COMPARATIVE ANALYSIS OF ARTIFICIAL NEURAL NETWORK MODELS FOR COMPLEX-VALUED LOAD PREDICTION

S. KRUNIC, Siemens d.o.o. Banja Luka, Bosnia and Herzegovina
I. R. KRCMAR, Faculty of Electrical Engineering, University of Banja Luka, Bosnia and Herzegovina

INTRODUCTION

Deregulated energy market puts a lot of pressure on every day operation of power utilities world wide. Increased energy demand and processes on the energy market require good planning of utility operation. Planning regards economic, as well as, technical issues. It should provide targeted financial margin and high availability of a power grid. Therefore, correct engagement of economical and technical resources, of power utilities, is very important. Due to these facts, short term load prediction is of great importance in deregulated energy market. Artificial neural network (ANN) based models are good choice for load prediction tasks, Bansal and Pandey (1), having in mind nonlinear and non stationary nature of a load signal Krunic (2), Khotanzad (3). Further, power utilities treat load information as a complex-valued signal. Thus, models for load prediction should be defined in complex domain.

Development of an ANN model in complex domain requires appropriate activation function (AF) at the neurons within the ANN. It would be good to have bounded and analytic complex function, on the whole complex plane C, as a neuron AF. However, the Liouville's theorem says that the only bounded and analytic function, on the complex set C, is a constant Mandic (4), Kim and Adali (5). Therefore, meromorphic functions are chosen as an AF within complex-valued ANNs. Meromorphic functions are analytic almost everywhere, with an exception of a discrete subset of the set C. These functions tend to infinity, at those singularities, thus avoiding the possibility of encountering essential singularities (4), (5).

There are reported results Krcmar (6), Krcmar and Krunic (7), where different neural adaptive structures were employed for complex-valued load prediction. In the paper (6), a simple, complex-valued, neural adaptive filter was applied to one step a head load prediction. Further, it recommends application of fully complex AFs. On the other hand, application of a multi layer perceptron (MLP), in short term complex-valued load prediction, was presented in (7). As in (6), load prediction depends only on the previous load data and AFs are fully complex (7).

MLP allows simultaneous processing of different data. So, each load data can be presented to a MLP with its time stamp. This fact can help to cope with non stationary load signals. Also, complex-valued time series prediction can be performed by application of dual univariate AF (DUAF). Within this setup,

real and imaginary parts of a signal are processed separately (4), (6). Through this approach, one can avoid problems associated with the choice of a complex-valued AF, and reduce complex-valued load prediction task to real-valued time series prediction. Thus, DUAF approach will give two predictors, with possible loss of information, due to decoupling of real and imaginary part of the load signal.
To this cause, we compare performance of two artificial neural network models. Load data, in both models, are presented to the ANN with its time stamp. The first model implements dual univariate approach, i.e. the model consists of two artificial neural networks. Training algorithm in this case is the backpropagation (BP) algorithm. The second model has fully complex AFs at network neurons, and it is trained by the complex backpropagation (CBP) algorithm. The analysis is carried out on the metered values of a complex-valued energy, in a distribution, medium voltage, grid.

STRUCTURE AND OPERATION OF MLP

The structure of a MLP, with one hidden layer, is given on Fig. 1, Haykin (8).

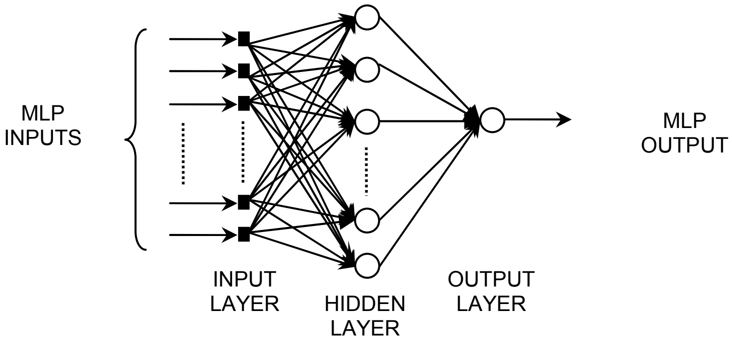

FIGURE 1 – STRUCTURE OF A MLP WITH ONE HIDDEN LAYER

Operation of the MLP, given on the Fig. 1, can be described by the following equations.

