International Journal of Forecasting 30 (2014) 928–933

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting intermittent demand by hyperbolic-exponential smoothing

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

[a] Insight Centre for Data Analytics, University College Cork, Ireland
[b] Institute of Population Studies, Hacettepe University, Ankara, Turkey
[c] University of Edinburgh Business School, Edinburgh, UK
[d] Computer Engineering Department, Izmir University of Economics, Turkey

ARTICLE INFO

Keywords:
Intermittent demand
Croston's method
Bayesian inference

ABSTRACT

Croston's method is generally viewed as being superior to exponential smoothing when the demand is intermittent, but it has the drawbacks of bias and an inability to deal with obsolescence, where the demand for an item ceases altogether. Several variants have been reported, some of which are unbiased on certain types of demand, but only one recent variant addresses the problem of obsolescence. We describe a new hybrid of Croston's method and Bayesian inference called Hyperbolic-Exponential Smoothing, which is unbiased on non-intermittent and stochastic intermittent demand, decays hyperbolically when obsolescence occurs, and performs well in experiments.

1. Introduction

Inventory management is of great economic importance to industry, but forecasting the demand for spare parts is difficult because it is *intermittent*: the demand is zero in many time periods. This type of demand occurs in several industries, for example in aerospace and military inventories, from which spare parts such as wings or jet engines are only required infrequently. Various methods have been proposed for forecasting, some simple and others statistically sophisticated, but relatively little work has been done on intermittent demand. Most of the work in this area is influenced by that of Croston (1972), who was the first to separate the forecasting of the demand size and the inter-demand interval.

Another difficult feature of some inventories is *obsolescence*, where an item is considered obsolete if it has seen no demand for a long time. When many thousands of items are being handled automatically, this may go unnoticed by Croston-style methods. One of the authors of this paper (Prestwich) has worked with an inventory company who used Croston's method, but were forced to resort to ad hoc rules such as: *if an item has seen no demand for* 2 *years then forecast* 0. This issue has been relatively neglected in the literature, though a method was designed recently for tackling it (Teunter, Syntetos, & Babai, 2011).

In this paper, we describe a new Croston-style forecasting method with a low bias and high forecasting accuracy on both intermittent and non-intermittent demand, which can also handle obsolescence. It is competitive with existing methods, and is more robust under changes to its smoothing factors. Its novelty is that its forecasts decay hyperbolically during periods of no demand (a property derived from Bayesian inference), whereas other methods decay exponentially or not at all. Section 2 provides a background on existing methods, Section 3 describes the new method, Section 4 evaluates it empirically, and Section 5

* Corresponding author.
E-mail address: s.prestwich@cs.ucc.ie (S.D. Prestwich).

http://dx.doi.org/10.1016/j.ijforecast.2014.01.006
0169-2070/

concludes the paper. For an extended version of this paper, see Prestwich, Tarim, Rossi, and Hnich (2013).

2. Background

We now briefly survey some relevant forecasting methods. *Single exponential smoothing* (SES) generates estimates $\hat{y}_t$ of the demand by weighting previous observations y exponentially via the formula

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

where $\alpha \in (0, 1)$ is a *smoothing parameter*. The smaller the value of α, the less weight is attached to the most recent observations. An up-to-date survey of exponential smoothing algorithms is given by Gardner (2006). They perform remarkably well, often beating more complex approaches (Fildes, Nikolopoulos, Crone, & Syntetos, 2008), but SES is known to perform poorly (under some measures of accuracy) on intermittent demand.

A well-known method for handling intermittency is *Croston's method* (Croston, 1972), which applies SES to the demand sizes y and intervals τ independently. Given the smoothed demand $\hat{y}_t$ and smoothed interval $\hat{\tau}_t$ at time t, the forecast is

$$f_t = \frac{\hat{y}_t}{\hat{\tau}_t}.$$

Both $\hat{y}_t$ and $\hat{\tau}_t$ are updated at each time t for which $y_t \neq 0$. According to Gardner (2006), it is hard to conclude from the various studies that Croston's method is successful, because the results depend on the data used and on the way in which forecast errors are measured. However, it is generally regarded as one of the best methods for intermittent series (Ghobbar & Friend, 2003), and versions of the method are used in leading statistical forecasting software packages such as SAP and Forecast Pro (Teunter et al., 2011).

