Pharmaceutical Inventory Management: A Comparative Analysis of Forecasting Techniques and Dynamic Reordering Policies

Antonio Maria Coruzzolo*, Francesco Lolli*,,***, Elia Balugani* , Laura Vaccari*, Benedetta Tronto*, Chiara Forgione******

** Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Via Amendola 2, 42122, Reggio Emilia, Italy (antoniomaria.coruzzolo@unimore.it, laura.vaccari@unimore.it, elia.balugani@unimore.it,)*

*** Interdepartmental Centre En&Tech, Piazzale Europa, 1, 42124, Reggio Emilia, Italy(francesco.lolli@unimore.it,)*

**** Interdepartmental Centre InterMech MO.RE, University of Modena and Reggio Emilia, Piazzale Europa, 1, 42124, Reggio Emilia, Italy*

*****Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Via Giuseppe Campi 287, 41125 – Modena, Italy (chiara.forgione@unimore.it).*

Abstract: The shortage of products, or stockout, is identified as a critical business issue, leading to disruptions in product flow and subsequent economic damage. In particular in the healthcare context, stockouts can pose risks to patients due to the inability to administer essential medications. The study presents an in-depth analysis of a set of pharmaceutical products based on a five-year database containing information on demand, stock, and orders placed. In particular the aim is to assess the performance of various demand forecasting techniques on this product set and subsequently find the most cost-effective dynamic reordering policy's parameters. The efficacy of the forecasting techniques is selected based on minimizing the Root Mean Square Error (RMSE). Subsequently, a periodic dynamic review policy is applied to determine the number of orders and resulting backorders, evaluating the total management cost for the item. This approach allows for the evaluation of the effectiveness of ad hoc forecasting methods for each product compared to using a uniform approach. The results of the analysis provide a detailed overview of the forecasting techniques' performance related to dynamic reordering policy parameters and demonstrate the benefits earnable with respect to the company classical management.

Keywords: inventory management, demand forecasting, pharmaceutical products.

1. Introduction

Healthcare logistics, also known as the healthcare supply chain, encompasses all activities contributing to the delivery of service to the patient(Villa, Giusepi & Lega 2012). These activities include the traditional functions of procurement, storage, and shipment of medical products, such as demand forecasting and inventory level definition. The accuracy of demand forecasting significantly affects safety stock and inventory levels, inventory holding costs, and customer service levels (Bon & Leng 2009). Product shortages, or stockouts, are identified as critical business issues, leading to disruptions in product flow and consequent economic damages. In the pharmaceutical context, stock shortages can pose a risk to patients due to the inability to administer essential medications. Despite the complexity and execution of forecasting processes across various activities, the intended purpose remains the same: to obtain a reasonably accurate estimate of future demand for a product or service based on historical data and plan and organize business accordingly. Forecast accuracy still poses a major challenge in the pharmaceutical sector (Merkuryeva, Valberga & Smirnov 2019) to ensure high service level. As reported in the literature, several forecasting methods have been developed based on two well-known approaches: qualitative and quantitative. Qualitative methods, such as executive opinions, the Delphi technique, sales force surveys, and customer services, generate forecasts based on judgments or opinions. On the other hand, quantitative techniques include forecasts based on historical data, such as moving averages, the Naive method, exponential smoothing, and Holt's method. Additionally, there are mixed or combined models that allow integration of both approaches. In the pharmaceutical industry, time series models are most commonly used (52%), followed by causal models (24%), while models based on expert opinions represent 19%. The remaining 5% corresponds to mixed or combined models (Jain 2003). This study will mainly focus on quantitative methods to forecast the demand for inventory management. The demand forecasting process involves a series of steps that entail analyzing historical data, in this case, concerning the inventory of 18 pharmaceutical products. This analysis is followed by selecting the most suitable forecasting model, validating it to assess its effectiveness using new data. In particular, historical series must be cleaned of any outlier values before being analyzed. To evaluate the performance of forecasting models various performance metrics can be used as the root mean square error (RMSE) to verify the quality and accuracy of a forecasting method (Yang et al. 2021). In addition, it will be studied how to outline and implement the revision policy based on previously identified forecasts, the opportunity to implement safety stock will also be evaluated. This process will involve a thorough analysis of forecast results and evaluation metrics to determine the need and effectiveness of maintaining safety stock to manage any unforeseen variations in demand. Factors such as demand variability and procurement lead time will be considered to make an informed decision on adopting safety stock and the optimal level to maintain. This evaluation will enable the integration of the revision policy with inventory management strategies to ensure optimal demand management and reduce the risk of stockouts. The paper is structures as follows: Chapter 2 presents forecasting methods exploited, Chapter 3 presents the dynamic inventory management policy chosen and some key performance indicators, Chapter 4 presents the first phase of the case study in the pharmaceutical sector regarding the forecasts, while Chapter 5 highlight the management phase of the case study. Finally, the conclusions will be stated containing future research directions.

