Principal Component Analysis based Strategy for Electrical Load Forecasting

Nikita Mittal[1], Akash Saxena[2]
[1]Department of Electrical Engineering, [2]Department of Electrical Engineering
Swami Keshvanand Institute of Technology Management & Gramothan, Jaipur
Email- [1]er.nikitamittal@gmail.com, [2]aakash.saxena@hotmail.com
Received 13 February 2016, received in revised form 27 March 2016, accepted 1 April 2016

Abstract: **Electrical load forecasting is a central and integral process for planning periodical operation and expansion of facilities in power system. Load demand patterns are getting complex with every passing day due to the deregulation of electricity markets. To develop a regression model which can map input and output variables in such a dynamic scenario is not easy. This paper presents an application of Principal Component Analysis (PCA) in the load forecasting of a large distribution network. Different models of ANNs are developed with the help of PCA based feature selection. Efficacy of the proposed approach is evaluated in terms of error indices namely Mean Square Error (MSE), Minimum Average Error (MAE), Mean Absolute Percentage Error (MAPE).**

Keywords: **Artificial Neural Network (ANN), Load Forecasting Mean Square Error, Principal Component Analysis.,Radial Basis Function Neural Network (RBFNN).**

1. INTRODUCTION

Modern power system has emerged as a complex interconnection of multiple utilities at transmission, distribution and generation ends. Load forecasting is an important exercise in planning. The reliable operation of the transmission as well as a generation grids depends upon the match of generation and load. Load forecasting is an important denominator used for planning.

In past, various methods have been applied by researchers in load forecasting. These methods were based on mathematical modeling [1,2], soft computing [3,4] and hybrid of different methods [5].

In [6] authors introduced random fuzzy logic neural network to forecast the load of micro grid. In the approach authors emphasized on the fuzzy and random nature of the load by modeled the same into random fuzzy variables. In [7] Random Weight Single Layer Feed forward Neural network (RWSLFN) utilized to forecast the load of Random Vector Function Link Network (RVFLN).Several Saptio temporal load forecasting technique proposed for residential units in [8]. Authors proposed decomposed forecast aggregate method to increase the efficiency of the forecast.

If the load forecasting is accurate the power system is economic and reliable. Electric utilities must have the information of load pattern in advance to meet the customer requirements. Accuracy of forecasting has very significant effect on power system, as forecasting errors can increase the risks of system breakdown [9].

Research works have been presented regarding electricity load prediction for micro-grids and buildings [10,11]. A few works have been published in the area of load forecasting. In [12] load forecasting of a residential area has been studied. In this reference Artificial neural network is compared with other models, including Grey model, regression model, polynomial method and polynomial regression model.

Load forecasting is based on several features like temperature, humidity, electricity price, dew point, hour of the day, day of the week, holidays and festivals. It is necessary to extract only those factors which are relevant for forecasting. Including all the factors into ANN will decrease the efficiency of network computing and cause large computational burden. Thus, the principal component analysis (PCA) is utilized to extract the necessary factors for forecasting [13].

Five features have been selected here for load forecasting i.e. dry-bulb, dew-point, wet-bulb, humidity and electricity price. Dry-bulb refers to the temperature of air measured by a thermometer freely exposed to the air but shielded from radiation and moisture. Dew point refers to the temperature at which the dew starts to form and is a measure of atmospheric moisture. Wet-bulb is the temperature a portion of air would have if it is cooled to saturation by the evaporation of water into it when the latent heat is supplied to it. Humidity is the amount of water vapor in the air.

In view of above discussion, the objectives of this paper are as follows:

1. To formulate different regression models on the basis of given data base of a distribution network.
2. To perform PCA for the intelligent feature selection and on the basis of PCA formulate supervised learning modules.
3. To present the comparison of the performances of supervised learning modules based on different standard error indices.

In the next section a brief description of different load forecasting methods are explained. Concepts of PCA are discussed in section III. In Section IV, the results and simulations are presented for different supervised learning modules. Highlights of this work along with future scope are presented in section V.

