Short Term Load Forecasting by Using Time Series Analysis through Smoothing Techniques

Mr. Dileshwar Prasad Patel [1], Assit.Prof.Amit Vajpayee [2] Assit.Prof. Jitendra Dangra [3]
1 LNCT, Indore (M.P), 2 LNCT, Indore (M.P), 3 LNCT, Indore (M.P)

Abstract— **as the accuracy of the electricity load forecast is crucial in providing better cost effective risk management plans; this paper proposes a Short Term Electricity Load Forecast (STLF) model with high forecasting accuracy.A lot of researchers and academicians are engaged in the activity of developing tools and algorithms for load and price forecasting. This paper provides a highly accurate yet efficient load forecasting tools based on time series analysis: smoothing techniques model. Results and discussions from real-world case studies based on the electricity markets of Maharashtra Electricity Board are presented.**

Index Terms—**Short-Term Load Forecasting, Maharashtra Data, Time Series Analysis**.

I. INTRODUCTION

Load forecasting is a vital and fundamental factor for a successful operation of energy system. In order to operate the electrical system effectively and efficiently, the system load should be correctly predicted. If the system load forecast is overstated, the system may over-commit the generation and leads to costly operation. On the other hand, if the system load forecast is understated, the reliability and security of the system may be compromised which may result in power interruptions. While load forecasting in a longer lead time is applied for long-term maintenance scheduling, power-sharing coordination or capacity expansion, short-term load forecasting is needed for economic scheduling of the generating units, security analysis and short-term maintenance.

Several different methods and techniques have been developed to forecast load demands, with the focus on improving the prediction accuracy. The approach using time series analysis is among the main areas with rich research effort, with specially formulated methods for data in various contexts.

In Maharashtra's context, research efforts in load forecasting have been devoted primarily to hybrid fuzzy-neural approach. A one-day ahead load forecasting technique using hybrid fuzzy-neural networks is explained in details in. Due to the vastly different behaviour of loads in weekends and public holidays from weekdays, special attention is paid to such daysfor performance improvement. Reference proposed the implementation and evaluation of the parallel neural network-fuzzy expert system for local load demand forecasting.

Two time series models, namely, the multiplicative decomposition model and the smoothing techniques model employed. The multiplicative decomposition technique has been profoundly used for forecasting tasks in the business sector, such as sales projection or financial forecasting.

However, it has not been commonly employed for electricity demand forecasting. This may due to the fact that load demands usually vary to a large extent across seasons with changing temperatures and it is challenging to fit a trend line. In Maharashtra, on the other hand, the weather varies within a small degree and trend lines are more easily predicted. Thus, this technique is proposed in this paper. The proposed models are implemented to predict one week demand data. The accuracy of the two models are calculated and compared. The paper utilizes the mean absolute percentage error (MAPE) as a measure of forecast accuracy.

The paper is organized as follows. Section II introduces the two abovementioned models and explains how they would be used in analysing and forecasting the load consumption with a brief note on forecasting error measurement. Section III is the case study whereby Maharashtra load demands are first plotted For visual comparison of weekday and weekend data. Then models would be applied with forecasts generated and compared with the amounts actually consumed.

II. TIME SERIES MODELS

A time series is defined as a set of data generated sequentially in time. The time series models assume that in the absence of major disruptions to critical factors of a recurring event, the data of this event in the future will be related to that of the past events and can be expressed via models developed from the past events. In this analysis, time series model: smoothing techniques Model, are employed and presented in the followings.

III.SMOOTHINGTECHNIQUE

This chapter we are going to cover the Different smoothing techniques and procedure to use those techniques.

INTRODUCTION TO SMOOTHING TECHNIQUES

Smoothing techniques are used to reduce irregularities

(random fluctuations) in time series data. They provide a clearer view of the true underlying behaviour of the series. Moving averages rank among the most popular techniques for the pre-processing of time series. They are used to filter random "white noise" from the data, to make the time series smoother or even to emphasize certain informational components contained in the time series.

