Dhaka Univ. J. Sci. 71(1): 26-35, 2023 (January)
ISSN 1022-2502, E-ISSN 2408-8528
DOI: https://doi.org/10.3329/dujs.v71i1.65269

Analyzing Business Strategies of a Company in Bangladesh by Comparing Three Rigorous Forecasting Techniques

M. Babul Hasan[1*], M Asadujjamanand[1] and M. Hasibul Haque[2]

[1]Department of Mathematics University of Dhaka, Dhaka-1000, Bangladesh
[2]Department of Mathematics Khulna University of Engineering and Technology Khulna-9203, Bangladesh

(Received: 30 August 2022; Accepted: 31 January 2023)

Abstract

Business organizations in Bangladesh are basically running their business through intuition based forecasting. But it is crucial to anticipate the near future as accurate as possible to make the business profitable. This helps the manager of a business organization to plan their resources properly and as a result the organization can minimize its cost and maximize profit. In this research paper, we will analyze the business strategies of a company in Bangladesh by comparing the results obtained from three different rigorous forecasting techniques such as Holt's method, Holt-Winter's method and Autoregressive Integrated Moving Average (ARIMA) method so that the business organization can select proper forecasting technique to run their business. For this, we will first illustrate and analyze basics of forecasting and time series analysis, usual forecasting methods, some rigorous methods e.g., Holt's method, Winter's method and Autoregressive Integrated Moving Average (ARIMA) models. We will carry out our analysis and calculation by using Microsoft Excel, statistical data analysis tool R and MATHEMATICA.

Keywords: Business strategies, Forecast, Holts's Method, Winter's Method, ARRIMA Method.

I. Introduction

The introduction section introduces the reader to our research work. We design this section with several sub sections such as a brief history of forecasting, definition, importance of forecast in everyday life and outline of the paper and methodology.

From ancient time to modern era people make forecasts. Ancient people were not aware of that they made forecast because of lacking the proper knowledge but they did it for their sake. They had to predict where they should go to collect their food, what would the weather be etc.In the twentieth century, local or international business expands all over the world especially after World War II. Business organizations are always in search of new market places and busy with estimating market demands[1,2]. The Operations Research workers plays a vital role to estimate the future demands, costs, and so on by implementing the theory of forecasting rigorously. The researchers are continuously trying to develop better forecasting models.

Importance of Forecasting

This section describes the importance of forecasting in the field of managements. Forecasting is an important and necessary tool which can help the business organizations to plan their resources effectively. Forecasting provides the relevant and reliable information so that the manager of the company can plan their future events accordingly[3,4,5].

However, the elements of doing business are constantly changing. As a result the business organizations has to change their decision making process[6,7,8].

Methodology

The data, we will use in our numerical example, are secondary type. We analyze our data using R (a statistical data analysis programming language), Microsoft Excel and we use MATHEMATICA to solve LP and R in analysis data pattern and suitable ARIMA model selection. Mean absolute deviation (MAD) and mean absolute percentage error (MAPE) are used as error evaluation[9,10, 11,12].

Holt's method, Winter's method and ARIMA (Auto Regressive Integrated Moving Average) method for time series data with trend and seasonality are noteworthy. Here in this Section we only use the plot to see its trend or seasonality for Holt's and winter's method. For ARIMA, we use a statistical programming language **R** to see the pattern and other test such as ACF[13,14] (Auto-Correlation Factor). We also try to find any shortcomings in the method used for forecasting for a particular data set. In the last section of this Section, we will present detail about ARIMA method for trend involved data and programming code in R for selecting better ARIMA model[15,16].

The rest of the paper is organized as follows. In Section 2, we will discuss some rigorous methods of forecasting such as Holt's method. In Section 3, we will analyze Winter's method. In Section 4, ARIMA method will be analyzed. We will give numerical examples for each method to expand our understanding.In Section 5, we will draw our conclusion about our research work.

II. Holt's Method

For the case of pattern which represents only trend but no

* Author of correspondence, e-mail: mbabulhasan@yahoo.com

seasonality, Holt's method performs better than the other methods. This method is also known as "Holt-Winters double exponential smoothing" method.

Holt's method suppose that the sequence of observations is $\{X_t\}$, which begins at time $t = 0$. Smoothing error for time t is $\{S_t\}$ and best estimate of the trend at time t is $\{B_t\}$. Then the forecast F_{t+m} is estimated of the value of x at time $t+m$, $m>0$ based on the raw data up to time t. Then the double exponential smoothing is:

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + B_{t-1}) \quad (1)$$

and

$$B_t = \beta(S_t - S_{t-1}) + (1-\beta) B_{t-1} \quad (2)$$

where α and β are smoothing constants such that $0 < \alpha, \beta < 1$. The initial value of S_t and B_t and for $t > 1$ are calculated as follows:

$$S_0 = X_0 \text{ And } B_0 = \frac{X_n - X_0}{n}, n > 1 \quad (3)$$

And the h-step forecast by this method is given by the following equation:

$$F_{t,h} = S_t + h * B_t \quad (4)$$

Here, $F_{t,h}$ is the forecasted value for the period $t+h$ at period t based on the available data for the first t periods. We will illustrate how this method works in practice by the next example cum mathematical problem.