2. LOAD FORECASTING METHODOLOGY

2.1 Problem Formulation

In this work, short term load forecasting of few minutes ahead is under consideration. Prediction is done on the basis

that future load relies on the historical load data, temperature, price of electricity, humidity, hour of the day. In the modeling process, data are selected from Australian energy market operator[14]. A set of 2000 data is selected for training and a set of another 500 points is used for model validation.

2.2 Modeling using RBF and FFNN

2.2.1 Radial basis function for forecasting

Radial Basis Function neural network (RBFNN) is very useful methodology for time series data forecasting. RBF neural network can be used to analyze the relationship between a major sequence and other comparative sequences in a given dataset. RBFNN have fast training velocity, better approximation properties and can solve the local minima problem.

This paper uses RBFNN for load forecasting of electric power system. Architecture of a typical RBFNN is shown in figure 1. It is a multi-input, multi-output system consisting of an input layer, a hidden layer and an output layer. During the data processing, the hidden layer performs the non-linear transforms for the feature extraction and the output layer gives a linear combination of output weights.

In RBFNN, each hidden neuron computes a Gaussian function in the following equation.

$$b_j(\bar{x}) = \exp\left[\frac{-\left(\bar{x}^T - \overline{\mu_j}\right)^2}{2\sigma_j^2}\right], for \quad j = 1,2,......,q \quad (1)$$

where μ_j and σ_j ,are the centre and width of the Gaussian potential function of the j^{th} neuron in the hidden layer.

Each output neuron of RBFNN computes a linear function in the following form:

$$o_k = \sum_{j=1}^{q} w_{kj} b_j(\bar{x}) - \theta_k, \quad for \quad k = 1,2,........, m \quad (2)$$

Where o_k is the output of k^{th} node in the output layer, w_{kj} is the weight between j^{th} node in the output layer, $b_j(x)$ is output from the j^{th} node in the hidden layer, θ_k is bias of the k^{th} node in the output layer.

The RBF neural network based forecasting method has been successfully implemented for load forecasting. The RBF neural network models were developed for 10 min. ahead load forecasting. The input layer has 5 neurons for the five input features; the hidden layer has 10 neurons and the output layer has one neuron for the load forecasting.

An RBFNN is non-linear if the basis functions can move or change size or if there is more than one hidden layer.

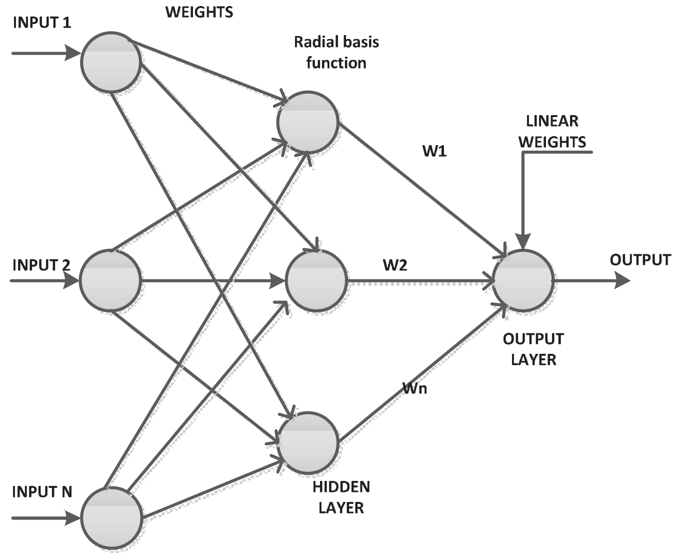

Figure 1 Architecture of RBF neural network

2.2.2 Feed Forward neural network for forecasting

A multi-layer feed-forward neural network (FFNN) consists of a layer of input layer, one or more hidden layers and a layer of output unit. The input layer consists of N_i inputs. Each i^{th} input is connected to j^{th} unit of hidden layer by a weighing factor, $W_{ij.}$ The output of each neuron is formulated as:

$$A_j = f_h\left(net_j\right) \quad (3)$$

Where f_h is a nonlinear activation function in the hidden layer and net_j can be formulated as :

$$net_j = \sum_{i=1}^{N_i} W_{ij} x_i + b_j \quad (4)$$

Where x_i is the input of unit i in the input layer; W_{ij} is the weighting factor between neuron i of the input layer and neuron j of the hidden layer and b_j is the bias term of the neuron. All the weighting factor and bias terms are adjusted during the training process.