DIFFERENT SMOOTHING TECHNIQUES

Many Smoothing Techniques are there but Moving Averages is the popular technique to smooth the data. So we are explaining the Moving Average Techniques below,

Moving Averages: A moving average (MA) is an average of data for a certain number of time periods. It "moves" because for each calculation, we use the latest x number of time periods' data. Using an average of prices, moving averages smooth a data series and make it easier to spot trends. This can be especially helpful in volatile markets.

Types of Moving Average: There are three major types of Moving Averages: Simple Moving average, Expontional Moving average and weighted moving average.

Simple Moving Average (SMA)

SMA is the most basic of the moving averages used for trading. A simple moving average (SMA) is formed by finding the average price of a currency or commodity over a set number of periods. Most often, the closing price is used to compute the moving average. A 5-day simple moving average is the five day sum of closing prices divided by five. As its name implies, a moving average is an average that moves. Old data is dropped as new data comes available.

Exponential Moving Averages (EMA)

Exponential moving averages reduce the lag by applying more weight to recent prices. The weighting applied to the most recent price depends on the number of periods in the moving average. An exponential moving average is similar to a simple moving average, but whereas a simple moving average removes the oldest prices as new prices become available, an exponential moving average calculates the average of all historical ranges, starting at the point you specify.

There are three steps to calculating an exponential moving average. First, calculate the simple moving average. An exponential moving average (EMA) has to start somewhere so a simple moving average is used as

the previous period's EMA in the first calculation. Second, calculate the weighting multiplier. Third, calculate the exponential moving average. The formula of EMA is shown below:

Multiplier: $(2 / (\text{Time periods} + 1))$

EMA: $\{\text{Close} - \text{SMA (previous day)}\} \times \text{multiplier} + \text{EMA (previous day)}$.

Or

We can also write this formula like:
$(\text{SMA} + \text{multiplier} (\text{Last value} - \text{SMA})$

Weighted Moving Average (WMA)

This average is calculated by multiplying each of the previous day's data by weight. Therefore it is designed to put more weight on recent data and less weight on past data. In other words, it is simply a moving average that is weighted so that more recent values are more heavily weighted than values further in the past. Evidence also indicates that the use of this type of moving average gives better volatility estimates than the simple moving average.

The formula of the WMA average is:

$((n * P_n) + ((n - 1) * P_{n-1}) + ((n - 2) * P_{n-2}) + ... ((n - (n - 1)) * p_{n-(n-1)}) / n + (n-1) + (n- (n-1))$.

Here n is the number of terms present; p is the values given there.

Data Set: I have the data of Maharashtra state electricity distribution co. ltd. we took the data from the site www.mahadiscom.in. The monthly data is available in that site. The data is available from December 2007 to April 2010

We selected the data from April 2008 to April 2010. Every month's data is divided like Zone wise, Circle wise and Hour wise. There are 4-5 weeks are available in every month and every week has an hourly data.

We had chosen pune zone, ganeshkhind (u) circle. From this zone, i had selected 5 areas of different feeders, but the division is same. The selected feeder names are: kasturba, sangam bridge, gk – i, proof range and may fair tower. here is some brief explanation of all the areas:

HR DATE 01-APR-08 02-APR-08 03-APR-08 04-APR-08 05-APR-08 06-APR-08 07-APR08

1	1.6118	1.6118	2.2539	1.1641 1.	1641	1.267	1.4327
2	1.6118	1.5223	2.191	1.1517	1.1641	1.1765	1.267
3	1.5223	1.5223	2.0805	1.0631	1.1641	1.1765	1.267
4	1.5223	1.6032	2.0157	1.0631	1.1641	1.1765	1.267
5	1.5384	1.6289	2.191	1.1517	1.1641	1.1765	1.267
6	1.7375	1.829	2.7149	1.7194	1.267	1.1888	1.3717
7	2.0119	2.1253	2.9264	1.5708	1.4632	1.2936	1.5708
8	2.0119	2.2177	2.9264	1.5708	1.4784	1.4784	1.6461
9	2.1948	2.3529	3.2236	1.7375	1.4784	1.5708	2.0119
10	2.5606	2.3529	3.4027	1.9204	1.8099	1.6461	2.1719
11	2.5968	2.5692	3.6323	2.1948	1.9909	1.7909	2.3529
12	2.5692	2.5415	3.1207	2.1491	1.949	1.8604	2.3282
13	2.5692	2.4539	3.1722	2.1262	1.949	1.6651	2.2386
14	2.5072	2.3034	3.3808	1.97	1.5947	1.6651	2.1491
15	2.392	2.2786	3.4675	1.9909	1.6832	1.5947	2.1491
16	2.4177	2.4177	3.3303	1.9909	1.7909	1.5947	2.0595
17	2.4806	2.4806	2.0376	2.1491	1.7909	1.6832	2.1262
18	2.5415	2.4539	1.9938	2.1262	1.8604	1.7528	2.1033
19	2.9797	2.8921	2.2786	2.4806	2.2786	2.191	2.4539
20	2.9797	2.9797	2.2786	2.4806	2.2786	2.2786	2.5415
21	2.6292	.0175	1.9938	2.191	2.1033	2.2786	2.1033
22	2.5415	3.4675	1.7718	1.8604	1.7528	1.8604	1.8604
23	2.2148	3.034	1.5061	1.5947	1.5061	1.7718	1.6118
24	1.8604	2.6006	1.3146	1.2403	1.286	1.4327	1.4175

Fig. (1) Raw data

This is the weekly data with 24 hours each, Like that I have a data of 2 years that is from April 2008 to April 2010 of 5 areas. The three variables of the data are:
HR – hour
Date
Day – Monday to Sunday

I used this data for the predictive analysis. Using the different smoothing techniques on the data, I had predicted the future values, their load and I also calculate their accuracy as well.

IV. SIMPLE MOVING AVERAGE

This Chapter explains how we can calculate the SMA and steeps to perform the SMA on our data set. How Pre-process the raw data, also learn to use MS- Excel for Time Series Analysis.

INTRODUCTION TO SIMPLE MOVINGAVERAGE

The moving average which serves as an estimate of the next period value of a variable given a period of length "n". For example, a 10-day simple moving average of closing price is the mean of the previous 10 days' closing prices. If those prices are then the formula is

$$SMA = \frac{p_M + p_{M-1} + \cdots + p_{M-9}}{}$$

STEPS USE TO CALCULATE THE SMA

Step 1: We down load the raw data from the web site www.mahadiscom.in. Which is an official web site of the Maharashtra Electric Board after that we extract the Zip file in the folder and we select our area and after getting that area we copy that data into the Note pad to convert it in to CSV file? Once the data is converted into the CSV file we use the MS-EXCEL for our work

