Advanced Artificial Neural Networks for Short-term Load Forecasting

K. Kalaitzakis, G. S. Stavrakakis, D. Kolokotsa*

Dept. of Electronics and Computer Eng.,
Technical Univ. of Crete, GR-73100, Chania, GREECE.

Abstract

This paper presents the development and application of advanced neural networks to face successfully the problem of the short-term electric load forecasting. Several approaches including gaussian encoding backpropagation, window random activation, radial basis function networks, real-time recurrent neural networks and their innovative variations are proposed, compared and discussed in this paper. The performance of each presented structure is evaluated by means of an extensive simulation study, using actual hourly load data from the power system of the island of Crete, in Greece. A parallel processing approach for 24 hour ahead forecasting is proposed and applied. Thus, acceptable accuracy load predictions are obtained without the need of weather data that increase the system complexity, storage requirement and cost.

1. Introduction

Load forecasting is a very crucial issue for the operational planning of electric power systems, especially for the isolated ones. Short-Term Load Forecasting (STLF) aims at predicting electric loads for a period of minutes, hours, days, or weeks. Short-term load forecasting plays an important role in the real-time control and the security functions of an energy management system. Short-term load forecasting applied to the system security assessment problem, especially in the case of increased Renewable Energy Sources (RES) penetration in isolated power grids, can provide, in advance, valuable information on the detection of vulnerable situations. Long- and the medium-term forecasts are used to determine the capacity of generation, transmission, or distribution system additions, along with the type of facilities required in transmission expansion planning, annual hydro and thermal maintenance scheduling, etc. Short-term forecasts are needed not only for power system control and dispatching, but also as inputs to load-flow study or contingency analysis.

A short-term load forecast for a period of one to twenty-four hours ahead is important for the daily operations of a power utility. It is used for unit commitment, energy transfer scheduling and load dispatch. With the emergence of load management strategies, the short-term forecast plays a broader role in utility operations, especially in the case of isolated power grids with increased RES penetration, as in the case of Crete island. Development of an accurate, fast and robust short-term load forecasting methodology is crucial to both, the electric utility and its customers.

An enormous upwelling of interest has grown in recent years in application of artificial intelligence techniques to industrial processes. Their advantage is that no complex mathematical formulation or quantitative correlation between inputs and outputs is required. Many years' data are also not necessary. The effective performance of artificial intelligence in the context of ill-defined processes has led to successful application in load forecasting procedures. As a consequence, pattern recognition [1], expert systems [2,3] and neural networks [4-9,11,12] have been proposed for electric load forecasting and wind power forecasting [10]. Expert system based methods capture the expert knowledge into a comprehensive database, which is then used for predicting the future load. These models exploit knowledge of human experts for the development of rules for forecasting. However, transformation of an expert knowledge to a set of mathematical rules is often a very difficult task.

---

* Corresponding author.

The Artificial Neural Network (ANN) based models are the most popular ones for load forecasting applications. The advantage of ANN over statistical models lies in its ability to model a multivariate problem without making complex dependency assumptions among input variables [4,6,8,10,12]. Furthermore, the ANN extracts the implicit non-linear relationship among input variables by learning from training data.

The ANN forecasting models, proposed in this paper, trace previous load patterns and predict a load pattern using recent load data, without the need for weather information. The ANNs structures implemented and illustrated in the present paper are: multi-layer perceptron, adaptive learning rate backpropagation, gaussian encoding backpropagation, random activation weight networks (RAWN), moving window regression trained RAWN, radial basis function networks, real-time recurrent networks and autoregressive recurrent networks.

The proposed neural network models are trained to identify the load model that reflects the stochastic behavior of the hourly load demand of the island of Crete in Greece. The results obtained are compared to those from the current standard utility practice (statistical method), providing useful comparative conclusions and observations for all described methods.

2. The Autoregressive STLF Method Used By The Utility

Although this work is primarily oriented to neural network approaches for STLF, the statistical method currently used in the utility of Crete Island is also modeled for comparison purposes. This method is the well-known stochastic autoregressive (AR) approach and is applied to the same load data.