2. Forecasting methods

Here we present some of the most exploited time-series-based forecasting methods that have been exploited. Before introducing the various models, a general notation is introduced of terms used along the paper. However, some specific notation will be introduced later for each model for clarity.

Notation

- t: generic time period in a reference horizon T.
- y_t: real demand at period-t.
- $\hat{y}_t$: forecast made at period-t.
- $RMSE$: Root Mean Squared Error.
- α, β, γ: smoothing parameters.
- $RMSE$: Root Mean Squared Error.
- ROP: Re-Order-Point
- RLT: Replenishment Lead Time.
- CSL: Cycle Service Level.
- SS_t: Safety Stock at period-t.
- IP_t: Inventory Position at period-t.
- Q_t: Order Quantity at period-t.
- SES: Simple Exponential Smoothing.
- MA: Moving Average.

2.1 Naïve

The Naive Method involves the naive forecast, which is obtained by assigning to all forecasts a value equal to the last observed as shown in Equation 1. The naive forecast is optimal when the historical series follows a random walk and is therefore also called a "random walk forecast."

$$\hat{\boldsymbol{y}}_t = \boldsymbol{y}_t \qquad (1)$$

2.2 Moving Average

The Moving Average (MA) is a procedure that allows sleek series and thus eliminates oscillations such as seasonal and erratic ones. The moving average of order-T is defined as shown in Equation 2. Thus, the moving average can capture the trend of the signal over a time interval T.

$$\hat{\boldsymbol{y}}_t = \frac{1}{T}\sum_{i=0}^{T-1} y(t-i) \qquad (2)$$

2.3 Simple Exponential Smoothing

The Simple Exponential Smoothing (SES) is suitable for forecasting data without a clear trend or seasonality. Therefore, forecasts are calculated using weighted averages, where the weights decrease exponentially as observations come from the past, and smaller weights are associated with older observations. Mathematically, it is represented as visible in Equation 3.

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

Where $0 \le \alpha \le 1$ is the smoothing parameter. The forward forecast for period $T+1$ is a weighted average of all observations in the series $y_1, \dots, y_t$. The speed at which the weights decrease is controlled by the parameter α. For each α between 0 and 1, the weights associated with the observations decrease exponentially. If α is small and therefore close to zero, more weight is given to observations from further in the past. If α is large and therefore close to 1, more weight is given to more recent observations. If $\alpha = 1$, then $\hat{y}_t = y_t$ the forecast provided is equal to the naïve one.

2.4 Holt Method

Holt Method was created by Holt in 1957 (Holt 2004), who extended simple exponential smoothing to allow for forecasting data exhibiting a trend. This method involves a prediction equation (Equation 4) and two smoothing equations (one for the level, Equation 5, and one for the trend, Equation 6).

- Prediction equation: $\hat{y}_t = l_t + h * b_t$ (4)

- Level equation: $l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$ (5)

- Trend equation: $b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$ (6)

Here, l_t represents an estimate of the series level at time t, b_t represents an estimate of the trend of the series at time t, α is the smoothing parameter for the level, where $0 \le$

$\alpha \leq 1$, and β is the smoothing parameter for the trend, where $0 \leq \beta \leq 1$.

