IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 4, Ver. IV (Jul – Aug. 2014), PP 56-64
www.iosrjournals.org

Comparative Study of Statistical Predictive Analytic Techniques

Aditya Vakaskar[1], Uday Joshi[2]
[1](Computer Engineering, K J. Somaiya College of Engineering/ Universityof Mumbai, India)
[2](Computer Engineering, K. J. Somaiya College of Engineering/ University of Mumbai, India)

Abstract: *"Prediction is very difficult, especially if it's about the future."- Niels Bohr*
Since the early days of mankind, man has always been fascinated by the idea of knowing future. Data is being captured at a rate never before seen in history. The retailer's goal is to translate that data into bottom line profits & Predictive analytics makes that possible. The data one captures about customers, or even consumers who interact with retail operation and don't make a purchase, is more revealing than one can think of. Customer data can provide insights on everything from large and systemic patterns of global markets, workflows, national infrastructures, and natural systems to the location, temperature, security, and condition of every item in supply chain. Predictive analytics offers access to reliable, timely information; understand customers, spot trends that drive better decisions to stay ahead in a competitive marketplace. Managers have many different decisions to make monthly, weekly, daily, sometimes even hourly. In 2012, worldwide Business Analytics software market grew 8.7% year over year with revenues reaching $34.9 billion [1] [2] and is also expecting accelerated growth which will be fueled by the quest to harness the power of big data. This paper gives Comparative Study of some of the Time Series Analytic Techniques which are the foundation blocks of Predictive Analytics.
Keywords: *predictive analytics, time series forecasting, demand planning, exponential smoothing, ARIMA*

I. Business Intelligence

"The answer to my problem is hidden in my data… but I cannot dig it up!"[3] This popular statement has been around for years as business managers gathered and stored massive amounts of data in the belief that they contain some valuable insight. But business managers eventually discovered that raw data are rarely of any benefit, and that their real value depends on an organization's ability to analyze them. Hence, the need emerged for software systems capable of retrieving, summarizing, and interpreting data for end-users. This need fueled the emergence of hundreds of business intelligence companies that specialized in providing software systems and services for extracting knowledge from raw data. These software systems would analyze a company's operational data and provide knowledge in the form of tables, graphs, pies, charts, and other statistics. For example, a business intelligence report may state that 90% of customers are between the ages of 20 and 30, or that one of their products sell much better in a particular geography than other.

A business intelligence system was responsible for collecting and digesting data, and presenting knowledge in a friendly way (thus enhancing the end-user's ability to make good decisions).
The following diagram illustrates the processes that underpin a traditional business intelligence system:

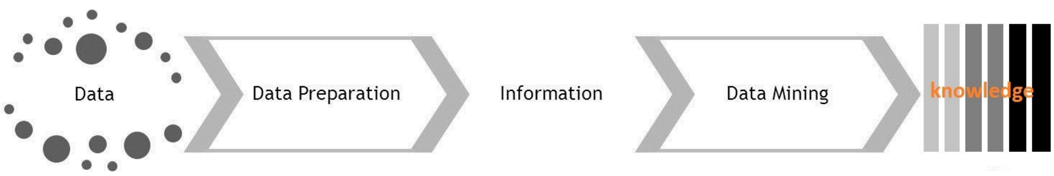


Figure 1: A traditional business intelligence system

1) Data is collected in the form of bits, numbers, symbols, and "objects."
2) Information is "organized data," which are preprocessed, cleaned, arranged into structures, and stripped of redundancy (i.e. Extract, Transform, Load).
3) Knowledge is "integrated information," which includes facts and relationships that have been perceived, discovered, or learned.

Because knowledge is such an essential component of any decision-making process, many businesses have viewed knowledge as the final objective. But it seems that knowledge is no longer enough. A business may "know" a lot about its customers – it may have hundreds of charts and graphs that organize its customers by age, preferences, geographical location, and sales history – but management may still be unsure of what decision to make. And here lies the difference between "decision support" and "decision making": all the knowledge in the world will not guarantee the right or best decision.[3]

Moreover, recent research in psychology indicates that widely held beliefs can actually hamper the decision-making process. For example, common beliefs like “the more knowledge we have, the better our decisions will be,” is not supported by empirical evidence. Having more knowledge merely increases confidence, but it does not improve the accuracy of decisions. Similarly, people supplied with “good” and “bad” knowledge often have trouble distinguishing between the two, proving that irrelevant knowledge decreases our decision-making effectiveness.