The original AR model is expressed by the equation:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + b_t$$

where $y_t, y_{t-1}, \ldots, y_{t-p}$ are the present and previous values of the time series, $a_1, a_2, \ldots, a_p$ are the unknown weighting coefficients of these values, a_0 is a constant term and b_t is the random noise term.

The STLF results of the AR model for the minimum load case (2h00) and the maximum load case (14h00) of the Crete Island utility are shown in Tables 1 and 2, respectively.

3. Advanced Backpropagation Algorithms

The Multi-Layer Perceptron (MLP) neural network, trained by the standard Back-Propagation (BP) algorithm, is the most widely used approach for complex mappings forming between input and output and its mathematical properties for non-linear function approximation are well-documented [11,13-15]. The generalized delta rule is applied to adjust the weights of the feed forward networks thus minimizing a predetermined cost error function.

Since the load forecasting procedure follows the dynamics of a non-linear system, a common approach is to configure and train a neural network to represent a non-linear autoregressive model structure. It is a natural perception to reflect the dynamic nature of the problem by sequential information processing.

It has been found that a major factor affecting the neural model prediction accuracy is the data coding method. The conventional data conditioning method is re-scaling and representing them using a single node at the input or output network layers. An alternative representation, called Gaussian Encoding (GE), a particular case of spread encoding, proved to ensure high degree of accuracy for the neural network [14].

According to the GE technique, each data value is represented as the mean value of a sliding gaussian pattern of excitation over several nodes at the network input and output. The reverse procedure is applied at the network output to decode the values back into the original variable range. The encoding procedure has similarities with data fuzzification techniques, where the scalar dimensional space of each variable is fuzzified to a higher dimensions space. Also, decoding of the network output using the GE method is performed computing a weighted summation of the node excitations, which resembles the conventional center of gravity defuzzification technique. Thus, a network featuring gaussian encoding can be considered as a fuzzy-neural-type network.

Formal techniques for determination of optimum number of nodes in the hidden layers are still under research. Currently, this task is often accomplished by experimentation. The resulting MLP network topology with gaussian encoding, applied to the case study load data, consists of 48 input nodes, 34 and 16 nodes for the two hidden layers and 6 output nodes. The statistical indices of the STLF results for the minimum load case (2h00) and the maximum load case (14h00) are illustrated in Table 1 and Table 2, respectively.

4. Random Activation Weight Neural Network (RAWN) And Moving Window Regression Trained Rawn (MWRAWN)

A general function approximation can be obtained by feed-forward neural networks consisting of just one hidden layer of non-linear neurons. The innovative idea behind RAWN networks is that training of the weights between the input and the hidden layer is not required. Initiating the activation weights as random numbers, the parameters estimation process can be considered of linear type, thus a linear least-squares estimator [16] can be used.

Iterative methods updating the estimate and accelerating the calculations, whenever information is available, can be used. For fast and accurate calculations, a moving window regression method for the RAWN network training is implemented here, which is described and proposed for fist time in [17].

For the case study under consideration, best results are obtained using an 8/20/1 RAWN and MWRAWN structure, while the statistical indices of the STLF results for the minimum load case (2h00) and the maximum load case (14h00) are illustrated in Table 1 and Table 2, respectively.

5. Radial Basis Functions Networks (RBFN)

Another type of hybrid network, which features the architecture of the instar-outstar model and uses the hybrid unsupervised and supervised learning scheme, is the radial basis function network (RBFN) suggested by Moody and Darken [11,15]. The present work adopts a systematic approach to the problem of center selection. Because a fixed center corresponds to a given regressor in a linear regression model, the selection of RBF centers can be regarded as a problem of subset selection. The orthogonal least squares method can be employed as a forward selection procedure, which constructs RBF networks in a rational way. The algorithm chooses appropriate RBF centers one by one from training data points until a satisfactory network is obtained. Each selected center minimizes the increment to the desired output variance, thus ill conditioned problems, frequently occurring when random center selection is used, can automatically be avoided. Orthogonal least squares learning procedure generally produces an RBF network smaller than a randomly selected RBF network [15,18]. Due to its linear computational procedure at the output layer, the RBFN is faster in training time compared to its back-propagation counterpart.

