Intermittency and obsolescence: A Croston method with linear decay

S.D. Prestwich [a,*], S.A. Tarim [b], R. Rossi [c]

[a] *Insight Centre for Data Analytics, University College Cork, Ireland*
[b] *Cork University Business School, University College Cork, Ireland*
[c] *Business School, University of Edinburgh, UK*

ARTICLE INFO

ABSTRACT

Only two forecasting methods have been designed specifically for intermittent demand with possible demand obsolescence: Teunter–Syntetos–Babai (TSB) and Hyperbolic-Exponential Smoothing (HES). When an item becomes obsolete the TSB forecasts decay exponentially while those of HES decay hyperbolically. Both types of decay continue to predict nonzero demand indefinitely, and it would be preferable for forecasts to become zero after a finite time. We describe a third method, called Exponential Smoothing with Linear Decay, that decays linearly to zero in a finite time, is asymptotically the best method for handling obsolescence, and performs well in experiments on real and synthetic data.

1. Introduction

Traditional forecasting methods based on exponential smoothing and moving averages fail to perform well on *intermittent demand*, characterised by time series with many zeroes. It occurs particularly in industries such as aerospace, car manufacturing, and the military, where some stock items are required in small numbers and relatively infrequently (known as *slow-moving parts*). As these parts are often expensive, accurate forecasting of their demand is vital. An early and influential work on forecasting intermittent demand is that of Croston (1972), which was the first to recognise the importance of separating demand *size* and the inter-demand interval. Since then, several variants of the approach have been proposed: Snyder et al. (2012) provides a survey of work in this area.

Stock items may also be at risk of *obsolescence*: that is, demand drops to zero and remains there forever,

because the item has become obsolete. This issue is not addressed by most methods designed for intermittency, which continue to forecast nonzero demand forever in such cases. This is a serious drawback, especially in the context of an automated system tracking thousands of stock items, where spurious forecasts are unlikely to be noticed quickly. Two methods have been designed to tackle obsolescence and intermittency: Teunter–Syntetos–Babai (TSB), named after its inventors (Teunter et al., 2011), and Hyperbolic-Exponential Smoothing (HES) (Prestwich et al., 2014a). These behave similarly on demand without obsolescence, but when demand ceases, the TSB forecasts decay *exponentially* while the HES forecasts decay *hyperbolically*.

This paper makes two contributions. Firstly, it addresses the problem of *sudden obsolescence* in which demand for an item abruptly drops to zero (without any warning, such as decreasing demand over time) and remains there forever. When this is combined with intermittent demand, we believe that no existing forecasting method is appropriate. The forecasts of both TSB and HES are affected by the obsolescent item for an unlimited

* Corresponding author.

E-mail addresses: s.prestwich@cs.ucc.ie (S.D. Prestwich), armagan.tarim@ucc.ie (S.A. Tarim), Roberto.Rossi@ed.ac.uk (R. Rossi).

time, and it would be better for the forecasts to adapt more quickly to abrupt change. Intermittency and sudden obsolescence occur together in real-world applications such as aircraft spare parts, which are slow-moving and may be replaced by new versions. To address this issue, we propose a new Croston variant called Exponential Smoothing with Linear Decay (ESLD) whose forecasts decay *linearly* to zero in a finite time, so that the effect of obsolescent items on forecasts quickly vanishes. We compare ESLD, TSB, and HES empirically on real and synthetic data using several error measures, and show that it has competitive forecasting accuracy. Secondly, the paper proposes a new asymptotic analysis that can be used to compare forecasting methods in the case of obsolescence. While forecast accuracy can be compared empirically, the results depend on the form of the data. We believe that it is also useful to have theoretical reasons for choosing between them. Under this analysis we show that ESLD handles sudden obsolescence better than other methods.

The paper is organised as follows: Section 2 surveys existing forecasting methods and presents the new method, Section 3 analyses the handling of obsolescence by forecasting methods, Section 4 compares them empirically using synthetic demand data, and Section 5 concludes the paper.

2. Forecasting for intermittency and obsolescence

In this section we survey relevant forecasting methods, especially those designed to handle intermittency and obsolescence, and we introduce our new method. We denote the observed demand at discrete time t by y_t, a smoothed estimate of y by $\hat{y}_t$, and a forecast by f_t.

2.1. Existing forecasting methods

Single exponential smoothing (SES) (Brown, 1956) generates forecasts $\hat{y}_t$ by weighting observations using the formula

$$\hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1} \tag{1}$$

where $\alpha \in (0, 1)$ is a *smoothing parameter*. The smaller the value of α, the less weight attached to the most recent observations. A statistical model called ETS(A,N,N) has been shown to underlie SES (Hyndman et al., 2008). SES are related methods, and are surveyed in Gardner (2006). They often perform well, even beating more sophisticated approaches (Athanasopoulos et al., 2011; Fildes et al., 2008; Makridakis et al., 1982, 1993; Makridakis & Hibon, 2000), but SES is known to perform poorly on *intermittent demand* (Croston, 1972; Willemain et al., 1994), because it leads to inappropriate stock levels: there is an upward bias in the forecast in the period directly after a nonzero demand.

