Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Forecasting Sales by Exponentially Weighted Moving Averages

Peter R. Winters,

To cite this article:
Peter R. Winters, (1960) Forecasting Sales by Exponentially Weighted Moving Averages. Management Science 6(3):324-342.
https://doi.org/10.1287/mnsc.6.3.324

FORECASTING SALES BY EXPONENTIALLY WEIGHTED MOVING AVERAGES*†

PETER R. WINTERS

Graduate School of Industrial Administration, Carnegie Institute of Technology

The growing use of computers for mechanized inventory control and production planning has brought with it the need for explicit forecasts of sales and usage for individual products and materials. These forecasts must be made on a routine basis for thousands of products, so that they must be made quickly, and, both in terms of computing time and information storage, cheaply; they should be responsive to changing conditions. The paper presents a method of forecasting sales which has these desirable characteristics, and which in terms of ability to forecast compares favorably with other, more traditional methods. Several models of the exponential forecasting system are presented, along with several examples of application.

1. Introduction

Many forecasts of the future are made by people in the course of running a business, although in a large number of cases the forecasts are not called that, and the methods of making the forecasts are not clearly known even to the people making them. Forecasts are made for a variety of purposes, for example, for cash budgeting, establishing sales quotas, setting expense budgets, planning capital expenditures, and for production and inventory planning and control. The methods used to make the forecasts differ widely, as one would expect them to, because their uses differ widely. The kind of forecast requirement that this paper is particularly concerned with is the last mentioned above: forecasting sales of individual products for inventory control and production scheduling.

The need for forecasts of individual product (or item) sales most frequently arises because of an inventory control system, or a production scheduling system, consisting of decision rules which specify when to produce or order more of a particular item (triggers or order points) and how much to produce or order (lot sizes or order quantities). These decision rules are based in part on a prediction of sales or usage of each item in the near future. The rules are applied on a routine basis to many products, often tens of thousands, or even hundreds of thousands of them. Forecasts must be made frequently, monthly or weekly, on a routine basis.

* Received June 1959.

† This paper will appear in substantially the same form as a chapter in the forthcoming book: *Planning Production, Inventories and Employment*, by Charles C. Holt, John F. Muth, Franco Modigliani and Herbert A. Simon; Prentice-Hall, publishers. Earlier models of the exponential forecasting system appeared in ONR Research Memorandum No. 52, "Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages", by Charles C. Holt, Carnegie Institute of Technology, April, 1957. William Gere assisted with the numerical calculations. This paper was written as part of the project "Planning and Control of Industrial Operations", sponsored by the Office of Naval Research, Contract No. ONR 27 T.O. 1, at Carnegie.

There are certain desirable characteristics of these forecasts that are implied by their use: they must be made quickly, cheaply, and easily; the forecasting technique must be clearly spelled out, so that it can be followed routinely, either manually or using an electronic computer. The number of pieces of information required to make a single forecast must be kept at a minimum, or else the total amount of information required for all products will be expensive to store and expensive to maintain. Finally, partly implicit in the above, is the need to be able to introduce the latest sales information easily and cheaply.

Quite a few forecasting techniques, or systems, are available or could be developed for predicting item sales. The ones discussed in this paper do not "predict" with a behavioral model of sales, but use an analysis of the sales time-series taken out of context. That is, the only input to the forecasting system is the past history of sales of the item, and not, for example, such information as what is happening in the market, the industry, the economy, sales of competing and complimentary products, price changes, advertising campaigns, and so on. A behavioral prediction model would have to be developed for each application, and would probably lack most of the desirable characteristics of a forecasting system that are listed above. Various forms of the exponential model have been used for a wide variety of forecasting applications with little modification.[1]

In its simplest form, the exponential system makes a forecast of expected sales in the next period by a weighted average of sales in the current period, and the forecast of sales for the current period made during the previous period. In the same way, the forecast for the current period was a weighted average of sales during the previous period and the forecast of sales for that period made in the period before. This same process continues back to the beginning of the sales data for the item. Thus a prediction made in any period is based on current sales data, and all the previous sales data for the item, but in such a way that only one number (the most recent estimate for the current period) must be retained to be combined with the latest incoming sales information.

This scheme obviously has the characteristics desired in a prediction method: current sales information is easily introduced, the forecast calculation is fast, and only a limited amount of information must be kept and maintained. For some products with stable sales rates and little seasonal influence, this simple exponential model proves quite satisfactory. Many products, however, have a marked trend in their sales, particularly when they are first introduced, or when competing products are introduced. And for many products there is a substantial seasonal pattern. It is usually worthwhile extending the exponential system to take into account long-run trends and seasonal effects. These two factors are handled in exactly the same way as the simple exponential system. More information is required with this more complete model, but the accuracy of prediction is also substantially increased for most kinds of products.

[1] For earlier types of exponential models see, for example, J. F. Magee, *Production Planning and Inventory Control*, McGraw-Hill, 1958. A more complete version of this model is given in *Statistical Forecasting*, by R. G. Brown, printed by Arthur D. Little, Inc., 1958.

Three forecasting models are used in a similar way in the paper to predict sales for three different time series. One of these is the (complete) exponential model; the other two are a naive model and a simple forecasting model of the more usual type. A comparison of "ability to predict" will be presented later in the paper to give a basis for evaluating the exponential model. The results of these comparisons show that the exponential system makes more accurate forecasts, requires less information storage, and requires slightly more time to compute, than the better of the two conventional models.

2. Development of the Exponential System

A number of variations of the exponential weighting method can be used. In this section we will present a sample of the possible variants, hopefully those which help demonstrate the method best and which are most generally useful. We will indicate other variants which might have more limited application.

2.1 *The Simplest Exponential Model*

The simplest application of an exponentially weighted moving average would be to the problem of making a forecast of the expected value of a stochastic variable whose mean (or expected value) does not change between successive drawings. This would correspond to predicting the expected sales for a product which had no definite seasonal pattern and no long-run trend. The following procedure is proposed: take a weighted average of *all* past observations and use this as a forecast of the present mean of the distribution, as

$$\bar{S}_t = AS_t + (1 - A)\bar{S}_{t-1} \tag{1}$$

where

S_t = actual sales during the t'th period
$\bar{S}_t$ = forecast of expected sales in the t'th period
$0 \leq A \leq 1$

Then

$$\bar{S}_{t-1} = AS_{t-1} + (1 - A)\bar{S}_{t-2},$$

so that

$$\bar{S}_t = AS_t + A(1 - A)S_{t-1} + (1 - A)^2\bar{S}_{t-2}. \tag{2}$$

Continuing this process, $\bar{S}_t$ can be expressed explicitly in terms of all the past observations of sales, that is, all the sales data available:

$$\bar{S}_t = A \sum_{n=0}^{M} (1 - A)^n S_{t-n} + (1 - A)^{M+1}\bar{S}_b \tag{3}$$

where $\bar{S}_b$ is the beginning value of $\bar{S}$. M is the number of observations in the series up to and including the current period, t. Even for relatively small A, if M is large enough, that is, if enough history is used, $(1 - A)^{M+1}$ becomes very small, and the last term can be ignored.