$$y(k) = \Phi(net_1(k)) \quad (1)$$

$$net_1(k) = \mathbf{x}_h^T(k)\mathbf{w}_h(k) \quad (2)$$

$$x_{hi}(k) = \Phi(net_{hi}(k)) \quad (3)$$

$$net_{hi}(k) = \mathbf{x}^T(k)\mathbf{w}_{ij}(k) \quad (4)$$

where $y(k)$ denotes MLP output, $\Phi(\bullet)$ denotes AF of the neurons, $net_1(k)$ is input to the MLP output neuron, $net_{hi}(k)$ is input to the i^{th} neuron in the hidden layer , k is a discrete time instant, $\mathbf{w}_h(k)=[w_{h1}(k), w_{h2}(k),\dots, w_{hN}(k)]^T$ is the vector of weights which connects neurons in the hidden layer and MLP output, $\mathbf{w}_{ij}(k)=[w_{ij1}(k), w_{ij2}(k),\dots, w_{ijM}(k)]^T$ is the vector of weights which connects MLP inputs and j^{th} neuron in the MLP hidden layer, N denotes number of neurons in the MLP hidden layer, M denotes length of the MLP input vector, $(\bullet)^T$ denotes vector transpose, $\mathbf{x}(k)=[x_1(k), x_2(k),\dots, x_M(k)]^T$ is the MLP input vector, and $\mathbf{x}_h(k)=[x_{h1}(k), x_{h2}(k),\dots, x_{hM}(k)]^T$ is vector that contains outputs of the neurons in the MLP hidden layer.

BACKPROPAGATION LEARNING ALGORITHM

The BP learning algorithm for a MLP, performs gradient descent search in the space of weights (8). Thus, it minimizes the criterion

$$J(k) = \frac{1}{2}\left|e(k)\right|^2 \quad (5)$$

where $|\bullet|$ denotes absolute value and

$$e(k) = d(k) - y(k) \quad (6)$$

defines error at the output neuron of MLP and *d(k)* denotes some teaching (desired) signal. Weight update is given by the following equation

$$w(k+1) = w(k) + \Delta w(k) \tag{7}$$

Computation of weight correction Δ*w(k)* depends on position of the weight in the network structure. Thus, for the weights, located between the hidden layer and the output, we have

$$\Delta w_{hi}(k) = \mu x_{hi}(k)(\Phi'(net_1(k)))e(k) \tag{8}$$

and in the case of the CBP learning algorithm it becomes

$$\Delta w_{hi}(k) = \mu x_{hi}^*(k)(\Phi'(net_1(k)))^* e(k) \tag{9}$$

where μ is the step size, (•)' denotes the first derivative and $(\bullet)^*$ denotes complex conjugate. For the weights between MLP input and hidden layer computation of the weight correction is as follows

$$\Delta w_{ijm}(k) = \mu x_m(k)(\Phi'(net_{hm}(k)))w_{hj}(k)(\Phi'(net_1(k)))e(k) \tag{10}$$

while for the CBP it becomes

$$\Delta w_{ijm}(k) = \mu x_m^*(k)(\Phi'(net_{hm}(k)))^* w_{hj}^*(k)(\Phi'(net_1(k)))^* e(k) \tag{11}$$

and $1 \leq j \leq N$. In order to stop the learning procedure it is necessary to define stopping criteria. Usual choices for the stopping criteria is minimum value of the index (5), averaged over the training set, flatness of the error surface, or combination of the previous ones.

