Journal of Energy

VOLUME 66 | 2017

journal homepage: http://journalofenergy.com/

NINOSLAV HOLJEVAC **CATARINA SOARES** **IGOR KUZLE**

ninoslav.holjevac@fer.hr catarina.soares@fer.hr igor.kuzle@fer.hr

University of Zagreb Faculty of Electrical Engineering and Computing

SHORT-TERM POWER SYSTEM HOURLY LOAD FORECASTING USING ARTIFICIAL NEURAL NETWORKS

SUMMARY

Artificial neural networks (ANN) have been used for many application in various sectors. The learning property of an ANN algorithm in solving both linear and non-linear problems can be utilized and applied to different forecasting problems. In the power system operation load forecasting plays a key role in the process of operation and planning.

This paper present the development of an ANN based short-term hourly load forecasting model applied to a real data from MIBEL – Iberian power market test case. The historical data for 2012 and 2013 ware used for a Multilayer Feed Forward ANN trained by Levenberg-Marquardt algorithm. The forecasted next day 24 hourly peak loads and hourly consumptions are generated based on the stationary output of the ANN with a performance measured by Mean Squared Error (MSE) and MAPE (Mean Absolute Percentage Error). The results have shown good alignment with the actual power system data and have shown proposed method is robust in forecasting future (short-term) hourly loads/consumptions for the daily operational planning.

Key words: Artificial neural networks; Short-term load forecasting; Electric power system operation and planning;

1. INTRODUCTION

The importance of load forecasting power system planning and operation is gaining on significance with the increase of stochastic element associated with both supply and consumption side [1], [2]. The forecasts are one of the most important inputs for system analysis tools e.g. economic dispatch and short term unit commitment. The main goal of these planning techniques is to maintain the stability and efficiency of the system. The deregulation of the power system and market structures both in the USA and Europe in recent years has made short term load forecasting increasingly important [3]. In a deregulated, competitive electricity market environment, the ability to accurately forecast load in the short term is of interest to all participants in the power system.

Different forecasting techniques for short term load forecasting have been used. There are models based on Kalman filtering [4], dynamic linear and non-linear model, and optimization techniques [5]. Generally speaking two most common models are the Autoregressive Integrated Moving Average (ARIMA) models and the multiple linear regression models [6]. ARIMA based methods have shown to achieve good mean absolute percentage errors (MAPE) [7]. These methods have been in use for a number of years, however, they are not able to adapt to changes in weather conditions or other load affecting factors [8]. Furthermore, these methods rely primarily on historical data but realized load demands and consumptions do not necessarily need to resemble and be similar to future load demands (especially in recent decade) and cannot efficiently consider weather data [9].

Artificial intelligence (AI) techniques have come up as a potential load forecasting techniques such as the fuzzy-neural model, the artificial neural network (ANN) model and the genetic algorithm method and out of the AI techniques, the most widely used model of feed-forward ANN [10]. The ANN model is suitable for load forecasting due to its ability to model the nonlinear relationships between variables without making prior assumptions on the functional relationship among the variables. Artificial neural networks are able to learn and adapt to the data. Artificial neural networks learn the interdependencies between the variables and reach the conclusion based on that information [11]. They can be used as an integral part of different applications, such as secondary regulation of voltage and reactive power [12] or a potential prediction module inside of central control energy management systems [13].

One of the most important factors influencing the performance of neural networks is the choice of input variables. Input variables commonly used in short term load forecasting are hour and indicators, recent load realizations and temperature and other weather variables. The use of recent hours load is of particular interest as the load series is strongly auto-correlated. To evaluate the correlation of the current hour load with previous hours load correlation analysis is frequently performed [14]. Other commonly used approach is using Euclidean norm approach [15]. In both methods there is a possibility to discard certain hours in the past and looking only at the most recent ones

In this paper, an hour ahead load forecasting method using an ANN is proposed. The paper presents a two-layered feedforward artificial neural network for performing a short-term power system load forecast. Two different models of ANN (ANN mode 1 and ANN mode 2) were tested and their results were compared. The main difference between developed artificial neural networks models is in the inclusion of different weather data (temperature, rainfall, wind). Both models were applied to the real data set to real data from the MIBEL - Iberian power market for years 2012 and 2013. Both models use the entire load profile of the previous day. Simulation results show that ANN mode 2 gives a slightly better performance than ANN model 1 and the overall use of ANN in load forecasting is robust and precise enough.

The paper is structured as follows: after the introduction in section II describes most commonly used components of the artificial neural network. Section III describes the developed models and forecasting procedure. Section IV presents the results of the preliminary correlation studies ("time-series study") and section V presents the ANN forecast results. Section VI concludes the paper and gives directions for potential future research work.

2. ARTIFICIAL NEURAL NETWORKS

An artificial neural network (ANN) consists of a series of interconnected systems of artificial neurons. Each neuron mimics the behaviour of a real neuron - receives data from multiple inputs, processes the data and responds with an output based on its characteristics as shown on figure below (Figure 1).

