1
Proceedings of the 8th IEEE Mediterranean Conference
on Control and Automation (MED 2000)
Rio, Patras, GREECE 17-19 July, 2000

LONG RANGE FORECASTING OF HOURLY POWER SYSTEM LOAD BY ARTIFICIAL NEURAL NETWORK VIA ORDERED WALSH TRANSFORM

S. G. BHIRUD[†] **and T. K. BASU**[‡,a]

† Department of Electronics and Computer Science, Shri Guru Gobind Singhji College of Engineering & Technology, Nanded-431 606, India.

‡ Department of Electrical Engineering, Indian Institute of Technology, Kharagpur-721 302, India

a Corresponding author e-mail: tkb@iitkgp.ernet.in

Abstract— Long range forecasts of hourly loads spanning 52 weeks (1 year) not only facilitates preparation of capital repair schedules of generating units for preventive maintenance in an integrated system, but may also obviate the need for medium range forecasts in certain situations. The varying nature of power system data having multiple periodicity of 24 hours/168 hours (1 day) /(1 week) makes it suitable for the application of Digital Image Processing technique. An attempt has been made to represent the data in the form of an image replacing the time variables by space variables. Thus the inter pixel gap of the image represents the sampling time of 1 hour along the horizontal axis and 24 hours along the vertical axis. Transforming the image by 2-D Ordered Walsh Transform (OWT) and modeling the OWTs of successive years via an Artificial Neural Network (ANN), forecasts are made by taking inverse of the forecast OWTs from the ANN. The dynamics of the process is captured by an input output relation for the ANN model, the parameters of which can be obtained by training the ANN with a given large set of sample input/output data (OWTs).

Keywords: Long Range Forecast, Artificial Neural Network, Ordered Walsh Transform

I. INTRODUCTION

Forecasting techniques have been in use in power system planning and control since early seventies [1-4]. The growing complexity of interconnected power system requires an accurate forecast of loads at different operational considerations like planning of routine maintenance schedules of unit auxiliaries, or capital repair of generating which are to be drawn up optimally to minimize the gap between demand and supply of power. Spinning reserve, hot reserve and hence the security of supply depend very much on such load demand pattern and hence apriori knowledge of this pattern is very much essential. It is therefore imperative, while modeling, to retain all the specific features of the demand pattern of each weekday in a system round the year. In a complex metropolitan load, industrial holidays are staggered due to generation deficiencies (and hence every day of the week has its own characteristic features of peaks and troughs in the load demand). There are seasonal variations too, in the load pattern of a particular weekday especially in regions with a wide variation of sunrise and sunset times and climatic conditions.

Various causal and noncausal time series models havebeen reported for short and medium range forecasting of power system loads employing different methods[5-21]. But a multistep ahead forecast for reflecting all the different elements of periodicities and trend in a complex power system can not be accomplished by the existing methods due to the facts that

(a) Multiplicative SARIMA models have parameters to be solved from a set of nonlinear equations and a large volume of data is required and

(b) The accuracy of forecast reduces with the step for which the prediction into the future is required.

Due to lack of data in fine details, a causal model is quite often discouraged. In references[22-28], grouping of data invarious forms/discrete fourier Transformer/walsh transformes have been used alongwith Box-Jenking in method to achieve inform level of errors for multistep ahead forecast. The different day groups (Saturday, Sunday) were modelled separately and forecasts were made for the individual days for the seven days group. The common features of different days were thus lost in such subgroup models. Efforts to forecast power system load using ANN have been made in recent years, for short and medium range where causal variables such as the weather conditions, social condition indicator, history of load data etc. are used as inputs, and hourly load for the next few hours have been forecast [29-34]. A good survey of application of ANN and Artificial Intelligence in power system has been made by Vidyasagar et al. [35] and Madan et al. [36].