Numerical Example

An established RMG factory, Dhaka has the revenue collection data given in Table1 over the past 23 years. We are to forecast the revenue collection for the next three years by Holt's method.

Table 1. Revenue (in thousands) BDT of a RMG factory

Period	Revenue	Period	Revenue	Period	Revenue
1991	591	1999	1301	2007	2146
1992	620	2000	1440	2008	2430
1993	699	2001	1661	2009	2746
1994	781	2002	1770	2010	3069
1995	891	2003	1851	2011	3649
1996	993	2004	1954	2012	4159
1997	1111	2005	2023	2013	4686
1998	1149	2006	2079		

Solution Step1: Plot the data so observe the pattern.

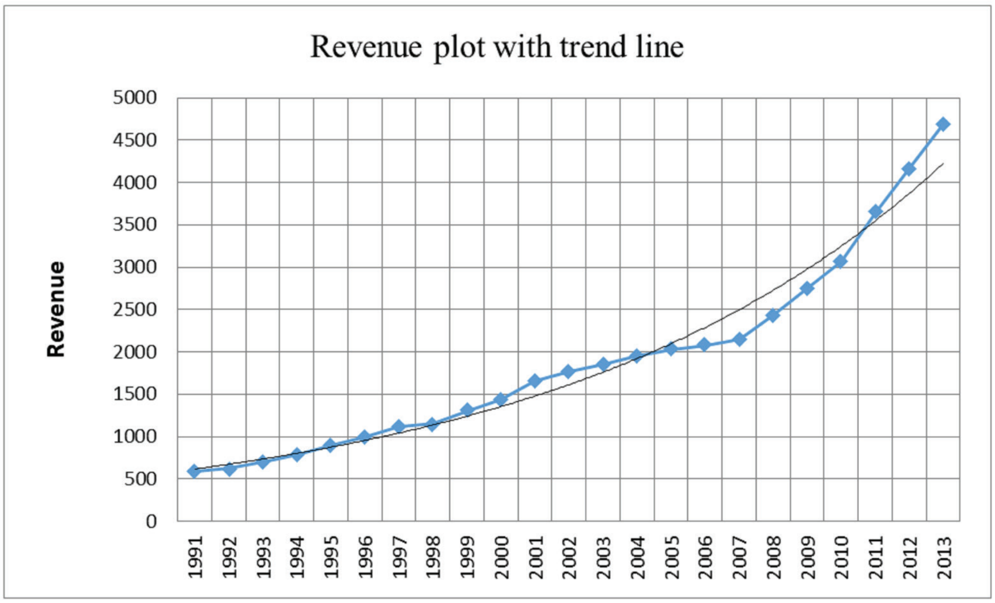

Fig. 1. Observed data VS trend line

From the Fig.1., we observe that there exists trend without seasonality and so we can apply Holt's exponential method. We note that a trend line is fitted to the observed data as plotted on same axes.

Step 2:Initialize $S_0 = X_0 = 591$, as the first observed data point and $B_0 = \frac{X_2 - X_0}{2} = \frac{699-591}{2} = 54$.

Analyzing several combination of different values of smoothing constants $and\ \beta$, we set the values for $\alpha\ and\ \beta$ as $\alpha = 0.7\ and\ \beta = 0.7$. We use Microsoft Excel to find such values.

Step 3: In this step, we manually calculate two values of smoothed level and trend each.

$S_1 = 0.7\,X_1 + 0.3 * (\ S_0 + \ B_0) = 0.7 * 620 + 0.3 * (591 + 54) = 627.5$

$B_1 = 0.7 * (\ S_1 - \ S_0) + 0.3 * \ B_0 = 0.7 * (627.5 - 591) + 0.3 * 54 \ = 41.75$

Forcast for period $1 = F_{0,1} = S_0 + B_0 = 591 + 54 = 645$

$S_2 = 0.7\,X_2 + 0.3 * (\ S_1 + \ B_1) = 0.7 * 699 + 0.3 * (627.5 + 41.75) = 690.075$

$B_2 = 0.3 * (\,S_2 - \ S_1) + 0.7 * \ B_1 = 0.3 * (690.075 - 627.5) + 0.7 * 41.75 = 56.3275$

$Forecast\ for\ period\ 2 = F_{1,1} = S_1 + B_1 = 627.5 + 41.75 = 669.25$

Similarly the rest of the values can be calculated. We use Microsoft Excel to manipulate these calculations and the obtained values are in the Table 2.

