FORECASTING WITH EXPONENTIAL SMOOTHING: SOME GUIDELINES FOR MODEL SELECTION

Everette S. Gardner, Jr., *U.S. Navy*
David G. Dannenbring, *Columbia University*

ABSTRACT

Despite the general acceptance of exponential smoothing, the choice of a specific smoothing model is often a difficult problem. Previous research involving smoothing-model comparisons and the penalties for selection of the wrong model has been limited. This paper evaluates the performance of a representative group of smoothing models over a variety of conditions in 9,000 simulated time series. Forecast-error results demonstrate that a major disadvantage of adaptive smoothing models is their tendency to generate unstable forecasts, even during periods when mean demand itself is stable. Several trend-adjusted smoothing models are shown to be robust forecasters, whether the time series actually display a trend or not.

Subject Areas: Forecasting and Production/Operations Management.

INTRODUCTION

Since the early research by Brown [2] [3], Holt [12], and Winters [20], forecasting models using exponential smoothing have gained widespread application in industry. Exponential smoothing is simple and inexpensive, and there is no evidence that the more complex and expensive forecasting models, such as Box-Jenkins, consistently provide better short-range forecast accuracy (see [9], for example, and the review in [1, ch. 7]).

Despite the general acceptance of exponential smoothing, the choice of a specific smoothing model is often a difficult problem. As shown by the model classifications in Figure 1, the decision maker must choose between models with fixed smoothing constants and a class of adaptive models that vary smoothing constants to shorten the response lag during periods of shifts in mean demand. Each type of model can also be enriched with either fixed or adaptive trend adjustments.

Unfortunately, in large-scale applications of exponential smoothing, it may not be feasible to make extensive model comparisons before selecting a specific model. In most time series, it is certainly difficult to determine whether incipient trends will actually develop or whether established trends will continue. Another problem is that it may be impossible to predict whether shifts in mean demand or trends will occur rapidly enough to require the use of adaptive smoothing models. Previous research has not investigated the costs or forecast-error penalties associated with selection of the wrong model. In particular, there appears to be a presumption in the literature that adaptive smoothing models are cost free (see [1] and [5], for example).

FIGURE 1
Classification of Exponential Smoothing Models

Trend Adjustment:	Smoothing Constant(s): Fixed	Smoothing Constant(s): Adaptive
None	1. Simple Smoothing [3]	4. Trigg and Leach [18] 5. Whybark [19]
Linear	2. Holt [12] 3. Double Smoothing [3]	6. Gilchrist [8] 7. Montgomery [15] 8. Roberts and Reed [17] 9. Chow [4]

The purpose of this paper is to compare the performance of the nine smoothing models in Figure 1 in order to establish some guidelines for model selection. On the basis of tests using 9,000 simulated time series, the following conclusions are offered. First, it is shown that adaptive smoothing models have a pronounced tendency to generate unstable forecasts. The advantage of the adaptive models' ability to react to sudden shifts in mean demand was offset by their tendency to overreact to purely random fluctuations in demand. Second, the nonadaptive, trend-adjusted models are shown to be robust forecasters under a variety of conditions. On time series with a constant mean and no trend, they duplicate the performance of the simple smoothing model. With appropriate smoothing constants, they can also do as well as the adaptive models in reacting to sudden shifts in demand.

MODELS TESTED

Although the models tested in this research are not a complete set of all exponential-smoothing models that have been proposed, they are representative of those most frequently encountered. The simple-smoothing, Holt, and double-smoothing models were included since they appear to be widely used in practice. The Whybark and the Trigg and Leach models were the two best performers in Whybark's comparison of adaptive models [19]. The four adaptive models with linear trend adjustments are based on interesting concepts that deserve more research to determine their suitability for practical applications. Other models were excluded on the grounds that they deal with special cases or with seasonal demand. In our opinion, exponential smoothing with seasonal demand deserves a separate research effort because of the problems inherent in separating the seasonal and trend components in many time series, the variety of types and strengths of seasonal patterns that could be encountered, and the numerous models that have been proposed [1] [2] [3] [8] [9] [12] [20]. The conclusions of the present research do, however, apply to seasonally adjusted time series.

The more complex models tested in this research are enrichments of the simple exponential smoothing model [3], a weighted-moving-average technique that gives unbiased forecasts if the time series has a constant mean. The model is repeated below for reference:

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

where S_t = estimate of the mean at period t and forecast for $t+1$,
X_t = actual demand in t, and
α = smoothing constant, with $0 \leq \alpha \leq 1$.