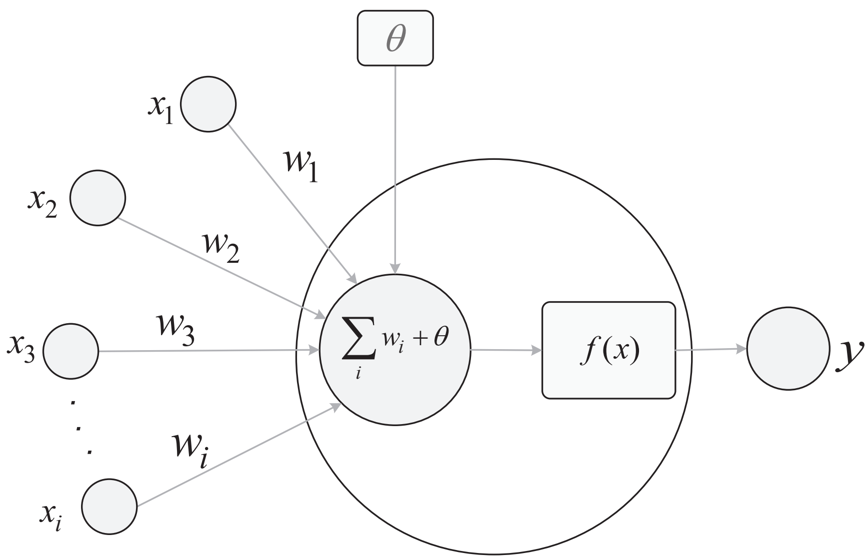


Figure 1. Basic schematic of the artificial neuron model (inputs, transfer function and output)

The output of the neuron can be given by equation (1):

$$y = f\left[\left(\sum_i w_i \cdot x_i\right) + \theta\right] \tag{1}$$

where x_i are neuron inputs, w_i are associated weights, θ represents the bias associated with the neuron and f is the transfer/activation function.

A feed-forward artificial neural network is used for the modelling of load forecasts. Basic schematic of feed-forward neural network (with the simple mark of the also commonly used back propagation) is represented on figure below (Figure 2). The choice of input variables and number of neurons in hidden layers affects the results.

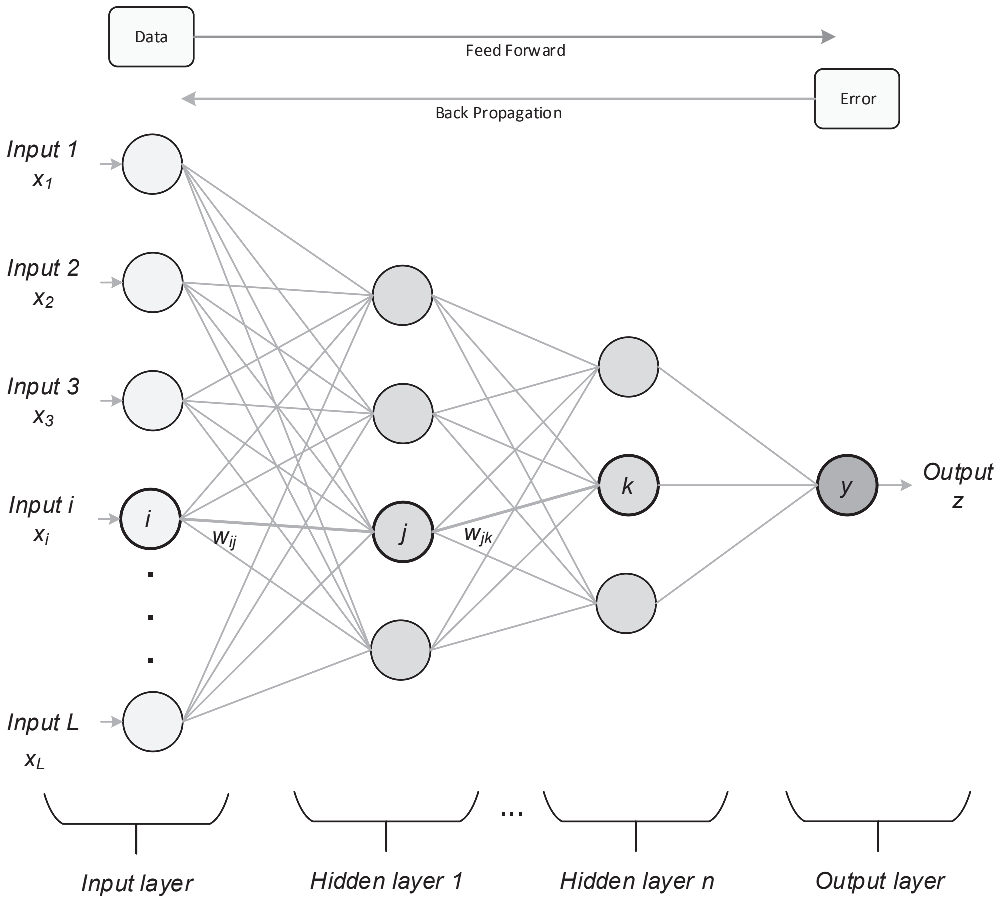


Figure 2. Artificial Neural Network (ANN) basic layout (with general multilayer structure)

Generally speaking there are three basic steps in creating (fitting) an artificial neural network model applied to load forecasting: selection of input variables, definition of neural network structure and determination of the training method.

