World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:8, No:11, 2014

Evaluation of Short-Term Load Forecasting Techniques Applied for Smart Micro Grids

Xiaolei Hu, Enrico Ferrera, Riccardo Tomasi, Claudio Pastrone

***Abstract*—**Load Forecasting plays a key role in making today's and future's Smart Energy Grids sustainable and reliable. Accurate power consumption prediction allows utilities to organize in advance their resources or to execute Demand Response strategies more effectively, which enables several features such as higher sustainability, better quality of service, and affordable electricity tariffs. It is easy yet effective to apply Load Forecasting at larger geographic scale, i.e. Smart Micro Grids, wherein the lower available grid flexibility makes accurate prediction more critical in Demand Response applications. This paper analyses the application of short-term load forecasting in a concrete scenario, proposed within the EU-funded GreenCom project, which collect load data from single loads and households belonging to a Smart Micro Grid. Three short-term load forecasting techniques, i.e. linear regression, artificial neural networks, and radial basis function network, are considered, compared, and evaluated through absolute forecast errors and training time. The influence of weather conditions in Load Forecasting is also evaluated. A new definition of Gain is introduced in this paper, which innovatively serves as an indicator of short-term prediction capabilities of time spam consistency. Two models, 24- and 1-hour-ahead forecasting, are built to comprehensively compare these three techniques.

***Keywords*—**Short-term load forecasting, smart micro grid, linear regression, artificial neural networks, radial basis function network, Gain.

I. Introduction

LOAD Forecasting (LF) is becoming a key feature for the electricity distribution industry in the Smart Grid age. In a scenario where electricity prices change dynamically and Renewable Energy Sources (RES) put the stability of the grid at stake, LF is an essential tool for energy utilities to organize operations and to support the planning of investments on electric power generation, infrastructure development and financial purchasing. LF is even more valuable in scenarios where Demand Response (DR) techniques are applied. LF tools can in fact enable the devices or customers to learn their operating decisions based on the load prediction for next hours or days, thus enhancing the DR mechanisms ability to shift reliably energy usage and shave load profiles. Short-term load forecasting (STLF) is based on statistical procedures which use past load and exogenous variables such as weather related variables to forecast one hour up to one day energy consumption. A large variety of statistical and artificial intelligence techniques have been developed for STLF, for instance, regression methods [1], [2], neural networks [3]-[6], radial basis function network [7]-[10], expert systems [11]-[14], and fuzzy logic [11], etc. Progress in LF can be achieved by providing STLF with probability distributions and the further direction should be artificial intelligence techniques with better understanding of the load dynamics and proper models [15], [16].

LF techniques can be applied by aggregating data at different spatial scales, which are e.g. single or few energy-consuming components or devices [17], all loads within a household, a building [18], [19], or the entire sections of the grid [20]. Due to this multi-scale applicability they are especially useful in multi-scale energy systems such as the one proposed by the GreenCom project [21], [25]. According to the GreenCom concept, currently being deployed in a real small-size pilot with actual users, loads information collected from single loads and households can be progressively combined, forecasted, and possibly controlled at different scales to pave the way towards (Virtual) Smart Micro Grids, which can aggregate loads, storage as well as generation capacity and act as a single cooperating entity towards the remaining of the Smart Grid.

The purpose of this paper is to analyze and evaluate load forecasting in such a concrete scenario comparing three different STLF techniques and the influence of weather conditions on forecasting process. The considered techniques are firstly introduced in Section II. Then programming methodologies including data preparation and forecast models are presented in Section III. Section IV describes the evaluation of those abovementioned three techniques and compares the results while conclusions and future works are drawn in Section V.

