

## **Exponential Smoothing with a Damped Multiplicative Trend**

James W. Taylor

*International Journal of Forecasting*, 2003, Vol. 19, pp. 715-725.

Address for Correspondence:

James W. Taylor  
Saïd Business School  
University of Oxford  
Park End Street  
Oxford OX1 1HP, UK

Tel: +44 (0)1865 288927  
Fax: +44 (0)1865 288805  
Email: [james.taylor@sbs.ox.ac.uk](mailto:james.taylor@sbs.ox.ac.uk)

## **Exponential Smoothing with a Damped Multiplicative Trend**

## **Exponential Smoothing with a Damped Multiplicative Trend**

### **Abstract**

Multiplicative trend exponential smoothing has received very little attention in the literature. It involves modelling the local slope by smoothing successive ratios of the local level, and this leads to a forecast function that is the product of level and growth rate. By contrast, the popular Holt method uses an additive trend formulation. It has been argued that more real series have multiplicative trends than additive. However, even if this is true, it seems likely that the more conservative forecast function of the Holt method will be more robust when applied in an automated way to a large batch of series with different types of trend. In view of the improvements in accuracy seen in dampening the Holt method, in this paper, we investigate a new damped multiplicative trend approach. An empirical study, using the monthly time series from the M3-Competition, gave encouraging results for the new approach at a range of forecast horizons, when compared to the established exponential smoothing methods.

*Key words:* Damped Trend Exponential Smoothing; Pegels Classification; Multiplicative Trend.

## 1. Introduction

The robustness and accuracy of exponential smoothing forecasting has led to its widespread use in applications where a large number of series necessitates an automated procedure, such as inventory control. Although Holt's method has tended to be the most popular approach for trending series, its linear forecast function has been criticised for tending to overshoot the data beyond the short-term. Gardner and McKenzie (1985) address this problem by including an extra parameter in Holt's method to dampen the projected trend. Empirical studies show that the damped method tends to offer improvements in accuracy (e.g. Makridakis *et al.*, 1993; Makridakis and Hibon, 2000). If the extra parameter is permitted to be greater than one, the method can also produce an exponential forecast function, which would seem to be useful for series with exponential trend. An alternative method for dealing with such series is multiplicative trend exponential smoothing, which is described by Pegels (1969) and Hyndman *et al.* (2002). It involves modelling the local growth rate by smoothing successive ratios of the local level, and leads naturally to a forecast function that is the product of the growth rate and level. The trend is thus modelled in a multiplicative way. By contrast, all the established exponential smoothing methods assume an additive trend.

The multiplicative trend method has received very little attention in the literature. This is a little surprising, given the preference for multiplicative, rather than additive, modelling of seasonality in the Holt-Winters approach. Pegels (1969) suggests that more real series have multiplicative trends than additive. Regardless of whether this is true, it seems likely that the more conservative forecast function of Holt's method will be more robust when applied in an automated way to a large batch of series with different types of trend. Motivated by the improvements seen in dampening the Holt method, in this paper, we introduce a damped Pegels exponential smoothing method. The method has the appeal of

modelling trends in a multiplicative fashion but includes a dampening term, which should lead to more robust forecasting performance.

In Section 2 of the paper, we review the literature on exponential smoothing methods for additive and multiplicative trends, and in particular the methods of Holt and Pegels. Section 3 introduces our new damped multiplicative trend exponential smoothing method. In Section 4, we use a large data set of real time series to compare the new method with existing approaches. The final section provides a summary and concluding comments.

## 2. Exponential Smoothing with Additive and Multiplicative Trends

Pegels' (1969) classification of exponential smoothing methods includes nine different methods. Each method is classified as being suitable for series with either constant level, additive trend or multiplicative trend, and with either no seasonality, additive seasonality or multiplicative seasonality. Hyndman *et al.* (2002) have recently extended this taxonomy to include damped additive trend with either no seasonality, additive seasonality or multiplicative seasonality. From among the various methods, simple, Holt's, damped Holt's and Holt-Winters' exponential smoothing have been very popular with practitioners and researchers. The Holt-Winters method with multiplicative seasonality has been widely used. By contrast, with the exceptions of Pegels (1969), Makridakis *et al.* (1998) and Hyndman *et al.* (2002), the multiplicative trend methods have been very largely ignored. Indeed, we are not aware of any software packages that include multiplicative trend exponential smoothing. Later in this section, we return to the multiplicative trend formulation but let us first consider the standard Holt additive trend method, which is given in expressions (1)-(3). The method estimates the local growth,  $T_t$ , by smoothing successive differences,  $(S_t - S_{t-1})$ , of the local level,  $S_t$ . The forecast function is the sum of level and projected growth.

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

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

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

where  $X_t$  is the actual observation,  $\hat{X}_t(m)$  is the  $m$ -step-ahead forecast, and  $\alpha$  and  $\gamma$  are smoothing parameters. Despite its popularity, empirical evidence has shown that the Holt linear forecast function tends to overestimate (Gardner and McKenzie, 1985). In view of this, Gardner and McKenzie (1985) describe how a dampening parameter,  $\phi$ , can be used within Holt's method to give more control over trend extrapolation. The damped Holt method is presented in expressions (4)-(6).