Table 2. Holt's method Calculation

T	Year	X_t	S_t	B_t	$F_{t-1,1}$	T	Year	X_t	S_t	B_t	$F_{t-1,1}$
0	**1991**	**591**	591.00	54.00	#NA	12	**2003**	**1851**	1878.81	116.19	1943.69
1	**1992**	**620**	627.50	41.75	645.00	13	**2004**	**1954**	1966.30	96.10	1995.00
2	**1993**	**699**	690.08	56.33	669.25	14	**2005**	**2023**	2034.82	76.80	2062.40
3	**1994**	**781**	770.62	73.28	746.40	15	**2006**	**2079**	2088.78	60.81	2111.62
4	**1995**	**891**	876.87	96.36	843.90	16	**2007**	**2146**	2147.08	59.05	2149.60
5	**1996**	**993**	987.07	106.05	973.23	17	**2008**	**2430**	2362.84	168.75	2206.13
6	**1997**	**1111**	1105.63	114.81	1093.12	18	**2009**	**2746**	2681.68	273.81	2531.59
7	**1998**	**1149**	1170.43	79.80	1220.44	19	**2010**	**3069**	3034.95	329.43	2955.49
8	**1999**	**1301**	1285.77	104.68	1250.24	20	**2011**	**3649**	3563.61	468.90	3364.38
9	**2000**	**1440**	1425.13	128.96	1390.45	21	**2012**	**4159**	4121.05	530.88	4032.51
10	**2001**	**1661**	1628.93	181.34	1554.09	22	**2013**	**4686**	4675.78	547.57	4651.93
11	**2002**	**1770**	1782.08	161.61	1810.27				$\alpha = 0.7\ and\ \beta = 0.7$		

From the Table 2, we observe that our smoothed values (sometimes called fitted values in Statistics) are well fitted for our choice of parameter values. To have a clear picture about the fitted value and one step ahead forecast value we plot them in the same axes shown in the following figure.

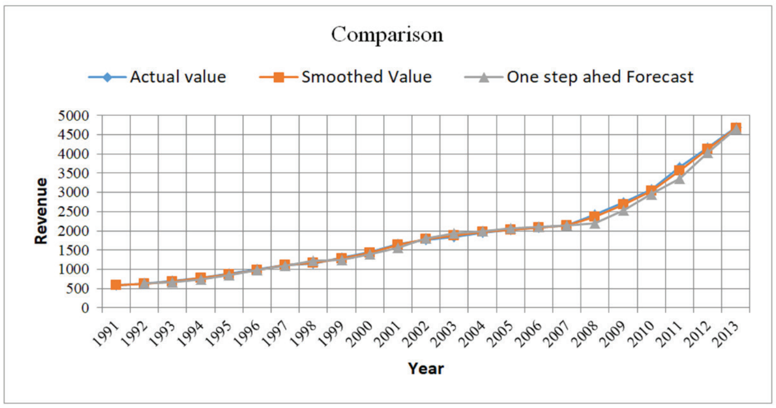

Fig.2. Comparison Plot

The Fig.2. makes us convince that our model is in good position enough that gives us well forecast as the smoothed aswell as one-step ahead forecast almost coincides with the actual revenue.

Step 4: Make the Forecast

Here we estimatethe forecasts for the next three periods namely year 2014, 2015, and 2016 period 23, 24, and 25 continued from the last given period 22. Hence forecast for h=1, 2, and 3 based on 22nd period are shown as follows.

$$F_{23} = F_{22,1} = S_{22} + 1 * B_{22} = 4675.78 + 1 * 547.57 = 5223.35$$

$$F_{24} = F_{22,2} = 4675.78 + 2 * 547.57 = 5770.92$$

$$F_{25} = F_{22,3} = 4675.78 + 3 * 547.57 = 6318.49$$

Thus we have our forecasts for period 23rd, 24th, and 25th which are BDT 5223.35, 5770.92, and 6318.49 thousand respectively.

III. Holt-Winter's Method

The generalized method of Holt's method by capturing seasonality is known as Holt-Winters' (HW) method of smoothing. Holt and Winter developed this method in 1960. HW method uses three smoothing constants and equations for base level, per period trend and seasonality in the given data series. This method is different from other forecasting methods because it uses iterative steps for forecasting.

HW method is classified in to two different versions such as additive or multiplicative process. The equations for multiplicative model are shown below.

Level: $L_t = \alpha\left(\frac{X_t}{I_{t-p}}\right) + (1 - \alpha)(L_{t-1} + T_{t-1})$ (5)

Trend: $T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$ (6)

Seasonality: $I_t = \gamma\left(\frac{X_t}{L_t}\right) + (1 - \gamma)I_{t-p}$ (7)

Forecast: $F_{t+k} = (L_t + k * T_t)I_{t-p+k}$ (8)

here, X_t is the observed series, p is the planning horizon, L_t gives the base level of the series, T_t represents the per period trend and α is the smoothing constant, 0 < α < 1, β is the trend smoothing constant, 0 < β < 1, and γ is the seasonal smoothing constant, 0 < γ < 1, and $F_{t,k}$ presents forecasts for k-periods ahead.

The Seasonal Additive HW equations are given as:

$$L_t = \alpha(X_t - I_{t-p}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (9)$$

$$Trend: T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (10)$$

$$Seasonality: I_t = \gamma(X_t - L_t) + (1 - \gamma)I_{t-p} \quad (11)$$

$$Forecast: F_{t+k} = (L_t + k * T_t) + I_{t-p+k} \quad (12)$$

Initialization of HW Method

The better the estimate of initial values of base, trend, and all seasonal factors the better the forecasts. Let L_0 be the Estimate of the base level at beginning of month / Quarter 1, T_0 be the estimate of the trend at beginning of month / Quarter 1, I_{1-p} be the estimate of January / Quarter (Q) 1 seasonal factor at beginning of the Month1 / Q1, I_{2-p} be the e stimate of February / Q2 seasonal factor at beginning of the month 1 / Q1. In a similar way, we get

$I_0 =$ Estimate of December / Q4 seasonal factor at beginning of Month 1 / Q1.