Today, most business managers realize that a gap exists between having the right knowledge and making the right decision. Because this gap affects management’s ability to answer fundamental business questions such as “What should be done to increase profits? Reduce costs? Or increase market share?” , the future of business intelligence lies in systems that can provide answers and recommendations, rather than mounds of knowledge in the form of reports. The future of business intelligence lies in systems that can make decisions. As a result, there is a new trend emerging in the marketplace called Predictive analytics & Adaptive Business Intelligence. [3]

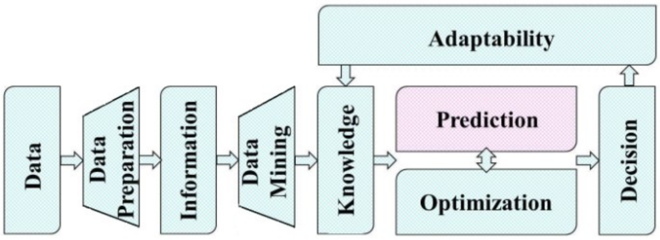


Figure 2: Adaptive business intelligence system

Forecasting is the process of making statements about events whose actual outcomes (typically) have not yet been observed. A commonplace example might be estimation of some variable of interest at some specified future date. Prediction is a similar, but more general term. Both might refer to formal statistical methods employing time series, cross-sectional or longitudinal data, or alternatively to less formal judgmental methods[4]. Risk and uncertainty are central to forecasting and prediction; it is generally considered good practice to indicate the degree of uncertainty attaching to forecasts. In any case, the data must be up to date in order for the forecast to be as accurate as possible. [5]

Demand planning is one such prediction problem that industry is facing since early days of industrial revolution. Demand planning and forecasting is a business process that involves predicting future demand for products and services and aligning production and distribution capabilities accordingly. Alternatively, it can be described as, using forecasts and experience to estimate demand for various items at various points in a supply chain. It involves a number of different business functions and requires the sharing of timely data, accurate processing of this data and agreement on joint business plans along the supply chain.

Greater competition, more frequent new product launches and shorter product life cycles have made forecasting increasingly complex. Organizations themselves have also become more complex in the past decade, and many now operate in a greater number of locations, business units and markets. Unprecedented levels of economic uncertainty, which have affected buying patterns and historical data, have added to this complexity.[6]

To address this industry problem statistical predictive analytic techniques come into picture.
Statistical predictive analytic techniques are used to forecast future data as a function of past data; they are appropriate when past data are available. These methods are usually applied to short- or intermediate-range decisions. Examples of quantitative forecasting methods which fall under statistical predictive analytics belong to time series analysis techniques where last period demand, simple and weighted N-Period moving averages, simple exponential smoothing, and multiplicative seasonal indexes.

A single technique might not be accurate for forecasting in all of the scenarios so a comparative study should done in order to find the best suitable technique for prediction. Here we will start form very basic techniques of time series forecasting and proceed towards complex techniques as we move further and compare them for accuracy. We will review techniques that are useful for analyzing time series data, that is, sequences of measurements that follow non-random orders. Unlike the analyses of random samples of observations that are discussed in the context of most other statistics, the analysis of time series is based on the assumption that successive values in the data file represent consecutive measurements taken at equally spaced time intervals.

II. Statistical predictive analytic methods

2.1 Naïve forecast

Naïve forecasts are the most cost-effective objective forecasting model, and provide a benchmark against which more sophisticated models can be compared. For time series data that are stationary in terms of first differences, the naïve forecast equals the previous period's actual value. [4]

In other words it is an estimating technique in which the last period's actuals are used as this period's forecast, without adjusting them or attempting to establish causal factors. It is generally used for comparison with the forecasts generated by the better (sophisticated) techniques. [7]

It is the simplest possible forecast, which emphasizes on "Tomorrow will be like today" and ignores any historical data previous to today. [8]

2.2 Smoothing methods

Inherent in the collection of data taken over time is some form of random variation. There exist methods for reducing of canceling the effect due to random variation. An often-used technique in industry is "smoothing". This technique, when properly applied, reveals more clearly the underlying trend, seasonal and cyclic components.
There are two distinct groups of smoothing methods
1) Averaging Methods
2) Exponential Smoothing Methods. [9]

2.2.1 Averaging methods

Averaging methods are those simple methods which do not take into account trends, seasonality or cycles into account while forecasting.

2.2.1.1 Moving Average

A moving average, also called rolling average, moving mean, rolling mean, sliding temporal average, or running average, is a type of finite impulse response filter used to analyze a set of data points by creating a series of averages of different subsets of the full data set.

Given a series of numbers and a fixed subset size, the first element of the moving average is obtained by taking the average of the initial fixed subset of the number series. Then the subset is modified by "shifting forward"; that is, excluding the first number of the series and including the next number following the original subset in the series. This creates a new subset of numbers, which is averaged. This process is repeated over the entire data series.[10]
Simple moving average (SMA) is the unweighted mean of the previous n datum points.

