

Short Term Load Forecasting Using Moving Averages

Mr. Sandeep Tiwari¹, Mr. Kamlesh Lahre²

¹M. Tech Student Dept. of CSE CVRU Bilaspur,

²Asst. Professor, Dept. of CSE CVRU Bilaspur

Abstract: *Electricity is an essential part of our life and without this nothing is possible an accurate and efficient Short term load forecasting plays a vital role for economic operational planning of both regulated power systems and electricity markets. Short term forecasting including energy purchasing and generation, infrastructure development and load switching . A large variety of mathematical methods have been developed for load forecasting. This paper proposed the multiplicative decomposition model and the smoothing techniques model. Forecasting errors of both models are computed and compared. Results show that both time series models can accurately predict the short-term demand.*

Keywords— *Short Term load forecasting, moving averages, exponential smoothing, mean Absolute Percentage Error (MAPE)*

I. INTRODUCTION

The short-term load forecast is of importance in the daily operations of a power utility. It is required for unit commitment, energy transfer scheduling and load dispatch. With the emergence of load management strategies, the short-term load forecast has played a broader role in utility operations. The development of an accurate, fast and robust short-term load forecasting methodology is of importance to both the electric utility and its customers. Several different methods, techniques and algorithm 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. 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. 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 above mentioned 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 where by 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. TIMESERIES MODELS

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, two time series models, the Multiplicative Decomposition Model and the smoothing techniques. Model are employed and presented in the followings.

A. Multiplicative decomposition model

Multiplicative Decomposition Model assumes that a time series can be described as (1). $x(t) = T(t) * S(t) * C(t) * R(t), t = \dots - 1, 0, 1, 2, \dots$ (1) Where $x(t)$ is the time series, $T(t)$ is trend component, $S(t)$ is seasonal component, $C(t)$ is cyclic component, and $R(t)$ represent the random component. The cyclic component is usually in the duration of one year to a few years and is not applicable to short-term load forecasting. Thus, we propose to simplify the above combination to only three terms as shown in (2). $x(t) = T(t) * S(t) * R(t), t = \dots - 1, 0, 1, 2, \dots$ (2) In order to apply this model, it requires that the trend of the time series be found and extended into the future. Intuitively, the trend component of a series made up of these three components can be found if the other two could be taken off the series. A typical load series contains daily seasonal indexes, which when used to divide a typical week's data, would

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

remove the seasonal component from the series. To find the indexes, data of the most recent week without any holiday are divided by the average of a few recent weeks and using the average of these weeks tends to minimize the random effect. With these two components removed or minimized, the series now contains mainly the trend component. The equation of this trend line is extrapolated to estimate the trend in the future. Seasonal effects can then be incorporated into these future trend forecasts to account for the intra-day variation and obtain reasonably comprehensive forecasts. In the following, the detail implementation is described.

- 1) *Intraday Seasonal Index*: This section analyses intra-day seasonal index which measures how the series evolve throughout a day in order to predict future load performance on a half-hourly basis. Due to distinct difference of the data between weekend and weekday in terms of magnitude and fluctuation, weekends will be taken off the series and analysed separately from weekdays. First, the average of all measurements from the immediately previous four weeks (weekdays) is calculated. The rationale for choosing a four-week period is to reduce the effect of the possible scenario that a holiday exists within the past week and distorts the weekly seasonal behaviour. In addition, data in the past weeks may not reflect the current situation well therefore only the latest four weeks are chosen. Next, each of the latest week's data is divided by the four-week average to give the weekly seasonal indexes. If a holiday is present in the latest week, the previous week is used instead to represent a typical weekly performance. The weekly seasonal indexes represent how each half-hourly measurement varies across the week and typically range from 80% to 120% of the four-week average.
- 2) *Depersonalized Weekly Trend*: A depersonalized weekly performance is obtained by dividing the latest week's data (or the previous week's, whichever is used) by the weekly seasonal indexes correspondingly. This is the depersonalized weekly trend which is expected to appear as a straight trend line with small variations due to random effects that we cannot remove completely.
- 3) *Forecasting*: Equation to the weekly trend can be easily obtained by mathematical derivation or software tools, such as MATLAB or WEKA. This will then be used to project into the future for future trend in a process called trend forecasting. The actual forecast is done by multiplying this forecasted trend with the weekly seasonal indexes. This process can be extended from a single week to two weeks or more. However, it should be noted that the past data is related more closely to the immediate next week and that the forecast tends to have more significant errors for data points further into the future. This is due to many unforeseen factors which are difficult to simulate accurately. When a holiday is anticipated in the forecasted duration, great care must be taken to consider that particular day's forecast to be similar to its form in the last occurrence. No universal seasonal indexes are possible for all holidays because during different holidays, the system load dynamics behave differently. For example, Chinese New Year holidays usually see a major drop in demand magnitude for two days while other single public holidays have lower demands only on that particular day. These holidays seem to resemble to a certain extent a typical Sunday's load curve but even more closely, the curve of the correspondingly holidays one year ago. In this analysis, previous year's holiday data is mapped into the forecasted year, with the magnitude appropriately adjusted for development in the year. After the forecast is done for weekdays, we will forecast for weekends data in a similar manner. Combining these two forecasts will give us the overall forecast of the future load.

