Available online at www.sciencedirect.com
ScienceDirect
IFAC PapersOnLine 55-10 (2022) 2737–2742

On the inventory performance of demand forecasting methods of medical items in humanitarian operations

Bahman Rostami-Tabar*, Marwa Hasni, M. Zied Babai*****

** Cardiff Business School, Aberconway Building, Colum Road, Cathays, Cardiff, CF10 3EU, UK (e-mail: rostami-tabarb@cardiff.ac.uk).*

*** IHEC Carthage, Laboratoire d'Economie et de Finance Appliquées (LEFA), Tunisia*

(e-mail: gmarwa92@gmail.com)

**** Kedge Business School, 680 Cours de la Liberation, 33400, Talence, France (e-mail: mohamed-zied.babai@kedgebs.com)*

Abstract: The inventory management of medical items in humanitarian operations is a challenging task due to the intermittent nature of their demand and long replenishment lead-times. While effective response to emergency results in inventory build-up which saves human lives, excess inventories could be intentionally burnt or donated which is costly for humanitarian organizations. Henceforth, linking demand forecasting to the inventory control task is shown to be a significant scope to offer a higher performance. In this vein, it is key to accurately select adequate forecasting methods. This paper investigates the effectiveness of parametric and non-parametric demand forecasting methods that are commonly considered to deal with stock keeping units (SKUs) characterized with an intermittent demand in industrial contexts. To do so, we conduct an empirical study by means of data related to 1254 SKUs managed in three warehouses of a major humanitarian organization based in Geneva, Middle-east and Africa. The investigation is carried out to compare the inventory performance of three parametric and two bootstrapping methods when used with an order-up-to-level inventory control policy. The results demonstrate the high performance of the bootstrapping methods in achieving higher service levels. The investigation enables to gain insights on the forecasting method that should be selected under particular assumptions on the demand and the lead-time value.

Keywords: Humanitarian operations, forecasting; inventory performance; intermittent demand; bootstrapping methods; parametric forecasting methods

1. INTRODUCTION

Life-saving items are crucial for millions of people especially during natural or human-made disasters. These medicines are typically granted by humanitarian logistics systems, which seek to provide them at the right time and with the right amount (Davydenko and Fildes, 2013). For that purpose, accurate visibility on the future required items is an important driving factor. In addition, efficient estimates allow more realistic decisions underpinning the production capacity and the amount of stock to set out, as well as the scaling of the required material and humanitarian resources to tie-up. On the other hand, inadequate demand forecasting system threatens humanitarian logistics systems to perform poorly. Furthermore, it could yield significant funding loss and misleading directions for research and developments (Thomas, 2003). These arguments support strong recommendation to accurately select the forecasting method given a particular demand pattern and under specific assumptions on the inventory control model being used.

Historically, the research investigating forecasts for intermittent demand has initially explored parametric procedures where the observed demand is fitted to some hypothesized probability distributions. Several adaptive variants of the existing parametric methods have been suggested in the academic literature to better fit the intermittent nature of the demand (e.g. Croston, 1972; SBA, Syntetos and Boylan, 2005; TSB, Teunter et al., 2011; Empirical-EVT, Zhu et al., 2017; Bayesian approach-based, Babai et al., 2021). It was in the late 1980s that small-scale research and case studies began to emerge advocating the use of nonparametric methods for more accurate intermittent demand forecasts. Bookbinder and Lordahl (1989) argue that while the lead-time demand distribution (LTD) is skewed, the parameters of an inventory control policy are highly affected by the shape of its hypothesized distribution. In fact, a wrongly assumed distribution would yield misleading estimates of the demand and inventory control parameters. Ergo, it is preferable to implement a "distribution-free" forecasting approach. Amongst all existing non-parametric approaches, the bootstrapping approach has been considered worthy of

10.1016/j.ifacol.2022.10.132

academic attention. Mainly, it consists in drawing life-like samples that enable to estimate the entire demand distribution. Therefore, within the recent decades, there has been a growing body of literature interested in developing and evaluating the accuracy of Bootstrap-based forecasting methods in predicting demand within a jeopardy context.