II. Techniques for Short-Term Load Forecasting

A. Linear Regression Method

Linear regression (LR) method is one of the most extensively used techniques for STLF. It analyzes the relationship between continuous dependent variables and one or more explanatory variables and uses the technique of weighted least-squares estimation to compute the regression coefficients according to the amount of historical data. The following model is applied for this analysis wherein t is the sampling time, y the total load, v the vector of adapted variables such as time, temperature, humidity, wind speed, etc., α vector of regression coefficients:

$$y_t = v_t\, \alpha_t + \varepsilon_t \tag{1}$$

The coefficients in the forecast model are calculated from the latest actual data prior to the forecasting day. Current observations are more important for the forecasting due to the

Xiaolei Hu, Enrico Ferrera, Riccardo Tomasi, and Claudio Pastrone are with the Istitutto Superiore Mario Boella, Via P.C. Boggio, 61, Turin, Italy (e-mail: {hu, ferrera, tomasi, pastrone}@ismb.it).

variability of load characteristics. The accuracy of forecast using regression models depends on the precision with which the regression function fits the data. Consequently, a pre-analysis of the load is essential for LR if faithful result of the analysis is expected.

B. Artificial Neural Networks

In the past two decades, there has been a great interest in the field known as artificial intelligence (AI) as it offers powerful and flexible methods for obtaining solutions to problems eluding traditional methods. Artificial neural networks (ANNs) and expert system are two major branches of AI. ANNs have been proven as a promising alternative to solve complex problems [3], [4], [15].

A feed-forward ANN is a supervised network organized in layers, which can have any number of layers, units per layer (neurons), inputs and outputs. Each single neuron is connected to other ones of the previous layer through adaptable weights. The neuron receives information through a number of input nodes, processes it internally where weights are adjusted so that the network attempts to produce the desired output, and eventually generates a response. Fig. 1 shows the architecture of a typical multilayer feed-forward neural network.

For processing of ANNs, the input values are linearly combined in the first stage. Then in the second stage, the result is applied as the argument of a nonlinear activation function such as bounded sigmoid functions.

When the network starts to be trained, all the information is supplied to it as a data set. After reading each pattern, the network produces an output by using the input data and then compares it with the training pattern. For any possible difference, the weights are changed to abate the error accordingly. The computation stops until all the errors are under the desired tolerance. The mostly implemented training algorithm is back-propagation (BP) which uses a steepest-descent technique based on the computation of the gradient of the loss function, changing the weights along its gradient, reducing the total error and improving the performance of the neural networks.

To design an ANN, the first step is to select an appropriate architecture, e.g. the multilayer perception (MLP) which is the most popular neural network with multiple hidden layers. Subsequently, the number of hidden layers, input nodes, neurons per layer, and the type of activation function should be determined.

C. Radial Basis Function Network

The radial basis function network (RBFN) is another kind of feed-forward ANN which is simple yet auspicious thanks to the utilization of extensional learning and high computing speed. An RBFN consists of an input layer, a hidden layer, and a linear output layer. The input layer determines the Euclidean distance amongst the input vector and the weight vectors of the hidden layer that is composed of units with Gaussian transfer function (radial bases) whose weight vectors form a vectored quantization of the input space. The output weights are linearly combined while the hidden layer utilizes the nonlinear transformation for feature extraction during the data processing.

Fig. 2 illustrates its structure. The learning process can be divided into two stages. First, the weights of the hidden layer are calculated by clustering techniques or are randomly assigned in the input space. Alternatively, an optimal method can be employed such as orthogonal least squares algorithm. Second, the weights of output layer are computed by linear regression method. The main difference between MLP and RBFN is the absence of hidden layer weights. The parameters adjusted in the learning process are only the linear mapping from hidden layer to output layer. So it is easier to interpret the hidden layer than the ones in an MLP.

The RBFN requires more neurons than standard BP network does, but it is optionally sub-dividable into parallel-training fractions of time which is comparatively shorter than it takes to train a BP network.