Since the process which generates the sales data is a stationary process, that is, there is no seasonal pattern and no trend, then $\bar{S}_t$ is an unbiased estimate of $E(S)$, the expected sales in any period:

$$E(\bar{S}_t) = E(S)A \sum_{n=0}^{M} (1 - A)^n + (1 - A)^{M+1}\bar{S}_b \tag{4}$$

As noted above for large M, and most A, $(1 - A)^{M+1}\bar{S}_b$ approaches zero. Under these same conditions $A \sum_{n=0}^{M} (1 - A)^n$ approaches one. Thus

$$E(\bar{S}_t) \cong E(S)$$

with the degree of approximation depending on the values of M and A.

The statistical properties of $\bar{S}_t$ as an estimate of $E(S_t)$ will not be rigorously spelled out when the stochastic process is not stationary. But a weighted moving average with exponential weights is evidently a sensible mode of behavior for this simple forecasting problem. If the distribution mean is subject to a variety of short- and long-run systematic changes then the exponential model has some intuitive advantages. If the distribution mean changes slowly, then A should be small so as to keep the effect of older observations. If the distribution mean changes quickly, then A should be large so as to attenuate quickly the effect of older observations, but not too large, or else $\bar{S}_t$ will be subject to too much random variation. The problem of finding a satisfactory value of the weighting parameter, A, will not be solved in detail for this simple model. In a later section, however, we will discuss in substantial detail the effect on prediction accuracy of the set of weighting parameters for the more complete exponential model, which includes seasonal and trend effects.

An exploration of those time series for which the exponential system is optimal will not be considered.[2] A more relevant question here is whether this approach to forecasting holds promise for a variety of applications, compared with other methods. Before going to the complete forecasting model which includes both seasonal and trend, we will study a simpler version which includes only a seasonal effect.

2.2 *Forecasting With Ratio Seasonals*

It is possible to develop a forecasting model with either a multiplicative or an additive seasonal effect. If the amplitude of the seasonal pattern is independent of the level of sales, then an additive model is appropriate. More often, however, the amplitude of the seasonal pattern is proportional to the level of sales. This would indicate using the multiplicative, or ratio, seasonal effect. Figure 1 shows the sales for an individual product over a period of time. The actual sales in period t is given by S_t. The estimate of the smoothed and seasonally adjusted sales rate in period t is given by $\bar{S}_t$. The periodicity of the seasonal effect is L;

[2] For discussion of this subject see J. F. Muth, "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, forthcoming.

ITEM SALES: SEASONAL PATTERN

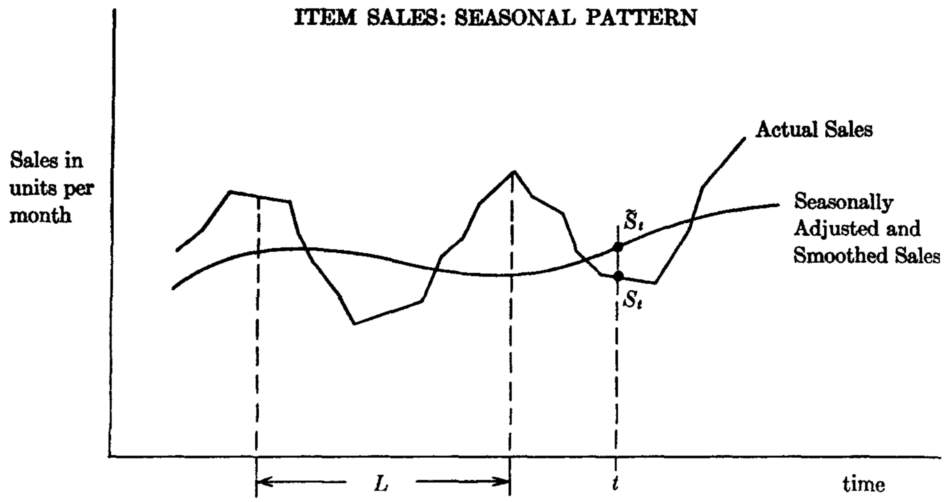


FIG. 1

if a period is a month, L would ordinarily be 12 months. The model is

$$(5) \qquad \tilde{S}_t = A \frac{S_t}{F_{t-L}} + (1 - A)\tilde{S}_{t-1}, \qquad 0 \leq A \leq 1$$

for the estimate of the expected deseasonalized sales rate in period t, and

$$(6) \qquad F_t = B \frac{S_t}{\tilde{S}_t} + (1 - B)F_{t-L}, \qquad 0 \leq B \leq 1$$

for the current estimate of the seasonal factor for period t. In equation 5, $\tilde{S}_t$ is a weighted sum of the current estimate obtained by deseasonalizing the current sales, S_t, and last period's estimate, $\tilde{S}_{t-1}$, of the smoothed and seasonally adjusted sales rate for the series. (Note that in deseasonalizing current sales by S_t/F_{t-L}, the most recent estimate of the seasonal effect for periods in this position in the cycle has been used; the seasonal factor computed for May last year would be used to seasonally adjust this year's May data.) The value of $\tilde{S}_t$ from Equation 5 is then used in forming a new estimate of the seasonal factor in Equation 6. This new estimate, F_t, is again a weighted sum of the current estimate, $S_t/\tilde{S}_t$, and the previous estimate, F_{t-L}. A *forecast* of the expected sales in the following period would then be made using the following:

$$(7) \qquad S_{t,1} = \tilde{S}_t F_{t-L+1}$$

where $S_{t,1}$ is the forecast made at the end of the current, or t'th period, for the following period. More generally, a forecast of expected sales T periods into the

[3] The forecasts can be readily extended beyond L periods in the future by *reusing* the L seasonal factors, $F_{t+1-L}, \cdots, F_t$.

future would be

$$S_{t,T} = \bar{S}_t F_{t-L+T}, \qquad T \leqq L^3 \tag{8}$$

The weighted averages in Equations 5 and 6 may be written in terms of past data and initial conditions:

$$\bar{S}_t = A \sum_{n=0}^{M} (1 - A)^n \frac{S_{t-n}}{F_{t-L-n}} + (1 - A)^{M+1} \bar{S}_b \tag{9}$$

and

$$F_t = B \sum_{n=0}^{J} (1 - B)^n \left(\frac{S_{t-nL}}{\bar{S}_{t-nL}}\right) + (1 - B)^{J+1} F_{bt} \tag{10}$$

where $\bar{S}_b$ is the initial value of $\bar{S}$, and F_{bt} is the initial value of F for the period in question. J is the largest integer less than or equal to M/L.

The forecast, then, is a function of all past observations of the variable, of the weights A and B, and of the initial conditions $\bar{S}_b$ and the set of F_{bt}, L in number. The effect of the initial conditions on the forecast depends on the size of the weights and the length of the series preceding the current period, t. The effect of $\bar{S}_b$ will be usually attenuated sooner than the effect of the initial F's, because $\bar{S}_t$ is revised every period, but the F's are revised only once per cycle.