A well-known method for handling intermittency is *Croston's method* (Croston, 1972), which explicitly separates the aspects of demand size y and the interval τ between nonzero demands ($\tau = 1$ for non-intermittent demand). It applies SES to these quantities independently

to obtain smoothed estimates $\hat{y}_t$ and $\hat{\tau}_t$ at time t, sometimes with different smoothing factors α for y and β for τ (introduced by Schultz, 1987). The forecast is given by

$$f_t = \frac{\hat{y}_t}{\hat{\tau}_t} \tag{2}$$

$$\hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1} \tag{3}$$

$$\hat{\tau}_t = \beta \tau_{t-1} + (1 - \beta)\hat{\tau}_{t-1} \tag{4}$$

where $\hat{y}_t$ and $\hat{\tau}_t$ are updated only at time t, for which $y_t \neq 0$. It is not universally accepted to be superior to SES (Gardner, 2006) and some contrary evidence has been presented (Bacchetti & Saccani, 2012; Syntetos & Boylan, 2005). But it is generally considered to be better (Ghobbar & Friend, 2003) and versions of the method are used in leading statistical forecasting software packages (Boylan & Syntetos, 2007).

However, Croston's method was shown by Syntetos and Boylan (2005) to be biased on intermittent demand. They decreased the bias by modifying the forecasts to

$$f_t = \left(1 - \frac{\beta}{2}\right)\frac{\hat{y}_t}{\hat{\tau}_t} \tag{5}$$

where β is the smoothing factor used for inter-demand intervals, and may differ from the α smoothing factor used for demand magnitude. (In Syntetos and Boylan (2005) this factor is denoted by α because it is used to smooth both $\hat{y}$ and $\hat{\tau}$.) We shall refer to this variant as SBA.

A drawback to SBA is that it is significantly biased on non-intermittent demand because its forecasts are exactly those of SES multiplied by $(1 - \beta/2)$. This problem was previously avoided by Syntetos (2001) by using a forecast

$$f_t = \left(1 - \frac{\beta}{2}\right)\frac{\hat{y}_t}{\hat{\tau}_t - \beta/2} \tag{6}$$

but this modification increases the forecast variance (Teunter & Sani, 2007) and SBA is better known. We shall use SBA to represent Croston's method.

Although the above variants successfully handle intermittency, they do not handle *obsolescence*, in which $y_t = 0$ for all $t > t_0$ for some t_0: when obsolescence occurs they continue to forecast a fixed nonzero demand for all $t > t_0$. The first method explicitly designed to handle obsolescence and intermittency is the TSB method of Teunter et al. (2011), which updates an estimate of the *demand probability* instead of the inter-demand interval (and hence it is not technically a variant of Croston's method): instead of a smoothed interval $\hat{\tau}_t$ it uses a smoothed probability estimate: $\hat{p}_t$

$$\hat{p}_t = \beta p_{t-1} + (1 - \beta)\hat{p}_{t-1} \tag{7}$$

where p_t is 1 when demand occurs at time t, and 0 otherwise. Demand size is smoothed using Eq. (3). $\hat{p}_t$ is updated every period, whereas $\hat{y}_t$ is only updated when demand occurs. The forecast is

$$f_t = \hat{p}_t \hat{y}_t \tag{8}$$

This method is unbiased and handles intermittency well (Teunter et al., 2011). It also solves the problem of obsolescence because, like SES but unlike Croston and related

methods, when an item becomes obsolete its forecasts decay exponentially to zero.

A Croston variant designed to handle obsolescence is the HES method of Prestwich et al. (2014a). HES also separates demands into demand size y_t and the inter-demand interval τ_t. Its forecasts are

$$f_t = \begin{cases} \hat{y}_t/\hat{\tau}_t & \text{if } y_t > 0 \\ \hat{y}_t/(\hat{\tau}_t + \beta\tau_t/2) & \text{if } y_t = 0 \end{cases} \qquad (9)$$

where $\hat{y}_t$ and $\hat{\tau}_t$ are computed as above. Between nonzero demands, τ increases linearly, producing a hyperbolic decay in the forecasts. This was justified by a Bayesian argument in Prestwich et al. (2014a) where it was also proved that HES has similar bias to SBA.

2.2. The new method

As pointed out in Section 1, a drawback to both TSB and HES is that their forecasts decay slowly in the event of sudden obsolescence. Consequently, such methods will continue to advise stocking up on an item indefinitely, while in reality no more stock of the product is needed. The forecast decay can be made steeper by choosing a larger smoothing constant, but this might reduce forecast accuracy on demand without obsolescence. A moving average would adapt quickly to obsolescence but would not perform well on intermittent demand.