For time series that display trend, the forecasts from the simple model will respond to the trend, but with a lag. The Holt model [12] uses two fixed smoothing constants to estimate the mean and slope, respectively, of the series at period t, correcting for the lag. Equations (2), (3), and (4) below are followed in sequence each period to generate one-period-ahead forecasts:

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + R_{t-1}) \tag{2}$$

$$R_t = \beta(S_t - S_{t-1}) + (1-\beta)R_{t-1} \tag{3}$$

$$F_{t+1} = S_t + R_t \tag{4}$$

where S_t is again an estimate of the mean of the series at period t, R_t is an estimate of the slope, and β is a separate smoothing constant for the apparent trend each period, $S_t - S_{t-1}$; F_{t+1} adds the two estimates to obtain the forecast for $t+1$.

Another approach to forecasting a linear trend is the double-smoothing model [3], which can be shown to be a special case of both the Holt model and the method of discounted least squares [8]. Double smoothing requires only one fixed smoothing constant and begins with equation (1) to generate the statistic S_t. Next, S_t itself is smoothed to obtain $S_t^{(2)}$:

$$S_t^{(2)} = \alpha S_t + (1-\alpha)S_{t-1}^{(2)} . \tag{5}$$

The forecast is then

$$F_{t+1} = [2S_t - S_t^{(2)}] + [\frac{\alpha}{1-\alpha}(S_t - S_t^{(2)})]. \tag{6}$$

The first term in (6) is an estimate of the mean, and the second is an estimate of the slope.

The two adaptive models without trend adjustments reflect different heuristic approaches to the problem of shortening the response lag of the simple model when rapid shifts in mean demand occur. The Trigg and Leach model [18] sets, in each period, the value of α in equation (1) equal to the absolute value of a tracking signal:

$$\text{Tracking Signal} = \frac{\text{Smoothed Forecast Error}}{\text{Smoothed Absolute Forecast Error}}. \tag{7}$$

The two error values are smoothed with equations similar to (1). If the model is in control, the value of (7), and hence α, will be small. If biased errors occur, α will become larger, approaching 1 as a limiting case.

The Whybark model [19] also involves continuous evaluation of forecast errors, but this model changes the smoothing constant only when errors exceed specified control limits. If the forecast error in a single period exceeds ± 4 standard deviations, or if two consecutive errors exceed ± 1.2 standard deviations, then α is increased from a base of .2 to .8 for one period, reduced to .4 for the next period, and then reset to .2.

With both the mean and slope of the time series subject to rapid shifts, four different approaches that have been suggested to the problem were tested. First, Gilchrist [8] argues that adaptive control of both smoothing constants in the Holt model can lead to unstable forecasts. He recommends adaptive control of the smoothing constant for the mean only, perhaps using the Trigg and Leach approach, with a fixed smoothing constant for the slope. Hence the Gilchrist model was formulated by using the tracking signal in (7) to generate α for the Holt model in equation (2). Except for the adaptive α, the Gilchrist model is identical to the Holt model.

Chow [4] has proposed a general methodology for adaptive control of models with a single smoothing constant, such as the double-smoothing model. In Chow's methodology, we compute three forecasts each period for the double-smoothing model. The first forecast is generated using a base value of α, and the other using $\alpha_H = \alpha + .05$ and $\alpha_L = \alpha - .05$. If the smoothed mean absolute deviation (MAD) of forecast errors using α is less than the MADs for α_H and α_L, no change is made. If the MAD from α_H or α_L is lower, α is set to α_H or α_L. All MADs are then reset to zero, new values of α_H and α_L are computed, and the process begins anew.

Chow's methodology was extended by Roberts and Reed [17], and later by Montgomery [15], to include adaptive control of two or more smoothing constants in the same model, in this case the Holt model. Both the Roberts and Reed and the Montgomery models are special cases of orthogonal first-order experimental designs. The Roberts and Reed model, also referred to as SAFT [17], is a two-level factorial design that is initialized with a base combination of values for α and β, as well as high and low values ($\pm .05$) around each base parameter. Each period, five forecasts are computed—one for the base conbination and one for each of the four possible combinations of high and low values. The base combination and high and low values are shifted together by fixed amounts (.05 is the usual amount and was used in this research) whenever one of the forecasts for the high and low values is significantly better than the base forecast.