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

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

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

Gardner and McKenzie explain that 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 method is identical to the standard Holt method presented in expressions (1)-(3). If  $\phi = 0$ , the method is identical to standard simple exponential smoothing. If  $\phi > 1$ , the forecast function has an exponential trend. Interestingly, Hyndman *et al.* (2002) suggest a slightly different damped formulation where there is no dampening of the trend for the first forecast period.

Gardner and McKenzie write that  $\phi > 1$  is probably a dangerous option in an automatic forecasting procedure. However, Tashman and Kruk (1996) show that there can be value in allowing  $\phi > 1$ , if it is applied only to strongly trending series. Indeed, as the method has an exponential forecast function, it would seem to be suited to series with exponential trends. However, we feel that it is a clumsy way to handle such series as the local slope is rather unsatisfactorily modelled as an additive trend by smoothing successive differences of the local level. We feel that if the process is exponential, then an appropriate method would be one that actually models this sort of trend, such as the method of Pegels (1969), which is presented in expressions (7)-(9). The method models the local growth rate,  $R_t$ , by smoothing successive ratios,  $(S_t / S_{t-1})$ , of the local level  $S_t$ .

$$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} R_{t-1}) \quad (7)$$

$$R_t = \gamma (S_t / S_{t-1}) + (1 - \gamma) R_{t-1} \quad (8)$$

$$\hat{X}_t(m) = S_t R_t^m \quad (9)$$

The method can be described as modelling the trend in a multiplicative way because the forecasts are formed from the product of the level and growth rate. A logarithmic transformation is sometimes used to convert a multiplicative trend into an additive trend. The resultant forecasts must then be transformed back into the original units, along with prediction intervals. The Pegels method has the appeal of simplicity in that the transformation is avoided.

The fact that the Pegels method expresses the trend in a unitless form can be very useful. For example, consider the introduction of a new product that is a variation on an existing product. The two products would probably share similar behaviours and could be classified together. When the product is first introduced, a forecast method would have to be implemented starting with no data. To help initialise the method, it would be reasonable to use information about the aggregate product class. Although the sales volume of the new product may be quite different to other products in the same class, they may share the same behaviour in terms of trend. With the Pegels multiplicative trend formulation, the trend for the new product could be initialised using the trend from the application of the method to a more aggregated product class, since the trend is unitless. By contrast, it is not possible to use an aggregated trend from an additive trend model.

### **3. Damped Multiplicative Trend Exponential Smoothing**

Pegels (1969) suggests that his multiplicative trend method may be more useful than the Holt additive trend method, as multiplicative trends “appear more probable in real-life applications”. If he is correct, then his method should perform well relative to the standard methods, which all use an additive trend. However, an obvious reason for using Holt’s additive trend method in preference to Pegels’ multiplicative trend method is that the more conservative trend extrapolation, provided by the additive trend method, may be more robust when applied to

a variety of different series in a large-scale, automated forecasting application. In view of this, there may be value in including an extra parameter in the Pegels formulation to dampen the extrapolated trend, in an analogous fashion to the dampening parameter in the damped Holt method. Our new damped Pegels multiplicative trend method is given in expressions (10)-(12).

$$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} R_{t-1}^\phi) \quad (10)$$

$$R_t = \gamma (S_t / S_{t-1}) + (1 - \gamma) R_{t-1}^\phi \quad (11)$$

$$\hat{X}_t(m) = S_t R_t^{\sum_{i=1}^m \phi^i} \quad (12)$$

The forecast function in expression (12) is analogous to the forecast function for damped Holt's exponential smoothing. The growth rate undergoes a further dampening for each period into the future. Hence, the forecast function is derived as:

$$\hat{X}_t(m) = S_t \left( R_t^\phi R_t^{\phi^2} R_t^{\phi^3} \dots R_t^{\phi^m} \right) = S_t R_t^{\sum_{i=1}^m \phi^i}$$

If  $0 < \phi < 1$ , the multiplicative trend is damped and the forecasts approach an asymptote given by the horizontal straight line  $S_t R_t^{\phi/(1-\phi)}$ . If  $\phi = 1$ , the method is identical to the standard Pegels method presented in expressions (4)-(6). If  $\phi = 0$ , the method is identical to standard simple exponential smoothing. If  $\phi > 1$ , the forecast function has a multiplicative inflationary term which exponentially increases over time. Our preference is to use the formulation with  $0 < \phi < 1$  in order simply to provide a damped version of Pegels' method.