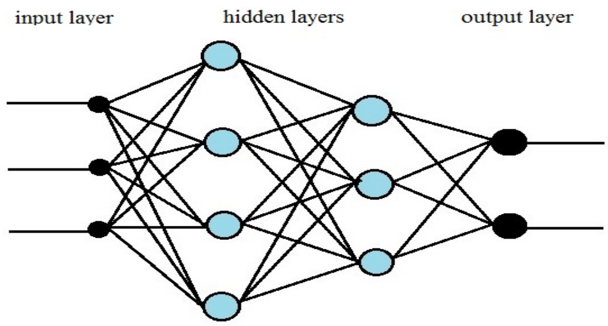

Fig. 1 Architecture of multilayer feed-forward neural network

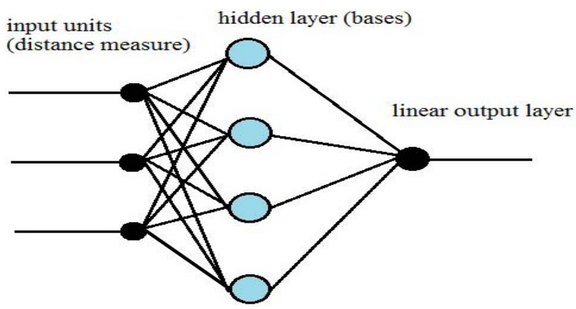

Fig. 2 Structure of a radial basis function network

III. PROGRAMMING METHODOLOGIES

A. Data Preparation

As it has been argued that the quality of the input data to the ANNs may strongly affect the results [22]-[24], data preparation in ANNs modeling is a critical step, especially for models with complex data analysis.

In the current work, the data were collected from single loads and households of GreenCom project from 2014/1/1 to 2014/6/30 and were prepared to comprise six different attributes,

1) L: peak load,
2) d: days of the week (Monday to Sunday),
3) h: hours of the day (0 to 24),

4) f: a flag indicating whether it is a holiday (0 indicates holiday and 1 indicates non holiday)
5) T: temperature,
6) H: humidity.

B. Forecasting Model and Error Analysis

A variety of papers is dependent solely on one model which is 24-hour-ahead forecasting to investigate the techniques for STLF [2], [4], [6]-[9], [11]. In order to more comprehensively compare the three techniques employed in this paper, a 1-hour-ahead forecasting is included in addition to the 24-hour one.

The 24-hour-ahead forecasting model is composed of all load variables from each of the previous 24-hour load and other five attributes described before and the model is like:

$$L(t)=F[L(t-24), L(t-25), ..., L(t-48), h(t), d(t), f, T, H] \quad (2)$$

The 1-hour-ahead forecasting model is similar with 24-hour-ahead forecasting model instead the load variables are from each of the previous one hour load:

$$L(t)=F[L(t-1), L(t-2), ..., L(t-24), h(t), d(t), f, T, H] \quad (3)$$

The forecasting error is measured by the Mean Absolute Percentage Error (MAPE) since it is the industry standard accepted for examining load forecasting performance. It is explained in the equation below where y is the actual load, yˆ estimated value, and N the number of observations,

$$MAPE\ \% = \frac{1}{N}\sum_{i=1}^{N}[abs(\hat{y}-y)/y] \times 100 \quad (4)$$

C. Forecasting Model and Error Analysis

The current paper introduces a new definition of Gain, which innovatively serves as an indicator of short-term predication capabilities, as larger gain implies more similar MAPE values for different forecasting time length (i.e. 1 hour and 24 hours) and hence the forecasting technique is more consistent in terms of time spam within the short-term range.

The Gain is defined as the ratio between the MAPEs of N-hour-ahead forecasting ($N \leq 24$) and 24-hour-ahead forecasting using the same technique,

Gain=[MAPE(N-hour-ahead)]/[MAPE(24-hour-ahead)](5)

In this paper, except 1-hour and 24-hour-ahead forecasting, we also simulate forecasting with other time length (i.e. 2 hours, 6 hours, and 12 hours) to have a good grasp of the gain.

IV. Results and Comparison

A. 1-Hour and 24-Hour-ahead Forecasting

For these two forecasting models, one month data (from 2014/05/01 to 2014/05/31) are acquired for training and the loads information collected on June 9th are forecasted using all three techniques.