In the present paper, the zero-order regularization (ZORRBF) algorithm proposed in [18] is employed to model the hourly demand load. For the case study under consideration, best results are obtained using 10 inputs. Although an elegant approach to the selection of the regularization parameter λ is to adopt Bayesian techniques, in this work this parameter was set by trial and error to small positive values, which satisfy the optimal problem solution.

The λ parameter is set equal to 0.0002 and 0.0008 for the maximum and minimum load cases, respectively. As a result, the corresponding centers are found by the orthogonal least-squares procedure to be equal to 27 and 13, respectively. For comparison purposes, it should be noted that for the case of maximum load, the original orthogonal least-squares algorithm, without the use of λ parameter, gives a network with similar accuracy but with the computational cost of 69 centers. The STLF results for the minimum load case (2h00) and the maximum load case (14h00) are illustrated in Tables 1 and 2, respectively.

6. Real-Time Recurrent Networks

One of the most popular training algorithms for recurrent networks based on the gradient descent is the real-time recurrent learning algorithm (RTRL) [14,15]. For the STLF case under consideration, the RTRL network is employed as one-step ahead predictor, similarly to the previous approaches. Best results are obtained using a 4/10/1 structure. Attempts to increase the input dimensionality, for the sake of improved accuracy, results in a severe deterioration of its performance. The statistical indices of the STLF results for the minimum load case (2h00) and the maximum load case (14h00) are illustrated in Table 1 and Table 2, respectively.

7. Autoregressive Recurrent Neural Networks (ARNN)

The Autoregressive Recurrent Neural Network (ARNN) is a hybrid type feed-forward/feedback neural network, with feedback represented by recurrent connections appropriate for approximating a load time series.

There are two hidden layers, with sigmoid transfer functions, and a single linear output node. The ARNN topology allows recurrence only in the first hidden layer. For this layer, the memoryless backpropagation model has been extended to include an autoregressive memory, a form of self-feedback, where the output depends also on the weighted sum of previous outputs.

For the STLF case under consideration, an ARNN with 8 inputs is employed, giving a structure of 8/20/14/1 nodes. The computer runs reveal that the proposed training procedure is faster compared with the standard MLP structures.

The STLF results for the minimum load case (2h00) and the maximum load case (14h00) are shown in Tables 1 and 2, respectively.

8. The 24 Hours Ahead Load Prediction Approach

The objective of STLF is to predict the n hourly loads ahead, where $n \leq 24$. A parallel processing forecasting procedure is proposed here, where n-neural blocks with a single output have been implemented and trained separately to provide the n hourly ahead load forecasts. Each neural block is fed by its precedent one. Hence, step-by-step, a n-hour ahead load prediction is obtained. According to this procedure, the requested load for each specific hour is forecasted, not only using the load time-series for this specific hour from the previous days, but also using the forecasted load data of the closer previous time steps for the same day. For the case studies of this work, the training data set consists of the hourly load data for the whole year 1994 (60% of the available data set). The testing set consists of the hourly load data of first four months of 1995 (40% of the available data set). The training and the testing sets are classified into 24 time-series, each one corresponding to an hour of the day.

The above-described parallel processing implementation improves significantly the results obtained from any forecasting model compared to those from the same forecasting model implemented without parallel processing. This is attributed to the proposed modular parallel architecture of 24 separate neural blocks, each one with a single output, featuring easier and faster training compared to traditional neural approaches, which treat the output as a unique 24×1 vector. Thus the accumulation and propagation of prior hours error is minimized. In addition, acceptable accuracy load predictions are obtained without the need of huge amounts of weather data that increase the system complexity, storage requirement and cost [2,4-7,12,19-23].

9. Results And Discussion

Forecast results and statistical properties obtained from application of the developed STLF ANN structures on the autonomous power system of the island of Crete (Greece) are presented and discussed in this section. Case studies for all proposed methods were carried out for a 24-hour load forecasting. The STLF results for the utility of Crete Island, produced by all ANN structures presented here, are analyzed and compared on the basis of the following well-known statistical indices, Standard deviation error: $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[y_i - \hat{y}_i]^2}$, Percent relative error:

$$\varepsilon = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|\cdot 100/y_i \quad \text{and Root mean square error (RMSE):} \quad \varepsilon_1 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{y_i - \hat{y}_i}{y_i}\right]^2}\cdot 100$$

The STLF results for the hours with minimum and maximum load consumption are summarized below in the Tables 1 and 2, respectively.