To remove at least some of the known bias of Croston's method on stochastic intermittent demand (in which demands occur randomly), a correction factor is introduced by Syntetos and Boylan (2005):

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

where β is the smoothing factor used for inter-demand intervals, which may be different to the α smoothing factor used for demands.[1] This works well for intermittent demand but is biased for non-intermittent demand, as its forecasts are those of SES multiplied by $(1 - \beta/2)$. This problem is avoided by Syntetos (2001), who uses a forecast

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

This removes the bias on non-intermittent demand, but increases the forecast variance (Teunter & Sani, 2007).

Another modified Croston method is given by Levén and Segerstedt (2004), who claim that it also removes the bias in the original method, but in a simpler way: they apply SES to the ratio of the demand size and interval length each time a nonzero demand occurs. That is, they update the forecast using

$$f_t = \alpha\left(\frac{y_t}{\tau_t}\right) + (1-\alpha)f_{t-1}.$$

However, this also turns out to be biased (Boylan & Syntetos, 2007).

A more recent development is the method of Teunter et al. (2011), which updates the demand probability instead of the demand interval. Instead of a smoothed interval $\hat{\tau}$, it uses exponential smoothing to estimate a probability $\hat{p}_t$, where p_t is 1 when demand occurs at time t and 0 otherwise. Different smoothing factors α and β are used for $\hat{y}_t$ and $\hat{p}_t$ respectively. $\hat{p}_t$ is updated every period, while $\hat{y}_t$ is only updated when the demand occurs. The forecast is

$$f_t = \hat{p}_t\hat{y}_t.$$

This method is unbiased. It also solves the problem of obsolescence, because, like SES but unlike other Croston variants, an item's forecasts decay when it becomes obsolescent.

Following convention, we shall use CR to denote the original method of Croston, SBA to denote the variant of Syntetos and Boylan, SY to denote that of Syntetos, and TSB to denote that of Teunter, Syntetos and Babai.

3. Hyperbolic-exponential smoothing

We take a Croston-style approach, separating demands into the demand size y_t and the inter-demand interval τ_t. As in most Croston methods, when a non-zero demand occurs, the estimated demand size $\hat{y}_t$ and inter-demand period $\hat{\tau}_t$ are both exponentially smoothed, using factors α and β respectively. The novelty of our method is what happens when there is no demand.

Suppose that at time t, we have a smoothed demand size $\hat{y}_t$ and an inter-demand period $\hat{\tau}_t$, up to and including the last non-zero demand, and that we have observed τ_t consecutive periods without demand since the last non-zero demand. What should be our estimate of the probability that a demand will occur in the next period? A similar question was addressed by Laplace (1814): given that the sun has risen N times in the past, what is the probability that it will rise again tomorrow? His solution was to add one to the count of each event (the sun rising or not rising) to avoid zero probabilities, and estimate the probability by counting the adjusted frequencies. So if we have observed N sunrises and 0 non-sunrises, in the absence of any other knowledge we would estimate the probability of a non-sunrise tomorrow as $1/(N + 2)$. This is known as the *rule of succession*. However, he noted that, given any additional knowledge about sunrises, we should adjust this probability. These ideas are generalised by the modern *pseudocount* (or *pseudo-observation*) method, which can be viewed as Bayesian inference with a Beta prior distribution. We base our discussion on a recent book by Poole and Mackworth (2010, Ch. 7), which describes the technique we need in the context of Bayesian classifiers, but similar material can be found in many publications.

[1] Syntetos and Boylan (2005) denoted this factor by α because it is used to smooth both $\hat{y}$ and $\hat{\tau}$.

For the two possibilities $y_t = 0$ and $y_t \neq 0$, we add non-negative pseudocounts c_0 and c_1 respectively to the actual counts n_0 and n_1 of observations.[2] In addition to addressing the problem of zero observations, pseudocounts also allow us to express the relative importance of prior knowledge and new data when computing the posterior distribution. By Bayes' rule, the posterior probability of a nonzero demand occurring is estimated by

$$p(y_t \neq 0) = \frac{n_1 + c_1}{n_0 + c_0 + n_1 + c_1}.$$

(This is actually a conditional probability that depends on the recent observations and prior probabilities, but we follow Poole & Mackworth, 2010, and write $p(y_t \neq 0)$ for simplicity.) In our problem, we have seen no demand for τ_t periods, so $n_1 = 0$ and $n_0 = \tau_t$:

$$p(y_t \neq 0) = \frac{c_1}{\tau_t + c_0 + c_1}.$$

We can eliminate one of the pseudocounts by noting that the prior probability of a demand found by exponential smoothing is $1/\hat{\tau}_t$, and that the pseudocounts must reflect this:

$$\frac{c_1}{c_0 + c_1} = \frac{1}{\hat{\tau}_t},$$

and hence $c_0 = c_1(\hat{\tau}_t - 1)$, so we can eliminate it. Like TSB, in order to obtain a forecast, we multiply this probability by the smoothed demand size:

$$f_t = \frac{\hat{y}_t}{\hat{\tau}_t + \tau_t/c_1}.$$

We can also eliminate c_1 by choosing a value that gives an unbiased forecaster for stochastic intermittent demand, as follows. 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, so our forecaster has an expected forecast

$$\mathbb{E}\left[\frac{\hat{y}_t}{\hat{\tau}_t + \tau_t/c_1}\right] = \mathbb{E}\left[\frac{\hat{y}_t}{\hat{\tau}_t}\left(\frac{1}{1 + \tau_t/\hat{\tau}_t c_1}\right)\right]$$
$$\approx \mathbb{E}\left[\frac{\hat{y}_t}{\hat{\tau}_t}\left(1 - \frac{\tau_t}{\hat{\tau}_t c_1}\right)\right] = \frac{\hat{y}_t}{\hat{\tau}_t}\left(1 - \frac{1}{c_1}\right).$$

The derivation used the linearity of expectations, the constancy of $\hat{y}_t$ and $\hat{\tau}_t$, the fact that $\mathbb{E}[\tau_t] = \hat{\tau}_t$, and the approximation $(1+\delta)^{-1} \approx 1-\delta$ for small δ. Choosing $c_1 = 2/\beta$, and therefore $c_0 = 2(\hat{\tau} - 1)/\beta$, we obtain a forecast

$$f_t = \frac{\hat{y}_t}{\hat{\tau}_t + \beta\tau_t/2},$$

with expected value

$$\mathbb{E}[f_t] = \left(1 - \frac{\beta}{2}\right)\frac{\hat{y}_t}{\hat{\tau}_t}.$$

Thus, our expected forecast over a substring is identical to the fixed SBA forecast over that substring. So, our forecaster has the same expected forecast as SBA on any substring, given the same values of $\hat{y}_t$ and $\hat{\tau}_t$. Moreover, it updates $\hat{y}_t$ and $\hat{\tau}_t$ in exactly the same way as SBA at the start of each substring, and therefore it has the same expected forecast as SBA over the entire demand sequence. Thus, according to Syntetos and Boylan (2005), it is unbiased for stochastic intermittent demand.

One drawback of this forecaster is that, like SBA, it is biased for non-intermittent demand. This can be overcome by a slight adjustment:

$$f_t = \frac{\hat{y}_t}{\hat{\tau}_t + \beta(\tau_t - 1)/2}.$$

By a similar derivation, we obtain:

$$\mathbb{E}\left[\frac{\hat{y}_t}{\hat{\tau}_t + \frac{\beta}{2}(\tau_t - 1)}\right] = \left(1 - \frac{\beta}{2}\right)\frac{\hat{y}_t}{\hat{\tau}_t - \frac{\beta}{2}}.$$

Now our expected forecast over a substring is identical to the fixed SY forecast over that substring. SY is unbiased on standard stochastic intermittent demand and also on non-intermittent demand (Syntetos, 2001), so (using the same arguments as above) our forecaster must be too. This is the forecaster we shall use, and we call it Hyperbolic-Exponential Smoothing (HES) because of its combination of exponential smoothing with hyperbolic decay.

One might ask: why not also apply the same Bayesian reasoning when demand *does* occur? We can do this by incrementing the count n_1 each time a demand occurs, and incrementing n_0 when no demand occurs (the pseudocount values c_0, c_1 can be set to 0). The drawback of this method is that it does not adapt quickly to changes in demand intermittency: it uses the entire demand history to estimate the probability of a demand as $n_1/(n_0 + n_1)$. Our proposed method discounts the early demand history in standard exponential smoothing fashion.

An illustration of the different behaviours of SY, TSB and HES is shown in Fig. 1. The demand is stochastic intermittent with probability 0.25, all nonzero demands (shown as impulses) are 1, and the forecasters use $\alpha = \beta = 0.1$. The HES forecasts are similar to those of SY but decay between demands. TSB has a much greater variation, though this difference could be reduced by using smaller smoothing parameters. When an item becomes obsolete, SY remains constant, TSB decays exponentially and HES decays hyperbolically.

4. Experiments

We now test HES empirically to evaluate its bias and forecasting accuracy. All results are given by Prestwich et al. (2013).