EXPERIMENTAL ANALYSIS

Structure of the input data

Structure of the input data is of high importance for the overall performance of the ANN model. Within this paper we have adopted structure of the input load data as in (2), i.e. in order to produce load prediction at time instant (*k*+1), ANN model uses load data at time instants (*k*), (*k* -1), (k-2), (*k* -24), (*k* -25), (k-26), (*k* -168), (*k* -169), (k-170), (*k* -192), (k-193) and (*k* -194). Further, each load data is accompanied with its time stamp. Time stamp is presented to the ANN model through 12 inputs. Identification of the hour is given by 5 binary inputs, while identification of the day is given by 7 inputs. If we numerate days of the week by numbers 1 to 7, then we set appropriate input to 1, while the rest of 6 inputs we set to 0. The load data is average hourly normalized power. The signals were normalized with respect to the maximum of the load absolute value. Load data is metered at the 10 kV feeder, in the Transformer station Banja Luka 2. The normalized test complex-valued load signal is shown on the Fig. 2 and Fig. 3.

Training set and training procedure

The training set contains load data for the last two weeks, prior to the time instant that marks the first prediction. Within the experiments, logistic function was chosen as AF at the network neurons, i.e.

$$\Phi(z) = \frac{1}{1+e^{-\beta z}} \tag{10}$$

Gain of the AF was set to β=4. Number of the neurons in the hidden layer was N=6 and learning rate