Two bootstrapping methods have been recently developed to deal particularly with intermittent demands, namely: the first, labelled the WSS, is developed in Willemain et al., (2004) and samples demand data by using a Markov chain to switch between no demand and demand periods and the second, labelled the VZ, is proposed in Zhou and Viswanathan (2011) and samples separately demand intervals and demand sizes. Recently, there has been an increasing amount of investigations that analyse the performance of parametric and bootstrapping methods. However, none of the existing studies in the literature deals with medical items in humanitarian operations, despite their demand patterns that are characterised with a high demand intermittence. To bridge this gap in the literature, we carry out a broad empirical study using the demand data of about 1254 medical items from three warehouses of a major humanitarian organization based in Geneva, Middle-east and Africa. More specifically, the inventory-service performance for the aforementioned bootstrapping methods along with three well-known parametric methods: SBA, Croston and the Simple Exponential Smoothing (SES) methods are analyzed. The results demonstrate the inventory outperformance of the bootstrapping methods with different preferences between the two counterparts under particular model features (i.e. data and parameters). Here, it is worth mentioning that the use of machine learning models is also a good scope to investigate. However, we find it judicious to investigate the boostrapping priory since a corroborative evidence of these methods for intermittent demand is available. Meanwhile, these sampling methods could be viewed as the very first research attempts for developing only data driven tools for forecasting decisions. So using Machine Learning techniques is the direct ensuing strategy to take forward bootstrap based studies.

The remainder of this paper is organized as follows. In Section 2, we present a detailed description of the forecasting methods included in this study. The empirical data and the experimental design are described in Sections 3, whereas the empirical results are reported in Section 4. Finally, the conclusions and some avenues for further research are presented in Section 5.

2. FORECASTING METHODS

In this section, we provide a detailed description of the forecasting methods being considered in our study.

1.1 Non-parametric bootstrapping methods

In our study, we consider two bootstrapping forecasting methods, which are the most advocated in the area of spa*re parts demand forecasting (Hasni et al., 2019)*. The first, the Willemain, Smart and Schwarz (hereafter WSS), is the most known bootstrapping method dedicated to intermittent demand inventory forecasting. The following steps, reported from Willemain et al. (2004), outline the WSS algorithm: (1) evaluate the transition probabilities using a two state Markov model of the available demand history data in a chosen time buckets (e.g. days, weeks, months); (2) generate a sequence of zero/non-zero demand values X over the lead-time, according to the obtained Markov model (3) substitute every non-zero state marker with one value, sampled with replacement from the available positive demands; (4) let each non-zero value undergo the 'jittering" process , which consists in generating its standard normal random deviate realization according to equation (1):

$$1 + INT(X + Z\sqrt{X}) \qquad (1)$$

(5) replace the positive non-zero demand by its jittered counterpart unless the result is less than or equal to zero, in which case the jittered value is simply X; This allows obtaining a life-like sample and has shown a direct practical relevance (Willemain et al., 21004) (6) obtain a single estimated lead-time demand (LTD) by summing all appearing forecast values over the sampling horizon (the lead-time); (7) replicate all of these steps 1000 times to obtain the bootstrapped distribution of the lead-time demand.

The second bootstrapping method included in our study is the Zhou and Viswanathan method (hereafter VZ) developed and introduced by Zhou and Viswanathan (2011). Being a variation of the WSS method, VZ is claimed to be simpler and faster than WSS. Indeed, instead of using a two-state Markov chain for generating the demand distribution, it samples directly from the observed demand size and inter-demand arrival distributions. In the following, we provide a brief summary of the underlying steps of the VZ method: (1) Obtain histograms of the demand sizes and demand intervals, from the historical demand, in a chosen time buckets; (2) Generate a random demand interval according to the corresponding histogram then, update the time passing by; (3) If the latter is equal or less than the lead-time, use the demand size histogram to randomly generate a demand size value, then, refer back to *step 2*. Else, sum the generated demand sizes over the lead-time to get the estimate of the lead-time demand;(4) After repeating these two steps a high number times (1000 replications), generate the resulting distribution of the lead-time demand.

1.2 Parametric forecasting methods

1.2.1. Notations

The following notations will be used for the purpose of describing the operating mode of the parametric methods under concern.