For load forecasting, there are three vital quantities:

i. the load shape which is measured by examining the error in each hour of a day
ii. the daily peak load which refers to the highest demand for the day
iii. daily energy which is the sum of all the hourly daily loads

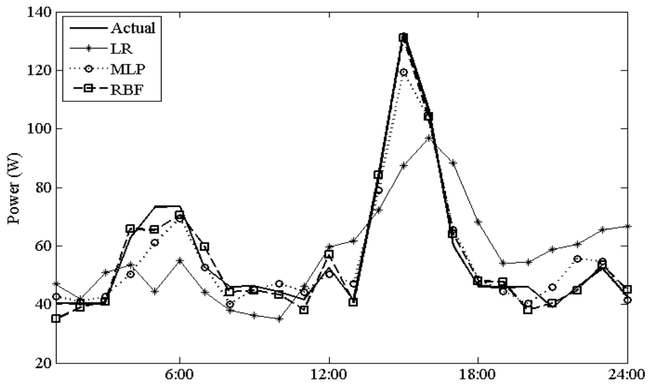

Fig. 3 Forecasting values of 24-hour-ahead loads by three techniques on June 9th

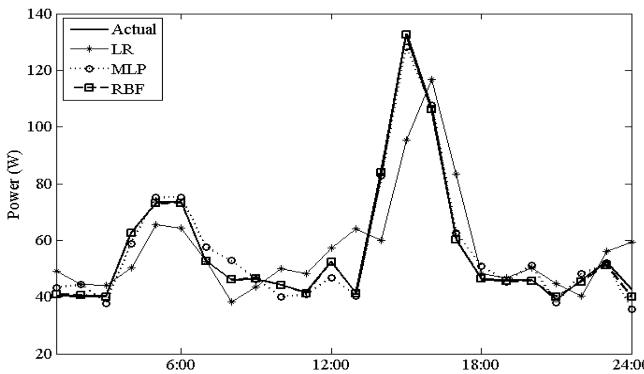

Fig. 4 Forecasting values of 1-hour-ahead loads by three techniques on June 9th

All these three quantities are presented in the figures below. Results are examined by plotting the actual loads and the forecast values, e.g. Fig. 3 illustrates the results for 24-hour-ahead forecasting on June 9th wherein the solid line represents the actual loads while the solid one with star the forecast value by using LR, the dashed one with circle the forecast value by MLP, and the dashed one with square the forecast value by RBF. Fig. 4 shows the 1-hour-ahead forecasting of the same day by three techniques.

For error analysis, MAPEs presented in Table I indicate that RBFN improves prediction accuracy in comparison with other two techniques. The main limit of LR method, which gives unsatisfactory results as MAPE=21.87% for 1-hour-ahead forecasting, and 32.55% for 24-hour-ahead forecasting, lies in the linear combination of the time series while the collected electric power signal has strong nonlinear behavior. Therefore it is unavailing to use LR method to forecast such loads. Instead, a nonlinear combination, as in MLP and RBF, possesses higher flexibility, which attributes to the improvement of results. Compared MLP with RBF, RBF has the smaller MAPEs than MLP does for both two models. Since

RBFN uses more hidden units than MLP does, it can be concluded that the utilization of more hidden units can improve the result.

TABLE I
MAPES(%) VALUES OF FORECASTING 1-HOUR AND 24-HOUR-AHEAD LOADS BY THREE TECHNIQUES

Model	LR	MLP	RBFN
1-hour	21.87	7.04	3.12
24-hour	32.55	19.34	7.92

In addition, another most important independent variable for load forecasting is weather information, which includes temperature, humidity, precipitation, wind speed, cloud cover, etc; especially temperature and humidity will affect the load prediction accuracy. In our models, temperature and humidity are also taken into consideration as one of the input variables. Fig. 5 shows the 1-hour-ahead forecasting values by using RBF with and without temperature and humidity as input variables wherein the solid line is the actual load, the dashed one the forecast value without temperature and humidity, and the dashed one with star the forecast value with those two parameters. It is observed from Table II that by adding temperature and humidity as input variables, the MAPE is reduced from 3.93% to 3.12%, thus temperature and humidity play a vital role in accurate load prediction.