In the present work the data has been represented in a matrix form and through a two dimensional orthogonal transform the horizontal as well as vertical dependence of the hourly data has been retained, in a very compact form by the fewer dominant components. One is left with the choice of modelling these coefficients over a span of a few years as the image matrix changes with time, through either time series analysis (if the data is available for sufficiently long period) or through Artificial Neural Network (for short period).

The present study has been made with an ANN model where the dominant coefficients of the transformed images of two consecutive years in particular location have been used as input-output sets for training the ANN model.

In most of the parts of India and many other countries, the weather dependence of load is marginal, except in extreme summer or winter conditions (e.g., in the northern parts of India). The present study has been made for Tata Power Company, Mumbai, where the variation of weather round the year is insignificant.

II 2-DIMENSIONAL ORDERED WALSH TRANSFORM

Let (x, y) be the co-ordinates of an N × N square array (image matrix) and f(x, y)be the value of the function (load demand) at this location.

The forward and inverse transforms are given by [37]

$$T(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)\ g(x,y,u,v) \qquad (1)$$

and

$$f(u,v) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} T(u,v)\ h(x,y,u,v) \qquad (2)$$

where $g(x,y,u,v)$ and $h(x,y,u,v)$ are the forward and inverse transformation kernels, which depend only on x, y, u, v and not on the values of $f(x,y)$ or $T(u,v)$.

If the forward kernels are separable into two identical kernals, then

$$g(x,y,u,v) = g_1(x,u)\ g_1(y,v) \qquad (3)$$

A transform with a separable Kernel can be computed in two steps, each requiring a 1-D Transform. and equation (1) can be expressed in matrix form as

$$T = AFA \qquad (4)$$

where,

$$\begin{aligned} F &= N \times N \text{image matrix} \\ A &= N \times N \text{symmetric transformation matrix}[g_1(i,j)]. \\ T &= \text{Resulting } N \times N \text{Transformed image.} \end{aligned}$$

Similarly, if B is the inverse transformation matrix of A then

$$F = BTB \qquad (5)$$

For Ordered Walsh Transform, the forward kernel is given by

$$g(x,u) = \frac{1}{N}\ (-1)^{\sum_{i=0}^{n-1} b_i(x) p_i(n)} \qquad \text{where } N = 2^n$$

and

$$\begin{aligned} p_0(n) &= b_{n-1}(n) \\ p_1(n) &= b_{n-1}(n) + b_{n-2}(n) \\ p_2(n) &= b_{n-2}(n) + b_{n-3}(n) \\ p_1(n) &= b_1(n) + b_0(n) \end{aligned} \qquad (6)$$

and $b_k(z)$ is the k^{th} bit in the binary representation of z. For example, if $n = 3$ and $z = 6$ (110 in binary) then, $b_0(z) = 0$, $b_1(z) = 1$, $b_2(z) = 1$, the summation in (6) are performed in modulo -2 arithmetic. The inverse kernel is

$$h(x,u) = (-1)^{\sum_{i=0}^{n-1} b_i(x) p_i(n)} \qquad (7)$$

The 2-D kernels are separable and identical and we can write

$$H(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) (-1)^{\sum_{i=0}^{n-1} b_i(x) p_i(n) + b_i(y) p_i(v)}$$

and

$$f(x,y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} H(u,v) (-1)^{\sum_{i=0}^{n-1} b_i(x) p_i(n) + b_i(y) p_i(v)} \qquad (8)$$

III. NEURAL NETWORK STRUCTURE

The ANN proposed here is constructed with one input layer with thirty-six neurons, one hidden layer with eleven neurons and an output layer with thirty-six neurons (Fig. 1). For each neuron j in the hidden layer and neuron k in the output layer, the net inputs are given by

$$net_j = \sum_{i-1}^{NI} W_{ji} O_i \qquad j = 1, ..., NJ \tag{9}$$

and

$$net_k = \sum_{j-1}^{NJ} W_{kj} O_j \qquad k = 1, ..., Nk \tag{10}$$

respectively. The neuron outputs are given by

$$\begin{aligned} O_i &= net_i \\ O_j &= \frac{1}{1+e^{-(net_j+\theta_j)}} = f_j(net_j, \theta_j) \\ O_k &= \frac{1}{1+e^{-(net_k+\theta_k)}} = f_k(net_k, \theta_k) \end{aligned} \tag{11}$$

where net_i is the input signal from the external sources to node i in the input layer [38].