We have several methods to estimate the parameters defined above. Among them one needs two years data monthly or quarterly. We suppose that we are provided with two years actual data. In this case the initial estimate of trend level and base level we can take as follows:

$T_0 = \frac{\bar{A}_1 - \bar{A}_2}{p}$ $L_0 = A_1 \frac{p}{}$ (median of all values assumed by p) * (13)

T_0

where, $\bar{A}_1 =$ Average monthly or quarterly data during Year 1

$\bar{A}_2 =$ Average monthly or quarterly data during Year 2 (previous year)

To estimate the seasonality index for a given period (say, January= I_{-11} or Q1= I_{-3}), we make an estimate of January or Q1 seasonality for Year 2 and Year 1 and average them. To do this first we divide the actual value of January or Q1 of corresponding year by the average value of that year and then we average them. Seasonal factors must be normalized so that their average in multiplicative case is one[13,14].

Numerical Example on HW Method

We have been given in Table 3 with quarterly data for five years. We are to forecast for the next year quarterly using appropriate HW method.

Table 3. Quarterly observed data

Year\ Quarter	Q1	Q2	Q3	Q4
1	318	380	358	423
2	379	394	412	439
3	413	458	492	493
4	461	468	529	575
5	441	548	561	620

Step 1: We plot the given data point to see its pattern that is whether it contains trend or seasonality or both of them. Microsoft Excels produces following graph.

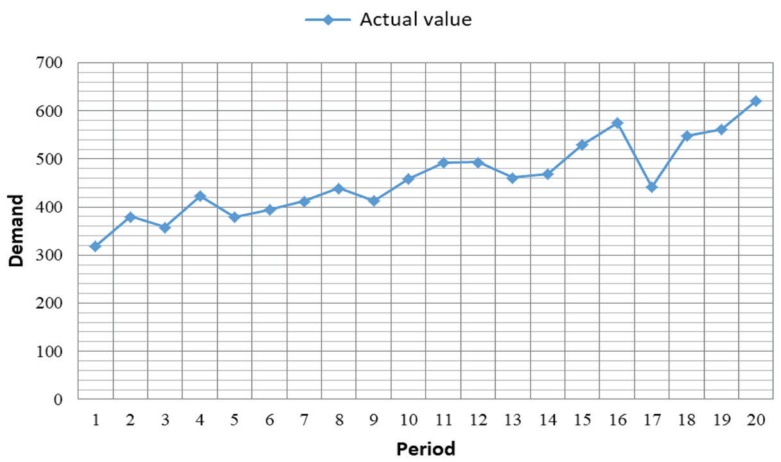

Fig. 3. Observed data plot

From Fig. 3.,we observe that it shows both seasonality and trend with an increasing tendency in seasonality and therefore we use multiplicative HW method to forecast.

Step 2: Initialization

Firstly, we calculate the average demand of each year and center of period of each year. Our calculation results are shown in the following table.

Table 4. Average values

Year\Quarter	Q1	Q2	Q3	Q4	Average	Center of avg. (t)
1	318	380	358	423	369.75	2.5
2	379	394	412	439	406	6.5
3	413	458	492	493	464	10.5
4	461	468	529	575	508.25	14.5
5	441	548	561	620	542.5	18.5

Now using equation 13 and 14 we calculate the initial trend and base as follows:

$$T_0 = \frac{Avg\ of\ year\ 2 - Avg\ of\ year\ 1}{p}$$

$$= \frac{406 - 369.75}{4} = 9.0625$$

$$L_0 = 369.75 - 2.5 * 9.0625 = 347.09375$$

Step 3: Seasonal Index Initialization

Here we divide the quarterly data by the respective yearly average to obtain the seasonal indices for the first two years.

Table 5. Initial Index Calculation

Quarter	Year 1 Index	Year 2 Index	Average Index (Y1+ Y2) ÷ 2	Normalized Index
Q1	318 ÷ 369.75=0.86	379 ÷ 406=0.93	0.895	0.9
Q2	380 ÷ 369.75=1.03	394 ÷ 406=0.97	1.00	1
Q3	358 ÷ 369.75=0.97	412 ÷ 406=1.01	0.99	0.99
Q4	423 ÷ 369.75=1.14	439 ÷ 406=1.08	1.11	1.11
			Total=3.995	Total=4

Normalized indices are obtained by dividing each average index by 3.995 and then multiplied by 4 e.g., (0.895 ÷ 3.995)*4=0.9. Therefore we have now our initial seasonal indices and we denote them as $I_{-3} = 0.9$ for Q1, $I_{-2} = 1$ for Q2, $I_{-1} = 0.99$ for Q3, and $I_0 = 1.11$ for Q4. Now we are in a position to make the first forecast, which is for the first quarter of year 1, according to the formula given by the equation 4.9:

$F_{0+1} = (L_0 + 1 * T_0) * I_{-3} = (347.09375 + 1 * 9.0625) * 0.9 = 320.54$ (14)

Step 4: Detail Calculations and Choosing Parameters Values

Detail calculation and parameter selection are cumbersome task and for our convenience we use Microsoft Excel to calculate required trend, level, and seasonality index for a choice of parameter values. After this, we change the parameter values by trial and error and have cautious eye on the change in the one step-ahead forecast and actual values.

