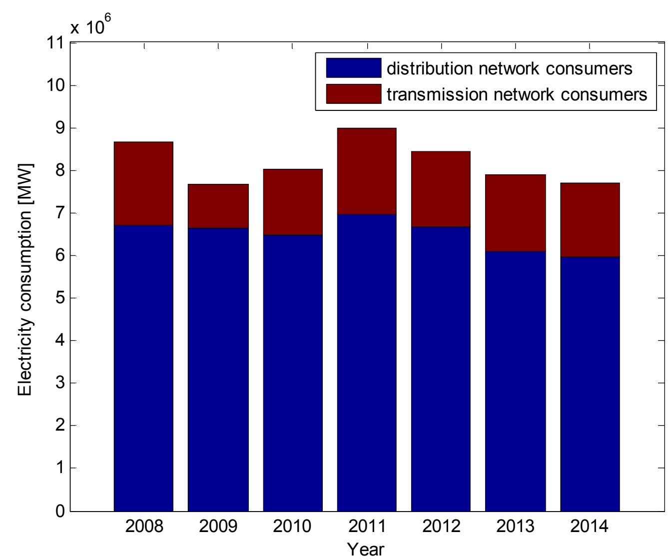



**Fig. 3.** Total electricity consumption (distribution network consumption + transmission network consumption) in the Republic of Macedonia for the 2008–2014 period.

In Fig. 4 the average hourly consumption of the consumers that are connected to the distribution network during work days is presented for 2011 and 2014. Throughout 2011 there are three peaks of the daily consumption. The first one starts at about 1 p.m., the second is around 5 p.m. and the last one starts at around 10 p.m. The first peak of consumption appears because in that period (more precisely, until August 2012) in Macedonia a cheap tariff period was in place from 1 p.m. to 4 p.m. on workdays. The second peak is due to consumption of the population when they come back from work, and the last peak is due to the second cheap tariff period which starts at 10 p.m. and ends at 7 a.m. (this cheap tariff period still

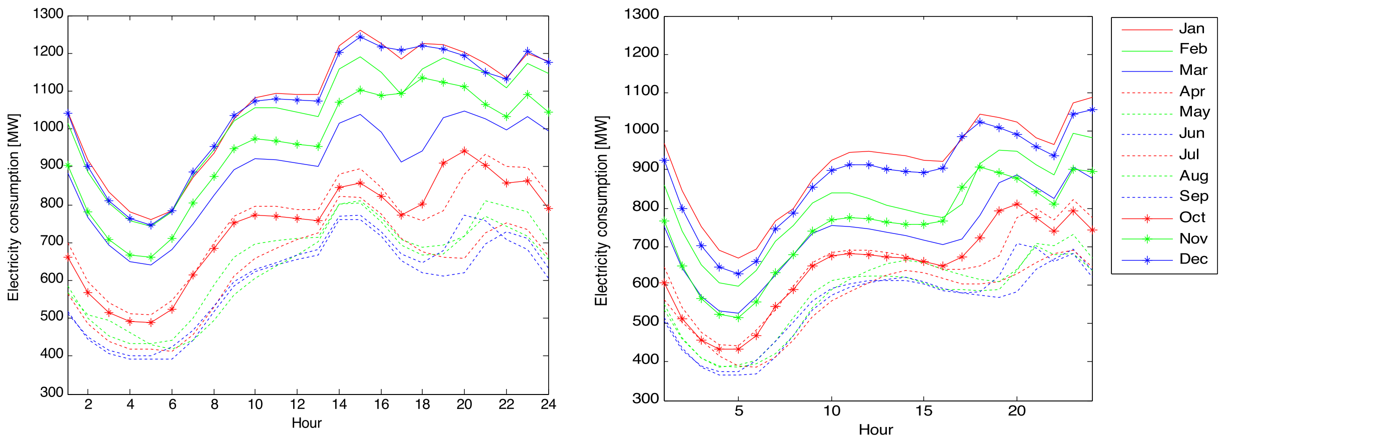



**Fig. 4.** Average hourly consumption of the consumers connected to the distribution network during work days for 2011 (on the left side) and 2014 (on the right side) for RM.

exists). However, in 2014 there are mainly two peaks during the day – the first one starting from 5 p.m. and the second one starting from 10 p.m. In fact, the first peak that was present in 2011 vanished as a result of the abolition of the daily cheap tariff period from 1 p.m. to 4 p.m. Both in 2011 and 2014 the consumption is minimal during the morning period from 3 a.m. to 6 a.m., and then rapidly increases until 10 a.m.

In Fig. 5 the average hourly consumption of the consumers that are directly connected to the transmission network (big industrial companies) during work days is presented for 2011 and 2014. It can be noticed that there are no strict regular patterns of the consumption as it was a case for the distribution network consumers. However, it can be concluded that most of these big industrial companies work more during the second and the third shift, so the consumption during the traditional working hours is lower. In 2014 a few big industrial companies worked using restricted working hours: second and third shift only (due to the introduction of a new policy for air pollution filters) and they stopped working during November and December 2014.

The average hourly consumption of all consumers in RM during the working days in 2011 and 2014 is presented in Fig. 6. It can be concluded that the total electricity consumption follows similar patterns as the distribution network consumers. This is due to the fact that the share of the distribution network consumption is about 77% of the total electricity consumption both in 2011 and 2014.

In Fig. 7 the average daily consumption for different days of the week is presented for the years from 2008 to 2014. As it is shown, the consumption is generally lowest on Saturdays, because it is nonworking day and also the big industrial companies do not work, but the highest consumption is on Sundays. Although it is also a nonworking day, on Sunday the cheap tariff applies during the whole day. The consumption on Sundays especially stands out in comparison to the other days in the weeks during and after 2012, or after the abolition of the daily cheap tariff in the working days.

### 3.3. *Selection of the input variables of the neural network for the case of Macedonia*

As mentioned in Section 2.3, there are two approaches for the selection of input variables. In this paper, both approaches are tested.

When selecting the input variables in the second approach, as there is no general rule that can be followed, the statistical analyses of the historical data is of great importance. In our case, this is the historical analyses of the electricity consumption, as it is shown in

