Integrated Approach for Short Term Load Forecasting using SVM and ANN

by

Amit Jain, B. Satish

in

TENCON 2008. IEEE Region 10 Conference 19-21 Nov. 2008 Page(s):1 - 6 Digital Object Identifier 10.1109/TENCON.2008.4766840

Report No: IIIT/TR/2009/61

Centre for Power Systems
International Institute of Information Technology
Hyderabad - 500 032, INDIA
March 2009

Integrated Approach for Short Term Load Forecasting using SVM and ANN

Amit Jain, *Member, IEEE* and B. Satish

***Abstract*-- A new hybrid technique using Support Vector Machines (SVM) and Artificial Neural Networks (ANN) to forecast the next '24' hours load is proposed in this paper. The forecasted load for the next '24' hours is obtained by using four modules consisting of the Basic SVM, Peak and Valley ANN, Averager and Forecaster and Adaptive Combiner. These modules try to extract the various components like Basic component, Peak and Valley components, Average component, Periodic component & random component of a typical weekly load profile. The Basic SVM uses the historical data of load and temperature to predict the next '24' hour's load, while the Peak and Valley ANN uses the past peak and valley data of load and temperatures respectively. The Averager captures the average variation of the load from the previous load behaviour, while the Adaptive Combiner uses the weighted combination of outputs from the Basic SVM and the Forecaster, to forecast the final load. The statistical and artificial intelligence based methods are conceptually incorporated into the architecture to exploit the advantages and disadvantages of each technique.**

***Index Terms*-- Artificial Neural Network, Back Propagation Algorithm, Short Term Load Forecasting (STLF), Support Vector Machines.**

I. INTRODUCTION

FORECASTING is an integral part of electric power system operations. It is the primary prerequisite for achieving the goal of optimal planning and operation of power systems [1, 2]. The forecast ranging from 5 to 20 years is termed as long term forecasting while the forecast ranging from few months to 5 years is termed as medium term forecasting. If the duration of the forecast varies from few hours to weeks, it is called as short term load forecasting [3]. The long and medium term forecasting are used to determine the capacity of generation, transmission or distribution system additions and the type of facilities required in transmission expansion planning, annual hydro thermal maintenance scheduling etc. The short term load forecast is needed for control and scheduling of power system and also as inputs to load flow study or contingency analysis. The system operators use the load forecasting result as a basis of off-line network analysis to determine if the system might be vulnerable. If so, corrective actions should be prepared, such as load shedding, power purchases and bringing peaking units on line. The purpose of very short term load forecasting (ranging from minutes to hours) is for real time control and security evaluation [4].

The amount of excess electricity production (or spinning reserve) required to guarantee supply, in the event of an underestimation, is determined by the accuracy of these forecasts and conversely, overestimation of the load leads to sub-optimal scheduling (in terms of production costs) of power plants (unit commitment) [5]. Accurate prediction of load results in economic, reliable and secure operation of the power system which in turn saves cost. Bunn [6] reported that 1% increase in the forecasting error leads to an increase of £10 million operating cost per year. The introduction of deregulation in the electricity industry made short term load forecasting and very short term load forecasting much more important. Because of its great economic importance and the high complexity of electric power systems, short term load forecasting has been subjected to constant improvements in which numerous techniques have been used [3, 6, 7]. In this paper, an attempt is made to predict the load by means of an integrated architecture.

The paper is organized as follows: Section – II discusses the need and suitability of artificial intelligence based short term load forecasting. Section – III explains the basics of ANN and SVM approaches briefly. The proposed architecture and solution methodology are explained in Section – IV. Section – V contain the conclusions.

II. NEED AND SUITABILITY OF ARTIFICIAL INTELLIGENCE BASED METHODS FOR SHORT TERM LOAD FORECASTING

There exist several conventional approaches such as Regression, Interpolation and complex algorithmic methods for forecasting which require heavy computational burden [8]. Broadly we can classify these approaches into two categories. One approach treats the load pattern as a time series signal and predicts the future load by using various time series analysis techniques. The second approach recognizes that the load pattern is heavily dependent on weather variables and finds a functional relationship between the weather variables and the system load.

