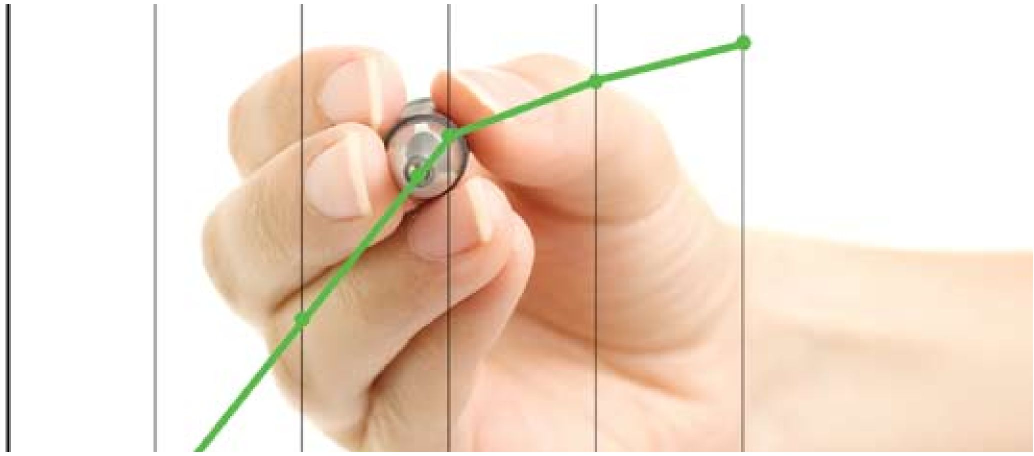
Observations on Some Exponential Smoothing Forecasting Methods



Forecasting is an important function in business management and other fields. From demands of products, stock market prices, economic trends, technological changes, weather – accurate forecasts would be of immense use in many situations. Like forecasting arises in many real life purposes, there are many methods available for the purpose of forecasting. The author in this article considers forecasting methods as the single, double and triple exponential smoothing methods in which additive trend and seasonality may be included. These are compared among themselves and also with last-value forecasting. This is done using a theoretical model and with data from an actual time series. It is observed that, single exponential smoothing method does not give consistent results. Last-value forecasting, like single exponential smoothing method, is not much accurate often. From the empirical evidence, among the methods considered, triple exponential smoothing method appears to be the most suitable for practical application.

Pritibhushan Sinha

(The author is Consultant (Operations Management & Research) based in Kolkata. He can be reached at p_sinha@live.com)

Introduction

Forecasting is an important function in business management and other fields. From demands of products, stock market prices, economic trends, technological changes, weather – accurate forecasts would be of immense use in many situations. Like forecasting arises in many real life purposes, there are many methods available for the purpose of forecasting. We may refer to Brown (1959), Box et al. (1994), Fildes et al. (2008) and others for discussions on the topic.

It is important that, given the situations, the forecasting method is selected correctly by the planners and the decision-makers. Forecasting methods may be broadly categorised as – (a) Experiential/ Judgmental Methods; (b) Causal Methods; (c) Time Series Methods. Experiential or judgmental methods are based on experience and perception. With such methods, forecasts may be subjective and inconsistent. In a causal method, factors (synonymously, input/ independent/control variables) which physically or economically affect the variable(s) (output/dependent variable) to be forecast are identified. Functional relationships, exact or approximate, are established among the independent and dependent variables. Such independent causal variables could be rainfall during the period considered, population, average *per capita* income in the population, etc. The relationships may be statistical, involving random deviations. We may, for instance, use a statistical linear regression equation between the independent and dependent variables. The relationship may also be a deterministic one, such as a neural network regression equation. In a causal method, we have to forecast future values of the independent variables mostly unless such future values are known in advance or we see present value of such a variable as causal to the future values of the dependent variable. In a time series method, a relationship is assumed, implicitly or explicitly, only between time and the dependent variable. The inherent assumption is that, each causal variable changes with time in a particular manner and this leads to a predictable pattern of change of the dependent variable with time. A time series method is the well-known autoregressive integrated moving average (ARIMA) (see, Box et al. 1994) technique. Many computer routines are available for forecasting.