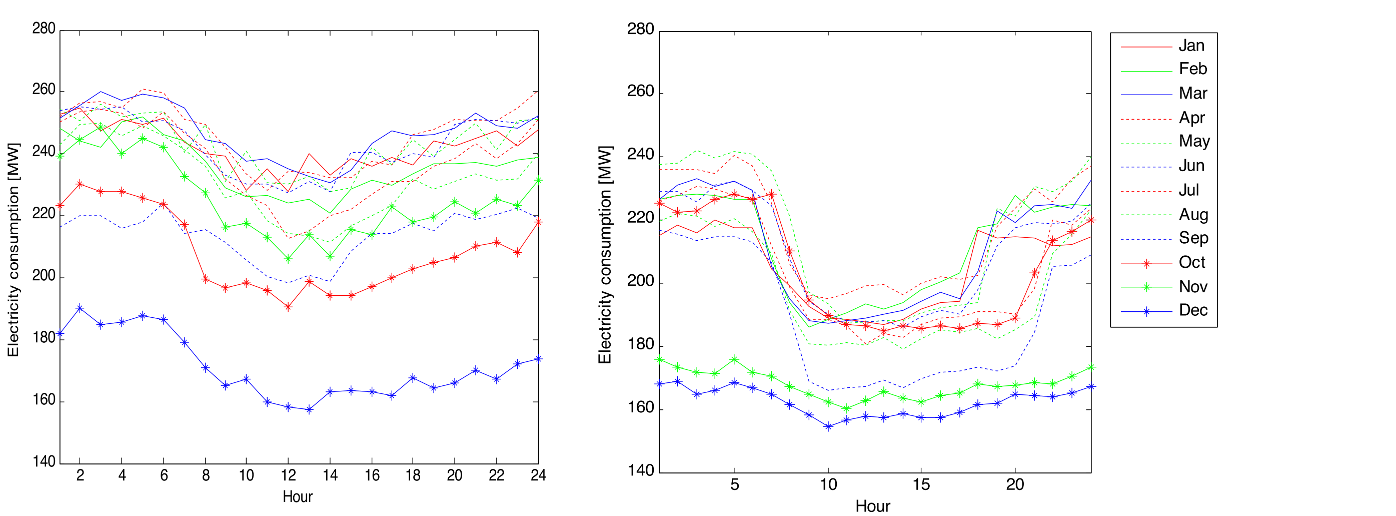



**Fig. 5.** Average hourly consumption of the consumers connected to the transmission network in working days for 2011 (on the left side) and 2014 (on the right side) for RM.

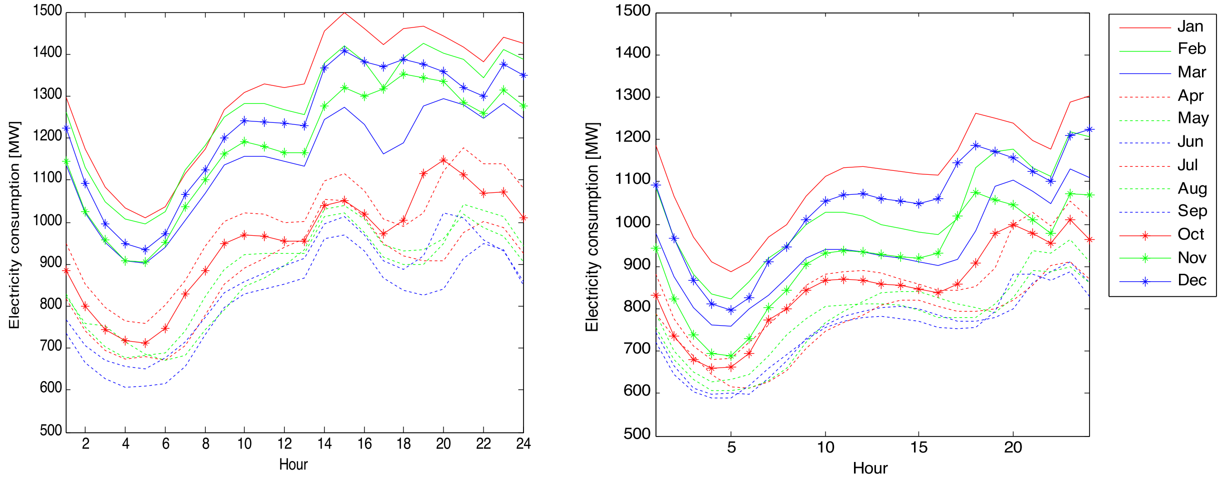



**Fig. 6.** Average hourly consumption of the total consumption in working days for 2011 (on the left side) and 2014 (on the right side) for RM.

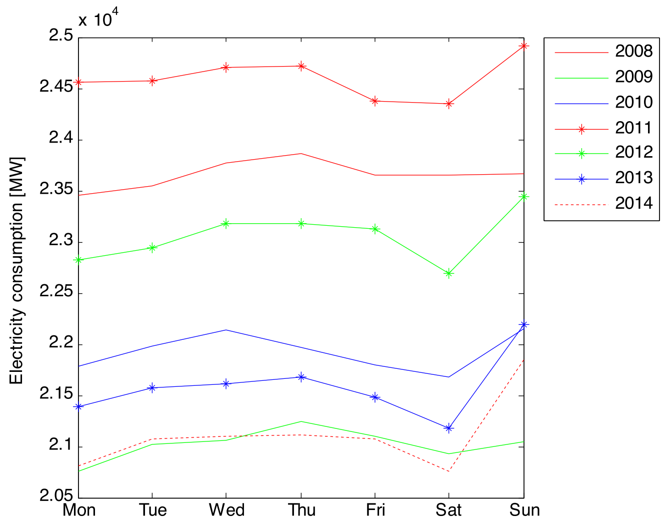



**Fig. 7.** Average daily consumption of the total consumption for the years from 2008 to 2014.

Section 3.2.

As it was presented, the electricity consumption shows strong patterns that are repeated on a daily basis. This leads to the conclusion that one of the most important input variables should be the hour of the day. Additionally, the consumption in a certain hour of the day is mostly similar to the consumption of the same hour, but the previous day. Also, the average load of the previous day is a significant feature, because it is close to the average load of the forecasted day. As day ahead forecasting is done in this paper, only the electricity consumption for the previous day (or earlier) can be considered, which means that input variables such as the electricity consumption from one or two hours ago cannot be included.

The load is also changing on a weekly basis. As previously stated the lowest consumption is on Saturdays, and the highest is on Sundays, and in the rest of the days the consumption is somewhere in the middle. So, a very important indicator is the day of week variable. Also, the same hour-day combination of the previous week is very significant. During the nonworking days, when there is no cheap tariff (which includes the Saturdays and the holidays that are not on Sundays) the consumption is lower than usual, so the information whether it is nonworking day (weekend or holiday) should be also considered as input variable. Naturally, according to the previous analyses of the data, the cheap tariff indicator is one of the most important input variables that has great impact on population behaviour, and consequently on the electricity consumption patterns.

The last input variable should be the temperature information. In Fig. 8 the average monthly temperatures in Macedonia for the period from 2008 to 2014 are shown. The correlation between the temperature and the electricity consumption in Macedonia (Fig. 2) for the period from 2008 to 2014 is presented in Fig. 9. The correlation is calculated using the Pearson's coefficient. A positive correlation (the correlation coefficient is greater than 0) means that the values of the variables increase or decrease together. Negative correlation means that as one variable is increasing, the other is decreasing and zero correlation means that the variables are not correlated. As it can be noticed there is a high correlation between the electricity consumption and the temperature, nearly −0.5, and the correlation is negative. This is due to the fact that during the winter period, when the temperatures are the lowest, the electricity consumption is at a highest level, and vice versa.