In Fig. 1, we show forecast profiles for the Holt method and for damped Holts with three different choices for the dampening parameter ( $0 < \phi < 1$ ). Decreasing the value of  $\phi$ , increases the degree of dampening but it does not alter the concave shape of the forecast profile. Fig. 2 presents forecast profiles for the Pegels method and for damped Pegels with three different choices for the dampening parameter ( $0 < \phi < 1$ ). As with the Holt method, decreasing  $\phi$ , increases the degree of dampening for the damped Pegels method. However, unlike the Holt method, the shape of the damped forecast profile varies. Although the profile



will eventually approach a horizontal asymptote, for relatively early forecast lead times, the profile can be convex, nearly linear or concave.

----- Figs. 1 and 2 -----

Although Hyndman *et al.* (2002) do consider trend dampening, their taxonomy of methods does not include a damped multiplicative trend. In Table 1, we present an extended version of their taxonomy, in which we use similar notation to theirs. The method N-N is simple exponential smoothing, A-N is Holt's, DA-N is damped Holt's, A-A and A-M are the Holt-Winters' seasonal methods, and M-N is Pegels' multiplicative trend method. M-A and M-M are seasonal versions of the multiplicative trend method. The bottom row of the table is the extension of the taxonomy of Hyndman *et al.* The method DM-N is damped Pegels, which is the focus of this paper. The methods DM-A and DM-M, also in the bottom row, are seasonal versions of damped Pegels. Although these seasonal versions are straightforward to formulate, we concentrate in this introductory paper on the non-seasonal method.

----- Table 1 -----

Many users of exponential smoothing methods find it easier to work with error-correction forms. For the damped Pegels method, the error-correction form is given by expressions (13)-(15):

$$S_t = S_{t-1} R_{t-1}^\phi + \alpha e_t \quad (13)$$

$$R_t = R_{t-1}^\phi + \gamma \alpha e_t / S_{t-1} \quad (14)$$

$$\hat{X}_t(m) = S_t R_t^{\sum_{i=1}^m \phi^i} \quad (15)$$

where  $e_t$  is the 1-step-ahead forecast error,  $e_t = X_t - S_{t-1} R_{t-1}^\phi$ .

An important issue for many forecasting applications is the estimation of prediction intervals to accompany point forecasts. For example, in inventory control, intervals enable the setting of appropriate levels of safety stock. Theoretical forecast error variance formulae are often derived for exponential smoothing methods by referring to the equivalent ARIMA model. However, there is no equivalent ARIMA model for either the Pegels or the new damped Pegels

methods. The lack of equivalent ARIMA models for various non-linear exponential smoothing methods has led to prediction intervals being based on the equivalent state-space model. Hyndman *et al.* (2001) derive theoretical forecast error variance formulae from the state-space models and Hyndman *et al.* (2002) generate prediction intervals by applying simulation to the models. In expressions (16)-(18), we present the state-space formulation for the damped Pegels method.

$$X_t = S_{t-1} R_{t-1}^\phi + \varepsilon_t \quad (16)$$

$$S_t = S_{t-1} R_{t-1}^\phi + \alpha \varepsilon_t \quad (17)$$

$$R_t = R_{t-1}^\phi + \gamma \alpha \varepsilon_t / S_{t-1} \quad (18)$$

where  $\varepsilon_t$  is a Gaussian white noise process. An alternative to the theoretical and model-based approaches is to use an empirical approach to estimate prediction intervals, such as that of Gardner (1988) or Taylor and Bunn (1999).

## 4. Empirical Results

### 4.1. Description of the Study

We carried out empirical analysis in order to address two main issues. Firstly, we wished to investigate whether the inclusion of a dampening term would lead to improvement in the accuracy of the Pegels method. Secondly, we wanted to compare the Pegels and damped Pegels methods with the established exponential smoothing methods.

The data used was the 1,428 monthly time series from the recent M3-Competition, which is described by Makridakis and Hibon (2000). The data set is a mixture of industry, demographic, meteorological, financial, microeconomic and macroeconomic series. The series vary in length from 48 to 126 with a median of 115. Since all the series consist of monthly observations and we wished largely to compare non-seasonal exponential smoothing methods, we deseasonalised the data prior to forecasting. We used the seasonal decomposition method based on ratio-to-moving averages, which was used in the M3-

Competition. By applying exponential smoothing methods in an automated way to a large number of series, we replicated common practice in inventory and productions management. We did not consider the quarterly or yearly series from the M3-Competition because automated forecasting procedures are rarely applied to data of such low frequencies. We should acknowledge at this stage that none of the 1,428 monthly series from the M3-Competition contain zeros. Multiplicative trend methods, such as Pegels and damped Pegels, are likely to be less suitable for series with embedded zeros (Hyndman *et al.*, 2002).

We produced forecasts using seven exponential smoothing methods. We derived parameter values by the common procedure of minimising the sum of squared one-step-ahead forecast errors, and we employed the constrained non-linear optimisation routine of the statistical programming language Gauss.

Method 1: Simple exponential smoothing with  $\alpha$  restricted to lie between zero and one.

Method 2: Holt's with  $\alpha$  and  $\gamma$  restricted to lie between zero and one.

Method 3: Damped Holt's with  $\alpha$ ,  $\gamma$  and  $\phi$  restricted to lie between zero and one.

