New Approach for Load Modeling

S. Chokri

Abstract—Load modeling is one of the central functions in power systems operations. Electricity cannot be stored, which means that for electric utility, the estimate of the future demand is necessary in managing the production and purchasing in an economically reasonable way. A majority of the recently reported approaches are based on neural network. The attraction of the methods lies in the assumption that neural networks are able to learn properties of the load. However, the development of the methods is not finished, and the lack of comparative results on different model variations is a problem. This paper presents a new approach in order to predict the Tunisia daily peak load. The proposed method employs a computational intelligence scheme based on the Fuzzy neural network (FNN) and support vector regression (SVR). Experimental results obtained indicate that our proposed FNN-SVR technique gives significantly good prediction accuracy compared to some classical techniques.

Keywords—Neural network, Load Forecasting, Fuzzy inference, Machine learning, Fuzzy modeling and rule extraction, Support Vector Regression.

I. INTRODUCTION

DURING the last four decades, a wide variety of techniques have been used for the problem of load forecasting [1], [2]. Such a long experience in dealing with the load forecasting problem has revealed time series modeling approaches based on statistical methods and artificial neural networks (ANNs). Statistical models include moving average and exponential smoothing methods such as multi-linear regression models, stochastic process, data mining approaches, autoregressive moving average (ARMA) models, Box-Jenkins' methods, and Kalman filtering-based methods [3]-[6]. Since, load time series are usually nonlinear functions of exogenous variables; therefore, to incorporate non-linearity, ANNs have received much attention in solving problems of load forecasting [7]-[10]. ANN-based methods have reported fairly good performances in forecasting. However, two major risks in using ANN models are the possibilities of less or excessive training data approximation, i.e., under-fitting and over-fitting, which increase the out of-sample forecasting errors. Hence, due to the empirical nature of ANNs their application is cumbersome and time consuming.

Recently, new machine learning techniques such as the support vector machines (SVMs) have been used for load prediction and electricity price forecasting, and have achieved good performances [11], [12]. SVM, namely, support vector regression (SVR) is a powerful machine learning technique used for regression, which is based on recent advances in statistical learning theory [13]. Established on the structural risk minimization (SRM) principle (estimate a function by minimizing an upper bound of the generalization error), SVMs have shown to be very resistant to the under-fitting and over-fitting problems caused by ANNs as in [11].

In this paper, we incorporated the concept of fuzzy set theory [14]. Into the SVM regression and Neural networks modeling [15], [16]. A novel fuzzy neural network combining with support vector learning mechanism called Fuzzy Neural Networks based on Support vector Regression (FNN-SVR) is proposed. The FNN-SVR combine the capability of minimizing the empirical risk (training error) and expected risk (testing error) of support vector learning in high dimensional data spaces and the efficient human-like reasoning of FNN.

A learning algorithm consisting of three learning phases is developed to construct the FNN-SVR and train the parameters. In the first phase, the fuzzy rules and membership functions are automatically determined by the clustering principle. In the second phase, the parameters of FNN are calculated by the SVR with the proposed adaptive fuzzy kernel function for time series prediction. In the third phase, the relevant fuzzy rules are selected by the proposed fuzzy rule reduction method.

For developing the forecasting models, we used the daily peak electrical load data provided by the Tunisia Electric and Gas Company (STEG) for the years 2002 through 2011.