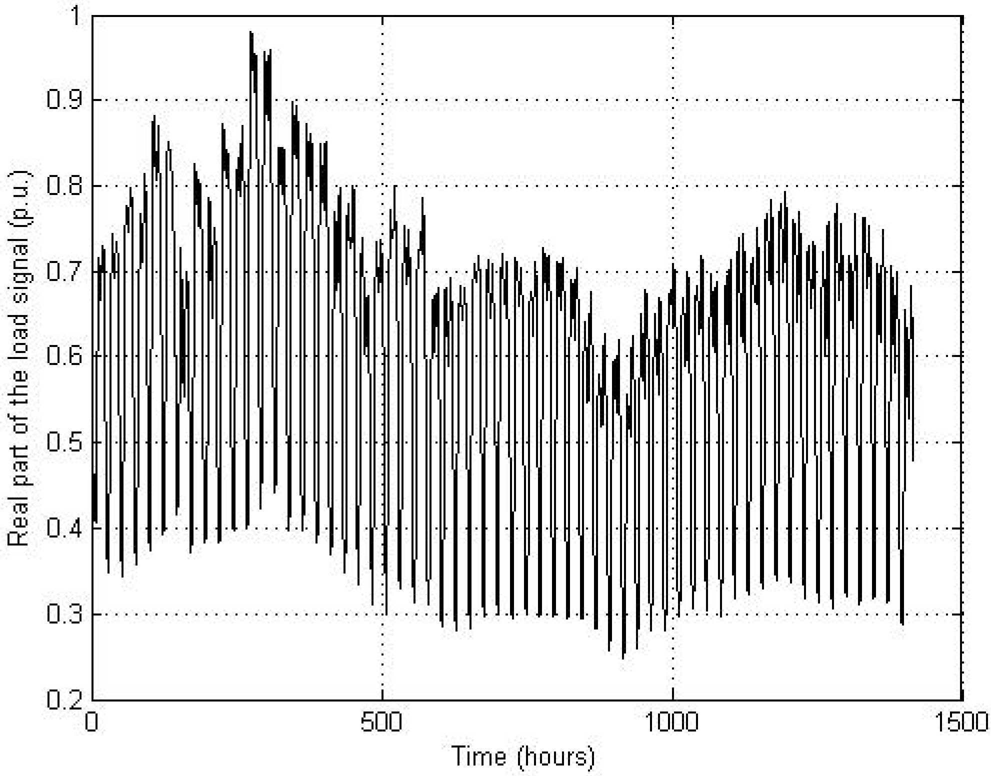

FIGURE 2 – REAL PART OF THE TEST LOAD SIGNAL

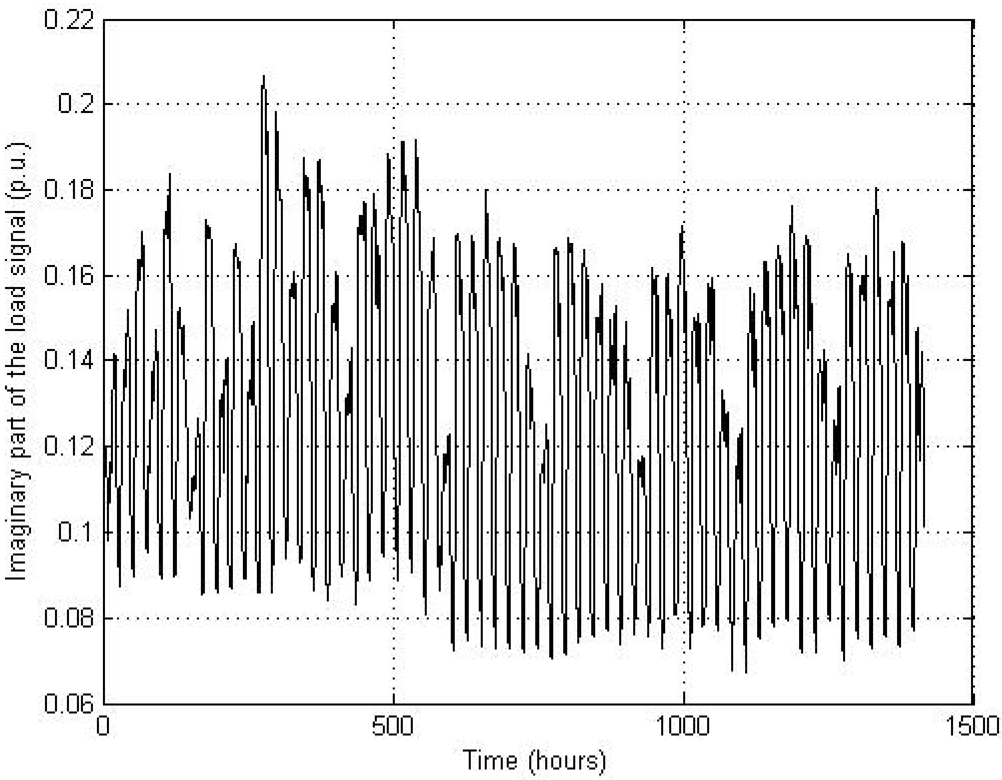

FIGURE 3 – IMAGINARY PART OF THE TEST LOAD SIGNAL

was μ=0.002. We applied modification of the BP algorithm to include momentum parameter, due to the fact that it stabilizes the learning process. The experiments were carried out with value of the momentum parameter α=0.8. The stopping criteria was defined by the minimum average value, over the epoch, of the index (5), J_{min}=0.0007.

Experimental results

The experiments were carried out, as short term complex-valued load prediction, through 100 independent runs, i.e. with 100 different sets of initial weights. Outputs of the network, i.e. predictions, after the training procedure, were fed back as network inputs in order to obtain next prediction. Prediction horizon was up to 48 hours. Average performance, of the analyzed ANN models, is shown on the Fig. 4 to 7.
Average absolute percentage error for the ANN model with fully complex AF, was 6.01%, for prediction horizon of 24 hours, and on prediction horizon of 48 hours, it was 6.62%. Regarding predictor based on the ANN model with DUAF, average absolute percentage error on the 24 hours prediction horizon was 7.43%, ad on the 48 hours prediction horizon it was 7.69%.