The AR method is the current standard utility practice used widely for STLF applications [2,4,7,12,13,20,21]. The AR approach along with the standard MLP combined with either a BP or a ALRBP learning algorithm, produce comparable results, thus they are considered as test bed cases in this work. On the other hand, the proposed GE structure provides a considerably higher degree of accuracy compared to the classic MLP structure. Although this performance improvement is generally at the expense of a larger network size, the use of the GE structure has significant advantages in applications requiring long-range predictions.

The RAWN algorithm is a least squares type method offering faster training. This algorithm uses a non-iterative training method for the computation of the set of weights and the corresponding results approximate those from a classic MLP using the BP learning rule. The proposed MWRAWN algorithm features the moving window regression method for network training, improving considerably the STLF accuracy.

An alternative STLF approach is a neural model employing radial basis functions (RBF). The presented ZORRBF algorithm improves both, training time and accuracy. An additional advantage of the specific algorithm is the elimination of the overfitting problem by the use of the regularized parameter λ. Because of this, the ZORRBF network has a very compact structure and achieves significant STLF accuracy improvement, compared to all other neural architectures presented here.

The most significantly accurate STLF results are obtained applying the ARNN and the ZORRBF structures.

A general remark concerning the STLF accuracy is that he forecasting model is trained using load data from the isolated power system of the island of Crete. The load profile includes industrial, commercial and domestic loads. Small power systems experience wider load fluctuations reflecting higher RMSE values. It is expected that the relevant RMSE of large interconnected systems would be in the range of 1.0% to 1.5%.

10. Conclusions

In this paper, a comparative analysis of artificial neural network based STLF techniques is presented, using the load data of the Greek island of Crete. These methods are applied for one-day-ahead prediction of the hourly electric load and employ a modular parallel architecture, based on a separate forecasting module for each hourly load. Several advanced neural architectures were tested including multilayer perceptrons, RAWNs, radial basis, recurrent neural networks and their innovative variations, presented in the present paper.

The case study proves that the zero-order regularization RBF (ZORRBF), the autoregressive recurrent neural network (ARNN) and the gaussian encoding (GE) structures lead to the most accurate STLF results. The STLF accuracy achieved by the above structures, applied using the proposed modular parallel architecture, is significantly satisfactory, thus the need for additional load and/or weather information is not necessary.

References

1. J. Hertz, A. Krogh, R.G. Palmer, Introduction to Theory of Neural Computation, Addison-Wesley, 1991.
2. K.L. Ho, et al, Short-term load forecasting of Taiwan power system using a knowledge based expert system, IEEE Trans. on Power Systems 5 (4) (1990), pp. 1214-1219.
3. S. Rahman, R. Bhatnagar, An expert system based algorithm for short-term load forecast, IEEE Trans. on Power Systems AS-101 (9) (1982).
4. K. Lee, Y. Cha, C. Ku, A study on neural networks for short-term load forecasting, Proc. of ANNPS '91, Seattle, July 1991, pp. 26-30.
5. C.N. Lu, H.T. Wu, S. Vemuri, Neural network based short-term load forecasting, IEEE Trans. on Power Systems 8 (1) (1993), pp. 337-342.
6. G.N. Kariniotakis, E. Kosmatopoulos, G. Stavrakakis, Load forecasting using dynamic high-order neural networks, Proc. of IEEE-NTUA Joint Int. Power Conf., Athens-Greece, 1993, pp. 801-805.
7. A. Bakirtzis, V. Petridis, S. Klartzis, A neural network short-term load forecasting model for the Greek power system, IEEE Trans. on Power Systems 11 (2) (1996), pp. 858-863.
8. A. Papalexopoulos, S. How, T.M. Peng, An implementation of a neural network based load forecasting model for the EMS, IEEE Trans. on Power Systems 9 (4) (1994), pp. 1956-1962.
9. Srinivasan, Dipti, Evolving artificial neural networks for short-term load forecasting, Neurocomputing 23 (1998), pp. 265-276.
10. G.N. Kariniotakis, G.S. Stavrakakis, E.F. Nogaret, Wind power forecasting using advanced neural networks models, IEEE Trans. on Energy Conversion 11 (4) (1996), 762-767.