The structure of FFNN output layer is similar to the hidden layer with the exception that the inputs of the output layer are the outputs of the hidden layer. The output of the neuron "k " in the output layer can be formulated as :

$$y_k = f_0\left(net_k\right) \quad (5)$$

Where f_0 is the non-linear activation function in the output layer, net_k is equal to:

$$net_k = \sum_{j=1}^{N_h} W_{jk} x_i + b_j \quad (6)$$

And W_{jk} is the weighting factor between neuron j in the hidden layer and neuron k in the output layer.

2.3 Evaluation criterion

For evaluating the performance of the model for load forecasting, define errors (e), the root mean square errors (RMSE), the mean absolute errors (MAE), and the mean absolute percentage errors (MAPE) as

$$e = \frac{[y_{act}(i) - y_{pre}(i)]}{y_{act}(i)} \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_{act}(i) - y_{pre}(i)\right)^2} \tag{8}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left| y_{act}(i) - y_{pre}(i)\right| \tag{9}$$

Where N=500 is the number of validation data, y_{act} is the real output and y_{pre} is the predicted output in this paper.

2.4 Description of supervised learning models

Supervised learning consists of modeling, on the basis of finite set of observations, the relation between a set of input variables and one or more output variables, which are somewhat dependent on the inputs. The general approach to model an input/output phenomenon, with a scalar output and vectorial input, relies on the availability of a collection of observed pairs typically referred as training set.

3. PCA-BASED FEATURE SELECTION

In this study, Principal Component Analysis (PCA) is investigated to reduce the dimensions of the input features. PCA performs an orthogonal linear transformation to the basis of correlation eigenvectors and projects onto the subspace spanned by those eigenvectors corresponding to the largest eigen values. Normally most of the original information is retained by the first PCs, reducing the dimensions of input features drastically. It would be extremely time consuming to train a NN with all input data as it would increase the dimension of the NN and make its response slow during testing. Therefore we select the optimum number of inputs by eliminating those input features that have no significant effect on its output. In this paper drybulb, dewpoint, wetbulb, humidity and electricity price are taken as input features. In using PCA, the initial feature matrix X is defined as below:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,k} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & x_{2,k} & \cdots & x_{2,N} \\ x_{3,1} & x_{3,2} & x_{3,k} & \cdots & x_{3,N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{M,1} & x_{M,2} & x_{M,k} & \cdots & x_{M,N} \end{bmatrix} \tag{10}$$

Where M is the number of observations.

$x_1, x_2, x_3, x_k \cdots, x_N$ Corresponds to the input features. The normalized data X is calculated as

$$x_{jk} = \frac{x_{jk} - \overline{x_k}}{\sqrt{\sigma_k}} \tag{11}$$

$$where \quad \overline{x_k} = \frac{1}{M}\sum_{j=1}^{M} x_{jk} \tag{12}$$

,

$$\sigma_k = \frac{1}{M-1}\sum_{j=1}^{M}\left(x_{jk} - \overline{x_k}\right)^2, \ \mathrm{k} = 1,2,..,\mathrm{N} \tag{13}$$

Thus, the covariance matrix of the feature matrix is calculated as

$$S = \begin{bmatrix} s_{1,1} & s_{1,2} & s_{1,3} & \cdots & s_{1,20} \\ s_{2,1} & s_{2,2} & s_{2,3} & \cdots & s_{2,20} \\ s_{3,1} & s_{3,2} & s_{3,3} & \cdots & s_{3,20} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_{20,1} & s_{20,2} & s_{20,3} & \cdots & s_{20,20} \end{bmatrix} \tag{14}$$

Where

$$s_{jk} = \frac{1}{M-1}\sum_{i=1}^{M}\left(x_{ij} \cdot x_{ik}\right) \tag{15}$$

$j, k = 1, 2, \ldots, 5.$

In general, the dimensionality of the dataset is reduced to the number of significant eigen values of the covariance matrix.

