Journal of Retailing and Consumer Services 62 (2021) 102662

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

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

Forecasting intermittent demand for inventory management by retailers: A new approach

Xin Tian [a,b], Haoqing Wang [a,b], Erjiang E [c,*]

[a] School of Economics and Management, University of Chinese Academy of Sciences, Beijing, China
[b] Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing, China
[c] Department of Industrial Engineering, Tsinghua University, Beijing, China

ARTICLE INFO

Keywords:
Intermittent demand
Forecasting
Retailing
Markov-combined method
Inventory

ABSTRACT

The forecasting of intermittent demand is a complex task owing to demand fluctuations and interval uncertainty. Intermittent demand is essentially random demand with a high percentage of zero values. In the retail industry, there are many products which face intermittent demand and this poses a problem of inventory management. This study proposes a Markov-combined method (MCM) for forecasting intermittent demand, which takes into account the inventory status and historical sales of products. We divide the prediction process into two stages. In the first stage, the transition probabilities of the four basic states of demand and inventory are calculated. In the second stage, the corresponding and appropriate prediction method is selected according to the predicted state. Further, using two large datasets from the two biggest e-commerce companies in China, we verify our results and show that the MCM forecasts more accurately than the Single Exponential Smoothing (SES), Syntetos-Boylan Approximation (SBA), and Croston (CR) methods. The MCM can be as an alternative method for forecasting intermittent demand because it is easy to compute and typically more accurate than the classical forecasting methods.

1. Introduction

For retailers, investment in inventory occupies a significant proportion of their resources. Inventory management has therefore been the subject of much discussion and many studies since the 1900s. In general, the reduction of inventory levels and achievement of high inventory turnover are goals that most retailers have been pursuing in their practical operations. Low liquidity caused by high inventory backlog or poor customer experiences ensuing from a shortage of inventory, arise from decisions or steps taken during inventory management. Such decisions usually depend on the inventory management system employed, especially the subsystem called the forecasting support system (FSS). Demand forecasting is an integral part of inventory management. Accurate demand forecasting can improve the competitiveness of the organization(Veiga et al., 2016) and is essential for appropriate decision-making (Lackes et al., 2020), that provides the basis for replenishment and distribution plans and supply chain management. Retail companies have continuously sought efficient and accurate forecasting methods over the past several decades.

In terms of volatility and continuity, demand can be classified into four different patterns: intermittent, lumpy, smooth, and erratic (Syntetos et al., 2005). Intermittent demand, which is characterized by a high proportion of zero values, is a common phenomenon in the real environment. Such demand patterns are common for spare parts SKUs in the military and aerospace and automotive industries (Babai et al., 2019). It also applies almost universally to the retail sector. Fig. 1 shows an example of an item selling on *Tmall.com*, the largest business-to-consumer (B2C) retail platform in Asia. Demand for the item is intermittent, with no demand for almost 87% of the time. Intermittent demand could lead to inappropriate stock levels in the store or the delivery center. Although there are many products exhibiting intermittent demand (Babai et al., 2014), this pattern has not been the subject of much research in the area of retail forecasting (Fildes et al., 2019).

In recent years, online retailing has been developing rapidly, especially with the widespread use of applications (Li et al., 2020). In China, online retail sales of goods and services crossed ten trillion CNY and online retail had a market penetration rate of 20.7% in 2019. The vast variety and number of SKUs, as well as the need for high-speed response seriously challenge and test online retailers' inventory management. Accurate forecasting of commodity demand is therefore, crucial for

* Corresponding author. Shuangqing Road 30, Haidian District, Beijing, 100084, China.
E-mail address: eej17@mails.tsinghua.edu.cn (E. E).

https://doi.org/10.1016/j.jretconser.2021.102662
Received 27 September 2020; Received in revised form 1 March 2021; Accepted 28 June 2021
Available online 7 July 2021
0969-6989/

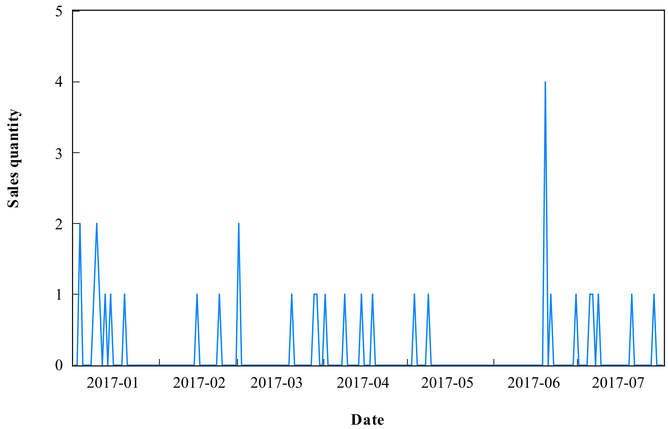

Fig. 1. Time series of an item.

online retailers.

The forecasting of intermittent demand is challenging. The difficulties arise mainly from two aspects. One is that the actual demand is sporadic in nature, and the second factor is the unpredictability of the timing of demand (Nikolopoulos, 2020). Thus, both the quantity and occurrence time of demand need to be predicted. Many forecasting methods of intermittent demand have been developed, such as the improvement of the Croston (CR) method (Croston, 1972; Syntetos and Boylan, 2001), the bootstrapping methods (Hasni et al., 2019), the machine learning methods (Lolli et al., 2017), and the aggregation approach methods (Nikolopoulos, 2020). These methods have mainly been applied to the forecasting of spare parts (Hua and Zhang, 2006; Pennings, van Dalen and van der Laan, 2017). To the best of our knowledge, research in forecasting intermittent demand for retailing is scarce.

In this study, we proposed the Markov-combined model (MCM) to predict intermittent demand. In addition to considering historical sales data, this model takes the inventory status of the product into account. For each product, there are two inventory states: in stock and out of stock; and two demand states of, in demand and no demand. Thus, in combination, there are four possible states: in stock and in demand, in stock and no demand, out of stock and in demand, out of stock and no demand. These states form a 2×2 probability transition matrix, which can be solved by the Markov model. If, from one state, the next one has no demand, the forecasting demand is zero. If the next state is, in stock and in demand, we can use ordinary prediction methods, such as exponential smoothing to forecast requirements. Finally, if the next state is, out of stock and in demand, the out of stock status may affect real demand. We compare the MCM with some classical intermittent demand forecasting models, including the Single Exponential Smoothing (SES), CR and Syntetos-Boylan Approximation (SBA) methods. The results show that the proposed MCM outperforms the classical intermittent demand forecasting models.

