Forecasting Trends in Time Series

Author(s): Everette S. Gardner, Jr. and Ed. McKenzie

Source: *Management Science*, Vol. 31, No. 10 (Oct., 1985), pp. 1237-1246
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/2631713
Accessed: 16/02/2011 15:22

MANAGEMENT SCIENCE
Vol. 31, No. 10, October 1985
Printed in U.S.A.

FORECASTING TRENDS IN TIME SERIES*

EVERETTE S. GARDNER, JR. AND ED. McKENZIE
Operations Analysis Department, Navy Fleet Material Support Office, P.O. Box 2010, *Mechanicsburg, Pennsylvania* 17055
Mathematics Department, University of Strathclyde, Glasgow G1 1XW, *Scotland, United Kingdom*

Most time series methods assume that any trend will continue unabated, regardless of the forecast leadtime. But recent empirical findings suggest that forecast accuracy can be improved by either damping or ignoring altogether trends which have a low probability of persistence. This paper develops an exponential smoothing model designed to damp erratic trends. The model is tested using the sample of 1,001 time series first analyzed by Makridakis et al. Compared to smoothing models based on a linear trend, the model improves forecast accuracy, particularly at long leadtimes. The model also compares favorably to sophisticated time series models noted for good long-range performance, such as those of Lewandowski and Parzen.
(FORECASTING—TIME SERIES)

1. Introduction

Research in time series analysis and forecasting has traditionally been concerned with modelling the autocorrelation structure in a stationary time series. However, as discussed in Fildes (1983), recent empirical work has shown this to be a relatively unimportant problem compared to the modelling of trends. For example, Makridakis et al. evaluated the post-sample accuracy of 21 automatic forecasting methods on a collection of 1,001 time series. The accuracy of all methods deteriorated badly at leadtimes more than a few steps ahead. This was particularly true of methods based on a linear trend which typically overshot the data at long leadtimes.

Makridakis also examined a subset of 111 time series taken from the 1,001. Several sophisticated methods were tested in this subset in addition to the 21 automatic methods. The sophisticated methods included the Box–Jenkins approach, the FORSYS system of Lewandowski (1982), and the ARARMA methodology of Parzen (1979, 1982). Like the automatic methods, Box–Jenkins did badly at long leadtimes. However, Lewandowski and Parzen were the most accurate at long leadtimes among all methods tested.

Lewandowski's FORSYS system is widely used in European companies. The distinguishing feature of FORSYS is that it damps the trend as the forecast leadtime increases. The rate of damping increases with the level of noise in the series. The rationale is that the more noise in the series the greater the risk in trend extrapolation. It is difficult to say more than this about FORSYS because the system is proprietary.

Parzen's approach attempts to classify the "memory" of the time series. "Short-memory" series are covariance-stationary and are modelled by conventional ARMA schemes. "Long-memory" series contain trends modelled by nonstationary autoregression. This approach produced models robust at all leadtimes.

The Makridakis study indicates the need for more research in trend extrapolation, particularly for relatively simple models which can be used in automatic forecasting systems. This paper develops and tests a generalization of the widely-used Holt model for exponential smoothing of a linear trend. The generalization adds a damping

*Accepted by Ambar G. Rao; received June 20, 1984. This paper has been with the authors 1 month for 1 revision.

0025-1909/85/3110/1237$01.25

parameter to the model to give more control over trend extrapolation. The damping parameter corresponds to an autoregressive term in the equivalent ARIMA process. For this reason, we refer to it as the autoregressive-damping (AD) parameter.

Exponential smoothing systems including AD-parameters are not new. They have been noted as members of certain larger classes of forecasting systems by Gardner (1985), Gilchrist (1976), and Roberts (1982). However, there has been no theoretical or empirical investigation of such systems. That is the purpose of this work. §2 discusses theoretical considerations in the use of AD-parameters, including alternative model formulations, stability regions, the effects of parameter choice, and equivalent ARIMA processes.

§3 is an empirical study of forecast accuracy. The generalized and standard Holt models are compared using the 1,001 time series from Makridakis et al. (1982). Comparisons are also made to the Lewandowski and Parzen approaches on the subset of 111 series. Conclusions from the empirical work enable us to make some positive recommendations in §4 about the practical application of the generalized Holt model.

An appendix on computational details is included. This should enable the empirical results to be replicated.