The Time series approach does not utilize weather information. Most Regression approaches use piece-wise linear relationship between weather variables and load without any justification. But the functional relationship between load

Amit Jain is with Power Systems Research Center, International Institute of Information Technology, Hyderabad, India. (e-mail: amit@iiit.ac.in).

B. Satish is with Power Systems Research Center, International Institute of Information Technology, Hyderabad, India. (e-mail: satish.stlf@gmail.com).

TABLE I
COMPARISON OF CONVENTIONAL AND ARTIFICIAL INTELLIGENCE BASED METHODS

Important features	Time series methods	Regression analysis	AI based methods
Load Information	Considered	Considered	Considered
Weather Information	Ignored	Considered	Considered
Functional Relationship between Load and Weather variables	Ignored	Required	Not required
Complex Mathematical Calculations	Required	Required	Not Required
Time required for Prediction	More	More	Less
Adaptability	Less	Less	More

and weather variables is not stationary, but depends on spatio-temporal patterns. These approaches are problem dependent to a large extent and converge slowly and even may diverge in certain cases. In addition to this, these approaches use either steady state component or average component or the peak component to predict the load. However the prediction of the load depends upon the weighted combination of these three components which varies dynamically. The comparison of various features between conventional and artificial intelligence based methods is provided in TABLE I. In the next section, a brief introduction of ANN and SVM for regression analysis is given.

III. BASICS OF ANN AND SVM

An ANN can be defined as highly connected array of elementary processors called neurons and is capable to perform non-linear modeling and adaptation. It uses previous load patterns as in the cases of Time series and Regression approaches and weather information as in the case of Regression approach. ANN has advantages of both of Time series and Regression methods. The feed forward back propagation algorithm, which updates the weights in such a way such that the error is minimized, is used to train the Neural Networks. The detailed explanation of back propagation algorithm is available in any standard neural network textbook.

Support Vector Regression (SVR) can be used for time series prediction. Given training data $(x_1, y_1), (x_2, y_2) \ldots \ldots (x_n, y_n)$ where x_i are input vectors and y_i are the associated output value of x_i, the Support Vector Regression is an optimization problem.

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{1}$$

subject to $y_i - (w^T \phi(x_i) + b) \leq \varepsilon + \zeta_i$,

$$(w^T \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i^*,$$

$$\zeta_i, \xi_i^* \geq 0,\ i = 1, 2, \ldots \ldots \ldots l,$$

where x_i is mapped to a higher dimensional space, ζ_i is the upper training error (ξ_i^* is the lower) subject to the ε-insensitive tube $|y - (w^T \phi(x) + b)| \leq \varepsilon$. The parameters which control the regression quality are the cost of error C, the width of tube ε and the mapping function, ϕ [9].

The constraints of (1) imply that we would like to put most data x_i in the tube $|y - (w^T \phi(x) + b)| \leq \varepsilon$. If x_i is not in the tube, there is an error ζ_i or ξ_i^* which we would like to minimize in the objective function. For traditional least-square regression, ε is always zero and data are not mapped into higher dimensional spaces. Hence SVR is a more general and flexible treatment on regression problems [9].

IV. PROPOSED ARCHITECTURE

In order to propose a solution to the STLF problem, an integrated architecture incorporating conceptually the statistical and artificial intelligence techniques is chosen to forecast the next day ‘24’ hours load. A typical weekly load profile consists essentially of 1) Basic (steady state) component, 2) Peak and Valley component, 3) Average component, 4) Periodic component and 5) Random component. The previous approaches to load forecasting generally consist of a single independent module, which uses historic data information as its inputs. Such approaches are unable to extract all these components from the past data in a well-defined manner for accurate load prediction. Hence an integrated architecture is proposed which incorporates the above features by means of four independent modules.