The rest of the paper is organized as follows: we briefly review research relevant to intermittent demand forecasting in Section 2; describe the proposed model for intermittent demand forecasting in Section 3; and examine the performance of the proposed model by using two real-world problems in Section 4. Finally, Section 5 concludes the study and discusses future research prospects.

2. Literature review

Over the past few decades, many models for demand prediction and forecasting have been proposed and thereafter, improved upon based on the specific application requirements. Nikolopoulos (2020) argues that much attention has been paid to modeling fast-moving time series, but limited attention has been paid to intermittent time series and intermittent demand forecasting. The CR method (Croston, 1972) is the seminal work on intermittent demand forecasting and has proven to be practically useful. However, it has the inherent problem of bias (Syntetos and Boylan, 2001). Instead, Syntetos and Boylan (2001) have proposed the SBA method, which is unbiased and found to be an effective improvement over the CR method. Prestwich, Tarim, Rossi, and Hnich (2014) also propose an unbiased model that combines the CR method with Bayesian inference. Intermittent demand predictions are further confounded by the issue of obsolescence, and this problem has aroused the attention of many scholars (Babai et al., 2019; Prestwich et al., 2014; Teunter et al., 2011).

The existing methods for intermittent demand forecasting are summarized as follows. First, the bootstrapping approaches, which do not rely on probability distribution (Hasni et al., 2019b). Secondly, the parametric methods, such as the CR and SBA methods that use historical data to estimate key parameters. Thirdly, machine learning methods, especially neural networks are widely applied to intermittent demand forecasting. For example, Kourentzes (2013) proves that neural networks are effective for predicting intermittent time series; Lolli et al. (2017) use feedforward single-hidden layer neural networks and consider different aggregation levels. Fourthly, some researchers use the correlation between demand and its intervals to predict intermittent demand. Pennings et al. (2017) consider the cross-correlation between inter-arrival time and demand quantity, and their experiment shows that the model performs well in spare parts management and bring financial benefit. Altay, Litteral, and Rudisill (2012) consider the different types of correlation between demand and interval, and find that different correlations lead to different service levels. Other interesting research, such as time aggregation methods (Kourentzes et al., 2017; Nikolopoulos, 2020; Petropoulos et al., 2016) and methods using expert judgments (Syntetos et al., 2009) also play an important role in intermittent demand forecasting. Willemain, Smart, and Schwarz (2004) propose the WSS algorithm, which includes the process of estimating transition probabilities for the two-state (zero and non-zero) Markov model. In this study, we also use the Markov model, but our model considers four-states (inventory states and demand states) instead of two. To the best of our knowledge, there is no other research that uses inventory-state information to forecast intermittent demand.

In recent years, the practical performance of forecasting models has also attracted the attention of scholars. Teunter and Duncan (2009) compare the performance of the moving average, SES, CR and bootstrapping methods. The results indicate that the CR and bootstrapping methods are better than the former two methods. Babai, Ali, and Nikolopoulos (2012) empirically study the importance of temporal aggregation in intermittent demand forecasting. Syntetos, Babai, and Gardner (2015) examine the performance of parametric and non-parametric methods using 7000 demand time series, and the conclusion is that simple parametric methods are more appropriate. Fu and Chien (2019) develop a UNISON-based data driven framework to forecast intermittent demand for electronics components. Ghobbar and Friend (2003) apply several forecasting methods to aircraft maintenance repair parts, and show that the Holt and CR methods are adaptable and well-suited to such analyses.

Additionally, little attention has been paid to intermittent demand in industrial retail. Fildes et al. (2019) review retail forecasting in detail, and point out that, “intermittent demand is a key problem where current research has not been adopted”. In this study, we review the intermittent demand pattern in the retail industry and propose the use of the MCM. We formulate inventory and demand states into a four-state Markov model, and adopt different forecasting methods according to the different states. Two datasets from the B2C platform are used to verify the performance of our model.

	(0, 0)	(0, 1)	(1, 0)	(1, 1)
(0, 0)	p_{11}	p_{12}	p_{13}	p_{14}
(0, 1)	p_{21}	p_{22}	p_{23}	p_{24}
(1, 0)	p_{31}	p_{32}	p_{33}	p_{34}
(1, 1)	p_{41}	p_{42}	p_{43}	p_{44}

Fig. 2. The transition probability matrix.

3. Methodology

3.1. Benchmark models

Three classical methods have been adopted as benchmark models, to compare with the proposed method. The three methods are SES, CR and SBA methods.

3.1.1. Simple exponential smoothing (SES)

The SES method is a compromise between the Naïve and moving average (MA) methods. The Naïve method equates predicted values to the last period's actual observed values. The MA method calculates the average of all historical data to get a predictive result, and treats each period equally. SES considers all the observed values and gives greater weight to recent observation data by using the smoothing parameter:

$$\hat{y}_t = \alpha y_{t-1} + \alpha(1-\alpha)y_{t-2} + \alpha(1-\alpha)^2 y_{t-3} + \ldots + \alpha(1-\alpha)^{t-2} y_1 \qquad (1)$$

where $0 \leq \alpha \leq 1$ is the smoothing parameter; a larger value of α indicates the increased importance of recently observed data. $y_1, \ldots, y_{t-1}$ are the observed historical data and $\hat{y}_t$ is the forecast result. SES is suitable for data without obvious time trends. Naïve, MA and SES are classical, widely-used statistical prediction methods.

3.1.2. Croston method (CR)

The CR method (Croston, 1972) is a pioneering work on intermittent demand. This method assumes that demand quantity follows a normal distribution and that the demand interval is stable. Letting z_t be the

Step 0 Obtain historical demand data and inventory state in chosen time buckets.

Step 1 Estimate transition probabilities for four-state (zero vs. nonzero for demand & in stock vs. out stock for inventory) Markov model.

Step 2 Conditional on last observed demand and inventory, use Markov model to generate a sequence of state values over forecast horizon.