DAY	DATE	1HR	2HR	3HR	4HR	5HR	6HR
TUESDAY	1-Apr-08	0.4448	0.4227	0.4074	0.4346	0.4444	0.4881
WEDNESDAY	2-Apr-08	0.463	0.4486	0.4469	0.427	0.4449	0.4876
THURSDAY	3-Apr-08	0.4048	0.41	0.4074	0.3932	0.4391	0.4731
FRIDAY	4-Apr-08	0.3388	0.3407	0.3218	0.3438	0.3224	0.4067
SATURDAY	5-Apr-08	1.46	1.3342	1.3377	1.2566	1.2754	1.5044
SUNDAY	6-Apr-08	1.4072	1.3311	1.2814	1.2777	1.3817	1.4444
MONDAY	7-Apr-08	1.3365	1.2991	1.2602	1.2951	1.3854	1.5131
TUESDAY	8-Apr-08	1.4295	1.3451	1.3626	1.3497	1.3264	1.7068
WEDNESDAY	9-Apr-08	1.5399	1.4605	1.3666	1.3457	1.3199	1.6919
THURSDAY	10-Apr-08	0.986	0.9514	0.9124	0.9584	1.0298	1.1214
FRIDAY	11-Apr-08	1.4923	1.3957	1.3519	1.342	1.279	1.5157
SATURDAY	12-Apr-08	1.4616	1.398	1.3487	1.3485	1.4058	1.4825
SUNDAY	13-Apr-08	1.4349	1.3097	1.3058	1.2754	1.3842	1.4586
MONDAY	14-Apr-08	1.4061	1.3605	1.2987	1.2754	1.364	1.4481
TUESDAY	15-Apr-08	1.4765	1.3929	1.3444	1.4175	1.4411	1.4784
WEDNESDAY	16-Apr-08	1.589	1.5202	1.514	1.4372	1.6159	1.7156
THURSDAY	17-Apr-08	1.6909	1.644	1.5829	1.5425	1.667	1.7234
FRIDAY	18-Apr-08	1.813	1.6387	1.6954	1.6449	1.548	1.7545
SATURDAY	19-Apr-08	1.0681	1.0183	0.9611	1.0444	0.9812	1.0589
SUNDAY	20-Apr-08	1.4306	1.2741	1.258	1.2802	1.2227	1.3983
MONDAY	21-Apr-08	1.422	1.3726	1.3246	1.334	1.297	1.3708
TUESDAY	22-Apr-08	0.7455	0.6946	0.659	0.6518	0.6728	0.6852
WEDNESDAY	23-Apr-08	0.7599	0.7099	0.6914	0.6523	0.6896	0.6511
THURSDAY	24-Apr-08	0.7446	0.7381	0.7341	0.6683	0.7302	0.7409
FRIDAY	25-Apr-08	0.8037	0.705	0.7336	0.6966	0.7256	0.7202
SATURDAY	26-Apr-08	0.8036	0.7965	0.7565	0.7056	0.7102	0.7262
SUNDAY	27-Apr-08	0.7771	0.7343	0.7148	0.6845	0.6741	0.6716
MONDAY	28-Apr-08	0.7333	0.682	0.6679	0.6464	0.6759	0.6572
TUESDAY	29-Apr-08	0.7735	0.7699	0.7663	0.6928	0.7189	0.7024
WEDNESDAY	30-Apr-08	0.6827	0.6416	0.6451	0.653	0.6511	0.6698

Fig. (2) Sangam nagar's data on Excel sheet

Step 2: After converting the data into the CSV format we use the MS-EXCEL for the calculating the SMA for that we have to arrange the same day data together for that we are using the FILTER option in the EXCEL with the help of the filter we are able to separate the specific data from the hole data. I had selected all the Tuesdayof 2 month is from December 2007 to January 2008.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
			1	2	3	4	5	6	7	8	9	10	11	12
1	Day	DATE												
2	Tuesday	04-Dec-07	0.051	0.051	0.048	0.048	0.048	0.034	0.044	0.067	0.067	0.067	0.067	0.067
3	Tuesday	11-Dec-07	0.64	0.64	0.64	0.675	0.817	0.034	0.034	0.067	0.908	0.943	0.873	0.873
4	Tuesday	18-Dec-07	0.652	0.652	0.652	0.652	0.686	0.72	0.067	0.067	0.067	1.063	0.873	0.067
5	Tuesday	25-Dec-07	0.048	0.048	0.048	0.048	0.048	0.05	0.057	0.042	0.057	0.05	0.076	0.05
6	Tuesday	01-Jan-08	0.857	0.823	0.823	0.823	0.857	1.029	0.05	0.042	0.05	1.029	1.029	1.458
7	Tuesday	08-Jan-08	0.084	0.084	0.084	0.084	0.088	0.095	0.095	0.05	0.05	0.172	0.181	0.172
8	Tuesday	15-Jan-08	0.069	0.069	0.069	0.069	0.069	0.172	0.172	0.172	0.206	0.172	0.206	0.309
9	Tuesday	22-Jan-08	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.257	0.343	0.429
10	Tuesday	29-Jan-08	1.029	0.96	0.943	1.029	1.2	1.972	1.972	1.972	1.972	1.543	1.8	1.543

Fig.(3) Sangam Nagar's data after using the Filter.

Step 3: SMA is Average value so we are using the simple formula to calculate the SMA the formula is SMA= sum of number of values / N, where N is the total numbers of values present in the sample. We are using the same formula to calculate the SMA in our data set.