We suggest that a new forecasting method is needed that behaves similarly to methods such as TSB and HES when obsolescence does not occur, yet adapts more quickly when it does occur. We propose a new Croston variant called Exponential Smoothing with Linear Decay (ESLD) whose forecasts decay *linearly* to zero in a finite time. Our new method is similar in form to HES but uses the following forecasts:

$$f_t = \begin{cases} \hat{y}_t/\hat{\tau}_t & \text{if } y_t > 0 \\ (\hat{y}_t/\hat{\tau}_t)(1 - \beta\tau_t/2\hat{\tau}_t)^+ & \text{if } y_t = 0 \end{cases} \qquad (10)$$

where x^+ denotes $\max(0, x)$, and $\hat{y}_t$, $\hat{\tau}_t$ are computed as in HES. When obsolescence occurs the forecasts decay linearly to zero at a rate controlled by β, and when they reach zero they remain there until further nonzero demands occur. The rate at which they decay can be controlled by adjusting β. This feature distinguishes it from TSB and HES, which only approach zero asymptotically. We call this forecasting method Exponential Smoothing with Linear Decay (ESLD).

We show in Appendix A that ESLD has similar bias to SBA on intermittent demand, under the assumption that $1 - \beta\tau_t/2\hat{\tau}_t \geq 0$ always holds. If this assumption does not hold (which may occur if we set β to a high value) then the term will be replaced by 0, causing a positive bias, but we show empirically in Section 4 that this effect is negligible under typical parameter settings.

The advantages of ESLD are that, as a Croston variant, it handles intermittency well, and that its linear decay under zero demand enables it to adapt more quickly to sudden obsolescence than TSB and HES (and indeed SBA and the original Croston method, which do not adapt at all).

3. Asymptotic obsolescence error

There are now three forecasting methods that are explicitly designed to handle obsolescence and intermittency: TSB, HES, and ESLD. They have qualitatively different behaviour when obsolescence occurs, respectively decaying exponentially, hyperbolically, and linearly. Each is approximately unbiased on intermittent demand, but which method handles obsolescence best? This is a difficult question because the answer depends on many factors: the type of demand data, how long we compare forecasting methods before and after obsolescence occurs, and which error measures we use for the comparison. In order to answer this question we first analyse the asymptotic behaviour of the different methods, in an attempt to obtain a definitive answer for a specific case.

We consider the case of sudden obsolescence, and compare the behaviour of the three forecasting methods from the time $T = 0$ just after the final nonzero demand, up to $T \to \infty$. We ignore the machine-dependent issue of arithmetic errors causing truncation to 0 as forecasts become small. All of the methods are approximately unbiased, so we assume they have the same forecast f_0 when obsolescence occurs at time 0. This represents the limiting case of sudden obsolescence followed by many forecasts, as might occur in an automated inventory control system: if a human does not notice that an item has become obsolete, then the best results will be obtained by a forecasting method that minimises the error as $T \to \infty$.

We would like to compare error measures for the methods over these times. However, many measures have been used in the literature and in forecasting competitions (Makridakis et al., 1982, 1993; Makridakis & Hibon, 2000), and there is no consensus on which is best. It is generally recommended to use several. We now discuss the suitability of all the measures listed in the surveys of De Gooijer and Hyndman (2006), Hyndman and Koehler (2006) and the article of Wallström and Segerstedt (2010).

The *scale-dependent measures* are based on the mean error $e_t = y_t - \hat{y}_t$ or mean square error e_t^2, and include the Mean Error, Mean Square Error, Root Mean Square Error, Mean Absolute Error, and Median Absolute Error. As $T \to \infty$, all these tend to zero, so they cannot be used for an asymptotic comparison.

The *percentage errors* are based on the quantities $p_t = 100e_t/y_t$ and include the Mean Absolute Percentage Error, Median Absolute Percentage Error, Root Mean Square Percentage Error, Root Median Square Percentage Error, Symmetric Mean Absolute Percentage Error, and Symmetric Median Absolute Percentage Error. As $y_t = 0$ for all $t > 0$, they are undefined.

The *relative error-based measures* are based on the quantities $r_t = e_t/e_t^*$, where e_t^* is the error from a baseline forecasting method, and include the Mean Relative Absolute Error, Median Relative Absolute Error, and Geometric Mean Relative Absolute Error. The baseline forecasting method is usually the *random walk* (or *naive method*), which simply forecasts that the next demand will be identical to the current demand. For all times, $e_t^* = 0$, so these measures are undefined. We could use another

baseline but we would still have the problem that the mean and median e_t are zero. As such, these cannot be used for a comparison.

The *relative measures* are mainly defined as the ratio of (i) an error measure, and (ii) the same measure applied to a baseline forecasting method. These include the Relative Mean Absolute Error, Relative Mean Squared Error, and Relative Root Mean Squared Error. Again, the baseline method is usually the random walk. Both measures tend to zero as $T \to \infty$, so these cannot be used for our comparison. A different form of relative measure is the Percent Better, which computes the percentage of times t a forecasting method has a smaller absolute error $|e_t|$ than a baseline method, again usually random walk. Random walk has asymptotically perfect performance, so the Percent Better cannot be used for our comparison. A related measure is the Percent Best, in which no baseline method is used: instead it computes the percentage of times t each method being tested has a smaller absolute error $|e_t|$ than the others. This does not have the random walk problem, and we shall use it below.

The *scaled errors* include the MAD/Mean Ratio (Kolassa & Schütz, 2007) and Mean Absolute Scaled Error (Hyndman & Koehler, 2006). The former cannot be used for our comparison because the denominator (the mean error) tends to zero, while the latter cannot be used because it is proportional to e_t, which tends to zero.