2.2.1.2 Weighted Moving Average (WMA)[10]

A weighted average is any average that has multiplying factors to give different weights to data at different positions in the sample window. Mathematically, the moving average is the convolution of the datum points with a fixed weighting function.

In technical analysis of financial data, a weighted moving average (WMA) has the specific meaning of weights that decrease in arithmetical progression.[11] In an n-day WMA the latest day has weight n, the second latest n − 1, etc., down to one.

$$\text{WMA}_M = \frac{np_M + (n-1)p_{M-1} + \cdots + 2p_{(M-n+2)} + p_{(M-n+1)}}{n + (n-1) + \cdots + 2 + 1} \quad (1)$$

The denominator is a triangle number equal to $\frac{n(n+1)}{2}$ (2)

In the more general case the denominator will always be the sum of the individual weights.

3.2.2 Exponential Smoothing

Exponential smoothing is a category of methods which have been widely used. The origins and formulation of these methods is based on original work of Brown (1959, [11]) and Holt (1957, [12]) who devised forecasting models for inventory control systems. Exponential smoothing is an intuitive forecasting method that weights the observed time series unequally.

Recent observations are weighted more heavily than remote observations. The unequal weighting is accomplished by using one or more smoothing parameters, which determine how much weight is given to each observation [13].

In other words, Exponential smoothing is a procedure for continually revising a forecast in the light of more recent experience. Exponential Smoothing assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more weight in forecasting than the older observations [14].

2.2.2.1 Simple Exponential Smoothing (Single Exponential Smoothing, SES)

The simplest technique of this type, simple exponential smoothing (SES), is appropriate for a series that moves randomly above and below a constant mean (stationary series). It has no trend and no seasonal patterns [15].

Generally, exponential smoothing is regarded as an inexpensive technique that gives good forecast in a wide variety of applications. In addition, data storage and computing requirements are minimal, which makes exponential smoothing suitable for real-time application [16].

The simple exponential smoothing (SES) model is usually based on the premise that the level of time series should fluctuate about a constant level or change slowly over the time [13].

Mathematical Formulation:

The SES model is given by the model equation

$$y(t)=\beta(t)+\varepsilon(t) \tag{3}$$

where $\beta(t)$ takes a constant at the time t and may change slowly over the time; $\varepsilon(t)$ is a random variable and is used to describe the effect of stochastic fluctuation.

Let an observed time series be y1,y2,…yn. In any case, in this simple model, to predict yt is merely to predict (estimate)β. To estimate, it makes sense to use all the past observations, but due to declining correlation as you go back into the past, to down-weight older observations.

Formally, the simple exponential smoothing equation takes the form of

$$Ft + 1 = \alpha yt + (1-\alpha)Ft, \tag{4}$$

where yt is the actual, known series value at the time t; Ft is the forecast value of the variable Y at the time t; Ft+1 is the forecast value at the time t + 1; α is the smoothing constant [17].

The forecast Ft+1 is based on weighting the most recent observation yt with a weight α and weighting the most recent forecast Ft with a weight of 1 − α.

To get started the algorithm, we need an initial forecast, an actual value and a smoothing constant.

Since F1 is not known, we can:

• Set the first estimate equal to the first observation. Further we will use F1 = y1.

• Use the average value of the first few observations of the data series for the initial smoothed value.

Smoothing constant α is a selected number between zero and one, $0 < \alpha < 1$.

Rewriting the model (4) we can see one of the neat things about the SES model

$$Ft+1-Ft = \alpha\,(yt -Ft) \tag{5}$$

change in forecasting value is proportional to the forecast error. That is

$$Ft+1 = Ft +\alpha\,\varepsilon t \tag{6}$$

where residual

$$\varepsilon t=yt- Ft \tag{7}$$

is the forecast error at the time t. So, the exponential smoothing forecast is the old forecast plus an adjustment for the error that occurred in the last forecast [18] [19].

By iterating formula (4) we get:

$F_1 = y_1$;

$F_2 = \alpha.\ y_1 + (1 - \alpha).\ F_1 = y_1$;

$F_3 = \alpha.\ y_2 + (1 - \alpha).\ F_2 = \alpha\ y_2 + (1 - \alpha)\ y_1 = \alpha.\ y_2 + \alpha\ (1 - \alpha).\ y_1 + (1 - \alpha)^2.\ y_1$;

$F_4 = \alpha.\ y_3 + (1 - \alpha).\ F_3 = \alpha\ y_3 + (1 - \alpha)\ (\alpha.\ y_2 + \alpha\ (1 - \alpha).\ y_1 + (1 - \alpha)^2\ .y_1)$

$= \alpha\ (y_3 + (1 - \alpha)\ y_2 + (1 - \alpha)^2\ y_1) + (1 - \alpha)^3.\ y_1$; …….