Now we use the Table 5 to calculate the MAPE. MAPE is determined just dividing the total in the last column of Table 6 by the number of data points i.e., 20 which gives us 4.21% and this should be taken into account when we forecast for the next year. Parameter values are taken so that the error is lowest as possible. We check several combinations of their different values and see that the values we have chosen yield the lowest error.

Step 5:Now we are in a position to forecast for the next year ornext four quarter that is for the period 21st, 22nd, 23rd, and 24th at 20th period.

According to Equation 4.9 we have,

$$F_{20+1} = \left(L_{20} + 1 * T_{20}\right) * I_{20-4+1} = \left(L_{20} + T_{20}\right) * I_{17}$$
$$= \left(562.5276 + 10.85729\right) * 0.906861 = 519.9802$$

$$F_{20+2} = \left(L_{20} + 2 * T_{20}\right) * I_{20-4+2} = \left(L_{20} + 2 * T_{20}\right) * I_{18}$$
$$= \left(562.5276 + 2 * 10.85729\right) * 0.998754 = 583.5144$$

$$F_{20+3} = \left(L_{20} + 3 * T_{20}\right) * I_{20-4+3} = \left(L_{20} + 3 * T_{20}\right) * I_{19}$$
$$= \left(562.5276 + 3 * 10.85729\right) * 1.006588 = 599.0202$$

$$F_{20+4} = \left(L_{20} + 4 * T_{20}\right) * I_{20-4+4} = \left(L_{20} + 4 * T_{20}\right) * I_{20}$$
$$= \left(562.5276 + 4 * 10.85729\right) * 1.093611 = 662.6811$$

Thus we have our forecasts values for the four quarter of next year made at the end of period 20 and they are 519.9802 for Q1 and 583.5144 for Q2 and 599.0202 for Q3 and 662.6811 for Q4. The very next figure shows the actual values as well as our forecast values.

Table 6. HW calculations with α = 0.1, β = 0.75, γ = 0.15

Period(t)	A_t	L_t	T_t	I_t	$F_{t,t-1}$	$\frac{\lvert A_t - F_t \rvert}{A_t} \times 100\%$
1	318	355.874	8.850781	0.899036	320.54	0.798742
2	380	366.2523	9.996426	1.00563	364.7247	4.019805
3	358	374.7854	8.898986	0.984782	372.4862	4.046426
4	423	383.4241	8.703735	1.108983	425.8897	0.683147
5	379	395.0713	10.91134	0.908079	352.5371	6.982291
6	394	404.5638	9.847194	1.000869	408.2685	3.621446
7	412	414.8065	10.14387	0.98605	408.1045	0.94552
8	439	422.0412	7.961969	1.098663	471.2626	7.349112
9	413	432.4835	9.822204	0.915109	390.4768	5.453559
10	458	443.8354	10.96945	1.005526	442.6901	3.342768
11	492	459.2204	14.28113	0.998849	448.4602	8.849556
12	493	471.0241	12.42307	1.090862	520.2184	5.520974
13	461	485.4789	13.9469	0.92028	442.4071	4.033167
14	468	496.0261	11.39707	0.996222	502.1856	7.30462
15	529	509.6418	13.06103	1.00472	506.8393	4.189162
16	575	523.1431	13.3913	1.092101	570.1964	0.835415
17	441	530.8012	9.091373	0.906861	493.7618	11.96412
18	548	540.9111	9.855301	0.998754	537.8528	1.851684
19	561	551.5263	10.42517	1.006588	553.3659	1.360806
20	620	562.5276	10.85729	1.093611	613.7078	1.014869

Total=84.16719 %

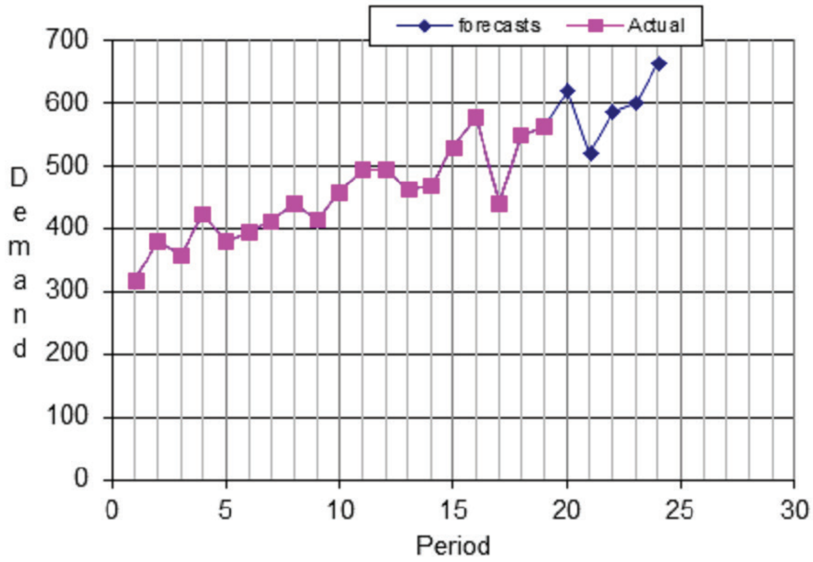

Fig. 4. Forecast with Actual value

IV. ARIMA Model

In this section, we discuss details about ARIMA model and illustrate with numerical example using statistical data analysis tool R.