The paper is organized as follows: the structure and the learning algorithm behavior of the proposed FNN-SVR are described in Section II. The proposed model is used to predict the Tunisia daily peak load and is compared with some and nonparametric techniques such as SVR and BNN in Section III. Conclusion is summarized in the last section.
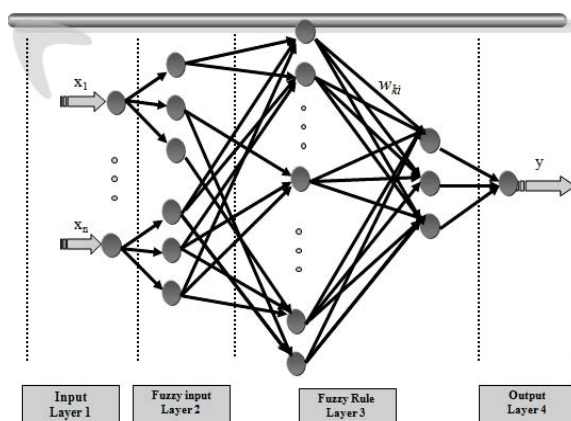

Fig. 1 Structure of the FNN-SVR

S. Chokri is with the MOCFINE Laboratory, Higher Institute of Accounting and Administration Enterprise, University of Manouba, Tunisia (phone: 0021671600705; fax: 0021671602404; e-mail: chokri.slim@ iscae.rnu.tn).

II. METHODOLOGY

A. Structure of the FNN-SVR

The proposed FNN-SVR is a four-layered FNN that is comprised of the input, membership function, fuzzy rules, and the output layer.

- Layer 1 accepts input variables, whose nodes represent input linguistic variables. No computation is done in this layer. Each node in this layer, which corresponds to one input variable, only transmits input values to the next layer directly. That is :

$$y^1 = x_i \tag{1}$$

where x_i, $i = 1, 2, \dots, n$ are the input variables of the network.

- Layer 2 is to calculate the membership values, whose nodes represent the terms of the respective linguistic variables. In other words, the membership value which specifies the degree to which an input value belongs to a fuzzy set is calculated in Layer 2:

$$y^2 = e^{-\frac{(x_i - \mu_{ij})^2}{\sigma_{ij}^2}} \tag{2}$$

where $e(.)$ is a Gaussian membership function

- Nodes at Layer 3 represent fuzzy rules. The links before Layer 3 represent the preconditions of fuzzy rules, and the links after Layer 3 represent the consequences of fuzzy rules

Here we use the AND operation for each Layer two nodes:

$$y^3 = e^{-[\mathbf{h}_j(\mathbf{x}-\boldsymbol{\mu}_j)^T][\mathbf{h}_j(\mathbf{x}-\boldsymbol{\mu}_j)]} \tag{3}$$

where $h_j = \text{diag } [\frac{1}{\sigma_{1j}}, \dots, \frac{1}{\sigma_{nj}}]$, $\boldsymbol{\mu}_j = [\mu_{1j}, \mu_{2j}, \dots, \mu_{nj}]^T$

$\mathbf{x} = [x_1, x_2, x_3, \dots, x_n]^T$ is the input vector of the network

- Layer 4 is the output layer : The single node y^4 in this layer is labeled with $\sum$, which computes the overall output and can be computed as:

$$y^4 = \sum_{j=1}^{m} w_j y^3 + w_0 = \sum_{j=1}^{m} w_j e^{-[\mathbf{h}_j(\mathbf{x}-\boldsymbol{\mu}_j)^T][\mathbf{h}_j(\mathbf{x}-\boldsymbol{\mu}_j)]} + w_0 \tag{4}$$

B. The Learning Algorithm

First the input datasets are partitioned. For each incoming pattern $\mathbf{b} = [\mathbf{x}, y]^T$ the strength a rule is fired. We can use the firing strength as this degree measure:

$$F^j(b) = e^{-[\mathbf{h}_j(\mathbf{b}-\boldsymbol{\mu}_j)^T][\mathbf{h}_j(\mathbf{b}-\boldsymbol{\mu}_j)]} \in [0 \ 1] \tag{5}$$