Some oft-used time series methods of forecasting are exponential smoothing (ES) type of methods. Relatively well-known of such methods are single (also called as, simple) ES (Brown 1959), double and triple ES (Holt 1957, Winters 1960). Such methods have comparatively less number of parameters, satisfying the principle of parsimony, valued for forecasting methods. These are easy to implement and are suitable for automated forecasting in system consisting of a large number of items, for example, a large retail store. There has been a substantial amount of research on the ES methods, as indicated by Gardner (1985) earlier and more recently (Gardner 2006). ES methods have been surprisingly successful in many actual applications and also in forecasting competitions, as described by Makridakis and Hibon (1979), Makridakis et al. (1982), Makridakis and Hibon (2000). Yet opinion is divided about the accuracy, practical usefulness and theoretical correctness of such methods. There is no conclusive answer to this till now.

From the observations in the numerical experiments, we may write that, single ES method may give good results sometimes, but the performance is not consistent. Last-value forecasting also should be avoided. Among the methods discussed, the best choice in most of the cases would be triple exponential smoothing method. If there is a marked seasonality pattern, period of which has been identified relatively precisely, and forecasts are for short terms, this method can give a forecasting accuracy adequate for practical purposes. At the same time, other methods may also be explored. ”

Another forecasting method that also finds frequent application is the last-value forecasting or naïve forecasting (see, for example, Hillier and Hillier 2005). In this method, the most immediate preceding value of the time series is used as the forecast. This article compares the three ES methods as mentioned and last-value forecasting using two test data sets. We offer an explanation and describe some conditions when single ES method would be the same as last-value forecasting. Of the test data sets one is obtained from an assumed model. Data for another are from a real-life situation. As far as we are aware, such comparisons and analysis have not been reported in the literature for these methods. The implications of the observations from the study are discussed.

Exponential Smoothing Forecasting Methods and Measurement of Accuracy

Let's talk about the three ES methods and usual ways to measure the accuracy of a forecasting method on the basis of related literature.

ES Methods:

In the single exponential smoothing method, a forecast, following the notation as used by Gardner (2006), is given as,

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

$$\bar{X}_t(m) = S_t, \quad (2)$$

where X_t is the observed (actual) value of the time series, S_t is the smoothed value of the series, computed after X_t is observed, in period t = 1, 2, …. A suitable value for S_0 is used; usually $S_0 = X_1$, yielding $S_1 = X_1$. $\bar{X}_t(m)$ is the forecast given for a period, m periods later from the t-th period; α (0 ≤ α ≤ 1) is a non-negative multiplier, the smoothing constant. To illustrate, suppose actual values (say, weekly demand of an item) of a time series are as: X_1 = 100, X_2 = 78, X_3 = 92, …, Take α = 0.9. With $S_0 = X_1$, $S_1 = X_1$ = 100. If forecasts are made two periods ahead (i.e., m = 2), then this is the forecast for the third period. S_2 = 0.9 × 78 + 0.1 × 100 = 80.2, S_3 = 0.9 × 92 + 0.1 × 80.2 = 90.8 are the forecasts for the fourth and the fifth week respectively.

The equations for double ES, which takes in account an additive trend component, are as follows.

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + T_{t-1}) \quad (3)$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1} \quad (4)$$

$$\bar{X}_t(m) = S_t + mT_t, \quad (5)$$

The smoothed trend component being represented with T_t and γ ($0 \le \gamma \le 1$) being another parameter in the model – the smoothing constant for trend. We may take $T_0 = 0$, or any other suitable initial value, S_0 being as before.

For triple ES, considering additive trend and seasonal component, the equations are as:

$$S_t = \alpha(X_t - I_{t-p}) + (1-\alpha)S_{t-1} \quad (6)$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1} \quad (7)$$

$$I_t = \delta(X_t - S_t) + (1-\delta)I_{t-p} \quad (8)$$

$$\bar{X}_t(m) = S_t + mT_t + I_{t-p+m} \quad (9)$$

where I_t represents smoothed seasonal component for period t, p is the period of seasonal cycles and δ ($0 \le \delta \le 1$) is the smoothing constant for the seasonal component. I_0 and earlier seasonal component values may be initialised with 0, other initialising values may be as before.

Forecasting Accuracy Criteria

The parameters in the above methods are often fixed by optimising some criterion, using past data. Commonly used criteria are sum of squared errors (SSE) (equivalently, root of the mean of sum of squared errors (RMSE)), mean of absolute error percentages (MAEP), etc. Let, F_i be the (proxy) forecast for the i-th period, X_i is the actual value for this period. Then SSE is $\sum(X_i - F_i)^2$, taking the sum over the periods considered. RMSE is its square root. MAEP is the average of the values $(|X_i - F_i| \,/\, |X_i|) \times 100\%$ ($X_i \ne 0$). These are indicators of the fitness of the model for the past data. Accuracy of a method would be measured for forecasts for future periods in the same manner, using such indicators.