In 1972, Box and Jenkins developed a method for analyzing stationary univariate time series data by implementing ARIMA. Their forecasting models are based on statistical concepts and principles and are able to model a wide spectrum of time series behavior.

Stationary series have some advantages over non-stationary series. Non-stationary series have gradually diminishing autocorrelations that can be function of time, whereas stationary series have stable but rapidly diminishing autocorrelations. For these reasons, a process must be weakly stationary for being modeled in time series using Box-Jenkins approach[22].

Examining stationarity in the time series data

If there is no growth or decline in the data then it is stationary. The data must be roughly horizontal along the time axis.

The formulation of the random walk process is

$$X_t = X_{t-1} + W_t \tag{15}$$

So that $$X_t - X_{t-1} = W_t \tag{16}$$

To make a time series stationary, it is necessary to difference the series. A log transformation can be taken before first differencing if non-stationarity in variance is evident. If log transformation is taken, the forecasts will come also in the form of log series for which an anti-log must be taken to get the forecasts for original series. The process which can be transformed into stationary by differencing is called difference stationary. Occasionally, the first differenced data does not appear stationary and second differencing is needed.

The stationarity can be checked visually from the time plot of the series or from the plots of ACF and PACF. If the ploted data are scattered horizontally around a constant mean, or equivalently, the ACF and PACF drop to near zero quickly. Otherwise, non-stationarity is implied.

Tests of stationarity

Portmanteau test is used to identify Stationarity of a data set. A common Portmanteau test is the Box-Pierce test which is based on the Box-Pierce Q statistic:

$$Q = n\sum_{k=1}^{h} r_k^2 \ , \tag{17}$$

where, h = maximum lag being considered, n = number of observations in the series and $r_k =$ autocorrelation at lag ***k***.

An alternative portmanteau test is the Ljung-Box test and is shown as follows.

$$Q^* = n(n+2)\sum_{k=1}^{h} \frac{r_k^2}{n-k} \ , \tag{18}$$

The white noise residuals indicates that the statistic Q^* has a chi-square distribution with degrees of freedom (h - m). Thus it is to conclude that the data are not white noise if the value of Q or Q^* lies in the extreme 5% of the right-hand tail of the distribution.

Another approach is Dickey-Fuller test. A simple autoregressive model, AR (1) is

$$X_t = \rho X_{t-1} + W_t \ , \tag{19}$$

Where, ρ is a coefficient A unit root is present if ρ = 1.

The model would be non-stationary in this case. This model can be written as

$$\Delta X_t = (\rho - 1)X_{t-1} + W_t = \gamma X_{t-1} + W_t \quad (20)$$

where, Δ be the difference operator. This model can be estimated and testing for a unit root is equivalent to testing Δ= 0 (where $\gamma = (\rho - 1)$).

An extension of Dickey-Fuller test is augmented Dickey-Fuller (ADF) test which is a test for unit root in a time series as shown below.

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \beta_1 \Delta Y_{t-1} + \dots\dots\dots + \beta_p \Delta Y_{t-p+1} + W_t, \quad (21)$$

Where α is a constant, β the coefficient on a time trend and p the lag order of the autoregressive process. The unit root testis then carried out under the null hypothesis $\gamma = 0$ against the alternative hypothesis of $\gamma < 0$. Once a value for the test statistic

$$DF_t = \frac{\hat{\gamma}}{s.e.(\hat{\gamma})}, \quad (22)$$

is computed. If the test statistic is less than the critical value, then the null hypothesis of $\gamma = 0$ is rejected and no unit root is present.

Formulation of ARIMA Model

The Box-Jenkins ARIMA model are classified as the autoregressive process, the integrated process and the moving average process. The general non-seasonal model is known as ARIMA (p, d, q) and applied to non-stationary data series. In the model

AR: p = order of the autoregressive part

I: d = degree of the first differencing involved

MA: q = order of the moving average part.

Using back shift notation, this model is written as,

$$\emptyset(B)(1 - B)^d X_t = \delta + \theta(B)W_t \quad (23)$$

where $\delta = \mu(1 - \varphi_1 - \varphi_p)$ for $E(\Delta^d X_t) = \mu$

Selection of best model

Spikes at lag 1 to p, and cuts off to zero in PACF indicates AR (p) model. Again, Spikes at lag 1 to q, and cuts off to zero in ACF indicates MA (q) model.

The Akaike information criterion (AIC) can be used to select the best model, which is given by

$$AIC = -2\ln L + 2m, \quad (24)$$

where L = maximum likelihood, m = number of parameters. We choose the model with the minimum AIC.