Let the connection weights W_{ji} and W_{kj} be updated after each presentation. They are changed after input/output patterns(pairs) have been presented. Then for a presentation , the sum of squared errors to be minimized is given by

$$E_p = \sum_{q=1}^{NQ} \sum_{k=1}^{NK} (t_{pqk} - O_{pqk})^2 \tag{12}$$

where t_{pqk} and O_{pqk} are the target output value and computed output value at output node k for the qth input/output pair in presentation p, respectively. By minimizing the error E_p using the technique of gradient descent, the connection weights between hidden unit j and output unit k can be updated.

IV. STATEMENT OF THE PROBLEM AND MATHEMATICAL MODEL

For power System data with periodicities of 24 hours/one week/or one year, (up to a few hours), to obtain a multistep ahead forecast with same variance of error the following ideas have been used.
(a) The hourly load data of a particular year 'k' are grouped together in a matrix form A^k ($365/366 \times 24$) as shown in Fig. 2(a).
(b) The matrix A^k is now modified to A_c^k in the following way

(i) At the end of 364 days (52 weeks), the extra one day (or 2 days in case of leap year) is transferred to the next year. This will be considered as the first day,(or days) of the next year, (accumulation of these extra days will cause a shift of one annual week after 5 or 6 years depending on the occurrence of leap-years; after this the location can be again readjusted), Fig. 2(b).

(ii) The matrix A_c^k is now partitioned into $52 \times 3(=156)$ segments $[A_{ij}]^k$ i= 1,2,.,52, j= 1,2,3 where A'_{ij}s are of dimensions (7×8), as shown in Fig. 2(c).

(iii) Each of the submatrices A_{ij} (7×8), Fig. 2(d) is now further modified to an (8×8) square matrix B_{ij} , Fig. 2(e), the last row containing the elements which are the averages of the 7 values in a particular column. Thus, if the original elements of a submatrix A_{ij} is given by a_{mn}, then $b_{8,n} = \sum_{m=1}^{7} \frac{a_{m,n}}{7}$, where $b_{8,n}$ are the elements of the 8th row of B_{ij}. This is average padding. Thus the submatrices are now complete matrices (2^N , N=3). This will ensure a square image size of 8×8. Thus we get a partitioned matrix as shown in Fig. 2(f).

(iv) Each submatrix is replaced by its transform which is sparse.

(v) A similar set of dominant coefficients d_m^{k+1} for the next year's corresponding block A_{ij}^{k+1} is obtained, and a two layer ANN is trained with d_m^k and d_m^{k+1} (m = 1,2,3) for a large number of pairs of matrices (B_i^k/B_i^{k+1}) selected randomly from among $156 \times$ N such pairs, where N = Number of years considered for modeling. We have considered five consecutive years from 1991 to 1995. After training the neural network with a few years data, the coefficients are stored. 36 dominant coefficients have been combined.

(vi) Now with a given year's data, following the step (i)-(v), the 156 submatrices of OWT coefficients d_m^{p+1} are obtained and for a particular value of p, all the 36 elements of d_m^p are fed as input to the ANN to obtain the corresponding output elements of d_m^{p+1}. Each of these submatrices is inverse transformed and we get one corresponding year's augmented data for 52 weeks. By removing the last row (8th day) of each submatrix we get the complete year's projection.

V. BAD DATA DETECTION

Criteria to detect a certain data to be bad and to apply a necessary correction require a lot of considerations.