We can obtain the following criterion for the generation of a new fuzzy rule:

$$H = \max_{1 \le j \le c(t)} F^j(b) \tag{6}$$

where c(t) is the number of existing rules at time t. The optimal parameters of FNN-SVR are trained by using the ξ insensitivity loss function SVR based on the fuzzy kernels. The dual quadratic optimization of SVR is solved in order to obtain an optimal hyperplane for any linear or non linear space:

$$\max \ L(a, a^*) = -\xi \sum_{i=1}^{q}(a_i^* + a_i) + \sum_{i=1}^{q}(a_i^* + a_i)y_i - \frac{1}{2}\sum_{i,j=1}^{q}(a_i^* - a_i)(a_j^* - a_j) \, K(\mathbf{x}_i, \mathbf{x}_j) \tag{7}$$

Subject to constraints:

$$\sum_{i=1}^{q} a_i^* = \sum_{i=1}^{q} a_i; \ \ 0 \le a_i^* \le C; 0 \le a_i \le C; \ \ i = 1, 2, \dots, q \tag{8}$$

where K(xi, xj) is the fuzzy kernel, ξ is a previously chosen nonnegative number for insensitive loss function, and C is a user-specified positive parameter to control the tradeoff between complexity of SVR and the number of non separable points. A solution a $= (a_1, a_2, \dots, a_{ns})$ and $a^* = (a_1^*, a_2^*, \dots, a_{ns}^*)$ can be obtained, where a_i and a_i^* are Lagrange coefficients. The corresponding support vectors s= $[s_1, s_2, \dots, s_i, \dots, s_{ns}]$ can be obtained, and the constant (threshold) w_0 in (4) is:

$$w_0 = \frac{1}{q}(\sum_{i=1}^{q}(y_i - \mathbf{x}_i^T \theta_0)) \text{ with } \theta_0 = \sum_{i=1}^{ns}(a_i^* - a_i) \, \mathbf{x}_i \tag{9}$$

The coefficients w_j in (4) can be calculated by:

$$w_j = y_j(a_j^* - a_j). \tag{10}$$

In this phase, the number of fuzzy rules learning in Phases 1 and 2 is reduced by removing some irrelevant fuzzy rules and the consequent parameters. The method reduces the number of fuzzy rules by minimizing the distance measure between original fuzzy rules and reduced fuzzy rules. To achieve this goal, we rewrite (4) as:

$$y^4 = \sum_{j=1}^{m} w_j y^3 + w_0 = \sum_{j=1}^{m} w_j \prod_{i=1}^{n} e^{-\frac{(x_i - \mu_{ik})^2}{\sigma_{ik}^2}} + w_0 \tag{11}$$

The reduced set is given by:

$$y^{re(4)} = \sum_{k=1}^{r} \gamma_k \prod_{i=1}^{n} e^{-\frac{(x_i - \mu_{ij}^{(r)})^2}{\sigma_{ij}^{2(r)}}} + w_0 \tag{12}$$

674

γ_k is the consequent parameters of the remaining fuzzy rules. The whole learning scheme is iterated until the new rules are sufficiently sparse.

III. EMPIRICAL STUDY

A. Data Set

The load and the temperature were obtained from the Tunisia Electric and Gas Company (STEG) for the years 2004 through 2011. The objective is to predict the daily peak electricity load using the given historical data. We divide the data into two sets: one with 70 percent of the source data, for training the models, and one with 30 percent of the source data, for testing the models. The peak load over the 2004-2005 periods is shown in Fig. 2. The seasonal trend on the load can be easily seen. Also, the weekly load structure can be seen in the form of lower load values on weekends than on working days.
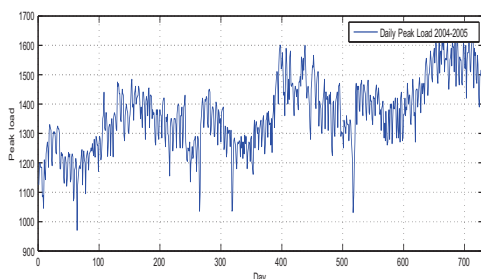

Fig. 2 Daily peak load 2004-2005

Observations regarding load data were investigated to determine the relationship between the load data and other information such as temperature and day types. The following observations are concluded for the given data.