Similar to simple exponential smoothing, the level equation shows that l_t is a weighted average of the observation y_t and the one-step-ahead forecast for time t based on the training data, provided by $l_{t-1} + b_{t-1}$. The trend equation shows that b_t is a weighted average of the trend estimated at time t based on $l_t - l_{t-1}$ and b_{t-1}, the previous trend estimate. The forecasting function is no longer flat but trended: the h-step-ahead forecast is equal to the last estimated level plus h times the last estimated trend value, hence, the forecasts are a linear function of h.

2.5 Holt-Winters method

Holt-Winters Method extends the Holt method to capture seasonality (Chatfield 1978). The Holt-Winters method includes the prediction equation and three smoothing equations: one for the level l_t, one for the trend b_t, and one for the seasonal component s_t, with corresponding smoothing parameters α, β e γ. m is used to indicate the seasonality period, i.e., the number of seasons in a year. There are two variants of this method differing because of the seasonal component. The additive method is preferred when seasonal variations are approximately constant throughout the series, while the multiplicative method is preferred when seasonal variations change proportionally to the level of the series. With the additive method, the seasonal component is expressed in absolute terms on the scale of the observed series, and in the level equation, the series is de-seasonalized by subtracting the seasonal component. Within each cycle, the seasonal component will sum approximately to zero. With the multiplicative method, the seasonal component is expressed in relative terms (percentages), and the series is de-seasonalized by dividing by the seasonal component. Within each cycle, the seasonal component will sum approximately to m.

2.6 Croston's method

Croston's method is intended for intermittent demand (Croston 1972), it constructs two new series from the original historical series by selecting time instants that contain null values and those that contain values other than zero. Let q_i be the i-th non-zero quantity and a_i be the time instant between q_{i-1} and q_i. Croston's method separates the forecasts produced by simple exponential smoothing into the two new series a and q. Since the method is usually applied to historical series related to the demand for items, q is often referred to as "demand" and a as "inter-arrival time." Let $\hat{q}_{i+1|i}$ and $\hat{a}_{i+1|i}$ be the one-step forecasts of the $(i+1)-th$ demand and the corresponding inter-arrival time based on the data up to demand i. Then, according to Croston's method, we have the Equation 7 and Equation 8.

$$\hat{q}_{i+1|i} = (1-\alpha_q)\hat{q}_{i|i-1} + \alpha_q q_i \quad (7)$$

$$\hat{a}_{i+1|i} = (1-\alpha_a)\hat{a}_{i|i-1} + \alpha_a a_i \quad (8)$$

The smoothing parameters α_a and α_q take values between 0 and 1. Let j be the time instant relative to the last positive observation. Then, the h-step ahead forecast for the demand at time $(T+h)$ is given by the ratio of the Equation 9.

$$\hat{y}_{T+h|T} = \frac{\hat{q}_{j+1|j}}{\hat{a}_{j+1|j}} \quad (9)$$

There are no algebraic results that allow the calculation of prediction intervals with this method because it does not correspond to any statistical model. The two smoothing parameters α_a and α_q are estimated from the data, which differs from how Croston originally intended to use this method. In the original approach, in fact, it was considered that $\alpha_a = \alpha_q = 0.1$, and α_0 and q_0 were set equal to the first observation of each of the two series.

2.7 ARIMA

The ARIMA model is a combination of differencing and autoregressive and/or moving average models. ARIMA stands for Auto Regressive Integrated Moving Average. The model can be defined as in the Equation 10.

$$y'_t = c + \phi y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (10)$$

Where y'_t is the d-differenced series, ε are the noise terms, θ are the coefficient of the moving average components, ϕ are the coefficient of the autoregressive components. The model can be denoted as $ARIMA(p,d,q)$ where:

- p is the order of the autoregressive component

- d is the number of differences

- q is the order of the moving average component

The same stationarity and invertibility conditions that apply to purely autoregressive and purely moving average models also apply to an ARIMA model.

2.8 Prediction error

A prediction error is the difference between an observed value and its forecast. Here, the term "error" should not be understood in the sense of a mistake, but rather as the unpredictable part of an observation. The root mean square error (RMSE) is the standard deviation of the residual values. The residual values are a measure of the distance of the data points from the regression line, representing the gap between the predicted and observed points. The formula for RMSE is represented in the Equation 11.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{y}_i - y_i)^2} \quad (11)$$

where $\hat{y}_i$ are the predicted values, y_i are the observed values, and T is the reference period. In this study, the RMSE metric will be exploited since is among the most commonly used metric in the literature (Spiliotis et al. 2021).