The relative humidity was also examined, but for the analyzed region it did not show any significant correlation with the

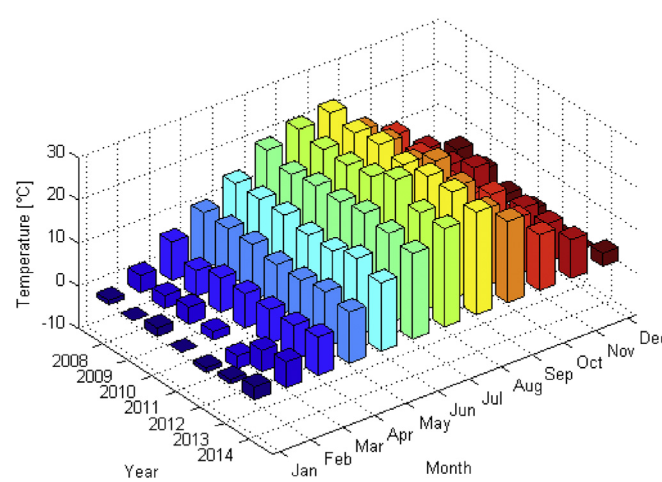



**Fig. 8.** Average monthly temperatures in Macedonia for the period from 2008 to 2014 [53].

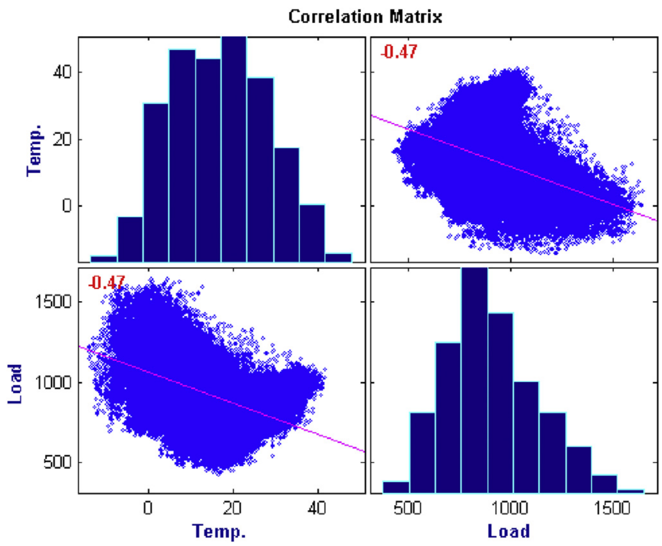



**Fig. 9.** Correlation between temperature and electricity consumption in Macedonia (2008–2014).

consumption of electricity, neither it improved the results of forecasting using the deep belief network.

Based on the previously stated analyses and the literature review [25,31,32], in Fig. 10 the final structure of the neural network (which is also used for the deep belief network) is shown. The correlation between each of these variables and the electricity consumption is calculated (using Pearson's and Spearman's correlation coefficients methods), and the absolute values of each of these correlation coefficients is greater than zero, which means that they are correlated (Detailed analyses of the correlation between these variables is presented in Ref. [54]). Additionally, each of these values is tested separately directly on the neural network, in order to verify that there is a positive effect by including it, and that it improves the results.

## 4. Results and discussion

The methods of neural and deep belief networks are implemented in MATLAB, and in order to reduce the execution time, the built-in functions for GPU utilization are used.

We have tested the two approaches for the selection of the input variables (as described in Section 2.3 and Section 3.3). Fig. 11 shows the mean absolute percent error for each of the three data sets analyzed. The results present the advantages of using the selected input variables instead of just using the historical data in the time series forecasting. Therefore, for the following analyzes the input variables presented in Fig. 10 are used.

By following the practical guidelines for training restricted Boltzmann machines provided in Ref. [50] the parameters of the deep belief network are obtained for each of the three data sets analyzed. As it is presented in Table 1, the best results are achieved when using four, four and three total number of layers (including input and output layers), for the distribution network consumers, transmission networks consumers and the total consumption, respectively. This means that there are a total of three, three and two restricted Boltzmann machines trained for each training set, respectively. Additionally, the optimal values for the number of epochs, the size of mini-batch, the learning rate and the momentum are also shown in Table 1.

Upon the calculation of the optimal parameters for deep belief networks, the same is performed for the traditional neural network, but only for two parameters. The first one is the number of layers of the network. The results show that the optimal number of layers is the same as in the deep belief networks: four for the distribution network consumers, four for the transmission network consumers and three for the total consumption forecasting.

The second parameter is the number of neurons in the hidden layer, which is determined for both the deep belief network and the traditional neural network. Fig. 12, Fig. 13 and Fig. 14 illustrate the mean average percent error for each of the three data set analyzed, depending on the number of neurons in the hidden layers for the deep belief network and traditional neural network. It can be concluded that in each of the three cases the minimum MAPE of the DBN is less than the minimum MAPE of the traditional NN. Moreover, it is interesting that DBN takes advantage of using more neurons in the hidden layers in order to reach the minimum error, unlike the traditional NN. However, generally at the beginning when there are fewer neurons, the traditional NN have better performances than DBN, but only to a certain number of nodes, after which the advantages of DBN start to stand out.

According to Fig. 12 the optimal number of neurons for the DBN is 25 neurons in the hidden layer of the DBN, and for the traditional NN the optimal number of neurons in the hidden layer is 8, for the