Let $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_{20}$ be the Eigen values of

S, the contribution rate of i^{th} principal component is given by

$$\lambda_i' = \frac{\lambda_i}{\sum_{i=1}^{20} \lambda_i} \quad (16)$$

The contribution rates of these principal components are shown in Table 1.

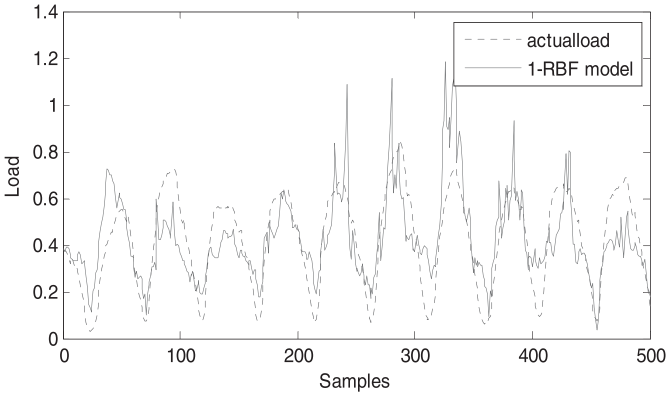

Fig.2 Actual and forecasted load patterns for 1-RBF model

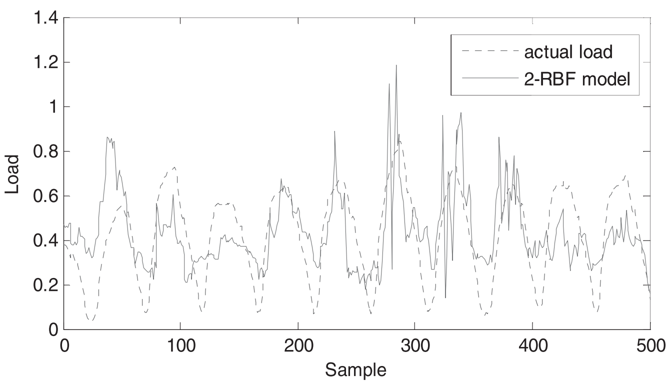

Fig.3 Actual and forecasted load patterns for 2-RBF model

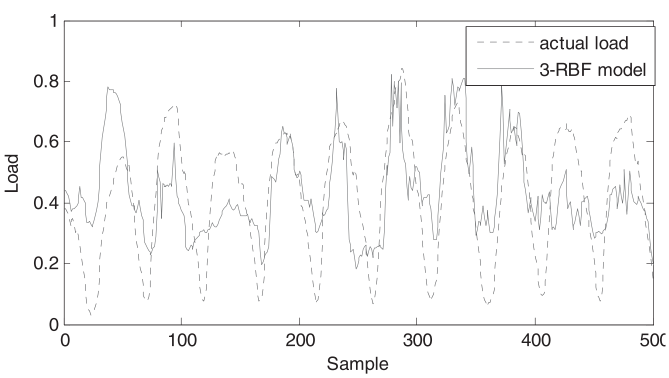

Fig.4 Actual and forecasted load patterns for 3-RBF model

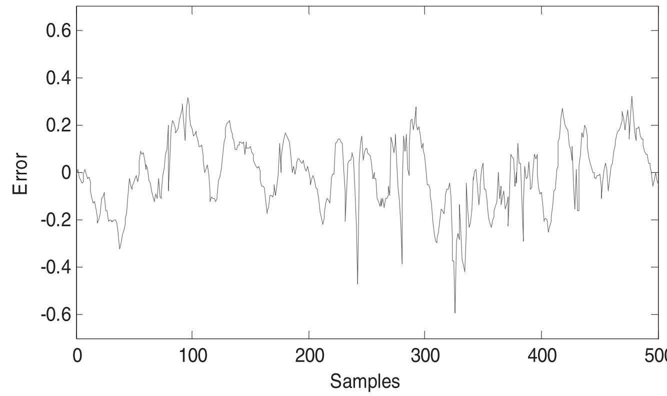