Module 1: Basic SVM,
Module 2: Peak and Valley ANN,
Module 3: Averager and Forecaster and
Module 4: Adaptive Combiner.

Out of these four modules, the ‘Basic SVM’ (tracks the steady state and random components of the daily load curve) predicts the next day ‘24’ hours load and the ‘Peak and Valley ANN’ (tracks the peak and valley components of the daily load curve) forecasts the peak and valley loads of the next day. The third module comprises of two blocks ‘Averager’ and ‘Forecaster’. The ‘Averager’ (tracks the average component) computes the hourly averaged load of the day to be forecasted. The ‘Forecaster’ calculates the next day ‘24’ hours load by using the predicted peak and valley loads obtained from the ‘Peak and Valley ANN’ and the hourly averaged load obtained from the ‘Averager’. Using ‘Adaptive Combiner’ the final forecast for the next ‘24’ hours load is done.

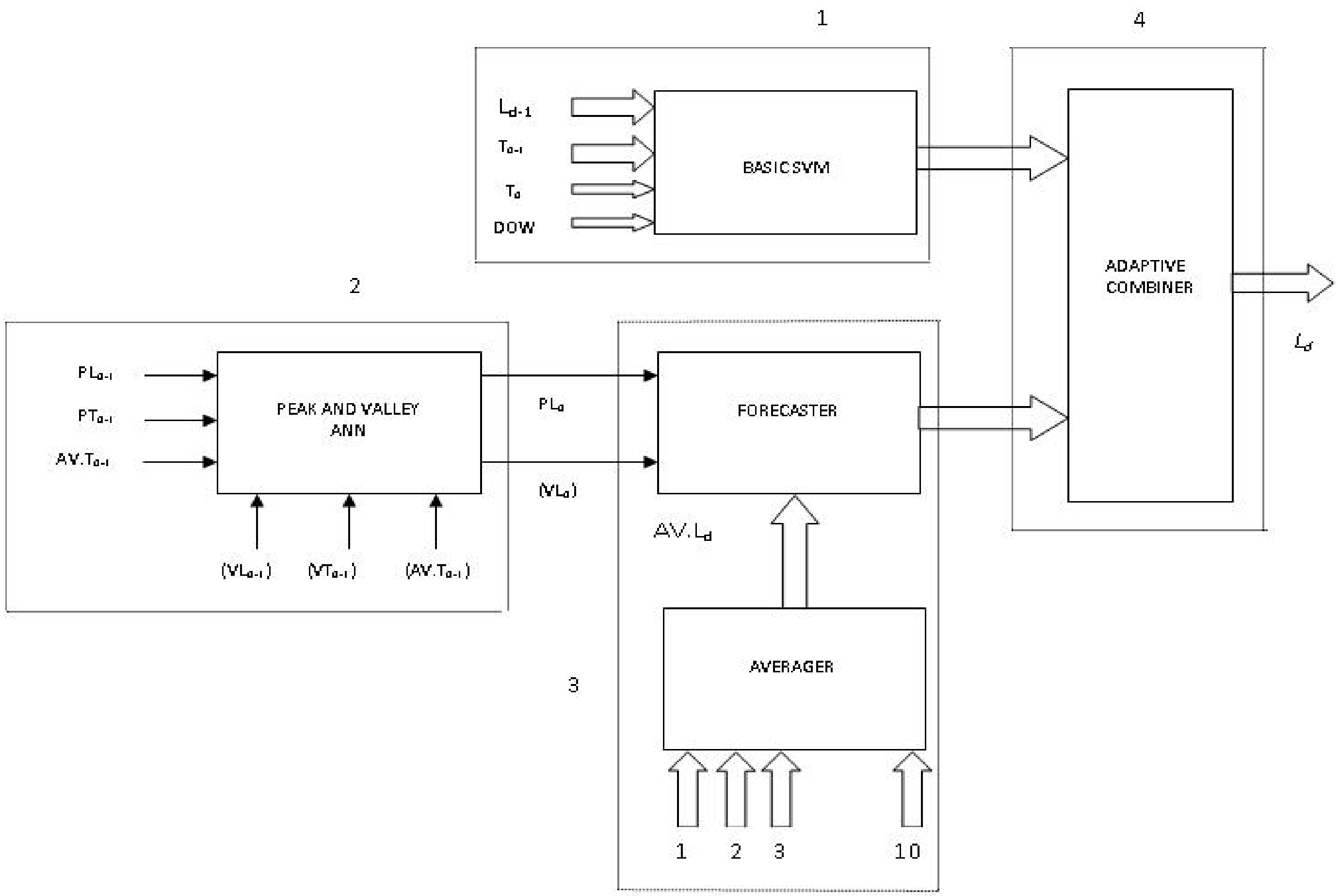


Fig. 1 Proposed Architecture for predicting next '24' hours of load

The proposed architecture is shown in Fig. 1. The power system load depends on several factors such as weather, type of day, hour, human, social and other activities etc., The objective of SVM and ANN based forecaster is to recognize these factors and predict the load accordingly. Thus a suitable architecture along with appropriate inputs is needed. There are no general rules to follow in the selection of input variables. It depends largely on