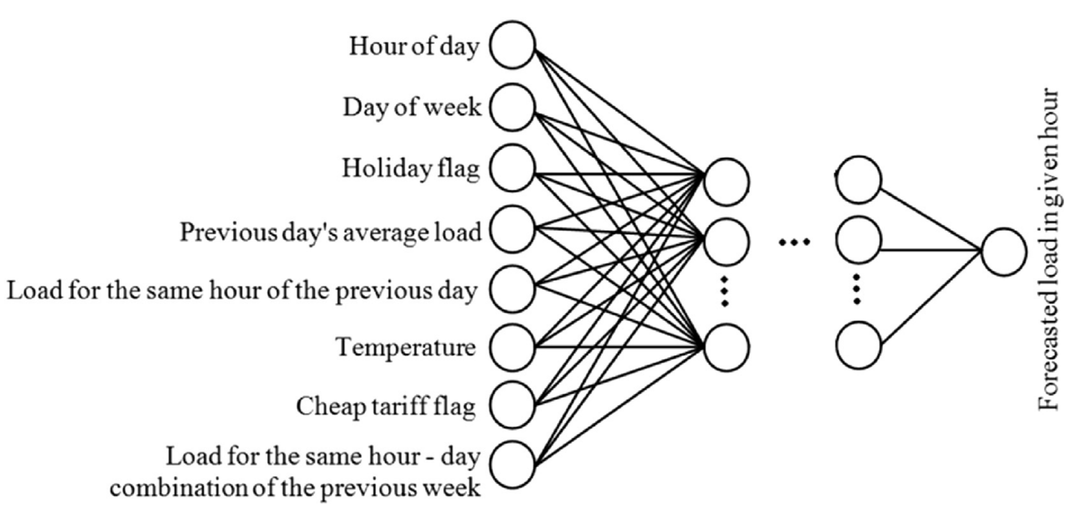



**Fig. 10.** Multilayer neural network for load forecasting.

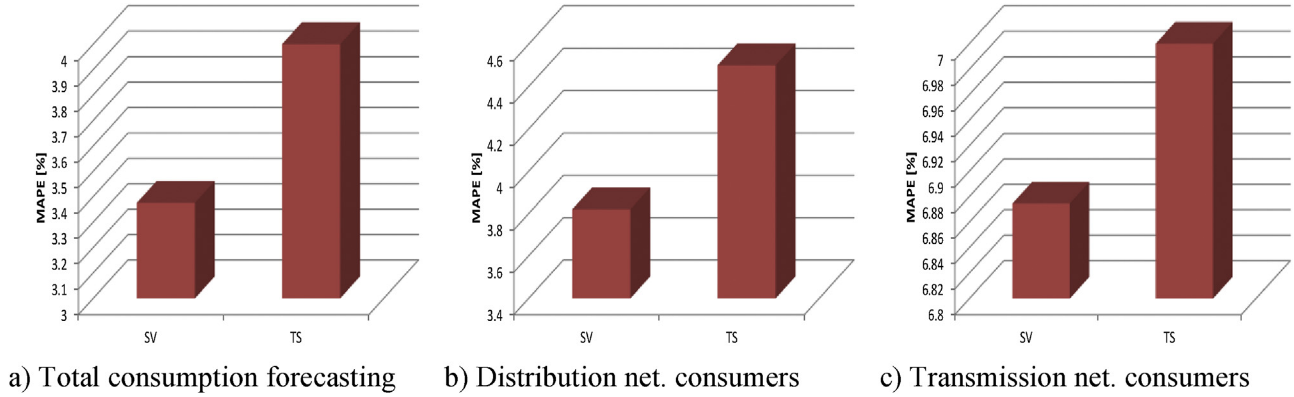



**Fig. 11.** MAPE for NN with selected variables (SV) and for time series forecasting (TS).

**Table 1**
Parameters used for the deep belief network.

| | Distribution net. consumers | Transmission net. consumers | Total consumption |
|---|---|---|---|
| Total number of layers | 4 | 4 | 3 |
| Number of epochs | [1,3,3] | [1,1,1] | [1,3] |
| Size of mini-batch | [2,2,2] | [1,1,1] | [2,2] |
| Learning rate | [1,0.8,1] | [1,1,1] | [1,1] |
| Momentum | [0,0,0] | [0,0,0] | [0,0] |

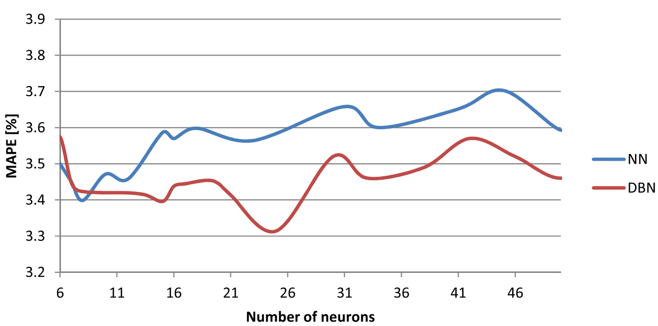



**Fig. 12.** Mean average percent error of the total consumption depending on number of neurons for traditional neural network and deep belief network.

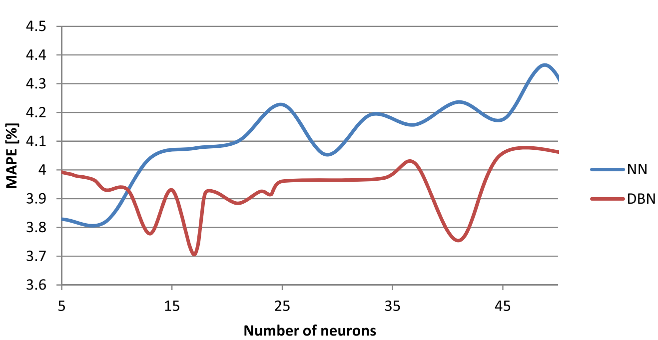



**Fig. 13.** Mean average percent error of the distribution network consumers depending on number of neurons for traditional neural network and deep belief network.

total electricity consumption forecasting.

For forecasting the distribution network consumers' load, the minimum MAPE for the DBN is achieved with 17 neurons in each of the hidden layers, and for the traditional NN there are 9 neurons in each of the hidden layers (Fig. 13).

The transmission network consumers' load is best forecasted by

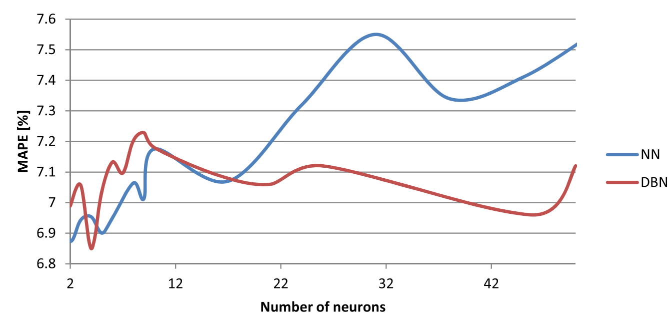



**Fig. 14.** Mean average percent error of the transmission network consumers depending on number of neurons for traditional neural network and deep belief network.

using 4 neurons in each of the hidden layers in a DBN, and 2 neurons in each of the layers in the traditional NN (Fig. 14).

All further results presented in this paper are obtained by using these optimal values for the parameters of the DBN and the traditional NN.

Fig. 15 presents the difference between the 24 h ahead electricity consumption forecasted by DBN and the real electricity consumption in the Republic of Macedonia at hourly level for the testing period, i.e. 2013 and 2014. Furthermore, zoomed representation is shown for randomly selected 150 h, in order to emphasize the small deviations of the forecasted from the actual consumption.

In addition to the comparison of the DBN forecasted data with the traditional NN forecasting and the actual data, the results are also compared to the forecasted data provided by the Macedonian system operator (MEPSO). In Fig. 16 the absolute percent error of the total electricity consumption is presented for MEPSO data, the forecasted data by the traditional NN model and the DBN model. Actually, the concrete numbers for the mean absolute percent error for each of the three cases are illustrated in Fig. 17a. It is very important to note here that for the transmission network