There are also three recent measures designed for intermittent demand (Wallström & Segerstedt, 2010). (i) The Cumulative Forecast Error is defined as the sum of all errors over the time periods under consideration. Not taking averages means that errors do not become vanishingly small, so this measure gives meaningful results. We shall use it and also the related Cumulative Squared Error (which was not mentioned in Wallström and Segerstedt (2010)): a motivation for using squared errors is that they penalise outliers more severely than absolute errors, giving a different perspective. (ii) The Number of Shortages at time t is defined as the number of periods in which the Cumulative Forecast Error is strictly positive and demand is nonzero. In our scenario demand is always zero after obsolescence occurs, so this is not meaningful. (iii) The Periods in Stock at time t is defined as

$$\sum_{i=1}^{t} (\hat{y}_i - y_i)(t + 1 - i) \tag{11}$$

In our scenario $y_i = 0$ for all $i > 0$, so this reduces to

$$f_0 \lim_{t \to \infty} \sum_{i=1}^{t} \hat{y}_i (t + 1 - i) \tag{12}$$

But as $t \to \infty$, the term $\hat{y}_1 t \to \infty$ and all other terms are positive, so this measure is also not meaningful here.

To summarise the above discussion: the Percent Best (PBt), Cumulative Forecast Error (CFE), and Cumulative Squared Error (CSE) are the only standard error measures we know of that can be used for our comparison. There has been recent work on devising new error measures for intermittent demand (Kolassa, 2016; Prestwich et al., 2014b) but these are less established, and at least some

Table 1
Asymptotic obsolescence errors.

	CFE	CSE	PBt
TSB	f_0/β	f_0^2/β^2	0%
HES	∞	$2f_0^2 \hat{\tau}_0/\beta$	0%
ESLD	$f_0 \hat{\tau}_0/\beta$	$2f_0^2 \hat{\tau}_0/3\beta$	100%

will yield infinities when used for asymptotic analysis. We shall therefore use the more classical PBt, CFE, and CSE.

The results of the comparison are shown in Table 1 and the derivations are given in Appendix B. Forecasts are made one step ahead in all cases. First we consider the CFE results. HES is worst with an infinite error, while TSB beats ESLD. However, the β of TSB is not smoothing the same quantity as the β of ESLD and other Croston variants, so these two terms are not strictly comparable. Teunter et al. (2011) recommend a smaller β for TSB than for Croston's method, and Babai et al. (2014) show that the TSB β should be approximately $1/\hat{\tau}$ times the Croston β to obtain similar variance. Under this assumption, TSB has an asymptotic obsolescence error $f_0\hat{\tau}_0/\beta$ that is the same as ESLD. However, this point is debatable, and if the reader is not convinced by the conversion of TSB β to Croston β then our results can be taken to show that TSB has a lower asymptotic error than ESLD. Next we consider the CSE results. TSB is not comparable to HES and ESLD because it is related to them by a factor of $\hat{\tau}_0\beta$, which might be less than or greater than 1, yet ESLD is three times better than HES. For sufficiently intermittent data (with large enough $\hat{\tau}_0$) the asymptotic performance of TSB will suffer unless β is increased, in which case it should beat ESLD and HES. But β cannot be increased beyond 1, so ESLD and HES will perform better on highly intermittent demand with sudden obsolescence. On the other hand, for only slightly intermittent demand with sudden obsolescence, TSB should be superior. Finally, we consider the PBt results. Here, ESLD beats both HES and TSB.

In summary, ESLD is not beaten by TSB or HES under CSE or CFE (though, as noted above, it can be argued that TSB beats ESLD under CFE), but beats both under PBt. Overall, we rank ESLD as the best method for handling obsolescence. TSB beats HES under CFE, draws with it under PBt, and is not comparable under CSE. Thus, we rank TSB second best, and HES third best.

4. Experiments

In this section we test the accuracy of ESLD using synthetic and real demand data. Our asymptotic analysis deals only with a special case, so it is important to verify that ESLD also performs well empirically. We only compare ESLD with TSB and HES, as other methods are not designed to handle both intermittency and obsolescence. SES and Croston's method were already shown to be inferior to TSB by Teunter et al. (2011), and the simple method of always forecasting zero is not useful in general, although it sometimes does well in experiments (Prestwich et al., 2014b).

4.1. Synthetic data

We reproduce the synthetic data experiments of Teunter et al. (2011). We use 1000 runs for each forecasting method, with 1000 time periods used for initialisation, and a further 1000 periods for evaluation. We use two error measures: the Mean Error (ME) to measure bias, and the Root Mean Square Error (RMSE).