experience, professional judgment and preliminary experimentation. The demand for electricity is known to vary by the time of the day, week, month, temperature and usage habits of the consumers. Though usage habit is not directly observable, it may be implied in the patterns of usage that have occurred in the past. For solving a STLF problem all of these inputs are not needed at the same time. Depending on the forecast to be made, whether daily or hourly; the choice of input variables changes.

A. Description of Proposed Architecture

MODULE 1: BASIC SVM

INPUTS: 54

Load (L_{d-1}): 24

Temperature (T_{d-1}): 24

Forecasted day's maximum, minimum and average Temperatures (T_d):3

Day Of the Week (DOW) to be forecasted: 3 (Sunday-001, Monday-010, Tuesday-011, Wednesday100, Thursday-101, Friday-110, Saturday-111)

OUTPUTS : 24

Forecasted Load (L_d^b): 24

MODULE 2: PEAK AND VALLEY ANN

PEAK LOAD PREDICTION

INPUTS: 3

Peak Load (PL_{d-1}):1

Peak Temperature (PT_{d-1}):1

Average Temperature ($AV.T_{d-1}$):1

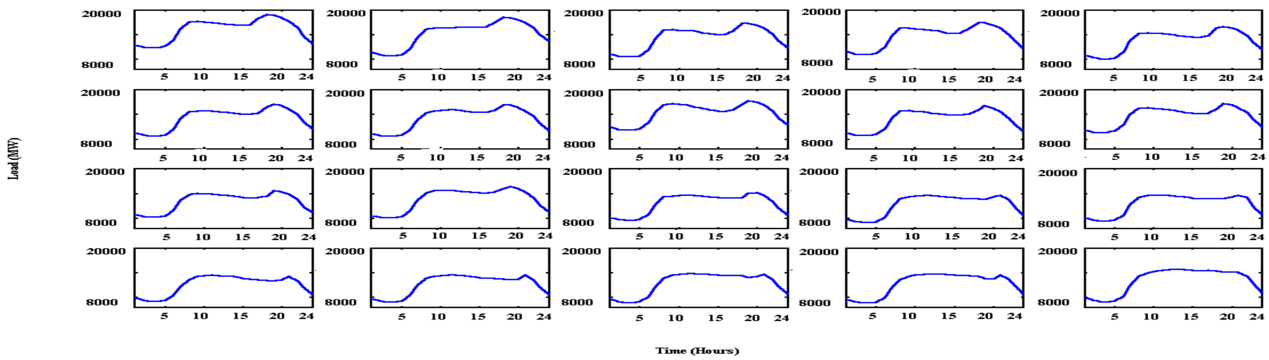


Fig. 2 Training patterns of Thursday

OUTPUT: 1

Forecasted Peak Load (PL_d):1

VALLEY LOAD PREDICTION

INPUTS: 3

Valley Load (VL_{d-1}):1

Valley Temperature (VT_{d-1}):1

Average Temperature ($AV.T_{d-1}$):1

OUTPUT: 1

Forecasted Valley Load (VL_d):1

MODULE 3: AVERAGER AND FORECASTER

AVERAGER [10]:

INPUTS:

Latest 10 patterns of '24' hours load of the day to be predicted.