The forecast equation in general form is

$$F_{t+1} = \alpha \sum_{k=0}^{t-1}(1-\alpha)^k y_{t-k} + (1-\alpha)^t y_1, \qquad t \in \mathbb{N}, \tag{8}$$

where Ft+1 is the forecast value of the variable Y at the time t +1 from knowledge of the actual series values yt , yt-1, yt-2 and so on back in time to the first known value of the time series, y1[17][20]. Therefore, Ft+1 is the weighted moving average of all past observations.

The series of weights used in producing the forecast Ft+1 is $\alpha, \alpha\ (1 - \alpha), \alpha\ (1 - \alpha)^2$ (9)

It is obviously from (9) that the weights are exponential; hence the name exponentially weighted moving average [18]. The exponential decline of the weights toward zero is evident.

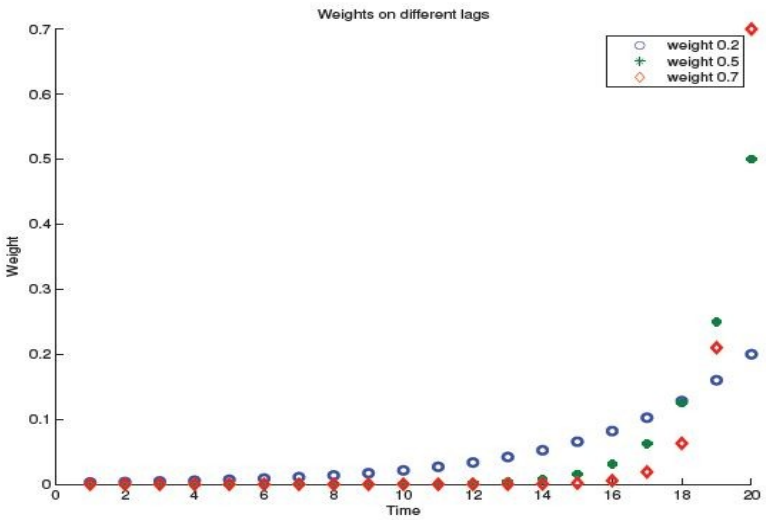


Figure 3: Exponentially declining weights [16]

As represented in the above figure, the decay is slower for small values of α. We can control the rate of decay by choosing α appropriately.

2.2.2.1.1 Choosing the Best Value for Smoothing Constant

The accuracy of forecasting of SES technique depends on smoothing constant. Choosing an appropriate value of exponential smoothing constant is very crucial to minimize the error in forecasting. Selecting a smoothing constant is basically a matter of judgment or trial and error, using forecast errors to guide the decision. The goal is to select a smoothing constant that balances the benefits of smoothing random variations with the benefits of responding to real changes if and when they occur. The smoothing constant serves as the weighting factor.

When α is close to 1, the new forecast will include a substantial adjustment for any error that occurred in the preceding forecast. When α is close to 0, the new forecast is very similar to the old forecast. The smoothing constant α is not an arbitrary choice but generally falls between 0.1 and 0.5. Low values of α are used when the underlying average tends to be stable; higher values are used when the underlying average is susceptible to change. In practice, the smoothing constant is often chosen by a grid search of the parameter space; that is, different solutions for α are tried starting, for example, with α = 0:1 to α =0:9, with increments of 0.1 [18] [19]. The value of with the smallest MAE, MSE, RMSE or MAPE is chosen for use in producing the future forecasts.

2.2.2.2 Double Exponential Smoothing (Holt's Method, DES) [14]

This method is used when the data shows a trend. Exponential smoothing with a trend works much like simple smoothing except that two components must be updated each period - level and trend. The level is a smoothed estimate of the value of the data at the end of each period. The trend is a smoothed estimate of average growth at the end of each period. The specific formula for simple exponential smoothing is:

$$S_t = \alpha * y_t + (1 - \alpha) * (S_{t-1} + b_{t-1}) \qquad 0 < \alpha < 1 \qquad (10)$$

$$b_t = \gamma * (S_t - S_{t-1}) + (1 - \gamma) * b_{t-1} \qquad 0 < \gamma < 1 \qquad (11)$$

Note that the current value of the series is used to calculate its smoothed value replacement in double exponential smoothing.

Initial Values

There are several methods to choose the initial values for St and bt.

S1 is in general set to y1. Three suggestions for b1

$$b_1 = y_2 - y_1 \qquad (12)$$

$$b_1 = [(y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3)]/3 \qquad (13)$$

$$b_1 = (y_n - y_1)/ (n - 1) \qquad (14)$$

2.2.2.3 Holt-Winters method (Triple Exponential Smoothing) [16] [21][22]

The Holt-Winters method, also referred to as triple exponential smoothing, is an extension of exponential smoothing designed for trended and seasonal time series. Holt-Winters smoothing is a widely used tool for forecasting business data that contain seasonality, changing trends and seasonal correlation [16]. This model is a widely used method in time series analysis. This popularity can be attributed to its simplicity, its computational efficiency, the ease of adjusting its responsiveness to changes in the process being forecast, and its reasonable accuracy.

