Order and Stock Costs Optimization in an Automotive Spare Parts Wholesaler

M. B. Cruz, S. F. Ramos, M. Pina, and R. Costa

Abstract We address a demand forecast problem posed by Nors, a Portuguese group working in the automotive sector and transport solutions. The aim is to define a procedure that efficiently forecasts the references demand and to obtain an optimized procedure for their ordering methodology. Nors works with several suppliers, resulting in a portfolio of hundreds of thousands of different parts number. Each supplier has its own lead-time and order periodicity. On an annual basis, and for each supplier, the yearly budget for purchases is agreed in order to define possible quantity discounts and the profitability of the business. Sales distribution is highly scattered, as there are high, medium and low rotation sales values for different references. Since the number of references and their total value is very high, the stock value and operational costs are non-negligible factors concerning the company costs. Nors goal is to reduce these costs while maintaining a high service level. In this paper, we implement a mixed methodology for the forecasting problem, present an alternative for the safety stock value, and discuss an optimization model to decide which and how many parts number should be ordered in each period. Finally, we present a set of computational and real implementation results, obtained with Nors data regarding two suppliers.

M. B. Cruz (✉)
LEMA - Engineering Mathematics Laboratory, School of Engineering, Polytechnic of Porto, Porto, Portugal
e-mail: mbc@isep.ipp.pt

S. F. Ramos
LEMA - Engineering Mathematics Laboratory, School of Engineering, Polytechnic of Porto and CEAUL - Faculty of Sciences, University of Lisbon, Lisbon, Portugal
e-mail: sfr@isep.ipp.pt

M. Pina · R. Costa
Nors Group, Porto, Portugal

1   Introduction and Problem Characteristics

It is widely recognized that the spare parts area is often the most profitable part of a corporation [13], especially when working at the repair, maintenance and aftersales business. In aftersales, the usage of appropriate forecasting methods reduces safety stocks, decreases several costs related with the stock level without penalizing the level of service and is one of the management key challenges [12]. In the aftermarket automotive sector this is even more critical for the wholesalers, as several aspects add complexity layers to the demand forecast and subsequent inventory management. Namely, the huge number of spare parts with heterogeneous sales patterns, the risk of stock obsolescence for some parts of the portfolio as well as the importance given by the costumers to the vehicle which lead them to have very small comprehension with long waiting times [1, 2, 4]. Despite the published work about this matter by many authors, there is a considerable gap between research and practice, as stated in [4] or [1]. In [11] it is presented a survey about the application of spare parts management methods at a national level, based on the work done by Rodrigues and Sirova in the Czech Republic. In this paper we address the practical implementation, in a Portuguese company, of a stock management process designed to meet the company constrains w.r.t. hardware and data availability, as well as to deliver accurate outputs and reliable Key Performance Indicators (KPI).

During this work, it was possible to observe very distinct sales behaviour across the references that compose the company portfolio. There are references with sporadic demand, in the sense that they are consumed in certain moments, followed by long and variable intervals with no demand, as well as there are references that are sold in high quantities for a long continuous period of time. In Fig. 1, it is presented the sales history for a given sub-family of 460 references (y-axis), across
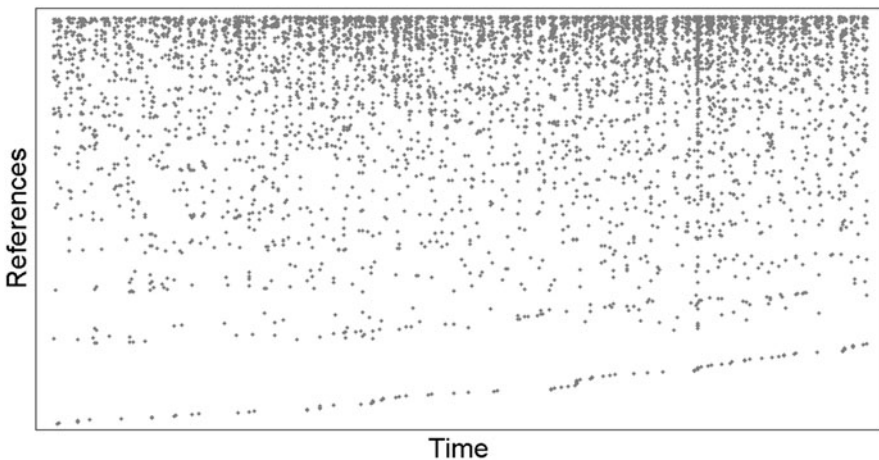

Fig. 1   Scatter of sales for a set of spare-parts references, during a 9-month period

a period of 9 months (x-axis). Each marker placed on (x, y) represents, at least, one sale of the y reference at time x.

Several facts may explain this behaviour. On the one hand, it is well known that the aftermarket spare parts sales for a given car model is, among other factors, correlated with its age. It is known that, some years after the model is launched there is an increase on the number of spare parts sold, followed by an equilibrium phase, and a subsequent decay on the end of the car's expected life [8, 9]. However, those trends depend also on many other factors as, for instance, the number of sold vehicles from that particular model, the number of years since the model was released, the expected life for each vehicle component. On the other hand, the same spare part may be applied in different car models, which may have distinct beginning and ending production periods, and a single car reference may have several different aftermarket suppliers. Finally, there are many additional factors that are related with the spare parts demands to a given company, as the mileage and the type of utilization of each car or the marketshare the company owns which is, in fact, time and reference dependent. All of these make very difficult for a wholesaler to accurately evaluate the parameters needed to deploy such a model.

Due to all the previous stated reasons, in the next section we present a model that uses several computationally cheap forecasting methods aiming at: being suitable for the different rotation type of references that compose the company portfolio; reacting quickly to market changes, and at maintaining the computational complexity at reasonable levels regarding its implementation on a company with regular hardware resources. Table 1 indicates the notation used in this paper, regarding the methods described in the next section, where, in its final subsections we present the results of their application to real databases and discuss these intermediate results. This will be the support to set the background regarding the

Table 1 Notation for the forecasting methods

Parameter	Notation	Method
d_t	Real demand for period t	All methods
$\hat{X}_t$	Demand forecast for period t	All methods
$\hat{S}_t$	Demand forecast for period t, given that the demand is positive	CR, SBA e TSB
$\hat{K}_t$	Forecast for period t, for the consecutive number of months with positive demand	CR e SBA
k_t	Number of months since the last positive demand, for period t	CR e SBA
$\hat{p}_t$	Probability forecast, for period t, of having positive demand	TSB
p_t	Binary variable, indicating positive demand for period t	TSB
a_t	Series level for period t	SAGA
b_t	Series trend for period t	SAGA
α, β	Smoothing constants ($0 \leq \alpha, \beta \leq 1$)	ES, CR, SBA, TSB, SAGA

optimization model that is presented on Sect. 3. Finally, we discuss the overall results on Sect. 4.

2   An Adjustable Model for Forecasting Spare Parts Demand

Several methods for the forecast of multi rotation references are presented in the literature [2, 6, 12]. Syntetos and Boylan [2] present a comprehensive review on forecasting spare parts demand, and the work of Romeijnders et al. in [12] is a very good survey on the subject this paper deals with, as it was based on the work done by the authors for Fokker Services, a company which deals with spare parts for aviation industry that has very similarities with the automotive business. Kennedy et al. [6] present also an overview on spare parts management that may be important for the second part of this paper.

2.1   Multirotation Reference Forecasting Methods

This section presents a brief review of methods for demand prediction, based on the description made by Romeijnders et al. in [12].

2.1.1   Zero Forecast

Zero Forecast method (ZF) forecasts the demand in the following period as 0 [12], that is:

$$\hat{X}_{t+1} = 0 \tag{1}$$

2.1.2   Moving Averages

The method of moving averages (MA) is one of the most famous methods for time series [7]. The demand on the next period is taken as the demand average of the previous N periods, eventually with respect to some weights $0 \leq \omega_i \leq 1$, where $\sum_{i=1}^{N} \omega_i = 1$. Unless stated otherwise, we use as default $\omega_i = \frac{1}{N}$:

$$\hat{X}_{t+1} = \sum_{i=1}^{N} \omega_i d_{t-N+i} \tag{2}$$

It is important to notice that Nors company already used MA to forecast the next period demand, with $N = 6$. We will denote this as MA(6) in further analysis.