If this forecasting model with a seasonal but no trend effect is applied to a sales series for which the mean is subject to long and short-run systematic changes, or trends, then the seasonal factors, the F's, will quickly cease to be simple seasonal factors, and will soon contain some of the trend effect. For example, in application to a series of monthly observations, with a long-run upward trend, the set of twelve F's will begin to sum to more than 12.0, and compensate for the lack of a trend factor in the model. If, on the other hand, the model is applied to a series which includes short-run trends, of shorter duration than L, then the trend effect that is incorporated in the seasonal factor, F, is made up of short-run trend effects from the years previous to the use of the factor. Consequently, erratic behavior may be introduced into the seasonal adjustment and smoothing of S_t, and into the forecast itself. If the series for which the method is intended does have trend effects, and many sales series seem to, then we must introduce a specific trend factor.

2.3 *Forecasting with Ratio Seasonals and Linear Trend*

As with the preceding section it is possible to develop a forecasting model with either a ratio trend, or an additive, or linear trend. Because of the combination of short and long-run systematic changes in expected sales, it is more generally useful to work with the latter case. The form of the model for this "complete" forecasting scheme is similar to that given in Equations 5 and 6. First,

$$\bar{S}_t = A \frac{S_t}{F_{t-L}} + (1 - A)(\bar{S}_{t-1} + R_{t-1}). \tag{11}$$

The only change in the definition of $\bar{S}_t$ is the addition of R_{t-1}, the most recent estimate of the additive trend factor, that is, the units per period that the expected sales rate, $\bar{S}_t$, is increasing (or decreasing). The expression for the revised estimate of the seasonal factor remains the same as it was in the previous section:

$$F_t = B\frac{S_t}{\bar{S}_t} + (1 - B)F_{t-L} \tag{12}$$

The expression for revising the estimate of the trend has the same form as Equations 11 and 12:

$$R_t = C(\bar{S}_t - \bar{S}_{t-1}) + (1 - C)R_{t-1} \tag{13}$$

weighting the estimate based on current data with the previous estimate. The forecast of sales T periods in the future would be obtained from the formula:

$$S_{t,T} = [\bar{S}_t + TR_t]F_{t-L+T} \qquad T = 1, 2, \cdots, L. \tag{14}$$

(Once again forecasts for future periods more distant than L can be made by reusing the appropriate F's.) Figure 2 illustrates the definition of variables.

In practice, the forecasting system would be used as follows to predict the sales of an individual product:

(1) At the end of the t'th (or current) period the actual sales of the product during the period, S_t, is recorded.
(2) Equation 11 is applied to evaluate $\bar{S}_t$, using $\bar{S}_{t-1}$ and A_{t-1} from the last period and the appropriate F_{t-L} computed during the previous cycle.
(3) Equation 12 is used to evaluate F_t, which can now replace F_{t-L}.
(4) Equation 13 is used to determine R_t, which can now replace R_{t-1}.
(5) Forecasts of future sales are made, using Equation 14.
(6) The value $\bar{S}_{t-1}$ is replaced by $\bar{S}_t$, and the data is ready for use at the end of the coming period.

ITEM SALES: SEASONAL AND TREND PATTERN

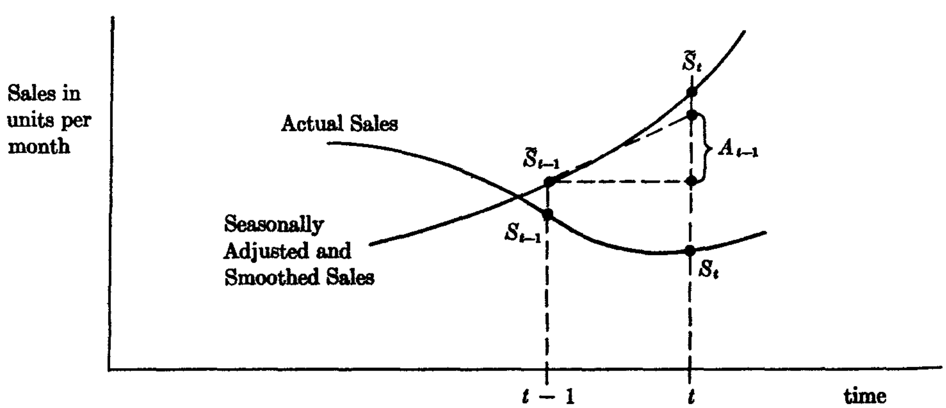


FIG. 2

Once again, the forecast is a function of past and current sales, of the weights A, B, C, and of the initial values of $\tilde{S}$, the F's, and R. And the quality, or accuracy, of forecasts depends upon these things, all of which except the sales history are at the command of the forecaster. In the next section we will discuss ways to select the weights and the initial conditions, and then we will demonstrate the use of the forecasting method.

3. Empirical Tests of Forecasting

3.1 *Selection of Weights and Initial Values*

First we will take up the problem of selecting initial values of $\tilde{S}$, the F's, and R, and the determination of the weights A, B, and C. Three series of sales are used to illustrate the application of the exponential system; these three are:

(1) Monthly sales of a cooking utensil, manufactured by Wearever, Inc. a subsidiary of the Aluminum Company of America; these are sales from the manufacturer's central warehouse to dealers; 7 years of data. See Figure 3.

(2) Bi-monthly sales of paint, one can-size of one color, manufactured by the paint division of Pittsburgh Plate Glass; sales are from one of the manufacturer's district warehouses to dealers; 5 years of data. See Figure 4.

(3) Monthly data of cellars excavated in one geographical area for the erection of prefabricated houses manufactured and sold by Admiral Homes, Inc. of West Newton, Pa.; 7 years of data. See Figure 5.

In each series the periods are numbered $t = 1, 2, \cdots, D$. For the cooking utensil, $D = 84$; for paint, $D = 30$; for cellars, $D = 84$. Each series has been divided into two parts, the first part of which was used to develop initial values of S, the F's, and R. The length of this first part of the series was H periods