4.1. Accuracy measures

In any comparison of forecasting methods, we must choose accuracy measures. de Gooijer and Hyndman (2005) list 17 measures, noting that a "bewildering array

[2] These pseudocounts are often denoted α, β from the Beta distribution hyperparameters, but we are already using these symbols for smoothing factors.

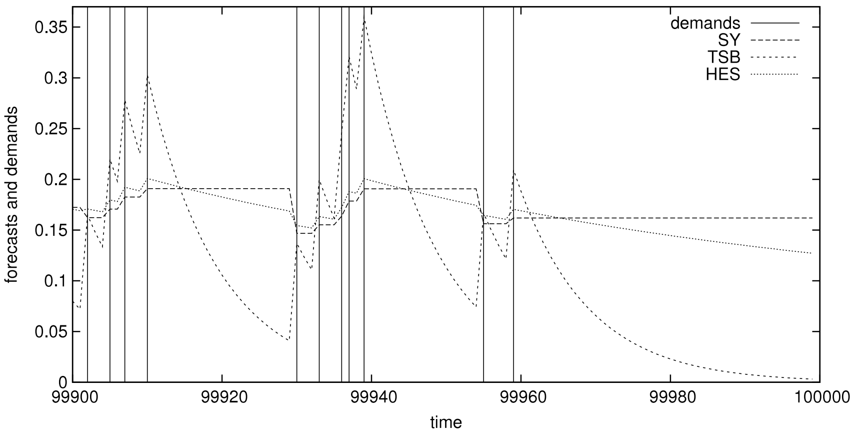

Fig. 1. Behaviour of SY, TSB and HES.

of accuracy measures have been used to evaluate the performance of forecasting methods", that no single method is generally preferred, and that some are not well-defined on data with intermittent demand. We shall use measures that have been recommended recently for intermittent demand. For measuring the bias, we use the MASE (Mean Absolute Scaled Error), recommended by Hyndman and Koehler (2006) and defined as mean($|q_t|$), where q_t is a scaled error defined by

$$q_t = \frac{e_t}{\frac{1}{n-1}\sum_{i=2}^{n} |y_i - y_{i-1}|},$$

e_t is the error $y_t - \hat{y}_t$, and $t = 1 \ldots n$ are the time periods of the samples used for forecasting, which we take to be the 10^4 samples used to initialise the smoothed estimates. We take these means over multiple runs. MASE effectively evaluates a forecasting method against the *naïve* (or *random walk*) forecaster, which simply forecasts that the next demand will be identical to the current demand.

As a measure of deviation, we use the MAD/Mean Ratio (MMR), which has been argued to be superior to several other methods used in forecasting competitions (Kolassa & Schütz, 2007), and is defined by

$$\frac{\sum_{t=1}^{n} |e_t|}{\sum_{t=1}^{n} y_t}.$$

Again, the summations are taken over multiple runs. As another measure of deviation, we also use the Relative Root Mean Squared Error, defined as RMSE/RMSE$_b$, where RMSE is measured on the method being evaluated and RMSE$_b$ on a baseline measure, both taken over multiple runs. When the baseline is a random walk, this is Theil's U2 statistic (Thiel, 1966), and this is the baseline we use. The motivation behind using these two particular measures of deviation is that MMR is based on absolute errors while U2 is based on root mean squared errors; the latter penalises outliers more than the former, so differences between them could be revealing.

4.2. Experimental details

We base our experiments on those of Teunter et al. (2011), in which demands occur with some probability in each period, and hence, the inter-demand intervals are distributed geometrically, and we use a logarithmic distribution for demand sizes. 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, for which there is theoretical and empirical evidence; see for example the recent discussion by Syntetos, Babai, Lengu, and Altay (2011).

Teunter et al. compare several forecasters on demands that are nonzero with probability p_0, where p_0 is either 0.2 or 0.5, and the sizes of which are distributed logarithmically. 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 $k \geq 1$. They use two values: $\ell = 0.001$ to simulate low demand and $\ell = 0.9$ to simulate lumpy demand. They use α values of 0.1, 0.2 and 0.3, and β values of 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3. We add SY and HES to these experiments but drop SES, as they found it to have large errors. They take mean results over 10 runs, each with 120 time periods, whereas we use 100 runs. They initialise each forecaster with "correct" values whereas we initialise with arbitrary values $\hat{y}_0 = \hat{\tau}_0 = 1$ then run them for 10^4 periods using the demand probability p_0. A final difference is that we use MASE, MMR and U2 instead of the mean error and mean squared error.