Additionally, we used 12 months MA, with a combination for ω_i that gives more importance to the last three periods, a little less to the remaining of the last semester, and a small one to the first trimester. This will be stated hereafter as K-Method.

2.1.3 Naive Forecast

The so-called Naive Method (NF) is a particular case of the moving averages with N=1, as it sets the value of demand during the last period as the forecast for the following period [12].

$$\hat{X}_{t+1} = d_t \tag{3}$$

2.1.4 Exponential Smoothing

The Exponential Smoothing method (ES) is robust and is very well known as it fits itself rapidly to changes in the demand. Unlike MA, this method uses decreasing exponential combinations of past observations, to estimate future demands [3, 12]. In fact, it uses the forecast on the last period, adjusted by the prediction error $d_t - \hat{X}_t$, through the expression:

$$\hat{X}_{t+1} = (1 - \alpha)\hat{X}_t + \alpha d_t \tag{4}$$

2.1.5 Croston Method

In the original paper [5], Croston showed that the ES and MA methods didn't fit to series with intermittent demands. He proposed to update the quantity $\hat{S}_{t+1}$ and the searching interval $\hat{K}_{t+1}$ separately. The ES model is used to estimate those components in the periods of positive demand. The Croston forecast (CR) for the search quantity in period t is

$$\hat{S}_{t+1} = \begin{cases} \hat{S}_t, & d_t = 0 \\ (1 - \alpha)\hat{S}_t + \alpha d_t, & d_t > 0 \end{cases} \tag{5}$$

and the forecast for the number of periods with positive demand may be calculated through

$$\hat{K}_{t+1} = \begin{cases} \hat{K}_t, & d_t = 0 \\ (1 - \beta)\hat{K}_t + \beta k_t, & d_t > 0 \end{cases} \tag{6}$$

The forecast suggested by Croston method is obtained through:

$$\hat{X}_{t+1} = \frac{\hat{S}_{t+1}}{\hat{K}_{t+1}} \tag{7}$$

2.1.6 SBA: Syntetos–Boylan Method

Syntetos and Boylan [14] presented in 2001 an update to the Croston method, the Syntetos–Boylan method (SBA). To avoid some bias, they proposed to include the factor $1 - \frac{\alpha}{2}$ in the Croston method, getting the following expression:

$$\hat{X}_{t+1} = \left(1 - \frac{\alpha}{2}\right) \frac{\hat{S}_{t+1}}{\hat{K}_{t+1}} \tag{8}$$

where $\hat{S}_{t+1}$ and $\hat{K}_{t+1}$ are evaluated using the expressions (5) and (6).

2.1.7 Teunter Method

Teunter method (TSB) [15] is an alternative to Croston method that updates the positive search probability, instead of updating the forecast for the periods with positive demand like the CR and TSB methods. The demand forecast for period t is achieved through:

$$\hat{S}_{t+1} = \begin{cases} \hat{S}_t, & d_t = 0 \\ (1 - \alpha)\hat{S}_t + \alpha d_t, & d_t > 0 \end{cases} \tag{9}$$

and the probability to have positive search is given by:

$$\hat{p}_{t+1} = (1 - \beta)\hat{p}_t + \beta p_t \tag{10}$$

As so, the forecast given by the Teunter method may be calculated by:

$$\hat{X}_{t+1} = \hat{p}_{t+1}\hat{S}_{t+1} \tag{11}$$

2.1.8 SAGA or Holt Method

The SAGA method, also known as Holt method, is suitable for series with linear trend. It uses the exponential smoothing technique to estimate the level a and the growing trend b [3]. This model uses the following update expressions:

$$a_t = \alpha d_t + (1 - \alpha)(a_{t-1} + b_{t-1}), \tag{12}$$

and

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1} \tag{13}$$

The forecast for the next period is obtained through:

$$\hat{X}_{t+1} = a_t + b_t \tag{14}$$

2.1.9 Poisson Method

Under certain well specified conditions, the Poisson distribution is used to count events that occur randomly in a given interval of time. This distribution can be used to get a forecast, as an alternative to methods based on time series [10]. This was adapted to this problem, using as rare event the quantity sold in a given period of time. As so, defining X_i as the random variable that represents the number of pieces sold for a given reference i on the time interval Δ, and supposing that $X_i \sim P_o(\lambda)$ with $\lambda = E(X_i) \approx \frac{1}{n}\sum_{j=1}^{n} d_{ij}$ where $\sum_{j=1}^{n} d_{ij}$ represents the total number of reference i articles sold in n periods Δ, the forecast, $\hat{X}_i$, for Poisson model (PO) is obtained solving the equation

$$P(X_i < \hat{X}_i) = 1 - \alpha \tag{15}$$

where α represents the desired confidence level.

Beside the methods described, many other could be inserted on this model. Indeed, several other methods were considered to be included in this work but, those failed to meet one of the requirements asked by the company: that its implementation should deliver results in reasonable time using the existing hardware. The Neural Networks approach and time series methods (like ARMA, ARIMA or SARIMA) didn't improve significantly our results in the several tests where this model was used and/or increased considerably the overall running-time to values that were not compatible with the company's available time to process the orders.

2.2 The Mixed Model: MM

The approach implemented in Nors was based on selecting, in each order period, the best method for each reference. With that purpose, we built a metric that uses recent periods to decide which method does fit better the present sales series and use it in the forthcoming forecasting period. More precisely, it receives the demand history for a given reference, and estimates which is the most suitable method for that series. This was done through the following error function (by default, the test

period uses $N = 18$ months)

$$\text{Error}_t = 0.6 \sum_{i=1}^{\lfloor N/2 \rfloor} (d_{t-i} - \hat{X}_{t-i})^2 + 0.4 \sum_{i=\lfloor N/2 \rfloor+1}^{N} (d_{t-i} - \hat{X}_{t-i})^2 \qquad (16)$$

where $\lfloor . \rfloor$ represents the floor function.

The methods MA, ES, CR, SBA, TSB e SAGA are parametric, implying that their use depends on the values assigned to each of their parameters. We made this using a brute-force parametrization test and selecting for the forecast of the next unknown period the combination that minimizes the Mean Squared Error (MSE) for the selected m test periods:

$$\text{MSE}_t = \frac{1}{m} \sum_{i=1}^{m} (d_{t-i} - \hat{X}_{t-i})^2 \qquad (17)$$

2.2.1 The Dynamic Safety Stock Approach

The methods described in the last subsection deliver the forecasted demand for the next period. Apart from this, another important quantity to determinate is the Safety Stock (SS). As stated in [16] this quantity may be understood as an *extra stock*, important to prevent eventual consumptions above the predicted demand on the following period. In order to determine the real quantities to be ordered for the next period, Nors used a safety margin based on Economic Order Quantity (EOQ) expression

$$\text{EOQ}_t = OP_t - Q(0)_t + \hat{X}_{t+1} \times \epsilon \qquad (18)$$

where OP_t is the order point, ϵ is a parameter given as function of both reference cost and next period forecast for each reference and $Q(0)_t$ represents the existing stock for the given reference at the time of placing the order (units). Those ϵ-values, that were pre-defined by the company commercial department, are presented in Table 2.