Here the method suggested in references[23] and [26] have been adopted to detect and filter out bad data.

VI. RESULTS

The yearly data of hourly load of TATA Power Company Mumbai for the years 1991-1996 have been used for the present study.

The input layer to hidden layer weights W11 (11×36) and hidden layer to output layer weights W22 (36×11) alongwith the bias B11 (1×11) and B22 (36×1) are computed after several iterations and the forecast is made for the whole year of 1996. Fig. 3 shows some typical days forecast and actual values. Comparison is made with the actual corrected data. For example, the worst case situation, Nov. 14, 1996 shows that the actual demand was much less than expected demand. This may be due to system outages or any special event, (actually it was the Children's day and 'Deepavali' festival). Since most of the Indian festivals do not fall on the same calendar day, their effects on power system load demands do not show any systematic pattern. Most of these festivals are associated with reduced industrial load demands as well. The best case is also shown which occurs on Jan. 25, 1996.

Sample forecasts for different days spread over the entire year show a uniform level of accuracy. The Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) for the year 1996 are found to be 6.6876% and 108.3380 MW respectively.

VII. CONCLUSION

In this paper the technique of Digital Image Processing technique have been applied for long range forecasting of hourly power system load. The power system data with multiple periodicity of 24 hours/168 hours have been represented in the form of image and transformed by Ordered Walsh Transform. Using OWT coefficients an artificial neural network has been trained to obtain a model for successive years. From the Figure 3 it is found that the approach developed in this paper for the multistep ahead forecast for long range, (spanning 1 year) by 2-dimentional Ordered Walsh Transform via Neural Network yields satisfactory results.

VIII. ACKNOWLEDGMENTS

The authors wish to thank the Tata Electric Power Company, Mumbai for their assistance in the collection of data. They also wish to thank the Director, I.I.T., Kharagpur and Principal, S.G.G.S. College of Engineering and Technology, Nanded for providing the necessary facilities.

References

[1] Schweppe, F C Wildes, J 'Power system static state estimations' Part I, II, III, IEEE Trans. Power Apparatus and Systems, PAS-89 (1970) 120-135.

[2] Mahalanabis, A K Handmandlu, M and Rana, R S 'Load modeling of An interconnected power system for short term prediction' Int. J. Sys. Sci., Vol 2 (1980) 445-453

[3] Hagan, M and Klein, R 'Identification techniques of Box-Jenkins, applied to the problem of short term load forecasting' IEEE Summer Power Meeting paper No a 77 618-2 (1977)

[4] Christiaanse, W R 'Short term load forecasting using exponential smoothing' IEEE Trans. PAS-90 No 2 (1971) 900-911

[5] Corpenning, S L Reppen, N D and Ringlee, R J 'Experience with weather sensitive load models for short and long term forecasting' IEEE Trans. PAS-92 No 6 (1973) 1966-1972

[6] Carta, G Grossl G and Manni, R 'Computer processing of data required for the probabilistic forecasting of daily peak power demand by means of Markovian processes' Proc. IFAC, New Delhi, (1979) 269-276

[7] Brubacher, S R and Wilson, G T 'Interpolating time series with application to the estimation of holiday effects on electricity demand' Appl. Statist., Vol 25 No 2 (1976) 107-116

[8] Abu-El-Magad, M A and Sinha, N K 'Two new algorithms for on-line modeling and forecasting of load demand of multinode power systems' IEEE Trans. Vol PAS-100 No 7 (1981) 3246-3252

[9] Abu-El-Magad, M A 'An interactive algorithm for short term load forecasting of multinode power systems' Faculty of Eng. McMaster Univ. Ontario, Canada, Rep. Soc-260 (1981)

[10] Abu-El-Magad, M A and Sinha, N K 'Short term load demand modeling and forecasting' IEEE Trans. Syst., Man Cybern., Vol SMC-12 No 3 (1982) 370-382