Method 4: Damped Holt's with  $\alpha$  and  $\gamma$  restricted to lie between zero and one, and  $\phi$  restricted to lie between zero and two. We refer to this as the 'generalised Holt' method.

Method 5: Pegels' with  $\alpha$  and  $\gamma$  restricted to lie between zero and one.

Method 6: Damped Pegels' with  $\alpha$ ,  $\gamma$  and  $\phi$  restricted to lie between zero and one.

Method 7: Holt-Winters for multiplicative seasonality with parameters restricted to lie between zero and one. As this method is suited to seasonal series, the data was not deseasonalised beforehand.

Chatfield and Yar (1988) describe how the use of different approaches for the derivation of initial values for the smoothed level, trend and seasonal components can give rise to substantially different optimised parameter values, which can lead to substantially different forecasts. Although there are theoretical arguments in favour of backcasting (Ledolter and Abraham, 1984), it frequently gave poor initial values in the M-Competition

(Makridakis *et al.*, 1982). Another approach is to use a simple linear regression on time to produce initial slope and trend (Gardner and McKenzie, 1985). Our experience is that this can give poor results for series that are notably different to a simple linear trend.

We feel that a more robust approach is to use simple averages of the first few data observations to calculate initial smoothed components, as in the work of Gardner (1999). We implemented the procedure of Williams and Miller (1999) who adapt the method of simple averages for monthly data. For all the exponential smoothing methods, except Pegels' and damped Pegels', the initial growth,  $T_0$ , was chosen as the average of (1) one-twelfth the difference between the mean of the first 12 and second 12 deseasonalised observations, and (2) the average of the first differences for the first 24 deseasonalised observations. The initial level,  $S_0$ , was chosen as the mean of the first 24 deseasonalised observations minus 12.5 times the initial growth. We used the same initial level,  $S_0$ , for the Pegels and damped Pegels methods and we set the initial growth rate,  $R_0$ , equal to  $(S_0+T_0)/S_0$ .

The initial value for each of the smoothed seasonal indices in the Holt-Winters method was set as the average of the ratios of actual observation to 12-point centred moving average taken from the corresponding month in each of the first two years of the time series. The same ratios were used to deseasonalise the data for the nonseasonal forecasting methods.

Using each of the seven methods, we produced forecasts for lead times from one to 18 for each of the 1,428 series. We chose these lead times because they had been used with the monthly data in both the M-Competition and M3-Competition.

#### *4.2. Results for All 1,428 Series*

Table 2 summarises the ex ante symmetric mean absolute percentage error (SMAPE) for each of the seven exponential smoothing methods applied to all 1,428 series. Although the SMAPE has received some criticism (see Goodwin and Lawton, 1999), it is the one summary error measure reported in numerical detail by Makridakis and Hibon (2000) in their presentation

of the results of the M3-Competition. They describe how the measure has the advantage over the more traditional MAPE of avoiding large errors when the actual,  $x_i$ , is close to zero and large differences between the absolute percentage error when  $x_i$  is greater than the forecast,  $f_i$ , and when  $f_i$  is greater than  $x_i$ .

$$SMAPE = \sum_i \frac{|x_i - f_i|}{(x_i + f_i)/2}$$

For simplicity, we do not show the SMAPE results for each of the 18 forecast horizons. Instead, Table 2 displays the average SMAPE for forecast horizons one to six (the short-term), for horizons seven to 12 (the medium-term), for horizons 13 to 18 (the long-term) and for all 18 horizons. The best results for each forecast horizon category are indicated in bold. The Holt-Winters method is the worst of all seven methods for all three forecast horizon categories. For Holt's, generalised Holt's and Pegels', the results are similar. Given that the Pegels and generalised Holt methods are suited to strongly trending series, it is not unexpected to see that they perform relatively poorly beyond the short-term, as the 1,428 series have a variety of different strength trends. Perhaps it is a little surprising to see that Holt's is slightly outperformed by Pegels' beyond the short-term. This suggests that the assumption of multiplicative trend is not as dangerous as one might have surmised. Holt's is also outperformed by generalised Holt's. This can, at least partially, be explained by the fact that the optimised  $\phi$  parameter for generalised Holt's was greater than one for only 203 of the 1,428 series, indicating that the method was the same as damped Holt's for a large proportion of the series. In the next subsection, we return to the issue of how well the methods perform on the 203 series for which the optimised generalised Holt's  $\phi$  parameter was greater than one. Table 2 shows that the best methods were simple exponential smoothing, damped Holt's and damped Pegels'. Interestingly, damped Pegels' outperforms all other methods for all three lead time categories with the improvement increasing with the lead time.