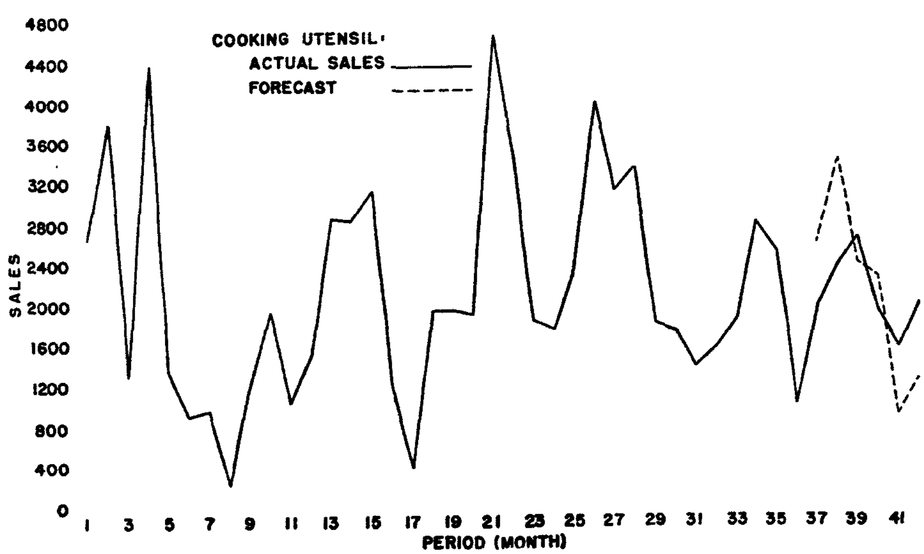


FIG. 3

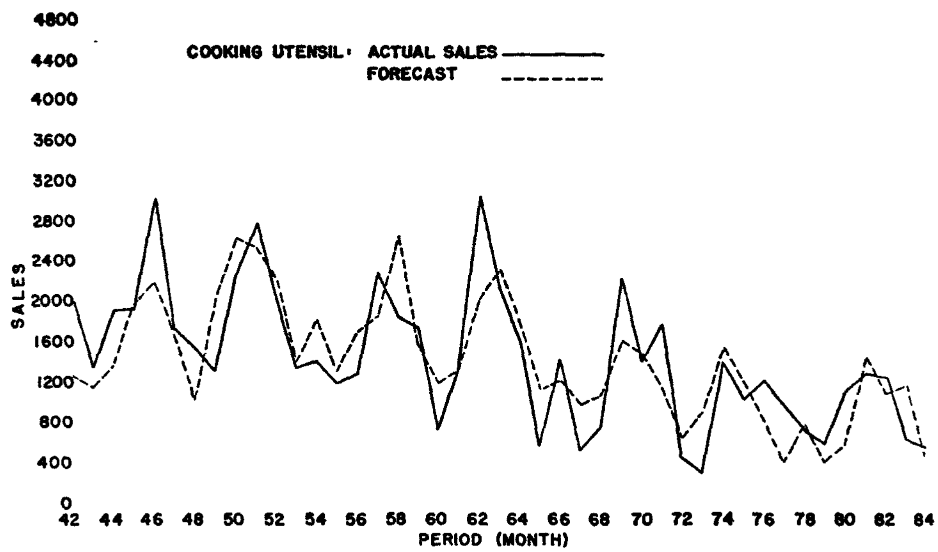


FIG. 3 (*Continued*)

($H = 36$, or 3 years, for the cooking utensil and cellars, and $H = 12$, or 2 years, for paint.) The second part of each series was then used to try out the forecasting method, by pretending that the future was unknown, making a forecast, moving along one period of actual sales data, comparing the forecast with actual sales, absorbing the actual sales data into the model, making another forecast, and so on. This is exactly as one would behave in practice, except for the advantage of collapsing several years of experience into a very short time, and the advantage of being able to experiment with a variety of values of the parameters, A, B, C.

The question immediately raised is how to compare one set of forecasts with another, that is, the effect of one set of parameters with the effect of another, or one forecasting method with another. Since the forecasts are predictions of expected values of sales it seems appropriate to choose as a criterion the standard deviation of forecast errors. If the forecast is unbiased, and this assumption is made here, then an estimate of the standard deviation of forecast errors is given by the sum of squared forecast errors. The forecast error is defined as

$$e_{t,T} = S_{t+T} - S_{t,T} \tag{15}$$

with S_{t+T} actual sales for the $(t + T)$ period, and $S_{t,T}$ the forecast for that period made in period t. If one is interested in forecasting only one period in the future, at any time, then the standard deviation of forecast errors is given by:

$$\sigma_e = \left\{ \frac{\sum_t e_{t,1}}{N - 1} \right\}^{1/2} \tag{16}$$

where t is summed over the set of N observations used for estimation. In practice forecasts are often made for several periods into the future, with the accuracy

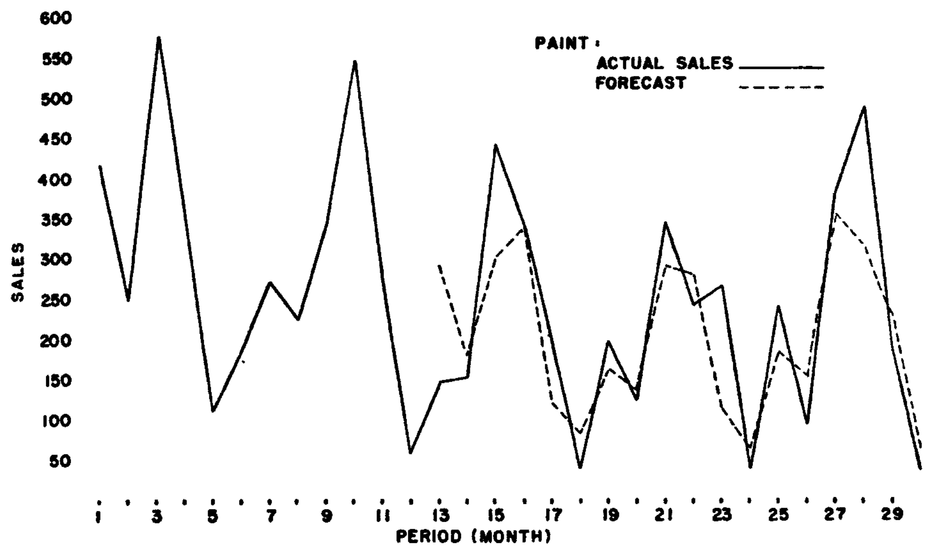


FIG. 4

of prediction most important for the coming period, and of declining importance for the later periods. In that case, a criterion function, U, which weights forecast errors such as the following can be used:

$$U = \sum_t (a_1 e_{t,1}^2 + a_2 e_{t,2}^2 + a_3 e_{t,3}^2) \tag{17}$$

with $a_1 > a_2 > a_3$, and possibly with $(a_1 + a_2 + a_3) = 1$. Investigation has shown that the set of weights (A, B, C) which given minimum σ_e is about the same as the set which gives minimum U for $a_1 = .6$, $a_2 = .3$, $a_3 = .1$. Because of this, we have used the simple criterion of Equation 16 for the empirical tests in the paper.

One can use any of a number of alternative methods to search for the set of (A, B, C) which minimizes σ_e. For example, we tried the method of steepest descent[4] and found it more or less satisfactory as a technique for finding the best weights for a single product. This method, however, gives only local information about the nature of σ_e, and it consumes enough time that it would not be feasible to use this method for each product, individually. We suspected that in fact σ_e is fairly "flat" near the minimum for any particular product, and that a single set of weights could be used for large classes of individual products. In order to find and present this kind of information, a grid of values of A, B, C was used for one series, the cooking utensil, and then additional points near the

[4] The method of steepest descent is sometimes called the gradient method. See, for example, *Modern Mathematics for the Engineer*, edited by E. F. Beckenbach, McGraw-Hill, 1956. Also, G. N. Lance, "Solution of Algebraic and Transcendental Equations on an Automatic Digital Computer," *Journal of the Association for Computing Machinery*, January, 1959.