D_t	the actual demand of an item at time period *t*
$\widehat{F}_t$	the estimated average demand per period for time $t + 1$ being made at time *t*
Z_t	the observed demand size at time *t*
$\widehat{Z}_t$	the estimated average demand size at time *t*
T_t	the observed demand interval at time *t*
$\widehat{T}_t$	the estimated demand interval at time period *t*
α	smoothing constant chosen in [0,1]
$MLTD_t$	Mean Lead-Time Demand at period *t*

1.2.2. Technical details

The Single Exponential Smoothing method (SES) is viewed as one of the earliest models that have been proposed for forecasting intermittent demand. This method calculates per period demand forecasts as outlined in equation (2):

$$\hat{F}_t = \hat{F}_{t-1} + \alpha\left(D_t - \hat{F}_{t-1}\right) \quad (2)$$

As may be noticed, all demands (zero and non-zero demands) are used when updating per period demand forecasts. Yet, a potential significant improvement to the SES method would be to take into account the interval demand profile as it has revealed an important driving factor to characterize different behaviours of intermittent demand. This hypothesis has been supported by Croston (1979) and gives rise to the Croston's method that uses simple exponential smoothing to separately estimate per period demand sizes and demand intervals according to equations (3) and (4), respectively:

$$\hat{Z}_t = \hat{Z}_{t-1} + \alpha\left(Z_t - \hat{Z}_{t-1}\right) \quad (3)$$
$$\hat{T}_t = \hat{T}_{t-1} + \alpha\left(T_t - \hat{T}_{t-1}\right) \quad (4)$$

The per period demand is then estimated by evaluating the quantity in equation (5)

$$\hat{F}_t = \hat{Z}_t / \hat{T}_t \quad (5)$$

Here, we shall note that the forecast updates are made only when a positive demand value D_t occurs. Afterwards, the mean of the demand distribution over $(L+1)$ periods is estimated as:

$$MLTD_t = (L+1) * \hat{F}_t \quad (6)$$

Later on in 2005, Syntetos and Boylon (2005) claimed the Croston's model to be biased and improved it by proposition a debiasing factor $(1-\alpha/2)$ to the demand estimate D_t according to the following expression (7) :

$$\hat{F}_t = (1-\alpha/2).\hat{Z}_t / \hat{T}_t \quad (7)$$

The so-called Syntetos and Boylan Approximation (hereafter SBA) method has been shown to be almost unbiased with good stock control performance (Syntetos and Boylan, 2005). Babai and Syntetos (2007) suggested estimating the variance of the mean demand per period using a smoothed Mean Squared Error (MSE) as follows:

$$MSE_t = \alpha\left(D_t - \hat{F}_{t-1}\right) + (1-\alpha)MSE_{t-1} \quad (8)$$

3. EMPIRICAL INVESTIGATION

3.1 Data

The historical demand information of 1254 items from a major international humanitarian organization based on Geneva, have been used for the purpose of our study. The dataset is related to the demand of medical items managed in warehouses located in four countries: Afghanistan, Nigeria, Mali and Switzerland. Demand is monthly recorded over 36 months which corresponds to the demand of medicines that can have a high degree of lumpiness. The lead-times vary from 1 month to 6 months. The detailed descriptive statistics on demand data are provided in Figures 1-2, which outline, respectively, the box and whisker plots of the periodic demand as well as the demand interval related to each data set. These are obtained by evaluating the logarithmic values of mean demand intervals, and the monthly demand for every single SKU/demand history. Subsequently, the minimum, 25th percentile, median, 75th percentile, and maximum across all SKUs are calculated. In order to undertake scale issues, we present the corresponding logarithmic values associated to these statistics, except those for the inter demand values of the Mali data set which have been divided by 10 to obtain adequate plot. As may be noticed from Figure 1, demand data from the Nigeria warehouse is found to be the most intermittent followed by that in Mali warehouse. It should be pointed out though, that following Figure 2, the demand intervals records within Mali warehouse are more dispersed leading to a significantly less observed demand than the remaining warehouses. Moreover, demand records of both the Switzerland and the Afghanistan warehouses are less intermittent The Switzerland data is more dispersed than the Mali's one. Moreover, we notice that Mali dataset encompasses SKUs with the highest variable intervals since 50% of them are comprised between within 1 and 31 months.

Overall, the simulated dataset encompasses demand records of high inter arrival values since minimum value is of about 6 months except for the Mali dataset so they could be viewed as slow movers. In addition, we found that monthly demands are highly variable since 25% of the histories show demand amounts higher than 182 items with about a 273-standard deviation value. Giving these different features, we believe our dataset to be enough rich to reflect the complexity of the real context under concern. This would be of interest to develop conclusions of a direct practical relevance for the forecasting of medical items in humanitarian operations.