The Montgomery model exploits the simplex experimental design technique [15] to reduce the number of combinations of smoothing parameters evaluated each period from five to three. The design matrix, D, containing the smoothing parameter values is defined as a 3×2 matrix for the Holt model:

$$D = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \end{bmatrix} \qquad (8)$$

Each row of D is used to compute one forecast each period. The smoothing parameters in D are adjusted according to the following rules:

1. Denote by c_i the current absolute forecast error for row i in D. Denote the *jth* row of D vectorially by d_j. Let the maximum value of c_i occur for row d_j'. Form a new matrix by deleting d_j' from D and substituting the new row:

$$d_j'^* = (d_1 + d_2 + d_3) - 2d_j' \qquad (9)$$

   Calculate the forecast for the next period using the smoothing parameters that are elements of the new row $d_j'^*$.
2. Apply rule 1 unless a row has occurred in three successive matrixes without being eliminated. Should this situation arise for the *ith* row, discard c_i and calculate the forecast for the next period using the smoothing parameters in d_i. Then apply rule 1.
3. Should c_i be the maximum absolute current error in the *nth* matrix and c_i^* be the maximum absolute current error in the (n+1)*st* matrix, do not return to the *nth* matrix. Instead, move from the (n+1)*st* matrix by discarding the second-largest absolute current error. This rule is designed to prevent oscillation.

Montgomery [15] reports some limited comparisons of the Roberts and Reed model to his model. When the signal-to-noise ratio of the time series is relatively small, the Roberts and Reed model seems to yield better forecasts. The Montgomery model, however, appears to possess superior trend-following ability. Both models can easily be extended to control three smoothing parameters, such as those used for seasonal demand in the Winters model [20].

RESEARCH DESIGN

Most previous research has dealt with empirical data [3] [4] [5] [7] or with a small number of simulated time series [15] [16] [17] [18] [19]. Comparisons of models with fixed smoothing constants to adaptive models have generally been limited to the comparison of a simple smoothing model with an α value of .1 or .2 versus several adaptive models [4] [5] [15] [16]]17] [19]. All time series used in this research were simulated to control the characteristics of the series. All nine models described were tested on all series, and a wide range of smoothing constants was enumerated for the models with fixed smoothing constants.

Demand was generated by the following model:

$$X_t = a_t + bt + e_t \qquad (10)$$

where X_t = demand for period t,
a_t = level or mean of the series at period t,
b = slope, and
e_t = independent observations on a population with a normal distribution, a mean of zero, and a constant variance of σ^2.

This model was used to generate 9,000 time series, each composed of seventy-five observations, with the general characteristics shown in Table 1. The time series can be divided into six categories. The time series in the first category had a constant level (a_t) and no trend ($b = 0$). The group of 300 includes fifty replications for each of six coefficients of variation (σ/a_0): .01, .025, .05, .10, .20, and .30. The second group is similar to the first, except that the level of the series was increased by 25 percent at period 51.

In each of the next four groups, 2,100 time series were necessary to obtain fifty replications of all forty-two combinations of the six coefficients of variation above with the following ratios of trend per period as a fraction of the level at time zero (b/a_0): .005, .01, .02, .03, .04, .05, and .10. The coefficients of variation specified above hold at time zero only. The true coefficients of variation become smaller each period as the trend is increased. Both the level and trend were held constant in group three. The level was increased by 25 percent at period fifty-one in group four while trend was held constant. Although it might be more realistic to expect trend to change slowly over a period of time, the adaptive models are designed to react to sudden changes in trend, which was investigated in groups five and six. In group five, the level was held constant while the trend was increased by 25 percent at period fifty-one. Finally, the level was held constant in the last group while the trend changed sign in period fifty-one.

A variety of forecast-error information was collected for each model, including the mean and variance of MAD, mean squared error (MSE), mean absolute error, mean absolute percentage error, mean percentage error (bias), and the number of tracking signal trips (the number of times each model's forecast error exceeded 95 percent probability limits). The first twenty-five periods of each time series were used for initialization of all model parameters except smoothing constants, which are discussed below. All models were started in period one with perfect knowledge of the mean at time zero, but with the assumption that the trend was always zero regardless of the time series. After period twenty-five, all forecast-error information was reset to zero to wash out the effects of initial conditions, and forecast-error information was then collected for the next fifty periods. This procedure also allowed twenty-five periods for the trend-adjusted models to develop an estimate of the trend. A complete set of runs was made in which the trend-adjusted models were started with perfect knowledge of the initial trend, but the forecast-error results were about the same as were obtained with the twenty-five-period initialization.