Diagnostic Checking

Before forecasting using the selected model, the diagnostic checks of the residuals should be performed. If the assumptions are not satisfied, then the selected model is no longer valid and a new model is to be chosen.To assess whether the residuals are white noise, the portmanteau tests such as Box-Pierce test and Ljung-Box test are useful. Box-Pierce test is useful only when sample size is large, but the Ljung-Box test can be used even when the sample size is small. For both the tests, if the p-value is large, then the residuals are thought to be white noise. The plots of standardized residuals and ACF of residuals can also be used for this purpose.

An approach to check the normality assumption of residuals is by using normal q-q plot. The assumption is satisfied when the plot is almost along a straight line.

An equivalent approach to check the normality assumption of residual considering hypothesis testing is Shapiro-Wilk test.

To check the randomness of the residuals can be checked by using run test. If the p-value is a large one, then the null hypothesis of randomness is not rejected.

If all the mentioned assumptions are satisfied by a model, only then the model can be used for the forecasting purposes.

Numerical Example

Suppose we are given a set of monthly sale data of a product offered by a newly started business organization for a year which is 4,6,8,10,14,18,20,22,24,28,31

and 34 lakh BDT from January 2013 to December 2013. We are to fit an ARIMA model to forecast for next six month.

Solution

Step-1: Plot the given data

From the figure 5, we see that a linear trend exists here.

Step 2:ACF and PACF test

ACF plot shows that it is decreasing exponentially which implies trend exist.

Step 3: Differencing data to remove trend and make the data stationary.

Step 4: Plot the differenced data, ACF and PACF of differenced data

After observing the figure 7, 8 and 9, we see that our differenced data is stationary as its ACF plot shows a random variation with 1 significant lag and PACF has no significant lag. Thus we have the p=0, d=1, and q=1 for our data set to build ARIMA (p, d, q) model.

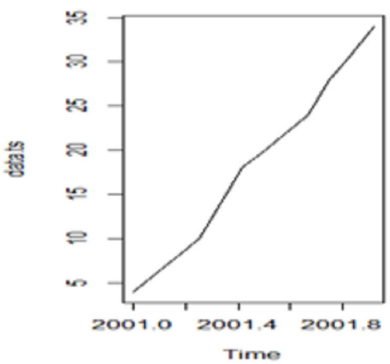

Fig.5. Time plot of original data

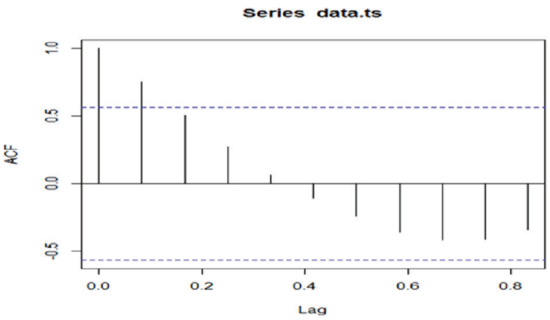

Fig. 6. ACF plot of original data

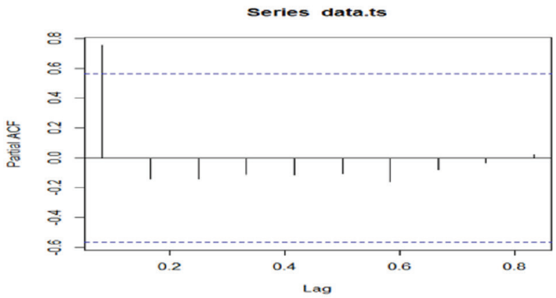

Fig. 7. PACF plot of original data

Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

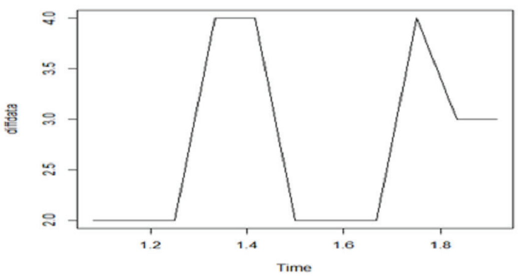

Fig. 8. Graph of differenced data

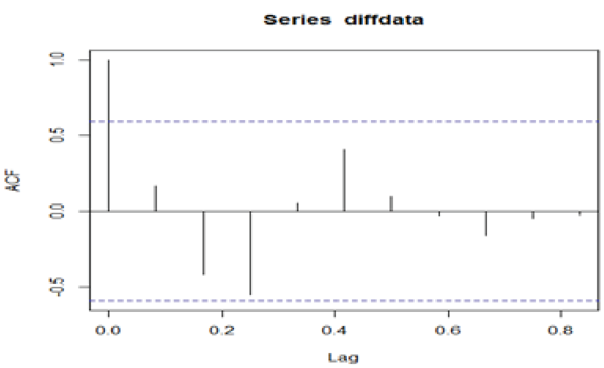

Fig. 9. ACF graph of differenced data

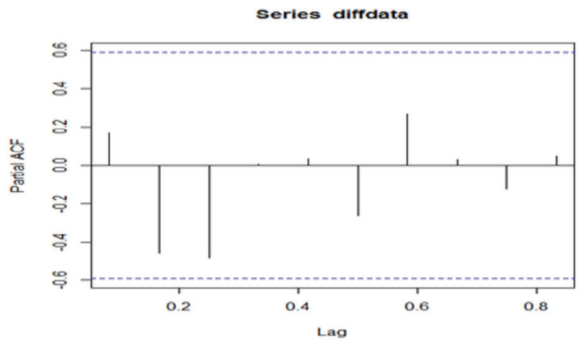

Fig. 10. PACF graph of differenced data

Step 5: Model Building

We run several model with various combination of p and q keeping d fixed in ARIMA (p, d, q) and nearest model as possible and list the AIC values in the following table.