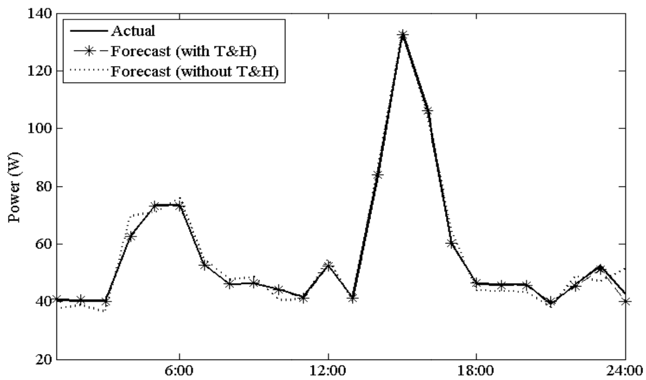

Fig. 5 Comparison of 1-hour-ahead forecasting with and without temperature and humidity by RBF on June 9th

TABLE II
MAPES(%) VALUES OF FORECASTING 1-HOUR-AHEAD LOADS BY RBF

	with T and H	without T and H
MAPE (%)	3.12	3.93

B. Training Time

In this work, we use BP as the training algorithm for MLP while for RBF we use unsupervised training algorithm which is kernel k-means clustering algorithm. Fig. 6 shows the recognition errors during the training process for MLP and RBF, which are plotted as a decreasing function of the number of epochs. For the first 100 epochs, RBF achieves faster convergence than MLP does. MLP with BP training algorithm gets trapped into local minima so the error does not decrease with further iteration and in this case we need to re-initialize the optimization of parameters during the training process. On the contrary, this disadvantage is absent in RBF networks.

The training time of 24- and 1-hour-ahead forecasting with MLP and RBFN is compared in Table III, in which the time of RBFN is about three times shorter than that of MLP, indicating that RBFN is a more powerful forecasting technique when fast learning is required.

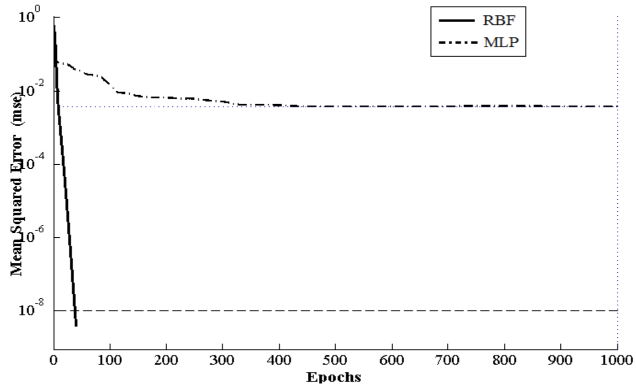

Fig. 6 NN errors as a function of the number of epochs during the training process for MLP and RBF for 1-hour-ahead forecasting

TABLE III
TRAINING TIME(S) FOR MLP AND RBF TECHNIQUES

	Model	MLP	RBFN
Training time (s)	24-hour	75	25
	1-hour	50	15

C. Gain

The current paper introduces the Gain served as an indicator of short-term prediction capabilities, as larger gain implies more similar MAPE values for different forecasting time lengths (i.e. 1, 2, 6, 12, and 24 hours) and hence the forecasting technique is more consistent in terms of time spam within the short-term range.