Table 2 Number of extra weeks to order, as function of reference's cost and demand forecast

		Next period demand forecast						
		A	B	C	D	E	F	G
Price		≥ 50	[10;50[	[5;10[	[2;5[	[1;2[	[0.5;1[	[0,0.5[
1	[0;1[	3	4	4	4	3	2	1
2	[1; 10[	2	2	3	4	3	2	1
3	[10; 50[	1	1	2	2	2	1	1
4	[50; 300[	1	1	1	2	2	1	1
5	≥ 300	1	1	1	1	1	1	1

Taking advantage of the data obtained from the tests made previously when selecting the best method, and given that the main reason for the safety stock is to cover the instances where the method underestimates the demand for the next period, we calculate the safety stock quantity in a different way. In fact, to evaluate the SS, we only use the periods in the past where the real demand was above the forecast

$$\text{Error}_{med_t} = \frac{1}{\#A_t} \sum_{i \in A_t} (d_{t-i} - \hat{X}_{t-i}), \ A_t = \left\{ i = 1, \cdots, n : d_{t-i} > \hat{X}_{t-i} \right\} \quad (19)$$

This value is upscaled by a risk factor,

$$\text{SS}_t = q_{1-\alpha'/2} \times \text{Error}_{med_t} \quad (20)$$

where q represents the $1 - \alpha'/2$ quantile of a suitable distribution for the errors series. The risk factor parametrization is intended to be settled using the company's technical knowledge (e.g., taking into account each reference critical importance).

As usual, the order point is defined as $OP_t = \hat{X}_t \times LT + SS_t \times LT$, where the Lead-Time (LT) is measured in the same temporal scale as the forecast $\hat{X}_t$.

2.3 Computational Results with Real Data

The previous model was implemented and applied with two real datasets, DS12016 and DS22016, consisting of 1172 and 4032 references, respectively, belonging to two suppliers for which this methodology has been tested and applied. The following results represent the comparison of the model within the year of 2016 (52 weeks) against the older model implemented in Nors that was based only in the MA(6). For this work, we also had access to the previous 30 months history of sales w.r.t. these references. The lead-time for this supplier is 28 days, and the orders were launched weekly, starting in January 4, and ending in December 26.

It should be noticed that although there are some references without sales (145 and 1010 for each dataset, respectively) for the whole year of 2016, we kept them in the datasets in order to reproduce real-world conditions in these tests, as it is impossible to know in advance if a reference will or not be sold in a given year. Those were only deleted when evaluating the relative errors. For the same reason, we also include the returned items in 2016, from 2015 sales. As so, the quantity sold per reference may be negative.

In Tables 3 and 4, the descriptive measures to characterize the datasets reinforces the idea of great amplitude of sales distribution per reference, introduced in the beginning this paper, as high, medium and low rotation products coexist in the same family of products. This factor may be one of the main factors that leads to the values presented on the last line of this table, representing the number of different

Table 3 Descriptive statistics regarding the application of MM in DS12016

Parameter	Min	First Q	Median	Mean	Third Q	Max	Std	CV
Sold qty per ref.	−6	1	5	23.8	17	1071	71.0	3.0
Forecasted qty per ref.	0	1.2	4.9	22.8	17.1	997.6	67.4	3.0
Av. nr. of days between sales	1.5	20.4	61.2	96.9	122.3	≥365	104.8	1.1
Nr. of different methods/ref:	1	2	4	3.7	5	9	2	0.55

Table 4 Descriptive statistics regarding the application of MM in DS22016

Parameter	Min	First Q	Median	Mean	Third Q	Max	Std	CV
Sold qty per ref.	−23	0	2	20.4	5	12,092	287.3	14.1
Forecasted qty per ref.	0	0.27	1.4	19.8	5.2	11,231.6	276.9	14.0
Av. nr. of days between sales	1.3	52.4	122.3	164.4	≥365	≥365	129.2	0.8
Nr. of different methods/ref:	1	2	3	3.3	5	9	1.8	0.6

Table 5 Impact by method in DS12016 and DS22016

Parameter	DS12016			DS22016		
	Global impact (%)	Max. best method (wk)	Ref w/most usage (%)	Global impact (%)	Max. best method (wk)	Ref w/most usage (%)
ZF	40.7	52	42.7	51.5	52	58.1
NF	1.3	10	0	1.0	5	0
Poisson	0.0	2	0	0.0	5	0
MA(6)	0.7	15	0	0.5	13	0
ES	0.4	6	0	0.2	4	0
MA	21.1	42	20.9	16.7	35	16.6
SAGA	1.5	10	0	1.0	3	0
CR	0.8	9	0	0.7	8	0
SBO	1.1	13	0	0.8	12	0
TSB	0.2	5	0	0.2	8	0
K	31.7	45	36.4	27.5	31	25.3

methods per reference used to forecast at least one period in 2016, which seems to justify by itself, the idea presented in Sect. 2.2.

As so, proposing a single method for each reference is not a good choice. On average, each reference had its best prediction using between 3 and 4 different methods across the whole year. With respect to the performance of each method, Table 5 and Fig. 2 present the main idea although some comments should be made with that respect.

First, all the methods are used somewhere across the year and, as the computation time is acceptable for the company, there is no reason why any of them should be removed. Second, the K-method clearly overperforms the MA method which was the choice prior to this new implementation on Nors. Finally, the high frequency of ZF may be explained by its usage in references that have very scattered sales, as also
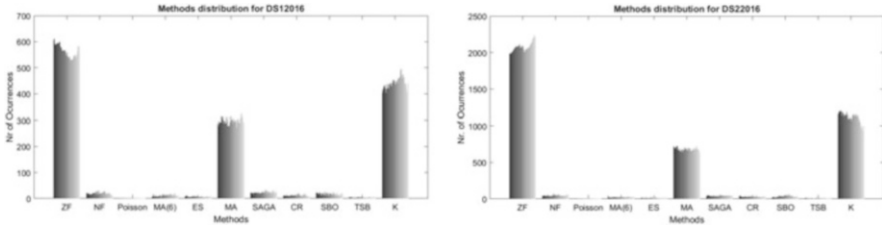
Fig. 2 Method's histogram regarding the number of times a method was selected as best, for each reference across the whole 52 weeks history

Table 6 Descriptive statistics for errors concerning the application of MM to both datasets

Dataset	Parameter	Min	First Q	Median	Mean	Third Q	Max	Std	CV
DS1	Real errors	−75	−1	0	−1.0	0	18	5.4	5.4
DS1	Relative errors (%)	−100	−13.0	0	−4.3	0	300	33.1	7.7
DS2	Real errors	−860	−1	0	−0.6	0	277	16.3	26.2
DS2	Relative errors (%)	−500	−8.0	0	−5.6	0	414	35.2	6.3

to references that have no sales at all. Nevertheless, it is impossible to know that in advance, and as such the decision to eliminate a given reference from the portfolio is a management decision, far beyond the scope of this work. We also highlight the fact that if a reference is predicted by means of ZF, that doesn't mean that the final ordered quantity will be 0, as the corresponding safety stock quantity may be positive.

For all these reasons, the MM mixed model is an upgrade to Nors prior state, as Table 6 clearly shows. The plots from Fig. 3 also allow to visualize the quality of predictions obtained by the model for both datasets.

2.4 Financial Results

As usual, the main quality indicators of the forecasting model in this industry are related with the service level and sales volume. In Table 7 we present those parameters obtained from simulating MM applied to both datasets against the simulation of old methodology, where we settled $q_{1-\frac{\alpha'}{2}} = 2$, in Eq. (20).