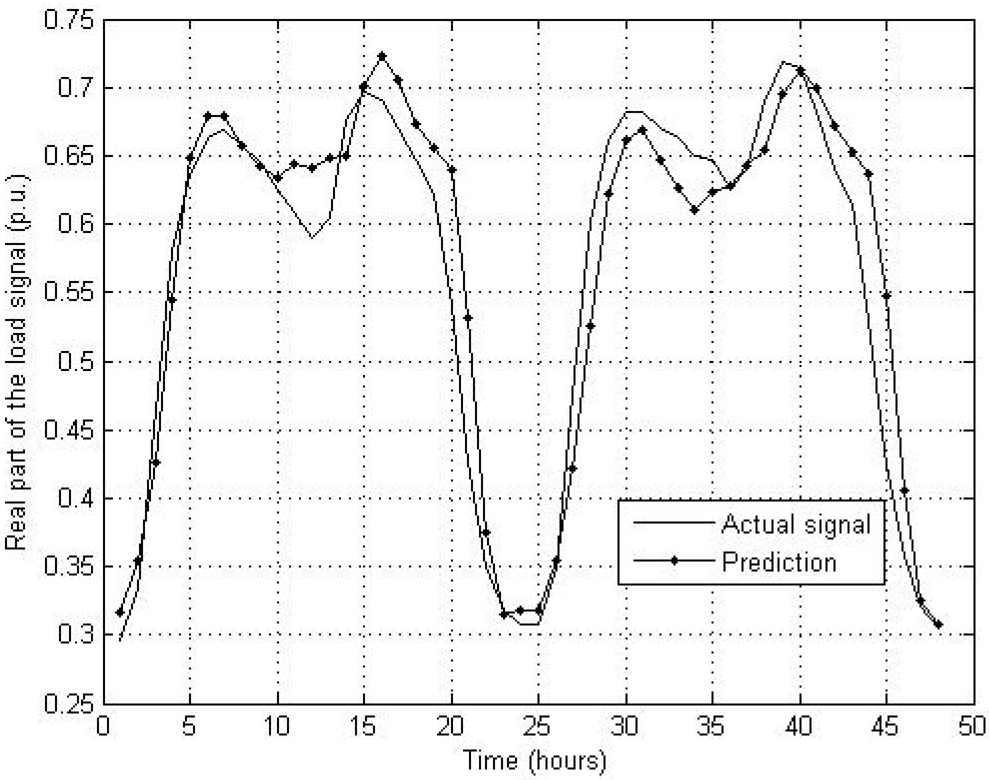

FIGURE 4– PERFORMEANCE OF THE ANN MODEL WITH FULLY COMPLEX AF, REAL PART OF THE LOAD SIGNAL

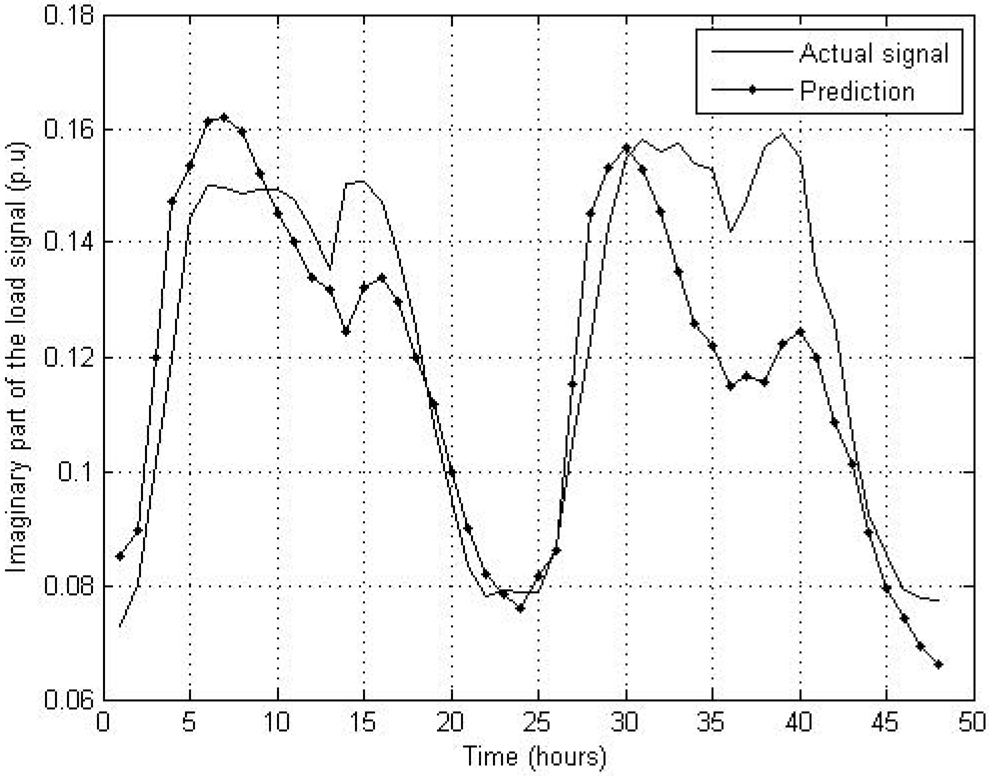

FIGURE 5– PERFORMANCE OF THE ANN MODEL WITH FULLY COMPLEX AF, IMAGINARY PART OF THE LOAD SIGNAL

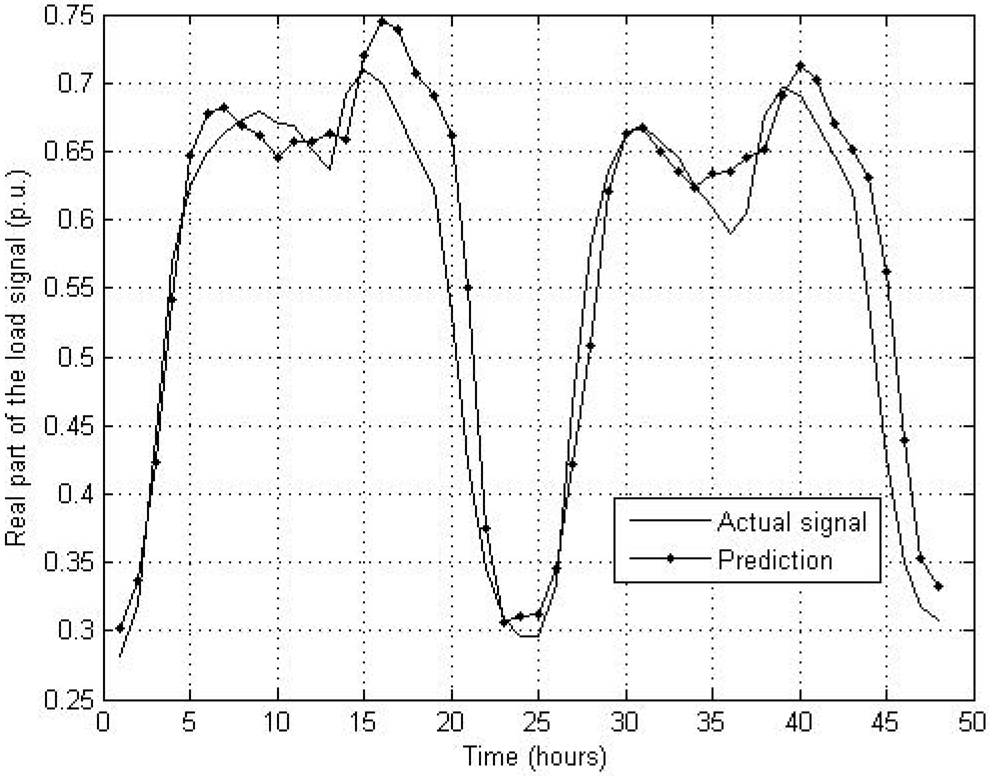

FIGURE 6– PERFORMANCE OF THE ANN MODEL WITH DUAF, REAL PART OF THE LOAD SIGNAL

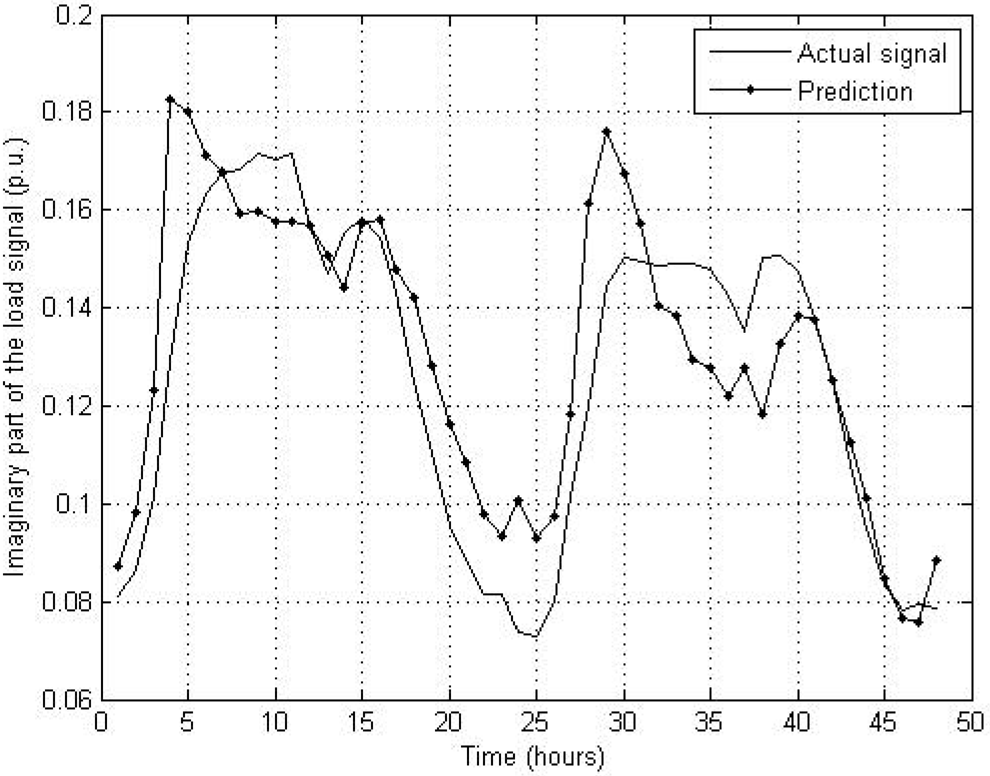

FIGURE 7– PERFORMANCE OF THE ANN MODEL WITH DUAF, IMAGINARY PART OF THE LOAD SIGNAL

CONCLUSION

Two ANN based models for complex-valued short term load prediction has been proposed and analyzed. The predictors have a MLP structure, with one hidden layer. The first model is fully complex, while the second one is based on the DUAF approach, i.e. it consists of two real-valued predictors. The predictions are based on the load data, accompanied with their time stamp, in order