A Comparison of the Methods

Let's discuss some theoretical points and present some experimental observations now.

- ***Last-Value Forecasting* vis-à-vis *Single ES Forecasting***

First we note a relationship, relevant to our discussions, between single ES method and last-value forecasting. From equations (1) – (3), the m-periods ahead forecast for the $(t+m)$-th period with single ES method is given as:

Suppose, the actual values in the time series are either increasing ($X_1 \le X_2 \le \ldots$) or decreasing ($X_1 \ge X_2 \ge \ldots$) and we take $S_0 = X_1$, which is a reasonable initialising value, particularly when the values of the parameters would not be changed for forecasting for a few periods. The value of the constant α is fixed optimising a criterion, based on past observations, as $\sum f((X_i - F_i)^2)$, where f is an increasing function. It is not difficult to see that, in such a situation, optimal value of α would be 1, implying that single ES method would give the same forecasts as last-value forecasting. Such a case may happen when there is a strong trend in the data, or time span as covered by the data is small, so that the observations are either increasing or decreasing.

- ***Experimental Observations***

The experiments have been done in the following manner. For two sets of test data, data for 20 consecutive periods are considered. We consider 1-period and 2-periods ahead forecasts. The parameters (α, γ, δ) in the methods are determined minimising RMSE, based on data for first a few periods, i.e., minimising $\sqrt{\sum(X_i - F_i)^2}$ for these periods. Forecasting accuracy in terms of RMSE and MAEP is then noted for the last 3 periods. For initialisation we use, $S_0 = X_1$, $T_0 = 0$, $I_t = 0$, for $t = 0$, or negative. For 1-period ahead forecasts, proxy forecasts are considered for 2nd to 17th period. This is done for 3rd to 17th period for 2-periods ahead forecasts. The experiments have been done in *MS Excel*, solving the optimisation problems with the *Solver* utility.

3.2.1. Data Set with an Assumed Model

Let $X_t = 52 + 10t + 5\sin(1.5t) + \varepsilon$, $t = 1, 2, \ldots$ We assume that, the random error ε follows independent, uniform distribution in (-2, 2), with average zero. It may be seen that, the series is increasing, even though there is seasonality and random error. So, single ES method would be only last-value forecasting. A particular realisation (Table 1) of the series is considered. Double and triple ES methods are applied to see the accuracy. For triple ES method we take period of seasonal

$$F_{t+m} = \bar{X}_t(m) = S_t = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \cdots + \alpha(1-\alpha)^{t-1}X_1 + (1-\alpha)^t S_0. \quad (10)$$

cycles, $p = 4$, approximately equal to the period of the sine component. The results are shown in Table 2. As anticipated, single ES method gives last-value forecasting. Triple ES method gives the best result both in terms of model fitness and predictive accuracy. In two-periods ahead forecasts, single ES method performance is not satisfactory.

Table 1: Time Series Values for a Realisation of the Theoretical Model

Obs. No.	Value	Obs. No.	Value	Obs. No.	Value	Obs. No.	Value
1	65.572	6	114.519	11	156.865	16	209.155
2	72.340	7	115.651	12	167.666	17	223.680
3	78.626	8	128.443	13	184.724	18	237.740
4	89.812	9	145.376	14	196.632	19	241.291
5	106.183	10	153.451	15	199.771	20	245.422

Table 2: Forecasting Accuracy for Theoretical Model Data

Obs. No.	Method	Optimal Parameter Values	RMSE (Model Fitness)	RMSE (Last 3 Periods)	MAEP (%) (Last 3 Periods)
1-Period Ahead Forecasts					
1	Single ES	$\alpha = 1.0$	11.01	8.71	3.02
2	Double ES	$\alpha = 1.0, \gamma = 0.342$	6.35	5.94	2.28
3	Triple ES	$\alpha = 0.638, \gamma = 0.388, \delta = 1.0$	6.03	4.31	1.5
2-Periods Ahead Forecasts					
1	Single ES	$\alpha = 1.0$	20.91	19.89	7.48
2	Double ES	$\alpha = 0.5, \gamma = 0.64$	10.93	9.03	3.07
3	Triple ES	$\alpha = 0.4, \gamma = 0.335, \delta = 1.0$	10.61	7.05	2.76

- ***Actual Time Series Data***

The series represents the closing share price of an energy organisation in the National Stock Exchange (www.nseindia.com), India. Share prices of 20 consecutive working days (in 15-03-2011 to 11-04-2011) are considered. Table 3 shows the prices on different days, which are also shown in a plot in Fig. 1, for better visualisation. Here also we take $p = 4$ (as a trial value) for triple ES method, although no seasonality is apparent. We may note that, for $p = \infty$, with the initialisation as here, triple ES method would be identical to double ES method. The results are given in Table 4. Single ES method, overall, gives the most satisfactory result. But, performance of triple ES method is comparable.