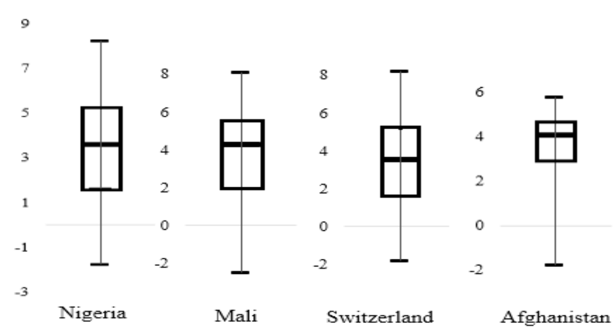

Figure 1: Box and Whisker plot of the monthly demand

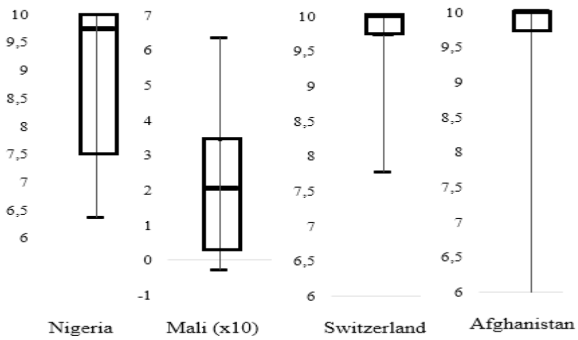

Figure 2: Box and Whisker plot of the demand interval

3.2. Experimental design

To measure the performance of the forecasting methods, we abide by a growing number of published studies (e.g. Babai et al., 2010; Syntetos et al., 2015; Hasni et al., 2019) which digress from traditional statistical measurements and evaluate forecasts by their direct effect on stock control performance. In this paper, the performance is evaluated by using a periodic order-up-to-level (T,S) inventory control policy where the review period T is equal to one month. The inventory control system is designed to meet a target fraction of replenishment cycles in which total demand can be delivered from stock. This fraction is called the cycle service level (CSL) (i.e. the probability of no stock-outs during a replenishment cycle). During the out-of-sample testing, the forecasting methods are used to estimate the lead-time demand and to compute the monthly order-up-to-levels. We vary the CSL targets between 85%, 90%, 95%, and 99%, i.e. in each case the order-up-to-level is computed by determining the value that corresponds to the CSL-quantile in the distribution. Table 1 summarizes these settings.

Table 1: Parameters settings

Study's parameters	**Value settings**
Inventory control Policy	(T, S) Policy
Lead-time (LT)	1,5
Smoothing Parameters	0.1 for sizes, intervals and MSE
Target CSLs	85%, 90%, 95%, and 99%
Performance measurements	Inventory efficiency: Inventory holding volumes vs. Achieved CSL

4. EMPIRICAL RESULTS

As previously pointed out, the competing forecasting methods are compared according to the inventory performance. The inventory performance is shown by means of efficiency curves that show the efficiency between the inventory holding volumes and the achieved CSL, i.e. we display the efficiency curves that show the inventory holding volumes versus the achieved CSL when varying the target CSL between the four values (85%, 90%, 95% and 99%). These efficiency curves can be interpreted as follows: for a certain inventory holding volume, the curve that is further from the x-axis (on the top of the other curves) implies more efficiency (by implying a higher achieved service level). In the following, we browse each of the four datasets included in our study, and report related results according to these performances along with associated interpretation. The efficiency curves are reported in Figures 3-1à for the four warehouses and the lead-times $L = 1$ and $L = 5$.

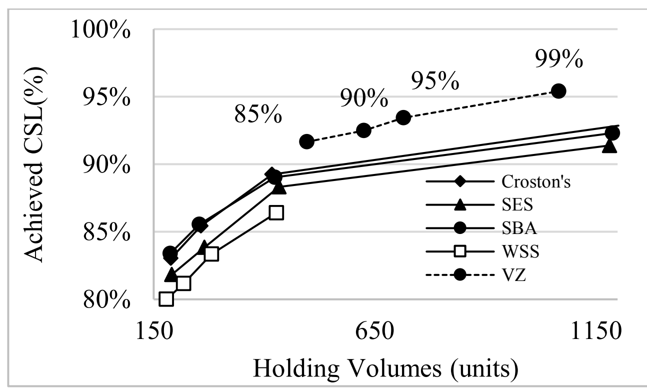

Figure 3: Afghanistan dataset (1-period lead-time)