to cope with non stationary load signals. Backpropagation learning algorithm, with momentum parameter, has been employed for model training. Also, the structure of the predictor inputs, size of the training set and number of neurons in the hidden layer has been given. Performance of the predictor has been tested on the metered values of complex-valued load, at a medium voltage grid. Proposed fully complex ANN model shows better performance.

REFERENCES

1. Bansal R, C,Pandey J. C, 2005, “Load forecasting using artificial intelligence techniques: a literature survey”,” International Journal of Computer Applications in Technology”, “vol. 22, no. 2/3”, 109-119
2. Krunic S, Krcmar I. and Rajakovic N, 2000, “An improved neural network application for short-term load forecasting in power systems”,” Electric Machines and Power Systems”,”vol. 28”, 703-721
3. Khotanzad A, Rohani R. A, Lu T. L, Abaye A, Davis M. and Maratukulam D. J, 1997, ”ANNSTLF-A neural-network-based electric load forecasting system”,”IEEE Transactions on Neural Networks”, “vol. 8”, 835-846
4. Mandic D. P., Goh S. I, 2009, “Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models”, “John Wiley & Sons
5. Kim T, Adali T, 2003, “Approximation by fully complex multilayer perceptrons”, “Neural Computations”, “vol. 15(7)”, 1641-1666

6. Krcmar I. R, Maric P. S. and Bozic M. M, 2010, “A Class of Neural Adaptive FIR Filters for Complex-Valued Load Prediction,” , “Proceedings of 10th Symposium on Neural Network Applications in Electrical Engineering NEUREL-2010”, 37-40,
7. Krcmar I, Krunic S, 2012, “Artificial neural network for complex-valued load prediction”,” Energija, ekonomija, ekologija – ENERGETIKA 2012”,”br. 3-4”, 127-130
8. Haykin S, 1994, “Neural Networks, A Comprehensive Foundation, Architectures and Stability”, “Prentice Hall”