4.1.1. Stationary demand

First we consider *stationary* intermittent demand (without obsolescence). Teunter et al. (2011) compared several forecasting methods on demand that is nonzero with probability p_0 where p_0 is either 0.2 or 0.5, and whose demand sizes y_t follow a logarithmic distribution. Geometrically distributed intervals are a discrete version of Poisson intervals, and the combination of Poisson intervals and logarithmic demand sizes yields a negative binomial distribution. There is theoretical and empirical evidence that these are realistic (Syntetos et al., 2011). The logarithmic distribution is characterised by a parameter $\ell \in (0, 1)$ and is discrete with $\Pr[X = k] = -\ell^k/k \log(1-\ell)$ for $k \geq 1$. Teunter et al. (2011) used two values: $\ell = 0.001$ to simulate low demand, and $\ell = 0.9$ to simulate lumpy demand. They used several combinations of smoothing factors: α values 0.1, 0.2, and 0.3, and β values 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, and 0.3.

4.1.2. Decreasing demand

In this experiment, demand sizes y_t again follow the logarithmic distribution, but the probability of a nonzero demand decreases linearly, from p_0 in the first period to 0 during the last period. As pointed out by Teunter et al. (2011), none of the forecasting methods use trending to model the decreasing demand, so all are positively biased.

4.1.3. Sudden obsolescence

This experiment is the same as that of Section 4.1.2 except that the demand probability is reduced instantly to 0 after half the time periods. Demand sizes y_t are again logarithmically distributed. This is the most important experiment from our point of view, as ESLD is designed for this situation.

4.1.4. Best-case results

The results are summarised in Table 2, in which each figure is optimised independently by testing different combinations of α, β values. In the table, 'S' denotes stationary demand, 'D' decreasing demand, and 'O' sudden obsolescence, and 1–4 respectively denote $\ell = 0.9, p_0 = 0.5$; $\ell = 0.9, p_0 = 0.2$; $\ell = 0.001, p_0 = 0.5$; and $\ell = 0.001, p_0 = 0.2$. It can be seen that the best-case error values of the three methods are quite similar, showing that ESLD is a reasonable method for intermittent demand without obsolescence.

The ME results also show that ESLD has low bias (though not the lowest), despite the fact that, as noted in Section 2, it will not be unbiased if the term $1 - \beta\tau_t/2\hat{\tau}_t$ becomes negative. This shows that the effect of negative values is negligible given reasonable smoothing factors.

Table 2
Best-case errors.

Case	ME			RMSE		
	TSB	HES	ESLD	TSB	HES	ESLD
S1	0.0018	−0.0002	0.0028	3.8432	3.8405	3.8405
S2	0.0001	0.0002	0.0004	2.5216	2.5190	2.5190
S3	−0.0006	0.0008	−0.0024	0.5006	0.4999	0.4999
S4	0.0004	−0.0003	−0.0028	0.4097	0.4084	0.4084
D1	0.0012	0.0407	0.0087	2.6774	2.6848	2.6843
D2	−0.0001	0.0103	0.0036	1.9825	1.9840	1.9839
D3	0.0013	0.0116	0.0018	0.4624	0.4326	0.4326
D4	0.0012	0.0084	0.0023	0.3061	0.3067	0.3067
O1	0.0074	0.0566	0.0038	2.9378	2.9494	2.9390
O2	0.0010	0.0370	0.0001	1.7470	1.7514	1.7477
O3	0.0013	0.0258	0.0008	0.3625	0.3720	0.3630
O4	0.0012	0.0179	−0.0010	0.2856	0.2887	0.2861

Table 3
Winning best-case forecasting methods.

Demand	ME	RMSE
S (stationary)	TSB+HES	HES+ESLD
D (decreasing)	TSB	TSB
O (sudden obsolescence)	ESLD	TSB

It is perhaps surprising that the differences between HES, TSB, and ESLD are so small in the case of sudden obsolescence. We believe that this is because of the demand patterns in the data. ESLD is designed to handle the case of highly intermittent demand, followed by sudden obsolescence, followed by a long period of forecasting. This case was analysed in Section 3. The data used here are not highly intermittent, nor do they continue for a long time after obsolescence occurs.

To distinguish between the three methods, Table 2 is further summarised in Table 3, which shows the winning forecasting method under each combination of demand pattern and error measure: a method is considered to 'win' if it gives the best results in at least one case. No clear dominance emerges, and all three perform quite similarly, as the differences are fairly small. However, TSB is the winner under decreasing demand, and ESLD wins more often under sudden obsolescence.

We conclude that ESLD is a competitive forecasting method, under three error measures and various combinations of intermittency and obsolescence, and when smoothing parameters are well-tuned. The most interesting result is for sudden obsolescence, for which ESLD is specifically designed. ESLD beats the other two methods under ME, while TSB wins under RMSE. ESLD comes in at a close second under RMSE, consistently beating HES by 0.0026–0.0104, whereas TSB consistently beats ESLD by 0.005–0.0012.

4.2. Real data

Finally, we tested ESLD on real data obtained from an inventory control company. The data are the same as those used in Prestwich et al. (2014a). The data come from a spare parts inventory and contain 24 monthly demands for each of 8727 products, with high intermittency: the mean probability of a nonzero demand is 0.238. We would

prefer to use long series, but it is common to have access to only relatively short series, and companies must often make forecasts based on short demand histories. This point was mentioned in Syntetos and Boylan (2005).