III. MULTIPLICATIVE DECOMPOSITION MODEL

In this case, the multiplicative decomposition technique is used to predict the loads for September 2008, data from the four weeks preceding September, namely the last four weeks in August 2008 (4th August to 31st August 2008) are used. Each day contains 24 data since they are hourly measurements. Similarly, 24 forecasts are to be made for each day in September. Also, weekdays and weekends are separately analyzed due to the difference in load pattern of these days, as could be observed. First, the overall average load of these 4 weeks (of weekdays) is calculated to be 3.00MW. Point average of each of the week's 120 data points (24 data per day in 5 days) is found by finding the average of corresponding data points in the four weeks. Dividing these 120 point averages by the overall average gives the weekly seasonal indexes. The last week of the 4 weeks is used to determine the current load demand trend since it does not contain any holiday. Dividing this week's 120 data by the weekly seasonal indexes leads to the deseasonalized weekly loads which, when plotted, closely resembles a trend line, with small fluctuations, as shown in the graph.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

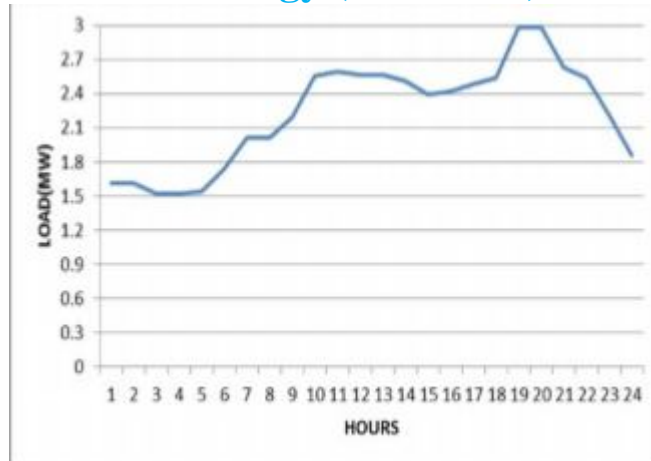


Fig:1 Depersonalized Load in hours

Multiplying them with the weekly seasonal indexes will give the final forecasts. Weekends are similarly analyzed and forecasted. Afterwards forecasts are joined together as overall forecasts for the entire month.

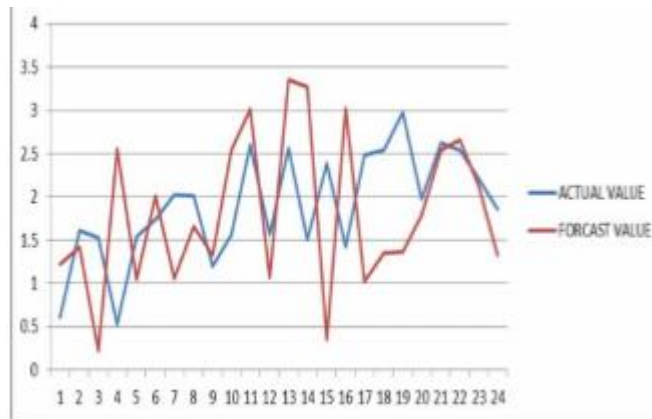


Fig:2 Forecast and Actual Load Comparison

A. Smoothing Techniques

- 1) *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.
- 2) *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, Exponential Moving average and weighted moving average.
- 3) *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. This causes the average to move along the time scale. Daily Closing Prices: 11,12,13,14,15,16,17 First day of 5-day SMA: $(11 + 12 + 13 + 14 + 15) / 5 = 13$ Second day of 5-day SMA: $(12 + 13 + 14 + 15 + 16) / 5 = 14$ Third day of 5-day SMA: $(13 + 14 + 15 + 16 + 17) / 5 = 15$ The first day of the moving average simply covers the last five days. The second day of the moving average