----- Table 2 -----

The median absolute percentage error (MedAPE) is another measure reported by Makridakis and Hibon (2000) in their presentation of the results of the M3-Competition. Here, the percentage error is defined simply as the ratio of error to actual. In the forecasting literature, the median is often preferred to the mean APE because it is more robust to outliers, and because the distribution of absolute percentage errors is often skewed. Table 3 summarises the median APE results for the seven exponential smoothing methods applied to all 1,428 series. Using the MedAPE criterion, the relative performances of the methods are broadly in line with those shown in Table 2 for the SMAPE measure. However, there are two notable differences. First, the Holt-Winters method is more competitive when judged by the MedAPE and, second, the simple exponential smoothing method was the weakest method at all horizons according to the MedAPE. The MedAPE results in Table 3 show Pegels' as being a little more accurate than Holt's. Both of these methods are improved by including the dampening parameter, and, of the two damped methods, damped Pegels' performed slightly better than damped Holt's.

----- Table 3 -----

The fact that damped Pegels' was able to match and, indeed, slightly outperform damped Holt's is very encouraging for the new method, given that damped Holt's is so widely used and respected. The similarity of the accuracy summary measures, particularly the MedAPE results, for damped Holt's and damped Pegels' prompted us to investigate further the estimated parameters and resultant forecasts. Comparison of the forecasts produced by damped Holt's and damped Pegels' indicated that, for many series, the forecasts were very similar. For these series, the change in the level of the series (or growth) tended to be much smaller than the level of the series. In the appendix, we show analytically that when this is the case, the damped Pegels smoothing equations and forecast function, given in expressions (10)-(12), approximate those of the damped Holt method, given in expressions (4)-(6). Intuitively, we feel that it is reasonable that the two methods are equivalent when the growth is much smaller than the level of the series; in this situation, it is probably difficult to decipher whether the trend is additive or

multiplicative. We also found that for series with growth much smaller than the level, simple exponential smoothing produced forecasts reasonably similar to the damped Holt and damped Pegels methods. This is not surprising given that damped Holt's and damped Pegels' are equivalent to simple exponential smoothing when applied to series with no growth.

#### *4.3. Results for the Series for which $\phi > 1$ in Generalised Holt's*

In large-scale forecasting applications, which require an automated forecasting procedure, such as inventory control, it is common to apply the same exponential smoothing method to each series. In view of this, there is clear value in empirical studies that compare the performance of a number of methods across a wide variety of series, such as our analysis of the 1,428 monthly series. However, certain methods, that have little value when applied across a variety of different series, can be very useful when selectively applied to appropriate series. In order to investigate further the standard Pegels and new damped Pegels methods, we felt that it would be interesting to analyse their respective performances for series for which they might be considered particularly suitable. Pegels (1969) and Gardner (1985) clearly feel that the Pegels method is suitable for series with strong trend.

The method selection protocol of Gardner and McKenzie (1988) provides a simple procedure for categorising the strength of the trend component in a series. It involves calculating the variance of the original series, the variance of the series differenced once and the variance of the series differenced twice. If the variance of the original series is the least of the three, the series is considered to have no trend and simple exponential smoothing is recommended; if it is least for the series differenced just once, the series is considered to have moderate trend and damped Holt's is recommended; and, if it is least for the series differenced twice, the series is considered to have strong trend in which case Holt's should be used. However, Tashman and Kruk (1996) found that the performance of damped Holt's was best for the series classified by the Gardner-McKenzie procedure as strongly trending. They also point out that the procedure

does not make a distinction between a linear and exponential trend. Unfortunately, therefore, we cannot expect this trend classification approach to be able to identify strongly trending series for which the Pegels and damped Pegels methods are more suitable than damped Holt's. Indeed, we found that the results for damped Holt's and damped Pegels' were similar for the 63 series, out of the 1,428 monthly series, which were classified by the Gardner-McKenzie procedure as strongly trending. We also found that some of the 1,428 series, which are clearly trending, are classified by the procedure as stationary. This seems to be the result of a reasonably large degree of variability in the original series. We agree with Tashman and Kruk that much work remains to be done on method selection rules for exponential smoothing.

In the previous subsection, we described how the generalised Holt's  $\phi$  parameter was larger than one for 203 of the 1,428 series. For these series, the constrained optimisation routine derived a value of one for the damped Holt's  $\phi$  parameter. A value greater than one for the generalised Holt's  $\phi$  parameter suggests that damped Holt's will not be able to satisfactorily forecast the trend in these series and that other forecasting methods may be preferable. We investigated whether the Pegels and damped Pegels methods are more suitable for these series. Tables 4 and 5 summarise the SMAPE and MedAPE measures, respectively, for the seven exponential smoothing methods applied to the subset of 203 series. The most alarming finding in Tables 4 and 5 is that generalised Holt's performs extremely poorly. Clearly, in Tables 2 and 3, the success of this method for the other 1,225 series hid the poor performance resulting when the  $\phi$  parameter was greater than one. According to the SMAPE results in Table 4, Pegels' is a little better than Holt's but the ranking is reversed for the MedAPE results in Table 5. Comparing the damped methods, we find that damped Pegels' comfortably outperformed damped Holt's for all forecast horizon categories according to the SMAPE in Table 4. The MedAPE results in Table 5 are less impressive but damped Pegels' was noticeably better for the long-term.