2. Theoretical Development

2.1. *Model Formulations*

We begin with the standard Holt (1960) formulation for exponential smoothing of a linear trend:

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

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}, \tag{2}$$

$$\hat{X}_t(m) = S_t + mT_t . \tag{3}$$

S_t is the local level of the series and T_t is the trend. $\hat{X}_t(m)$ is the forecast at origin t for m steps ahead. The smoothing parameters α and γ are usually restricted to the range 0 to 1 in practice although the model is stable over a wider range.

The trend estimate can be modified with an AD-parameter ϕ. The revised model is:

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + \phi T_{t-1}), \tag{4}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\phi T_{t-1}, \tag{5}$$

$$\hat{X}_t(m) = S_t + \sum_{i=1}^{m} \phi^m T_t . \tag{6}$$

This generalization includes four possibilities for trend depending on the value of ϕ. If $\phi = 0$, there is no trend in the forecasts—the model is equivalent to simple smoothing. If $0 < \phi < 1$, the trend is damped and the forecasts approach an asymptote given by the horizontal straight line $S_t + T_t\phi(1-\phi)$. If $\phi = 1$, the model is equivalent to the standard version of the Holt model and the trend is linear. Finally, if $\phi > 1$, the trend is exponential, which is probably a dangerous option in an automatic forecasting model.

In many forecasting systems, it is common practice to apply the same exponential smoothing model to every time series. This practice is unavoidable in large inventory systems where thousands of forecasts are needed each time period. If the generalized model is fitted over the range 0 to 1 for each parameter, accuracy should be better on average than standard Holt. Apart from wider applicability, the rationale for using the generalized model as an automatic forecaster is similar to that of Lewandowski. With a strong trend in the data, ϕ should be fitted at a value near 1 and the forecasts should be nearly the same as Holt. If the data are extremely noisy or if the trend is erratic, the model should damp the trend with a ϕ less than 1.

The model in (4)–(6) is cumbersome but can be simplified considerably. Using the one-step-ahead forecast error $e_t = X_t - S_{t-1} - \phi T_{t-1}$, we can rewrite (4) and (5) as

$$S_t = \hat{X}_{t-1}(1) + h_1 e_t, \tag{7}$$

$$T_t = \phi T_{t-1} + h_2 e_t, \tag{8}$$

where $h_1 = \alpha$, $h_2 = \alpha\gamma$.

Further simplification can be achieved by separating the level and trend components of the model. First define $A_t = S_t + T_t\phi/(1-\phi)$, and $B_t = -T_t\phi/(1-\phi)$. A_t is the asymptote. Note that B_t has an opposite sign to the trend in (8). These components can be smoothed with:

$$A_t = A_{t-1} + g_1 e_t, \tag{9}$$

$$B_t = \phi B_{t-1} + g_2 e_t, \tag{10}$$

$$\hat{X}_t(m) = A_t + \phi^m B_t. \tag{11}$$

The forecasts from (9)–(11) are identical to those of (6)–(8) when $\phi < 1$ and $g_1 = \alpha + \alpha\gamma\phi/(1-\phi)$, $g_2 = -\alpha\gamma\phi/(1-\phi)$. The advantage of this form of the model is that we avoid the need for a summation in the forecast equation. A direct estimate of the asymptote is also provided. The disadvantage is that separate provision must be made for the case of $\phi = 1$.

2.2. *Model Stability and Parameter Choice*

The model in (6)–(8) is stable over a wide range for h_1 and h_2 defined by:

$$(\phi - 1)/\phi < h_1 < (\phi + 1)/\phi, \qquad \phi h_2 + (1-\phi)h_1 > 0, \qquad \phi h_2 + (1+\phi)h_1 < 2(1+\phi).$$

Within the region of stability, the search for parameters can be constrained in many ways. Exhibit 1 shows several regions of choice for h_1 and h_2 (with the restriction that $0 \leqslant \phi \leqslant 1$). The three-parameter model discussed above constrains h_1 and h_2 to lie within the triangle (above the diagonal of the unit square).