3. Inventory management policy

In standard inventory theory, two main types of multiperiod reorder policies exists: the continuous review policy (ROP, Q) and the period policy (T, S). In the (ROP, Q) policy, a fixed order of quantity Q is placed whenever the inventory position reaches or falls below the reorder point (ROP). In contrast, in the (T, S) reorder policy, a variable quantity order Q is placed at intervals of time T, corresponding to the forecasting interval, to raise the inventory level S (Gamberini et al. 2014). This paper refers to dynamic sizing policy of ordered quantities according to the (T, S) approach since the case study involves only a periodic review management of the stocks. It is essential to consider the Replacement Lead Time (RLT) whenever optimizing a reorder policy, which corresponds to the time it takes to replenish goods from the moment an order is placed. Initially, zero safety stocks are assumed, where safety stocks are calculated to mitigate the risk of stockouts. An order is placed at time-t if the inventory position IP_t falls below the order up to level S_t, which is the inventory level at which stocks need to be replenished to meet demand during lead times. Specifically, the order placed will be the difference between S_t and IP_t. Since initially safety stocks are not considered, the order up to level is the demand forecasted for the review period. Therefore, S_t is given by:

$$S_t = y_t * (T + RLT) \quad (12)$$

Where y_t is the actual demand in period t, T and RLT are the review period and Replenishment Lead Time respectively. The inventory position at period t, denoted as IP_t is given by Equation 13.

$$IP_t = Q_{t-RLT} + IP_{t-1} - y_t \quad (13)$$

Where Q_{t-RLT} is the quantity ordered which is expected to arrive at period-t, IP_{t-1} is the inventory position from the previous period. The order up to level must be set to ensure a Cycle Service Level target (CSL_{target},) within $T + RLT$, considering that the ordered quantity is not available at the time of ordering but only after the RLT. Consequently, safety stocks, SS, must be set to avoid stockouts during the $T + RLT$ period. It is hypothesized, therefore, that the demand during the $T + RLT$ period, which corresponds to the period when stocks are replenished, follows a normal probability distribution with mean μ_{T+RLT_t}, standard deviation σ_{T+RLT_t} and probability density function $g(y)$ as shown in Equation 14.

$$g(y_{T+RLT}) \sim N(\mu_{T+RLT_t}, \sigma_{T+RLT_t}) \quad (14)$$

At this point, it is possible to find the order up to level as the inverse of the cumulative probability distribution of demand during $T + RLT$,y_{T+RLT}, calculated at the CSL_{target}, as shown in Equation 15.

$$S_t = G^{-1}{}_{T+RLT}(CSL_{target}) \quad (15)$$

Therefore, the order up to level will be equal to the average demand during $T + RLT$, increased by the safety stocks, as shown in Equation 16.

$$S_t = \mu_{T+RLT_t} + SS_t \quad (16)$$

Safety stocks are calculated as in Equation 17.

$$SS_t = k * \sigma_{T+RLT_t} \quad (17)$$

Where k is defined as the inverse of the standardized normal calculated at $CSL_{targhet}$. Hence, S_t is calculated in Equation 18.

$$S_t = \mu_{T+RLT_t} + k * \sigma_{T+RLT_t} = \mu_{T+RLT_t} + \varphi^{-1}(CSL_{target}) * \sigma_{T+RLT_t} \quad (18)$$

It is important to emphasize that k can be calculated in this way given the assumption of the demand being normally distributed. Now it is possible to define the two forecasters $\hat{\mu}_{T+RLT_t}$ and $\hat{\sigma}_{T+RLT_t}$, estimated as in Equation 19 and Equation 20.

$$\hat{\mu}_{T+RLT_t} = \hat{y}_t * (T + RLT) \quad (19)$$

$$\hat{\sigma}_{T+RLT_t} = \sqrt{\widehat{MSE}_t * (RLT + T)} \quad (20)$$

In the second forecaster, MSE_t refers to the Mean Squared Error, which provides a measure of the estimator's accuracy, as it calculates the average of the squares of the differences between the possible values of the estimator and the parameter to be estimated, which in this case correspond respectively to the actual and predicted demand, as calculated in Equation 21. Where γ corresponds to the damping coefficient, which is set to 0.1 by default. Therefore, the order up to level at period t is given by the Equation 22.