Table 7 Financial details for new (MM) and old (MA(6)) methodology for both datasets

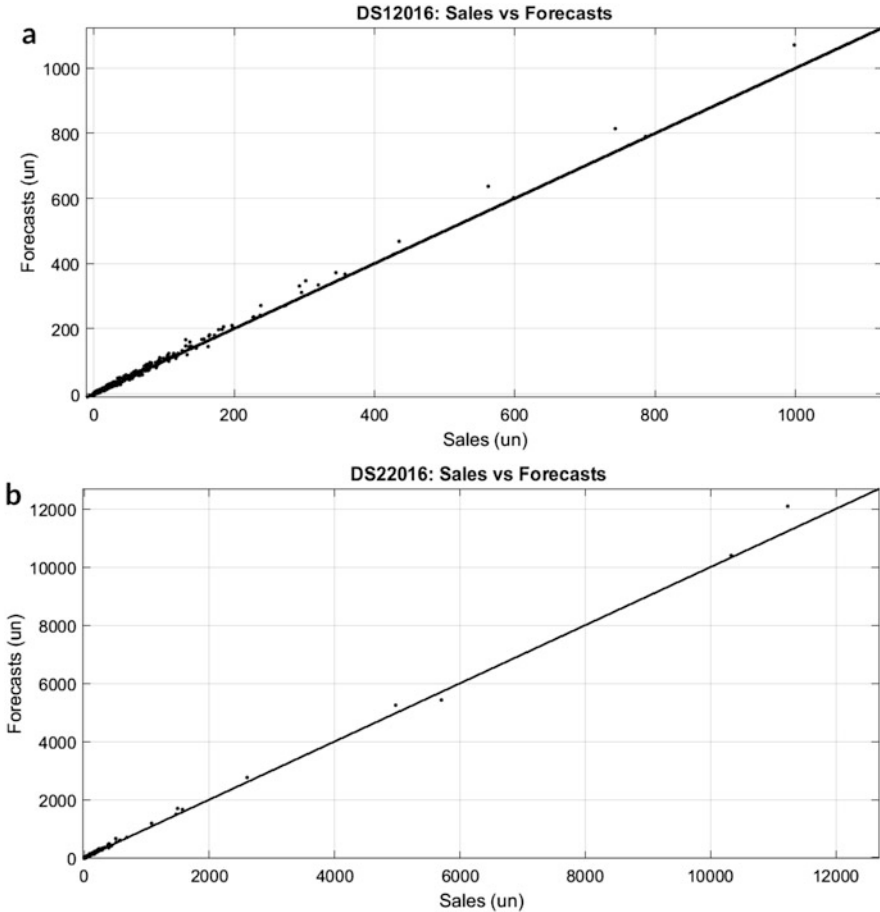| Dataset	Method	Service level	Sales volume	Initial stock	Final stock
DS12016	MA(6)	94.0%	264,152 €	68,112€	35,471€
DS12016	MM	95.0%	267,240€	68,112€	38,999€
DS22016	MA(6)	93.6%	1,027,987€	556,806€	260,897€
DS22016	MM	95.1%	1,052,687€	556,806€	286,083€

Fig. 3 Real errors applying MM to DS12016 (**a**) and DS22016 (**b**)

Given that it is common knowledge in this industry that increasing the service level towards values above 90% is very expensive, as well as that the presented sales volume is measured in cost price (in particular, the increase in sales has a higher return than the one stated in Table 7's third column) the previous results are good outputs of this model from a financial point of view. Is also important to notice that in the past, the reliability of the prediction model wasn't accurate enough, as they didn't used the SS notion, but only the safety margin given by the ABC procedure in Table 2. We made the simulations with the SS approach for all the methodology. If not, the service levels for MA(6) would drop to levels below 85%. In the past, to tackle that problem, the product managers increased the sales prediction by some factor related with their experience. Obviously, this had several negative aspects such as: depending on each manager experience and making the company's performance very vulnerable to changes in the staff. Also, the effect

Stock Behaviour(2015–2016) results
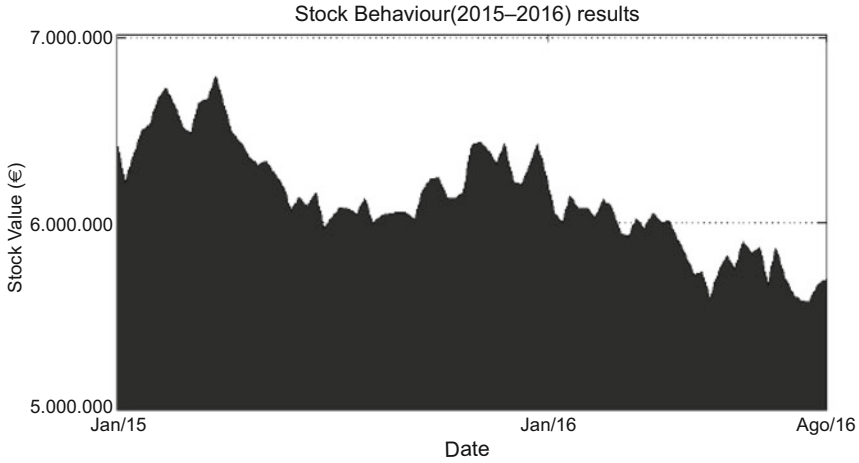
	2014	**2015**	**Variation**	
Mean Daily Stock (m€)	7.792	6.383	−1.409	−18%
Sales Volume (m€)	19.721	21.633	+1.712	+9%
Service Level (%)	91.9	92.8	+1%	
Cover rate (months)	5.7	4.2	−1.5 months	

Fig. 4 Evolution of stock value from Jan 2015 till Aug 2016 (top) and some important indicators from the previous year till the year after this model was implemented (bottom)

of trying to protect the company against out-of-stock episodes led it to reinforce manually the safety margins. In the end, the effects of applying this methodology to the whole set of the company's products in late 2014, led to the global results presented on Fig. 4.

3 MORS: A Lightweight Model for the Supplier Order Processing System

Due to the success of the MM approach explained in the previous section, it was decided to develop a new mathematical model to optimize the order processing system. One of the reasons to develop this optimization model—denoted as MORS in the forthcoming references—was related with the fact that often, product managers decide to place orders of considerable amount when the demand forecasts don't indicate so. Sometimes, this behaviour aims to accomplish contracts and to get some discounts agreed with the supplier when negotiating the rappel in the
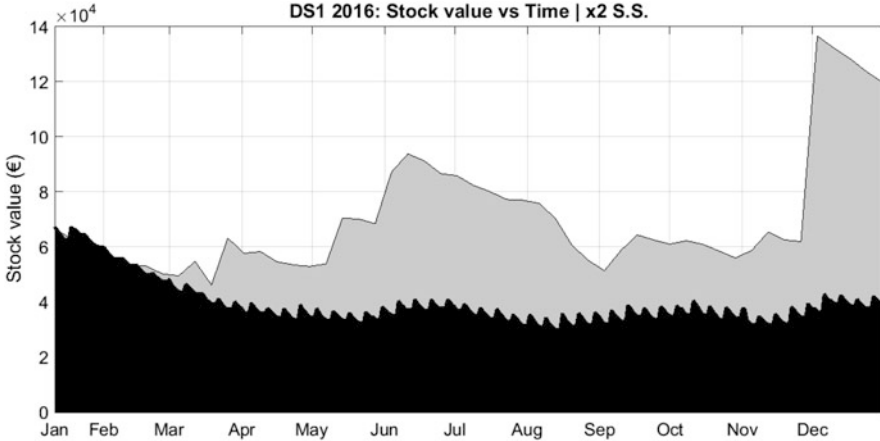
Fig. 5 Theoretical (black) vs Real Stock (grey) in DS12016 family

beginning of the year. This leads to several stock unbalances across the year, with low levels of stock during some parts of the year, that are compensated with big amounts of stock in the end of the year in order to reach the agreed amounts to accomplish the contracts or to reach some level of discount. An example may be seen in Fig. 5, where the daily stock value for the MM model (black) is plotted against the real stock (grey).

Table 8 presents the notation used for the optimization model regarding a supplier of a set comprising N references ($i = 1, \ldots, N$), supposed to be ordered in a periodic T-days scheme ($t = 1, \ldots, T$).

As those thresholds are known in the beginning of the year, our idea is to distribute the buying quantities all across the year, introducing the variable total budget (TB) and/or minimum budget to be spent (tb) as a parameter on the optimization model. This quantity is intended to be defined and evaluated across the year by the buyer manager, allowing him to decide the minimum (and/or maximum) amount to be spent in each order, as defined in constrains (30, 31). In our implementation, we distributed that value uniformly across the year, but this may be easily adapted to other types of requirements. It is also important to notice that, in grounds of a considerable loss in sales w.r.t. the goods supplied by a given supplier or some changes in the contract, the existence of this variable allows for a quick change/adaption to the buying strategy at any time of the process.