A three-parameter nonseasonal model may be cumbersome in some applications. Discounted least squares (DLS) can be used to reduce the number of parameters to two. With discount factor β, the function to be minimized is: $\sum_{j=0}^{\infty} \beta^j [X_{t-j} -$

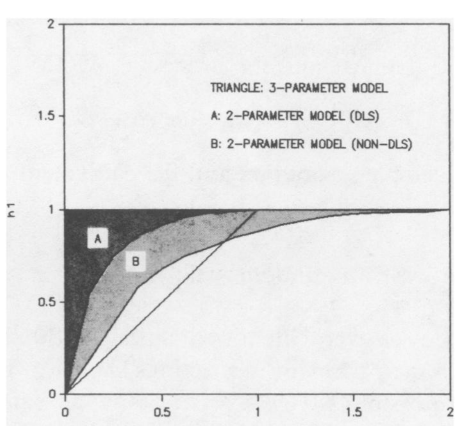

EXHIBIT 1. Regions of Parameter Choice.

$\hat{X}_t(-j)]^2$. Following McKenzie (1976), the solution is given by (6)–(8) with:

$$h_1 = \left[1 - (\beta/\phi)^2\right], \qquad h_2 = \left[1 - (\beta/\phi)\right]\left[1 - (\beta/\phi^2)\right]. \tag{12}$$

When $\phi = 1$, (12) reduces to the DLS solution for a linear trend recommended by Brown (1963). Region A in the exhibit corresponds to (12) and lies within the region for the three-parameter model.

Another way to reduce the number of parameters to two is to use:

$$h_1 = \alpha(2 - \alpha), \qquad h_2 = \alpha(\alpha - \phi + 1). \tag{13}$$

This is region B in the exhibit. (13) is an heuristic designed like (12) to give the DLS solution for a linear trend when $\phi = 1$. Note that (13) allows h_2 to be larger than h_1, which is not possible in the three-parameter model or in (12).

An empirical comparison of the forecast accuracy resulting from choosing parameters in these regions is given in the Appendix. The three-parameter model appears to be the most accurate. This should be expected because its region of choice includes all of the region for (12) and most of that for (13). However, the differences in accuracy among the three regions are small and may be insignificant to many users. Regardless of how parameters are selected, model (6)–(8) is more accurate than models based on a linear trend.

Using either two or three nonseasonal parameters, it is straightforward to incorporate seasonality in model (6)–(8). A selection of three-parameter seasonal models is given in Gardner (1985).

2.3. *Equivalent ARIMA Processes*

It is now well known that linear forecasting systems often have equivalent ARIMA processes. By equivalent, we mean that the forecasts are minimum mean squared error (MSE) forecasts for the corresponding ARIMA processes. For a fuller discussion and examples, reference may be made to Gardner (1985) or McKenzie (1984). The forecasting system given by (6)–(8) has at least six equivalent processes. These processes provide a statistical rationale for the model. In particular they can be used to compute the variance of the errors in order to set confidence limits around the forecasts.

If $0 < \phi < 1$, the trend is damped and the equivalent process is ARIMA (1, 1, 2) which can be written as:

$$(1 - B)(1 - \phi B)X_t = \left[1 - (1 + \phi - h_1 - \phi h_2)B - \phi(h_1 - 1)B^2\right]e_t. \tag{14}$$

We can also obtain an ARIMA (1, 1, 1) process by setting $h_1 = 1$. With $h_1 = h_2 = 1$, the process is ARIMA (1, 1, 0).

When $\phi = 1$, the trend is linear and the process is ARIMA (0, 2, 2):

$$(1 - B)^2 X_t = \left[1 - (2 - h_1 - h_2)B - (h_1 - 1)B^2\right]e_t. \tag{15}$$

When $\phi = 0$, we have simple smoothing and the equivalent ARIMA (0, 1, 1) process:

$$(1 - B)X_t = \left[1 - (1 - h_1)\right]e_t. \tag{16}$$

The ARIMA (0, 1, 0) process, or random walk model, can be obtained from (16) by choosing $h_1 = 1$.

All six ARIMA processes were identified amongst the 1,001 time series of the Makridakis data. Due to the parameter restrictions shown in Exhibit 1, these processes are only a subset of the possible ARIMA processes of the same order. To illustrate, in (14) we chose ϕ from the range 0 to 1. In the general ARIMA (1, 1, 2), ϕ can range from -1 to 1.

The parameter restrictions do not appear to be a practical disadvantage. As shown in the next section, model (6)–(8) is robust, which is a major consideration in the design of exponential smoothing systems.