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

drops the first data point (11) and adds the new data point (16). The third day of the moving average continues by dropping the first data point (12) and adding the new data point (17). Notice that the moving average also rises from 13 to 15 over a three day calculation period. Also notice that each moving average value is just below the last price. For example, the moving average for day one equals 13 and the last price is 15. Prices the prior four days were lower and this causes the moving average to lag

- 4) *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 below is for a 10-day EMA.

SMA: 10 period sum / 10

Multiplier: $(2 / (\text{Time periods} + 1)) = (2 / (10 + 1)) = 0.1818$ (18.18%)

EMA: {Close-SMA (previous day)} x multiplier + EMA (previous day).

Or we can also write this formula like (SMA +multiplier (Last value-SMA)

- 5) *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: $(1 * P_n - (n-1)) / (n + (n-1) + \dots + (n-(n-1))) W ((n * P_n) + ((n - 1) * P_{n-1}) + ((n - 2) * P_{n-2}) + \dots + ((n - (n - 1)))$ here n is the number of terms present, p is the values given there. The WMA of prices of 1.2900, 1.2900, 1.2903, and 1.2904 would give a moving average of 1.2903 using the calculation $((4 * 1.2904) + (3 * 1.2903) + (2 * 1.2900) + (1 * 1.2900)) / (4 + 3 + 2 + 1) = 1.2903$

- 6) *Data Set*: I have the data of MAHARASHTRA STATE ELECTRICITY DISTRIBUTION CO. LTD. I 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 I 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. I 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

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

Figure 3 Raw data

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

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

7) *Simple Moving Average*: 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.

B. Introduction to Simple Moving Average

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

$$SMA = \frac{P_M + P_{M-1} + \dots + P_{M-9}}{10}$$

1) Steps Use to Calculate The Sma

a) *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

Figure 4 Sangam nagar’s data on Excel sheet

b) *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 Sundays of 2 years that is from April 2008 to April 2010