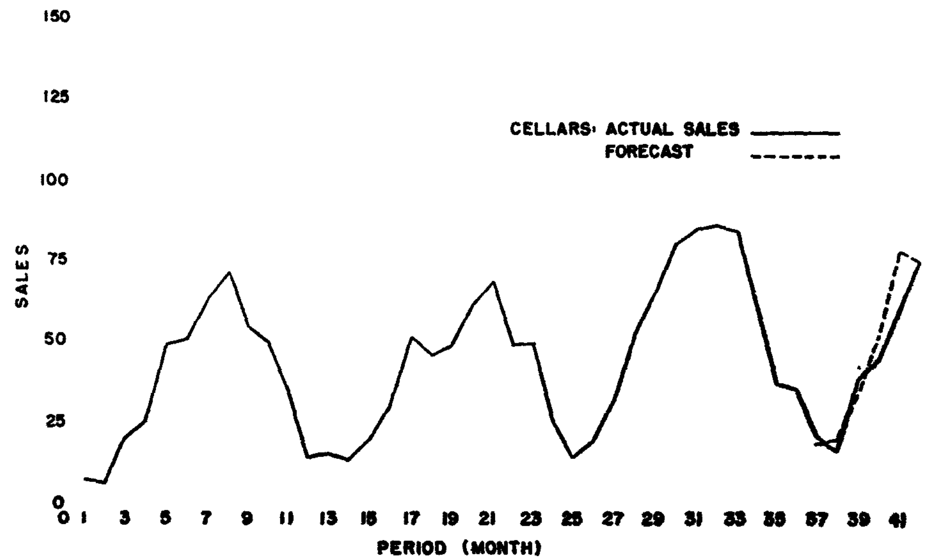


FIG. 5

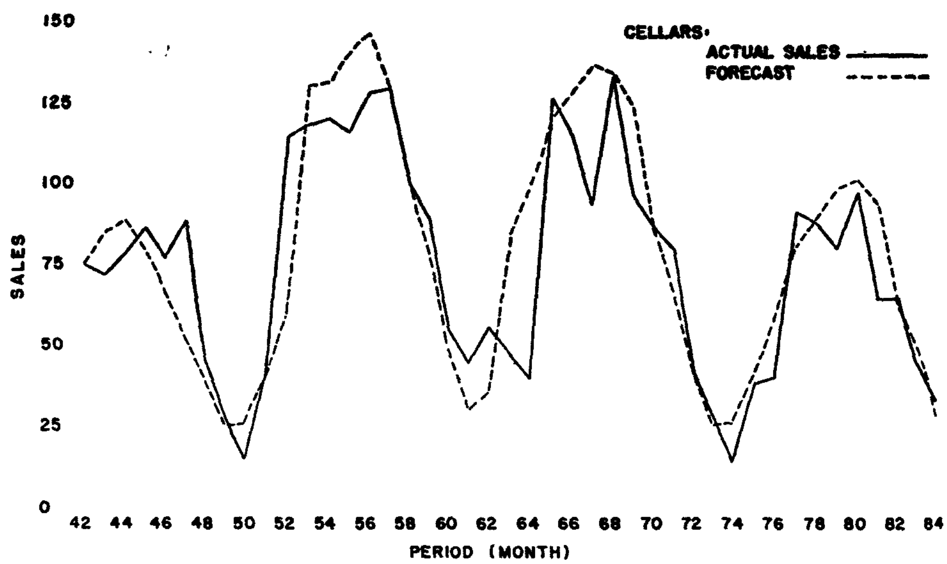


FIG. 5 (*Continued*)

minimum were used for each of the three series. The grid was made up of all possible combinations of 0, .2, .4, .6, .8, 1.0. The values of σ_e for these sets is given in Table 1. Table 2 shows the areas around the minimums of each series; the smallest division of the weights was a single decimal digit, i.e., (.1, .3, .2), but not (.13, .28, .24). It did in fact turn out that σ_e is flat near the minimum, but a detailed discussion of the results will be left until the method of calculation has been more fully described.

TABLE 1

Standard Deviations of Forecast Errors: Grid Results for the Cooking Utensil

B	C											
	0	.2	.4	.6	.8	1.0	0	.2	.4	.6	.8	1.0
	A = 0						A = .6					
0	1607	1607	1607	1607	1607	1607	630	663	717	781	855	940
.2	1249	1249	1249	1249	1249	1249	592	626	676	734	801	879
.4	1006	1006	1006	1006	1006	1006	567	603	655	713	780	864
.6	856	856	856	856	856	856	557	598	659	725	803	908
.8	778	778	778	778	778	778	566	617	698	791	915	1084
1.0	764	764	764	764	764	764	597	670	796	970	1294	1832
	A = .2						A = .8					
0	640	574	600	614	641	672	676	736	815	907	1009	1115
.2	570	518	539	557	599	649	650	710	784	871	967	1067
.4	531	488	510	541	622	752	630	689	761	845	938	1034
.6	523	*487**	522	605	849	4719	614	674	746	830	924	1019
.8	544	520	598	877	3592	20850	604	666	740	828	930	1029
1.0	603	617	840	2776	9282	11135	600	666	746	844	963	1077
	A = .4						A = 1.0					
0	607	609	641	678	719	768	738	825	930	1049	1178	1320
.2	559	565	597	637	679	726	738	825	930	1049	1178	1320
.4	532	543	582	637	698	767	738	825	930	1049	1178	1320
.6	532	550	607	698	816	958	738	825	930	1049	1178	1320
.8	558	595	691	863	1122	1483	738	825	930	1049	1178	1320
1.0	619	699	896	1578	4088	4424	738	825	930	1049	1178	1320

Values of $\sigma_e(A, B, C)$ were calculated on the IBM 650 computer. A flow chart of the program appears in Figure 6. The general procedure was as follows: the first part of a series was used in a common-sense way to get initial values of $\bar{S}$, the F's, and R. The exponential model was then used on the *first part* of the series, $(t = 1, 2, \cdots, H)$, in the same manner as it was used in the second part, except that no forecasts were made, and thus no measure was made of forecast error. The values of $\bar{S}$, the F's, and R that remained at period H were then considered initial values for the second half of a series. The added complication of using the first part of the series twice was incurred in order to wash out as far as possible the effects of the arbitrarily chosen initial values. This is particularly important in the case of the seasonal factors, the F's, because each is re-estimated only once a year; and it becomes increasingly important for each of the effects (smoothing, seasonal, and trend) as the weights become smaller. The best values of A, B, C are in fact relatively small so that this is a real problem. The method of obtaining values to start the first part of the series was the following.

(1) The average sales per period for each year, V_i, was computed. The i subscript refers to the year.