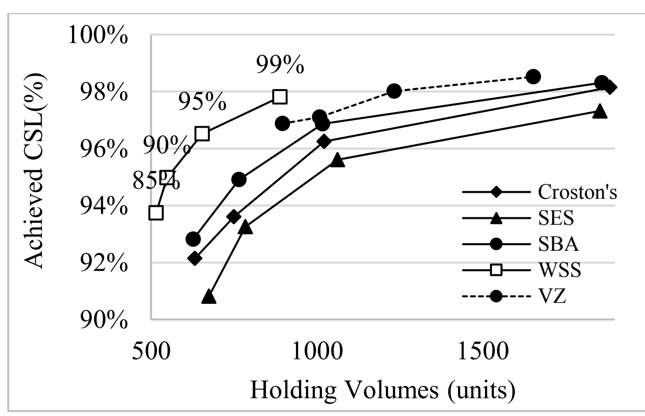

Figure 4: Afghanistan dataset (5-period lead-time)

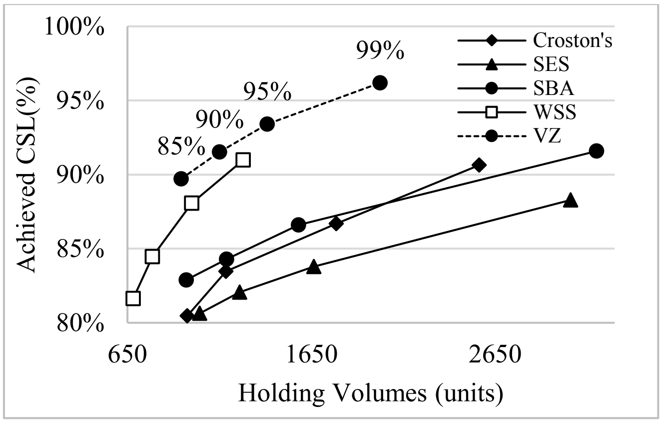

Figure 5: Mali dataset (1-period lead-time)

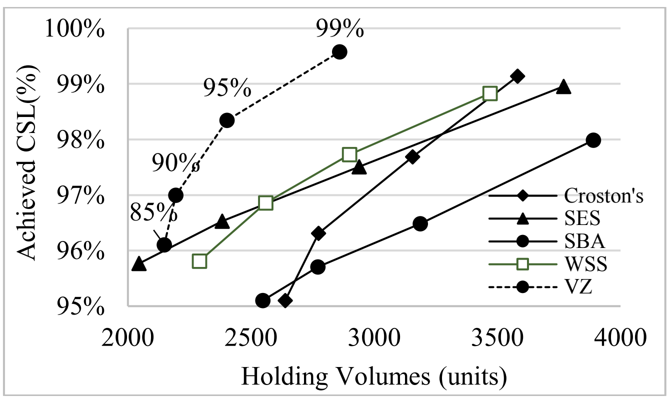

Figure 6: Mali dataset (5-period lead-time)

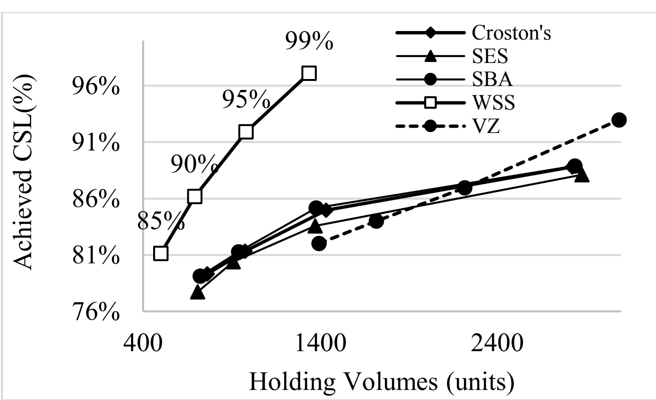

Figure 7: Nigeria dataset (1-period lead-time)