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

DAY	DATE	1HR	2HR	3HR	4HR	5HR	6HR	7HR	
SUNDAY	6-Apr-08	1.4072	1.3311	1.2814	1.2777	1.3817	1.4444	1.5223	
SUNDAY	13-Apr-08	1.4349	1.3097	1.3058	1.2754	1.3842	1.4586	1.4871	
SUNDAY	20-Apr-08	1.4306	1.2741	1.258	1.2802	1.2227	1.3983	1.4169	
SUNDAY	27-Apr-08	0.7771	0.7343	0.7148	0.6845	0.6741	0.6716	0.6008	
SUNDAY	4-May-08	0.8066	0.7799	0.7462	0.7552	0.7256	0.7128	0.6941	
SUNDAY	11-May-08	0.9007	0.8947	0.8817	0.8557	0.893	0.8704	0.8295	
SUNDAY	18-May-08	0.9336	0.9037	0.87	0.8804	0.8716	0.8558	0.8566	
SUNDAY	25-May-08	0.8433	0.8125	0.779	0.7607	0.7871	0.7586	0.7128	
SUNDAY	1-Jun-08	0.907	0.8967	0.8516	0.8282	0.8148	0.7806	0.7222	
SUNDAY	8-Jun-08	1.0516	0.9869	0.9448	0.9549	0.9628	0.9637	0.9781	
SUNDAY	15-Jun-08	0.971	0.9054	0.8889	0.907	0.9078	0.9611	1.0336	
SUNDAY	22-Jun-08	1.08	1.083	0.9924	0.9942	1.0139	1.0683	1.1348	
SUNDAY	6-Jul-08	0.3871	0.37	0.3748	0.3564	0.3515	0.3609	0.3443	
SUNDAY	13-Jul-08	1.1523	1.1183	1.1235	1.08	1.081	1.0458	1.1169	
SUNDAY	20-Jul-08	1.349	1.349	0.8155	0.8889	0.9024	0.9674	0.986	
SUNDAY	27-Jul-08	0.6425	0.6063	0.6086	0.627	0.6644	0.668	0.7049	
SUNDAY	3-Aug-08	0.6088	0.5926	0.5909	0.5545	0.6117	0.6475	0.7134	
SUNDAY	10-Aug-08	0.573	0.5633	0.554	0.5597	0.553	0.5855	0.5725	
SUNDAY	17-Aug-08	0.6366	0.6105	0.6008	0.608	0.6168	0.6608	0.6511	
SUNDAY	24-Aug-08	1.3015	1.3211	1.1656	1.1951	1.2037	1.385	1.5328	
SUNDAY	7-Sep-08	1.0323	0.9734	0.9698	0.9268	0.9806	1.0536	1.0936	
SUNDAY	14-Sep-08	0.9502	0.9162	0.8901	0.8906	0.9379	1.1324	1.1602	
SUNDAY	21-Sep-08	0.854	0.8223	0.8185	0.8254	0.8955	0.9991	1.0536	
SUNDAY	28-Sep-08	0.7505	0.7174	0.6765	0.6816	0.7323	0.8171	0.8528	
SUNDAY	5-Oct-08	1.1613	1.1592	1.0856	1.108	1.1204	1.2709	1.3704	
SUNDAY	12-Oct-08	0.3928	0.3765	0.3721	0.3731	0.3535	0.3741	0.3697	
SUNDAY	19-Oct-08	0.3838	0.3656	0.3473	0.3619	0.3619	0.4162	0.3619	
SUNDAY	26-Oct-08	0.3361	0.3361	0.3174	0.3174	0.3361	0.3734	0.3734	
SUNDAY	2-Nov-08	0.3584	0.3584	0.3395	0.3395	0.3206	0.3395	0.3584	
SUNDAY	9-Nov-08	0.3597	0.3389	0.3239	0.3206	0.338	0.3617	0.0191	
SUNDAY	16-Nov-08	0.3395	0.3206	0.3206	0.3206	0.3395	0.3584	0.3584	

Figure 5 Sangam Nagar's data after applying the SMA

- c) Step 3: SMA is Average value so we are using the simple formula to calculate the SMA the formula is $SMA = \frac{\text{sum of number of values}}{N}$, 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.

DAY	DATE	1HR	3PSMA	4PSMA	5PSMA	6PSMA	7PSMA
SUNDAY	6-Apr-08	1.4072					
SUNDAY	13-Apr-08	1.4349					
SUNDAY	20-Apr-08	1.4306	1.424233				
SUNDAY	27-Apr-08	0.7771	1.2142	1.26245			
SUNDAY	4-May-08	0.8066	1.004767	1.1123	1.17128		
SUNDAY	11-May-08	0.9007	0.828133	0.97875	1.06998	1.126183	
SUNDAY	18-May-08	0.9336	0.8803	0.8545	0.96972	1.04725	1.098671
SUNDAY	25-May-08	0.8433	0.892533	0.87105	0.85226	0.94865	1.018114
SUNDAY	1-Jun-08	0.907	0.894633	0.89615	0.87824	0.861383	0.9427
SUNDAY	8-Jun-08	1.0516	0.933967	0.933875	0.92724	0.907133	0.888557
SUNDAY	15-Jun-08	0.971	0.976533	0.943225	0.9413	0.934533	0.916257
SUNDAY	22-Jun-08	1.08	1.0342	1.0024	0.97058	0.964417	0.955314
SUNDAY	6-Jul-08	0.3871	0.8127	0.872425	0.87934	0.873333	0.881943
SUNDAY	13-Jul-08	1.1523	0.873133	0.8976	0.9284	0.924833	0.913186
SUNDAY	20-Jul-08	1.349	0.9628	0.9921	0.98788	0.9985	0.985429
SUNDAY	27-Jul-08	0.6425	1.047933	0.882725	0.92218	0.930317	0.947643
SUNDAY	3-Aug-08	0.6088	0.866767	0.93815	0.82794	0.86995	0.884386
SUNDAY	10-Aug-08	0.573	0.6081	0.793325	0.86512	0.78545	0.827529
SUNDAY	17-Aug-08	0.6366	0.606133	0.615225	0.76198	0.827033	0.764186
SUNDAY	24-Aug-08	1.3015	0.837033	0.779975	0.75248	0.8519	0.894814
SUNDAY	7-Sep-08	1.0323	0.990133	0.88585	0.83044	0.799117	0.877671
SUNDAY	14-Sep-08	0.9502	1.094667	0.98015	0.89872	0.8504	0.8207
SUNDAY	21-Sep-08	0.854	0.9455	1.0345	0.95492	0.891267	0.850914
SUNDAY	28-Sep-08	0.7505	0.851567	0.89675	0.9777	0.92085	0.871157
SUNDAY	5-Oct-08	1.1613	0.921933	0.929	0.94966	1.0083	0.9552
SUNDAY	12-Oct-08	0.3928	0.7682	0.78965	0.82176	0.85685	0.920371