OUTPUTS: 24

Averaged load of '24' hours.

FORECASTER:

INPUTS: 26

Predicted Peak Load (PL_d):1

Predicted Valley Load (VL_d):1

Averaged Load ($AV.L_d$):24

OUTPUTS: 24

$$L_d^f = VL_d + (PL_d - VL_d)\,(AV.L_d) \qquad (2)$$

Where

VL_d =Forecasted Valley Load

PL_d =Forecasted Peak Load

$AV.L_d$ =Averaged Load

L_d^f =Forecasted '24' hours Load

MODULE IV: ADAPTIVE COMBINER [11]

INPUTS: 48

It is a weighted combination of outputs from the 'Basic SVM' and the outputs from the 'Forecaster', which produces the final load forecast. For weekdays (Tuesday, Wednesday, Thursday and Friday) the '24' hours load curve is divided into four parts of six hours each and the final load is predicted by using the following equations

For 1 to 6 hours, $L_d = 0.15\ L_d^b + 0.85\ L_d^f$ (3)

For 7 to 12 hours, $L_d = 0.10\ L_d^b + 0.90\ L_d^f$ (4)

For 13 to 18 hours, $L_d = 0.50\ L_d^b + 0.50\ L_d^f$ (5)

For 19 to 24 hours, $L_d = 0.15\ L_d^b + 0.85\ L_d^f$ (6)

For weekends (Saturday and Sunday) and the day after holiday (Monday) the final load is predicted by

For 1 to 24 hours, $L_d = 0.50\ L_d^b + 0.5\ L_d^f$ (7)

Where L_d^b is the forecasted load obtained from the 'BASIC SVM' and

L_d^f is the forecasted load obtained from the 'Forecaster'.

OUTPUTS: 24

Forecasted load (L_d): 24

B. Solution Methodology

(i) Creating the Sample Set:

The SVM and ANN are to be trained with the historic data before testing them. The first step for training them is obtaining an accurate historical data. The data should be chosen that is relevant to the model. How well the data is chosen is the defining factor in how well the networks output will match the event being modeled. There should be some correlation between the training data and the testing data. In the load data, in general all the Sunday's load data look alike, all the Monday's data look alike and this holds good for all the days of the week. Hence for testing a day, the training data considered is the past data same as that of the testing day.

Fig 2 shows the 20 training patterns of Thursday. In all the patterns, the two peaks are occurring almost at the same time as well as the valley load. The load at each hour is also nearly same for all the Thursdays. Hence for predicting the load for Thursday, the Basic SVM and Peak and Valley ANN are trained with the past 20 Thursday patterns.

(ii) Data Preparation:

(a) Basic SVM:

In this stage, the typical (raw) input data has to be arranged as input and output pattern pairs for training the SVM. 'Seven' months of past data is used for training and testing the proposed architecture. The 54 inputs and the 24 outputs for the Basic SVM to be arranged as one column vector and the '24'outputs are to be arranged as another column vector. This is to be done for all the days of the past data.

(b) Peak and Valley ANN:

For predicting the Peak load, the data is rearranged as pattern pairs consisting of 3 input and 1 output vectors respectively. Similarly, for Valley load forecasting, the data is arranged as pattern pairs consisting of 3 inputs and 1 output vectors respectively.

(c) Averager:

For Averager, the '24' hours load data of the previous latest 10 patterns same as that of the testing day are averaged hour by hour. This forms the column vector of '24' hours, hourly averaged load of the latest 10 patterns.