$$\widehat{MSE}_t = \gamma * (\hat{y}_t - y_t)^2 + (1 - \gamma) * \widehat{MSE}_{t-1} \quad (21)$$

$$S_t = \hat{\mu}_{T+RLT_t} + \varphi^{-1}(CSL_{target}) * \hat{\sigma}_{T+RLT_t} \quad (22)$$

3.1 Key performance indicators

In the context of inventory management policy evaluation, it is essential to consider a diverse range of key performance indicators, such as Fill Rate (FR) and Cycle Service Level (CSL). The FR represents an important metric for assessing the responsiveness of the inventory management system to customer demands. A high FR indicates that most requests are immediately fulfilled from available inventory, ensuring efficient and timely service. Conversely, a low FR could signal product availability issues and may negatively impact the overall customer experience. The FR is calculated as shown in Equation 23. The CSL provides a measure of the probability that customer demand will be fulfilled within an order cycle, as it is calculated in Equation 24. A high CSL is indicative of an effective inventory management policy that ensures reliable delivery times and a low incidence of backorders. However, a low CSL could indicate potential delays in product delivery and increased uncertainty for the customer. The addition of safety stock, therefore, helps to reduce the risk of backorders and increases product availability for the customer, thereby improving their overall experience.

$$Fill\ Rate = 1 - \frac{Backordered\ demand}{Total\ demand} \quad (23)$$

$$CSL = 1 - \frac{Number\ of\ orders\ with\ backorders}{Number\ of\ orders} \quad (24)$$

4. Case study

A set of 18 products has been selected from the Centralized Logistics Unit (CLU), focusing only on those managed with stocks. For each of these products, data on total daily requests from the departments were extracted via the CLU management system in the period from 01/01/2018 to 08/10/2023. For each code, daily demand, the quantity of product in stock, the order placed, and any urgent orders are recorded. The unit of measurement for each code is the posological unit, i.e., the standard quantity of a pharmaceutical product prescribed for a single dose, whether expressed in terms of tablets, capsules, sachets, injectable vials, kits, or tools administered to the customer. The lead time was determined by averaging the lead times during the first year and the mean value has been used as deterministic since a very low dispersion was identified. Subsequently, using the R programming language, daily demand was transformed and aggregated on a weekly time scale into 301 weeks. The aggregation has been done since the reference minimum review period is two weeks. The first phase aimed to identify the forecasting model that yields the best predictions for each selected code. Each historical series is divided into a training set and a test set. The training set consists of 249 weeks, which is 83% of the data, while the test set corresponds to the last 52 weeks of the historical series, representing the remaining 17% of the data. This division serves to evaluate the effectiveness of the implemented models in making accurate predictions on future data and follows the classical 80-20 rule to benchmark different forecasting methods (Mukhopadhyay, Solis & Gutierrez 2012). The training set is used to train the prediction models, allowing them to learn patterns and relationships in the historical data. Once trained, the models are tested using the test set, which represents temporal data unseen by the models during training. The methods seen previously are exploited for the forecasting: Naïve, Moving Average (MA), Simple Exponential Smoothing (SES), Holt's Trend Method, and Holt-Winters Seasonal Method, Croston, ARIMA. During the training the parameters for the models have been chosen to minimize errors. In particular, smoothing parameters have been optimized with a grid search in the interval [0.05-0.3] with a step of 0.01 to minimize MSE, the best order for the moving average have been searched in the interval [2-14] weeks to minimize MSE while the parameters of ARIMA have been chosen in order to minimize the Akaike information criterion (AIC) using the 'autoarima' package in R (Hyndman & Khandakar 2008). Once the parameters of the models are chosen during the training, they are used to generate demand forecasts for each of the 52 weeks of the test set, through a one-step forecast process, which involves predicting only the next value in the time series at each step, using the observed data up to that point. To compare the accuracy of the predictions $RMSE$ is calculated. Table 1 provides information for evaluating the performance of the considered forecasting models for each product it contains: the name of the best forecaster in terms of $RMSE$, its value, and best smoothing parameters for SES and best MA's order since these two emerged as best forecasters. However, for only 4 products ARIMA provides the best forecasts instead of SES and MA. For this reason, Table 2 reports the comparison of best forecasting models in terms of $RMSE$ where minimum $RMSE$ are highlighted in bold. As visible from Table 2 models achieve very similar results. In particular, the last column of Table 2 reports the percentage difference between ARIMA's performance and the performance of the best method between SES and MA. The mean percentage variation of ARIMA's performance and SES and MA performances is less than 3%, consequently, the Arima method has been substituted with MA and SES. This choice is further supported by the fact that implementing and interpreting an ARIMA model is statistically more difficult than managing SES and MA.