Q	R	S	T	U	V	W	X	Y	Z	AA	AE
15	16	17	18	19	20	21	22	23	24		
0.061	0.057	0.072	0.067	0.072	0.065	0.061	0.061	0.048	0.046	1.4004	
0.061	0.926	1.029	1.315	1.187	1.102	1.102	1.022	0.846	0.846	18.323	
0.873	0.96	0.995	1.372	1.303	1.2	0.995	0.755	0.789	0.686	17.081	
0.067	0.067	0.076	0.086	0.086	1.102	0.057	0.076	0.095	0.095	2.542	
0.067	1.509	1.372	0.549	0.686	0.652	0.583	0.514	0.48	0.446	17.471	
0.873	0.133	0.141	0.13	0.13	0.122	0.114	0.103	0.095	0.088	4.9107	
0.12	0.051	0.051	0.051	0.051	0.122	0.103	0.069	0.069	0.069	2.917	
0.343	0.343	0.343	0.343	0.343	0.429	0.343	0.257	0.172	0.172	5.4865	
1.543	1.115	1.286	1.629	1.972	1.715	1.543	1.458	1.286	1.2	35.768	

Fig. (4) Sangam Nagar's data after applying the SMA.

Here we are showing SMA techniques used in the all the Tuesdays of the Sangam Nagar.

Step 4:
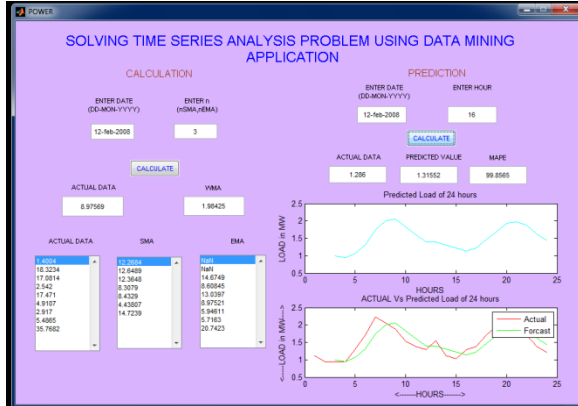
Fig (5)Matlab GUI represents the calculated SMA and Predicted value.

This is our MATLAB GUI which is used in our data set for calculating the 2PSMA, 3PSMA, 4PSMA, 5PSMA, 6PSMA, 7PSMA to predict the load of upcoming Tuesday. We are taking the data from our data set to solve these calculations.

3PSMA = (N1+N2+N3)/N

For 4PSMA =(N1+N2+N3+N4)/N

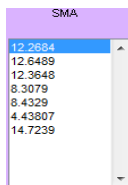Like this I apply this formula up to 7PSMA in our data set.

Fig. (6) Calculated values of SMA.

This is the calculated 3PSMA.

7.4 Calculated values of SMA.

SUM= It is the additions of the total numbers of values present on the data set.

MEAN =Sum of vales / Total number

MAPE=Mean absolute percentage error is measure of accuracy in a fitted time series.

$$M = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

Where A_t is the actual value and F_t is the forecast value.

ACTUAL VALUE= It the original value

PREDICATED VALUE = It is the calculated value.

ACCURACY= It is a measurement of a correctness; the accuracy of the model is near about 98%.
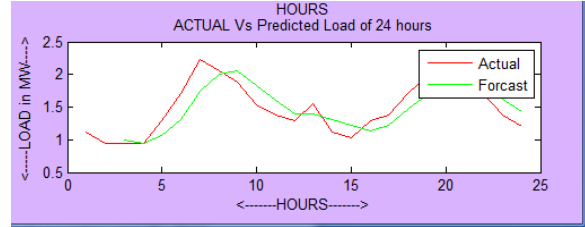
ACCURACY = MAPE*100

Fig (7) Graph between actual load and predicted load.

This is the graph which shows the comparison between 'Actual Value' and Predicted value. Where, Red line denotes the Actual data and Light Green line denotes the Predicted value.

In this way, we can calculate Simple Moving Average for all 24 hours.

HOW TO CALCULATE THE PREDICTED VALUES:

We need to calculate the simple moving average of the data. Then we need to calculate the accuracy of all the models 3PSMA gives us the Accuracy up to 98%.