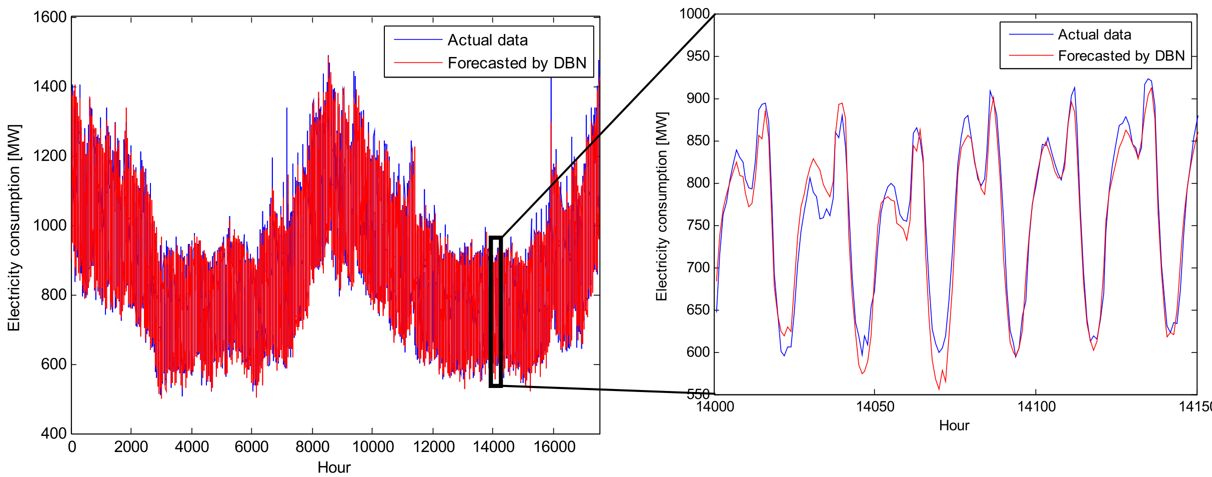



**Fig. 15.** Actual and forecasted by DBN total consumption in RM for the testing period (2013 and 2014).

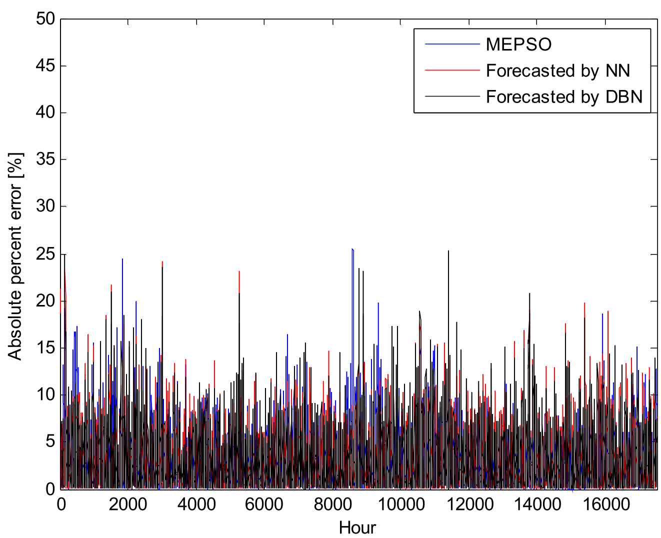



**Fig. 16.** Absolute percent error of the total electricity consumption for the testing period.

consumers' load, MEPSO does not use any software tool in order to forecast it, but it calculates it according to the work plans for the following day submitted by the consumers (the big industrial companies) themselves. As shown in Fig. 17b, the plans of large industrial companies for 2013 and 2014 are fairly more accurate than the predictions obtained by DBN and NN. However, there is an improvement of the MAPE when using the DBN and NN total consumption forecast for up to 2.3% in relation to the MEPSO predictions, though they use more accurate data (i.e. plans) for determining the load of transmission network consumers (Fig. 17a). In order to make a realistic comparison, in the DBN and NN models for forecasting of the total electricity consumption, the data for the plans of the transmission network consumers' load from MEPSO is used instead of the forecasting. This comparison is shown in Fig. 17c, where now clearly the improvement of the forecasting by using our NN model is highlighted when compared to MEPSO predictions, as well as the big advantage of using the DBN model which improves the MAPE for 8% compared to the MEPSO data.

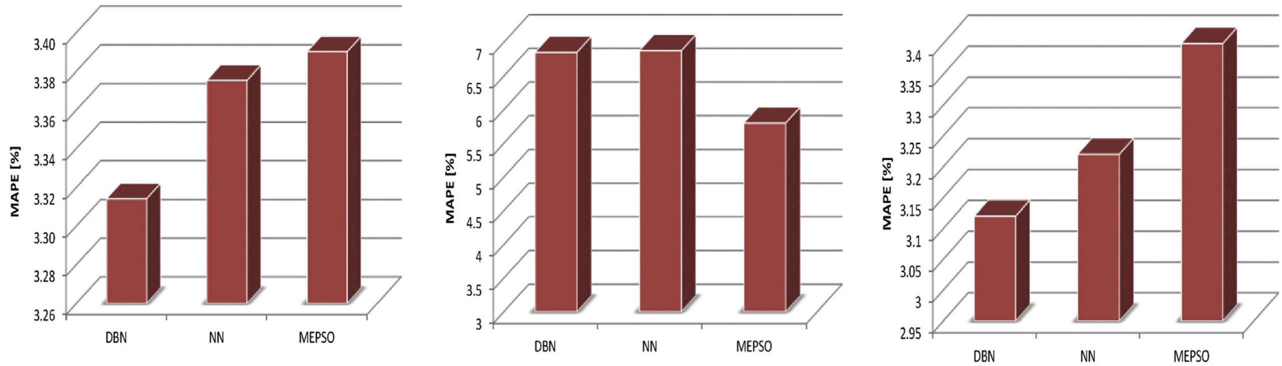
The advantages of using DBN model for electricity forecasting are mostly emphasized when using the data for the consumers that are connected to the distribution network (mainly represented by the household sector). In Fig. 18 the DBN forecasting is compared to the real data of the distribution network consumers in the Republic of Macedonia for 2013 and 2014, as well as a large scale representation of randomly selected 150 h. According to the data presented in Fig. 19, it is shown that there is an improvement of the mean absolute percent error by 8.6% when using DBN model compared to MEPSO forecasting data.

In addition of the hourly forecasting of the following 24-h electricity consumption, also the peak consumption of the following day is forecasted, using the same model (Figs. 1 and 10). Fig. 20 illustrates the improvement of the daily peak forecasting when using DBN compared to the traditional NN model and MEPSO data, for each of the three subsets of data. (Again, for NN and DBN models the data for forecasting of the transmission network consumers' load is used and not the plans). The greatest improvement is achieved for the distribution network consumers, or the DBN daily peak MAPE is improved for 21% when compared to the MEPSO data. However, there is also great improvement of the daily peak MAPE for the other two data sets, and in both cases the DBN forecasted daily peak MAPE is 10% better than the MEPSO daily peak MAPE.

Fig. 21 shows the average hourly absolute percent error depending on the hour of the day, when using the DBN model, for each of the three data sets. It is interesting to note that the highest error is in the afternoon period from 1 to 6 p.m. for the distribution network consumers and the total consumption, which is mainly due to the unpredictable peoples' behaviour when they come back from work and the abolition of the daily cheap tariff from 1 to 4 p.m. On the other hand, for the transmission network consumers this error is fairly even distributed during the whole day because most of the big industrial companies also work in second and third shift. The error is highest for the first work shift, because during the testing period, as stated in Section 2.2 few big industrial companies worked using restricted working hours: second and third shift only, which was not a case during the training period of the DBN model.

The average absolute percent error depending on day of week when forecasting each of the three data sets analyzed by using DBN model is presented in Fig. 22. For the distribution network consumers and so for the total consumption, the most unpredictable day is Monday as a first work day of the week. The smallest error is presented in the middle of the week, i.e. on Thursdays and Fridays. For transmission network consumers the smallest error is on Saturdays and Sundays, because most of them do not work during these two days.