Table 1: Performance of forecasting models

Code	Best RMSE %	RMSE %	SES α *	MA order *
1	ARIMA	0.867	0.059	6
2	MA	0.714	0.51	7
3	SES	0.740	0.048	9
4	ARIMA	0.576	0.218	5
5	MA	0.550	0.268	3
6	SES	0.600	0.192	3
7	MA	0.540	0.297	7
8	SES	0.694	0.233	12
9	ARIMA	0.464	0.236	5
10	SES	0.336	0.270	6
11	SES	0.479	0.386	3
12	MA	0.417	0.267	7
13	SES	0.502	0.134	10
14	SES	0.280	0.159	7
15	SES	0.356	0.630	3
16	ARIMA	0.414	0.349	5
17	MA	0.179	0.369	5
18	SES	0.218	0.525	4

Table 2: Comparison of best forecasting models

Code	RMSE % SES	RMSE % MA	RMSE % ARIMA	ARIMA Δ %

1	0.8806	0.9264	**0.8670**	0,0155
2	0.7493	**0.7136**	0.7334	0
3	**0.7402**	0.8203	0.7615	0
4	0.5762	0.6000	**0.5759**	0.0006
5	0.5659	**0.5499**	0.5559	0
6	**0.5996**	0.6000	0.6602	0
7	0.5470	**0.5403**	0.5460	0
8	**0.6945**	0.7660	0.7157	0
9	0.4755	0.4974	**0.4635**	0.0258
10	**0.3358**	0.3380	0.3366	0
11	**0.4790**	0.5071	0.4965	0
12	0.4465	**0.4166**	0.4339	0
13	**0.5021**	0.5196	0.5205	0
14	**0.2799**	0.2910	0.2866	0
15	**0.3561**	0.3913	0.4289	0
16	0.4288	0.4196	**0.4141**	0.0132
17	0.1859	**0.1789**	0.1888	0
18	**0.2180**	0.2207	0.2207	0

As an example, Figure 1 shows the actual demand and the demand predicted by SES and MA in 52 weeks for product nr°2.

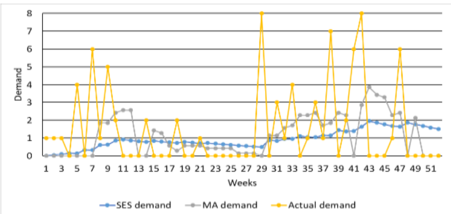

Figure 1: Actual demand and forecasting made with SES and MA (order 7) for code nr°2.