Prestwich et al. (2014b) found large differences between the mean demands over the first and second years that biased the results of their experiments. They therefore applied a data-cleaning phase, discarding any series whose first-year mean demand was more than three times that of its second year, or vice-versa, leaving 2202 series. We show results using both the cleaned and raw data: the former are (we believe) more typical, but the latter contain examples of obsolescence, which is of interest here. We also show results for the series that exhibit obsolescence, which we define here as those whose first-year mean was greater than 0.1 and whose second-year mean was less than 0.1. In the data, 5145 series exhibited obsolescence.

Again the question arises of which error measures to use. Prestwich et al. (2014b) used the Percent Better (PB, with random walk as the baseline), the U2 statistic of Thiel (1966) (equivalent to Relative RMSE with random walk as the baseline), and the Relative Mean Absolute Error (RelMAE). On this data, RelMAE ranked the all-zero method (in which all forecasts are zero) as better than SBA. This is not a useful result, and we do not include it here.

We used the same smoothing factor values α as Syntetos and Boylan (2005) (0.05, 0.1, 0.15, and 0.2) and the same β values as Teunter et al. (2011) (0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, and 0.3). Table 4 is the result of testing every parameter combination and reporting the best results over all series, for each combination of forecasting method and error measure independently. We treat our series in a similar way to Syntetos and Boylan, who used several thousand series, each with 24 demands: initialise the forecasting methods using the first year and evaluate them using the second year. The estimates of demand size and the inter-demand interval (and the TSB smoothed demand probability) are initialised to their averages over the first year. We also show results for SES, as it is a simple universal benchmark that is not always beaten on intermittent demand.

The results can be summarised as follows. On the cleaned data under U2, ESLD and HES jointly rank first, under PB, ESLD ranks first. SES was worst under both measures. On the raw data under U2, surprisingly, SES ranks first with ESLD second; under PB, HES ranks first, ESLD second, and SES last. On the obsolescence data under U2, TSB and HES jointly rank first, and SES ranks last; under PB, HES ranks first, ESLD second, and SES last. However, many of the differences between TSB, HES, and ESLD are very small. These results confirm that ESLD is a reliable method for intermittent demand with and without obsolescence.

5. Conclusion

The problem of sudden obsolescence is an important one faced by companies in real life when a stock item is replaced by a new product. Existing intermittent demand

Table 4
Best results on real data.

Measure	SES	TSB	HES	ESLD
	Cleaned data			
U2	0.728	0.722	0.721	0.721
PB	31.58	31.61	31.64	31.65
	Raw data			
U2	0.768	0.782	0.783	0.774
PB	13.04	15.11	15.13	15.12
	Obsolescence data			
U2	0.787	0.793	0.793	0.785
PB	21.72	25.24	25.27	25.25

methods have not been developed for this case, and do not perform well when presented with it. We proposed a new method that has properties making it ideal for this case, and showed it to be almost unbiased on intermittent demand. We proposed a form of asymptotic analysis to compare how well forecasting methods handle obsolescence, based on a worst-case scenario in which a highly intermittent item becomes obsolete and forecasts continue forever. This analysis ranked ESLD as the best method, followed by TSB, and then HES.

We also performed empirical experiments using synthetic and real demand data, and found ESLD to be highly competitive compared to TSB and HES. TSB has previously been shown to have lower bias and deviation than other methods on intermittent demand (Teunter et al., 2011) so ESLD will also compare well against these forecasting methods. We found that ESLD has the highest bias but the lowest deviation, that TSB has the least error under decreasing demand, and that ESLD has the least error under sudden obsolescence. The last result agrees with our asymptotic analysis. In future work we hope to find other forms of analysis to compare these forecasting methods.

Based on all our experiments, we recommend setting both smoothing factors α and β in the range of 0.05–0.3. This is in broad agreement with recommendations in the literature for Croston and SES-based methods.

Appendix A. Derivation of the forecasting method

This derivation follows a pattern similar to that of HES (Prestwich et al., 2014a). The ESLD method uses a forecast of the form

$$f_t = \begin{cases} \hat{y}_t/\hat{\tau}_t & \text{if } y_t > 0 \\ (\hat{y}_t/\hat{\tau}_t)(1 - k_t \tau_t)^+ & \text{if } y_t = 0 \end{cases} \qquad (13)$$

for some value k_t, and we choose k_t to make ESLD unbiased on intermittent demand. First we derive the expectation $\mathbb{E}[f_t]$. Consider the demand sequence as a sequence of *substrings*, each starting with a nonzero demand: for example the sequence $(5, 0, 0, 1, 0, 0, 0, 3, 0)$ has substrings $(5, 0, 0)$, $(1, 0, 0, 0)$ and $(3, 0)$. Within a substring, $\hat{y}_t$ and $\hat{\tau}_t$ remain constant (because they are only updated when $y > 0$), and if an item has not become obsolescent and k_t is sufficiently small then $1 - k_t \tau_t > 0$, so ESLD has the following expected forecast:

$$\mathbb{E}\left[\left(\frac{\hat{y}_t}{\hat{\tau}_t}\right)(1 - k_t \tau_t)\right] = \left(\frac{\hat{y}_t}{\hat{\tau}_t}\right)(1 - k_t \mathbb{E}[\tau_t]) \tag{14}$$

