Open access • Journal Article • DOI:10.1002/FOR.918

Smooth Transition Exponential Smoothing — Source link

James Taylor

Institutions: University of Oxford

Published on: 01 Sep 2004 - Journal of Forecasting (John Wiley & Sons, Ltd.)

Topics: Smoothing, Exponential smoothing, Additive smoothing, Outlier and Logistic function

Smooth Transition Exponential Smoothing

James W. Taylor

Saïd Business School

University of Oxford

Journal of Forecasting, 2004, Vol. 23, pp. 385-394.

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

Smooth Transition Exponential Smoothing

SMOOTH TRANSITION EXPONENTIAL SMOOTHING

Abstract

Adaptive exponential smoothing methods allow a smoothing parameter to change over time, in order to adapt to changes in the characteristics of the time series. However, these methods have tended to produce unstable forecasts and have performed poorly in empirical studies. This paper presents a new adaptive method, which enables a smoothing parameter to be modelled as a logistic function of a user-specified variable. The approach is analogous to that used to model the time-varying parameter in smooth transition models. Using simulated data, we show that the new approach has the potential to outperform existing adaptive methods and constant parameter methods when the estimation and evaluation samples both contain a level shift or both contain an outlier. An empirical study, using the monthly time series from the M3-Competition, gave encouraging results for the new approach.

Keywords: Adaptive exponential smoothing; Smooth transition; Level shifts; Outliers

INTRODUCTION

Exponential smoothing is a simple and pragmatic approach to forecasting, whereby the forecast is constructed from an exponentially weighted average of past observations. The literature generally recommends that the smoothing parameters should be estimated from the data, usually by minimising the sum of *ex post* 1-step-ahead forecast errors (Gardner, 1985). Some researchers have argued that the parameters should be allowed to change over time, in order to adapt to the latest characteristics of the time series. For example, if there has been a level shift in the series, the exponentially weighted average should adjust so that an even greater weight is put on the most recent observation. A variety of adaptive exponential smoothing methods have been developed to deal with this problem. However, these methods have been criticised for leading to unstable forecasts (e.g. Fildes, 1979), and, indeed, empirical studies have shown that they are less successful than the simpler, traditional procedure of constant optimised parameters (e.g. Makridakis *et al*., 1982).

This paper presents a new adaptive exponential smoothing method, which enables a smoothing parameter to be modelled as a logistic function of a user-specified variable. One simple choice for this variable is the magnitude of the forecast error in the previous period. The new method is analogous to that used to model the time-varying parameter in smooth transition models (see Teräsvirta, 1998). Using simulated and real data, we show that the new adaptive method challenges the performance of established adaptive approaches and constant parameter methods.

In the next section, we review the literature on adaptive exponential smoothing methods. We then introduce our new approach. The fourth section illustrates the potential of the approach using two simulated series. In the fifth section, we perform a large-scale simulation study in order to compare the forecast accuracy of the new method with that of other methods. We then use a large data set of real time series to evaluate the robustness of the new method. The final section provides a summary and concluding comments.

ESTABLISHED ADAPTIVE EXPONENTIAL SMOOTHING METHODS

There have been many different proposals for enabling the exponential smoothing parameters to adapt over time according to the characteristics of the series. Williams (1987) writes that it is widely accepted that in multi-dimensional smoothing, such as the Holt and Holt-Winters methods, only the smoothing parameter for the level should be adapted in order to avoid instability. It is, therefore, perhaps not surprising that the literature has focused largely on adaptive simple exponential smoothing. The forecasts from the adaptive simple exponential smoothing method are given by

$$f_{t+1} = \alpha_t \, y_t + (1-\alpha_t) f_t \qquad (1)$$

where f_{t+1} is the 1-step-ahead forecast made at period t and α_t is the adaptive parameter.

Although there is no consensus as to the most useful adaptive approach, the best known and most widely-used procedure was developed by Trigg and Leach (1967). Their method defines the smoothing parameter as the absolute value of the ratio of the smoothed forecast error to the smoothed absolute error:

$$\begin{aligned} \alpha_t &= \left| \frac{A_t}{M_t} \right| \\ A_t &= \phi \, e_t + (1-\phi) \, A_{t-1} \\ M_t &= \phi \left| e_t \right| + (1-\phi) \, M_{t-1} \end{aligned}$$

where e_t is the forecast error at period t ($e_t = y_t - f_t$), and ϕ is set arbitrarily, with 0.2 being a common choice. Trigg and Leach explain that this formulation enables α_t to vary according to the degree to which biased forecasts are obtained. Unfortunately, the approach sometimes delivers unstable forecasts. A number of researchers have tried to overcome this problem by restricting α_t to take one of a set of discrete values when certain control limits have been broken. Whybark (1973) defines the control limits in terms of multiples of the forecast error standard deviation, σ. An indicator variable, δ_t, is defined as

$$\delta_t = \begin{cases} 1 \ \ if \ |e_t| > 4\sigma \\ 1 \ \ if \ |e_t| > 1.2\sigma \ and \ |e_{t-1}| > 1.2\sigma \ and \ e_t \, e_{t-1} > 0 \\ 0 \ \ otherwise \end{cases}$$

The value of δ_t then determines whether α_t takes a base value, B, a medium value, M, or a high value, H. Whybark suggests $B = 0.2$, $M = 0.4$ and $H = 0.8$, with

$$\alpha_t = \begin{cases} H \ \ if \ \delta_t = 1 \\ M \ \ if \ \delta_t = 0 \ and \ \delta_{t-1} = 1 \\ B \ \ otherwise \end{cases}$$

With Dennis' (1978) method, α_t changes when the number, N_t, of consecutive errors with the same sign has broken a control limit, L. Formally, N_t and α_t are defined as

$$N_t = \begin{cases} 1 & if \ e_t e_{t-1} \leq 0 \\ N_{t-1} + 1 & otherwise \end{cases}$$

$$\alpha_t = \begin{cases} B & if \ N_t < L \\ \min\{\alpha_{t-1} + \lambda, 1\} & otherwise \end{cases}$$

where B is a base value for α_t, and λ is an incremental change. Common choices for the parameters are $B = 0.2$, $L = 2$ and $\lambda = 0.6$. Despite the attempts of Whybark, Dennis and others to overcome the problem of unstable forecasts, the experience of many users has been that instability remains (Fildes, 1979).

The Kalman Filter has been used to adapt the parameter in simple exponential smoothing (Bunn, 1981; Enns *et al.*, 1982; Snyder, 1988). However, the empirical results have not been particularly supportive (e.g. Ekern, 1981, 1982), and there is no single established best approach. A related area where the Kalman filter has become established is adaptive forecasting using weighted least squares, where the weighting incorporates limited memory on the estimation algorithm (see, for example, Young, 2003). The most popular such weighting scheme is the exponentially-weighted-past algorithm. The relationship to exponential smoothing can be seen by considering discounted least squares which reduces to simple exponential smoothing (see, for example, Harvey, 1990, §2.2.3). However, Young

(2003) shows that much improved results can be obtained if the parameters are modelled as simple stochastic processes (e.g. random walks), with the hyper-parameters (input variances or variance ratios), which effectively act as smoothing parameters, optimised by maximum likelihood, thus removing some of the subjectivity in the choice of more conventional exponential smoothing parameters (see also Young, 1999; Young *et al*., 1999).

Mentzer (1988) proposes a simple, albeit *ad hoc*, adaptive method. He suggests using the absolute percentage forecast error from the most recent period as α_t. In order to restrict α_t to the interval [0,1], if the absolute percentage error is greater than 100%, α_t is given a value of 1. Empirical evaluation gave encouraging results (Mentzer and Gomes, 1994).