Since AIC value for ARIMA (1, 1, 0) is the lowest we choose it our primarily fitted model**.**

Step 6: Forecast

Step 7: Error analysis

Table 7. AIC values

ARIMA (p, d, q)	AIC
ARIMA(0, 1, 1)	49.94
ARIMA(1, 1, 1)	39.62
ARIMA(1, 1, 0)	39.33

Table 8. Forecast values with 95% confidence level

Point	Forecast	Lo 95	Hi 95
January 2014	36.73752	34.55905	38.91599
February 2014	39.23553	34.53402	43.93704
March 2014	41.51498	33.90787	49.12208
April 2014	43.59499	32.81679	54.37318
May 2014	45.49301	31.35732	59.62871
June 2014	47.22498	29.60178	64.84817

Table 9. Errors from different methods

Error measures	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.3487024	1.064168	0.732045	2.905805	4.630305	0.08873272

V. Conclusion

This paperfocused on basic concept of time series forecasting and analyzed some rigorous forecasting techniques. We analyzed the business strategies of a company in Bangladesh by comparing the results obtained from three different rigorous forecasting techniques such as Holt's method, Holt-Winter's method and Autoregressive Integrated Moving Average (ARIMA) method. For this, we first illustrated and studied basics of forecasting and time series analysis, usual forecasting methods, some rigorous methods e.g., Holt's method, Winter's method and Autoregressive Integrated Moving Average (ARIMA) models. We carried out our analysis and calculation by using Microsoft Excel, statistical data analysis tool R and MATHEMATICA. We hope that, this research will help the business organization to select perfect forecasting method for minimizing the cost and maximizing the profit.

Acknowledgement

This research work was funded by the Dhaka. University centennial research grant

References

1. Berggren J.L.2000, Reviewed Work(s): Encyclopedia of the History of Arabic Science by RoshdiRashed, *Journal of the American Oriental Society*, **120**, 282-283.
2. Richardson, Lewis Fry, 1922, Weather Prediction by Numerical Process, Cambridge University Press, England.
3. Maurice O.A., T.O Francis. and D.O Samuel. 2013, Determining the Better Approach for Short-Term Forecasting of Ghana's Inflation: Seasonal ARIMA vs. Holt-Winters, *International Journal of Business, Humanities and Technology*, **3(1)**, 69-79.
4. Pissal Y., A. Pinnoi and A. Charoenthavornying, 2001, Demand Forecasting and Production Planning for Highly Seasonal Demand Situations, *ScienceAsia*, **27**, 271-278.
5. Mentzer J.T. and Cox, 1984, Familiarity Applications and Performance of Sales Forecasting Techniques, *Journal of Forecasting*, **3(1**), 27-36.
6. Lazim Abdullah, 2012, ARIMA Model for Gold Bullion Coin Selling Prices Forecasting*, International Journal of Advances in Applied Sciences*, **1(4)**, 153-158.
7. Armostrong J.S. 2001, Principles of Forecasting, Kluwer Publishers, Massachsetts, USA.
8. Makridakis S., C.S, Wheelwright and J.R. Hyndman, 1998, Forecasting: Methods and Applications, 3rd Edition, Wiely, New Jersey.
9. Makridakis, S. and M.Hibbon, 1979, Accuracy of Forecasting: An Empirical Investigation, *Journal of the Royal Statistical Society*, Series A,**142**, 97-145.
10. Robert S.H., and D.S. Stoffer, 2011, Time Series Analysis and its Application, 3rd Edition, *Springer*, New York.
11. Gardner, E.S, Jr ,1985, Exponential Smoothing: The State of Art, *Journal of forecasting***, 4**, 1-28.
12. Gardner E.S. and E.A. Anderson, 1997, Focus Forecasting Reconsidered, *International Journal of Forecasting*, **13(4)**, 501-508.
13. Frank, M. B. 1969, A New Product Growth for Model Consumer Durables, *Management Science*, **15(5)**, 215-227.
14. Hyndman, R.J, B.K. Anne, J. Keith, and D. Snyder, 2008, Forecasting with Exponential Smoothing: The State Space Approach, 1st Edition, Springer,Berlin.
15. Chatfield C., and M. Yar, 1988, The Statistician, *Journal of the Royal Statistical Society*, **37(2)**, 129-140.
16. Bun D. 1989, Forecasting with more than one models, *Journal of Forecasting***, 8**, 161-166.