On one hand, as shown in Fig. 7, LR has the largest gain among three techniques and its gain increases smoothly. Therefore LR is seemingly the most consistent technique for forecasting. Nevertheless, LR controversially trains its own linear combination algorithm using the previously mentioned electric signal with strong nonlinearity, and consequently is characterized by questionable prediction capabilities resulting in the MAPE, albeit invariably significant in absolute value in comparison with those of MLP and RBF, which can be observed from Figs. 3 and. 4, insensitive to neither N-hour-ahead forecasting is intended. Therefore, the Gain of LR is ascribable to the consistent inaccuracy and is omitted from the comparison specific to the training signals collected in this work.

On the other hand, comparison between MLP and RBF, forecasting techniques with more appropriate nonlinear combination algorithms, shows that the Gain for RBF increases first, then decreasing while forecasting 6-hour-ahead loads, and finally increases again. For 1-hour and 2-hour-ahead forecasting, RBF has the larger gain than MLP does, but starting from 6-hour to 12-hour-ahead forecasting, it has the smaller gain than MLP does. The Gain of MLP is increasing positively than the one of RBF.

TABLE IV
GAIN FOR N-HOUR-AHEAD FORECASTING BY USING THREE TECHNIQUES

N HOURS	LR	MLP	RBF
1	0.67	0.36	0.39
2	0.91	0.48	0.77
6	0.94	0.73	0.43
12	0.97	0.95	0.8
24	1	1	1

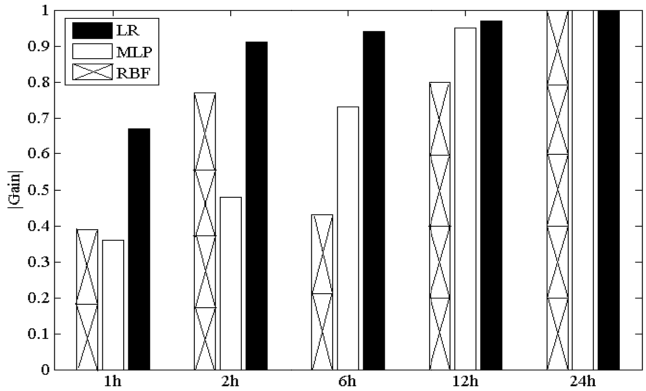

Fig. 7 Absolute values of Gain for N-hour-ahead forecasting of three techniques

D. Comparison and Discussion

First, the MAPEs of the 24- and 1-hour-ahead forecasting models using all three techniques are compared, which is presented in Table I. Comparatively, using each technique to forecast 24- and 1-hour-ahead loads, lower MAPE values are obtained for the last model since load variables of very short-term are taken into consideration. In addition, by comparing the MAPEs for 24-hour-ahead forecasting using three techniques, it can be observed that RBFN provides the lowest MAPE, being 7.92%, substantiating that RBFN gives the best prediction accuracy for this forecasting model. The same result can be also observed from the MAPEs for 1-hour-ahead forecasting model.

Thus it can be concluded that the prediction accuracies using RBFN are enhanced as absolute values of MAPEs for two models using RBFN are both the lowest among all the three techniques. By comparison to the results of [10], we implement a more comprehensive and suitable model regarding to our signal, which contributes to a lower value of MAPE. Also we compared the forecasting results by adding temperature and humidity data as one of the input variables. It indicates that this weather information can also affect the prediction accuracy and plays a vital role in short-term load forecasting.

Second, the training time of 24- and 1-hour-ahead forecasting with MLP and RBFN is compared in Table III. The training time of RBFN is three times shorter than that of MLP, indicating that RBFN is a more powerful forecasting technique when fast learning is required. This is because RBFN uses radial basis functions whose value depends only on the distance from a center point as activation functions instead of sigmoid functions used by MLP. The typical activation function for RBF is a Gaussian, that is, there is only a small effect for input values that are far away from its center if parameters of that neuron are changed. The main difference between MLP and RBFN is the absence of hidden layer weights. The parameters adjusted in the learning process are only the linear mapping from hidden layer to output layer. Therefore, the error is ensured by linearity and has a single minimum which is easily to be found and hence the processing is rapid. Besides, RBF is a locality type of learning so there is no doubt that RBF will not be trapped into a local minima and there is no need to restart the training in case of no error decreasing with further iteration sometimes happened by MLP during the training process, thus accelerating training.