5. Products management

The total cost for the management of a product comprehends the purchasing of the product (C_u), the cost for the orders ($C_{eo} = 3.4 \frac{€}{order}$), backorder cost for each unit in backorder each time is not available ($C_{bo} = 20 \frac{€}{days*product}$), shortage cost for each unit in backorder the first time is not available ($C_s = \frac{2€}{product}$), and holding cost ($h = 0.0038 \frac{€}{week*product}$). In particular, regarding purchasing cost to maintain confidentiality the products have been grouped into five distinct classes based on demand characteristics and an average of their purchasing cost has been made. Class 1 (products 1-11, 14), Class 2 (products 12, 15, 17) and Class 3 (products 13, 16, 18) have an average purchasing cost of €27.40, €8.60 and €5.76 respectively. Order Emission Cost (C_{eo}), refers to the expenses associated with preparing and issuing reorder orders for pharmaceutical products, considering both administrative and logistical costs. Unit Backorder Cost (C_{bo}), is the cost associated with the unavailability of a product at the time period, which can cause delivery delays or compromise customer satisfaction. Unit Shortage Cost (C_s), represents the cost resulting from inventory shortages, when the requested quantity of a product exceeds the available quantity, resulting in lost sales opportunities. Percentage on Purchase Cost for Holding (h), is the percentage of the purchase cost added as inventory holding cost, including costs such as storage, depreciation, and the cost of tied-up capital. Knowing all the incurred costs allows us to calculate the weekly management cost for each product as shown in Equation 25. Where Q_g is the average inventory level in units, n_b is the number of backordered orders, n_o is the number of orders placed during the review period, Q_s is the quantity of units in shortage, n is the number of weeks in the reference period, which in the analysis performed coincides with 52 weeks that is test set.

$$Weekly\ Cost = Q_g * C_u * h + \frac{n_b * C_{bo}}{n} + \frac{n_o * C_{eo}}{n} + \frac{Q_s * C_s}{n}$$

(25)

5.1 Cost analysis with MA and SES

In the detailed analysis of total management costs, different configurations have been implemented and evaluated. In particular, we tested how total management cost for the different products is impacted by different forecasting models, different review period ($T = 2$ or $T = 4$) and by the absence or presence of safety stocks (SS) found exploiting Equation 18 imposing a CSL_{target} of 95%. Different forecasting models have been tested choosing between SES, MA and a mixed version where the best model in terms of $RMSE$ is exploited for each product. This focus on forecasting allowed for a more in-depth assessment of differences in total management costs based on the choice of demand forecasting method. Results for the different scenarios are shown in Figure 2.

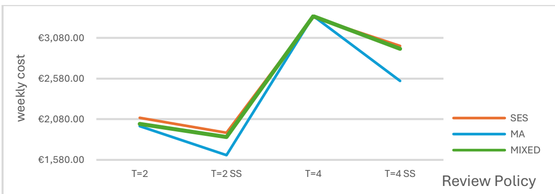

Figure 2: Management cost for each scenario

As visible from the Figure it can be observed that the lowest total management cost is achieved exploiting a reordering policy with a review period of two weeks ($T = 2$) with safety stock and using for each product MA as a forecaster. In addition, the figure shows how longer review period increases the management cost no matter which forecaster is used or the presence or not of safety stocks. This clearly indicates how to manage these types of products shorter review period are favorable. The best management is obtained using MA with $T = 2$ and SS which cost is 16.83% less than the management using SES and 13.62% less than using a mixed policy. This is due to the fact that for most of the products the best forecaster was SES (50% of the products) and in addition for 3 of the 4 products where the initial best forecaster was ARIMA it was substituted by SES. The overall better performance of MA in terms of cost can be attributed to the Mean Percentage Error (MPE) of MA that is over all the products of 26% respect to the overestimation of SES that is of about 24%. This overdemand forecasted by MA protects against unpredictable events that can cause high backorder costs. The best management of items leads to a mean weekly cost of 1,637.68 $\frac{€}{week}$ to manage all the products. However, the individual cost to manage products with MA and a (T, S) policy with CSL_{target} of 95%. varies a lot among the different products. In fact, the lowest cost, found for product nr°2, is of 1.97 $\frac{€}{week}$ while the highest, found for product nr° 13, if of 452.77 $\frac{€}{week}$. This difference is mainly guided by the different quantities requested for each product. Looking at the decomposition of the total cost we found that 84% of the cost can be attributed to the stocking of products, 14% to shortage and the remaining for orders and backorders costs. The high contribution of the stocking cost is due to the high CSL imposed to protect the system from stockouts.