A relatively recent addition to the literature is the method presented by Pantazopoulos and Pappis (1996). They argue that, since the ideal value for α_t would deliver $f_{t+1} = y_{t+1}$, then this ideal value can be derived by substituting y_{t+1} for f_{t+1} in (1) and solving for α_t to give

$$\alpha_t = \left(\frac{y_{t+1} - f_t}{y_t - f_t} \right).$$

Since y_{t+1} is unknown at time t, the ideal value of α_t for period t-1 is used for period t to give

$$\alpha_t = \left(\frac{y_t - f_{t-1}}{y_{t-1} - f_{t-1}} \right). \tag{2}$$

In order to restrict α_t to the interval [0,1], Pantazopoulos and Pappis propose that if α_t is found to be outside the interval [0,1], then it is replaced by the limit value to which it is closer. Unfortunately, brief consideration of (2) suggests that the approach is unlikely to be of use. Since the 1-step-ahead forecast is also the multi-step ahead forecast for simple exponential smoothing, the numerator in (2) is a 2-step-ahead forecast error, while the denominator is a 1-step-ahead forecast error. This suggests that the expression in (2) will very often deliver values greater than 1. Following the rule of Pantazopoulos and Pappis, α_t would then take a value of 1. The result is that α_t very often takes a value of 1. Perhaps, the explanation for the impressive empirical results reported by the authors is that they applied their approach to random walk series; $\alpha_t = 1$ is optimal for such series.

Although there have been a large number of adaptive methods proposed, empirical evidence does not conclusively favour any one method. There is, therefore, scope for further comparison, and potential for an improved method that avoids the problems of instability.

A NEW ADAPTIVE EXPONENTIAL SMOOTHING METHOD

Our new adaptive exponential smoothing method can be viewed as a type of smooth transition model. Granger and Teräsvirta (1993) and Teräsvirta (1998) provide useful reviews of these models and describe various applications. The essence of smooth transition models is that at least one parameter is modelled as a continuous function of a transition variable, V_t. A simple example is the following smooth transition regression model (STR)

$$y_t = a + b_t x_t + e_t \ ,$$

where
$$b_t = \frac{\omega}{1 + \exp(\beta + \gamma V_t)} ,$$

and a, ω, β and γ are constant parameters. If, for example, $\gamma < 0$ and $\omega > 0$, then b_t is a monotonically increasing function of V_t, which varies between 0 and ω. Smooth transition autoregressive models (STAR) are of a similar form, except x_t is replaced by lagged dependent variable terms, such as y_{t-1}. The parameters of smooth transition models can be estimated using non-linear least squares (Granger and Teräsvirta, 1993, §7.4).

In this paper, we propose a smooth transition adaptive exponential smoothing method, with smoothing parameter, α_t, defined as a logistic function of a user-specified transition variable, V_t. Although the approach is applicable to all exponential smoothing methods, for simplicity, we focus, in this introductory paper, on the case of simple exponential smoothing. Smooth transition exponential smoothing (STES) is then written as

$$f_{t+1} = \alpha_t \ y_t + (1 - \alpha_t) f_t \ ,$$

where $$\alpha_t = \frac{1}{1 + \exp(\beta + \gamma V_t)}. \tag{3}$$

If $\gamma < 0$, α_t is a monotonically increasing function of V_t. Hence, as V_t increases, the weight on y_t rises, and correspondingly the weight on f_t decreases. The logistic function restricts α_t to lie between 0 and 1. Although a wider range can be justified (Gardner, 1985), the restricted range has strong intuitive appeal, and is the convention of most of the popular forecasting packages. A feature of the approach is that unlike the existing adaptive methods, historical data is used to calibrate the adaptive smoothing parameter, α_t, through the estimation of β and γ in (3).

The choice of the transition variable, V_t, is of crucial importance to the success of the method. Consideration of the adaptive methods described in the previous section leads to a number of different possible transition variables. The value of the smoothing parameter in all of the existing adaptive methods depends to varying degrees on the magnitude of the most recent period's forecast error. Obvious candidates for the transition variable are, therefore, the square or the absolute value of the forecast error from the most recent period. An alternative to focusing solely on the most recent period's forecast error is to use the mean squared error or mean absolute error from several recent periods. One could also use a percentage error measure as transition variable.

Another possible choice of transition variable is the adaptive parameter from one of the existing methods. For example, the popular Trigg and Leach parameter could be used directly as the transition variable. The derived values of the STES constant parameters, β and γ in (3), will then govern the degree to which the variation in the Trigg and Leach parameter influences the STES smoothing parameter. Clearly, if the data dictate that the Trigg and Leach smoothing parameter is of no value, then γ will be 0 and the smoothing parameter will be constant. The STES method, therefore, enables recalibration of the existing adaptive methods.

Van Dobben de Bruyn (1964) briefly proposes the use of a logistic function as smoothing parameter. He suggests a more general function than that in expression (3) with the absolute error as transition variable. He does not present empirical results, but, instead, focuses on a procedure for which "computation is simpler". Obviously, computational constraints are now much reduced.

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 the new STES method. 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 (4)-(6), we present the state-space formulation for the STES method.

$$X_t = \mu_{t-1} + \varepsilon_t \tag{4}$$

$$\mu_t = \mu_{t-1} + \alpha_t \, \varepsilon_t \tag{5}$$

$$\alpha_t = \frac{1}{1 + \exp\left(\beta + \gamma V_t\right)}\,. \tag{6}$$

where ε_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). It is also worth noting that the statistical framework of expressions (4)-(6) enables standard errors to be estimated for the model parameters and, therefore, significance testing of the transition variable, V_t.

EMPIRICAL ILLUSTRATION USING TWO SIMULATED TIME SERIES

The adaptive methods reviewed in the second section of this paper were designed to enable the forecast function to adapt to level shifts. This has also been the main motivation in our development of the STES method. However, another feature of the method is that it can be useful for series with outliers. In this section, we use a simulated time series with level shifts and another with outliers to illustrate the potential usefulness of the method. We use the controlled environment of simulated data in order to create series with known properties.

Description of the Study

Simple exponential smoothing with a constant smoothing parameter, α, is optimal for an ARIMA(0,1,1) process with parameter θ=1-α. In view of this, it seems reasonable to generate the data from an ARIMA(0,1,1) process, interrupted by level shifts and (additive) outliers. We generated series of 100 observations. The first 80 were used to estimate method parameters and the remaining 20 were used to evaluate 1-step-ahead forecast performance. We focussed solely on 1-step-ahead prediction because adaptive exponential smoothing methods are designed to improve very short-term forecasting. Observations were generated from the following process:

$$\begin{aligned} y_t &= u_t + outlier_t \\ u_t &= u_{t-1} - \theta\, e_{t-1} + e_t + shift_t \end{aligned} \tag{7}$$

where e_t ~ N(0,1), θ = 0.8 and u_0 = 20. The variables $outlier_i$ and $shift_i$ were set as zero for all periods, except those for which we wanted to simulate a level shift or outlier. We modelled level shifts and outliers as increases of 25% of the underlying level. We chose 25% because it was the median level shift considered by Williams and Miller (1999). An in-sample level shift was represented by $shift_{40}$ = 0.25 u_{39}; an in-sample outlier by $outlier_{40}$ = 0.25 u_{39}; a post-sample level shift by $shift_{90}$ = 0.25 u_{89}; and a post-sample outlier by $outlier_{90}$ = 0.25 u_{89}. We generated a time series from each of the following two processes:

Process A – Expression (7) with an in-sample level shift and a post-sample level shift.

Process B – Expression (7) with an in-sample outlier and a post-sample outlier.

We implemented the STES method with the squared error from the previous period, e_t^2, used as transition variable, V_t. We also applied constant parameter simple exponential smoothing and the Trigg and Leach adaptive parameter method. The constant parameter method is an obvious benchmark because of its widespread use. We chose the Trigg and Leach method because it is the best known, and probably most widely used, adaptive method.

For the constant parameter exponential smoothing methods used in this paper, we calculated initial values for the smoothed components using the approach of Williams and Miller (1999), which is based on averages of the first few observations. To be consistent with our analysis of monthly data later in the paper, we calculated averages from the first 24 periods. Since adaptive methods are used when it is believed that the characteristics of the series may vary over time, we simply used the first observation as the initial smoothed level for these methods. We derived parameter values for all methods in this paper by the common procedure of minimising the sum of squared 1-step-ahead forecast errors, and we employed the constrained non-linear optimisation module of the Gauss programming language.