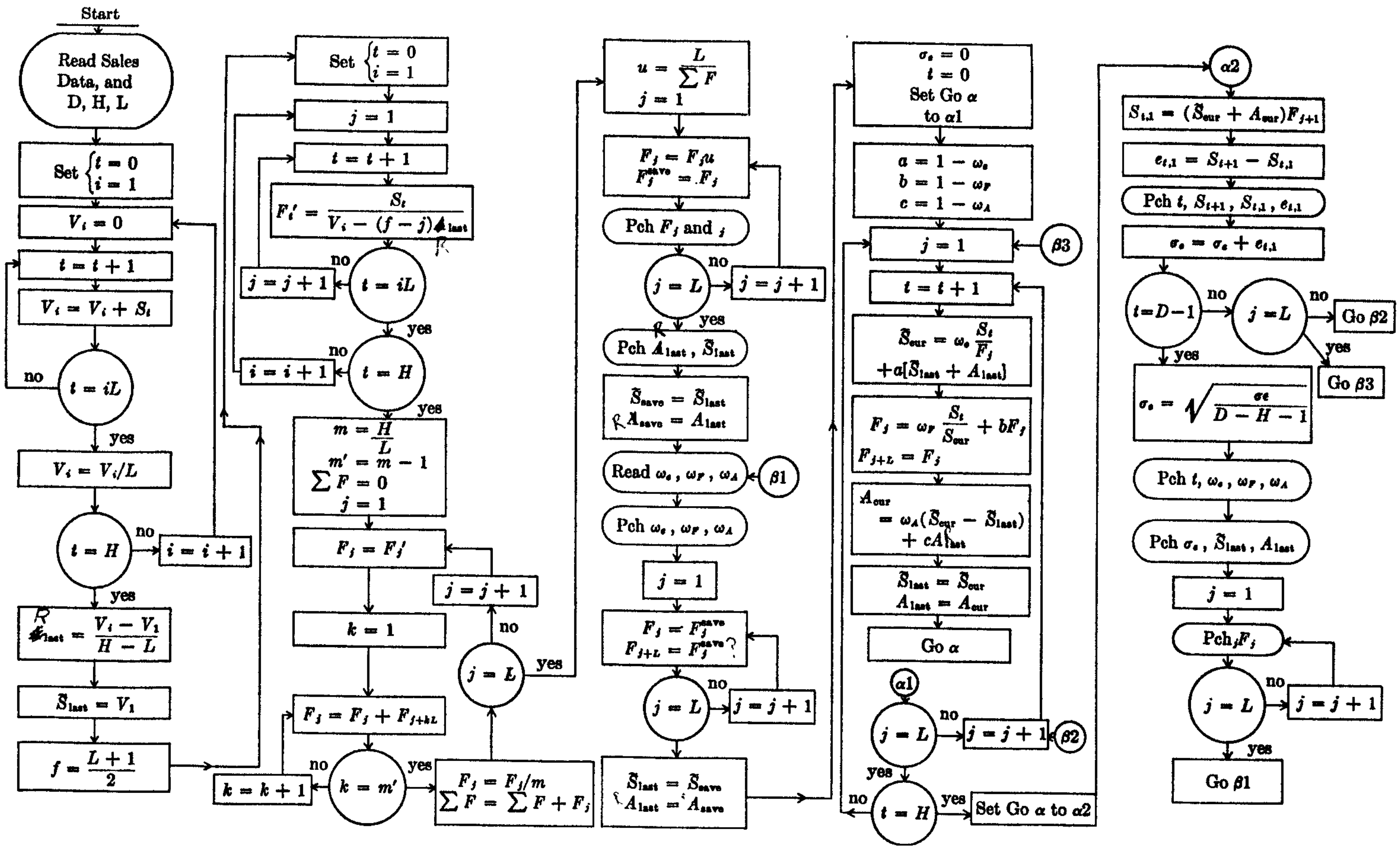
(2) R_{last}, the "previous" estimate of R (corresponds to R_{t-1} when in the t'th period) is $R_{last} = [V_{(H/L)} - V_1]/[H - L]$. In other words, R_{last} is

TABLE 2

Standard Deviations of Forecast Errors: Three Products and Composite Results

B	C																							
	Cooking Utensils						Paint						Cellars						Composite Rating—Three Products					
	0	.1	.2	.3	.4	.5	0	.1	.2	.3	.4	.5	0	.1	.2	.3	.4	.5	0	.1	.2	.3	.4	.5
A = .5																								
0	614	617	634				97.2	98.4	100.8				19.0*	19.8*	20.7				—	—	—			
.1	591	595	611				95.8	96.8	99.2				19.4*	20.4*	21.2				42	—	—			
.2	572	577	593				94.5	95.5	97.9				19.9*	20.9	21.9				39	—	—			
.3	557	563	579				93.4	94.3	96.7				20.4*	21.6	22.6				37	45	—			
A = .4																								
0	607	598	609				95.4	95.7	97.0				19.0*	19.8*	20.5*				42	45	—			
.1	580	574	584				93.7	93.8	95.1				19.4*	20.3*	21.1				38	41	—			
.2	559	554	565				92.2	02.2	93.4				19.9*	20.9	21.8				33	38	45			
.3	543	539	550				90.8	90.7	91.9				20.4*	21.6	22.6				31	36	45			
.4	532	530	543				89.7	89.5	90.7				20.9	22.2	23.4				30	36	—			
.5	529	527	542				88.8	88.6	89.8				21.5	23.0	24.4				32	40	—			
A = .3																								
0	611	584	589				93.9	93.2	93.3				19.5*	20.4*	20.9				44	42	45			
.1	580	557	562				91.9	91.1	91.0				19.7*	20.8*	21.4				37	37	39			
.2	555	535	540				90.2	89.2	89.2				20.1*	21.2	21.9				31	31	36			
.3	536	519	524	537	552		88.7	87.7	87.7	88.2	89.6		20.4*	21.7	22.5	23.5	24.9		26*	30	35	43	—	
.4	525	509	516	530	548		87.6	86.6	86.6	87.2	88.6		20.8	22.2	23.1	24.3	26.2		25*	28*	34	44	—	
.5	520	506	515	532	556		86.9	85.9	86.0	86.8	88.4		21.3	22.7	23.7	25.1	27.5		26*	29	36	—	—	
.6	523	510	522	546	577		86.5	85.7	86.0	87.0	89.0		21.8	23.2	24.3	25.9	28.9		29	33	40	—	—	
.7	533	522	540				86.7	86.1	86.8				22.3	23.8	24.9				34	39	—			
A = .2																								
0	640	583					93.3	92.1					20.8	21.8					—	—				
.1	602	552					91.2	89.8					20.6	21.8					45	39				
.2	570	526					89.4	88.0					20.6	21.8					36	31				
.3	547	506	500	510	520	528	88.0	86.5	85.7	84.8*	84.1*	84.1*	20.7	22.0	22.8	23.1	23.3	23.9	30	26*	28*	31	33	38
.4	531	493	488*	499	510	521	87.0	85.5	84.8*	84.0*	83.5*	83.6*	20.9	22.2	23.0	23.3	23.5	24.1	26*	24*	26*	29	31	37
.5	523	487*	484*	496	510	527	86.5	85.0*	84.4*	83.9*	83.6*	83.9*	21.2	22.6	23.4	23.6	23.6	24.3	26*	24*	26*	30	33	40
.6	523	488*	487*	503	522	550	86.5	85.1*	84.7*	84.5*	84.5*	85.1*	21.7	23.1	23.9	24.0	23.9	24.8	29	26*	30	34	37	—
.7	530	497	499	520	550	597	87.1	86.0	85.9	85.9	86.4	87.3	22.3	23.8	24.5	24.5	24.5	25.9	34	33	37	41	—	—
A = .1																								
.3	620	537	518	501	502	533	89.7	87.9	88.0	88.0	87.2	85.9	22.3	23.8	24.5	25.5	25.9	25.4	—	45	44	—	—	—
.4	594	521	502	486*	494	532	88.8	87.0	87.0	86.9	86.2	85.1*	22.3	23.9	24.7	25.7	26.0	25.6	—	40	40	43	45	—
.5	578	512	492	479*	494	538	88.5	86.6	86.5	86.4	85.8	84.8*	22.5	24.2	25.2	26.1	26.3	26.0	—	39	40	43	—	—
.6	569	509	488*	479*	502	551	88.6	86.8	86.6	86.5	86.0	85.3	22.9	24.8	25.8	26.7	26.8	26.6	—	42	42	45	—	—
.7	569	514	491	486*	517	573	89.3	87.6	87.4	87.3	87.1	86.6	23.5	25.5	26.7	27.4	27.6	27.7	—	—	—	—	—	—