2.1. Selection of input variables

The input variables used typically for short-term load forecasting are loads for previous hours, moment (time) indicators and weather conditions. In the process of ANN creation selection of different combination of input variables achieves different results. The performance is usually measured by means of mean squared error (MSE) given by equation (2) and then finally forecasts measured by means of absolute percentage error (MAPE) given by equation (3):

$$MSE = \frac{1}{N} \cdot \sum_{i=1}^{N} \left(l_i - a_i\right)^2 \tag{2}$$

$$MAPE = \varepsilon = \frac{1}{N} \cdot \sum_{i=1}^{N} \left| \frac{l_i - a_i}{a_i} \right| \tag{3}$$

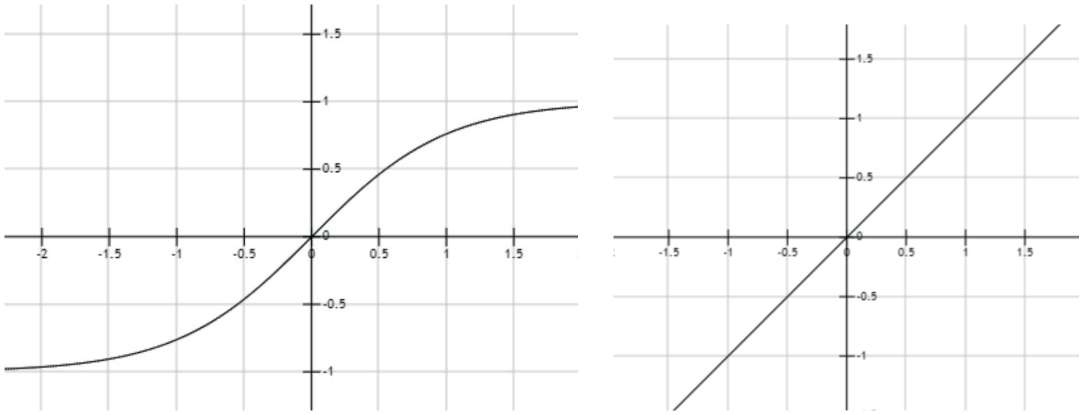
where N is the number of data samples, l_i is the target output (actual historical load), a_i is the artificial neural network output. Furthermore usually it is a recommendation in accordance to parsimony rule that networks with smaller number of input nodes and similar performance are better to be chosen to prevent over parametrization. The most important input series are loads from realized/past hours that usually have a strong correlation with the predicted load. There is a trade-off between accuracy of the model and computational burden when selecting the "look back window" of realized loads that is used for the ANN. Second important input is the day identifier. It is used to present the seasonal, weekly and daily cycles. It can be a binary input variable that for example indicates weekdays (0 for workdays, 1 for weekends) and an input variable ranging from 1 to 12 to represent the months of the year, January through December and 0 to 24 for hours of the day. Third important input parameter are weather variables, mainly temperature with others like wind, clouds and humidity/rainfall have smaller influence [16]. The list and definition of possible inputs and outputs is given in the table below (Table I).

Table I. List of possible ANN inputs and outputs

No.	Inputs	Description
1	Time of day (hour)	Time of the day identifier
2	Time of Week	Week day identifier
3	Time of Month	Month identifier
4	Loads of the day D-1	Loads of the previous day
5	Temperature of the day D-1	Temperatures of the previous day
6	Loads of day the D-6	Loads of the day D-6
7	Temperatures of the D+1*	Temperature forecasts
8	Rainfall of the D+1	Rainfall forecasts
9	Wind of the D+1	Wind forecast
*time D+1 is the day for which the loads are to be forecasted		

2.2. Artificial neural network structure

When fitting and ANN model it is important to define the artificial neural network structure. A feed forward networks is used. A total of 17 hidden neurons is used. Neural networks are made up of process units (neurons) that are bind in layers, whose number is set by the user and with different weights in the links. Inputs that are used are as mentioned realized load hour, week, month identifier, loads of for the past day and weather conditions inputs. The transfer function used for the hidden layer is sigmoid transfer function and for the output layer is linear transfer function (Figure 3).



a) *b)*

Figure 3. a) Tan-Sigmoid function $\frac{2}{(1+e^{-2x}-1)}$ b) Linear transfer function

2.3. Training method selection

The training algorithm determines the way the ANN adjusts the weights associated with different node connections. The algorithm that is used is Levenberg-Marquardt backpropagation and is used in combination with minimization of MSE. This method is quasi-Newton [17] method that is designed to approach second-order training without having to compute the Hessian matrix ($\mathbf{H}$) that can be approximated when the performance function has the form of a sum of squares (equation 4).