[11] Ganesino, R Vicino, A and Aneilli, R 'Bivariate load temprature modelling for short term forecasting' Proc. 2nd Forecasting Symp. Istambul. Turkey, (1982)

[12] Koh, H S and Chang, J K 'Long term maximum power demand forecasting in consideration of dry bulb temperature' Trans. Konean Inst. Electr. Eng., Vol 34 No 10 (1985) 398-398

[13] Zhu, A 'Short term forecasting of grid load using time series analysis adaptive model' Autom Elect. Power System (China) Vol 8(2) 12-19 (1982)

[14] Bakar, A B 'Load forecasting for scheduling generation on a large interconnected system', in Comparative Models for Electrical Load Forecasting, ed. Bunn, D. W. and Farmer, E. D., (1985)

[15] Goh, T N and Choi, S S 'Short term forecasting of electricity demand by decomposition analysis' Eur. J. Opl. Res., Vol 17 (1984) 79-84

[16] Goh, T N Choi, S S Tan, C H and Tan, K E 'Comparative study of short term load forecasting of energy and peak demand' EPSR, Vol 5 (1982) 63-71

[17] Gross, G and Galiana, F D 'Short term load forecasting' Proc. IEEE, Special issue on computers in power system operations, (1987) 1558-1573

[18] Bunn,D R and Seigal, J P 'Forecasting the effects of Television programming upon electricity loads' J. Opl. Res. Soc. Vol 34 (1983) 17-25

[19] Caprio, U et. al. 'Short term load forecasting in electric power systems: a comparison of ARMA model and extended Wiener filtering' J. Forecasting, Vol 12 (1984) 59-76

[20] Hubele, N F and Cheng,C S 'Identification of seasonal short term load forecasting models using statistical design function' IEEE Trans. PAS Vol 5 No 1 (1990) 40-43

[21] Bajoy, S V 'Long term electricity demand forecasting models: a review of methodologies' EPSR (Switzerland) Vol 5 No 4 (1983) 243-247

[22] Sarma, U K Basu, T K and Sinha, S ''Medium range forecasting of power system load (energy) demand' Int. J. Syst. Sci., Vol 9 (1987) 1691-1702

[23] Sarma, U K Basu, T K and Sinha, S 'Medium range forecasting of power system load via Discrete Fourier Transform' Int. Conf. On System Science, held at Wrocklaw, Poland, Sept. 16-19th (1987)

[24] Sarma, U K Basu, T K and Sinha, S 'Medium range forecasting for power system load by Fast Fourier Transform' J. of Indian Institute of Science, (1987) 241-247

[25] Basu, T K Bhattacharya, T K and Purkayastha, P 'Medium range forecasting of hourly power system load by time series analysis using the Walsh transform' Int. J. Electr. Power and Energy Syst., Vol 13 No 4 (1991) 193-200

[26] Bhattacharya, T K and Basu, T K 'Medium range forecasting of power system load using modified Kalman filter and Walsh transform' Int. J. Electr. Power and Energy Syst., Vol 15 No 2 (1993) 109-115

[27] Bhattacharya, T K and Basu, T K 'A new method of medium range forecasting by Walsh spectrum analysis' First symp. on modeling and simulation of lumped and distributed parameter system, Lille-France,(1986)

[28] Bhattacharya, T K and Basu, T K 'Long range forecasting of peak load demand by Walsh spectrum analysis' International conference on systems science, Wroclaw, Poland, (1986)

[29] Srinivasan, D Liew, A C and Chen, J S P 'A Novel approach to elect. load forecasting based on neural network' Proc. IEEE Int. Joint Conf. Neural N/W. (IJCNN) Singapore, Vol 2 (1991) 1172-1177

[30] Park, D C El-Sharkari, M A Mark, R Allas, J L E and Daonborg, M J 'Electrical load forecasting using an artificial neural network' IEEE Trans. Power System. Vol 6 (1991) 442-449