It is also natural for the manager to decide whether to restrain the number of days for which, in regular conditions, the system is allowed to maintain stock. Denoting this quantity by $Tmax_i$, the inequality (32) guarantees this possibility.

Obviously, non-feasibility of the model may be a strong indicator that the management decisions on TB and $Tmax$ may be incoherent. This may become an additional advantage of the model, as it may alert the managers to the existence of strategies that may be simultaneously unfeasible.

Table 8 Notation for the optimization model

Parameter	Description
Sets	
S	Set of all references from the given supplier
S_M	Set of the mandatory references to be ordered at the present time from the supplier
S_A	Set of the admissible references, that is, the ones that may optionally be included in the present time order to the supplier
$S' = \{S_M \bigcup S_A\}$	Set of the references allowed to be ordered in the present time from a given supplier
Decision variables	
x_{ij}	(binary decision variable) reference i is ordered for a period of $j \times P$ days
Quantities	
$Q_i(0)$	Existing stock for reference i at the time of placing the order (units)
$\hat{D}_i = \frac{\hat{X}_L}{t}$	Expected daily demand for reference i (units/day)
O_i	Reference i quantity already ordered but not delivered at the present time
R_i	The maximum inventory level for the i-th product (units)
SS_i	Reference i safety stock (units)
OP_i	Order point for reference i (units)
$I_i(t)$	The expected inventory level for the i-th product at day t (units)
Costs and incomes	
MC	Maximum inventory monetary value allowed for the existing stock w.r.t. the supplier (€)
TB	Total available budget for the supplier (€/year)
tb	Minimum budget to be spent with the supplier (€/year)
IR	Interest rate
pc_i	Reference i purchasing cost (€)
hc_i	Reference i holding cost per day (€/day)
fc_i	Reference i fixed ordering cost (€)
vc_i	Reference i variable ordering cost (€/piece)
Time quantities	
P	Supplier order periodicity (days)
LT	Supplier Lead-Time (days)
$Tmax_i$	Maximum number of days allowed to hold stock
Tf_i	Expected time-period needed to sold out all the reference i inventory (days)
Volumes	
wva	Maximum warehouse volume available for the whole stock from the supplier (m^3)
v_i	Reference i volume (m^3)
Other parameters	
SL_i	Reference i service level

Another side effect of this model, is the minimization of distinct buys per reference, as decreasing the number of different references purchased in each order pulls an additional positive effect at the warehouse level. Indeed, handling and put-away timings are important drawbacks within the warehouse management. The effect of a economy of scale, using deeper and narrower orders, may have an important impact on the global efficiency results. We noticed, prior to this model application, that there were a non-negligible number of occurrences where products, although at the warehouse, were not available to be sold as their status was "in conference" leading to worst performances at the level-of-service KPI.

Regarding the number of days in stock, we also added a constraint (33) to assure that, if the forecast model is accurate, the order will provide enough pieces of each reference to prevent that no stockout takes place before receiving the order placed in the next period.

From the implementation point of view, we ran a pre-processing analysis on the references list, splitting the references set in three different subsets. The mandatory subset, S_M, which comprises the references that must be ordered, as their stock is forecasted to end before the arrival of the next order. The admissible subset, S_A, comprehends all the references that don't belong to S_M, but whose stock is forecasted not to last more than a threshold $\tau \geq P + LT$ number of days, where P stands for the supplier order periodicity. These references, may or not be ordered within the current order (28, 29). All the references that don't belong to $S_M \cup S_A$, are discarded from the optimization model, and we will denote the resulting list as S'. In the scope of this real-world implementation, this fact may have important effects, as the number of different references by supplier may reach several thousands resulting in a significant reduction in the number of variables, which impacts the resources (memory/time) needed to run the model. We also restricted the maximum number of days for which the reference stock may be ordered to $Tmax$. We only allow our model to order the quantity of a given reference needed for multiples of the order periodicity. This option allow us to transform the integer decision variables for this type of problem $Q_i \in \mathbb{N}$, for $i \in S'$, that represent the reference i replenishment quantity, into the set of binary variables $x_{ij} \in \{0, 1\}$, $\forall i \in S'$, indicating the number of days $j \times P$, $j = 0, 1, \ldots, \lfloor \frac{Tmax}{P} \rfloor$ in which each reference i may be ordered. This transformation allow us to determine the replenishment quantity, using the identity

$$Q_i = \sum_{j=0}^{\lfloor \frac{Tmax.}{P} \rfloor} j x_{ij} \hat{D}_i, \ \forall i \in S' \tag{21}$$

where $\hat{D}_i$ represents the expected daily demand for reference i. The warehouse space availability when receiving the order may also be added to the model through the inequality (34).

As our goal is to minimize the handling, storing and interest costs, the following components should be included in the objective function:

- The financial costs,[1] associated to purchase the references needed to fulfil the demand throughout a given period. Defining the reference i purchasing cost (pc_i) and the interest rate (IR), this may be represented as a function of the number of days (n) we decide to place the order,

$$f_1(n) = \sum_{j=1}^{n} pc_i \, \hat{D}_i \, ((1 + IR/360)^j - 1)$$

$$= pc_i \, \hat{D}_i \, (1 + IR/360) \frac{1 - (1 + IR/360)^n}{1 - (1 + IR/360)} - (pc_i \, \hat{D}_i)n \qquad (22)$$

$$= pc_i \, \hat{D}_i \, ((1 + IR/360) \frac{(1 + IR/360)^n - 1}{IR/360} - n)$$

- The fixed cost fc_i associated with the administrative process of placing the order,

$$f_2(n) = fc_i \qquad (23)$$

- The cost associated with the handling and put-way processes, when the goods arrive at the warehouse, where vc_i represents the variable cost per unit ordered,

$$f_3(n) = n \, vc_i \, \hat{D}_i \qquad (24)$$

- Storage costs, which include warehouse rent and stock insurance, where hc_i represents the reference i unitary holding cost per day (€/day)

$$f_4(n) = \sum_{j=1}^{n} hc_i \, \hat{D}_i \, (1 + IR/360)^j$$

$$= hc_i \, \hat{D}_i \, (1 + IR/360) \frac{1 - (1 + IR/360)^n}{1 - (1 + IR/360)} \qquad (25)$$

$$= hc_i \, \hat{D}_i \, (1 + IR/360) \frac{(1 + IR/360)^n - 1}{IR/360}$$

In order to make the comparison between different strategies simpler, a normalization is made with respect to the number of days for which a order is placed. As a consequence, the cost function with respect to a given reference i in the order placed

---

[1]We focused only on the interest value, as we don't consider the obsolescence rate.

to fulfil the needs for an n-days period, is defined by the following function:

$$C_i(n) = \frac{hc_i \hat{D}_i(1 + IR/360)((1 + IR/360)^n) - 1}{n\,IR/360} +$$

$$+ pc_i\,\hat{D}_i\left((1 + IR/360)\frac{(1 + IR/360)^n - 1}{n\,IR/360} - 1\right) + \tag{26}$$

$$+ vc_i\,\hat{D}_i + \frac{fc_i}{n}\ (\text{€/day})$$

where, w.r.t. the solution for the optimization model, the parcel $vc_i\,\hat{D}_i$ may be dropped.

The plot of $C_i(n)$ for a given set of references is presented in Fig. 6, showing that for each reference there are very distinct optimal order periods.