3. Results of the Empirical Study

3.1. *Data and Experimental Design*

The accuracy of the generalized Holt model in (6)–(8) was evaluated using the collection of 1,001 time series first analyzed by Makridakis et al. (1982). This collection includes 181 yearly, 203 quarterly, and 617 monthly series. Another classification of the series is by level of aggregation. There are 302 series of company sales, 236 of industry sales, 319 macroeconomic series, and 114 demographic series. About two-thirds of the series are seasonal. The collection includes a wide range of starting and ending dates.

In the Makridakis study, each series was divided into two segments. Suppose there are N observations in each series. For the annual series, models were fitted to the first $N - 6$ observations; postsample forecasts were made at origin $N - 6$ for horizons 1 to 6. No postsample observations were used to generate forecasts. For the quarterly and monthly series, forecasts were made at origins $N - 8$ and $N - 18$ for the last 8 and 18 observations, respectively. Postsample forecast errors were compiled by horizon and averaged over all series.

When models were fitted to deseasonalized data, the seasonal indexes were computed with the ratio-to-moving average method. Again no postsample observations were used in computing the seasonal indexes. Forecasts were reseasonalized before the errors were compiled.

Procedures identical to Makridakis were used to compile postsample errors in this research. The three-parameter version of model (6)–(8) was used with deseasonalized data. To keep the results strictly comparable, Makridakis' original seasonal indexes were used. Parameters were selected from the triangular region in Exhibit 1 on the basis of minimum MSE (one-step-ahead).

3.2. *Comparisons to Holt's Linear Model*

Graphical comparisons to Holt's linear model, also based on deseasonalized data where necessary, are presented in Exhibits 2–4. Tables of forecast errors and computational details are given in the Appendix. In the Exhibits, "Holt-D" is the generalized model (6)–(8) which allows damping of trends. "Holt-L" is the standard model (1)–(3) with a linear trend. As discussed in the Appendix, the Holt–L results are better than

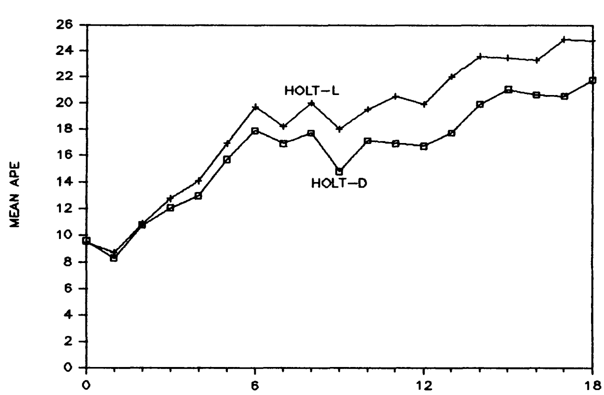

EXHIBIT 2. Mean APE by Horizon (1,001).

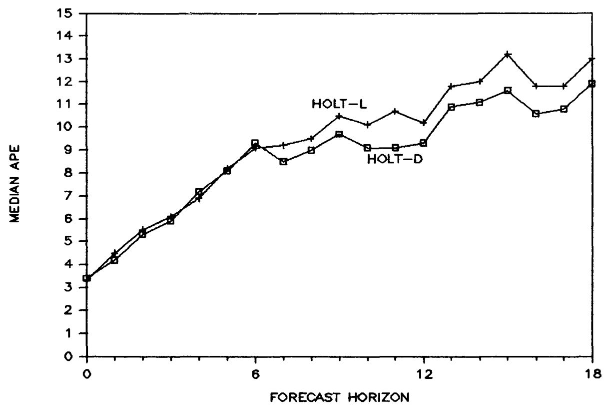

EXHIBIT 3. Median APE by Horizon (1,001).

those originally reported by Makridakis. The difference lies in the initialization procedure. Makridakis used backcasting to obtain initial forecasts whereas a simple linear regression on time was used in this work.

In Exhibit 2, the mean absolute percentage error (APE) by forecast horizon for all 1,001 series is plotted. Horizon zero refers to the fitted mean APE. At horizons of one or two steps ahead, there was little difference between the two models. As the horizon increased, Holt-D became substantially more accurate.