$$\mathbf{H} = \mathbf{J}^T\mathbf{J} \quad (4)$$

The gradient is then computed as:

$$\mathbf{g} = \mathbf{J}^T\mathbf{e} \quad (5)$$

where $\mathbf{J}$ is the Jacobian matrix that contains first derivatives of the network errors with respect to weights and biases and $\mathbf{e}$ is a vector of network errors. This algorithm uses the approximated Hessian matrix in the following cyclic update procedure:

$$\mathbf{x_{k+1}} = \mathbf{x_k} - \left[\mathbf{H} + \mu\mathbf{I}\right]^{-1} \mathbf{g} \tag{6}$$

where $\mathbf{x_k}$ is the vector of current weights and biases, μ is the adaptive scalar (when small the method behaves like Newton's method and when big the algorithm acts as a gradient descent with a small step size. μ is decreased as long as the performance function of the network is improved (reduction in size) with each step and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm and ensures efficient training step. The Levenberg-Marquardt algorithm was the chosen, and is recommended for most problems because it appears to be the fastest method for training moderate-sized feedforward neural networks. For larger problems, however, Scale Conjugate Gradient algorithm is recommended as it uses gradient calculations which are more efficient than the Jacobian calculations.

3. CORRELATION STUDY RESULTS

Performing the preliminary study of the information available (inputs) is essential to the forecasting process, so that final forecasts are as close as possible to possible realizations. With that aim, so called "time-series" analysis is intended to determine what is the most relevant information contained in the input variables, in order to build the prediction model. The study can be performed using the analysis of correlations between specific data enabling the construction of correlation charts and offering the trend lines and the values of Pearson correlation coefficient r defined as shown in equation (7) that is a zero-dimensional index that ranges from -1 to 1, inclusive and reflects the existence of a linear relationship between two data sets.

$$r = \frac{\sum_{i=1}^{n}\left[(x_i - \bar{x}) \cdot (y_i - \bar{y})\right]}{\sqrt{\sum_{i=1}^{n}\left[(x_i - \bar{x})^2\right]} \cdot \sqrt{\sum_{i=1}^{n}\left[(y_i - \bar{y})^2\right]}} \tag{7}$$

where two sets are being compared, actual data $\{x_1, ..., x_n\}$ and model output data $\{y_1, ..., y_n\}$ with their mean values marked with $\bar{x}, \bar{y}$.

The Pearson coefficient was calculated for all variables (Table II) that could be used in the forecasting model for predicting consumption for the next day (D+1). The whole data set available (both the training and validation parts) for the MIBEL data for years 2012 and 2013.

The results show the best correlation was as expected for hour, consumption (D-1) and consumption (D-6). Based on this information, the correlation graphs were made and are presented in the following figures (Figure 4, Figure 5 and Figure 6), confirming the relation between variables.

Table II. List of possible ANN inputs and outputs

Input data set		*Pearson coefficient*
Time data	**Hour**	**0.6540**
	Day week (D-1)	-0.3120
	Day Week (D+1)	-0.1869
	Month (D-1)	-01293
	Month (D+1)	-0.1348
	Year (D+1)	-0.0519
Production data	Wind production (D-1)	0.0997
	Wind production (D-6)	0.1001
	Hydro production (D-1)	0.3889
	Hydro production (D-6)	0.5254
	Cogeneration + Solar (D-1)	0.4427
	Cogeneration + Solar (D-6)	**0.5705**
	Coal production (D-1)	0.1663
	Coal production (D-6)	0.3088
	Nuclear production (D-1)	0.0449
	Nuclear production (D-6)	0.0644
	Natural Gas production (D-1)	0.3929
	Natural Gas production (D-6)	0.5866
Weather data	Temperature (D-1)	0.1234
	Temperature (D+1)	0.0975
	Wind (D+1)	0.1735
	Direction (D+1)	0.0534
	Irradiance (D+1)	0.3620
	Rainfall (D+1)	0.0159
Price data	Price (D-1)	0.4111
	Price (D-6)	0.4614
Export data	Export (D-1)	0.1423
Consumption data	**Hourly consumption (D-1)**	**0.7623**
	Hourly consumption (D-6)	**0.9122**

The more precise statistical analysis could include larger amounts of data (e.g. 10 years of data) but most impactful values would remain the same with the weather forecast for the forecasting period always playing a significant role regardless of the statistical analysis results. The impact of the weather inputs in the artificial neural network will be shown in the following chapter that presents the results of the wholesome forecasting process.