Fig 5 Error of 1-RBF model

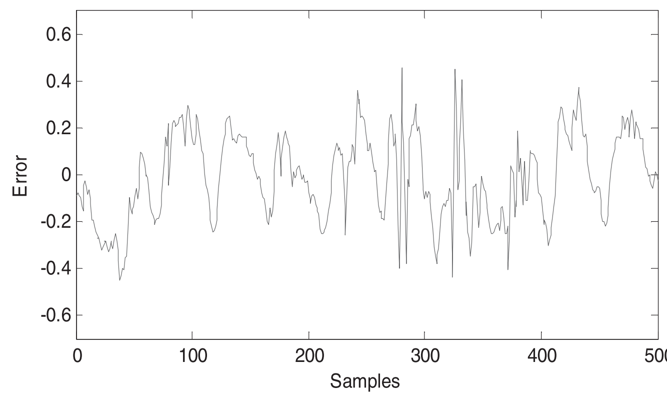

Fig 6 Error of 2-RBF model

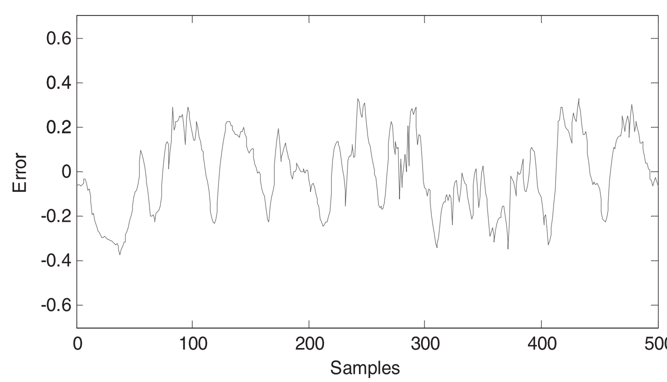

Fig 7 Error of 3-RBF model

4. SIMULATION RESULTS

With this work two simulation models using RBFNN and FFNN are established initially. Comparisons were then made among the proposed networks with the same input/output data. Figure 8 and 9 shows the model output and plant output over the validation. Best results are obtained with RBFNN which can be observed in the figure as the errors are reduced. From calculation of errors it can be observed that accuracy of RBFNN is higher than FFNN as shown in table 2.