Step 3 Choose predictive models based on the predictive states. (1) If state value is (0, 0) or (0, 1), choose Naïve method and then the predictive demand is zero. (2) Else if state value is (1, 0), choose Average method and then the predictive demand is the average value of historical demand state (1, 0). (3) Else state value is (1, 1), choose Single Exponential Smooth method and the training data is from the set of historical state (1, 1).

Step 4 Repeating Steps 2-3 over forecast horizon. Obtain the forecasting results.

Fig. 3. A brief summary of the steps in the Markov-combined method.

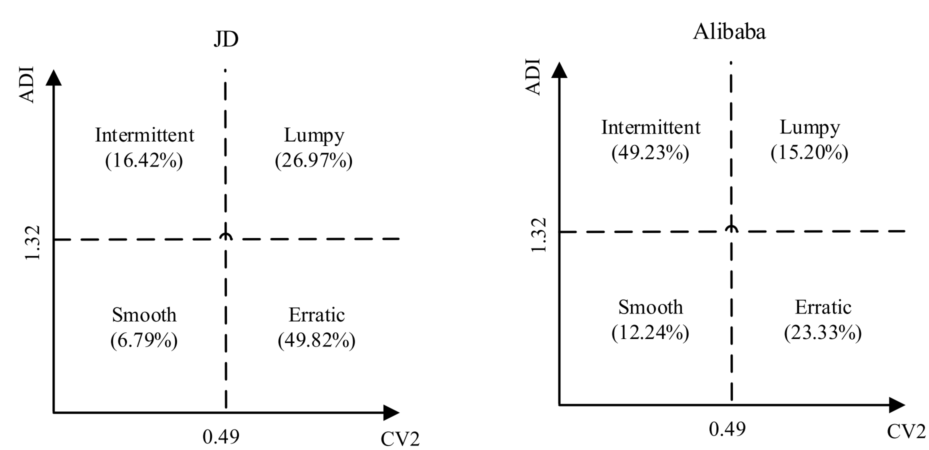

Fig. 4. The classification of demand patterns.

Table 1
Characteristics of the sales of both online retailers of JD and Alibaba.

	JD	Alibaba
Data features		
No. series	963	3106
Mean obs./series	209.369	142.341
% Zero values		
Mean	0.796	0.853
S.D.	0.105	0.110
Maximum	0.989	0.991
Minimum	0.567	0.550
Average of nonzero demand		
Mean	1.546	1.491
S.D.	0.813	0.717
Maximum	48.000	413.000
Minimum	1.000	1.000
CV of nonzero demand		
Mean	0.476	0.438
S.D.	0.185	0.185
Maximum	0.700	0.700
Minimum	0	0
ADI		
Mean	7.901	21.860
S.D.	12.543	35.814
Maximum	172.000	210
Minimum	1.333	1.326
% Vendibility		
Mean	0.619	0.914
S.D.	0.323	0.144
Maximum	1.000	1.000
Minimum	0.000	0.009

Notes: ADI: average demand interval; CV: the coefficient of variation; S.D.: standard deviation.

estimated value of non-zero demand at period t, and x_t be the actual demand value in the equation below:

$$z_t = \alpha x_{t-1} + (1-\alpha) z_{t-1} \tag{2}$$

Let v_t be the average of demand intervals, and q be the consecutive zero demand (i.e., the time series {2, 0, 0, 0, 0, 1}, then $q = 4$):

$$v_t = \alpha q + (1-\alpha) v_{t-1} \tag{3}$$

If x_{t-1} is positive, the final estimated value is $y_t = \frac{z_t}{v_t}$; otherwise $y_t = y_{t-1}$, $z_t = z_{t-1}$, $v_t = v_{t-1}$.

3.1.3. Syntetos-Boylan Approximation (SBA) method

The SBA method (Syntetos and Boylan, 2001) is considered an effective improvement over the CR method. SBA introduces the corrective factor, which gives the final estimated value as:

$$y_t = \left(1 - \frac{\beta}{2}\right) \frac{z_t}{v_t} \quad \text{if} \quad x_{t-1} \quad \text{is} \quad \text{positive.}$$

Both CR and SBA methods are highly regarded methods in the field of intermittent demand forecasting.

3.2. The markov-combined method (MCM)

The stochastic process X is a discrete-time Markov chain with state space S if $j \in S$ and $k = 0, 1, 2...$ in equation (4):

$$P[X_{k+1} = j | X_0 = i_0, X_1 = i_1, ..., X_k = i_k] = P[X_{k+1} = j | X_k = i_k] \tag{4}$$

for any states $i_0, ..., i_k$ in the state space S.

When the actual problem can be formulated by the Markov chain, we must first determine its state space, and then estimate the transition probability. The estimation of this probability can be obtained from the objective laws of the problem, past experience, or based on observation data. Here we use observation data to calculate transition probability.

Supposing we know the historical data of goods sales and inventory vendibility; then letting y_t be the sales of goods at the end of day t ($y_t \geq$ 0), q_t be the inventory level ($q_t \geq 0$), x_{1t} be the sales state of goods at the end of day t ($x_{1t} = 0$ *or* 1), x_{2t} be the inventory state of goods ($x_{2t} = 0$ *or* 1), we get:

Table 2
Forecast performance for rolling forecast testing – JD' dataset.

	MAE	RMSE	sMAPE	% improvement of MCM		
				MAE	RMSE	sMAPE
Horizon = 1 (Obs. = 927*1*7)						
SES	0.474	0.598	1.823	30.802	−2.508	77.619
	(0.565)	(0.713)	(0.204)			
Croston	0.491	0.610	1.816	33.198	−0.492	77.533
	(0.569)	(0.697)	(0.206)			
SBA	0.495	0.617	1.800	33.737	0.648	77.333
	(0.469)	(0.620)	(0.268)			
MCM	0.328	0.613	0.408			
	(0.560)	(0.880)	(0.363)			
Horizon = 3 (Obs. = 927*3*7)						
SES	0.476	0.613	1.823	31.092	23.589	77.784
	(0.593)	(0.748)	(0.195)			
Croston	0.494	0.626	1.816	33.603	26.210	77.698
	(0.599)	(0.743)	(0.198)			
SBA	0.478	0.625	1.840	31.381	20.640	77.989
	(0.633)	(0.750)	(0.177)			
MCM	0.328	0.496	0.405			
	(0.719)	(1.001)	(0.494)			
Horizon = 5 (Obs. = 927*5*7)						
SES	0.476	0.624	1.821	45.566	8.333	77.814
	(0.593)	(0.755)	(0.198)			
Croston	0.496	0.637	1.814	51.682	10.204	77.729
	(0.598)	(0.760)	(0.200)			
SBA	0.479	0.637	1.839	31.733	10.204	78.032
	(0.629)	(0.776)	(0.169)			
MCM	0.327	0.572	0.404			
	(0.650)	(0.973)	(0.403)			
Horizon = 7 (Obs. = 927*7*7)						
SES	0.479	0.635	1.819	31.106	2.205	77.680
	(0.607)	(0.771)	(0.201)			
Croston	0.499	0.647	1.812	33.868	4.019	77.594
	(0.613)	(0.785)	(0.203)			
SBA	0.483	0.651	1.834	31.677	4.608	77.863
	(0.641)	(0.829)	(0.185)			
MCM	0.330	0.621	0.406			
	(0.654)	(0.971)	(0.358)			