Selection of smoothing constants for the nonadaptive models presented a difficult problem in research design. Several alternatives were considered. One common approach in forecasting research is to use the first portion of each time the table. As discussed below, the Holt model had nearly the same mean forecast error as the simple-smoothing and Whybark models for the time series in these two categories. Also, all the Whybark model successes were with coefficients of variation of .01 and .025, which are perhaps unrealistically small.

Another perspective on the results is given by the MSE indices in Tables 2 and 3. To obtain this MSE index, the MSE of each model is divided by the MSE of the expected-value model; this value is then multiplied by 100. Thus, an MSE index value of 100 means that the MSE of the model in question is the same as the MSE of the expected-value model. Each MSE-index value in the tables represents the mean of fifty replications of time series with the coefficient of variation and trend characteristics as shown. Tests for statistical significance were made using standard analysis of variance procedures for differences among means. For each category of time series (i.e., row in the table), MSE-index values designated by asterisks are significantly smaller than those without asterisks in the same row at the .01 probability level. Among each group of values with asterisks there is no significant difference. Other comments about statistical significance in the discussion below also refer to the .01 probability level.

In Table 2, under the conditions of a constant mean and no trend, there are two important observations. First, there is a significant penalty in MSE for the use of adaptive smoothing models. All models with fixed smoothing constants gave significantly smaller MSE's than the adaptive models, which tended to overreact to purely random fluctuations in demand and gave unstable forecasts. Second, there is little or no penalty for the use of models with trend adjustments on time series without trend. The Holt model, for example, matched the performance of both the simple-smoothing and the expected-value model. Double smoothing was only marginally worse.

The relative standing of the models on the other measures of forecasting performance (not reported herein) was about the same except for bias and the number of tracking signal trips. None of the models showed any significant bias, and all models averaged less than one trip per time series. The relative performance of the models was not sensitive to changes in the coefficient of variation of the time series. As expected, the best smoothing constants were all low, in the .01 to .10 range, including the β parameter in the Holt model.

Table 2 shows that the models with fixed smoothing constants can do as well or better than the adaptive models when there is a sudden step increase in mean. Except for the Roberts and Reed model, there is no significant difference in MSE among all models with the step increase in mean. The instability of the adaptive models before and after the increase in mean offsets their advantage in reacting to the increase.

The smoothing-constant values used by the nonadaptive models to follow the increase in mean were moderate. The α values for the simple-smoothing and Holt models were in the .20 to .40 range about 80 percent of the time with the lower coefficient of variation. At the higher coefficient, they were in the .10 to .30 range about 90 percent of the time. The Holt β value was always .01 or .05. The double-smoothing model used α values of .10 or .20 about 95 percent of the

TABLE 2
Comparative MSE Indices for Simulated Time Series Without Trend†

			Nonadaptive Models			Adaptive Models					
			No Trend	With Trend		No Trend		With Trend			
Type of Series	Coefficient of Variation	Expected Value Model	Simple Smoothing	Holt	Double Smoothing	Trigg and Leach	Whybark	Gilchrist	Montgomery	Roberts and Reed	Chow
Constant Mean	.10	100	100*	100*	101*	121	121	122	106	125	114
	.20	100	100*	100*	101*	122	123	124	106	130	115
25% Step Increase in Mean at Period 51	.10	100	145*	145*	151*	151*	146*	152*	154*	180	151*
	.20	100	119*	119*	121*	130*	130*	131*	127*	145	124*

†Values shown: (MSE Model/MSE E-V Model)(100).
*Significantly lower (.01 probability level) MSE than for other models for stated type of series and coefficient of variation.

TABLE 3
Comparative MSE Indices for Simulated Time Series With Linear Trend†

			Nonadaptive Models			Adaptive Models					
			No Trend	With Trend		No Trend		With Trend			
Type of Series	Slope/ Intercept	Expected Value Model	Simple Smoothing	Holt	Double Smoothing	Trigg and Leach	Whybark	Gilchrist	Montgomery	Roberts and Reed	Chow
Constant Slope and Intercept	.01	100	117*	105*	111*	122*	126	123	128	134	114*
	.05	100	157	110*	115*	151	171	125	129	140	115*
Constant Slope. 25% Increase in Intercept at Period 51	.03	100	143	117*	122*	136	148	128	130	147	122*
Constant Intercept. 25% Increase in Slope at Period 51	.03	100	153	128*	132*	146	159	137*	146	155	132*
Constant Intercept. Change in Sign of Slope at Period 51	.03	100	135*	128*	135*	134*	142	134*	152	148	134*

†Values shown: (MSE Model/MSE E-V Model)(100); σ/Intercept = .20 for all series.
*Significantly lower (.01 probability level) MSE than for other models for stated type of series and ratio of slope to intercept.

time at the lower coefficient of variation and always used values of .05 or .10 at the higher.