It was first suggested by Holt's student, Peter Winters, in 1960.Suppose we have a sequence of observations (xt,) beginning at time t = 0 with a cycle of seasonal change of length L. The method calculates a trend line for the data as well as seasonal indices that weight the values in the trend line based on where that time point falls in the cycle of length L. There are 2 variants of this method, one which takes into account additive seasonality and the other which takes into account multiplicative seasonality.

st represents the smoothed value of the constant part for time t.

bt represents the sequence of best estimates of the linear trend that are superimposed on the seasonal changes.

ct is the sequence of seasonal correction factors. ct is the expected proportion of the predicted trend at any time t mod L in the cycle that the observations take on. To initialize the seasonal indices ct-L there must be at least one complete cycle in the data.

The output of the algorithm is again written as Ft+m, an estimate of the value of x at time t+m, m>0 based on the raw data up to time t. Triple exponential smoothing is given by the formulas

$$s_0 = x_0 \tag{15}$$

$$s_t = \alpha \frac{x_t}{c_{t-L}} + (1-\alpha)(s_{t-1} + b_{t-1}) \tag{16}$$

$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1} \tag{17}$$

$$c_t = \gamma \frac{x_t}{s_t} + (1-\gamma)c_{t-L} \tag{18}$$

$$F_{t+m} = (s_t + mb_t)c_{t-L+((m-1) \pmod L)}, \tag{19}$$

where α is the data smoothing factor, 0 < α < 1, β is the trend smoothing factor, 0 < β < 1, and γ is the seasonal change smoothing factor, 0 < γ < 1.

The general formula for the initial trend estimate b is:

$$b_0 = \frac{1}{L}\left(\frac{x_{L+1} - x_1}{L} + \frac{x_{L+2} - x_2}{L} + \ldots + \frac{x_{L+L} - x_L}{L}\right) \tag{20}$$

Setting the initial estimates for the seasonal indices ci for i = 1,2,...,L is a bit more involved. If N is the number of complete cycles present in your data, then:

$$c_i = \frac{1}{N}\sum_{j=1}^{N} \frac{x_{L(j-1)+i}}{A_j} \quad \forall i \ = 1,2,\ldots,L \tag{21}$$

Where

$$A_j = \frac{\sum_{i=1}^{L} x_{L(j-1)+i}}{L} \quad \forall j \ = 1,2,\ldots,N \tag{22}$$

Note that Aj is the average value of x in the jth cycle of your data [22].

2.3 ARIMA (Auto Regressive Integrated Moving Average)

Autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data to predict future points in the series (forecasting) by projecting the future values of a series based entirely on its own inertia[23][24].

ARIMA models are, in theory, the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging. In fact, the easiest way to think of ARIMA models is as fine-tuned versions of random-walk and random-trend models: the fine-tuning consists of adding lags of the differenced series and/or lags of the forecast errors to the prediction equation, as needed to remove any last traces of autocorrelation from the forecast errors [25].

Lags of the differenced series appearing in the forecasting equation are called "auto-regressive" terms, lags of the forecast errors are called "moving average" terms, and a time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series. Random-walk and random-trend models, autoregressive models, and exponential smoothing models (i.e., exponential weighted moving averages) are all special cases of ARIMA models.

A nonseasonal ARIMA model is classified as an "ARIMA(p,d,q)" model, where:

p is the number of autoregressive terms,

d is the number of nonseasonal differences, and

q is the number of lagged forecast errors in the prediction equation [25].

To identify the appropriate ARIMA model for a time series, the first step is to check for stationarity. "Stationarity" implies that the series remains at a fairly constant level over time. If a trend exists, the data is said to be nonstationary. The data should also show a constant variance in its fluctuations over time. If a graphical plot of the data indicates nonstationarity, then you should "difference" the series. By differencing we can transform a nonstationary series to a stationary one. This corresponds to "I" or Integrated part of the model. It is done by subtracting the observation in the current period from the previous one. If this transformation is done only once to a series, you say that the data has been "first differenced". This process essentially eliminates the trend if your series is growing at a fairly constant rate. If it is growing at an increasing rate, you can apply the same procedure and difference the data again [23].

The next stage is to determine the p and q in the ARIMA (p, I, q) model (the I refers to how many times the data needs to be differenced to produce a stationary series). To determine the appropriate lag structure in the AR part of the model, the PACF or Partial correlogram is used, where the number of non-zero points of the PACF determine where the AR lags need to be included. To determine the MA lag structure, the ACF or correlogram is used, again the non-zero points suggest where the lags should be included. Seasonal dummy variables may also need to be included if the data exhibits seasonal effects [26].