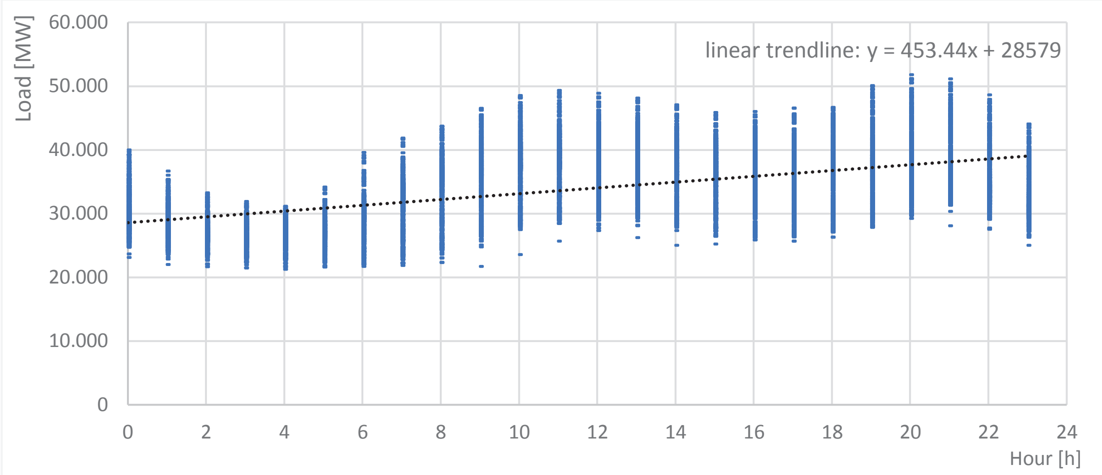


Figure 5. Hourly consumption (D+1) and hour identifier correlation coefficient

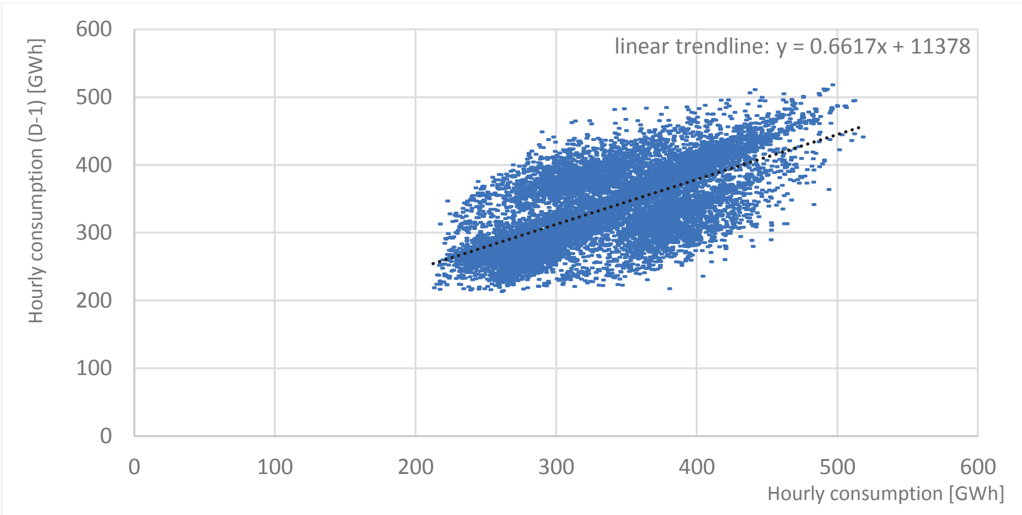


Figure 6. Hourly consumption/load (D+1) and hourly consumption/load (D-1) correlation coefficient (Pearson coefficient)

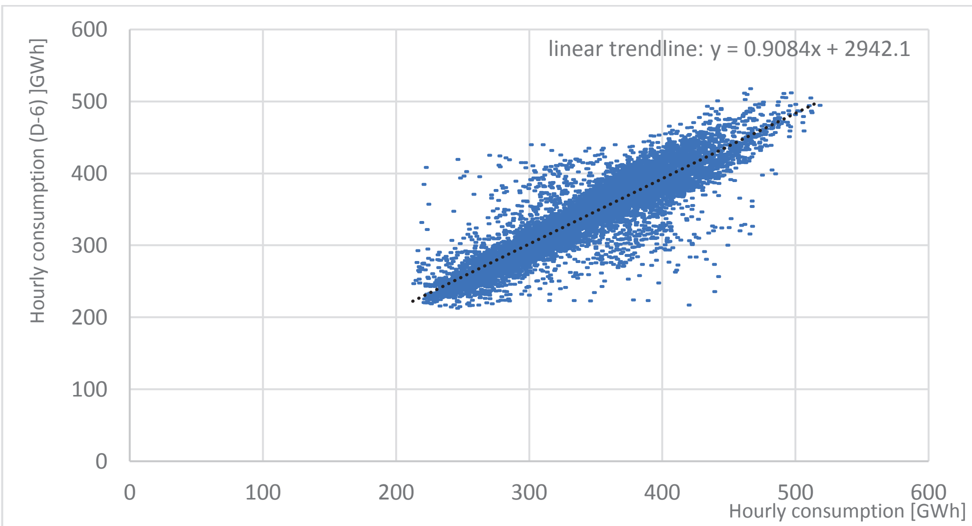


Figure 7. Hourly consumption/load (D+1) and hourly consumption/load (D-6) correlation coefficient (Pearson coefficient)