Despite the increase in mean, the bias of most models was small at the lower coefficient of variation, less than 1 percent for all models except the Montgomery model. Curiously, the Montgomery model was biased high at + 1.6 percent. At the higher coefficient, the nonadaptive models displayed significantly less bias than the adaptive models. The simple-smoothing, Holt, and double-smoothing models all had bias values around -1.5 percent. The Trigg and Leach and the Whybark models were biased at -3.3 percent and -2.7 percent, respectively. The Chow model was biased at -2.9 percent, the Roberts and Reed model at -3.7 percent, and the Gilchrist model at -3.4 percent. The adaptive models were often unable to sort out the increase in mean from the purely random fluctuations in demand, which caused their forecasts to lag the increase.

There was a slight advantage to the adaptive models in minimizing the number of tracking signal trips. The Trigg and Leach, Gilchrist, and Whybark models averaged about .25 trips per series, or about 12.5 trips per fifty replications of each coefficient of variation. The other models were all significantly larger than this, although still in the range of .60 to 1.0 trips per time series.

Table 3 gives forecast errors in terms of the MSE index (as defined above) for time series with linear trend. The adaptive models, with the exception of the Chow model, again gave unstable forecasts. This instability was magnified when the step changes in slope and intercept were introduced. Although only one ratio of σ to intercept is given, the relative positions of the models in MSE were about the same under the other ratios tested as well. Relative model performances also were not sensitive to the ratio of slope to intercept.

The best smoothing constant for the simple-smoothing model was always .50 or .75 to keep up with the trends. The Holt α value was usually in the .10 to .20 range, except when the step change in mean was introduced, where .30 was most frequently used. The Holt β value was usually low, in the .05 to .10 range. Double-smoothing α values ranged from .10 to .20.

No useful generalizations can be made about the performance of the models on bias or tracking signal trips. The relative performance of all models varied with the coefficient of variation and trend, although not in a consistent way. The adaptive models definitely showed no significant advantage in reducing either bias or the number of tracking signal trips in time series with linear trend.

CONCLUSIONS

For time series that display a reasonably stable mean, with no apparent trend, the smoothing constant for the simple-smoothing model should be maintained in the .01 to .10 range. Higher values of the smoothing constant could be used to hedge against the development of trends. A better alternative, however, is to employ routinely one of the trend-adjusted models, since there is no apparent penalty in forecast error if there is no trend. In cases of doubt about trend, a safe course for the Holt model is to maintain both the α and β values in the .01 to .10

range. With a definite trend, the recommended Holt α range is .10 to .20, with β set at .05 or .10. For double smoothing when there is doubt about trend, α should be set at .05 or .10. Double smoothing with a definite trend should employ a smoothing constant in the .10 to .20 range, the same values recommended by Brown [3].

The best single-smoothing constant value in the ranges discussed above depends on the stability of the time series for any model. The lower values should be employed with the more unstable time series in order to filter out random fluctuations in demand before generating forecasts.

The choice between the Holt and double-smoothing models is equivocal. Although there was rarely any significant difference between the two models, the Holt model was always by some margin better than the double-smoothing model. This result is to be expected since the two models are quite similar, with the exception that the Holt model has the advantage of an additional parameter value to forecast trend. The criterion of parsimony would favor double smoothing, however, since there is no need to bother with selection of a separate smoothing constant for trend. Another advantage of the double-smoothing model is its preference for a few α values, which considerably simplifies the problem of selecting the best one.

For most time series in which some rapid shift in mean demand (and/or trend) is predicted, we recommend against the use of adaptive smoothing models. It is always difficult to predict that a rapid shift will occur. If the prediction is wrong, the adaptive models will tend to mistake purely random fluctuations in demand for shifts in the true mean demand. The result will be unstable forecasts and forecast errors significantly higher than those of the nonadaptive models. This conclusion contradicts previous research, particularly [1] and [5].