Given the previous considerations, the optimization model may be written as:

$$\min \sum_{i \in S'} \sum_{j=0}^{\lfloor \frac{Tmax.}{P} \rfloor} C_i(jP)x_{ij} \tag{27}$$

s.t.

$$\sum_{j=0}^{\lfloor \frac{Tmax.}{P} \rfloor} x_{ij} = 1,\ \forall i \in S_M \tag{28}$$

$$\sum_{j=0}^{\lfloor \frac{Tmax.}{P} \rfloor} x_{ij} \leq 1,\ \forall i \in S_A \tag{29}$$

$$\sum_{i \in S'} pc_i\,Q_i \geq \frac{tb}{\frac{365}{P}} \tag{30}$$

$$\sum_{i \in S'} pc_i\,Q_i \leq \frac{TB}{\frac{365}{P}} \tag{31}$$

$$Q_i(0) + O_i + Q_i - LT * \hat{D}_i - SS_i \leq Tmax_i * \hat{D}_i,\ \forall i \in S' \tag{32}$$
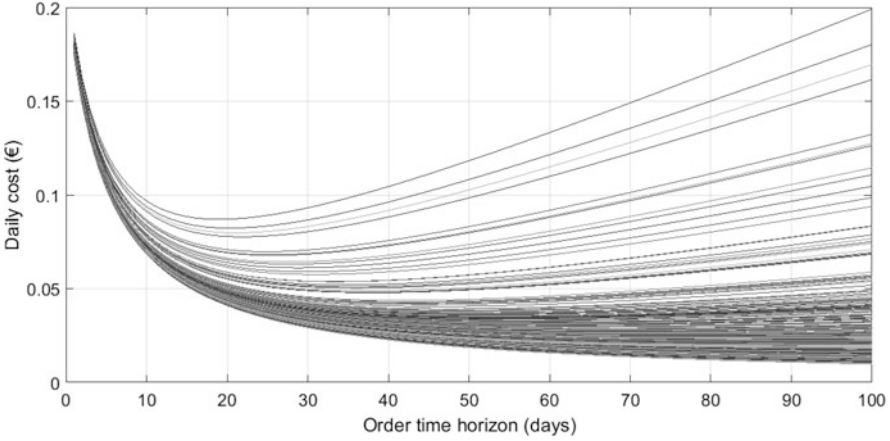
Fig. 6 Cost per day, for a given set of references, if the order is settled to cover the demand for the next x days

$$Q_i(0) + O_i + Q_i - LT * \hat{D}_i \geq (LT + P) * (\hat{D}_i + \frac{SS_i}{P}), \ \forall i \in S' \tag{33}$$

$$\sum_{i \in S'} v_i(Q_i + I_i(LT)) \leq wva \tag{34}$$

$$x_{ij} \in \{0, 1\} \tag{35}$$

4 Computational Study and Implementation Results

This optimization model was implemented using Matlab R2016b with the optimization toolbox. For both the DS12016 and DS22017 datasets, real sales and predictions given by the model described on Sect. 2.2 and respective results were compared with the existing rules for the Nors group (denoted as MM), as well as with the real order placed to the supplier. As in the previous section, we added the MA(6) method in order to be possible to check additional quality and financial indicators.

For the DS12016, we made 52 orders, resulting in a mean processing time of 20 s per order in a virtual machine with OS Windows 10, 8Gb RAM and a processor 2.8 GHz Intel Core i7. With the same hardware the solution for DS22016 dataset was exported, on average, after 46 s of computation time. Those times include the data importation and the solution exportation to an *Excel* worksheet, which is then imported by Nors native's ERP. We settled the total budget (TB) to the real amount bought by Nors in 2016 with each of the suppliers, and the order periodicity to 7
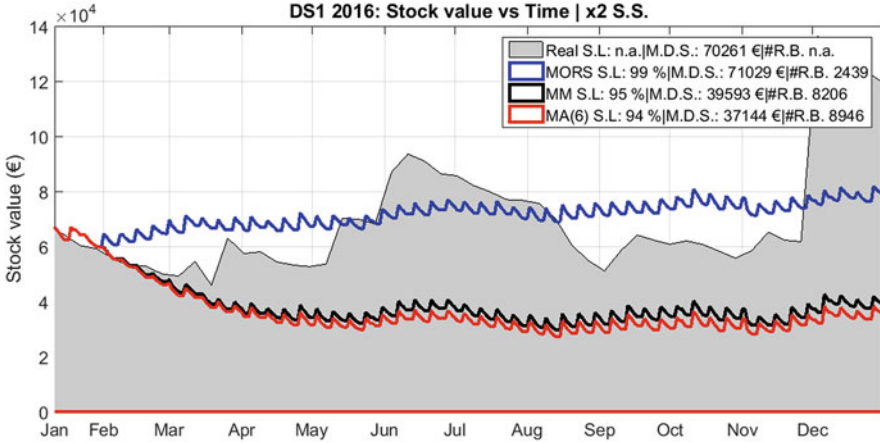
Fig. 7 DS12016: Comparative performance of three methods vs real data, using $2\times$ Safety Stock and $T_{max} = 364$ days

days. We didn't set the *tb* parameter in (30), as was our purpose to demonstrate that the budget for DS12016 was clearly excessive even when $T_{max} = 364$, and that was not an issue regarding the other dataset. The maximum order value per week was defined as TB/52 although it may be settled in other ways (e.g. the remaining yearly budget divided by the number of remaining weeks till the end of the year). The maximum number of days to hold stock was defined in 364 days or 126 days (two different simulations) for DS12016, because the annual budget to this supplier was clearly excessive (Fig. 5). For DS22016 that value was settled to 94 days, as the yearly budget was much tighter w.r.t. the company needs. Both suppliers had an expected lead time of 28 days. For the MM and MA(6) the number of weeks for which the reference is ordered is one month plus the value indicated by EOQ (Table 2), and all the simulations used $Z_{1-\frac{\alpha'}{2}} = 2$ in (20).

In terms of daily stock, the overall behavior of the method is presented in Figs. 7, 8, and 9, with the following legend abbreviations:

- Real—Real stock on the company.
- MORS—Stock behaviour using MORS optimization process.
- MM—Stock behaviour using the MM model cf. Sect. 2.2.
- MA(6)—Stock behaviour using Nors original forecast model.
- S.L.—Service level (%).
- M.D.S.—Mean Daily Stock (€).
- #R.B.—Total of references unique buy's.

Analysing Table 9, where the main KPIs for the whole set of simulations are presented, some interesting facts are highlighted. The total amount really bought by Nors w.r.t. the DS12016 dataset is clearly excessive for the defined maximum period to hold stock. At the same time, the comparison between the MORS optimized
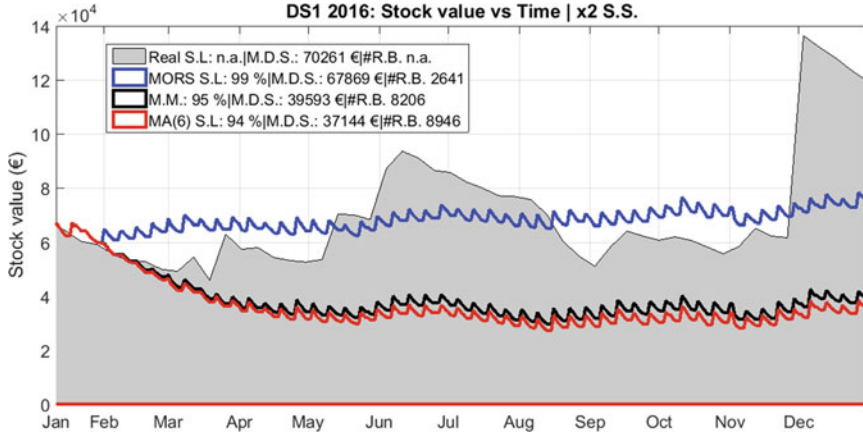
Fig. 8 DS12016: Comparative performance of three methods vs real data, using $2\times$ Safety Stock and $T_{max} = 126$ days
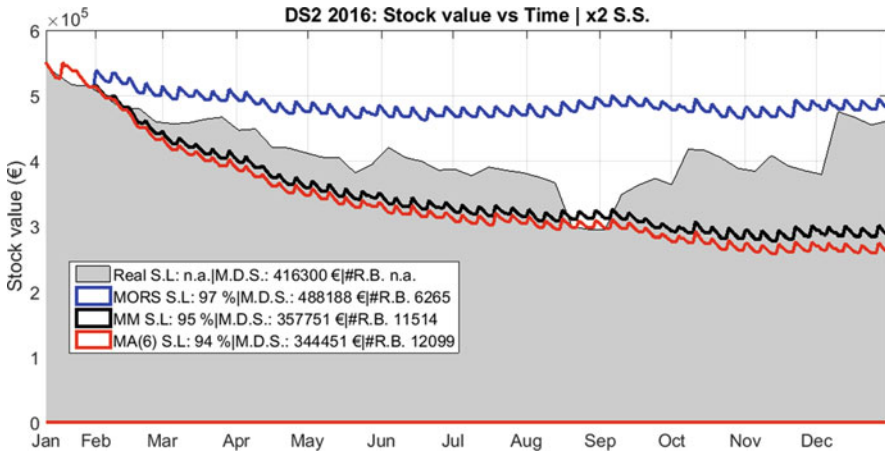