4.3. Results

The results are given by Prestwich et al. (2013) and summarised here. Because CR, SBA and SY use only one smoothing factor α, we do not perform experiments on these methods with $\beta \neq \alpha$. Comparing the MASE best-cases: TSB and SY are least biased, followed by HES, then CR and SBA. Comparing the MMR best-cases: SBA is best, followed by TSB and HES, then CR and SY. Comparing the U2 best-cases: HES is always at least as good as TSB, SBA is best in some cases, and CR and SY are the worst. Comparing the MASE worst-cases: the ranking is SY, TSB, HES, CR, SBA (best first). Comparing the MMR worst-cases: neither TSB nor HES dominates the other, though HES seems slightly better. SBA gives the best results, CR and SY generally the worst. Comparing the U2 worst-cases: HES beats TSB and is more robust under different smoothing factors. SBA again gives the best results, while CR and SY have variable performances.

These results agree with the known fact that SY has a lower bias but higher variance than SBA. In terms of bias (as measured by MASE), HES lies between SY and SBA, while TSB is similar to SY; in terms of variance (as measured by MMR), HES and TSB lie between SY and SBA and are similar; in terms of variance (as measured by U2) HES and TSB lie between SY and SBA, but HES is more robust under parameter change. This shows that, although we modified HES to be more like SY than SBA, in order to gain its advantage of being unbiased on non-intermittent demand, HES has not inherited the high variance of SY. In this sense it is a good compromise between SBA and SY, with a low bias *and* low variance.

We performed further experiments as follows. Firstly, under two forms of nonstationary demand with obsolescence, TSB beats HES, which in turn beats the other variants. Secondly, we compared HES and TSB using stationary demand with geometrically distributed demand sizes. The relative performances of TSB and HES were similar to those with logarithmically distributed demand sizes. Thirdly, we compared TSB and HES on stationary demand using two relative measures: the percentage of times better (recommended by Kolassa & Schütz, 2007) and the Relative Geometric Root Mean Squared Error (recommended by Armstrong & Collopy, 1992, and also called the Geometric Mean Relative Absolute Error). Tuning TSB and HES using U2 as recommended by Ghobbar and Friend (2003), there was little difference in relative performance, while TSB beat HES when we tuned using MMR.

Finally, Teunter et al. (2011) note that although TSB is unbiased when the bias is computed over all points in time, it is nevertheless biased if we only compare forecasts with the expected demand at issue points only (that is, when demand occurs). SES is similarly biased, but Croston methods such as SBA or SY are not. This is due to the decay in forecast size between demands, and HES will clearly suffer from a similar bias. This bias is relevant because of the way in which forecasts are used in real inventory control systems: they are often made only when demands occur. We repeated the stationary demand experiments with logarithmically and geometrically distributed demand sizes, and measured the biases of TSB and HES based on issue points only. Both had larger biases than SY (as expected) but neither dominated the other.

4.4. Summary of results

For stationary demand, HES and TSB are both good compromises between the low bias of SY and the low variance of SBA, to some extent achieving the best of both worlds. No single method dominated all others over all experiments, and which is best depends on the demand pattern and the accuracy measure. However, comparing TSB and HES over all experiments with stationary demand, we observe a clear pattern: TSB wins (best-case and worst-case) using MASE, HES wins (best-case and worst-case) using U2, TSB wins (best-case only) using MMR, and HES wins (worst-case only) using MMR. Thus, neither HES nor TSB is best under all measures, and they have similar biases when measured at issue points only. However, the greater robustness of HES (shown by its better worst-case performance) means that we can recommend smoothing factors that behave reasonably well, on both stationary and non-stationary demand: $\alpha = \beta = 0.1$. In all cases, the results are not very much worse than those with optimally-tuned factors.

However, TSB wins in the obsolescence experiments, followed by HES. As has been found by other researchers, and as is intuitively clear, large smoothing factors are best at handling changes in demand pattern, and here TSB's faster reaction times serve it well. Nevertheless, HES is still much better than the other Croston methods when faced with obsolescence.

5. Conclusion

We have presented a new forecasting method called Hyperbolic-Exponential Smoothing (HES), which combines exponential smoothing with an application of Bayesian inference when no demand occurs. It is only the second method to be proposed for handling intermittent demand with possible obsolescence, and it is qualitatively different to the other existing method (TSB), providing practitioners with an alternative.