(iii) Normalization:

Normalization is an important stage for training the SVM and the neural network. The data is normalized in such a way that the higher values should not suppress the lower values in order to retain the activation function [12].

(iv) Training & Discussion of Results:

The input and output pattern pairs are presented for both the Basic SVM and the Peak and Valley ANN. The SVM and ANN modules are trained with the latest 20 patterns, the optimal learning sequence for STLF [13] and tested with the 21st pattern. The training is done in the

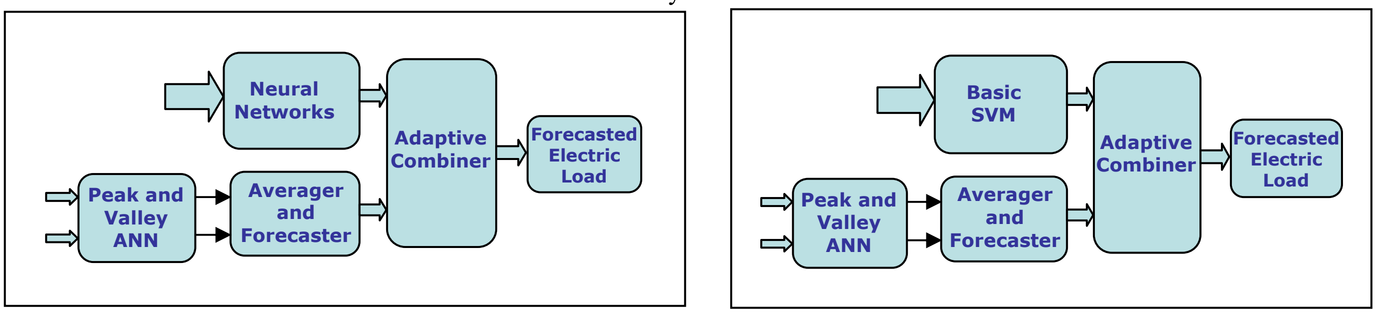


Fig. 3 Two Architectures for comparison of results

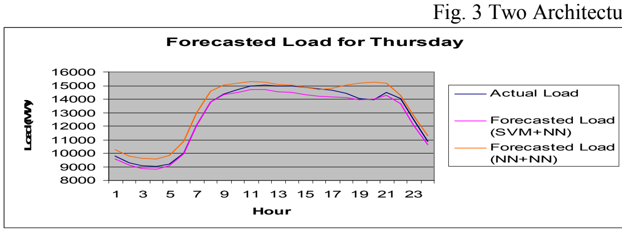


Fig. 4 Comparison of results of the two architectures for Thursday

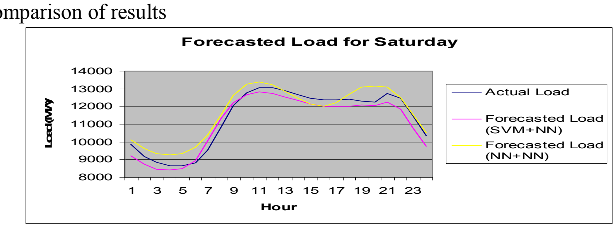
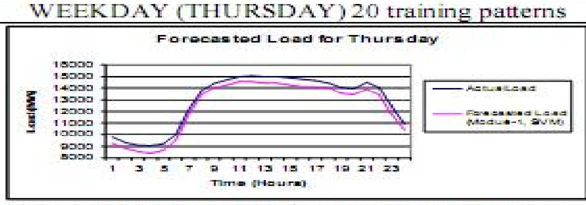
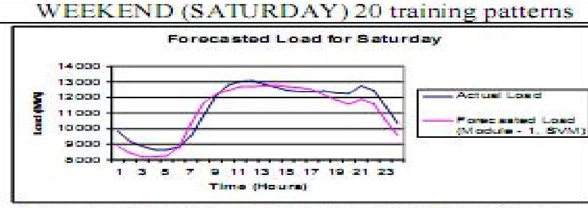
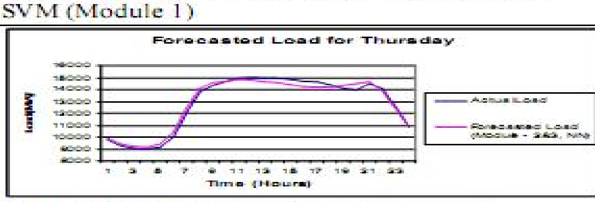
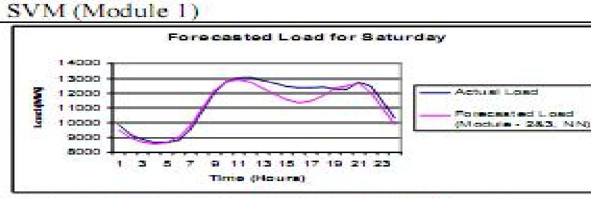
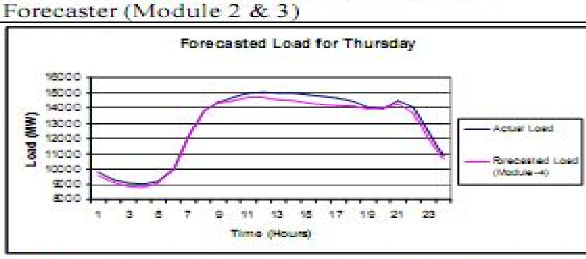
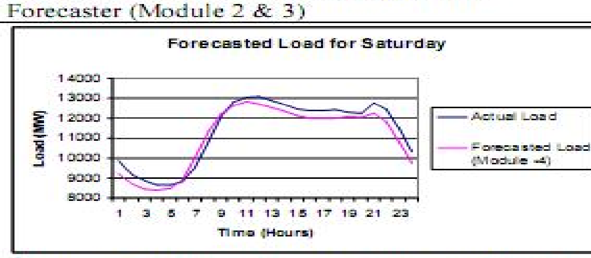


Fig. 5 Comparison of results of the two architectures for Saturday

WEEKDAY (THURSDAY) 20 training patterns	WEEKEND (SATURDAY) 20 training patterns
 Figure 6.1 Comparison between Actual and Basic SVM (Module 1)	 Figure 7.1 Comparison between Actual and Basic SVM (Module 1)
 Figure 6.2 Comparison between Actual and Forecaster (Module 2 & 3)	 Figure 7.2 Comparison between Actual and Forecaster (Module 2 & 3)
 Figure 6.3 Comparison between Actual and Adaptive Combiner (Module 4)	 Figure 7.3 Comparison between Actual and Adaptive Combiner (Module 4)