If the decision maker's prediction that a rapid shift in demand will occur is correct, the forecast-error penalty must be paid during periods of demand stability both before and after the shift occurs. Another consideration is that the double-smoothing model offers a simple and very effective alternative to hedge against shifts in demand. With the smoothing constant set in the .05 to .20 range, depending on the coefficient of variation of the time series, the double-smoothing model is an extremely robust forecaster. The double-smoothing forecast errors recorded in this research were never significantly higher than the adaptive models, and they were usually significantly lower.

We cannot entirely rule out adaptive smoothing models on the basis of this research. Especially perverse time series, such as the sequence of frequent, rapid up and down shifts in mean demand tested by Whybark [19], were not investigated. We do believe, however, that our conclusions will assist in the selection of the proper smoothing model and parameters for time series likely to be encountered in practice. [Received: June 11, 1979. Accepted: December 12, 1979.]

REFERENCES

[1] Armstrong, J. S. *Long-Range Forecasting.* New York: John Wiley and Sons, 1978.
[2] Brown, R. G. *Statistical Forecasting for Inventory Control.* New York: McGraw-Hill, 1959.

[3] Brown, R. G. *Smoothing, Forecasting, and Prediction of Discrete Time Series*. Englewood Cliffs, N.J.: Prentice-Hall, 1962.

[4] Chow, W. M. "Adaptive Control of the Exponential Smoothing Constant." *The Journal of Industrial Engineering*, Vol. 16, No. 5 (1965), pp. 314-317.

[5] Dancer, R. E., and C. F. Gray. "An Empirical Evaluation of Constant and Adaptive Computer Forecasting Models for Inventory Control." *Decision Sciences*, Vol. 8, No. 1 (1977), pp. 228-238.

[6] Dennis, J. D. "A Performance Test of a Run-Based Adaptive Exponential Forecasting Technique." *Production and Inventory Management*, Vol. 19, No. 2 (1978), pp. 43-46.

[7] Eilon, S., and J. Elmaleh. "Adaptive Limits in Inventory Control." *Management Science*, Vol. 16, No. 8 (1970), pp. B-533-548.

[8] Gilchrist, W. *Statistical Forecasting*. London: John Wiley and Sons, 1976.

[9] Groff, G. K. "Empirical Comparison of Models for Short Range Forecasting." *Management Science*, Vol. 20, No. 1 (1973), pp. 22-31.

[10] Groff, G. K., and J. F. Muth. *Operations Management: Analysis for Decisions*. Homewood, Ill.: Richard D. Irwin, Inc., 1972.

[11] Gross, D., and J. L. Ray. "A General Purpose Forecast Simulator." *Management Science*, Vol. 11, No. 6 (1965), pp. B119-B135.

[12] Holt, C. C. *Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages*. Pittsburgh, Penn.: Carnegie Institute of Technology, 1957.

[13] McClain, J. O. "Dynamics of Exponential Smoothing with Trend and Seasonal Terms." *Management Science*, Vol. 20, No. 9 (1974), pp. 1300-1304.

[14] McClain, J. O., and L. J. Thomas. "Response-Variance Tradeoffs in Adaptive Forecasting." *Operations Research*, Vol. 21, No. 2 (1973), pp. 554-568.

[15] Montgomery, D. C. "Adaptive Control of Exponential Smoothing Parameters by Evolutionary Operation." *AIIE Transactions*, Vol. 2, No. 3 (1970), pp. 268-269.

[16] Raine, J. E. "Self-Adaptive Forecasting Reconsidered." *Decision Sciences*, Vol. 2, No. 2 (1971), pp. 181-191.

[17] Roberts, S. D., and R. Reed. "The Development of a Self-Adaptive Forecasting Technique." *AIIE Transactions*, Vol. 1, No. 4 (1969), pp. 314-322.

[18] Trigg, D. W., and A. G. Leach. "Exponential Smoothing with an Adaptive Response Rate." *Operations Research Quarterly*, Vol. 18, No. 1 (1967), pp. 53-59.

[19] Whybark, D. C. "A Comparison of Adaptive Forecasting Techniques." *The Logistics and Transportation Review*, Vol. 9, No. 1 (1973), pp. 13-26.

[20] Winters, P. R. "Forecasting Sales by Exponentially Weighted Moving Averages." *Management Science*, Vol. 6, No. 3 (1960), pp. 324-342.