Notes: Obs.: Observations = item number * forecasting horizon(s) * rolling rounds. The first three columns show means (standard deviations).

$$x_{1t} = \begin{cases} 1, & \text{if} \quad y_t > 0; \\ 0, & \text{if} \quad y_t = 0. \end{cases}, \quad x_{2t} = \begin{cases} 1, & \text{if} \quad y_t < q_t; \\ 0, & \text{if} \quad y_t \geq q_t. \end{cases} \tag{5}$$

In summary, there are four states:

State 1 (0,0), which means no demand or lost sale; State 2 (0,1): which means no demand; State 3 (1,0): which means sold out at the end of the day; State 4 (1,1): which is a perfect state.

The transition probability matrix is shown in Fig. 2. Using historical data, we calculate this transition probability matrix. For each state, we compare the four probabilities and choose the highest one as the next state. Different forecasting methods such as Naïve, average method and SES methods are used based on the next predictive states. Since this is a method that combines the Markov model and other predictive methods, we name it the Markov-combined method (MCM). A brief summary of the steps in MCM is shown in Fig. 3.

4. Empirical analysis

In this section, we conduct empirical analysis with the aim of answering the following research questions:

RQ1 Does our proposed MCM model outperform the classical methods including SES, SBA, and CR?

RQ2 Under what demand patterns is the MCM model superior to benchmark models?

RQ3 What are the differences in the performance of the MCM model

Table 3
Forecast performance for rolling forecast testing – Alibaba' dataset.

	MAE	RMSE	sMAPE	% improvement of MCM		
				MAE	RMSE	sMAPE
Horizon = 1 (Obs. = 2994*1*7)						
SES	0.295	0.386	1.871	25.763	−1.036	82.416
	(0.283)	(0.360)	(0.237)			
Croston	0.374	0.464	1.881	41.444	15.948	82.509
	(1.257)	(1.273)	(0.205)			
SBA	0.311	0.397	1.875	29.582	1.763	82.453
	(0.341)	(0.395)	(0.222)			
MCM	0.219	0.390	0.329			
	(0.315)	(0.450)	(0.384)			
Horizon = 3 (Obs. = 2994*3*7)						
SES	0.286	0.377	1.877	29.021	22.281	84.017
	(0.286)	(0.358)	(0.231)			
Croston	0.363	0.454	1.887	44.077	35.463	84.102
	(1.260)	(1.274)	(0.199)			
SBA	0.309	0.400	1.884	34.304	26.750	84.076
	(0.354)	(0.413)	(0.207)			
MCM	0.203	0.293	0.300			
	(0.387)	(0.506)	(0.482)			
Horizon = 5 (Obs. = 2994*5*7)						
SES	0.286	0.381	1.874	29.720	12.598	84.152
	(0.285)	(0.358)	(0.244)			
Croston	0.363	0.458	1.884	44.628	27.293	84.236
	(1.260)	(1.274)	(0.216)			
SBA	0.309	0.404	1.880	34.951	17.574	84.202
	(0.354)	(0.412)	(0.225)			
MCM	0.201	0.333	0.297			
	(0.342)	(0.483)	(0.417)			
Horizon = 7 (Obs. = 2994*7*7)						
SES	0.287	0.386	1.869	30.314	7.513	84.109
	(0.284)	(0.356)	(0.257)			
Croston	0.363	0.462	1.880	44.904	22.727	84.202
	(1.260)	(1.274)	(0.233)			
SBA	0.310	0.407	1.875	35.484	12.285	84.160
	(0.353)	(0.413)	(0.240)			
MCM	0.200	0.357	0.297			
	(0.319)	(0.466)	(0.386)			

itself under different demand patterns?

4.1. Empirical data

Two datasets used are from JD and Alibaba, which are the two biggest e-commerce companies in China. JD's inventory data ranges from January 1, 2016 to December 31, 2017, and Alibaba's inventory data ranges from January 1, 2017 to July 1, 2017. The datasets of JD and Alibaba respectively contain 5973 items and 7883 items of different categories and various time horizons. It should be noted that JD's data in June and November are excluded to prevent the influence of promotional activities. Demand can be divided into four patterns: smooth, erratic, intermittent, and lumpy. The indicators, on which the classification is based, are the coefficient of variations (CV) and average demand interval (ADI). CV reflects the volatility in demand. ADI is the average demand interval between non-zero demands. According to Syntetos et al. (2005), if $CV^2 < 0.49$ (or $CV < 0.7$) and ADI>1.32, the demand pattern is intermittent. According to this rule, we calculate CV^2 and ADI for the two datasets. Finally, there are 963 items in JD's data and 3106 items in Alibaba's data that meet the intermittent demand pattern, which account for 16.42% and 49.23% respectively of the original items. The classification of demand patterns is shown in Fig. 4, which indicates that a considerable proportion of goods exhibit intermittent demand. The detailed descriptive statistics of ADI and CV on these intermittent demand items are shown in Table 1. The ADI of JD and Alibaba are 7.019 and 21.860 respectively, and the CV of nonzero demand of JD and Alibaba are 0.476 and 0.438 respectively. These fully meet the standard/criteria for classification as intermittent demand.