We have shown theoretically that HES is approximately unbiased on stochastic intermittent and non-intermittent demand, and compared it empirically with four other Croston variants, including TSB. For stationary demand, we found that HES was best under one measure (U2), TSB was best under another measure (MASE), and neither was best under a third measure (MAD/Mean Ratio). Overall, HES was more robust than TSB under smoothing parameter changes, while TSB handled obsolescence better than HES. These results show that HES is a competitive forecaster. Its robustness allows us to recommend default smoothing factor values of 0.1, making it highly suitable for use in automated systems.

Regarding the practical use of HES, we note that it is no harder to implement, and no more computationally expensive, than other Croston variants; this can be seen by comparing the formulae for forecasts in HES (see Section 3) with those of other methods (see Section 2).

In future work, we hope to explore the issue of robustness under smoothing factor changes further, using other demand patterns and accuracy measures. In particular, we would like to obtain real data, and to devise other artificial data with obsolescence. For example, TSB handles obsolescence better than HES, but how would the two compare on demand that appears to be obsolescent temporarily? More experiments along these lines are needed.

Acknowledgments

Thanks to Aris Syntetos for helpful advice, and to the anonymous referees for useful criticism. This work was partially funded by Enterprise Ireland Innovation Voucher IV-2009-2092.

References

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: empirical comparisons. *International Journal of Forecasting*, *8*, 69–80.

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

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

de Gooijer, J. D., & Hyndman, R. J. (2005). *25 years of IIF time series forecasting: a selective review*. Tinbergen Institute Discussion Paper No. 05-068/4, Tinbergen Institute.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. A. (2008). Forecasting and operational research: a review. *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 and 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.

Kolassa, S., & Schütz, W. (2007). Advantages of the MAD/mean ratio over the MAPE. *Foresight*, *6*, 40–43.

Laplace, P.-S. (1814). *Essai philosophique sur les probabilités*. Paris: Courcier.

Levén, E., & Segerstedt, A. (2004). Inventory control with a modified Croston procedure and Erlang distribution. *International Journal of Production Economics*, *90*(3), 361–367.

Poole, D., & Mackworth, A. (2010). *Artificial intelligence: foundations of computational agents*. Cambridge University Press.

Prestwich, S. D., Tarim, S. A., Rossi, R., & Hnich, B. (2013). Forecasting intermittent demand by hyperbolic-exponential smoothing. http://arxiv.org/abs/1307.6102.

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

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*, 606–615.

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

S.D. Prestwich is a lecturer in Computer Science at University College, Cork, Ireland. He holds an M.A. in mathematics from Oxford University, UK, and an M.Sc. and Ph.D. in computer science from the University of Manchester, UK. He has worked for several companies, including the European Computer-Industry Research Centre, Munich, Germany. His research interests include forecasting, reasoning under uncertainty, constraint programming, metaheuristics, artificial intelligence, operations research and hybrid algorithms.

S.A. Tarim is Professor of Operations Research at Hacettepe University, Ankara, Turkey. He holds a B.Sc. in electronic engineering (1989) from the Middle East Technical University, Ankara, and a Ph.D. in operational research (1996) from Lancaster University, UK. He has also held positions at the Turkish Court of Accounts as Chief Advisor to the President, the University of York, UK, the National University of Ireland in Cork, Ireland, and the University of Nottingham, UK. His research interests include energy planning, stochastic modeling and risk management, and modeling of production/inventory systems.

R. Rossi holds a B.Eng. and a M.Eng. from the University of Bologna, Italy, and a Ph.D. from University College, Cork, Ireland. He is currently an Assistant Professor at Wageningen University, the Netherlands. He spent four years at the Cork Constraint Computation Centre, Department of Computer Science, University College, Cork, as a research assistant, Ph.D. student and postdoctoral researcher. His research is focused on automated reasoning and on the development of systems that aim to be robust and scalable in such a way as to enable computers to act intelligently in increasingly complex real world settings and in uncertain environments.

B. Hnich is Professor of Computer Science at the Department of Computer Engineering, Izmir Economic University, Turkey. He holds a B.Sc. in computer engineering from Bilkent University, Ankara, Turkey, and a Ph.D. from Uppsala University, Sweden. He was a Research Fellow at the Cork Constraint Computation Centre, University College, Cork, Ireland, has a Habilitation from the University of Montpellier II, France, and is a Docent at Uppsala University, Sweden. He is also an Associate Editor of the *Artificial Intelligence Journal*. His primary area of research is in artificial intelligence.