B. Attribute Selection

Our proposed load forecasting model is based on the past daily peak loads (historical consumption data) as one of the candidate input variables. The best input features for our forecasting model are those which have the highest correlation with the output variable (i.e., peak load of the next day) and the highest degree of linear independency. Thus, the most effective candidate inputs with minimum redundancy are selected as the model attributes.

Through simple analysis of the graphs representing the yearly load data, it is observed that the electricity load follows seasonal patterns, i.e., high demand of electricity in the winter (September through March) while low demand in the summer (April through August) as in [11]. This pattern implies the relationship between the electricity usage and weather conditions in different seasons, as indicated in Fig. 2. Secondly, another load pattern is observed, where load periodicity exists in the profile of every week, i.e., the load demand on the weekend (Saturday and Sunday) is usually lower than that on weekdays (Monday through Friday), as shown in Fig. 2. In addition, electricity demand on Saturday is

a little higher than that on Sunday, and the peak load usually occurs in the middle of the week, i.e., on Wednesday.

The tree attributes used in the FNN-SVR modeling process are:

1) Daily peak load: Since, past load demand affects and implies the future load demand, therefore, including the past daily peak load as a key attribute, will greatly influence improvement in the forecasting performance.
2) Daily temperature: The electricity load and the temperature have a causal relationship (high correlation) between each other. Therefore, the daily temperature is used as an attribute in the forecasting model.
3) Type of day: Weekly periodicity of the electricity load is noticed through load data analysis as shown in Fig. 2. As the electricity demand on holidays is observed to be lower than on non-holidays, therefore, encoding information of the type of day (calendar indicator) into the forecasting model will benefit performance of the model.

There are also many other factors that make the models different from each other.

These differences can be for example in:

- the use of the weather data
- the other input variables
- network architecture
- training algorithm
- selection of the training data

For the input variables, the following symbols are used:

- Lmax(i)=maximum load of day i
- Tmax(i)=maximum temperature of day i
- Tmin(i)=minimum temperature of day i
- Lavg(i)=average temperature of day i

To forecast the maximum load of a certain day, at least the maximum loads of the previous day and the corresponding day from the previous week, can be considered potential input variables. Also the temperature data of those days may be useful if temperature forecasts for the target day are available. Maximum, minimum and average temperatures are considered for this purpose. Eight different input structures are tested separately for peak load. These are numbered from 1 to 8, and are listed in Table I for maximum load forecasting Lmax(i).

TABLE I
INPUT STRUCTURES

Input	Structures
1	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tavg(i), Tavg(i-1), Tavg(i-7), Tavg(i-8)
2	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tavg(i), Tavg(i-1), Tavg(i-7)
3	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tavg(i), Tavg(i-1)
4	Lmax(i-1), Lmax(i-7), Tavg(i), Tavg(i-1)
5	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tmax(i), Tmax(i-1), Tmax(i-7), Tmax(i-8)
6	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tmax(i), Tmax(i-1)
7	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tmin(i), Tmin(i-1), Tmin(i-7), Tmin(i-8)
8	Lmax(i-1), Lmax(i-7), Lmax(i-8), Tmin(i), Tmin(i-1),

In addition to inputs listed above, each input structure contains four extra nodes. These get binary values and inform the network of the day type of the target day. The day type

classes are: 1) Mondays, 2) Tuesdays-Fridays, 3) Saturdays, and 4) Sundays.

Informing the network about the day type is important, because Saturdays and Sundays have much lower peak loads than working days.

C. Performance Criteria

Although the MSE is a perfectly acceptable measure of performance, in practice the ultimate goal of any testing strategy is to confirm that the results of models are robust and capable of measuring the profitability of a system. It is important, therefore, to design a test from the outset. This is not always carried out with the level of rigor that it merits, partly because of unfamiliarity with the established methods or practical difficulties intrinsic to non-linear systems. Consequently, we designed test sets to evaluate the effects of the models. The prediction performance is evaluated using the following statistics: Mean Squared Error (MSE), Normalized Mean Squared Error (NMSE), Mean Absolute Error (MAE), Mean Absolute Percent Error (MAPE).