Results for the Series with Level Shifts

Figure 1 shows the time series generated from Process A with 1-step-ahead forecasts produced by constant parameter simple exponential smoothing. The optimised parameter of 0.59 is not close to 0.2, which is the optimal value for all periods except the two with level shifts. The level shift in the estimation period has caused the parameter to be relatively high in order that the forecasts adjust reasonably quickly to the new level of the series.

---------- FIGURES 1, 2 & 3 ----------

Figure 2 shows the same time series with 1-step-ahead forecasts produced by the Trigg and Leach method. The adaptive smoothing parameter, α_t, is plotted against the secondary y-axis. The forecast function can be seen to react a little more quickly to the level shift than the constant parameter forecast function in Figure 1. The plot of the adaptive parameter shows particularly high values around the level shifts in periods 40 and 90; the parameter increases in order to lift the forecast function to the new level. However, the high volatility in the Trigg and Leach parameter is a worrying feature; the parameter seems to overreact to variation in the time series. It is, therefore, not surprising that the Trigg and Leach forecasts have been criticised for being unstable.

Figure 3 shows the same series with forecasts produced by the STES method. As with the method of Trigg and Leach, the forecast function swiftly adapts to the level shifts. The plot of α_t shows clear increases around the level shifts, but noticeably lower volatility than the Trigg and Leach parameter. Apart from the two level shift periods, α_t takes a value of approximately 0.3. This is quite appealing since it is close to the value of 0.2, which is optimal for all periods except the two with level shifts. It is important to appreciate that the method's parameters were estimated using only the first 80 observations. The reaction to the level shift in period 90 shows that the method is able to adapt in the post-sample period. The median absolute percentage error (MedAPE) for the 1-step-ahead forecasts for the 20 post-sample periods was 1.8% for the STES method and 2.4% for both the constant parameter method and the Trigg and Leach method.

Expression (8) shows the STES adaptive parameter plotted in Figure 3. The negative coefficient for the transition variable, e_t^2, implies that an increase in e_t^2 will result in an increase in α_t. This enables the forecast function to adapt swiftly to the level shifts. The intuitive appeal of having a negative coefficient for series with level shifts motivates the inclusion of a non-positive constraint on the coefficient. We did not impose this constraint in this particular case, but we return to this issue later in the paper.

$$\alpha_t = \frac{1}{1 + \exp\left(0.57 - 0.039\, e_t^2\right)} \tag{8}$$

Maximum likelihood estimation of the state-space formulation of the STES method, in expressions (4)-(6), simplifies to the same least squares criterion we used to estimate the parameters in expression (8). Using the maximum likelihood module of the Gauss programming language, we derived: a standard error of 0.48 for the constant term, 0.57; a standard error of 0.019 for the transition variable coefficient, -0.039; and a value of –0.50 for the correlation between the estimation error in the two parameters. The standard error on the transition variable suggests that the variable is significant and that it is, therefore, worth using the STES method, rather than constant parameter simple exponential smoothing. To investigate the robustness of the formulation in expression (8) to parameter estimation error, we performed a Monte Carlo simulation, based on the parameter standard errors and correlation. We investigated how parameter estimation error affects the change in the adaptive parameter immediately after period 90, where the post-sample level shift occurred. A positive change is suitable for a level shift. The simulation results revealed that the probability of the change being negative was less than 2%, which is an encouraging result with regard to the robustness of the method.

Results for the Series with Outliers

Figure 4 shows the time series generated from Process B with 1-step-ahead forecasts produced by simple exponential smoothing with a constant parameter. The relatively low value of 0.25 for the optimised parameter enables the method to be reasonably robust to the outliers, but there is still a noticeable rise in the forecast function following each outlier.

---------- FIGURES 4, 5 & 6 ----------

The Trigg and Leach adaptive parameter and 1-step-ahead forecasts are shown in Figure 5. The forecast function shows considerable reaction to the two outliers. It rises up as it did for the level shifts in Process A, and then returns to the level of the series very slowly. As for the series with level shifts, the high volatility in the parameter is a worrying feature.

Figure 6 shows the 1-step-ahead forecasts produced by the STES method. Again, the squared error from the previous period, e_t^2, was used as transition variable. The forecast function shows little if any reaction to the outliers. The plot of the adaptive parameter shows a decrease around the two outliers in order to put a reduced weight on the outlier. It is important to appreciate that the reaction to the outlier of period 90 is a post-sample result. Although the adaptive parameter has more variation than it did in Figure 3 for Process A, the variation is noticeably lower than that shown in Figure 5 for the Trigg and Leach parameter. The MedAPE for the 1-step-ahead forecasts for the 20 post-sample periods was 2.7% for the STES method, 3.4% for the constant parameter method, and 3.9% for Trigg and Leach.

Expression (9) shows the adaptive smoothing parameter plotted in Figure 6. The positive value for the coefficient of the transition variable, e_t^2, implies that an increase in e_t^2 will result in a decrease in α_t. This avoids the forecast function reacting to the outliers.

$$\alpha_t = \frac{1}{1 + \exp\left(-0.29 + 0.32 e_t^2\right)} \tag{9}$$

For this series, we also implemented maximum likelihood estimation of the state-space formulation in expressions (4)-(6). We derived: a standard error of 0.97 for the constant term, -0.29; a standard error of 0.46 for the transition variable coefficient, 0.32; and a value of –0.89 for the correlation between the estimation error in the two parameters. The standard error for the transition variable suggests that it is not significant and it would, therefore, be preferable to use constant parameter simple exponential smoothing, rather than the STES method. The finding clearly differs from that for the series with level shifts discussed in the previous section. We feel this is reasonably intuitive because a method that is robust to outliers will improve accuracy for just the period in which the outlier occurs. By contrast a method that is able swiftly to pick up a level shift will be more accurate for more than just the period in which the shift took place due to the permanent nature of the shift. This reasoning suggests that the new STES method will be more useful for level shifts than outliers. The large-scale simulation study in the next section will address this issue.

We also performed a Monte Carlo simulation to investigate the robustness of the formulation in expression (9) to parameter estimation error. We assessed how the parameter estimation error affects the change in the adaptive parameter immediately after period 90, where the post-sample outlier occurred. Although a negative change is suitable for an outlier, we found that the probability of the change being positive was 24%, which again suggests that the method will not perform as well for outliers than for series with level shifts.

LARGE-SCALE SIMULATION STUDY

The previous section showed that the STES method has the potential to outperform other exponential smoothing methods for series in which the there is a level shift in both the estimation and evaluation periods, or an outlier in both. In this section, we extend this analysis in three ways. Firstly, we consider two additional data generating processes; one contains an in-sample level shift and a post-sample outlier, and the other has an in-sample outlier and a post-sample level shift. Secondly, in order to get a better understanding of forecasting performance, we evaluate accuracy for 1,000 series generated from each of the four processes. Thirdly, we include 16 different methods in the comparison.

Description of the Study

As in the previous section, we generated series of 100 observations from the process in expression (7). The first 80 were used to estimate method parameters and the remaining 20 were used to evaluate 1-step-ahead forecast performance. In addition to Processes A and B described in the previous section, we generated series from the following two processes:

Process C – Expression (7) with an in-sample level shift and a post-sample outlier.

Process D – Expression (7) with an in-sample outlier and a post-sample level shift.

We again modelled level shifts and outliers as increases of 25% of the underlying level. However, we noticed that for some series, the 25% increase was not sufficiently large for the observation to be a clear level shift or outlier. In view of this, and the fact that Williams and Miller (1999) considered shifts of up to 50% of the level of the series, we decided to also consider shifts of 50%.