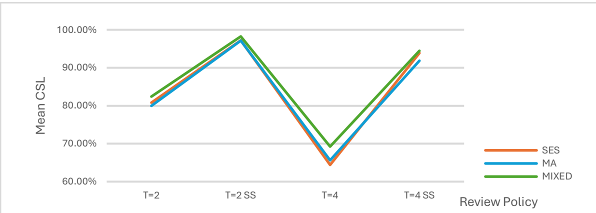

Figure 3: Mean CSL for each scenario

Regarding the CSL in Figure 3 we reported the mean CSL obtained in each scenario. As visible, with a review period of two-week, safety stocks and a CSL_{target} of 95% all the forecasting methods obtained similar results: MA achieve a CSL of 97.15%, SES of 97.22% while the mixed strategy of 98.25%. However, it has to be noted that the mean CSL is higher than the one imposed. This is due to the bias of the forecaster that overestimated the demand. Future research should focus on studying the relationship between forecasters, series characteristics and compliance with imposed CSL.

6. Conclusion

The objective of inventory management is to reduce the discrepancy between forecasted and actual demand to improve supply chain efficiency, customer satisfaction, and business profitability. In this study different forecasting models have been applied for the management of 18 pharmaceutical products. In particular, a time span of 249 weeks has been considered divided between a training set of 83% of the data and a test set of 52 weeks representing the remaining 17% of the data. From the initial phase the best forecasting models have been selected for each code, selecting at the end as best forecaster only SES and MA since provided very similar results to ARIMA the third optimal forecasting method individuated. Then, the management of the products has been evaluated exploiting a dynamic period review policy under different review period and the presence/absence of safety stocks. Results showed that lower revision period was preferred no matter the forecasting model exploited. However, the lowest cost was obtained using MA with safety stocks and a revision period of two weeks. At the same time different products show very different management costs with most of them related to the stocking cost, 84%, and shortage cost 14%. We also evaluate the obtained CSL having imposed a CSL of 95% noticing how it is always higher than the one imposed. This can be explained by the positive bias, around 25%, of the forecaster but future research should focus on explaining this difference based on series characteristics, type of forecaster and re-ordering policy. The results of the study are limited by the size of the initial dataset, but a similar approach can be applied to a wider dataset in the pharmaceutical sector to achieve generalizable results.

References

Bon, A. & Leng, C.Y., 2009, 'The Fundamental on Demand Forecasting in Inventory Management', *Australian Journal of Basic and Applied Sciences*, 3, 3937–3943.

Chatfield, C., 1978, 'The Holt-Winters Forecasting Procedure', *Applied Statistics*, 27(3), 264.

Croston, J.D., 1972, 'Forecasting and Stock Control for Intermittent Demands', *Journal of the Operational Research Society*, 23, 289–303.

Gamberini, R., Lolli, F., Regattieri, A. & Rimini, B., 2014, 'Dynamic Re-order Policies for Irregular and Sporadic Demand Profiles', *Procedia Engineering*, 69, 1420–1429.

Holt, C.C., 2004, 'Forecasting seasonals and trends by exponentially weighted moving averages', *International Journal of Forecasting*, 20(1), 5–10.

Hyndman, R.J. & Khandakar, Y., 2008, 'Automatic Time Series Forecasting: The forecast Package for R', *Journal of Statistical Software*, 27(3), 1–22.

Jain, C.L., 2003, 'Benchmarking Forecasting Practices in Pharmaceutical Industry', *Proceedings of Pharmaceutical SAS Users Group*, 22.3, 3–6.

Merkuryeva, G., Valberga, A. & Smirnov, A., 2019, 'Demand forecasting in pharmaceutical supply chains: A case study', *Procedia Computer Science*, 149, 3–10.

Mukhopadhyay, S., Solis, A.O. & Gutierrez, R.S., 2012, 'The Accuracy of Non-traditional versus Traditional Methods of Forecasting Lumpy Demand', *Journal of Forecasting*, 31(8), 721–735.

Spiliotis, E., Makridakis, S., Kaltsounis, A. & Assimakopoulos, V., 2021, 'Product sales probabilistic forecasting: An empirical evaluation using the M5 competition data', *International Journal of Production Economics*, 240, 108237.

Villa, S., Giusepi, I. & Lega, F., 2012, *Gli assetti organizzativi a supporto della gestione operativa: esperienze a confronto*.

Yang, Y., Ding, C., Lee, S., Yu, L. & Ma, F., 2021, 'A modified Teunter-Syntetos-Babai method for intermittent demand forecasting', *Journal of Management Science and Engineering*, 6(1), 53–63.