TABLE II
PERFORMANCE CRITERIA

Metrics
$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(La_i - Lp_i)^2$
$\text{NMSE} = \frac{1}{\delta^2 n}\sum_{i=1}^{n}(La_i - Lp_i)^2;$
$\delta^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(La_i - \overline{L}a)^2$
$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|La_i - Lp_i\right|$
$\text{MAPE } \frac{100}{n}\sum_{i=1}^{n}\left|\frac{La_i - Lp_i}{La_i}\right|$

These criteria are defined in Table II. MSE, RMSE, MAE and MAPE measure the correctness of a prediction in terms of levels and the deviation between the actual (La_i) and predicted values (Lp_i). The smaller the values, the closer the predicted time-series values will be to the actual values.

1. Results

A comparative study of our proposed model with other machine learning techniques was performed using the train and the test data. Our proposed FNN-SVR method was compared with two different prediction techniques: (1) SVR and (2) ML-BPNN. The eight different input structures were tested in all cases. The input structure 4 gives the best results in all the cases. The prediction accuracy of our proposed model compared with the different prediction techniques for predicting the peak load of train and test data are shown in Tables III and IV. Figs. 3 and 4 show the comparison plot of the predicted peak load of train and test data using the FNN-SVR, SVR and ML-BPNN techniques.

TABLE III
COMPARISON OF THE FORECASTING ACCURACY USING DIFFERENT PREDICTION TECHNIQUES (TRAINING SET)

Models	MSE	NMSE	MAE	MAPE
BPNN	2399.9	0.054	32.10	2.126%
SVR	958.278	0.0216	26.05	1.394%
FNN-SVR	931.281	0.0209	25.71	1.085%

TABLE IV
COMPARISON OF THE FORECASTING ACCURACY USING DIFFERENT PREDICTION TECHNIQUES (TEST SET)

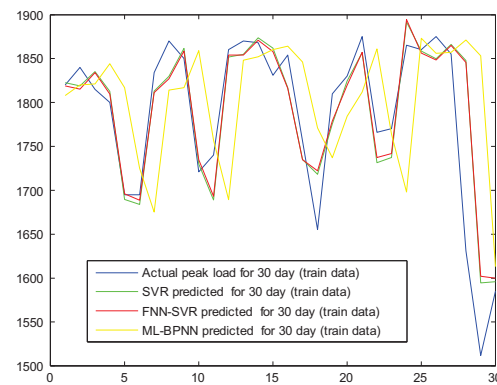| Models	MSE	NMSE	MAE	MAPE
BPNN	981.741	31.333	13.457	0.882%
SVR	725.432	26.934	11.582	0.767%
FNN-SVR	709.707	26.640	11.519	0.762%

Fig. 3 Comparison plot of predicted peak load using FNN-SVR, SVR, and ML-BPNN techniques- 30 days train data
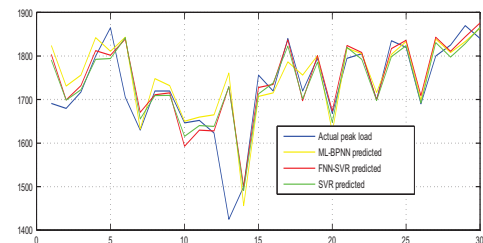

Fig. 4 Comparison plot of predicted peak load using FNN-SVR, SVR, and ML-BPNN techniques- 30 days test data

Results obtained in Tables III and IV indicates that the prediction accuracy of the ML-BPNN is not satisfactory. This is due to the problems of local minima and over-fitting associated with ANNs, which tends to decrease the generalization performance for unseen data. Our proposed FNN-SVR model proves to be superior in terms of the all performance criteria compared to the SVR and ML-BPNN models. This is due to the presence of the Fuzzy component in our model, which fits the clustered training data into the appropriate SVRs based on the Euclidean distance.