For each series, we produced 1-step-ahead forecasts using 16 different methods. We implemented three constant parameter exponential smoothing methods: simple, Holt's and damped Holt's. As a naïve benchmark, we produced random walk forecasts. We implemented the adaptive exponential smoothing methods of Pantazopoulos and Pappis, Mentzer, Trigg and Leach, Whybark and Dennis, which we presented in the second section of the paper. We produced forecasts from the STES method using the following transition variables: the absolute value of the error from the previous period, $|e_t|$; the squared error from the previous period, e_t^2; the Trigg and Leach adaptive parameter; the Whybark adaptive parameter; and the Dennis adaptive parameter. We also considered the STES method with $|e_t|$ or e_t^2 as transition variable and the coefficient constrained to be less than or equal to zero. As we discussed in relation to expression (8), this constraint ensures that the method will be sensibly calibrated for level shifts.

Results

Tables I and II present the MedAPE results for the four processes, with level shifts and outliers simulated as 25% and 50%, respectively, of the underlying level. A percentage measure is suitable because we are evaluating across different series. We also calculated a variety of other percentage measures but we do not present these here because the relative performance of the methods for these measures was very similar to that for the MedAPE. Since we have 20 post-sample observations for each of the 1,000 series, each MedAPE value summarises 20,000 post-sample 1-step-ahead forecast errors. For each process, we also present, in Tables I and II, the ranking of the MedAPE result for each of the 16 methods. The first point to note is that, although it seemed important to consider level shifts and outliers of different magnitudes, the relative performance of the methods in Table I is broadly similar to that in Table II. Our comments in the remainder of this section are based on both tables.

---------- TABLES I & II ----------

For Process A, STES with $|e_t|$ or e_t^2 as transition variable performed very well. Constraining the coefficient of the transition variable to be less than or equal to zero had no effect on the results, indicating that the constraint was not needed to ensure a non-positive coefficient. Although the Trigg and Leach, Whybark and Dennis methods performed reasonably for Process A, none of the established adaptive methods, apart from the Whybark method, performed particularly well for the outliers in Process B. These methods are designed for level shifts, and so presumably they mistake the post-sample outlier in Process B for a level shift. By contrast, the STES method can be calibrated for outliers, as shown in expression (9). For Process B, the STES method with $|e_t|$ or e_t^2 as transition variable performed very well. Including the non-positive constraint on the transition variable coefficient led to weaker performance in Table II, because, as we saw in expression (9), a positive coefficient is appropriate for series with outliers. The results for the constrained case were the same as constant parameter simple exponential smoothing, and this is due to the constrained coefficient of the STES transition variable taking a value of zero for series generated from Process B. Using the Trigg and Leach, Whybark or Dennis adaptive parameter as transition variable in the STES method was an improvement on the straightforward use of these established adaptive parameters for Process B. However, this was not the case for the other three processes.

Processes C and D provide useful examinations of the STES method because one would assume that the method would perform poorly when evaluated on a post-sample period with different characteristics to that of the estimation sample. However, Tables I and II show that the STES method, using $|e_t|$ or e_t^2 as transition variable, performs very well for Process C. Detailed analysis of several series revealed that this is because the post-sample outlier is treated by the STES method as level shifts in successive periods. We found that the STES adaptive parameter rose sharply in reaction to the sudden rise in the series, and then maintained its high value for a further period because the level in the series falls sharply in the period following the outlier.

The results for Process D show that the STES method, with no constraint on the transition variable coefficient, performs relatively poorly when there is an outlier in the estimation sample and a level shift in the post-sample period. When the post-sample level shift occurs, the unconstrained STES method treats it as an outlier and, consequently, lowers the value of the adaptive parameter, which leads to the forecast function being very slow to adapt to the new level of the series. However, the results for the STES method were considerably improved by constraining the coefficient of the transition variable to be non-positive. The constraint does not allow the STES method to become calibrated for outliers, and should thus lead to a similar performance to that of constant parameter simple exponential smoothing. Curiously, the constrained STES method actually outperformed this method for Process D. The established adaptive methods performed very well for Process D. This is because they rely far less on the behaviour of the series in the estimation sample. This is also the reason for the impressive performance of the random walk.

In the second section of this paper, we commented that the Pantazopoulos and Pappis method is likely to very often produce a smoothing parameter equal to 1. This is supported by the fact that the results for this method are very similar to those for the random walk, which is equivalent to simple exponential smoothing with parameter equal to 1.

Although we do not report detailed results here, we also considered simulated series with both a level shift and an outlier in-sample, and either a level shift or an outlier post-sample. Perhaps not surprisingly, for these series, the STES method could not be suitably calibrated for the post-sample level shift or outlier, and as a result it performed similarly to constant parameter simple exponential smoothing.

In summary, the simulation results show that there is strong potential for the new STES method when $|e_t|$ or e_t^2 is used as the transition variable. However, the method will perform poorly if the estimation sample contains an outlier and the post-sample period contains a level shift. A solution to this problem is to constrain the transition variable coefficient to be non-positive, in order to avoid the adaptive parameter being calibrated towards outliers. The downside to this is that the constrained method will clearly not be appropriately calibrated if the post-sample period contains an outlier. However, it is important to note that, in this case, the performance of the constrained STES method will be no worse than that of constant parameter exponential smoothing.

EMPIRICAL COMPARISON USING REAL DATA

The main criticism levelled against adaptive exponential smoothing methods has been that they deliver unstable forecasts. Researchers have found that the instability can offset any response advantage when changes in structure occur (Gardner and Dannenbring, 1980). Indeed, empirical evidence is not favourable for adaptive exponential smoothing (e.g. Makridakis *et al.*, 1982). Presumably as a result of this, no such methods were included in the recent M3-Competition (see Makridakis and Hibon, 2000). To investigate the stability and usefulness of the STES method, in this section, we evaluate forecasting performance using real time series.

The data used was the 1,428 monthly time series from the M3-Competition. By applying exponential smoothing methods in an automated way to a large number of different 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. The monthly series vary in length from 48 to 126 with a median of 115. For each series, a further 18 observations were available to evaluate post-sample forecasting from 1 to 18 steps-ahead. As 1-step-ahead estimation is the focus of this paper, we calculated a 1 step-ahead forecast for each of the 18 post-sample observations. The results are discussed in the next section. They are not directly comparable with those from the M3-Competition because Makridakis and Hibon only used the first of the 18 post-sample periods from each series were to evaluate 1 step-ahead performance. We briefly consider multi-step-ahead forecasting in the section after next.

We implemented the 16 methods considered in our large-scale simulation study plus multiplicative Holt-Winters, which was omitted from our earlier analysis because it is suitable only for seasonal data. We calculated initial values for the smoothed level, trend and seasonal components of the Holt-Winters method using the approach of Williams and Miller (1999). For all methods, apart from Holt-Winters, 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.

Results for 1-Step-Ahead Forecasting

Table III summarises post-sample performance for the 17 methods. By contrast with the simulation study, the relative ranking of the methods for the real data varied according to the summary error measure used. In Table II, we report four popular percentage measures: MedAPE; mean absolute percentage error (MAPE); mean symmetric absolute percentage error (MSymAPE) (as used in the M3-Competition); and the root mean square percentage error (RMSPE). Since we have 18 post-sample observations for each of the 1,428 series, each MedAPE value summarises 25,704 post-sample 1-step-ahead forecast errors. For each process and each error measure, we also present in Table III, the ranking of each of the 17 methods. The final column in Table III shows the mean of the four ranks for each method.

---------- TABLE III ----------

Overall, simple exponential smoothing was the most successful of the constant parameter exponential smoothing methods. Of the five established adaptive methods, the Whybark and Dennis methods were the best. The RMSPE values for these two methods were the lowest of all 17 methods, while the MAPE results were also impressive. The other three established adaptive methods were, on the whole, no better than the random walk.

All seven versions of the STES method outperformed all five of the established adaptive methods according to the MedAPE and MSymAPE. Using the parameter from an established adaptive method as transition variable in the STES method led to improved accuracy according to three of the four error measures. Indeed, the final column in Table III shows that using the Whybark adaptive parameter as transition variable in the STES method resulted in the best mean rank of all 17 methods. Despite this, there is perhaps stronger intuitive appeal in using simpler transition variables, such as $|e_t|$ or e_t^2. Of the two, the mean ranks suggest that e_t^2 is better. The superiority of e_t^2 over $|e_t|$ is particularly apparent for the MAPE and RMSPE measures. Interestingly, the mean ranks for both methods are a little better when the non-positive constraint on the transition variable coefficient is used.