Finally, we analyze the Gain for N-hour-ahead forecasting models by using three techniques, as shown in Table IV, as larger gain implies more similar MAPE values for different forecasting time lengths and hence the forecasting technique is more consistent in terms of time spam within the short-term range. We conclude that MLP is characterized by a more consistent behavior with short-term range in terms of forecast error. It is observed from Fig. 7 that the Gain of RBF is not increasing positively like the one of MLP and it is sometimes larger than the Gain of MLP and sometimes smaller, which indicates that RBF is not consistent in any time spam length of forecasting. This is mainly because we use kernel k-means algorithm as the training algorithm for RBF network, in which data points are mapped from input space to a higher dimensional feature space through a nonlinear transformation. The drawbacks of this algorithm are that its final solution is dependent on the initial position of the cluster centers, and also the clusters must be separated linearly. Besides, compared with MLP structure which has one or more hidden layers, RBF only has one layer so it requires more hidden neurons and hence it has curse of dimensionality and difficulties with large number of units. In addition, since our data are collected from the real small pilot with actual users, it is reasonable that there exist missing data. With its high sensitivity to the initial condition, especially when the dataset has noise, it is not reliable to use RBF technique to forecast different time spam randomly.

V. CONCLUSIONS

Demand Response is a valuable feature in current and future of Smart Grids and Smart Micro Grids. Accurate LF can help the DR mechanisms to shift reliably energy usage to periods of low demand or high availability of renewable energy. Therefore, LF plays a fundamental role especially in short-term time spam wherein variability is higher. Different techniques are applied for STLF over the years but there are few papers comparing their performances.

In this study, three techniques, i.e. LR, ANNs, RBF, are assessed and compared for 24- and 1-hour-ahead load forecasting. It can be concluded that electric signals considered in the evaluation have strong nonlinear behavior so LR is useless to forecast the loads and the forecast errors are large. Also we compare both 24- and 1-hour-ahead forecasting models in terms of absolute forecast errors and training time. We can conclude that RBF is the best technique to perform LF in the chosen settings because not only has it the smallest absolute error, but also it has highest computing speed which can enable RBFN to be potentially applied to attaining real-time LF which in principle requires fast learning. In addition, we introduce the gain to analyze the consistency of the forecasting techniques and consequently conclude that it is particularly encouraging as MLP may be extended to forecast the loads of any time length within this range for our signals. Although RBF has the highest prediction accuracy, it is sensitive and not consistent if we need to forecast the loads of different time spam randomly. In contrary, MLP is the best choice among these three forecasting techniques for our signal from the concrete scenario if we need to do forecasting with different time spam randomly.

ACKNOWLEDGMENT

The authors are grateful for the helpful comments provided by Prof. Roberto Garello (Politecnico di Torino). This work described in this paper has been partially supported by the European FP7 Project GreenCom (Contract No. 318213).