IV. CONCLUSION

In this paper a computational intelligence scheme based the FNN and SVR is applied to reconstruct the dynamics of

electricity load forecasting using a time series approach. The proposed FNN-SVR technique is applied on the Tunisia load data to predict the peak load, which demonstrates the effectiveness and efficiency of the prediction technique in contrast to others.

Results obtained indicated that the proposed FNN-SVR model outperforms the other two approaches in terms of all evaluation criteria used in this research. This can be explained by the formulation of the FNN-SVR, SVM and BPNN networks. FNN-SVR and SVM methods use a quadratic programming problem which is convex and has a global optimum solution. BPNN networks use the backpropagation algorithm to minimize the network error, the problem is non convex and it is hard to find the global optimum.

REFERENCES

[1] E. Gonzalez-Romera, M.A. Jaramillo-Moran, D. Carmona-Fernandez, Monthly electric energy demand forecasting based on trend extraction, IEEE Transactions on Power Systems 21 (4) (2006) 1946--1953.
[2] A.D. Papalexopoulos, T.C. Hesterberg, A regression-based approach to short term load forecasting, IEEE Transactions on Power Systems 5 (4) (1990) 1535--1550.
[3] S. Rahman, O. Hazim, A generalized knowledge-based short term load forecasting technique, IEEE Transactions on Power Systems 8 (2) (1993) 508--514.
[4] S.J. Huang, K.R. Shih, Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, IEEE Transactions on Power Systems 18 (2) (2003) 673--679.
[5] H. Wu, C. Lu, A data mining approach for spatial modeling in small area load forecast, IEEE Transactions on Power Systems 17 (2) (2003) 516--521.
[6] G.E.P. Box, G.M. Jenkins, Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, CA, 1976.
[7] A. Khotanzad, R.C. Hwang, A. Abaye, D. Maratukulam, An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities, IEEE Transactions on Power Systems 10 (3) (1995) 1716--1722.
[8] T. Czernichow, A. Piras, K. Imhof, P. Caire, Y. Jaccard, B. Dorizzi, A. Germond, Short term electrical load forecasting with artificial neural networks, Engineering Intelligent Systems for Electrical Engineering and Communications 2 (1996) 85--99.
[9] L.M. Saini, M.K. Soni, Artificial neural network-based peak load forecasting using conjugate gradient methods, IEEE Transactions on Power Systems 12 (3) (2002) 907--912.
[10] S. Fan, C.X. Mao, L.N. Chen, Peak load forecasting using the self-organizing map, in: Advances in Neural Networks (ISNN) 2005, Pt. III, Springer-Verlag, New York, 2005, pp. 640--649.
[11] B.-J. Chen, M.-W. Chang, C.-J. Lin, Load forecasting using support vector machines: a study on EUNITE competition 2001, IEEE Transactions on Power Systems 19 (4) (2004) 1821--1830.
[12] S. Fan, C.X. Mao, L.N. Chen, Next-day electricity price forecasting using a hybrid network, IET Generation, Transmission and Distribution 1 (1) (2007) 176--182.
[13] C. Cortes, V. Vapnik, Support-vector network, Machine Learning 20 (1995) 273--297.
[14] L. A. Zadeh, The concept of linguistic variable and its application to approximate reasoning, Inform. Sci., vol. 8, pp. 199-249, 1975.
[15] C Slim (2009) : Hybrid Approach in Neural Network design Applied to Financial Time Series Forecasting, The Journal of American Academy of Business Cambridge, Vol. 15, Num. 1,September 2009, 294-300.
[16] C. Slim (2007) : Fuzzy Neural Model for Bankruptcy Prediction, The Journal of Business Review Cambridge, Vol. 8, Num. 2, December 2007, 117-122.