[31] Shrinivasan, D Liew, A C and Chang, C S 'A neural network short-term load forecaster' EPSR, No 28 (1994) 227-234

[32] Kiartzis, S J Bakritzis, A G and Petridis, V 'Short term load forecasting using neural network' ESPR, No 33 (1995) 1-6

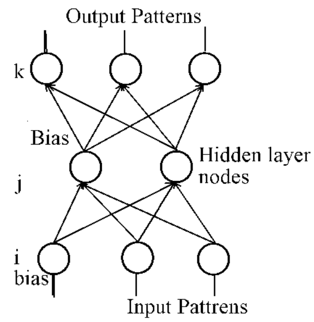
[33] Chen, S T Yu, D C and Maghaddamjo, A K 'Weather sensitive short term load forecasting using nonfully connected artificial neural network' IEEE Trans. Power System. Vol 7 (1991) 10984-1105

[34] Dash, P K Liew, A C and Rahman, S 'Peak load forecasting using a fuzzy neural network' EPSR, No 32 (1995) 19-23

[35] VidyaSagar, S and Rao, ND 'Artificial neural networks and their applications to power systems: a bibliographical survey' EPSR, No 28 (1993) 67-73

[36] Madan, S and Bollinger, K E 'Applications of artificial intelligence in power systems' EPSR, No 41 (1997) 117-131

[37] Gonzalez, R C and Woods, R E Digital Image Processing, Addison Wesely Publication Company, USA.



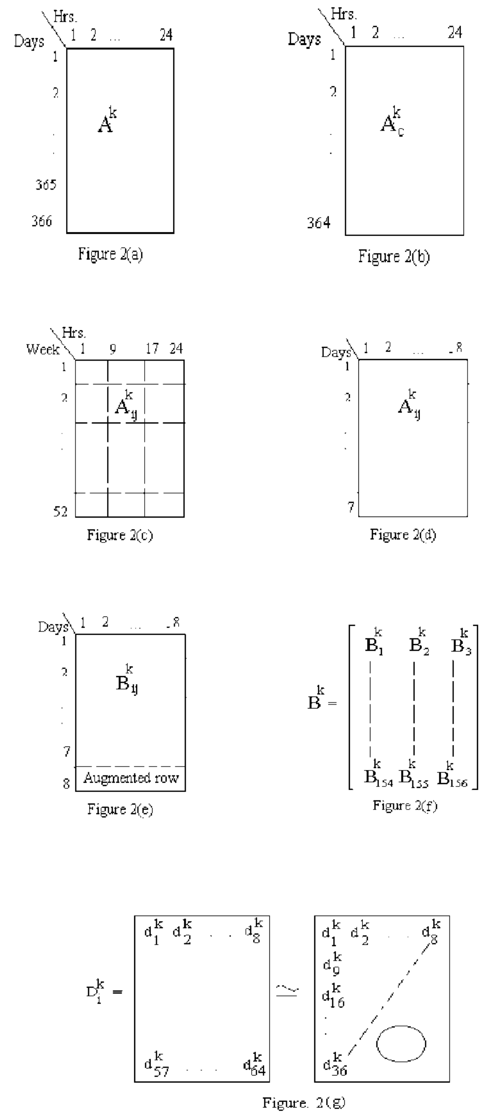


Fig.2. Different structures of hourly load data matrix for 1 year. (a) hourly load data of a particular year k (A^k), (b) modified hourly load data of a particular year k (A_c^k), (c) partitioned A_c^k (156 segments), (d) one of the segments of A_c^k (A_{ij}^k), (7x8), (e) augmented A_{ij}^k, (8x8), (f) complete submatrices of A_c^k, (g) prominent OWT coefficients of B_{ij}^k