III. Measuring Forecast Error

After the model specified, its performance characteristics should be verified or validated by comparison of its forecast with historical data for the process it was designed to forecast. This is no consensus among researchers as to which measure is best for determining the most appropriate forecasting method. Accuracy is the criterion that determines the best forecasting method; thus, accuracy is the most important concern in evaluating the quality of a forecast. The goal of the forecast is to minimize error [27].

Some of the common indicators used to evaluate accuracy are MAE (Mean absolute error), MSE (Mean squared error), RMSE (Root mean squared error) or MAPE (Mean absolute percentage error):

Accuracy Indicators

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|e_t|$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}e_t^2$$

$$RMSE = \sqrt{MSE}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{y_t}.100\ \%$$

where yt is the actual value at the time t; e_t is residual at the time t; n is the total number of the time periods. MAE is a measure of overall accuracy that gives an indication of the degree of spread, where all errors are assigned equal weights. If a method fits the past time series data very good, MAE is near zero, whereas if a method fits the past time series data poorly, MAE is large. Thus, when two or more forecasting methods are compared, the one with the minimum MAE can be selected as most accurate [27].

MSE is also a measure of overall accuracy that gives an indication of the degree of spread, but here large errors are given additional weight. It is a generally accepted technique for evaluating exponential smoothing and other methods [28].

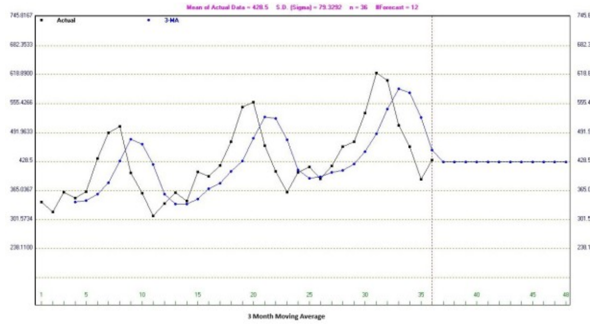
Often the square root of MSE, RMSE, is considered, since the seriousness of the forecast error is then denoted in the same dimensions as the actual and forecast values themselves. MAPE is a relative measure that corresponds to MAE. It is the most useful measure to compare the accuracy of forecasts between different items or products since it measures relative performance. It is one measure of accuracy commonly used in quantitative methods of forecasting [29]. If MAPE calculated value is less than 10 %, it is interpreted as excellent accurate forecasting, between 10–20 % good forecasting, between 20–50 % acceptable forecasting and over 50 % inaccurate forecasting [30]. Selection of an error measure has an important effect on the conclusions about which of a set of forecasting methods is most accurate [16].

IV. Comparison of the Statistical Time Series Methods

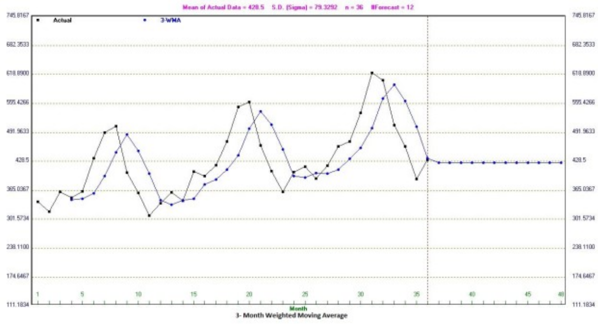
The Data Set

The time series data set used for comparing the above methods is "Airline data". The last 3 years of data was used out of the data set.