4. FORECASTING PROCESS AND RESULTS

The general forecasting process includes the creation of the artificial neural network model in the first phase and the usage of the model in the second phase to generate final outputs - next day load forecasts (Figure 4)

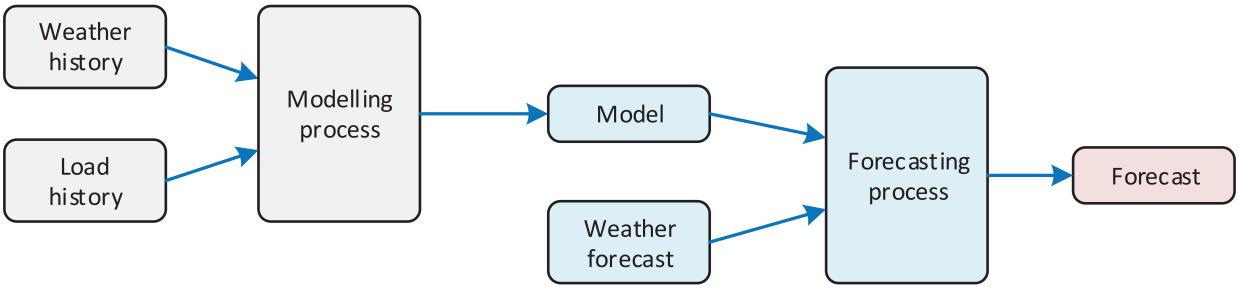


Figure 8. Load forecasting process diagram

With the purpose of obtaining best results different variations of input variables (Table I) can be considered. As was discussed some have more significant influence than the others and inclusion of some can improve the results. Furthermore pre-processing of the data was done. Loads and weather data for years 2012 and 2013 were obtained and prepared and as mentioned earlier part of it will be used for training part and part as a validation and testing set. Regarding the final implementation of neural network, it is necessary to pay attention to overfitting and the high computational effort. Overfitting happens when the error associated to the training set decreases in the same time as the error associated to the validation set increases. This means that the network over fits the training set which leads to a lack of generalization. Finally, weights are initiated randomly and backpropagation algorithm is used for the training and the network performance was tested on 3 different test sets.

From the used input data available 70% was used for training purposes (12264 samples), 15% (2628) or validation and 15% (2628) for testing of the developed artificial neural network. With these settings, the input vectors and target vectors are divided into three mentioned sets, 70% for training of the network where networks weights are adjusted, 15% to validate that the network is generalizing satisfactory and to stop training before overfitting (when generalization stops improving), 15% to test the network performance(this set has no effect on performance during training). The data set used in this paper is load data for years 2012 and 2013 of the MIBEL – Iberian power market. The neural network was trained using the neural network package in MATLAB 2013a [18].

The results (expressed as overall MAPE in Table III) of the artificial neural network design show the best performance of the test network 3. The validation performance of the network is shown on the Figure 9. It can be seen that after approximately 60 epochs the performance of the algorithm improves insignificantly and that it can be stopped.

Table III. Used variables/inputs for different test sets

No.	*Inputs*	*MAPE (%)*
Test 1	Time of day (hour)	2.67%
	Time of Week (day)	
	Time of Month (season)	
	Consumption (D-1)	
	Consumption (D-6)	
Test 2	Test 1 + Price (D-1)	2.68%
Test 3	Test 1 + Rainfall (D+1) + Temperature (D+1)	2.37%

The quality of neural fitting can be observed from the figure above (Figure 9). It can be seen that around 60 cycles/epochs are required for the algorithm to reduce the total amount of MSE to a point where no significant reductions are visible.