Median APEs by horizon are shown in Exhibit 3 on an enlarged scale. The APE distributions were skewed left, a common finding in empirical studies, with median APEs about half the size of the means. There was little difference in medians through horizon 6, although Holt-D was more accurate beyond that point.

Makridakis also analyzed a sample of 111 series taken from the population of 1,001. Overall comparisons between Holt-D and Holt-L were about the same in this sample, although there were important differences in accuracy on different types of data. The sample includes 20 yearly, 23 quarterly, and 68 monthly series. Exhibit 4 shows the mean APE (average over all horizons) by type of data within the 111 series.

Most of the yearly series contained a strong trend. In the Holt-D model, the trend was damped (with $\phi < 1$) for only about 20% of the yearly series. Thus the average accuracy was about the same as Holt-L. In the quarterly and monthly series, trends were far more erratic. Holt-D improved accuracy by damping the trend more than 70% of the time.

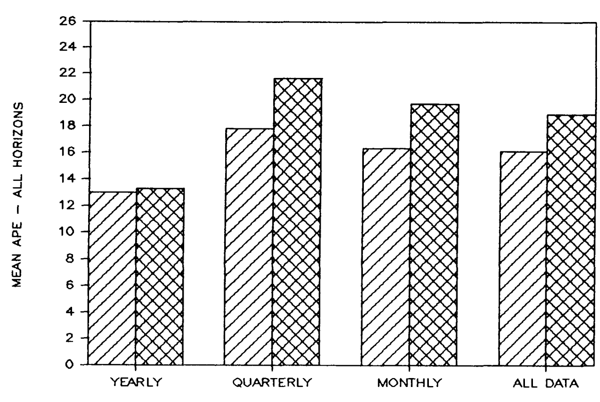

EXHIBIT 4. Mean APE by Type of Data (111).

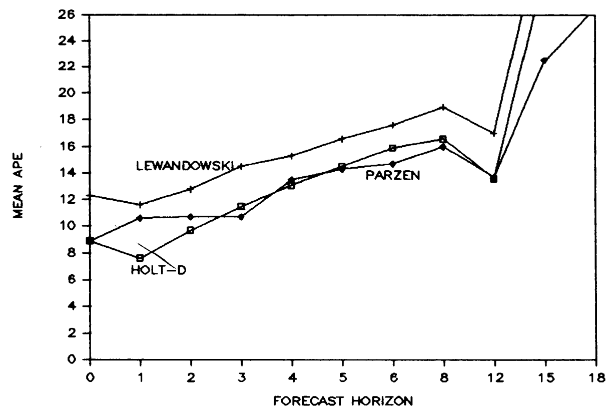

EXHIBIT 5. Mean APE by Horizon (111).

3.3. *Comparisons to Lewandowski and Parzen*

In the Makridakis study, Lewandowski was generally the best performer on the median APE criterion. However, the method did not yield a good mean APE. The reason was that the strategy of damping the trend in every time series resulted in some extremely large errors. Parzen's method behaved differently, giving a better mean APE than Lewandowski but a worse median.

The Holt-D model is compared to Lewandowski and Parzen in Exhibits 5 and 6. Both exhibits give results for the sample of 111 series. The Lewandowski and Parzen results are taken from Makridakis et al. (1982). In interpreting the exhibits, it should be noted that the X-scale contains gaps beyond horizon 6 (Makridakis did not report the errors at horizons 7, 9–11, 13–14, and 16–17).

In Exhibit 5, Holt-D gave a better mean APE than Lewandowski at all horizons through 15 steps ahead. Compared to Parzen, Holt-D did better at one step ahead, then about the same through 12 steps ahead. Thereafter Parzen was more accurate.

In Exhibit 6 (median APE), Holt-D did better at one step ahead than both sophisticated methods, then about the same through 8 steps ahead. Thereafter, both sophisticated methods were more accurate.

Thus the Holt-D model compared favorably overall to both sophisticated methods except at the longest leadtimes. Whether the forecasts from any of the three approaches are of any practical value beyond say 12 steps ahead is a matter of opinion. It is perhaps asking too much of any time series method to forecast at that range.