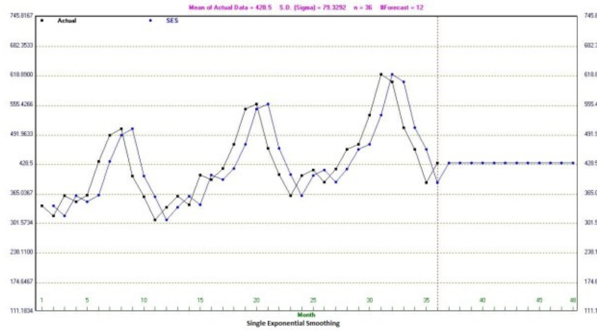
Outputs



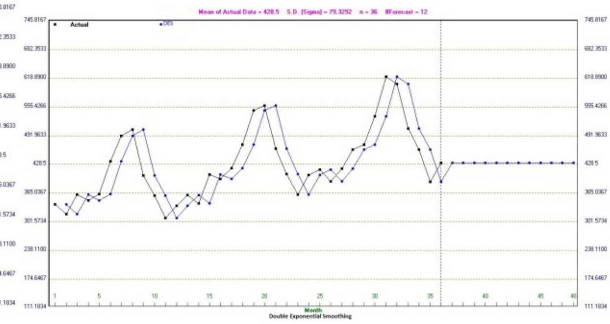
Graph 1: **3** Month Moving Average



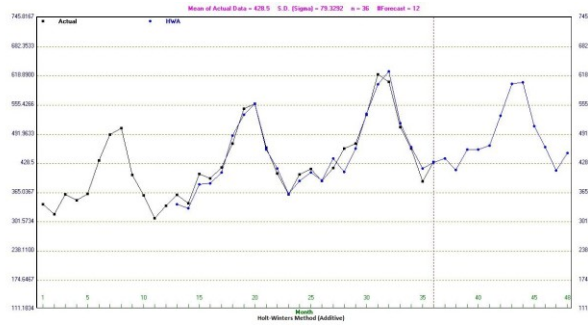
Graph 2: **3** Month Weighted Moving Average



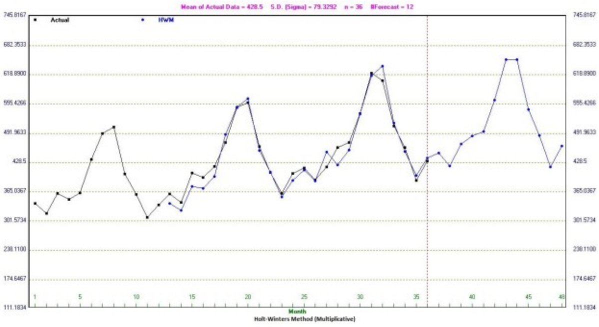
Graph 3: Single Exponential Smoothing



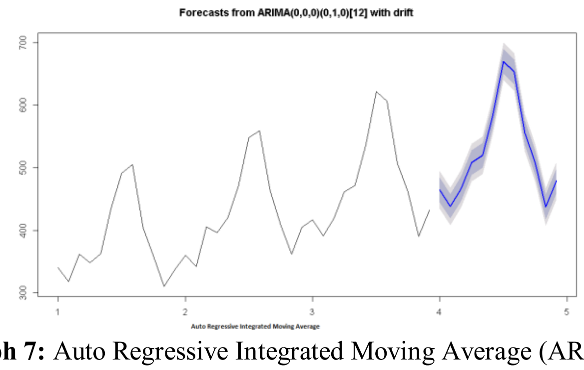
Graph 4: Double Exponential Smoothing



Graph 5: Holt-Winters Additive Method



Graph 6: Holt-Winters Multiplicative Method



Graph 7: Auto Regressive Integrated Moving Average (ARIMA)

Comparison

The moving average methods provide a very short term forecast. As the number of periods in average increases the forecast becomes less sensitive to changes. The moving average methods do not forecast trends well. They require sufficient historical data i.e. at least number of periods to be considered in average for forecasting a single period.

When we tested averaging methods on our dataset, they were found to be least accurate in comparison to other methods. Weighted moving average (WMA) with 3 months average and coefficients as 0.2, 0.3, 0.5 performed slightly better than 3 month moving average with MSE = 4519.241 as compared to MSE=5730.044 of 3 month moving average. The responsiveness of averaging methods were as follows

1) Forecast lags with increasing demand
2) Forecast leads with decreasing demand

Moving averages and weighted moving averages are effective in smoothing out sudden fluctuations in demand pattern in order to provide stable estimates but require maintaining extensive records of past data.

Single Exponential smoothing requires little record keeping of past data. This model can be seen as form of weighted moving average where weights decline exponentially and most recent data weighted most. It requires smoothing constant (α) which ranges from 0 to 1 and is subjectively chosen.

While the Single Exponential smoothing and Double Exponential smoothing methods performed a lot better than Averaging methods but both yielded similar results for accuracy (MSE). Moreover the trend was captured by Double Exponential smoothing.

Both variants of Holt-Winter's methods performed i.e. Multiplicative (MSE = 305.4738) and Additive (MSE = 294.5886) along with ARIMA model performed exceptionally well yielding very near to actual results.

These three methods could capture trend as well as seasonality. Also both the variants of Holt-Winter's method needed at least one complete cycle of data to extract seasonality before they could forecast, but once initialized they could yield medium to long term forecasts with good accuracy.

Holt-Winter's Additive method provides to lowest MSE value, hence provides best forecasts among the methods tested for this dataset.

V. Conclusion

Moving Average & Weighted Moving Average provides a fair idea of forecast but are not very accurate. The Single & Double exponential smoothing methods provide an idea that the most recent observations usually give the best guide to future, these values react to changes quickly if smoothing constants are chosen close to unity. Holt-Winters exponential model is generally performs good when data exhibits trend as well as seasonality as in our case. ARIMA has good accuracy if p,d,q i.e. the model selectors are chosen carefully. Model parameters play an important role in enhancing the accuracy of smoothing methods and are best left to be determined by a good software package.