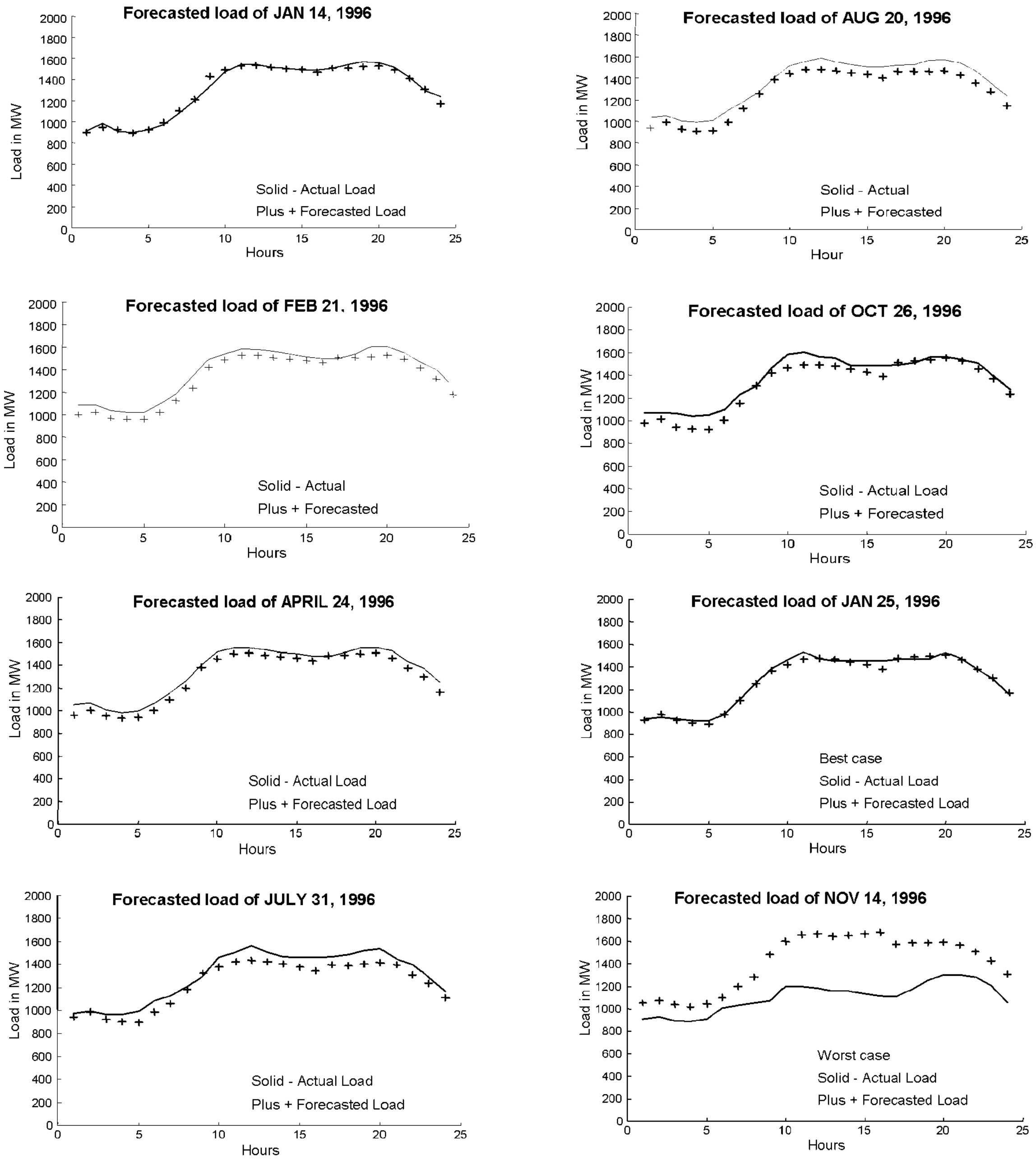


Fig.3. Forecast and actual values of some typical days including best and worst cases for the year 1996