The RMSPE and MAPE are less robust to outliers and level shifts than the MSymAPE and MedAPE. This has led to the latter two measures becoming more widely quoted than the former two (see, for example, Makridakis and Hibon, 2000). However, in this paper, we are particularly interested in performance for level shifts and outliers, and so the MAPE and RMSPE are particularly relevant. Although our results show that constant parameter simple exponential smoothing is a good default, it is interesting to see that the method was outperformed according to MAPE and RMSPE by the Whybark method and by the STES method with the Whybark parameter or e_t^2 used as transition variable.

The main aim of this real data study was to investigate the stability and robustness of the various adaptive methods across a large number of series with different characteristics. It would also be interesting to compare the performance of the methods for the particular series, out of the 1,428, that have level shifts and outliers. However, with such a large number of series, automated procedures are needed to identify those series. Although Adya *et al.* (2001) describe heuristics for this purpose, they are considerably less convincing than the rigorous tests of, amongst others, Tsay (1988) and Balke (1993), but these tests have not yet been implemented in an automated procedure. Much more confidence is needed in the efficiency of such heuristics, before they can be used as a basis for comparing the usefulness of methods for series with level shifts and outliers. The development of automated procedures for identifying level shifts and outliers is an interesting area for future research. Indeed, there is strong potential for the use of the STES method as part of an automated expert system, which selects a method according to the characteristics of each time series.

Results for Multi-Step-Ahead Forecasting

Although adaptive exponential smoothing methods are designed to improve short-term forecasting, it is interesting to consider their performance for longer lead times. We used the final 18 observations for each of the 1,428 M3-Competition monthly series to evaluate forecasts for lead times from 1 to 18 steps-ahead. Using each of the 17 methods considered in the previous section, one forecast was produced for each lead time for each series. The same lead times were used with the monthly data in the M3-Competition, and so the results in this section are comparable with those reported by Makridakis and Hibon (2000).

Table IV summarises post-sample performance for the 17 methods. For brevity, we report only the MedAPE results as the relative rankings of the methods were broadly the same using the other error summary measures considered in the previous section. The table displays the average MedAPE for forecast horizons 1 to 6, for horizons 7 to 12, for horizons 13 to 18, and for all 18 horizons. The results show that the various implementations of the STES method produced very similar results to constant parameter simple exponential smoothing. Table IV confirms the findings of Makridakis and Hibon (2000) that it is very hard to beat the constant parameter damped Holt method for multi-step-ahead prediction. The success of damped trend exponential smoothing is also supported by the results of our recent analysis of damped multiplicative trend exponential smoothing (Taylor, 2003a).

---------- TABLE IV ----------

SUMMARY AND CONCLUDING COMMENTS

In this paper, we have introduced a new smooth transition adaptive parameter exponential smoothing method. A large-scale simulation study showed that there is strong potential for the new STES method when $|e_t|$ or e_t^2 is used as the transition variable. The method performs well when the estimation sample and evaluation sample both contain a level shift or both contain an outlier. However, if the estimation sample contains an outlier and the post-sample period contains a level shift, the method will perform poorly. A solution to this problem is to constrain the transition variable coefficient to be non-positive, in order to avoid the adaptive parameter being calibrated towards outliers.

Adaptive methods have been criticised for producing unstable forecasts and for performing poorly in empirical studies. Using 1,428 real time series, we found that the new STES method, with e_t^2 as transition variable, was able to outperform constant parameter exponential smoothing in terms of 1-step-ahead prediction according to RMSPE and MAPE. The new approach also performed particularly well when the Whybark adaptive parameter was used as transition variable. The use of an adaptive parameter, produced by another adaptive method, as transition variable demonstrates how the new approach enables recalibration of the existing adaptive methods.

A potential extension of the method is to incorporate judgemental information into the approach, as in the work of Williams and Miller (1999). They reformulate the exponential smoothing forecast function to allow for judgemental estimates of the timing and magnitude of a future level shift. These estimates could be used as transition variables in the smooth transition exponential smoothing method to ensure that the model adapts swiftly to the changing structure in the series.

The popularity of exponential smoothing for forecasting the volatility in financial returns and the recent introduction of smooth transition GARCH volatility models prompted us to investigate the use of the STES method for predicting stock index volatility (Taylor, 2003b). We found that the new method gave encouraging results when compared to fixed parameter exponential smoothing and a variety of GARCH models.

ACKNOWLEDGEMENTS

The author would like to acknowledge the helpful comments of an editor and an anonymous referee.

REFERENCES

Adya M, Collopy F, Armstrong JS, Kennedy M. 2001. Automatic identification of time series features for rule-based forecasting. *International Journal of Forecasting* **17**: 143-157.

Balke NS. 1993. Detecting level shifts in time series. *Journal of Business and Economic Statistics* **11**: 81-92.

Bunn DW. 1981. Adaptive forecasting using the Kalman filter. *Omega* **9**: 323-324.

Dennis JD. 1978. A performance test of a run-based adaptive exponential smoothing. *Production and Inventory Management* **19**: 43-46.

Ekern S. 1981. Adaptive exponential smoothing revisited. *Journal of the Operational Research Society* **32**: 775-782.

Ekern S. 1982. On simulation studies of adaptive forecasts. *Omega* **10**: 91-93.

Enns PG, Machak JA, Spivey WA, Wroblesk WJ. 1982. Forecasting applications of an adaptive multiple exponential smoothing model. *Management Science* **28**: 1035-1044.

Fildes R. 1979. Quantitative forecasting – the state of the art: extrapolative models. *Journal of the Operational Research Society* **30**: 691-710.

Gardner ES Jr. 1985. Exponential smoothing: the state of the art. *Journal of Forecasting* **4**: 1-28.

Gardner ES Jr. 1988. A simple method of computing prediction intervals for time-series forecasts. *Management Science* **34**: 541-546.

Gardner ES Jr, Dannenbring DG. 1980. Forecasting with exponential smoothing: Some guidelines for model selection. *Decision Sciences* **11**: 370-383.

Granger CWJ, Teräsvirta T. 1993. Modelling Nonlinear Economic Relationships, Advanced Texts in Econometrics, Oxford University Press: New York.

Harvey AC. 1990. Forecasting, structural time series models and the Kalman filter. Cambridge University Press: New York.

Hyndman RJ, Koehler AB, Ord JK, Snyder RD. 2001. Prediction intervals for exponential smoothing state space models. Working Paper, Department of Econometrics and Business Statistics, Monash University, Australia.

Hyndman RJ, Koehler AB, Snyder RD, Grose S. 2002. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* **18**: 439-454.

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.

Mentzer JT. 1988. Forecasting with adaptive extended exponential smoothing. *Journal of the Academy of Marketing Science* **16**: 62-70.

Mentzer JT, Gomes, R. 1994. Further extensions of adaptive extended exponential smoothing and comparison with the M-Competition. *Journal of the Academy of Marketing Science* **22**: 372-382.

Pantazopoulos SN, Pappis CP. 1996. A new adaptive method for extrapolative forecasting algorithms. *European Journal of Operational Research* **94**: 106-111.

Snyder RD. 1988. Progressive Tuning of Simple Exponential Smoothing Forecasts. *Journal of the Operational Research Society* **39**: 393-399.

Taylor JW, Bunn DW. 1999. A quantile regression approach to generating prediction intervals. *Management Science* **45**: 225-237.

Taylor JW. 2003a. Damped multiplicative trend exponential smoothing. Forthcoming *International Journal of Forecasting*.

Taylor JW. 2003b. Volatility forecasting with smooth transition exponential smoothing. Forthcoming *International Journal of Forecasting*.

Teräsvirta T. 1998. Smooth transition models. In Heij C, Schumacher H, Hanzon B, Praagman K. (Eds.), (with discussion), *System Dynamics in Economic and Financial Models* (pp. 109-136). Wiley: New York.