For intermittent demand the inter-demand interval is a random variable with geometric distribution and mean $1/p$. We estimate $p \approx 1/\hat{\tau}_t$ so $\mathbb{E}[\tau_t] \approx \hat{\tau}_t$ and the expected forecast over the string is

$$\left(\frac{\hat{y}_t}{\hat{\tau}_t}\right)(1 - k_t \hat{\tau}_t) \tag{15}$$

This coincides with SES on non-intermittent demand, so ESLD is unbiased on non-intermittent demand whatever the value of k_t because it generates the same forecasts. To make it unbiased on intermittent demand we choose k_t so that it has the same expected forecast as SBA's fixed forecast over each string, which is

$$\left(\frac{\hat{y}_t}{\hat{\tau}_t}\right)\left(1 - \frac{\beta}{2}\right) \tag{16}$$

So $k_t = \beta/2\hat{\tau}_t$, and the forecast when $y_t = 0$ is

$$f_t = \left(\frac{\hat{y}_t}{\hat{\tau}_t}\right)\left(1 - \frac{\beta \tau_t}{2\hat{\tau}_t}\right)^+ \tag{17}$$

Moreover, ESLD updates $\hat{y}_t$ and $\hat{\tau}_t$ in exactly the same way as SBA at the start of each substring. Therefore, it has the same expected forecast as SBA over the entire demand sequence.

Appendix B. Derivation of asymptotic errors

In this Appendix we derive asymptotic obsolescence errors for the three forecasting methods.

B.1. Cumulative forecast error

The CFE is the sum of all errors for $t \geq 0$, used for example in Wallström and Segerstedt (2010). In our scenario all forecasts are positive and all demands are zero, so the CFE coincides with the Cumulative *Absolute* Error. The TSB CFE is

$$f_0[1 + (1 - \beta) + (1 - \beta)^2 + \cdots] = \frac{f_0}{\beta} \tag{18}$$

HES's CFE is derived from the $y_t = 0$ case of Eqs. (9) (moving the $\hat{\tau}_t$ term from the first denominator to the second one):

$$f_0 \sum_{t=0}^{\infty} \frac{1}{1 + t\beta/2\hat{\tau}_0} \tag{19}$$

This is a special case of the general harmonic series, which diverges to ∞. ESLD's CFE is

$$f_0 + f_0\left(1 - \frac{\beta}{2\hat{\tau}_0}\right) + f_0\left(1 - \frac{2\beta}{2\hat{\tau}_0}\right) + \cdots + f_0\left(\frac{\beta}{2\hat{\tau}_0}\right) \tag{20}$$

Under the simplifying assumption that $2\hat{\tau}_0/\beta$ is an integer ℓ, the series contains ℓ terms, so the CFE is

$$\frac{f_0}{\ell}[\ell + (\ell - 1) + (\ell - 2) + \cdots + 1]$$
$$= \frac{f_0(\ell + 1)}{2} \approx \frac{f_0 \ell}{2} = \frac{f_0 \hat{\tau}_0}{\beta} \tag{21}$$

We justify the simplifying assumption as follows. For highly intermittent demand, $\hat{\tau}_0$ will be large, and β should be set to a small value to compensate for this. So $2\hat{\tau}_0/\beta$ will be a non-integer that is much larger than 1, and rounding it to an integer will have a negligible effect.

B.2. Cumulative squared error

The CSE is the sum of all squared errors. The TSB CSE is

$$f_0^2[1 + (1 - \beta)^2 + (1 - \beta)^4 + \cdots] = \left(\frac{f_0}{\beta}\right)^2 \tag{22}$$

The HES CSE is (like the CFE) derived from the $y_t = 0$ case of Eqs. (9):

$$f_0^2 \sum_{t=0}^{\infty} \frac{1}{(1 + t\beta/2\hat{\tau}_0)^2} \tag{23}$$

To evaluate this summation we use the digamma function. It is known that

$$\psi^{(1)}(z) = \sum_{i=0}^{\infty} \frac{1}{(z + i)^2} \tag{24}$$

where $\psi^{(1)}$ is the first derivative of the digamma function. Replacing z by $1/x$,

$$\psi^{(1)}(1/x) = \sum_{i=0}^{\infty} \frac{1}{(1/x + i)^2} = x^2 \sum_{i=0}^{\infty} \frac{1}{(1 + ix)^2} \tag{25}$$

so

$$\sum_{i=0}^{\infty} \frac{1}{(1 + ix)^2} = \frac{\psi^{(1)}(1/x)}{x^2} \tag{26}$$

Using an asymptotic expansion for large z,

$$\psi^{(1)}(z) = \sum_{i=0}^{\infty} \frac{B_i}{z^{i+1}} \tag{27}$$

where the B_i are Bernoulli numbers. Taking a first term approximation we get $B_0/z = B_0 x = x$. Therefore,

$$\frac{\psi^{(1)}(1/x)}{x^2} \approx \frac{1}{x} \tag{28}$$

Substituting $\beta/2\hat{\tau}_0$ for x, HES's CSE is $2f_0^2\hat{\tau}_0/\beta$. ESLD's CSE is

$$\left(\frac{f_0}{\ell}\right)^2 \left[\ell^2 + (\ell-1)^2 + (\ell-2)^2 + \cdots + 1\right]$$
$$= \left(\frac{f_0}{\ell}\right)^2 \left(\frac{\ell^3}{3} + \frac{\ell^2}{2} + \frac{\ell}{6}\right) \qquad (29)$$

Recall that $\ell = 2\hat{\tau}/\beta$, which for highly intermittent demand is a large number, so the ℓ^2 and ℓ terms are dominated by the $\ell^3 term$ and we can ignore them. The result is $2f_0^2\hat{\tau}_0/3\beta$.