4.2. Settings and performance evaluation

SES is a simple and effective method that is widely used in forecasting. The CR and SBA methods are regarded as classical models for intermittent demand forecasting, which have previously proved to be effective. Thus, we choose SES, CR and SBA methods to be the benchmark models. For each method, we conduct four experiments, including seven rounds of rolling forecast by setting the horizon to one, three, five and seven day(s).

The evaluation of intermittent demand forecasting is a key problem, since the actual value is always zero, which causes the denominator of many accuracy metrics to be zero. The metrics, including mean square error (MSE), mean absolute error (MAE), and Periods in Stock (PIS) are often used to evaluate the performance of predictive models (Ghobbar and Friend, 2003; Kourentzes, 2014; Syntetos and Boylan, 2005). In this study, three accuracy measures including MAE, root mean square error (RMSE), and symmetric mean absolute percentage error (sMAPE) are used to evaluate the performance of the forecasting methods.

The MAE for item y at time period t is given by:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \hat{y}_t\right) \tag{6}$$

The RMSE for item y at time period t is calculated as:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_t - \hat{y}_t\right)^2} \tag{7}$$

The sMAPE was first proposed by Armstrong (1985) as a modified MAPE, and its formulation is:

$$sMAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{|y_t - \hat{y}_t|}{(|y_t| - |\hat{y}_t|)/2} \times 100\% \tag{8}$$

Note that sMAPE is invalid when the real value and the predicted value are both equal to zero. Thus, we set sMAPE to zero when this happens, which means perfect prediction.

4.3. Empirical results

4.3.1. Basic research results (RQ1)

Based on the two datasets, we use the SES, CR, SBA methods and MCM to conduct four forecasting experiments and calculate the three-accuracy metrics: MAE, RMSE, and sMAPE. We test the model performance for seven rounds with different horizons (Huang et al., 2011; Freeman et al., 2019). Rolling forecast testing have more stringent requirements on the length of historical data. For example, data for the last 13 days needs to be intercepted if we want to conduct the experiment with seven rounds and seven horizons. Thus, we only use items with the length of historical data being greater than 30 days. Finally, there are 927 items in JD's data and 2994 items in Alibaba's data that are used to conduct rolling forecast testing. The forecast accuracy comparisons for different methods on the testing datasets, with different forecasting horizons are shown in Table 2 and Table 3. In general, the forecast results indicate that the proposed MCM performs better than the other three benchmark models. For instance, for JD's items when the horizon equals one, the sMAPE of SES was 1.823%, that of CR was 1.816%, that of SBA was 1.8%, while for MCM it was 0.408%. We also calculate the improvement of the proposed method over the three benchmark models. Fig. 5 shows the performance of the various methods on four different forecasting horizons. As is evident, MCM performs better. But there is no obvious rule of prediction accuracy between different horizons.

Statistical test is essential to determine whether the better performance is statistically significant. We run KSPA test (the Kolmogorov-Smirnov Predictive Accuracy) to provide statistical evidence. KSPA test enables distinguishing the forecast distribution of the two models

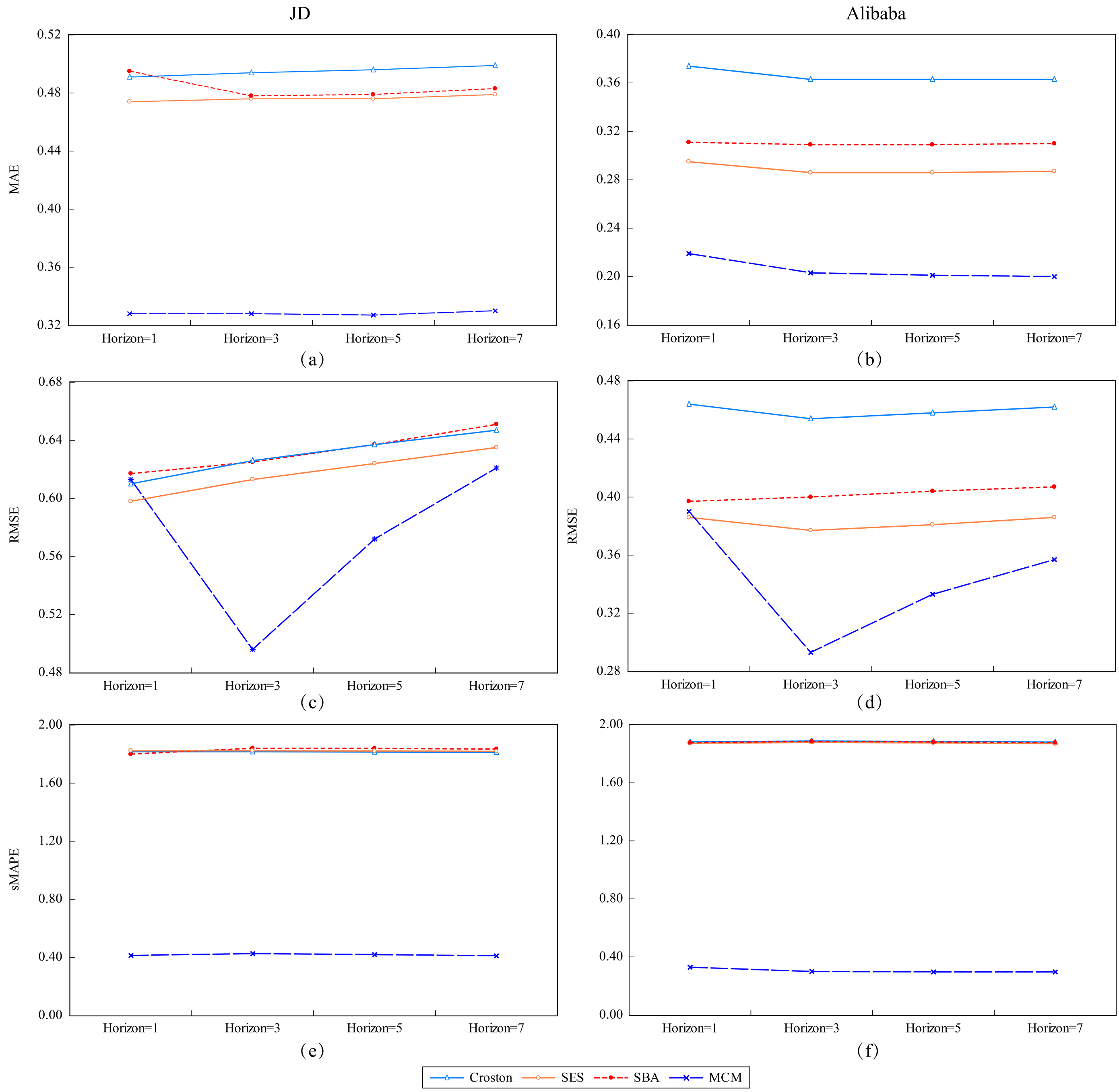

Fig. 5. The results of different accuracy metrics – JD′ dataset and Alibaba′ dataset.