Trigg DW, Leach AG. 1967. Exponential smoothing with an adaptive response rate. *Operational Research Quarterly* **18**: 53-59.

Tsay RS. 1988. Outliers, level shifts, and variance changes in time series. *Journal of Forecasting* **7**: 1-20.

van Dobben de Bruyn CS. 1964. Prediction by progressive correction. *Journal of the Royal Statistical Society Series B* **26**: 113-122.

Whybark DC. 1973. Comparison of adaptive forecasting techniques. *Logistics Transportation Review* **8**: 13-26.

Williams TM. 1987. Adaptive Holt-Winters forecasting. *Journal of the Operational Research Society* **38**: 553-560.

Willaims DW, Miller D. 1999. Level-adjusted exponential smoothing for modelling planned discontinuities. *International Journal of Forecasting* **15**: 273-289.

Young PC. 1999. Nonstationary time series analysis and forecasting. *Progress in Environmental Science* **1**: 3-48.

Young PC, Pedregal DJ, Tych W. 1999. Dynamic harmonic regression. *Journal of Forecasting* **18**: 369-394.

Young PC. 2003. Identification of time varying systems. Forthcoming in *Encyclopaedia of Life Support Systems*, UNESCO.

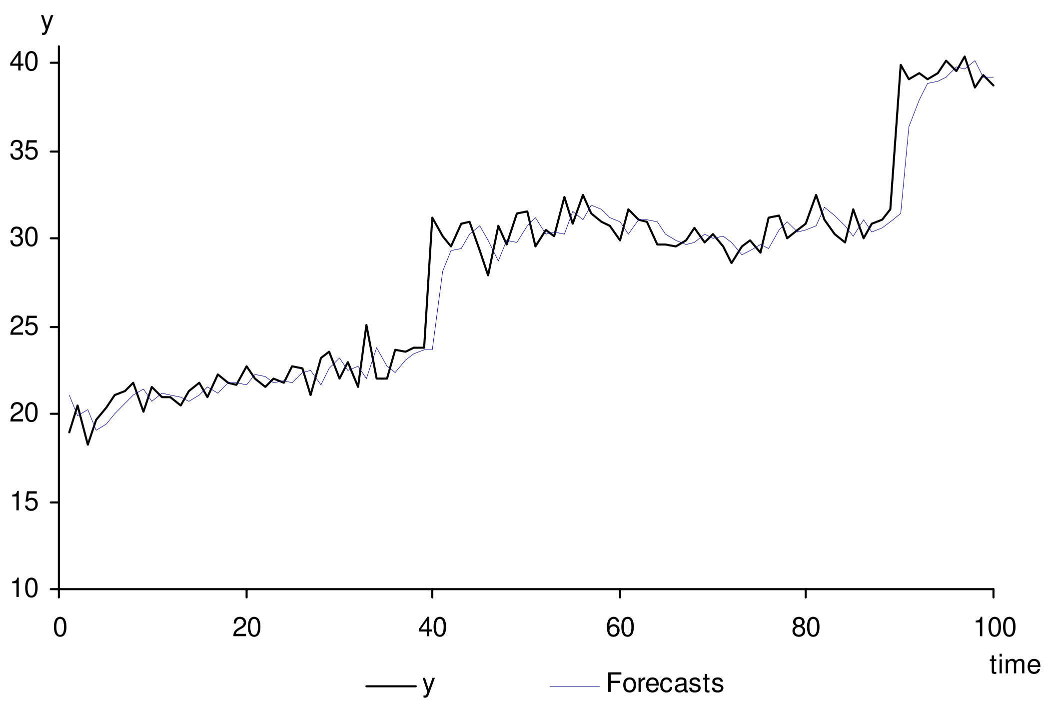

Figure 1. Simple exponential smoothing with constant parameter of 0.59. First 80 observations used to estimate parameter. Simulated series has level shifts of 25% of underlying level.

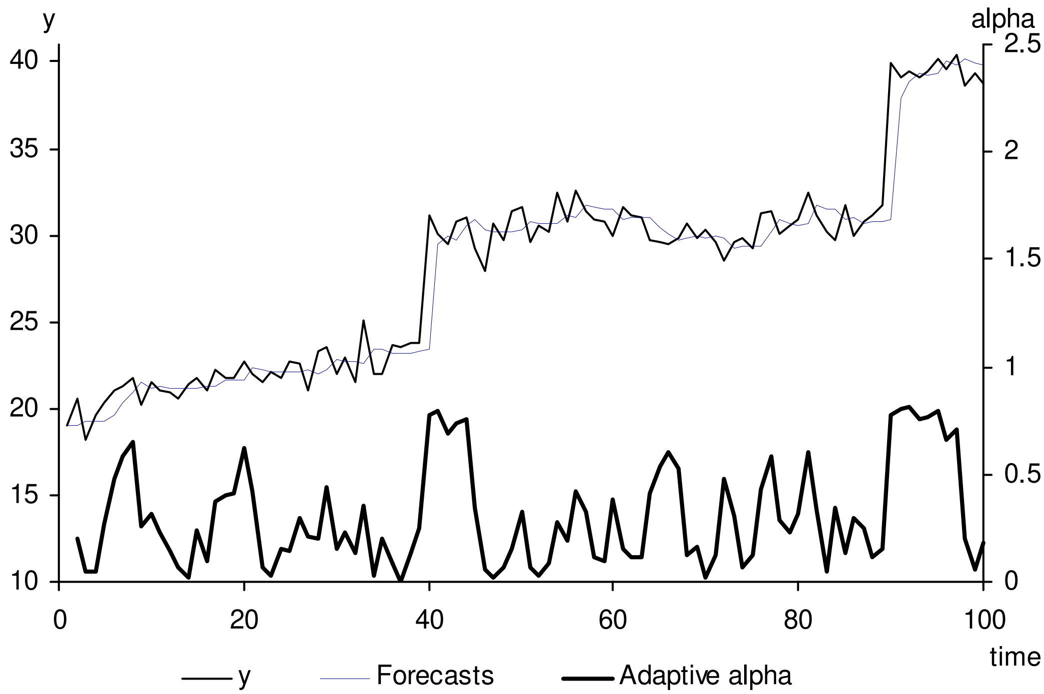

Figure 2. Trigg and Leach adaptive exponential smoothing.
Alpha plotted on secondary y-axis. Simulated series
has level shifts of 25% of underlying level.

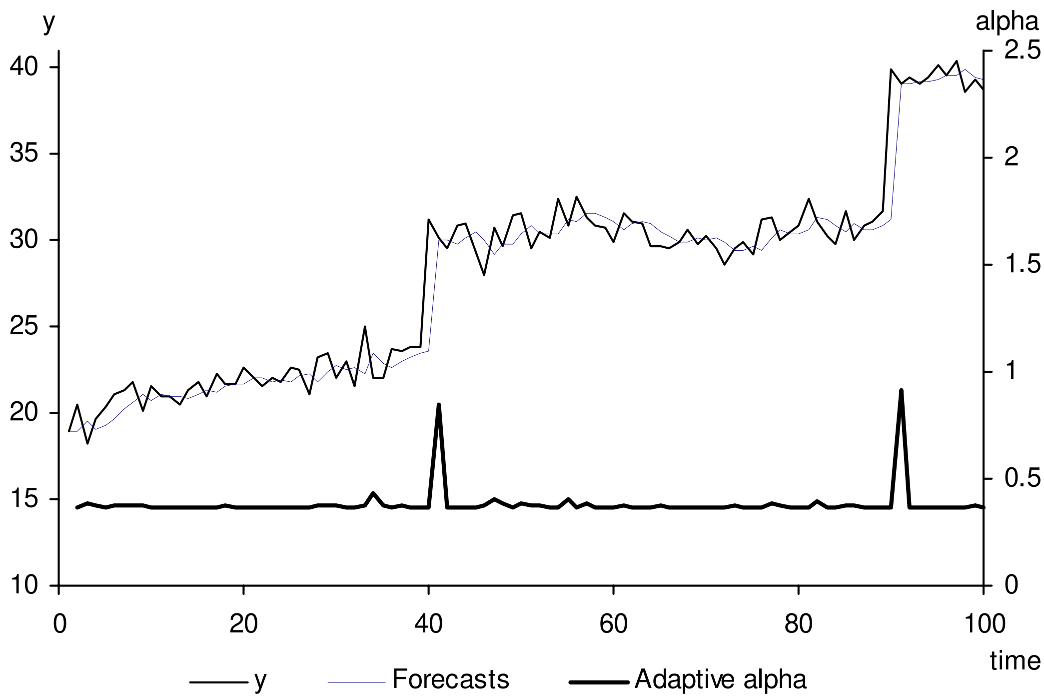

Figure 3. STES using squared error as transition variable. Alpha plotted on secondary y-axis. First 80 observations used to estimate parameters. Simulated series has level shifts of 25% of underlying level.