Table 3: Share Prices of a Company

Obs. No.	Value	Obs. No.	Value
1	304.8	11	311.3
2	304.15	12	314.5
3	310.65	13	310.3
4	303.55	14	316.7
5	297.25	15	318.95
6	299.25	16	320.1
7	306.25	17	321.7
8	308.15	18	318
9	309.85	19	307.85
10	310.1	20	313.5

Figure 1: Plot of the Share Prices

Table 4: Forecasting Accuracy for Share Prices Data

Obs. No.	Method	Optimal Parameter Values	RMSE (Model Fitness)	RMSE (Last 3 Periods)	MAEP (%) (Last 3 Periods)
1-Period Ahead Forecasts					
1	Single ES	$\alpha = 0.969$	4.12	7	2.06
2	Double ES	$\alpha = 0.935, \gamma = 0.018$	4.12	7.12	2.08
3	Triple ES	$\alpha = 0.0, \gamma = 0.008, \ \delta = 0.461$	4.05	7.2	2.11
2-Periods Ahead Forecasts					
1	Single ES	$\alpha = 0.969$	5.93	8.48	2.2
2	Double ES	$\alpha = 0.24, \gamma = 0.395$	5.44	13	3.85
3	Triple ES	$\alpha = 0.0, \gamma = 0.0016, \ \delta = 1.0$	5.85	8.61	2.21

Concluding Remarks

The three exponential smoothing forecasting methods, namely, single ES, double ES and triple ES methods have been compared with additive trend and seasonality. These are also compared with last-value forecasting, often resorted to by practitioners. Two test data sets are used, and for each case one-period and two-periods ahead forecasts are noted. For the first set, a theoretical model, which has a strong trend component, is used. As anticipated, single ES method gives last-value forecasting for the data. Use of double and triple ES method improves forecasting accuracy, particularly for two-periods ahead forecasts. The methods are also applied for an actual time series, share prices of an organisation. Single ES method again is near to last-value forecasting, but shows the most satisfactory results. Performance of triple ES method, although, is comparable.

From the observations in the numerical experiments, we may write that, single ES method may give good results sometimes, but the performance is not consistent. Last-value forecasting also should be avoided. Among the methods discussed, the best choice in most of the cases would be triple exponential smoothing method. If there is a marked seasonality pattern, period of which has been identified relatively precisely, and forecasts are for short terms, this method can give a forecasting accuracy adequate for practical purposes. At the same time, other methods may also be explored.

References:

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., (1994), Time Series Analysis: Forecasting & Control, Prentice-Hall, Upper Saddle River, NJ.

Brown, R. G. (1959), Statistical Forecasting for Inventory Control, McGraw Hill, New York.

Fildes, R., Nikolopoulus, K., Crone, S. F. and Syntetos, A. A. (2008), "Forecasting and operational research: a review," Journal of the Operational Research Society, 59, pp. 1150-1172.

Gardner, E. S. (2006), "Exponential smoothing: the state of the art – Part II," International Journal of Forecasting, 22, pp. 637-666.

Gardner, E. S. (1985), "Exponential smoothing: the state of the art," Journal of Forecasting, 4, pp. 1-28.

Hillier, F.S. and Hillier, M. S. (2005), Introduction to Management Science: A Modeling and Case Studies Approach with Spreadsheets (TMH ed.), Tata McGraw-Hill, New Delhi.

Holt, C. C. (1957), "Forecasting Trend and Seasonal by Exponentially Weighted Averages," ONR Memorandum No. 52, Carnegie Institute of Technology, Pittsburgh, USA (reprinted in Int. J. of Forecasting, 2004, 20, pp. 5-13)

Makridakis, S. and Hibon, M. (2000), "The M-3 Competitions: Results, conclusions and implications," International Journal of Forecasting, 16, pp. 451-476.

Makridakis, S., Andersen, A., Carbone, R., Filders, R., Hibon, M. and Lewansowski, R. (1982), "The accuracy of exploration (time series) methods: Results of a forecasting competition," Journal of Forecasting, 1, pp. 111-153.

Makridakis, S. and Hibon, M. (1979), "Accuracy of forecasting: an empirical investigation (with discussion)," Journal of the Royal Statistical Society A, 142, pp. 97-145. ■