The * marks indicate values of σ within 2% of minimum value for Cooking Utensils, within 2% for Paint, and within 8% for Cellars; they also indicate the best Composite



Start
Read Sales Data, and D, H, L
Set $\begin{cases} t = 0 \\ i = 1 \end{cases}$
$V_i = 0$
$t = t + 1$
$V_i = V_i + S_t$
$t = iL$
no
yes
$V_i = V_i/L$
$t = H$
no
$i = i + 1$
yes
$A_{last} = \frac{V_i - V_1}{H - L}$
$S_{last} = V_1$
$f = \frac{L+1}{2}$
Set $\begin{cases} t = 0 \\ i = 1 \end{cases}$
$j = 1$
$t = t + 1$
$F_t' = \frac{S_t}{V_i - (f - j)A_{last}}$
$t = iL$
no
$j = j + 1$
yes
$t = H$
no
$i = i + 1$
yes
$m = \frac{H}{L}$
$m' = m - 1$
$\sum F = 0$
$j = 1$
$F_j = F_j'$
$j = j + 1$
no
$j = L$
yes
$k = 1$
$F_j = F_j + F_{j+kL}$
$k = m'$
no
$k = k + 1$
yes
$F_j = F_j/m$
$\sum F = \sum F + F_j$
$u = \frac{L}{\sum F}$
$j = 1$
$F_j = F_j u$
$F_j^{save} = F_j$
Pch F_j and j
$j = L$
no
$j = j + 1$
yes
Pch A_{last}, S_{last}
$S_{save} = S_{last}$
$A_{save} = A_{last}$
Read ω_s, ω_F, ω_A
$\beta 1$
Pch ω_s, ω_F, ω_A
$j = 1$
$F_j = F_j^{save}$
$F_{j+L} = F_j^{save}$
$j = L$
no
$j = j + 1$
$S_{last} = S_{save}$
$A_{last} = A_{save}$
$\sigma_e = 0$
$t = 0$
Set Go α to $\alpha 1$
$a = 1 - \omega_s$
$b = 1 - \omega_F$
$c = 1 - \omega_A$
$j = 1$
$\beta 3$
$t = t + 1$
$S_{cur} = \omega_s \frac{S_t}{F_j} + a[S_{last} + A_{last}]$
$F_j = \omega_F \frac{S_t}{S_{cur}} + bF_j$
$F_{j+L} = F_j$
$A_{cur} = \omega_A(S_{cur} - S_{last}) + cA_{last}$
$S_{last} = S_{cur}$
$A_{last} = A_{cur}$
Go α
$\alpha 1$
$j = L$
no
$j = j + 1$
$\beta 2$
yes
$t = H$
no
yes
Set Go α to $\alpha 2$
$\alpha 2$
$S_{t,1} = (S_{cur} + A_{cur})F_{j+1}$
$e_{t,1} = S_{t+1} - S_{t,1}$
Pch t, S_{t+1}, $S_{t,1}$, $e_{t,1}$
$\sigma_e = \sigma_e + e_{t,1}$
$t = D - 1$
no
$j = L$
no
Go $\beta 2$
yes
Go $\beta 3$
yes
$\sigma_e = \sqrt{\frac{\sigma e}{D - H - 1}}$
Pch t, ω_s, ω_F, ω_A
Pch σ_e, S_{last}, A_{last}
$j = 1$
Pch $_j F_j$
$j = L$
no
$j = j + 1$
yes
Go $\beta 1$

Fig. 6

the average trend between the first and last years, considering only these years.

(3) $\tilde{S}_{\text{last}}$, the "previous" estimate of $\tilde{S}$ (corresponds to $\tilde{S}_{t-1}$ when in the t'th period) is $\tilde{S}_{\text{last}} = V_1$, the average sales for the first year.

(4) Seasonal factors were computed for each period, $t = 1, \cdots, H$, as the ratio of actual sales for the period to average seasonally adjusted sales for that year, further adjusted by the trend, R_{last}:

$$F_t = \frac{S_t}{V_i - \left(\frac{L+1}{2} - j\right) R_{\text{last}}}$$

where V_i is the average sales in the appropriate year, and j is the position of the period within the year, e.g., for January, $j = 1$; for February, $j = 2$; etc.

(5) Seasonal factors for corresponding periods in each of the initial years were averaged to obtain one seasonal factor for each period in a year. For example, the F's were averaged for all Januarys to get one January seasonal factor.

(6) Finally, the seasonals were normalized so that they added to L; for 12 periods a year, $\sum_{j=1}^{12} F_j = 12$.

$$F_j = \text{Ave } F_j \left(\frac{L}{\sum_{j=1}^{L} \text{Ave } F_j} \right)$$

This last step was made to ensure that over a cycle the seasonal factors would make only seasonal adjustments, and not increase or decrease the average level of sales.

This process gave values of $\tilde{S}$, the F's, and R to use in period $t = 1$. At this point a set of values of (A, B, C) was read into the machine. The exponential system was applied starting as $t = 1$ and run through $t = H$ without making forecasts, and then through the rest of the data, $H + 1$ to L, making forecasts, and computing forecast errors. See Figure 6 for details.

3.2 *Results of Testing the Exponential System*

The results of the grid of (A, B, C) for the cooking utensil series are given in Table 1. The lowest σ_e is 487, at $(A = .2, B = .6, C = .2)$. The function σ_e is convex in the region of the minimum. Further search was conducted in the neighborhood of (.2, .6, .2), by taking a finer grid. These results are presented in Table 2. Slightly better values of σ_e were found, with the new minimum of 479 at two points: (.1, .5, .3) and (.1, .6, .3). Values of σ_e within 2% of the minimum are starred, and the minima are starred and underlined. The function σ_e is fairly "flat" in the neighborhood of the minima, as we suspected. Undoubtedly we could find lower values of σ_e, but the probable gain doesn't appear worth the effort.

A search was made, using the same finer grid size, for the minimum σ_e's of the other two series, using the assumption that they, too, were convex in a substantial region surrounding the minima. The results of this search is shown in Table 2, along with the cooking utensil values. Again, the best values are underlined, and values within 2 % of the minimum for paint, and within 8 %, for cellars, are starred. The minimum for paint is 83.5, at (.2, .4, .4); the minimum σ_e for cellars is 19.0, at (.4, 0, 0).[5] Values of zero for B and C mean that the original estimates of the seasonal and trend factors are never changed. The functions σ_e are, as in the case of the cooking utensil, quite "flat" in the neighborhoods of the minima, and are convex in these regions.