Figure 6 Sangam Nagar's data after applying the SMA

Here we are showing SMA techniques used in the 1st hour data of all the Sundays of the Sangam Nagar. I applied period function to

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

the data like 3PSMA, 4PSMA, 5PSMA, 6PSMA, 7PSMA. Results are constant so up to 7PSMA. Therefore no need to use more periods.

d) *Step 4:* In our data set we have calculated the 3PSMA, 4PSMA, 5PSMA, 6PSMA, 7PSMA to predict the load of upcoming Sunday. We are taking the data from our data set to solve these calculations. To calculate the 3PSMA we need three values so we are taking three values from our data set these values are 1.6041, 1.5199, and 1.475

$$3PSMA = (N1+N2+N3)/N \quad 3PSMA = (1.6041+1.5199+1.475)/3$$

3PSMA=1.533 For 4PSMA we take four values from our data set like 1.6041, 1.5199, 1.475 and 0.769 and apply the formula

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

4PSMA = (1.6041+1.5199+1.475+0.769)/4 4PSMA=1.342 Like this I apply this formula up to 7PSMA in our data set.

e) *Result*

	Actual Value	Predicted Values				
SUM	77.7605	75.10573	73.82948	72.56248	71.45563	70.372
MEAN		0.790587	0.78542	0.780242	0.776692	0.773319
MAPE		0.035196	0.052116	0.068914	0.083588	0.097955

Figure 7 Calculated values of SMA.

This is the result of the 1st hour data, after applying SMA. We got one actual value and five predicted values

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, here the Actual value is 77.7605

PREDICATED VALUE = It is the calculated value here the Predicated value is 75.10573, 73.829, 72.56248, 71.45563, 70.372.

ACCURACY= It is a measurement of a correctness; the accuracy of the model is near about 97%. ACCURACY = MAPE*100

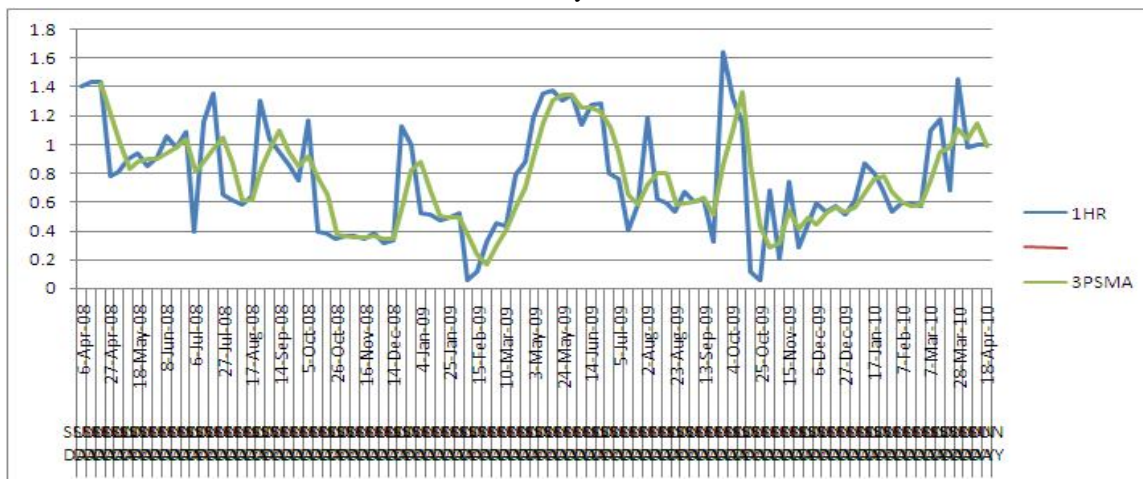


Figure 8 Graph between actual load and predicted load.