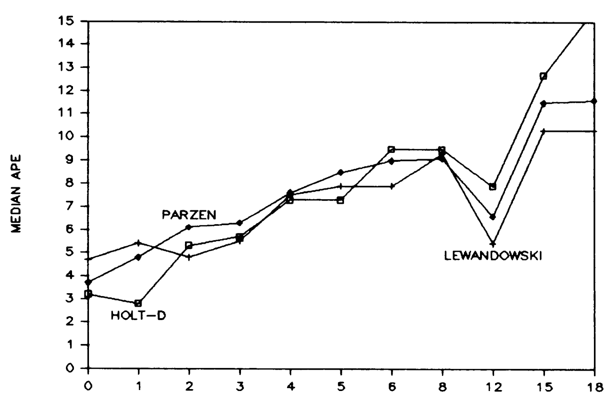

EXHIBIT 6. Median APE by Horizon (111).

4. Conclusions

Although a linear trend is commonly assumed in time series forecasting, empirical research shows this is a reasonable assumption only at short horizons. As the horizon increases, a linear trend frequently overshoots the data.

One alternative approach to trend extrapolation is Lewandowski's FORSYS system, which damps the trend in every time series according to the level of noise. This strategy appears to be overly conservative. When the time series actually contains a strong trend, Lewandowski's forecasts track well below the data at long horizons. Another difficulty with Lewandowski is that the method achieves good long-range performance at some cost in short-range accuracy.

Another alternative is the Parzen methodology, which may be the most robust approach to time series forecasting reported to date. However, Parzen may be too complex for use in large forecasting systems.

A third alternative is to modify Holt's linear exponential smoothing model with an AD-parameter. The result is a simple model structure which includes a variety of useful special cases. If the trend in the data is erratic, the model is based on the first differences of the data. The AD-parameter is fitted at a value less than one and damps the trend. When the trend is persistent, the model is based on second differences. The AD-parameter is fitted at a value near one and the model behaves much like Holt's linear model. On average, the model improves long-range forecast accuracy compared to models which assume a linear trend. This is achieved at no apparent cost in short-range performance. The model also compares favorably with Lewandowski and Parzen and should be suitable for routine use in large forecasting systems, such as in inventory control.[1]

Appendix—Computational Details

Initial forecasts for both the Holt-L and Holt-D models were computed from a simple linear regression on time, with the beginning level equal to the intercept and the trend equal to the slope. In the Makridakis study, initial forecasts were selected by backcasting, starting with the level equal to the last observation and the trend equal to the average difference between the last four observations. This procedure frequently resulted in negative initial forecasts which in turn distorted the model-fitting process. Exhibits 7 and 8 compare Makridakis' results to the Holt-L and Holt-D models.

Exhibit 9 shows the effects of parameter choice on the Holt-D model. The mean APE for the three regions

EXHIBIT 7. APE Comparisons (1,001).

Forecast Horizon	Mean APE			Median APE		
	Holt-D	Holt-L	Mak. Holt	Holt-D	Holt-L	Mak. Holt
1	8.3	8.7	8.7	4.2	4.5	4.5
2	10.8	10.9	11.0	5.3	5.5	5.3
3	12.1	12.8	13.3	5.9	6.1	5.6
4	13.0	14.1	15.2	7.2	6.9	7.3
5	15.7	16.9	19.1	8.1	8.2	8.1
6	17.9	19.7	21.6	9.3	9.1	9.2
8	17.7	20.0	24.8	9.0	9.5	9.8
12	16.7	19.9	23.9	9.3	10.2	9.9
15	21.0	23.5	33.7	11.6	13.2	12.2
18	21.7	24.8	48.3	11.9	13.0	13.6
Overall	16.2	18.1	22.9	8.4	8.8	8.8

[1] This research was supported by the Navy Regional Data Automation Center, Norfolk, Virginia and by the Navy Fleet Material Support Office, Mechanicsburg, Pennsylvania. The second author is also pleased to acknowledge the support of a National Research Council Associateship at the Naval Postgraduate School where his share of this work was carried out.

EXHIBIT 8. APE Comparisons (111).