----- Tables 4 and 5 -----



## 5. Summary and Conclusions

In this paper, we have introduced a new damped exponential smoothing method. The method follows the multiplicative trend formulation of Pegels (1969) but includes an extra parameter to dampen the projected trend. We used the 1,428 monthly time series from the M3-Competition to compare the method to the standard Pegels method and the established exponential smoothing methods. The performance of the standard Pegels method was similar to that of the standard Holt method. This is an interesting result as there have been no previous empirical studies comparing the post-sample forecasting accuracy of the standard Pegels method with that of other exponential smoothing methods. It suggests that the assumption of a multiplicative trend is not as dangerous as might have been expected. We found that the damped Pegels method comfortably outperformed the standard Pegels method at all forecast horizons. Furthermore, the new damped version of the method also slightly outperformed the popular damped Holt method.

The generalised Holt formulation is identical to damped Holt's except that the  $\phi$  parameter is permitted to take values greater than one. This occurred for 203 of the 1,428 series. An optimised value greater than one for the generalised Holt's  $\phi$  parameter suggests that damped Holt's will not be able to satisfactorily forecast the trend in these series and that other forecasting methods may be preferable. We investigated whether the multiplicative trend formulation of the standard Pegels and damped Pegels methods is preferable for these series. We compared the accuracy of these methods to the established exponential smoothing methods for the subset of 203 series. The standard Pegels method outperformed standard Holt's according to the Symmetric Mean APE summary error measure but not according to the Median APE. Of all the seven methods considered, the best results were achieved for both error measures using the damped Pegels method. This suggests that the damped Pegels method could at least be useful as an alternative to the popular and successful damped Holt method for series for which the latter seems unsuitable. In view of this, there would seem to be strong appeal in including the damped

Pegels method as a candidate in automated method selection approaches, such as that of Hyndman *et al.* (2002). In conclusion, we feel that the results for the 1,428 series and for the subset of 203 suggest that the new damped Pegels method is a considerable improvement on the standard Pegels method, and that it is a potentially useful alternative to the established exponential smoothing methods.

### Acknowledgements

We would like to acknowledge the helpful comments of two anonymous referees.

### Appendix

In this appendix, we show that if the growth,  $T_t$ , in the series is much less than the level of the series,  $S_t$ , the forecasts from the damped Holt and damped Pegels methods will be very similar. If we write the growth rate as  $R_t = 1 + \frac{T_t}{S_t}$ , expression (10) in the damped Pegels formulation becomes

$$S_t = \alpha X_t + (1 - \alpha) S_{t-1} \left( 1 + \frac{T_{t-1}}{S_{t-1}} \right)^\phi$$

If the growth,  $T_{t-1}$ , is less than the level,  $S_{t-1}$ , we can expand this expression to give

$$\begin{aligned} S_t &= \alpha X_t + (1 - \alpha) S_{t-1} \left( 1 + \phi \left( \frac{T_{t-1}}{S_{t-1}} \right) + \frac{1}{2!} \phi(\phi - 1) \left( \frac{T_{t-1}}{S_{t-1}} \right)^2 + \frac{1}{3!} \phi(\phi - 1)(\phi - 2) \left( \frac{T_{t-1}}{S_{t-1}} \right)^3 + \dots \right) \\ &= \alpha X_t + (1 - \alpha) (S_{t-1} + \phi T_{t-1}) + \frac{1}{2} (1 - \alpha) \phi(\phi - 1) T_{t-1}^2 \left( \left( \frac{T_{t-1}}{S_{t-1}} \right) + \frac{1}{3} (\phi - 2) \left( \frac{T_{t-1}}{S_{t-1}} \right)^2 + \dots \right) \end{aligned}$$

Comparing this with the damped Holt smoothing equation for the local level in expression (4), we can see that the two expressions are approximately the same if  $T_{t-1} \ll S_{t-1}$ .

If we write  $R_t = 1 + \frac{T_t}{S_t}$ , expression (11) of the damped Pegels formulation becomes

$$1 + \frac{T_t}{S_t} = \gamma \left( \frac{S_t}{S_{t-1}} \right) + (1-\gamma) \left( 1 + \frac{T_{t-1}}{S_{t-1}} \right)^\phi$$

Rearranging this, we get

$$T_t = \frac{S_t}{S_{t-1}} \left( \gamma S_t - S_{t-1} + (1-\gamma) S_{t-1} \left( 1 + \frac{T_{t-1}}{S_{t-1}} \right)^\phi \right)$$

If the growth,  $T_{t-1}$ , is less than the level,  $S_{t-1}$ , we can expand this expression to give

$$T_t = \frac{S_t}{S_{t-1}} \left( \gamma(S_t - S_{t-1}) + (1-\gamma)\phi T_{t-1} + \frac{1}{2}(1-\gamma)\phi(\phi-1)T_{t-1} \left( \left( \frac{T_{t-1}}{S_{t-1}} \right) + \frac{1}{3}(\phi-2) \left( \frac{T_{t-1}}{S_{t-1}} \right)^2 + \dots \right) \right)$$

Comparing this with the damped Holt smoothing equation for the local growth in expression

(5), we can see that the two expressions are approximately the same if  $T_{t-1} \ll S_{t-1}$ . Note that the

condition  $T_{t-1} \ll S_{t-1}$  implies that  $S_t \approx S_{t-1}$ .