B.3. Percent best

To compute the Percent Best (PBt) we take a collection of forecasting methods and count the percentage of times at which each gives the smallest error. The PBt is popular because it is scale-free and easy to understand. Comparing the three forecasting methods in this way, ESLD has a PBt of 100% while the others have a PBt of 0%. This is because the ESLD forecasts reach zero at some finite time T after obsolescence occurs, whereas TSB and HES never reach zero. Therefore, for all times $t \geq T$ the ESLD error is smaller than the TSB and HES errors. So in the limit, as $t \to \infty$, the PBt tends to 100% for ESLD and to 0% for TSB and HES.

References

Athanasopoulos, G., Hyndman, R. J., Song, H., & Wu, D. C. (2011). The tourism forecasting competition. *International Journal of Forecasting*, *27*, 822–844.

Babai, M. Z., Syntetos, A., & Teunter, R. (2014). Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. *International Journal of Production Economics*, *157*, 212–219.

Bacchetti, A., & Saccani, N. (2012). Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. *Omega*, *40*(6), 722–737.

Boylan, J. E., & Syntetos, A. A. (2007). The accuracy of a modified Croston procedure. *International Journal of Production Economics*, *107*, 511–517.

Brown, R. G. (1956). *Exponential smoothing for predicting demand* (p. 15). Cambridge, Massachusetts: Arthur D. Little Inc.

Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, *23*(3), 289–304.

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, *22*(3), 443–473.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. A. (2008). Forecasting and operational research: A review. *The Journal of the Operational Research Society*, *59*, 1150–1172.

Gardner, E. S., Jr. (2006). Exponential smoothing: The state of the art — Part II. *International Journal of Forecasting*, *22*(4), 637–666.

Ghobbar, A. A., & Friend, C. H. (2003). Evaluation of forecasting methods for intermittent parts demand in the field of aviation: A predictive model. *Computers & Operations Research*, *30*, 2097–2114.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: The state-space approach*. Springer.

Kolassa, S. (2016). Evaluating predictive count data distributions in retail sales forecasting. *International Journal of Forecasting*, *32*, 788–803.

Kolassa, S., & Schütz, W. (2007). Advantages of the MAD/Mean ratio over the MAPE. *Foresight: The International Journal of Applied Forecasting*, *6*, 40–43.

Makridakis, S., Andersen, 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 M-2 competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, *9*, 5–23.

Makridakis, S., & Hibon, M. (2000). The M3-competition: Results, conclusions and implications. *International Journal of Forecasting*, *16*(4), 451–476.

Prestwich, S. D., Tarim, S. A., Rossi, R., & Hnich, B. (2014). Forecasting intermittent demand by hyperbolic-exponential smoothing. *International Journal of Forecasting*, *30*(4), 928–933.

Prestwich, S. D., Tarim, S. A., Rossi, R., & Hnich, B. (2014). Mean-based error measures for intermittent demand forecasting. *International Journal of Productions Research*, *52*(22), 6782–6791.

Schultz, C. R. (1987). Forecasting and inventory control for sporadic demand under periodic review. *The Journal of the Operational Research Society*, *38*(5), 453–458.

Snyder, R. D., Ord, J. K., & Beaumont, A. (2012). Forecasting the intermittent demand for slow-moving inventories: A modelling approach. *International Journal of Forecasting*, *28*, 485–496.

Syntetos, A. A. (2001). *Forecasting for intermittent demand* (Ph.D. thesis), Buckinghamshire Chilterns University College, Brunel University, Unpublished.

Syntetos, A., Babai, Z., Lengu, D., & Altay, N. (2011). Distributional assumptions for parametric forecasting of intermittent demand. In N. Altay, & A. Litteral (Eds.), *Service parts management: Demand forecasting and inventory control* (pp. 31–52). NY, USA: Springer Verlag.

Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting*, *21*, 303–314.

Teunter, R., & Sani, B. (2007). On the bias of Croston's forecasting method. *European Journal of Operational Research*, *194*, 177–183.

Teunter, R., Syntetos, A. A., & Babai, M. Z. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, *214*(3), 606–615.

Thiel, H. (1966). *Applied economic forecasting*. Chicago: Rand McNally.

Wallström, P., & Segerstedt, A. (2010). Evaluation of forecasting error measurements and techniques for intermittent demand. *Internation Journal of Production Economics*, *128*, 625–636.

Willemain, T. R., Smart, C. N., Shockor, J. H., & DeSautels, P. A. (1994). Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal of Forecasting*, *10*, 529–538.