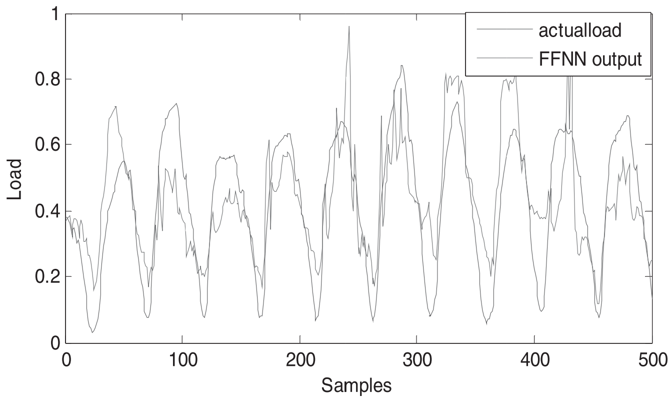

Fig 8. FFNN real and predicted output

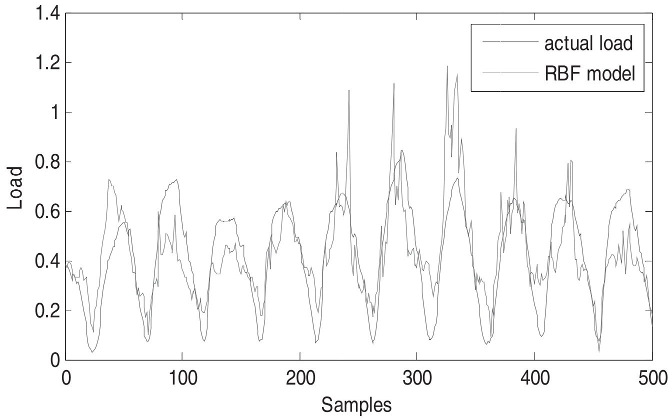

Fig 9. RBFNN real and predicted output

Initially the networks are trained with all the five inputs. PCA is performed for both networks and contribution rates of these components are calculated. Table below shows the eigen values and contribution rates of every component. It is observed that Dry bulb has the highest contribution rate.

Table 1 Eigen values and contribution rates of components

Principal components	Eigen values	Contribution rates %
Drybulb	2.2576	0.4515
Dewpoint	1.7505	0.3501
Wetbulb	0.9566	0.1913
Humidity	0.0011	2.25E-04
Electricity Price	0.0342	0.0068

Table 2 Errors of FFNN and RBFNN

MODEL	MAE	RMSE	MAPE
FFNN	0.1198	0.1433	0.5183
RBF	0.1176	0.1452	0.4779

As the contribution of last two components is very less they only contribute to computational burden and decrease the efficiency. Now taking the first three features into consideration PCA is performed for RBFNN. Table shows the MAE and RMSE for the RBFNN. Figure – shows the real and predicted outputs for RBFNN. Figure – shows the modelling errors. It is clear after applying PCA that first factor i.e. Dry bulb plays the dominant role as its contribution is highest.

Table 3 Errors in different RBF models

MODEL	MAE	RMSE
1-RBF MODEL	0.1378	0.1634
2-RBF MODEL	0.1438	0.1682
3-RBF MODEL	0.1446	0.1711

In 1-RBF Model only dry bulb is chosen as the input variable and its contribution rate is highest. Least error is produced with 1-RBF Model. In 2-RBF Model dew point and electricity price are chosen as input parameters. As contribution rate of electricity price is very low in forecasting, it introduces errors in the network. In 3-RBF Model dry bulb, wet bulb and electricity price are chosen as input variables and errors are calculated.