Fig. 9 DS22016: Comparative performance of three methods vs real data, using $2\times$ Safety Stock and $T_{max} = 91$ days

model and the MM model, both based on the same forecasting model, clearly gives much better service-levels to our model. Also, the number of orders of different references made all over the year, reduces by a factor of 0.32 (0.54 for DS22016) when compared with the MM model. The purchase patterns of the models presented in Fig. 10, clearly emphasizes advantages at the warehouse level. The same pattern is found when looking at the mean percentage of times that the references are ordered in the whole year: 2.2 (per reference) for MORS versus 7.0 for the MM method, which emphasizes that impact, as it represents a very significant reduction on the number of different references that must be manipulated in the warehouse over the
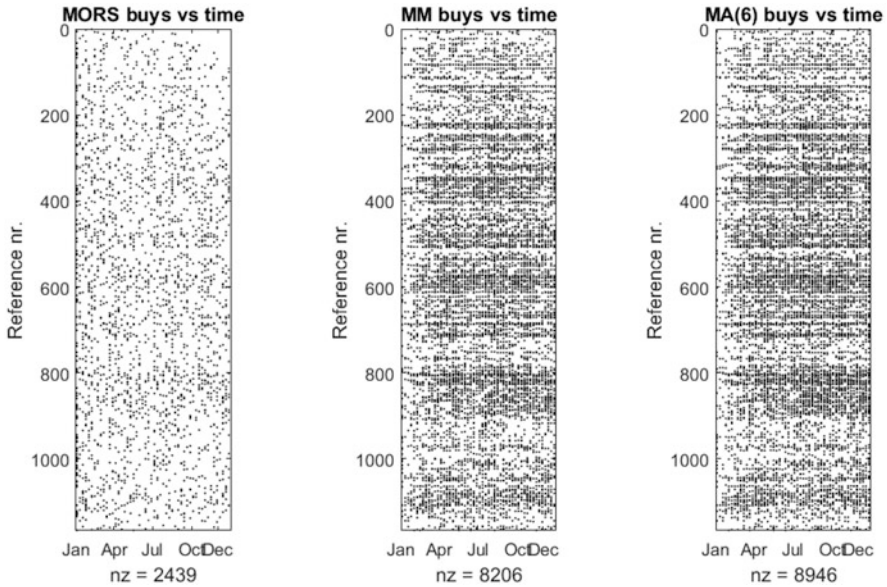
Fig. 10 DS12016: Purchases pattern for three methods with $T_{max} = 364$ days

year. These values are 1.6 and 2.9 for DS22016, resulting in 54% of reduction on the optimized model.

It is also important to notice that our model's Objective Function (27), had a global cost reduction factor between 0.5 (DS12016) and 0.71 (DS22016), which reflects very important savings within the logistics costs. The absolute values are not presented because of confidentiality related aspects. A final remark should be made w.r.t. the stock and/or purchases amounts as we just had access to the existing stock for those families in a weekly base (the day where the order was placed), as well as the orders placed in December 2015, that should had been delivered in January 2016. As so, some minor inconsistencies in this section may occur, regarding service levels in the first 28 (Lead-time) days of January or real stock levels. Nevertheless, independently of the goods being delivered in the beginning of 2017 due to lead time, we accounted for all the orders placed during 2016 in the financial results summary presented in Table 9. Concerning the lost sales KPI (Figs. 11 and 12), we evaluated these values based on the real sales recorded by the company. This happened because Nors didn't keep the record of lost sales in the system, and, as so it was impossible to evaluate that KPI accurately afterwards. The lost sales value is reduced in more than 50%, although based on different levels of purchases as the MM model doesn't take into account the available budget.

Figure 13 presents the order periodicity for the whole set of references, where a significant heterogeneity may be seen, resulting of minimizing order's cost according to the optimization model.

Table 9 Results main indicators for DS12016 and DS22016

| Parameter	DS1 2016	DS2 2016
# References	1166	4032
Total budget (€)	333,283 €	1,051,848
Periodicity	7	7
Min. days stock	7	7
Max. days stock	126/364	91
Real M.D.S.(€)	70,261	416,299
MORS M.D.S. (€)	67,869/71,030	488,188
M.M. M.D.S (€)	39,594	357,751
MORS S.L. (%)	98.7/98.6	97.5
M.M. S.L. (%)	95.1/95.1	95.1
MORS # unique buys	2641/2439	6265
M.M. # unique buys	8206	11,514
Reduction in # buys (%)	32/30	54
Reduction in O.F. (%)	50/56	71
Initial stock (€)	68,112	556,806
Real final stock (€)	119,450	461,057
MORS final stock (€)	75,340/78,991	479,628
M.M. final stock (€)	38,999	286,084
MORS lost sales (€)	4028/4226	38,946
M.M. lost sales (€)	14,706	94,910
MORS purchase vol. (€)	293,308/297,131	1,040,644
M.M. purchase vol. (€)	243,061	776,754
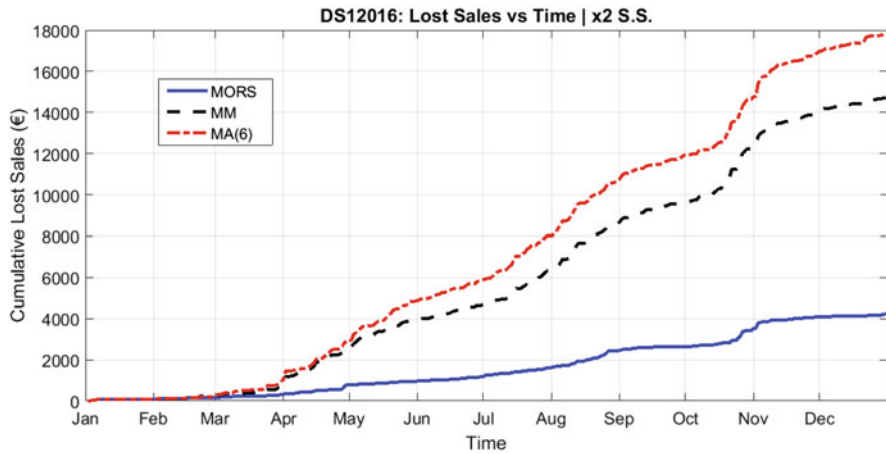
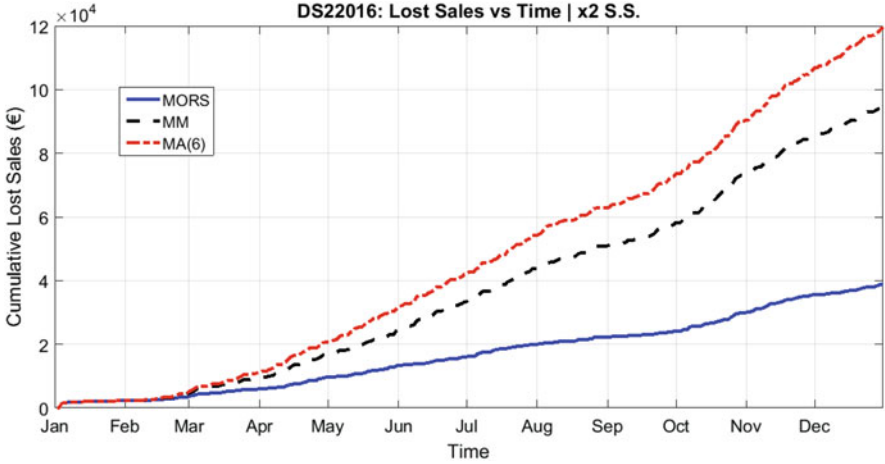**Fig. 11** DS12016: Lost Sales Monetary Value for three methods with $T_{max} = 364$ days