Please cite this article in press as: Dedinec A, et al., Deep belief network based electricity load forecasting: An analysis of Macedonian case, Energy (2016), http://dx.doi.org/10.1016/j.energy.2016.07.090




a) Total consumption forecasting b) Transmission net. consumers c) Total consumption (forecast of distribution net. consumers+plans for transmission net. consumers)

Note: The plans for the transmission network consumption provided by MEPSO is not forecasted by using a software tool, but is calculated according to the following day work plans submitted by the big industrial companies themselves.

**Fig. 17.** MAPE of the forecasting by using DBN, NN and MEPSO plans.

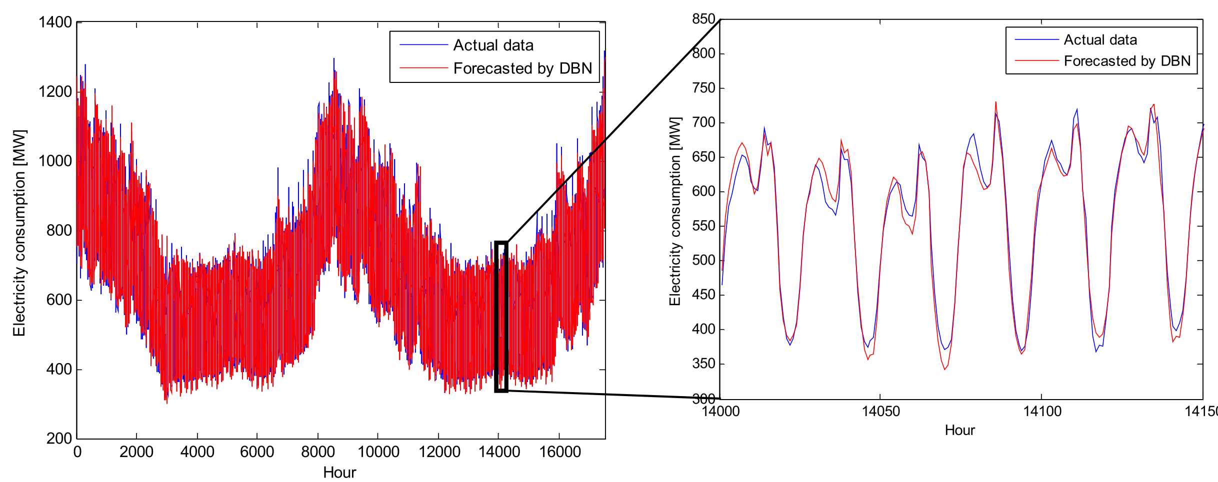



**Fig. 18.** Actual and forecasted by DBN electricity consumption of the consumers connected to the distribution network in RM for the testing period (2013 and 2014).

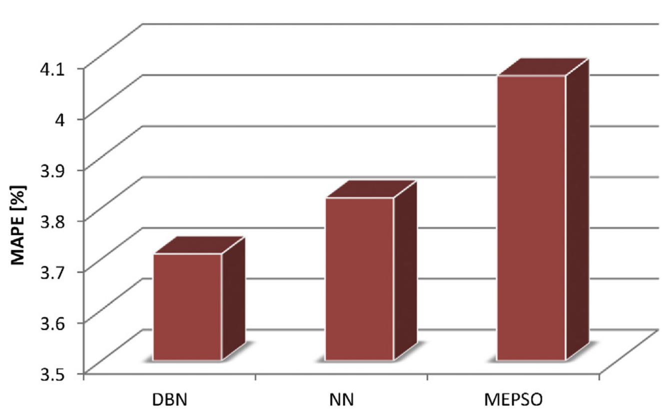



**Fig. 19.** MAPE of the forecasting of the distribution net. consumption by using DBN, NN and of the forecasts provided by MEPSO.

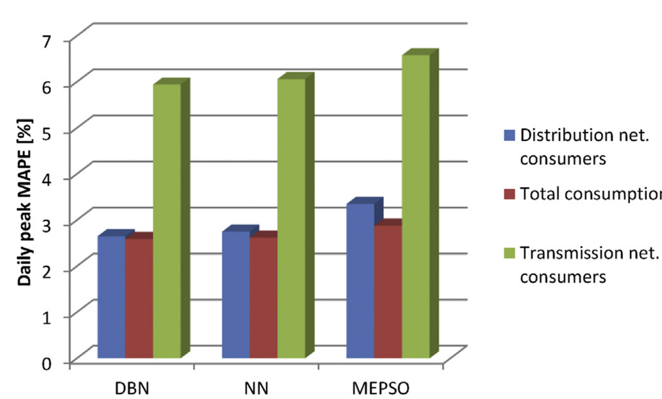



**Fig. 20.** Daily peak forecast mean average percent error of the three data sets analyzed, by using DBN, NN and data provided from MEPSO.

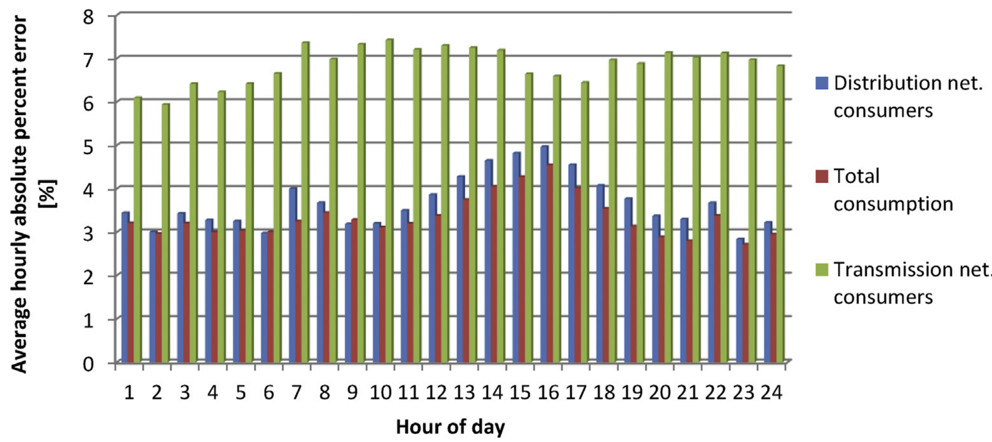



**Fig. 21.** Average absolute percent error by hour of day (by using DBN).

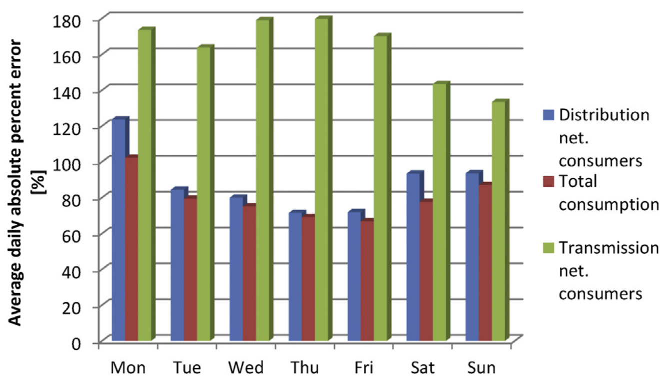



**Fig. 22.** Average absolute percent error by day of week (by using DBN).

## 5. Conclusion

In this paper a model of deep belief network composed of multiple restricted Boltzmann machines was incorporated into a feed-forward multi-layer perceptron neural network model in order to make short-term electricity load forecasting. The model is tested on the Macedonian hourly electricity consumption data for a period of 7 years (2008–2014). The results of the deep belief network model are compared to the results of using only feed-forward multi-layer perceptron neural network model. Furthermore, a comparison is made with the data provided by the Macedonian system operator (MEPSO) for electricity consumption predictions.