For calculating the predicted values, we can select the last three values of 3PSMA of the 16th hour. We can calculate the 3PSMA of last three values
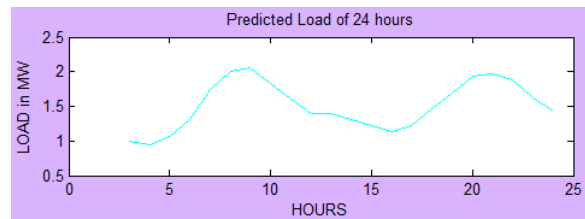

Fig (8) Graphical representation of Predicted load.

This is the graph, which shows the load of the predicted values of all the 24 hours of 12-feb-2008, after using the SMA technique.
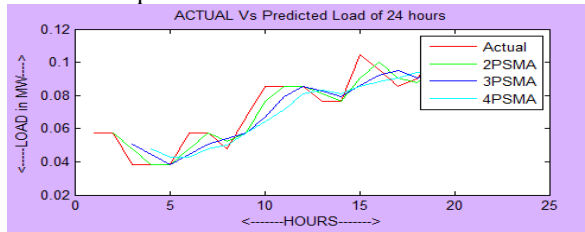

Fig (9) Comparison graph b/w 2PSMA, 3PSMA, and 4PSMA

V. CONCLUSION

Smoothing techniques used in this approach is very effective, efficient and helpful in any organization (Companies) for load forecasting and also the method Simple Moving Average used in this approach has better accuracy.The results also concluded that when we increase the no of periods the efficiency of load forecasting gets decreases as shown in result for comparison between 2PSMA, 3PSMA and 4PSMA technique.

REFERENCES:

[1]"Jianguang Deng and Panida Jirutitijaroen" "Short Term Load Forecasting   Using Time series Analysis: A Case Study for Singapore", IEEE, International Journal of Electrical and Electronics Engineering (IJEEE), Vol-1, Issue-3, 2012, Pages: 38 – 41.

[2]M. T. Hagan and S. M. Behr, "The time series approach to short term load forecasting," IEEE Trans. Power Syst., Vol. PWRS-2, No. 3, Aug.1987.

[3]A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," IEEE Trans. Power Syst., Vol. 5, No. 4, Nov. 1990.

[4] J. Shafer, R. Agrawal, and M. Mehta. SPRINT: A scalable parallel classifier for data mining. VLDB'96.

[5] L. Kaufman and P. J. Rousseeuw. Finding Groups in Data: an Introduction to Cluster Analysis. John Wiley & Sons, 2003

[6] Han and M. Kamber, "Data Mining Concepts and Techniques" Second edition, Morgan Kaufmann, San Francisco, CA 94111, 2005

[7]A. J. Conejo, M. A. Plazas, R. Espinola and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet tr7nsform and ARIMA models," IEEE Trans. Power Syst., Vol. 20, No. 2, May 2005.

[8]H. Witten and E. Frank, Data Mining Practical Machine Learning Tools and Techniques Second edition Morgan Kaufmann, San Francisco, CA, 2006.

[9] J.W. Taylor and P. E. McSarry,"Short Term Load Forecasting Method: An Evaluation Based on European Data", IEEE Transaction on Power System, Vol- 22, 2213-1-2219, 2008

[10] "Milos B. Stojanovic, Milos M. Bozic","Short Term Load Forecasting Using Recursive Time Series Prediction Strategy with Support Vector Machines", Vol-23, No-3, 2010, December, Pages: 287– 298.

[11]"Irina Asenova and Dimiter Georgiev" "Short Term Load Forecasting In Electric Energy System in BULGARIA", Advances in Electrical and Electronic Engineering (AEEE), Vol-8, No-3, 2010, December, Pages: 25 – 31.

[12]"Alexandra Kotillova" "Very Short Term Load Forecasting Using Exponential Smoothing and ARIMA, Journal of Information, Control and Management System,Vol-9, No-2, 2011, December, Pages: 85– 92.

[13] "Yuhang Yang, Yao Meng, Yingju Xia, Yingliang Lu, and Hao Yu", An Approach of Short Term Load Forecsting, IMECS Vol- 1 March 2011,