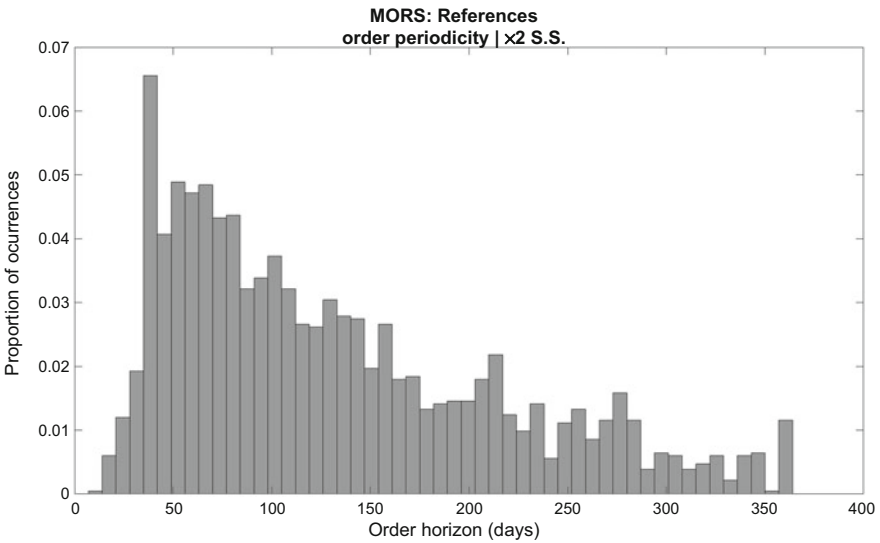**Fig. 12** DS22016: Lost Sales Monetary Value for three methods

Fig. 13 DS12016: Histogram of relative frequencies vs orders periodicity. Only ordered references were considered
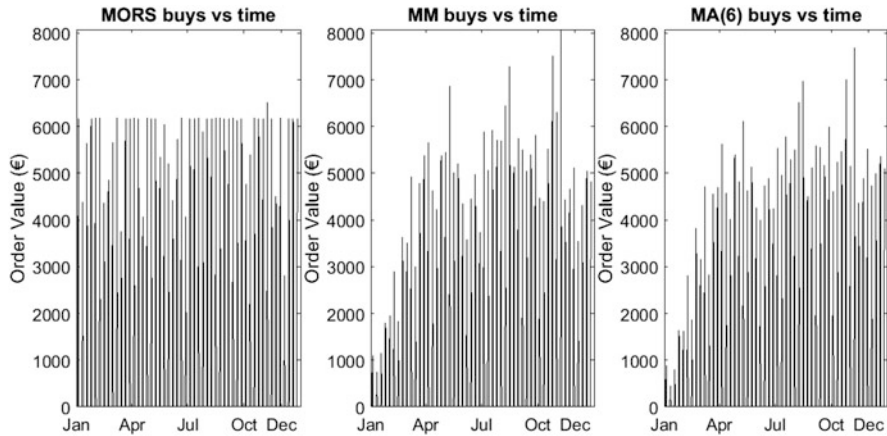
Fig. 14 DS12016: Buying amounts per order with $T_{max} = 364$ days

It is important to notice that although the MORS mean daily stock may increase as a direct consequence of the values used for T_{max}, the volume of buys is kept below the maximum allowed (that is, the real volume for 2016). Also, when compared with MM and with the original Nors model, the amount per order is made much more uniform across the year, as it may be seen in Fig. 14.

5 Conclusions

In this work we modelled, developed, and implemented a new methodology in Nors group for the spare parts forecasting, improving the accuracy in predictions. This was tested with two different suppliers with distinct orders of magnitude regarding their budget. Afterwards, those forecasts were used as input on a new optimization model to decrease the financial costs of the ongoing orders. The proposed methodology had two different outcomes when tested in those two distinct suppliers for the 2016 year. On the one hand, and regarding the forecasting accuracy, it increased the service level for more than 1% in both suppliers. On the other hand, applying the MORS optimization model with the same budget that the company used in 2016 for those suppliers, we found that: the service level increased in the range of 2.4–3.6%; the number of different reference purchases along the whole year was reduced between 30% and 54% which has a strong impact on the warehouse management and the stock level was kept smooth across the whole year. Finally, the objective function of the optimization model, that is, the indirect costs of the purchases (handling, interests, and warehouse renting) were reduced by a factor of between 0.5 and 0.71. As expected, the mean daily stock value was increased, but the costs of that raise were already contemplated in the objective function definition. All

of these were achieved in a model that is fully parametrizable by the management regarding factors like the maximum period to hold stock or the total budget amount, making it flexible to encompass any further changes in the management policy of the group. Despite the considerable gains achieved with this implementation, the model is being updated on the forecasting methods described on Sect. 2. In fact, the group is now using several other information available and new techniques of artificial intelligence, in order to obtain better forecasts for the references demands.

References

1. Bacchetti, A., Saccani, N.: Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. Omega **40**(6), 722–737 (2012)
2. Boylan, J.E., Syntetos, A.A.: Spare parts management: a review of forecasting research and extensions. IMA J. Manag. Math. **21**, 227–237 (2010)
3. Caiado, J.: Métodos de Previsão em Gestão. Edições Sílabo, Lisboa (2011)
4. Cohen, M.A., Agrawal, N., Agrawal, V.: Winning in the aftermarket. In: Harvard Business Review. https://hbr.org/2006/05/winning-in-the-aftermarket. Visited on 12 Feb 2017
5. Croston, J.D.: Forecasting and stock control for intermittent demands. Oper. Res. Q. **23**, 289–303 (1972)
6. Kennedy, W.J., Patterson, J.W., Fredendall, L.D.: An overview of recent literature on spare parts inventories. Int. J. Prod. Econ. **72**, 201–215 (2002)
7. Makridakis, S.G., Wheelwright S.C.: Forecasting Methods for Management. Wiley, London (1989)
8. Masoud, V., Manbir, S.: Application of game theory on inventory level decision making. Int. J. Bus. Econ. Res. **3(6)**, 211–219 (2014). https://doi.org/10.11648/j.ijber.20140306.12
9. Matsumoto, M., Ikeda, A.: Examination of demand forecasting by time series analysis for auto parts remanufacturing. J. Remanuf. **5**(1), 1–20 (2015)
10. Neves, G., Diallo, M., Lustosa, L.: Initial electronic spare parts sock and consumption forecasting. Investig. Oper. **28**, 45–58 (2008)
11. Rodrigues, M., Sirova, E.: Application of spare parts management methods in the companies in the Czech Republic. ACC J. **23**, 136–145 (2017)
12. Romeijnders, W., Teunter, R., Jaarsveld W.: A two-step method for forecasting spare parts demand using information on component repairs. Eur. J. Oper. Res., **220**, 386–393 (2012)
13. Suomala, P., Sievänen,M., Paranko, J.: The effects of customization on spare part business: a case study in the metal industry. Int. J. Prod. Econ. **79**(1), 57–66 (2002)
   %bibitemSyntetosBoylan
14. Syntetos, A.A., Boylan, J.E.: On the bias of intermittent demand estimates. Int. J. Prod. Econ. **71**, 457–466 (2001)
15. Teunter, R.H., Syntetos, A.A., Zied Babai, M.: Intermitent demand: linking forecasting to inventory obsolescence. Eur. J. Oper. Res. **214**, 606–615 (2011)
16. Zermati, P.: A Gestão de Stocks. Editorial Presença, Lisboa (2000)