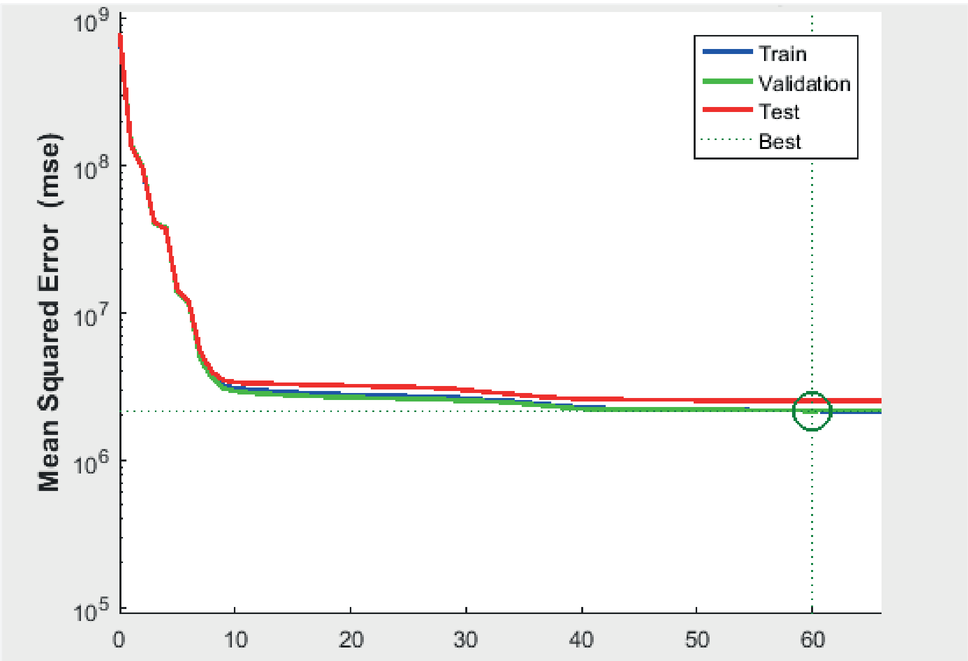


Figure 9. Ranges of error through neural network process epochs for different sets of data (training, validation, test)

The final forecast results are shown for test network 3 on the following figures ((Figure 10 and Figure 11).

As it can be seen from the comparison between historic data and output forecasts of the ANN for educational purposes it serves as a valid enough result since the curves align well with certain differences in forecasts of the peak values of some of the days (especially noticeable for Friday).

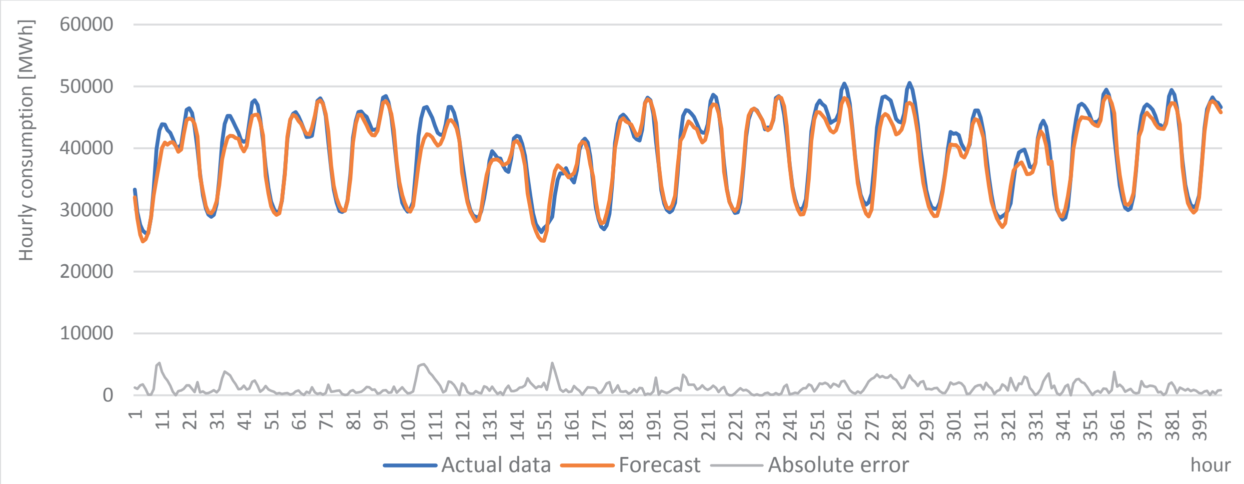


Figure 10. Comparison between actual data and ANN forecasts for a segment of test data set

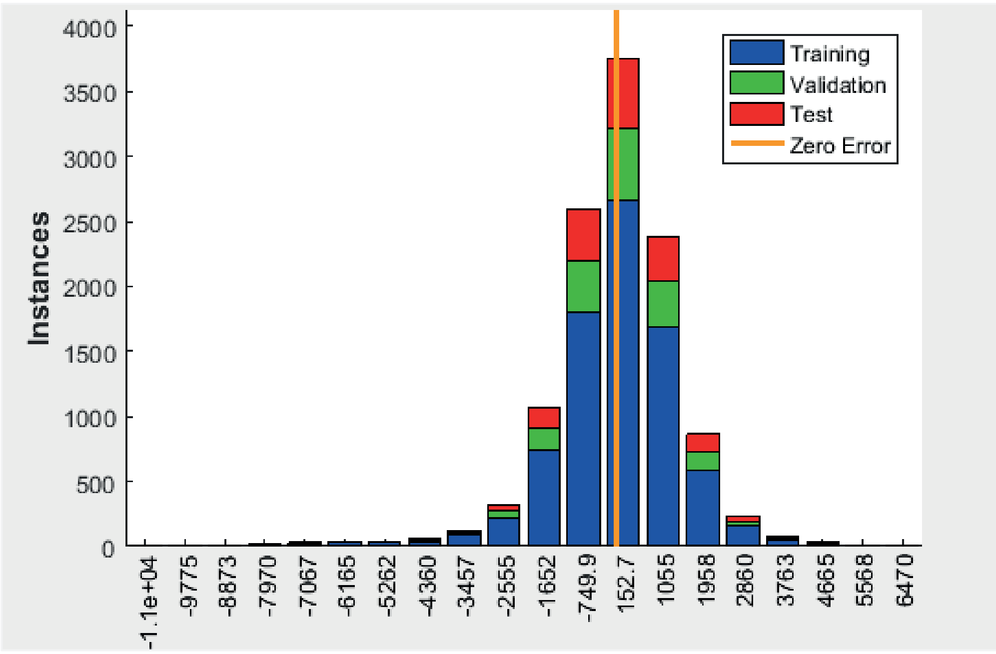


Figure 11. Error histogram showing error distribution between historic data and ANN outputs (Targets-Outputs)