The following conclusions are drawn:

- The results present the importance of using the selected input variables instead of just using the historical data in the time series forecasting problems.
- In each of the three data sets analyzed in this paper the minimum mean average percent error of the DBN model forecasting is less than the minimum MAPE of the traditional NN model. Additionally, DBN takes advantage of using more neurons in order to reach the minimum error, unlike the traditional NN. When there are fewer neurons, the traditional NN have better performances than DBN, but only to a certain number of nodes, after which the advantages of DBN start to stand out.
- It is shown that there is an improvement of the mean absolute percent error by up to 8.6% when using DBN model for 24-h ahead electricity forecasting compared to MEPSO forecasting data.
- There is also great improvement of the following day peak consumption forecasting. Actually, the DBN daily peak MAPE is improved for up to 21% when compared to the MEPSO data.
- The obtained results show that the forecasting is much more accurate for the distribution network consumers' load (mainly household sector) than the transmission network consumers' load (represented mainly by the big industrial companies)
- Additionally, the results illustrate the most critical period of day and most critical day of week for forecasting electricity consumption when using DBN.

## Acknowledgement

This work was partially financed by the Faculty of Computer Science and Engineering at the “Ss. Cyril and Methodius” University.

## References

[1] Friedrich L, Afshari A. Short-term Forecasting of the Abu Dhabi electricity load using multiple weather variables. Energy Procedia 2015;75:3014–26. http://dx.doi.org/10.1016/j.egypro.2015.07.616.
[2] Loulou R, Goldstein G, Noble K. Documentation for the MARKAL family of models. 2004. p. 1–389. http://www.iea-etsap.org/web/MrklDoc-I_StdMARKAL.pdf.
[3] Victor N, Nichols C, Balash P. The impacts of shale gas supply and climate policies on energy security: the U.S. energy system analysis based on MARKAL model. Energy Strateg Rev 2014;5:26–41. http://dx.doi.org/10.1016/j.esr.2014.10.008.
[4] Dedinec A, Taseska-Gjorgievska V, Markovska N, Pop-Jordanov J, Kanevce G, Goldstein G, et al. Low emissions development pathways of the Macedonian energy sector. Renew Sustain Energy Rev 2016;53:1202–11. http://dx.doi.org/10.1016/j.rser.2015.09.044.
[5] Connolly D, Lund H, Mathiesen BV, Werner S, Möller B, Persson U, et al. Heat Roadmap Europe: combining district heating with heat savings to decarbonise the EU energy system. Energy Policy 2014;65:475–89. http://dx.doi.org/10.1016/j.enpol.2013.10.035.
[6] Fernandes L, Ferreira P. Renewable energy scenarios in the Portuguese electricity system. Energy 2014;69:51–7. http://dx.doi.org/10.1016/j.energy.2014.02.098.
[7] Huang Y, Bor YJ, Peng CY. The long-term forecast of Taiwan's energy supply and demand: LEAP model application. Energy Policy 2011;39:6790–803. http://dx.doi.org/10.1016/j.enpol.2010.10.023.
[8] Xiong P, Dang Y, Yao T, Wang Z. Optimal modeling and forecasting of the energy consumption and production in China. Energy 2014;77:623–34. http://dx.doi.org/10.1016/j.energy.2014.09.056.
[9] Hamzacebi C, Es HA. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 2014;70:165–71. http://dx.doi.org/10.1016/j.energy.2014.03.105.
[10] Bianco V, Manca O, Nardini S, Minea AA. Analysis and forecasting of nonresidential electricity consumption in Romania. Appl Energy 2010;87:3584–90.