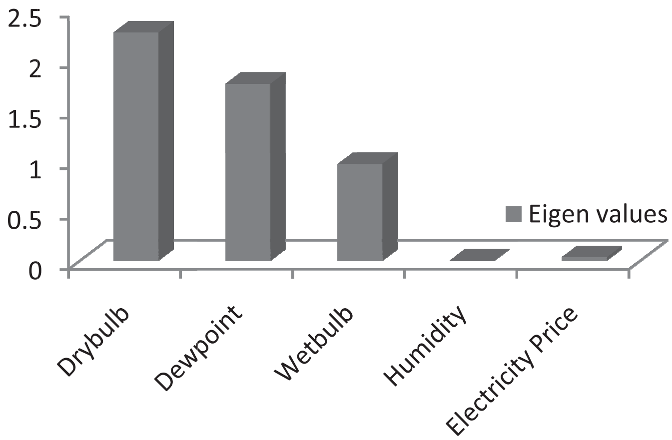

Fig 10. Eigen values of features

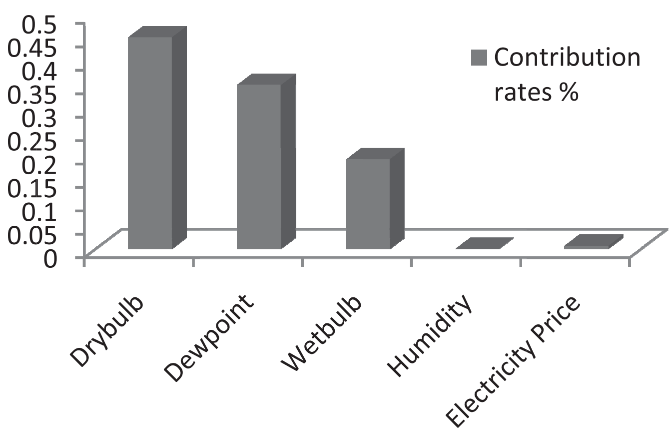

Fig 11. Contribution rates of features

5. CONCLUSION

This paper presents and efficient methodology based on supervised learning models to forecast the electrical load. Principal component analysis is used to obtain the major contributing features for building the supervised learning algorithm. Results from RBFNN are promising. Following are the major contribution of this manuscript:

a. Principal Component Analysis is used to reduce the size of input feature matrix. Major principal components are temperature, historical data of electricity price and weather conditions. Contribution rates are shown in fig.11

b. Developed regression models are based on ANNs (RBFNN and FFNN). Efficacy of the proposed model is tested by standard error indices (MAE, RMSE and MAPE).

Application of Design of Experiment based regression models for price forecast with supervised learning models lays in future scope.

REFERENCES

[1] V. Vahidinasab, S. Jadid, A. Kazemi, Day-ahead price forecasting in restructured power systems using artificial neural networks, Electr. Power Syst. Res. 78 (8)(2008) 1332–1342.

[2] N. Ding, C. Benoit, G. Foggia, Y. Besanger, F. Wurtz, Neural network-based model design for short-term load forecast in distribution systems, IEEE Trans. Power Syst. (2015) (in press).

[3] D. Ranaweera, N. Hubele, G. Karady, Fuzzy logic for short term load forecasting,Int. J. Electr. Power Energy Syst. 18 (4) (1996) 215–222.

[4] G. Li, J. Shi, On comparing three artificial neural networks for speed forecasting, Appl. Energy 87(7) (2010) 2313-2320.

[5] Al-Hamadi H, Soliman S. Short Term Electric Load Forecasting based on Kalman Filtering Algorithm with moving window weather and load model. Electr Power syst Res. 2004; 68(1):47-59.

[6] Chin Wang Lou, Ming Chui Dong "A novel random fuzzy neural networks for tackling uncertainties of electric load forecasting", Electrical Power and Energy systems, vol. 73,December 2015.

[7] Ye Ren, P.N. Suganthan, N. Srikanth, Gehan Amaratunga, Random Vector Functional Link Network for Short-term Electricity Load Demand Forecasting, Information Sciences, Available online 6 January 2016

[8] Akin Tascikaraoglu, Borhan M. Sanandaji, Short-term residential electric load forecasting: A compressive spatio-temporal approach, Energy and Buildings, Volume 111, 1 January 2016, Pages 380-392.

[9] A.S. Khwaja, M. Naeem, A. Anpalagan*, A. Venetsanopoulos, B. Venkatesh, Improved short term loas forecasting using bagged neural networks. Electr power system research, volume 125, 2015 pages 109-115, ISSN 0378-7796.

[10] G. Escriva-Escriva, C. Alvarez-Bel, C. Roldan-Blay, M. Alcazar-Ortega, New arti-ficial neural network prediction method for electrical consumption forecastingbased on building end-uses, Energy Build. 43 (2011) 3112–3119.

[11] S. Farzana, M. Liu, A. Baldwin, M.U. Hossain, Multi-model prediction and sim-ulation of residential building energy in urban areas of chongqing, south westchina, Energy Build. 81 (2014) 161–169.

[12] S. Farzana, M. Liu, A. Baldwin, M.U. Hossain, Multi-model prediction and sim-ulation of residential building energy in urban areas of chongqing, south westchina, Energy Build. 81 (2014) 161–169.

[13] Xiaobing Kong, Xiangjie Liu, Ruifeng Shi , Kwang Y. Lee. Wind speed prediction using reduced support vector machines with feature selection.

[14] http://www.iso-ne.com/

◆ ◆ ◆