Substituting  $R_t = 1 + \frac{T_t}{S_t}$  in expression (12) of the damped Pegels formulation, we get

$$\hat{X}_t(m) = S_t \left( 1 + \frac{T_t}{S_t} \right)^{\sum_{i=1}^m \phi^i}$$

If the growth,  $T_t$ , is less than the level,  $S_t$ , we can expand this expression to give

$$\hat{X}_t(m) = S_t + T_t \sum_{i=1}^m \phi^i + \frac{1}{2!} \sum_{i=1}^m \phi^i (\sum_{i=1}^m \phi^i - 1) T_t \left( \frac{T_t}{S_t} \right) + \frac{1}{3!} \sum_{i=1}^m \phi^i (\sum_{i=1}^m \phi^i - 1) (\sum_{i=1}^m \phi^i - 2) T_t \left( \frac{T_t}{S_t} \right)^2 + \dots$$

Comparing this with the damped Holt forecast function in expression (6), we can see that the

two expressions are approximately the same if  $T_t \ll S_t$ .

## References

- Chatfield, C., & Yar, M. (1988). Holt-Winters forecasting: Some practical issues, *The Statistician*, 37, 129-140.
- Gardner, E.S., Jr. (1985). Exponential smoothing: The state of the art, *Journal of Forecasting*, 4, 1-28.
- Gardner, E.S., Jr. (1988). A simple method of computing prediction intervals for time-series forecasts, *Management Science*, 34, 541-546.
- Gardner, E.S., Jr. (1999). Rule-based forecasting vs. damped trend exponential smoothing, *Management Science*, 45, 1169-1176.
- Gardner, E.S., Jr., & McKenzie, E. (1985). Forecasting trends in time series, *Management Science*, 31, 1237-1246.
- Gardner, E.S., Jr., & McKenzie, E. (1988). Model identification in exponential smoothing, *Journal of the Operational Research Society*, 39, 863-867.
- Goodwin, P., & Lawton, R. (1999). On the asymmetry of the symmetric MAPE, *International Journal of Forecasting*, 15, 405-408.
- Hyndman, R.J., Koehler, A.B., Ord, J.K., & Snyder, R.D. (2001). Prediction intervals for exponential smoothing state space models, Working Paper, Department of Econometrics and Business Statistics, Monash University, Australia.
- Hyndman, R.J., Koehler, A.B., Snyder, R.D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, 18, 439-454.
- Ledolter, J., & Abraham, B. (1984). Some comments on the initialization of exponential smoothing, *Journal of Forecasting*, 3, 79-84.
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications, *International Journal of Forecasting*, 16, 451-476.
- Makridakis, S, Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition, *Journal of Forecasting*, 1, 111-153.
- Makridakis, S, Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L.F. (1993). The M2-Competition: A real-time judgementally based forecasting study, *International Journal of Forecasting*, 9, 5-22.
- Makridakis, S., Wheelwright, S.C. & Hyndman, R.J. (1998). *Forecasting Methods and Applications*, 3rd edition, New York: Wiley.
- Pegels, C.C. (1969). Exponential forecasting: Some new variations, *Management Science*, 15, 311-315.

Tashman, L.J., & Kruk, J.M. (1996). The use of protocols to select exponential smoothing procedures: a reconsideration of forecasting competitions, *International Journal of Forecasting*, 12, 235-253.

Taylor, J.W., & Bunn, D.W. (1999). A quantile regression approach to generating prediction intervals, *Management Science*, 45, 225-237.

Williams, D.W., & Miller, D. (1999). Level-adjusted exponential smoothing for modeling planned discontinuities, *International Journal of Forecasting*, 15, 273-289.

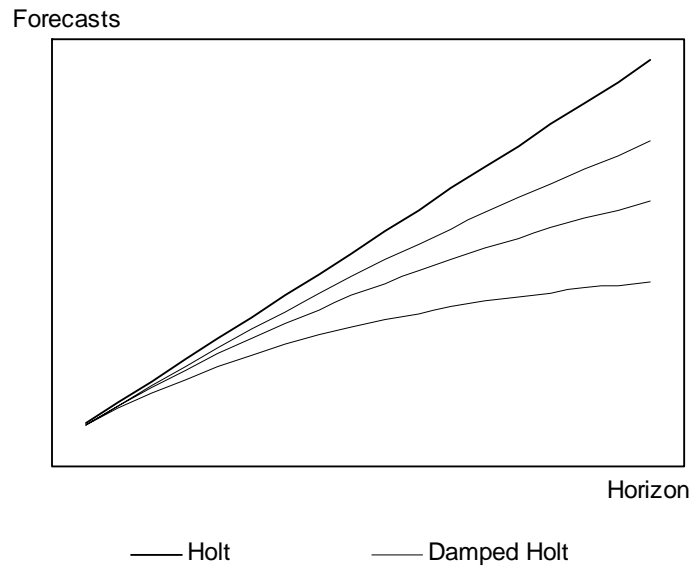


Fig. 1. Forecast profiles for the Holt method and three damped Holt methods with differing dampening parameters.