and determine whether lower error also represents lower stochastic error (Hassani and Silva, 2015). The resulting output from KSPA test comparing the prediction errors of MCM and SES, MCM and Croston, MCM and SBA with different horizons are shown in Table 4 and Table 5. The results show that there exists a statistically significant difference between the distribution of prediction errors from MCM and other benchmark models at a 99% confidence level.

4.3.2. The results of different data characteristics (RQ2)

To explore the MCM further, we divide the two datasets into four patterns based on the median of ADI and CV^2. Pattern1 represents the group that ADI $<$ $Median_{ADI}$ and $CV^2 < Median^2_{CV}$. Pattern2 represents the group that ADI $>$ $Median_{ADI}$ and $CV^2 < Median^2_{CV}$. Pattern3 represents the group that ADI $>$ $Median_{ADI}$ and $CV^2 > Median^2_{CV}$. Pattern4 represents the group that ADI $<$ $Median_{ADI}$ and $CV^2 > Median^2_{CV}$. We calculate the average forecast error of RMSE, and the results are shown in Table 6. Both datasets show that Pattern2 and Pattern3 have better prediction accuracy by using the proposed model, which indicates the fact that commodities with more stable demand pattern are relatively easy to predict. The results in Table 6 also prove that considering inventory information could improve forecast accuracy, especially for products with long demand intervals. In retail inventory management, companies usually update inventory quickly. The proposed model can capture the state of zero inventory when the demand interval is long. Thus, when there exists long interval and the inventory is emptied, MCM can determine that the sales volume is zero by obtaining inventory information instead of other time series models to predict sporadic demand through historical data. When stockout occurs, it is difficult to say

Table 4
The results of KSPA test – JD' dataset.

	Two-Sided (*p*-Value)	Greater (*p*-Value)
Horizon = 1 (Obs. = 927*1*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*
Horizon = 3 (Obs. = 927*3*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*
Horizon = 5 (Obs. = 927*5*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*
Horizon = 7 (Obs. = 927*7*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*

Note: * indicates results are statistically significant based on a p-value of 0.001.

Table 5
The results of KSPA test – Alibaba' dataset.

	Two-Sided (*p*-Value)	Greater (*p*-Value)
Horizon = 1 (Obs. = 2994*1*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*
Horizon = 3 (Obs. = 2994*3*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*
Horizon = 5 (Obs. = 2994*5*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*
Horizon = 7 (Obs. = 2994*7*7)		
MCM vs SES	<0.001*	<0.001*
MCM vs Croston	<0.001*	<0.001*
MCM vs SBA	<0.001*	<0.001*

Note: * indicates results are statistically significant based on a p-value of 0.001.

Table 6
The average RMSE forecast error of four demand patterns.

	SES	Croston	SBA	MCM	Obs
JD					
All	0.617 (0.736)	0.630 (0.736)	0.632 (0.666)	0.576 (0.870)	927
Pattern1	0.645 (0.401)	0.647 (0.400)	0.649 (0.407)	0.633 (0.400)	154
Pattern2	0.338 (0.240)	0.357 (0.230)	0.352 (0.208)	0.259 (0.264)	310
Pattern3	0.630 (0.486)	0.642 (0.480)	0.635 (0.502)	0.520 (0.488)	153
Pattern4	0.876 (1.107)	0.888 (1.112)	0.903 (0.962)	0.892 (1.340)	310
Alibaba					
All	0.383 (0.350)	0.459 (1.272)	0.402 (0.401)	0.343 (0.412)	2994
Pattern1	0.404 (0.185)	0.407 (0.186)	0.404 (0.179)	0.358 (0.260)	316
Pattern2	0.188 (0.169)	0.202 (0.161)	0.196 (0.159)	0.135 (0.169)	1182
Pattern3	0.258 (0.301)	0.355 (0.814)	0.293 (0.293)	0.185 (0.288)	314
Pattern4	0.605 (0.396)	0.758 (1.931)	0.637 (0.500)	0.589 (0.501)	1182

Notes: Pattern1 represents the group that ADI $<$ Median$_{ADI}$ and CV^2 $<$ Median$^2_{CV}$. Pattern2 represents the group that ADI $>$ Median$_{ADI}$ and CV^2 $<$ Median$^2_{CV}$. Pattern3 represents the group that ADI $>$ Median$_{ADI}$ and CV^2 $>$ Median$^2_{CV}$. Pattern4 represents the group that ADI $<$ Median$_{ADI}$ and CV^2 $>$ Median$^2_{CV}$.

Table 7
The average forecast error of four demand patterns.

	MAE_AVG	RMSE_AVG	sMAPE_AVG	Obs
JD				
All	0.328(0.587)	0.576(0.870)	0.406(0.322)	927
Pattern1	0.364(0.241)	0.633(0.400)	0.521(0.283)	154
Pattern2	0.128(0.141)	0.259(0.264)	0.205(0.205)	310
Pattern3	0.259(0.250)	0.520(0.488)	0.276(0.229)	153
Pattern4	0.545(0.929)	0.892(1.340)	0.613(0.324)	310
Alibaba				
All	0.206(0.297)	0.343(0.412)	0.306(0.356)	2994
Pattern1	0.206(0.182)	0.358(0.260)	0.376(0.316)	316
Pattern2	0.065(0.094)	0.135(0.169)	0.126(0.168)	1182
Pattern3	0.093(0.170)	0.185(0.288)	0.136(0.190)	314
Pattern4	0.377(0.383)	0.589(0.501)	0.511(0.417)	1182

Table 8
Heterogeneity in four demand patterns.