References

[1]. http://www.idc.com/getdoc.jsp?containerId=prUS24194613 , Last Accessed On: 24-06-2014
[2]. http://www.idc.com/getdoc.jsp?containerId=241689 , Worldwide Business Analytics Software 2013–2017 Forecast and 2012 Vendor Shares , Last Accessed On: 24-06-2014
[3]. Michalewicz, Z., Schmidt, M., Michalewicz, M., Chiriac, C, Adaptive Business Intelligence (Springer-Verlag Berlin Heidelberg 2007)
[4]. http://en.wikipedia.org/wiki/Forecasting, Last Accessed On: 24-06-2014
[5]. http://www.forecastingprinciples.com/index.php/faq, Last Accessed On: 24-06-2014
[6]. http://www.igd.com/our-expertise/Supply-chain/Strategy-planning-technology/3586/Demand-Planning--Forecasting/ , Last Accessed On: 24-06-2014
[7]. http://www.businessdictionary.com/definition/na-ve-forecasting.html , Last Accessed On: 26-06-2014
[8]. http://prospectlearning.com/210/1NSLectures/NS06_Forecast_MovingA.pdf , Ted Mitchell, Last Accessed On: 26-06-2014
[9]. http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc42.htm , Last Accessed On: 26-06-2014
[10]. http://en.wikipedia.org/wiki/Moving_average , Last Accessed On: 26-06-2014
[11]. R. G. Brown, Statistical Forecasting for Inventory Control, McGraw-Hill: New York, 1959.
[12]. C. C. Holt, Forecasting trends and seasonals by exponentially weighted averages, O.N.R. Memorandum 52/1957, Carnegie Institute of Technology, 1957.
[13]. Li Zhi-Peng, Yu Hong, Liu Yun-Cai, Liu Fu-Qiang, An Improved Adaptive Exponential Smoothing Model for Short Term Travel Time Forecasting of Urban Arterial Street, Acta automatica sinica, Vol. 34, No. 11, 1404–1409, 2008
[14]. Prajakta Kalekar, Time series Forecasting using Holt-Winters Exponential smoothing, Kanwal Rekhi School of Information Technology, Mumbai, 2004.
[15]. V. Yorucu, The Analysis of Forecasting Performance by Using Time Series Data for Two Mediterranean Island, Review of Social, Economic & Business Studies, Vol. 2, 175–196, 2003.
[16]. Eva Ostertagov'a, Oskar Ostertag, Forecasting Using simple Exponential Smoothing Method, Acta Electrotechnica et Informatica, Vol. 12, No. 3, 2012, 62–66, DOI: 10.2478/v10198-012-0034-2, 2012.
[17]. R.G. Brown, R. F. Meyer, The Fundamental Theory of Exponential Smoothing, Operations Research,9, 673–685, 1961
[18]. A. D. Aczel, Complete Business Statistics, Irwin, 1989, ISBN 0-256-05710-8
[19]. Eva Ostertagov'a, Applied Statistics, [in Slovak], Elfa, Koˇsice, 2011, ISBN 978-80-8086-171-1
[20]. Eva Ostertagov'a, Oskar Ostertag, The Simple Exponential Smoothing Model, Proceedings of the 4th International Conference on Modelling of Mechanical and Mechatronic Systems, Technical University of Kosice, Slovak Republic, 380–384, 2011.
[21]. http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc435.htm , Last Accessed On: 03-07-2014
[22]. http://en.wikipedia.org/wiki/Exponential_smoothing , Last Accessed On: 03-07-2014
[23]. http://www.forecastingsolutions.com/arima.html , Last Accessed On: 07-07-2014
[24]. http://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average , Last Accessed On: 07-07-2014
[25]. http://people.duke.edu/~rnau/411arim.htm#pdq , Last Accessed On: 07-07-2014
[26]. http://people.bath.ac.uk/bm232/EC50162/ARIMA%20Model.ppt , Last Accessed On: 07-07-2014
[27]. Kisang Ryu, Alfonso Sanchez , The Evaluation of Forecasting Methods at an Institutional Foodservice Dining Facility, The Journal of Hospitality Financial Management, Vol. 11, No. 1, 2003
[28]. Jeffrey Jarrett, Business Forecasting Methods, Cambridge, MA: Basil Blackwell, 1991.
[29]. Spyros Makridakis, Steven Wheelwright, Rob Hyndman, Forecasting Methods and Applications, New York, Wiley, 1998.
[30]. C. D. Lewis, Industrial and Business Forecasting Methods, London, Butterworths, 1982.