Forecast Horizon	Mean APE Holt-D	Mean APE Holt-L	Mean APE Mak. Holt	Median APE Holt-D	Median APE Holt-L	Median APE Mak. Holt
1	7.6	8.0	7.9	2.8	3.2	3.4
2	9.7	10.8	10.5	5.3	5.5	5.0
3	11.5	12.9	13.2	5.7	6.9	6.4
4	13.1	14.6	15.1	7.3	7.5	8.1
5	14.5	17.1	17.3	7.3	7.4	7.4
6	15.9	19.1	19.0	9.5	8.7	9.0
8	16.6	21.0	23.1	9.5	11.1	10.6
12	13.6	19.0	16.5	7.9	7.5	7.4
15	29.0	31.9	35.6	12.7	14.3	12.4
18	29.5	32.5	35.2	15.5	13.6	15.0
Overall	16.1	18.9	19.7	8.2	8.6	8.7

EXHIBIT 9. Effects of Parameter Choice on the Holt-D Model (111).

Forecast Horizon	Mean APE Number of Parameters 3	2	2(DLS)	Median APE Number of Parameters 3	2	2(DLS)
1	7.6	7.6	7.8	2.8	2.9	3.0
2	9.7	9.7	9.7	5.3	4.6	5.2
3	11.5	11.3	11.6	5.7	5.2	5.5
4	13.1	13.1	13.2	7.3	7.6	8.0
5	14.5	15.0	15.0	7.3	7.8	7.4
6	15.9	16.4	16.7	9.5	9.8	9.2
8	16.6	17.6	18.1	9.5	9.9	11.2
12	13.6	14.1	15.0	7.9	7.6	7.7
15	29.0	30.4	32.4	12.7	12.8	12.9
18	29.5	30.4	31.8	15.5	15.6	15.4
Overall	16.1	16.5	17.0	8.2	8.5	8.6

of choice (see Exhibit 1) was about the same through 6 steps ahead. Thereafter, the three-parameter model was more accurate, although the differences were small.

Parameters in the Holt-L and Holt-D models were selected by a grid search. To replicate Exhibits 7–9, the following routine should be followed. For Holt-L, compute the fitted one-step-ahead MSE for four combinations of parameters (0.33 and 0.67 for α and γ). Next, compute the MSE at points ± 0.17 around the best combination and change the parameters if a better MSE is found. Continue the search from this point with progressively smaller values ($\pm 0.08, 0.04, 0.02, 0.015, 0.005$) until the change in MSE is less than 0.001. Note that the parameters are allowed to reach 0 or 1. The same procedure should be used for Holt-D except that eight combinations of parameters are needed initially in the three-parameter version.

The computer program used in this research was written in BASIC for the IBM Personal Computer. Run time for the Holt-D model with a compiled program was about 40 minutes for the 111 series and about 5 hours for the 1,001. The times are for the standard 8088 CPU with the data stored on floppy diskettes. The 1,001 series are available on tape from the International Institute of Forecasters (IIF), c/o Faculty of Management, McGill University, Montreal, Canada H3A 1G5. Diskettes (formatted with IBM PC DOS 1.1) of the time series were contributed to the IIF and may be obtained from that organization.

References

BROWN, R., *Smoothing, Forecasting, and Prediction of Discrete Time Series*, Prentice-Hall, Englewood Cliffs, N.J., 1963.

FILDES, R., "An Evaluation of Bayesian Forecasting," *J. Forecasting*, 2 (1983), 137–150.

GARDNER, E., "Exponential Smoothing: The State of the Art" (with commentary), *J. Forecasting*, 4 (1985), 1–38.

GILCHRIST, W., *Statistical Forecasting*, Wiley, London, 1976.

HOLT, C. ET AL., *Planning Production, Inventories, and Work Force*, Prentice-Hall, Englewood Cliffs, N.J., 1960.

LEWANDOWSKI, R., "Sales Forecasting by FORSYS," *J. Forecasting*, 1 (1982), 205–214.

MAKRIDAKIS, S. ET AL., "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," *J. Forecasting*, 1 (1982), 111–153.

MCKENZIE, E., "An Analysis of General Exponential Smoothing," *Oper. Res.*, 24 (1976), 131–140.

——— "General Exponential Smoothing and the Equivalent ARMA Process," *J. Forecasting*, 3 (1984), 333–344.

PARZEN, E., "Time Series and Whitening Filter Estimation," *TIMS Studies in the Management Sciences*, 12 (1979), 149–165.

——— "ARARMA Models for Time Series Analysis and Forecasting," *J. Forecasting*, 1 (1982), 67–82.

ROBERTS, S., "A General Class of Holt–Winters Type Forecasting Models," *Management Sci.*, 28 (1982), 808–820.