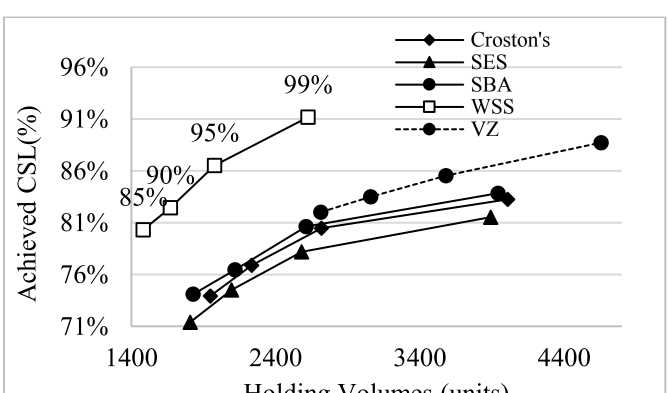

Figure 8: Nigeria dataset (5-period lead-time)

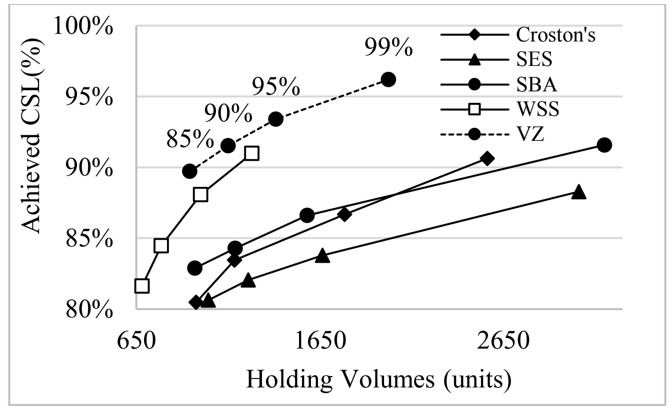

Figure 9: Switzerland dataset (1-period lead-time)

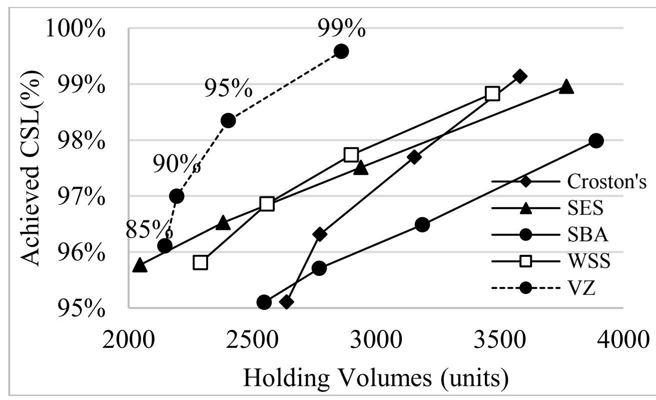

Figure 10: Switzerland dataset (5-period lead-time)

Recall from the data description in Section 3.1 that both historical data samples related to Mali and Nigeria warehouses show the highest intermittence of the demand per period with a more dispersed distribution for the latter one. Thereby, one might consider these demand profiles as intermittent to Lumpy according to the Syntetos and Boylan (2005) classification. On the question of the remaining datasets, we have noticed that demand intervals are rather stable but the values of demand per period have been observed to be relatively highly variable. Therefore, these samples could be described as erratic to intermittent according to the same classification grid. Focussing on the achieved CSL performances under a particular lead-time value, we notice that for L = 1, all forecasting methods fail to achieve the 85% target CSL, except the VZ bootstrapping method. Note that even the latter is not able to reach the 85% target CSL with the Nigerian demand data, which may be explained by the fact that this warehouse is characterised with a higher variability. Also, for the same target points, we notice that for L = 5 all the forecasting methods succeed in achieving the 85% target CSL though remaining quite difficult for the highest variable data (i.e. Nigeria dataset). Note also that the highest target CSL, i.e. CSL = 99%, is difficult to reach in most of the cases. Overall, it could be stated that all forecasting methods operate better under the highest lead-time values and target CSLs. For the bootstrapping, this might be justified by the fact that higher lead-times imply longer data samples which tones down total data variances, whereas for the parametric models this might be explained by the fact that they are upwardly biased. For humanitarian logistic systems, this might be favourable as operations are typically defined within a jeopardy context. Here, we shall note that these findings go in accordance with results that have been found in previous studies dealing with industrial spare parts (e.g. Hasni et al., 2019). Going into more details, we suggest visualizing the results emanating from the least to the most intermittent datasets. As such, we notice the VZ bootstrapping method is best performing. It achieves target CSLs in 75% of cases. In addition, the achieved values are higher than those offered by the remaining forecasting models for a fixed amount of inventory. This property is verified under different target CSL and for all ranges of lead-times. Here, it is worthy to mention that the competitiveness between VZ and the WSS variant gets clearer with the increase of the lead-time. This might be explained by the advantage of the jittering process in capturing demand noise that becomes more visible for long lead-times. As far as demand variability rises, the competitiveness between both bootstrapping variants becomes more visible. Indeed, the VZ kept best ranked when dealing with data from the Switzerland and the Mali warehouses. Nevertheless, undeniable competitiveness of the WSS is observed for the highest lead-times, which has frequently led mixed findings. Lastly, the WSS compared more favorable for the Nigerian data whereby demand is highly intermittent. These results confirm previous findings on the outperformance of the WSS under highly variable data and longer lead-times. However, a newly emerging finding is the fact that the bootstrapping methods have particularly shown significant preference to the parametric variants, most notably the SBA. Interestingly, this overcomes the commonly encountered mixed finding issue risen by similar comparative studies. Furthermore, it triggers further analysis to particular features underpinning the demand profile of medical items that might explain these findings and generalize them to other demand items.