http://dx.doi.org/10.1016/j.apenergy.2010.05.018.
[11] Bianco V, Manca O, Nardini S. Electricity consumption forecasting in Italy using linear regression models. Energy 2009;34:1413–21. http://dx.doi.org/10.1016/j.energy.2009.06.034.
[12] Ardakani FJ, Ardehali MM. Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types. Energy 2014;65:452–61. http://dx.doi.org/10.1016/j.energy.2013.12.031.
[13] Kaytez F, Taplamacioglu MC, Cam E, Hardalac F. Forecasting electricity consumption: a comparison of regression analysis, neural networks and least squares support vector machines. Int J Electr Power Energy Syst 2015;67:431–8. http://dx.doi.org/10.1016/j.ijepes.2014.12.036.
[14] World Energy Statistics | World Energy Consumption & Stats n.d. https://yearbook.enerdata.net/ (accessed November 18, 2015).
[15] Deihimi A, Orang O, Showkati H. Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction. Energy 2013;57:382–401. http://dx.doi.org/10.1016/j.energy.2013.06.007.
[16] Raviv E, Bouwman KE, van Dijk D. Forecasting day-ahead electricity prices: utilizing hourly prices. Energy Econ 2015;50:227–39. http://dx.doi.org/10.1016/j.eneco.2015.05.014.
[17] Keles D, Scelle J, Paraschiv F, Fichtner W. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. Appl Energy 2016;162:218–30. http://dx.doi.org/10.1016/j.apenergy.2015.09.087.
[18] Dudek G. Pattern-based local linear regression models for short-term load forecasting. Electr Power Syst Res 2016;130:139–47. http://dx.doi.org/10.1016/j.epsr.2015.09.001.
[19] Lauret P, Fock E, Randrianarivony RN, Manicom-Ramsamy J-F. Bayesian neural network approach to short time load forecasting. Energy Convers Manag 2008;49:1156–66. http://dx.doi.org/10.1016/j.enconman.2007.09.009.
[20] Andersen FM, Larsen HV, Gaardestrup RB. Long term forecasting of hourly electricity consumption in local areas in Denmark. Appl Energy 2013;110:147–62. http://dx.doi.org/10.1016/j.apenergy.2013.04.046.
[21] Khwaja AS, Naeem M, Anpalagan A, Venetsanopoulos A, Venkatesh B. Improved short-term load forecasting using bagged neural networks. Electr Power Syst Res 2015;125:109–15. http://dx.doi.org/10.1016/j.epsr.2015.03.027.
[22] Tang N, Zhang D-J. Application of a load forecasting model based on improved grey neural network in the smart grid. Energy Procedia 2011;12:180–4. http://dx.doi.org/10.1016/j.egypro.2011.10.025.
[23] Ilic S, Vukmirovic S, Erdeljan A, Kulic F. Hybrid artificial neural network system for short-term load forecasting. Therm Sci 2012;16:215–24. http://dx.doi.org/10.2298/TSCI120130073I.
[24] Amber KP, Aslam MW, Hussain SK. Electricity consumption forecasting models for administration buildings of the UK higher education sector. Energy Build 2015;90:127–36. http://dx.doi.org/10.1016/j.enbuild.2015.01.008.
[25] Mahmoud TS, Habibi D, Hassan MY, Bass O. Modelling self-optimised short term load forecasting for medium voltage loads using tunning fuzzy systems and Artificial Neural Networks. Energy Convers Manag 2015;106:1396–408. http://dx.doi.org/10.1016/j.enconman.2015.10.066.
[26] Hernández L, Baladrón C, Aguiar JM, Carro B, Sánchez-Esguevillas A, Lloret J. Artificial neural networks for short-term load forecasting in microgrids environment. Energy 2014;75:252–64. http://dx.doi.org/10.1016/j.energy.2014.07.065.
[27] Hsiao Y-H. Household electricity demand forecast based on context information and user daily schedule analysis from meter data. IEEE Trans Ind Inf 2015;(11):33–43. http://dx.doi.org/10.1109/TII.2014.2363584.
[28] Javed F, Arshad N, Wallin F, Vassileva I, Dahlquist E. Forecasting for demand response in smart grids: an analysis on use of anthropologic and structural data and short term multiple loads forecasting. Appl Energy 2012;96:150–60. http://dx.doi.org/10.1016/j.apenergy.2012.02.027.
[29] Qiu X, Zhang L, Ren Y, Suganthan PN. Ensemble deep learning for regression and time series forecasting. comput intell ensemble Learn (CIEL). In: IEEE symp 2014; 2014. p. 1–6.
[30] Hrasko R, Pacheco GC, Krohling RA. Time series prediction using restricted boltzmann machines and backpropagation. Procedia Comput Sci 2015;55:990–9. http://dx.doi.org/10.1016/j.procs.2015.07.104.
[31] Jovanović S, Savić S, Bojić M, Djordjević Z, Nikolić D. The impact of the mean daily air temperature change on electricity consumption. Energy 2015;88:604–9. http://dx.doi.org/10.1016/j.energy.2015.06.001.
[32] Badri A, Ameli Z, Birjandi AM. Application of artificial neural networks and fuzzy logic methods for short term load forecasting. Energy Procedia 2012;14:1883–8. http://dx.doi.org/10.1016/j.egypro.2011.12.1183.
[33] Khan AR, Mahmood A, Safdar A, Khan ZA, Khan NA. Load forecasting, dynamic pricing and DSM in smart grid: a review. Renew Sustain Energy Rev 2016;54:1311–22. http://dx.doi.org/10.1016/j.rser.2015.10.117.
[34] Goel A, Goel A. Regression based forecast of electricity demand of New Delhi. Int J Sci Res Publ 2014;4:1–7.
[35] Huang SR. Short-term load forecasting using threshold autoregressive models. Gener Transm Distrib IEE Proceedings 1997;144:477–81. http://dx.doi.org/10.1049/ip-gtd:19971144.
[36] Rao MSS, Soman SA, Menezes BL, Chawande P, Dipti P, Ghanshyam T. An expert system approach to short term load forecasting for Reliance energy limited, Mumbai. In: IEEE power India conf 2006; 2006. p. 763–8. http://dx.doi.org/10.1109/POWERI.2006.1632604.
[37] Hippert HS, Taylor JW. An evaluation of Bayesian techniques for controlling model complexity and selecting inputs in a neural network for short-term load forecasting. Neural Netw 2010;23:386–95. http://dx.doi.org/10.1016/j.neunet.2009.11.016.
[38] Quan H, Member SSSS, Srinivasan D, Member SSSS, Khosravi A, Member SSSS, et al. Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans Neural Netw Learn Syst 2014;25:303–15. http://dx.doi.org/10.1109/TNNLS.2013.2276053.
[39] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural Comput 2006;18:1527–54. http://dx.doi.org/10.1162/neco.2006.18.7.1527.
[40] Sanchez-Riera J, Hua K-L, Hsiao Y-S, Lim T, Hidayati SC, Cheng W-H. A comparative study of data fusion for RGB-D based visual recognition. Pattern Recognit Lett 2016;73:1–6. http://dx.doi.org/10.1016/j.patrec.2015.12.006.
[41] Tang B, Liu X, Lei J, Song M, Tao D, Sun S, et al. DeepChart: combining deep convolutional networks and deep belief networks in chart classification. Signal Process 2015. http://dx.doi.org/10.1016/j.sigpro.2015.09.027.
[42] Chen Y, Huang T, Liu H, Zhan D. Multi-pose face ensemble classification aided by Gabor features and deep belief nets. Opt – Int J Light Electron Opt 2015:1–9. http://dx.doi.org/10.1016/j.ijleo.2015.10.179.
[43] Abdel-Zaher AM, Eldeib AM. Breast cancer classification using deep belief networks. Expert Syst Appl 2015. http://dx.doi.org/10.1016/j.eswa.2015.10.015.
[44] Li H, Li X, Ramanathan M, Zhang A. Identifying informative risk factors and predicting bone disease progression via deep belief networks. Methods 2014;69:257–65. http://dx.doi.org/10.1016/j.ymeth.2014.06.011.
[45] Tran VT, AlThobiani F, Ball A. An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks. Expert Syst Appl 2014;41:4113–22. http://dx.doi.org/10.1016/j.eswa.2013.12.026.
[46] Shen F, Chao J, Zhao J. Forecasting exchange rate using deep belief networks and conjugate gradient method. Neurocomputing 2015. http://dx.doi.org/10.1016/j.neucom.2015.04.071.
[47] Fagiani M, Squartini S, Gabrielli L, Spinsante S, Piazza F. Neurocomputing A review of datasets and load forecasting techniques for smart natural gas and water grids: analysis and experiments. Neurocomputing 2015;170:448–65. http://dx.doi.org/10.1016/j.neucom.2015.04.098.
[48] Kuremoto T, Kimura S, Kobayashi K, Obayashi M. Time series forecasting using a deep belief network with restricted Boltzmann machines. Neurocomputing 2014;137:47–56. http://dx.doi.org/10.1016/j.neucom.2013.03.047.
[49] He W. Deep neural network based load forecast. Comput Model New Technol 2014;3:258–62.
[50] Hinton G. A Practical Guide to Training Restricted Boltzmann Machines n.d. https://www.cs.toronto.edu/~hinton/absps/guideTR.pdf (accessed November 25, 2015).
[51] Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. Adv Neural Inf Process Syst 2007;19:153–60.
[52] MEPSO (Electricity Transmission System Operator of Macedonia) n.d. http://mepso.com.mk/en-us/Default.aspx (accessed November 24, 2015).
[53] Weather Undergraund - Weather History & Data Archive n.d. http://www.wunderground.com/history/ (accessed November 24, 2015).
[54] Dedinec A, Dedinec A. Correlation of variables with electricity consumption data. In: 6th Int. Conf. Inf. Soc. Technol; 2016. p. 118–23.