It is possible to get some idea of the accuracy of prediction from the standard deviation of forecast errors. The coefficient of variation of the distribution of sales rates, given the forecasting method, can be approximated by the ratio of σ_e to the average sales rate for that part of the series in which forecasts are made, i.e.,

$$\text{(18)} \qquad \text{Coefficient of variation} \cong \sigma_e \Big/ \left[\sum_{t=H+1}^{D} S_t/(D - H) \right] = \sigma_e/\bar{S}$$

For each of the three series, this value is

Cooking utensils	.31
Paint	.38
Cellars	.26

This means that, as an order of magnitude, about 65 % of the forecasts will be within ±31 %, 38 %, and 26 %, respectively, of the sales rates that actually occur.

To give some idea of the effect of different sets of values of (A, B, C) on all three series, a composite rating has been devised as follows. Given the minimum σ_e, all the other values can be expressed as a percentage above the minimum. This has been done for most of the sets of weights evaluated in Table 2, and percentages added across the three series. For example, for (.2, .4, .2) the percentage above the minimum is 2 % for cooking utensils, 2 % for paint, and 22 % for cellars. The composite rating is the sum of these three, 26 %, and appears in position (.2, .4, .2) in Table 2. The best composite rating is 24 %, for (.2, .4, .1) and for (.2, .5, .1). Other good composite ratings are starred.

3.3 *General Comments on the Exponential Weights*

It is possible to make some general statements about the optimal weights that we would expect to find associated with various kinds of sales series. The best weights for cellars of (.4, 0, 0) seem strange at first. We might expect zero weights, which mean no weight to current estimates, in a situation where the random effect in current data is substantial. This would imply that accuracy of prediction would be low, but that is not the case with cellars. In fact, the coefficient of variation for cellars is the lowest of the three, indicating the best prediction.

[5] A comment on zero values will be made later in the section.

In order to shed light on this apparent paradox, some examples have been chosen which are a little extreme in order to make the point clear. One of the justifications for the exponential system is that if any of the basic parameters of a series changes or drifts over time, the forecasting system will soon pick up the new value; another justification is the ability to filter out substantial random effects in the observations. It is possible to imagine, however, a series in which some of the parameters are subject to little or no drift; and also series where there is little random effect. If there is no change in a parameter over the series then even if the random effect is small, the weight associated with that parameter would be very small, or even zero, because there is no value in changing, very much or at all, the original, and still accurate estimate of the parameter. If the drift in a parameter is large over the series, even to the point of abrupt changes in its value from time to time, there are two possibilities: a) little random effect would lead to large weights, weighting current estimates heavily; b) a large random effect would yield substantially smaller weights, depending on the relative importance of the changes in the parameter *versus* the random element. These intuitive conclusions are summarized in the following table:

Size of Weights Expected

Drift or change in parameters	Random Effect		
	None	Little	Large
None	indeterminate	very small	small
Little	large	very small to small	small
Large	large	medium	small to medium

Evidently, the cellars series is one in which there is not much change in the seasonal or trend parameters over the series, and at the same time, the random element in the series is not large. Also, note that A is somewhat larger for this series than for the other two; some of the change in $\bar{S}$, not accounted for by a trend factor, is made up by the faster re-estimation of $\bar{S}$ itself.

4. Comparison with Other Forecasting Models

As we mentioned in the first section of the paper, two other forecasting models were applied to the three series, cooking utensils, paint, cellars excavated, to make a comparison with the exponential system. We call these two more conventional models Comparison Model No. 1, and Comparison Model No. 2. Comparison Model No. 1 is quite simple. A prediction for any period is made by averaging sales in the two preceeding periods:

$$S_{t,1} = \tfrac{1}{2}(S_t + S_{t-1}) \tag{19}$$

Detailed results of this forecasting model will not be given, but a summary of the results appears in Table 3, along with the best forecast results of the exponential system. The standard deviations of forecast errors for Comparison Model No. 1

TABLE 3

Comparison of Forecasting Models

Exponential Forecast: Best Values			Standard Deviation of Forecast Errors		
A	B	C	Kitchen Utensils	Paint	Cellars Excavated
.1	.5	.3	479	—	—
.2	.4	.4	—	83.5	—
.4	0	0	—	—	19.0
Comparison Model No. 1			733	177.9	28.6
Comparison Model No. 2			512	97.5	22.6
Exponential Forecast Best Overall Weights					
A	B	C			
.2	.3	.1	507	86.5	22.0
.2	.4	.1	494	85.5	22.2
.3	.4	.1	509	86.6	22.2
Next Best Weights					
A	B	C			
.2	.3	.2	500	85.7	22.8
.2	.4	.2	488	84.8	23.0
.2	.5	.1	487	85.0	22.6
.2	.6	.1	488	85.1	23.1
.3	.3	.1	518	87.7	21.7
.3	.5	0	520	86.9	21.3
.3	.5	.1	506	85.9	22.7

are substantially higher for every series, and outstandingly so for the Paint series.

Comparison Model No. 2 is a more serious contender. It is in some ways similar to the variant of the exponential model with ratio seasonals and no trend factor. A forecast is made for a period by multiplying by a seasonal factor the average sales over the preceeding L periods:

$$S_{t,1} = \frac{1}{L}\left[\sum_{j=0}^{L-1} S_{t-j}\right] F_{t+1-L} \tag{20}$$

The seasonal factor, F, for the period in question was developed during the previous cycle, and is thus the latest estimate of a seasonal factor for the period being forecast.

The three series are handled in the same way for this model as for the exponential system: the first part of the data, through period H, is used to calculate seasonal factors. These seasonal factors are then used to forecast sales in the second part of the data ($H + 1$ to L), and are adjusted each year by weighting the current estimate by $\frac{1}{3}$, the previous estimate by $\frac{2}{3}$, and adding, exactly as with the exponential system. Initial F's are calculated by simply reversing Equation 20:

$$F_t = LS_t \Big/ \left[\sum_{j=1}^{L} S_{t-j} \right] \tag{21}$$

Because of this definition of the seasonal factor however, three years of data yields only two seasonal factors for corresponding periods, and two years of data given only one seasonal factor. Two or more estimates of an F are averaged. These seasonal factors are not normalized, of course, because they are intended to include some trend effect. If there is a consistent long-run trend, not normalizing leads to better forecasts.

Results of Comparison Model No. 2 are given in Table 3. The standard deviations of forecast errors for this model are higher for each of the three series than for the exponential model using the best set of weights for each series individually. There are three sets of weights which, when used for all of the series, yield standard deviations for each series lower than those from Comparison Model No. 2. These are shown as Best Overall Weights in Table 3. Finally, there are a number of sets of weights which yield better results for two out of the three series; these are shown as Next Best Weights in Table 3. In all of these latter cases, the exponential result that is worse for the one series is not much worse.

We conclude, then, that the exponential forecasting model has several advantages over more conventional forecasting models: (1) it gives better forecasts, (2) it requires less information and storage space, (3) it responds more rapidly to sudden shifts in the time series so that is routinely protects the forecaster.