6. CONCLUSIONS

In this paper, we investigate the inventory control performances of parametric and bootstrapping forecasting methods in the context of humanitarian operations. To the best of our knowledge, this is the first endeavour to enact using sophisticated forecasting methods when dealing with medical items' demands, as well as exploring inventory control indicators instead of the statistical counterparts. Using data records of about 1254 SKUs supplied from four warehouses of a major humanitarian organization, we come to develop preliminary insights. To start with, the bootstrapping methods are found to outperform the parametric variants. More specifically, the VZ is recommended for moderately intermittent to erratic data, whereas the WSS compares more favourable under highly erratic data and longer lead-times.

To take forward this study, it is interesting to assess the relative performances under broader range of parameter settings and demand variability in order to get further insights on the cut-off values that split the domains of effectiveness of each of the VZ and the WSS forecasting methods. This would be of interest to redirect managers toward the right forecasting method to be selected given particular features of an inventory control system.

REFERENCES

Babai, M. Z., Chen, H., Syntetos, A. A., and Lengu, D. (2021). A compound-Poisson Bayesian approach for spare parts inventory forecasting. *International Journal of Production Economics*, 232, 107954.

Babai, M. Z., Syntetos, A. A., and Teunter, R. (2010). On the empirical performance of (T, s, S) heuristics. *European Journal of Operational Research*, 202(2), 466-472.

Bookbinder, J. H., and Lordahl, A. E. (1989). Estimation of inventory re-order levels using the bootstrap statistical procedure. IIE transactions, 21(4), 302-312.

Boylan, J. E., Syntetos, A. A., and Karakostas, G. C. (2008). Classification for forecasting and stock control: a case study. *Journal of the operational research society*, 59(4), 473-481.

Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Journal of the Operational Research Society*, 23(3), 289-303.

Davydenko, A., and Fildes, R. (2013). Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. International Journal of Forecasting, 29(3), 510-522.

Hasni, M., Babai, M. Z., Aguir, M. S., and Jemai, Z. (2019). An investigation on bootstrapping forecasting methods for intermittent demands. *International Journal of Production Economics*, 209, 20-29.

Syntetos, A. A., Babai, M. Z., and Gardner Jr, E. S. (2015). Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping. *Journal of Business Research*, 68(8), 1746-1752.

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

Thomas, A. (2003). Humanitarian logistics: enabling disaster response. *Fritz Institut*e, 15.

Willemain, T. R., Smart, C. N., and Schwarz, H. F. (2004). A new approach to forecasting intermittent demand for service parts inventories. *International Journal of forecasting*, 20(3), 375-387.

Zhou, C., and Viswanathan, S. (2011). Comparison of a new bootstrapping method with parametric approaches for safety stock determination in service parts inventory systems. *International Journal of Production Economics*, 133(1), 481-485.

Zhu, S., Dekker, R., Van Jaarsveld, W., Renjie, R. W., and Koning, A. J. (2017). An improved method for forecasting spare parts demand using extreme value theory. *European Journal of Operational Research*, 261(1), 169-1.