	MAE	RMSE	sMAPE
JD			
Pattern1	−0.181*(0.055)	−0.259*(0.082)	−0.092*(0.026)
Pattern2	−0.417*(0.045)	−0.633*(0.067)	−0.408*(0.021)
Pattern3	−0.286*(0.055)	−0.372*(0.082)	−0.337*(0.026)
Observations	927	927	927
R-squared	0.088	0.090	0.316
Alibaba			
Pattern1	−0.171*(0.016)	−0.231*(0.022)	−0.135*(0.019)
Pattern2	−0.311*(0.011)	−0.454*(0.015)	−0.385*(0.013)
Pattern3	−0.284*(0.017)	−0.404*(0.023)	−0.375*(0.019)
Observations	2994	2994	2994
R-squared	0.233	0.257	0.259

Note: * indicates results are statistically significant based on a p-value of 0.001.

if there is a potential demand loss. MCM could help estimate whether the sales of the day are real demand. For example, we need to know whether it is caused by no consumer demand or inventory shortage when sales records are zero. However, the other three benchmark models do not consider the reasons behind.

4.3.3. The intrinsic of MCM (RQ3)

We calculate the average forecast error of the three evaluation indicators. The results in Table 7 show that all indicators are performed better in Pattern2. Referring to Gu and Zhu (2021), we create dummy variables to place demand patterns into four categories: *Pattern1* = *Pattern2* = *Pattern3* = 1, and *Pattern4* = 0. It can be inferred from Table 7 that if the forecast error gap is statistically significant, the regression result of the evaluation indicators of other groups should be significantly decreased compared to the baseline group. Table 8 reports the results, which proves that MCM performs best in the group of pattern2, followed by the group of pattern3, the group of pattern1, and finally the baseline group. The items in baseline group are relatively unpredictable with characteristic of short intervals and high volatility. The items in the group of pattern2 have long intervals and low volatility, which are easier to capture the laws of inventory and sales changes. This result is in line with general perception that the more stable demand pattern has the more accurate forecasting result.

5. Discussion and conclusion

Our research complements studies in the retail forecasting area by analyzing intermittent demand. We propose a combined forecasting method– the MCM, which takes inventory information into account. The performance of MCM is evaluated and compared with three benchmark models by calculating three accuracy metrics using the datasets from JD and Alibaba. The empirical results show that the MCM delivers better/ more accuracy at various forecast horizons. It is worth mentioning that the empirical analysis is conducted on two big datasets, which eliminates the influence of other uncertain factors to an extent and thus, lends more credibility to our results. We also conduct statistical test which explain the robustness of the results further.

The study indicates that considering inventory information can promote more accurate forecast results. Inventory management plays a vital role in the supply chain, and maintaining appropriate inventory levels can simultaneously reduce operating costs and meet customers' demand. Due to supply chain issues or inaccurate replenishment planning, shortages often occur in real life. However, if the inventory levels are sufficient, then the sales can represent/satisfy actual demand. Conversely, when an item is out of stock, sales do not reflect/mirror real demand. Thus, it is reasonable and necessary to consider inventory information when forecasting demand. The better performance of MCM can illustrate this point.

Our research mainly has the following management implications. First, since MCM comprises a combination of simple methods, it has more practical application value compared with many other complex intelligent algorithms. Therefore, we recommend the MCM as an alternative method for forecasting intermittent demand because MCM is not computationally expensive and typically more accurate than the classical forecasting methods including SES, CR, and SBA. Secondly, MCM could improve the prediction accuracy of intermittent demand products, thereby helping to reduce inventory management costs and improve the economic benefits of enterprises. Thirdly, the types of SKUs are more varied and customers' requirements are gaining refinement; this can make intermittent demand more common in today's retail environment. This study tells us that it is necessary for companies to consider inventory information when making sales forecasts.

Although our study provides some managerial insights on forecasting intermittent demand, additional research is required. We have not investigated the essential causes of intermittent demand for retail goods, and have not ascertained whether external factors such as discounts could influence demand and inventory levels. Further research can examine the categories of retail products that display the characteristics of intermittent demand and study their influencing factors. Another interesting problem is to study how to implement effective inventory management plans according to the different kinds of demand.

Declaration of competing interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by Beijing Natural Science Foundation (9212020), the National Natural Science Foundation of China (71932002), and the Fundamental Research Funds for the Central Universities.

References

Altay, N., Litteral, L.A., Rudisill, F., 2012. Effects of correlation on intermittent demand forecasting and stock control. Int. J. Prod. Econ. 135 (1), 275–283. https://doi.org/10.1016/j.ijpe.2011.08.002.

Armstrong, J.S., 1985. Measures of Accuracy in Long-Range Forecasting: from Crystal Ball to Computer. Wiley.

Babai, M.Z., Ali, M.M., Nikolopoulos, K., 2012. Impact of temporal aggregation on stock control performance of intermittent demand estimators: empirical analysis. Omega 40 (6), 713–721. https://doi.org/10.1016/j.omega.2011.09.004.

Babai, M.Z., Dallery, Y., Boubaker, S., Kalai, R., 2019. A new method to forecast intermittent demand in the presence of inventory obsolescence. Int. J. Prod. Econ. 209, 30–41. https://doi.org/10.1016/j.ijpe.2018.01.026.

Babai, M.Z., Syntetos, A.A., Teunter, R.H., 2014. Intermittent demand forecasting: an empirical study on accuracy and the risk of obsolescence. Int. J. Prod. Econ. 157, 212–219. https://doi.org/10.1016/j.ijpe.2014.08.019.

Croston, J.D., 1972. Forecasting and stock control for intermittent demands. Oper. Res. Q. 23 (3), 289–303. https://doi.org/10.1057/jors.1972.50.

Fildes, R., Ma, S., Kolassa, S., 2019. Retail forecasting: research and practice. Int. J. Forecast. https://doi.org/10.1016/j.ijforecast.2019.06.004.

Freeman, N.K., Narayanan, A., Burns, G.P., 2019. A decision support system for empty hopper car management. Interfaces 49 (1), 173–188. https://doi.org/10.1287/inte.2019.0987.

Fu, W., Chien, C., 2019. UNISON data-driven intermittent demand forecast framework to empower supply chain resilience and an empirical study in electronics distribution. Comput. Ind. Eng. 135, 940–949. https://doi.org/10.1016/j.cie.2019.07.002.