'batch' mode where all the pattern pairs are recursively applied till the Sum of the Squared Error (SSE) for all the patterns is less than the specified error tolerance value.

The modules are presented with the unknown input pattern (21st pattern) and the output is calculated. The error between the actual output and the calculated output should be with in the tolerance limits.

For comparison of results, the Module -1 is implemented with Neural Networks and the load is predicted for the next '24' hours by keeping the remaining modules intact. The inputs and outputs for the Module -1 with Neural Networks are similar to the Module-1 with Basic SVM. The two architectures for comparison of results are shown in Fig 3.

Fig. 4 and Fig. 5 show the comparison of the actual load, final forecasted loads obtained from the two architectures (shown in Fig. 3) for Thursday and Saturday. The graphs show that the final forecasted load obtained from the architecture containing SVM and neural networks is more accurate compared to the architecture containing neural networks as the modules.

Fig. 6 and Fig. 7 show the load forecasting for the weekday (ex. Thursday) and weekend (ex. Saturday) for outputs of different modules ('Basic SVM', 'Forecaster, 'Adaptive Combiner') for the optimum learning sequence of 20 patterns. The plots are comparison of forecasting results between the actual load with the outputs of different modules It can be observed that for the optimum learning sequence of 20 patterns, the error in load forecasting corresponding to the "Adaptive Combiner" is the least , indicating that its forecasted output is more accurate than either the 'SVM' or the "NN based Forecaster".