5. CONCLUSIONS

In this paper, a feed-forward neural network model used for short term load forecasting is presented. This model makes use of past hours hourly consumptions, temperature, rainfall, day indices, weekend indices and month indices as the input variables. The MAPE of the load forecasted by this neural network was found to be between 2% and 3% depending on the selection of input variables. The test-case data was for MIBEL - Iberian power market. The result presented in this paper suggests that the use of carefully selected input historic data can have a positive impact on results. For example, the structure of the neural network can be improved upon to take into account the effects of holidays or usage of more input samples (more years of past consumptions/load).

6. ACKNOWLEDGEMENT

The work of the authors is a part of the H2020 project CROSSBOW – CROSS BOrder management of variable renewable energies and storage units enabling a transnational Wholesale market (grant no.773430). This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of authors and can under no circumstances be regarded as reflecting the position of the European Union. Additionally, this work has been supported in part by the Croatian Science Foundation under the project FENISG – Flexible Energy Nodes in Low Carbon Smart Network (grant no. IP-2013-11-7766).

7. REFERENCES

[1] E. Ilseven and M. Gol, "Medium-Term Price Forecasting Considering Demand and Supply Uncertainty", *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1-1, 2017.

[2] I. Kuzle, D. Bošnjak, H. Pandžić, “Comparison of Load Growth Prediction Methods in Distribution Network Planning”, *The 20th International Conference and Exhibition on Electricity Distribution*, Prague, Check Republic, Paper 0765, pp. 1-4, 2009.

[3] W. Taylor and P. E. McSharry, "Short-Term Load Forecasting Methods: An Evaluation Based on European Data", *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2213-2219, 2007.

[4] S. Sargunaraj, D. Gupta, and S. Devi, “Short-Term Load Forecasting for Demand Side Management,” *IEE Proceedings-Generation, Transmission and Distribution*, vol. 144, no. 1, pp. 68–74, 1997.

[5] Z. Yu, “A Temperature Match Based Optimization Method for Daily Load Prediction Considering DLC Effect”, *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 728–733, 1996.

[6] H. Steinherz, C. Pedreira, and R. Castro, “Neural Networks for Short- Term Load Forecasting: A Review and Evaluation”, *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, 2001.

[7] C. Y. Tee, J. B. Cardell and G. W. Ellis, "Short-term load forecasting using artificial neural networks," *41st North American Power Symposium*, Starkville, MS, USA, pp. 1-6, 2009.

[8] E. A Feinberg, D Genethliou, “Chapter 12 - Load forecasting”, *Applied mathematics for power systems*, State University of New York, Stony Brook, pp 269-282, 2016.

[9] D. Park, M. El-Sharkawi, R. Marks, L. Atlas, and M. Damborg, “Electric Load Forecasting Using an Artificial Neural Network,” *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 442–449, 1991.

[10] D. Bunn and L. Sch, “Forecasting Loads and Prices in Competitive Power Markets”, *Proceedings of the IEEE*, vol. 88, no. 2, pp. 163–169, 2000.

[11] E. Banda and K. A. Folly, "Short Term Load Forecasting Using Artificial Neural Network", *2007 IEEE Lausanne Power Tech*, Lausanne, pp. 108-112, 2007.

[12] I. Kuzle, T. Plavšić, S. Tešnjak, “Primjena neuronskih mreža za sekundarnu regulaciju napona i jalove snage”, *Journal of Energy*, vol. 47, pp. 133-140, 1998.

[13] N. Holjevac, T. Capuder, N. Zhang, I. Kuzle, C. Kang, “Corrective receding horizon scheduling of flexible distributed multi-energy microgrids”, *Applied Energy*, Volume 207, pp. 176-194, 2017.

[14] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, “One-Hour-Ahead Load Forecasting Using Neural Network”, *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 113–118, 2002.

[15] N. Kandil, R. Wamkeue, M. Saad, and S. Georges, “An Efficient Approach for Short Term Load Forecasting Using Artificial Neural Networks”, *International Journal of Electrical Power and Energy Systems*, vol. 28, no. 8, pp. 525–530, 2006.

[16] A. Al-Shareef, E. Mohamed, and E. Al-Judaibi, “One Hour Ahead Load Forecasting Using Artificial Neural Network for the Western Area of Saudi Arabia”, *International Journal of Electrical Systems Science and Engineering*, vol. 3, no. 1, pp. 35–40, 2008.

[17] MATLAB and Simulink User’s guide, R2013a, The MathWorks Inc., 2014.

[18] https://uk.mathworks.com, available online, accessed 6th Dec 2017.