Ghobbar, A.A., Friend, C.H., 2003. Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model. Comput. Oper. Res. 30 (14), 2097–2114. https://doi.org/10.1016/s0305-0548(02)00125-9.

Gu, G., Zhu, F., 2021. Trust and disintermediation: evidence from an online freelance marketplace. Manag. Sci. 67 (2), 794–807. https://doi.org/10.1287/mnsc.2020.3583.

Hasni, M., Aguir, M.S., Babai, M.Z., Jemai, Z., 2019a. On the performance of adjusted bootstrapping methods for intermittent demand forecasting. Int. J. Prod. Econ. 216, 145–153. https://doi.org/10.1016/j.ijpe.2019.04.005.

Hasni, M., Babai, M.Z., Aguir, M.S., Jemai, Z., 2019b. An investigation on bootstrapping forecasting methods for intermittent demands. Int. J. Prod. Econ. 209, 20–29. https://doi.org/10.1016/j.ijpe.2018.03.001.

Hassani, H., Silva, E.S., 2015. A Kolmogorov-Smirnov based test for comparing the predictive accuracy of two sets of forecasts. Econometrics 3 (3), 590–609. https://doi.org/10.3390/econometrics3030590.

Hua, Z., Zhang, B., 2006. A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts. Appl. Math. Comput. 181 (2), 1035–1048. https://doi.org/10.1016/j.amc.2006.01.064.

Huang, L.T., Hsieh, I.C., Farn, C.K., 2011. On ordering adjustment policy under rolling forecast in supply chain planning. Comput. Ind. Eng. 60 (3), 397–410. https://doi.org/10.1016/j.cie.2010.07.018.

Kourentzes, N., 2013. Intermittent demand forecasts with neural networks. Int. J. Prod. Econ. 143 (1), 198–206. https://doi.org/10.1016/j.ijpe.2013.01.009.

Kourentzes, N., 2014. On intermittent demand model optimisation and selection. Int. J. Prod. Econ. 156, 180–190. https://doi.org/10.1016/j.ijpe.2014.06.007.

Kourentzes, N., Rostami-Tabar, B., Barrow, D.K., 2017. Demand forecasting by temporal aggregation: using optimal or multiple aggregation levels? J. Bus. Res. 78, 1–9. https://doi.org/10.1016/j.jbusres.2017.04.016.

Lackes, R., Siepermann, M., Vetter, G., 2020. What drives decision makers to follow or ignore forecasting tools - a game based analysis. J. Bus. Res. 106, 315–322. https://doi.org/10.1016/j.jbusres.2019.02.036.

Li, X., Zhao, X., Xu, W., Pu, W., 2020. Measuring ease of use of mobile applications in e-commerce retailing from the perspective of consumer online shopping behaviour patterns. J. Retailing Consum. Serv. 55 https://doi.org/10.1016/j.jretconser.2020.102093.

Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T., Gucci, S., 2017. Single-hidden layer neural networks for forecasting intermittent demand. Int. J. Prod. Econ. 183, 116–128. https://doi.org/10.1016/j.ijpe.2016.10.021.

Nikolopoulos, K., 2020. We need to talk about intermittent demand forecasting. Eur. J. Oper. Res. https://doi.org/10.1016/j.ejor.2019.12.046.

Pennings, C.L.P., van Dalen, J., van der Laan, E.A., 2017. Exploiting elapsed time for managing intermittent demand for spare parts. Eur. J. Oper. Res. 258 (3), 958–969. https://doi.org/10.1016/j.ejor.2016.09.017.

Petropoulos, F., Kourentzes, N., Nikolopoulos, K., 2016. Another look at estimators for intermittent demand. Int. J. Prod. Econ. 181, 154–161. https://doi.org/10.1016/j.ijpe.2016.04.017.

Prestwich, S.D., Tarim, S.A., Rossi, R., Hnich, B., 2014. Forecasting intermittent demand by hyperbolic-exponential smoothing. Int. J. Forecast. 30 (4), 928–933. https://doi.org/10.1016/j.ijforecast.2014.01.006.

Syntetos, A.A., Babai, M.Z., Gardner, E.S., 2015. Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping. J. Bus. Res. 68 (8), 1746–1752. https://doi.org/10.1016/j.jbusres.2015.03.034.

Syntetos, A.A., Boylan, J.E., 2001. On the bias of intermittent demand estimates. Int. J. Prod. Econ. 71 (1), 457–466. https://doi.org/10.1016/S0925-5273(00)00143-2.

Syntetos, A.A., Boylan, J.E., 2005. The accuracy of intermittent demand estimates. Int. J. Forecast. 21 (2), 303–314. https://doi.org/10.1016/j.ijforecast.2004.10.001.

Syntetos, A.A., Boylan, J.E., Croston, J.D., 2005. On the categorization of demand patterns. J. Oper. Res. Soc. 56 (5), 495–503. https://doi.org/10.1057/palgrave.jors.2601841.

Syntetos, A.A., Nikolopoulos, K., Boylan, J.E., Fildes, R., Goodwin, P., 2009. The effects of integrating management judgement into intermittent demand forecasts. Int. J. Prod. Econ. 118 (1), 72–81. https://doi.org/10.1016/j.ijpe.2008.08.011.

Teunter, R.H., Duncan, L., 2009. Forecasting intermittent demand: a comparative study. J. Oper. Res. Soc. 60 (3), 321–329. https://doi.org/10.1057/palgrave.jors.2602569.

Teunter, R.H., Syntetos, A.A., Babai, M.Z., 2011. Intermittent demand: linking forecasting to inventory obsolescence. Eur. J. Oper. Res. 214 (3), 606–615. https://doi.org/10.1016/j.ejor.2011.05.018.

Veiga, C.P.d., Veiga, C.R. P.d., Puchalski, W., Coelho, L.d.S., Tortato, U., 2016. Demand forecasting based on natural computing approaches applied to the foodstuff retail segment. J. Retailing Consum. Serv. 31, 174–181. https://doi.org/10.1016/j.jretconser.2016.03.008.

Willemain, T.R., Smart, C.N., Schwarz, H.F., 2004. A new approach to forecasting intermittent demand for service parts inventories. Int. J. Forecast. 20 (3), 375–387. https://doi.org/10.1016/s0169-2070(03)00013-x.