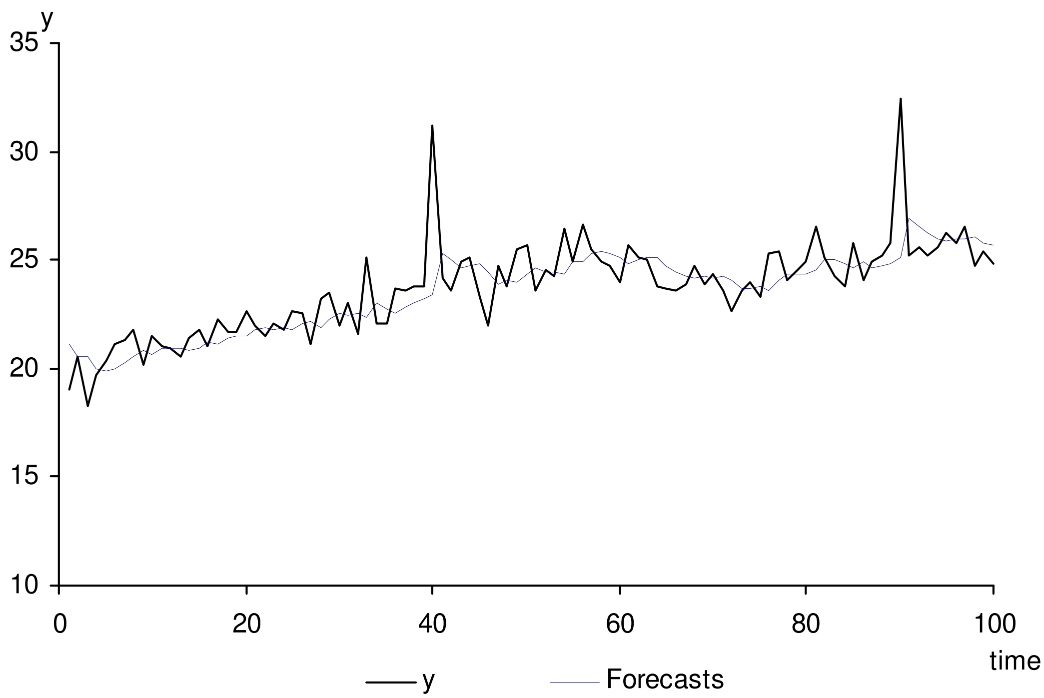

Figure 4. Simple exponential smoothing with constant parameter of 0.25. First 80 observations used to estimate parameter. Simulated series has outliers of 25% of underlying level.

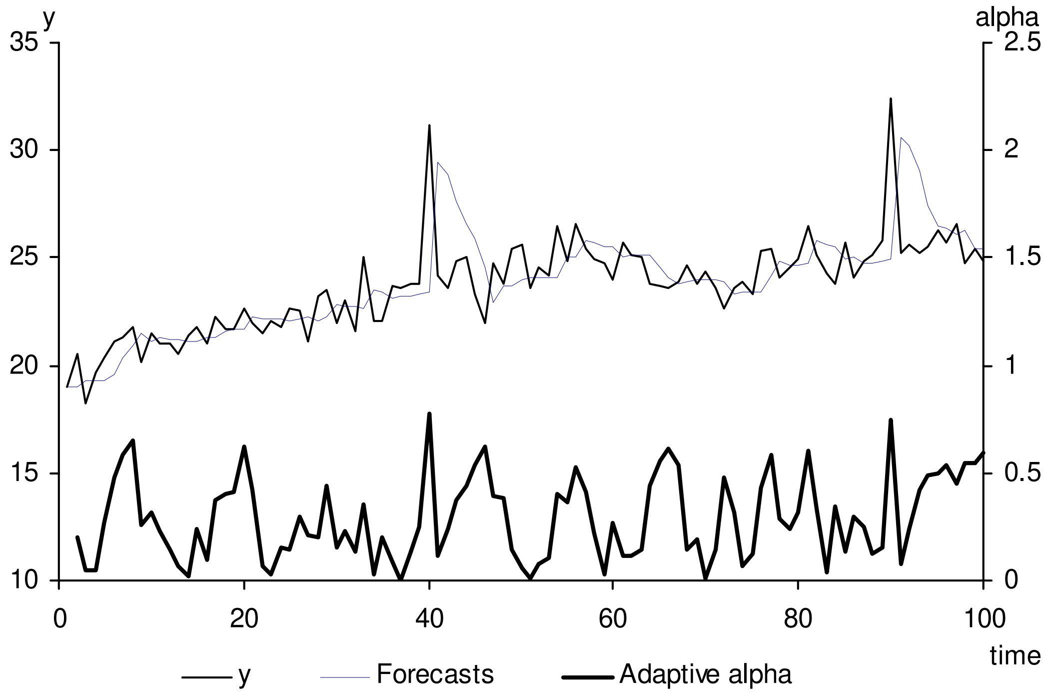

Figure 5. Trigg and Leach adaptive exponential smoothing.
Alpha plotted on secondary y-axis. Simulated series
has outliers of 25% of underlying level.

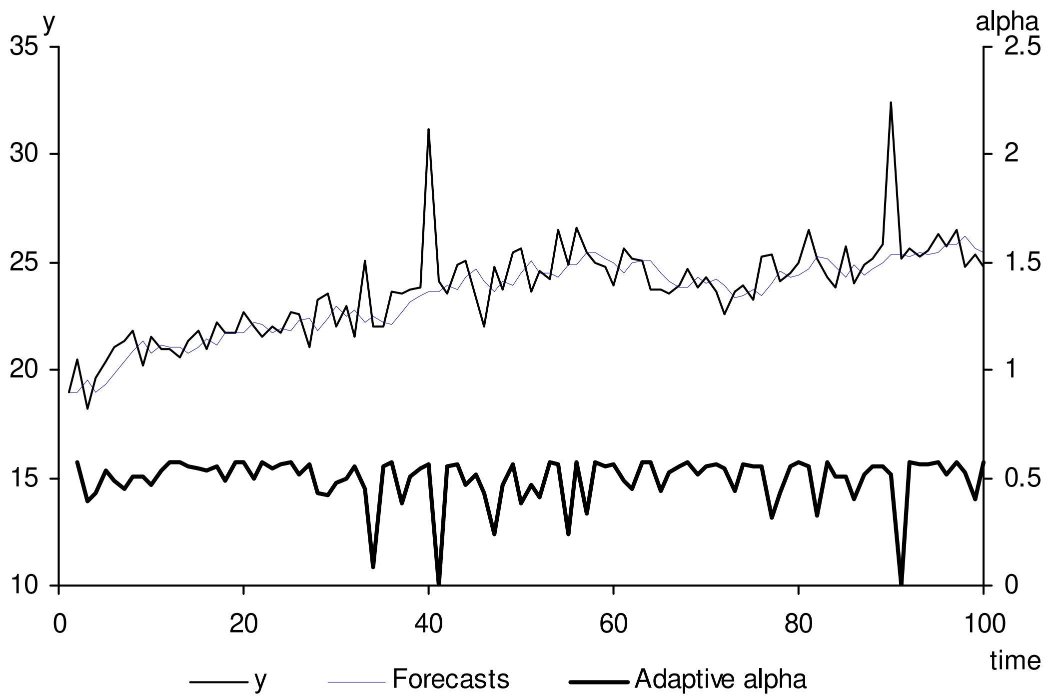

Figure 6. STES using squared error as transition variable. Alpha plotted on secondary y-axis. First 80 bservations used to estimate parameters. Simulated series has outliers of 25% of underlying level.