This is the graph which shows the comparison between 'Actual Value' and '3PSMA' of the 1st hour. Where, Blue line

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

denotes the Actual data and Light Green line denotes the 3PSMA.

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

C. 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 97% so we are using the 3PSMA to calculate the load for the upcoming day.

For calculating the predicted values, we can select the last three values of 3PSMA of the 1st hour. We can calculate the 3PSMA of last three values

$$\text{Ex } 0.9709+0.9997+0.9997 / 3 = 0.9901.$$

It is the predicted value of 1st hour. For predicting 2nd, 3rd up to 24th hour value, we can use 3PSMA technique on the last three values.

HR	Forecast Value
1	0.9901
2	0.9712
3	0.949366667
4	0.939966667
5	0.9336
6	0.930266667
7	0.873766667
8	0.920833333
9	0.899
10	1.005766667
11	1.030833333
12	1.0489
13	1.039866667
14	1.042466667
15	1.0525
16	1.062533333
17	1.056266667
18	1.087666667
19	1.138
20	1.2197
21	1.1655
22	1.138
23	1.0965
24	1.0536

Figure 9 Forecasted load for next day.

These are all the predicted values of 24 hours data of Sunday 25th April.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

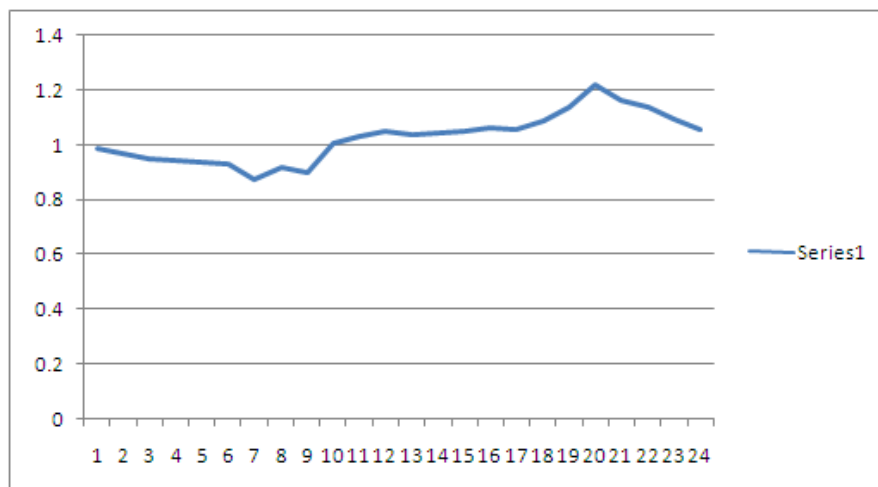


Figure 10 Graphical representation of forecasted load.

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

REFERENCES

- [1] I.H. Witten and E. Frank, Data Mining Practical Machine Learning Tools and Techniques Second edition Morgan Kaufmann, San Francisco, CA, 2005.
- [2] J. Han and M. Kamber, "Data Mining Concepts and Techniques" Second edition, Morgan Kaufmann, San Francisco, CA 94111, 2005
- [3] J. Shafer, R. Agrawal, and M. Mehta. SPRINT : A scalable parallel classifier for data mining. VLDB'96
- [4] K. Liu, S. Subbarayan, R. R. Shoultz, M. T. Manry, C. Kwan, F. L. Lewis and J. Naccarino, "Comparison of very short-term load forecasting techniques," IEEE Trans. PowerSyst., Vol. 11, No. 2, May 2002. <http://archive.ics.uci.edu/ml/datasets.html>
- [5] L. Kaufman and P. J. Rousseeuw. Finding Groups in Data: an Introduction to Cluster Analysis. John Wiley & Sons, 1990
- [6] objhyndman.com/researchtips/forecasting-and-time-series-books
- [7] M. S. Bartlett. "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series." Journal of the Royal Statistical Society, B8, 27 (1946).
- [8] <http://archive.ics.uci.edu/ml/datasets/Auto+MPG>
- [9] Weka 3.6.3 Data set
- [10] <http://www.mahadiscom.in/>
- [11] <http://www.mahadiscom.in/>
- [12] http://ebookey.org/Spectral-Analysis-and-Time-Series-2-Vol-set_215763.html