11. M. El-Sharkawi, D. Niebur, Artificial neural networks with applications to power systems, IEEE Catalog No 96TP112-0, IEEE Service Center, NJ, 1996.

12. I. Drezga, S. Rahman, Short-term load forecasting with local ANN predictors, IEEE Trans. on Power Systems 14 (3) (1999).

13. A.G. Parlos, B. Fernandez, A. Atiya, J. Muthusami, W.K. Tsai, An accelerated learning algorithm for multilayer perceptron networks, IEEE Trans. on Neural Networks 5 (3) (1994), pp. 220-238.

14. Kevin Swingler, Applying neural networks – a practical guide, Academic Press, London, 1996.

15. Chin-Teng Lin, G.S.G. Lee, Neural fuzzy systems, Prentice Hall, NJ, 1996.

16. H.A.B. Te Braake, G. Van Straten, Random activation weight neural net (RAWN) for fast non-iterative training, Engineering Applications of Artificial Intelligence 8 (1) (1995), pp. 71-80.

17. G.S. Stavrakakis, A. Pouliezos, Fatigue life prediction using a new moving window regression method, Mechanical Systems and Signal Processing 5 (4) (1991), pp. 327-340.

18. S. Chen, E. Chng, Regularized orthogonal least squares algorithm for constructing radial basis function networks, Int. J. Control 64 (5) (1996), pp. 829-837.

19. I. Moghram, S. Rahman, Analysis and evaluation of five short-term load forecasting techniques, IEEE Trans. on Power Systems 4 (4) (1989), pp. 1484-1491.

20. G. Irisarri, S. Widergren, On-Line load forecasting for energy control center applications, IEEE Trans. on Power App. And Sys. 101 (1) (1982), pp. 71-78.

21. K. Jabbour, J.F.V. Riveros, D. Landsbergen, ALFA: automated load forecasting assistant, IEEE Trans. PWRS 3 (1988), pp. 908-914.

22. M.S. Kandil, S.M. El-Debeiky, N.E. Hasanien, Overview and comparison of long-term forecasting techniques for a fast developing utility: part I, Electric Power Systems Research 58 (2001), pp. 11-17.

23. Bai-Linz Zhang, Zhao-Yang Dong, An adaptive neural-wavelet model for short term load forecasting, Electric Power Systems Research 59 (2001), pp. 121-129.

Tables

Table 1. Statistical indices for the 2h00 hour load forecasting (min load case)			
Models	AR	BP	ALRBP
Relative Error (%)	3.86297	4.22628	2.73460
Standard Dev. Error	5.90894	5.16192	3.91741
RMSE (%)	5.80199	5.18811	3.80634
Models	GE	RAWN	MWRAWN
Relative Error (%)	1.51737	6.82535	2.75417
Standard Dev. Error	2.13450	9.24337	4.19438
RMSE (%)	2.03138	8.98179	4.06330
Models	ZORRBF	RTRL	ARNN
Relative Error (%)	1.35119	2.98850	1.22330
Standard Dev. Error	1.96127	4.35816	1.83768
RMSE (%)	1.86722	4.08409	1.71771

Table 2. Statistical indices for the 14h00 hour load forecasting (max load case)			
Models	AR	BP	ALRBP
Relative Error (%)	11.18372	13.40970	11.9674
Standard Dev. Error	25.60678	27.20622	24.3568
RMSE (%)	17.20486	19.04599	17.6778
Models	GE	RAWN	MWRAWN
Relative Error (%)	3.40010	11.81696	10.59024
Standard Dev. Error	7.71597	26.29459	21.74416
RMSE (%)	4.61941	17.54215	15.19347
Models	ZORRBF	RTRL	ARNN
Relative Error (%)	2.79244	11.5430	2.73292
Standard Dev. Error	5.77855	24.2722	5.92847
RMSE (%)	3.60609	16.4966	3.68308