Table I. MedAPE for evaluation of 20 post-sample 1-step-ahead forecasts for 1,000 simulated series from each of 4 processes. Level shifts and outliers simulated as 25% increases from the underlying level of each series. Holt-Winters not included because series are not seasonal.

Process	**A**		**B**		**C**		**D**	
In-sample	**level shift**		**outlier**		**level shift**		**outlier**	
Post-sample	**level shift**		**outlier**		**outlier**		**level shift**	
Method	**%**	**rank**	**%**	**rank**	**%**	**rank**	**%**	**rank**
Simple	3.1	10	3.8	1	3.4	2	6.3	11
Holt	3.1	10	3.9	7	3.4	2	6.2	10
Damped Holt	3.1	10	4.0	10	3.4	2	7.2	14
Random walk	3.3	13	4.9	15	3.9	14	4.1	4
Pantazopoulos and Pappis	3.3	13	5.3	16	4.3	16	4.1	4
Mentzer	6.1	16	4.1	12	3.5	11	6.4	12
Trigg and Leach	2.9	6	4.7	14	4.0	15	3.6	2
Whybark	2.7	1	4.0	10	3.2	1	3.4	1
Dennis	3.0	7	4.4	13	3.6	12	3.8	3
STES using Trigg and Leach	3.0	7	3.9	7	3.7	13	5.9	9
STES using Whybark	3.0	7	3.8	1	3.4	2	6.4	12
STES using Dennis	3.3	13	3.9	7	3.4	2	5.8	8
STES using $\|e_t\|$	2.8	2	3.8	1	3.4	2	7.7	15
STES using $\|e_t\|$ with □≤ 0	2.8	2	3.8	1	3.4	2	5.6	6
STES using e_t^2	2.8	2	3.8	1	3.4	2	7.9	16
STES using e_t^2 with □≤ 0	2.8	2	3.8	1	3.4	2	5.6	6

Table II. MedAPE for evaluation of 20 post-sample 1-step-ahead forecasts for 1,000 simulated series from each of 4 processes. Level shifts and outliers simulated as 50% increases from the underlying level of each series. Holt-Winters not included because series are not seasonal.

Process	**A**		**B**		**C**		**D**	
In-sample	**level shift**		**outlier**		**level shift**		**outlier**	
Post-sample	**level shift**		**outlier**		**outlier**		**level shift**	
Method	**%**	**rank**	**%**	**rank**	**%**	**rank**	**%**	**rank**
Simple	2.4	8	4.2	5	3.1	6	10.2	12
Holt	2.4	8	4.3	9	3.2	8	9.2	10
Damped Holt	2.4	8	4.4	11	3.2	8	11.7	13
Random walk	2.5	12	4.9	13	3.3	12	3.7	4
Pantazopoulos and Pappis	2.5	12	5.4	15	3.6	14	3.7	4
Mentzer	6.7	16	5.9	16	5.5	16	7.2	6
Trigg and Leach	2.2	5	5.3	14	3.7	15	3.3	2
Whybark	2.2	5	4.2	5	2.9	5	3.2	1
Dennis	2.4	8	4.6	12	3.2	8	3.5	3
STES using Trigg and Leach	2.3	7	4.1	4	3.5	13	10.1	11
STES using Whybark	2.6	14	3.7	1	3.2	8	21.2	16
STES using Dennis	2.7	15	4.3	9	3.1	6	7.7	7
STES using $\|e_t\|$	2.1	1	3.8	2	2.8	3	17.5	15
STES using $\|e_t\|$ with □≤ 0	2.1	1	4.2	5	2.8	3	9.0	9
STES using e_t^2	2.1	1	3.8	2	2.7	1	15.9	14
STES using e_t^2 with □≤ 0	2.1	1	4.2	5	2.7	1	8.8	8

Table III. Evaluation of 18 post-sample 1-step-ahead forecasts for each of the 1,428 monthly series from the M3-Competition. All methods, except Holt-Winters, were applied to deseasonalised data.

	MedAPE		MAPE		MSymAPE		RMSPE		mean
Method	**%**	**rank**	**%**	**rank**	**%**	**rank**	**%**	**rank**	**rank**
Simple	3.7	1	15.6	6	10.8	1	253.6	10	4.5
Holt	3.7	1	15.6	6	11.1	7	284.0	14	7.0
Damped Holt	3.7	1	15.5	5	11.1	7	266.5	13	6.5
Holt-Winters	3.8	7	16.4	11	11.5	11	298.6	17	11.5
Random walk	4.2	13	17.2	16	13.2	15	169.3	6	12.5
Pantazopoulos and Pappis	4.3	15	18.4	17	13.4	16	196.1	7	13.8
Mentzer	6.8	17	16.6	14	13.6	17	151.8	4	13.0
Trigg and Leach	4.2	13	17.1	15	11.9	13	251.6	9	12.5
Whybark	4.4	16	15.2	3	11.5	11	129.1	2	8.0
Dennis	4.1	12	15.6	6	12.1	14	111.1	1	8.3
STES using Trigg and Leach	3.7	1	15.6	6	10.9	2	263.8	12	5.3
STES using Whybark	3.7	1	15.0	2	11.0	4	198.5	8	3.8
STES using Dennis	3.8	7	15.6	6	10.9	2	254.0	11	6.5
STES using $\|e_t\|$	3.8	7	16.4	11	11.2	9	291.3	15	10.5
STES using $\|e_t\|$ with $\gamma \leq 0$	3.7	1	16.4	11	11.0	4	291.4	16	8.0
STES using e_t^2	3.9	11	14.5	1	11.2	9	148.2	3	6.0
STES using e_t^2 with $\gamma \leq 0$	3.8	7	15.3	4	11.0	4	167.5	5	5.0

Table IV. MedAPE for evaluation of forecasts for lead times from 1 to 18 steps-ahead for each of the 1,428 monthly series from the M3-Competition. All methods, except Holt-Winters, were applied to deseasonalised data.

	Forecasting Horizon							
	1-6		7-12		13-18		1-18	
Method	**%**	**rank**	**%**	**rank**	**%**	**rank**	**%**	**rank**
Simple	5.3	4	7.5	4	9.8	6	7.5	4
Holt	5.1	2	6.9	2	9.5	2	7.2	2
Damped Holt	5.0	1	6.8	1	8.9	1	6.9	1
Holt-Winters	5.2	3	7.0	3	9.5	2	7.2	2
Random walk	6.0	15	8.3	15	10.6	15	8.3	15
Pantazopoulos and Pappis	6.1	16	8.4	16	10.8	16	8.4	16
Mentzer	7.4	17	9.2	17	11.3	17	9.3	17
Trigg and Leach	5.7	12	7.5	4	10.0	12	7.7	12
Whybark	5.8	13	7.8	13	10.1	13	7.9	13
Dennis	5.9	14	7.9	14	10.4	14	8.1	14
STES using Trigg and Leach	5.3	4	7.6	11	9.7	4	7.5	4
STES using Whybark	5.3	4	7.5	4	9.8	6	7.6	7
STES using Dennis	5.3	4	7.5	4	9.7	4	7.5	4
STES using $\lvert e_t \rvert$	5.4	10	7.5	4	9.9	9	7.6	7
STES using $\lvert e_t \rvert$ with $\gamma \leq 0$	5.3	4	7.5	4	9.9	9	7.6	7
STES using e_t^2	5.4	10	7.5	4	9.9	9	7.6	7
STES using e_t^2 with $\gamma \leq 0$	5.3	4	7.6	11	9.8	6	7.6	7