V. Conclusions

In this paper, we have presented a new methodology for short-term load forecasting using SVM and ANN based hybrid model, which consists of four modules. The Basic SVM tracks the steady state component of the load. The predicted Peak and Valley loads obtained from the Peak and Valley ANN are accurate compared to the Basic SVM. It tracks the Peak and Valley components of the load patterns. The final forecasted load obtained from the Adaptive Combiner is more accurate than obtained from the Basic SVM and the Forecaster and is following the actual load. So, a conclusion can easily be drawn that the forecasting accuracy of the hybrid model performs better than the individual ones in general. When comparing the hybrid model with either SVM based module-1 or ANN based module-1, it was possible to realize that the hybrid model with SVM approach has showed a better performance than hybrid ANN model, when dealing with the load forecasting topic. Pre-processing of data is a must to get better results. The type of normalization of input and output data has an effect on the forecasting accuracy. The proposed model is shown to be useful for predicting the load without much complexity and showed this model's feasibility in practical application.

VI. References

[1] K. L. Ho, Y. Y. Hsu, C. C. Yang, "Short term load forecasting using a multilayer neural network with an adaptive learning algorithm ", IEEE Trans. on PAS, Vol. 7, No. 1, Feb. 1992, pp. 141-149.

[2] K. Y. Lee, Y. T. Cha, J. H. Park, "Short term load forecasting using an artificial neural network", IEEE Trans. on PAS, Vol. 7, No. 1, Feb. 1992, pp. 124-131.

[3] G. Gross, F. D. Galiana, "Short term load forecasting", Proc. IEEE, Vol. 75, No. 12, Dec. 1987, pp. 1558-1573.

[4] K. S. Swarup, S. Yamashiro, "Neural network based forecasting of daily load ", FANATIC - 99, REC-WARANGAL, pp. 1-11.

[5] Damien Fay, John V. Ringwood, Marissa Condon and Michael Kelly, "24-h electrical load data—a sequential or partitioned time series?", Neurocomputing, 55 (2003), pp. 469 – 498.

[6] Bunn, D. W., "Short Term Forecasting: A review of procedures in the electricity supply industry", Journal of the Operational Research Society, Vol. 33, 1982, pp.533-545.

[7] I. Moghram and S. Rahmq "Analysis and evaluation of five short-term load forecasting techniques", IEEE Transactions on Power Systems, Vol. 4, No. 4, October 1989, pp.1484-1491.

[8] D. C. Park, M. A. Elsharkawi, R. J. Marks, L. E. Atlasm and J. Damborg, "Electric load forecasting using an artificial neural network", IEEE Trans. PAS, Vol. 6, No. 2, May 1991, pp. 442-448.

[9] Ming-Wei Chang et. al, "Eunite Network Competition: Electricity Load Forecasting", 2001, pp.1- 8.

[10] Y. Y. Hsu, C. C. Yang, "Design of artificial neural networks for short term load forecasting, part-2, multilayer feed-forward networks for peak and valley load forecasting", Proceedings IEEE, Vol. 138, No. 5, Sep. 1991, pp. 414-418.

[11] A. Khotanzad, R. A. Rohani, D. Maratukulam, "ANNSTLF-Artificial neural network short term load forecaster-Generation Three", IEEE Trans. PAS, Vol. 13, No. 4, Nov. 1998, pp. 1413-1422.

[12] D. D. Prabhakar, S. Viswanathan, M. G. Raoot, P. Pentayya and N. Nallarsan, "Implementation aspects of artificial neural networks based short term load forecasting method-Experience of WRLDC", National power system conference, Indian Institute of Technology, Kanpur, 1986, pp. 355-359.

[13] T. Dillon and S. Sestito, "'Short term load forecasting using neural networks".