REFERENCES

[1] M. T. Hagan, S. M. Behr, "The time series approach to short term load forecasting" IEEE Trans. Power Sys., vol. PWRS-2, no.3, Aug. 1987, pp. 785-791.
[2] Papalexopoulos, A. D., Hesterberg, T.C., "A regression-based approach to short-term load forecasting" IEEE Trans. Power Sys., vol.5, no.4, Nov. 1990, pp 1535-1537.
[3] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation" IEEE Trans. Power Sys., 16:44-55, 2001.
[4] J. W. Taylor, R. Buizza, "Neural network load forecasting with weather ensemble predictions" IEEE Trans. Power Sys., 17:626-632, 2002.
[5] W. Charytoniuk, M. S. Chen, "Very short-term load forecasting using aritificial neural networks" IEEE Trans Power Sys., 2000, pp15(1): 263-268.
[6] X. F. Li, E. Zou, T. S. Zhang, "Optimization design of feed-forward neural network structure based on chaos wariables" Control and Decision, 2003, 18(6):703-707.
[7] G. Bugmann, "Normalized Gaussian radial basis function networks" Neurocomputing (Special Issue on Radial Basis Function Networks), vol. 20, no.1, pp.97-110, 1998.
[8] G. B. Huang, P. Saratchandran, N. Sundararajan, "A generalized growing and pruning RBF neural network for function approximation" IEEE Trans. on Neural Networks, vol. 16, no.1, pp. 57-67, 2005.
[9] K. Z. Mao, G. B. Huang, "Neural slection for RBF neural network classifier based on data structure preserving criterion" IEEE Trans. on Neural Networks, vol. 16, no.6, pp. 1531-1540, 2005.
[10] Gontar Z., Hatziargyriou N., "Short-term load forecasting with radial basis function network" IEEE Porto Power Tech Conference, vol. 3, Sep 2001.
[11] S. J. Kiartzis, A. G. Bakirtzis, "A fuzzy expert system for peak load forecasting: application to the Greek Power System" Proceedings of the 10th Mediterranean Electrotechnical Conference, 3:1097-1100, 2000.
[12] Bretthauer G., Handschin E., Hoffmann W., "Expert systems application to power systems--state-of-the-art and future trends" IFAC Symposium on Control of Power Plants and Power Systems, Munich, Germany, 1992, p.463-68.
[13] Adapa R., "Expert system applications in power system planning and operations" IEEE Power Eng. Rev, 1994, 14(2):15-8.
[14] Liu C., Ma T., Liou K., Tsai M., "Practical use of expert systems in power systems" Int J Eng Intelligent Sys. Electr. Eng. Commun 1994; 2(1):11-22.
[15] Bann J., Irisarri G., Kirschen D., Miller B., Mokhtari S., "Integration of artificial intelligence applications in the EMS: issues and solutions" IEEE Trans. Power Sys., 1996, pp11(1):475-82.
[16] Madan S., Bollinger K., "Applications of artificial intelligence in power systems" Electric Power Sys. Res., 1997, pp 41:117-31.
[17] Barbato, A., Capone, A., Rodolfi, M., Tagliaferri D., "Forecating the usage of household appliances through power meter sensors for demand management in the smart grid," Smart Grid Communications, 2011 IEEE International Conference on, vol., no., pp.404,409, 17-20 Oct. 2011.
[18] Fernandez, I., Borges, C.E., Penya, Y.K., "Efficient building load forecasting," Emerging Technologies & Factoy Automation (ETFA), 2011 IEEE 16th Conference on, vol., no., pp.1,8,5-9 Sep. 2011.
[19] Berardino, J., Nwankpa, C.O., "Interval-specific building load forecasting models for demand resource planning," Power System Technology (POWERCON), 2012 IEEE International Conference on, vol., no., pp.1.5. Oct. 30 2012-Nov. 2 2012.
[20] Canbing Li, Meiping Fu, Jincheng Shang, Peng Cheng, "A novel probabilistic short-term load forecasting method for large power grid," Power and energy Engineering Conference (APPEEC), 2010 Asia-Pacific, vol., no., pp.1,4,28-31 Mar. 2010.
[21] GreenCom project http://www.greencom-project.eu.
[22] Lean Y., Shouyang W., K. K. Lai, "An integrated data preparation scheme for neural network data analysis" IEEE Tansactions on Knowledge and Data Engineering, vol. 18, no.2, Feb. 2006.
[23] X. Hu, "DB-H reduction: a data preprocessing algorithm for data mining applications" Applied Math. Letters, vol. 16, pp. 889-895, 2003
[24] M. Lou, "Preprocessing data for neural networks" Technical Analysis of Stocks Commodities Magazine, Oct. 1993.
[25] GreenCom pilot, Island of Fur. http://www.greencom-project.eu/island-of-fur.html.