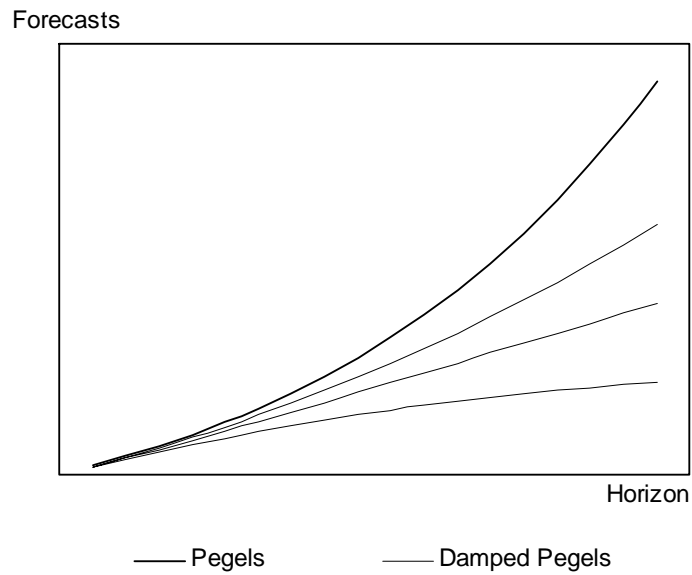


Fig. 2. Forecast profiles for the Pegels method and three damped Pegels methods with differing dampening parameters.

Table 1  
 Classification of exponential smoothing methods (adapted from Hyndman *et al.*, 2002)

<b>Trend Component</b>	<b>Seasonal Component</b>		
	None (N)	Additive (A)	Multiplicative (M)
None (N)	N-N	N-A	N-M
Additive (A)	A-N	A-A	A-M
Damped Additive (DA)	DA-N	DA-A	DA-M
Multiplicative (M)	M-N	M-A	M-M
Damped Multiplicative (DM)	DM-N	DM-A	DM-M

Table 2  
Symmetric Mean APE for the 1,428 Monthly Series from the M3-Competition

Method	Forecasting Horizon			Overall
	1-6	7-12	13-18	
Simple	12.5	14.1	17.4	14.7
Holt	12.8	15.3	20.2	16.1
Damped Holt	12.4	14.2	17.7	14.7
Generalised Holt	12.8	15.2	19.3	15.7
Pegels	12.8	15.1	19.5	15.8
Damped Pegels	<b>12.3</b>	<b>13.8</b>	<b>17.1</b>	<b>14.4</b>
Holt-Winters	13.3	15.6	20.4	16.4

Table 3  
Median APE for the 1,428 Monthly Time Series from the M3-Competition

Method	Forecasting Horizon			Overall
	1-6	7-12	13-18	
Simple	5.3	7.4	9.8	7.5
Holt	5.1	6.9	9.5	7.2
Damped Holt	<b>5.0</b>	<b>6.8</b>	8.9	6.9
Generalised Holt	5.1	7.1	9.3	7.2
Pegels	<b>5.0</b>	<b>6.8</b>	9.4	7.1
Damped Pegels	<b>5.0</b>	<b>6.8</b>	<b>8.8</b>	<b>6.8</b>
Holt-Winters	5.2	7.0	9.4	7.2



Table 4

Symmetric Mean APE for the 203 Monthly Series from the M3-Competition for which  $\phi > 1$  in the Generalised Holt Method

Method	Forecasting Horizon			Overall
	1-6	7-12	13-18	
Simple	13.8	15.7	18.0	15.8
Holt	14.1	16.9	21.1	17.4
Damped Holt	13.9	16.6	20.5	17.0
Generalised Holt	16.8	23.9	32.0	24.2
Pegels	14.2	16.5	20.1	16.9
Damped Pegels	<b>13.7</b>	<b>15.3</b>	<b>17.8</b>	<b>15.6</b>
Holt-Winters	<b>13.7</b>	15.4	19.5	16.2

Table 5

Median APE for the 203 Monthly Series from the M3-Competition for which  $\phi > 1$  in the Generalised Holt Method

Method	Forecasting Horizon			Overall
	1-6	7-12	13-18	
Simple	5.4	8.9	11.4	8.5
Holt	<b>4.8</b>	7.1	9.2	7.0
Damped Holt	<b>4.8</b>	6.9	8.9	6.9
Generalised Holt	5.8	9.7	14.3	9.9
Pegels	<b>4.8</b>	7.3	9.7	7.3
Damped Pegels	4.9	<b>6.8</b>	<b>8.5</b>	<b>6.7</b>
Holt-Winters	5.2	7.2	<b>8.5</b>	7.0