Improved spare parts inventory management: a case study

L.W.G. Strijbosch\textsuperscript{1}, R.M.J. Heuts\textsuperscript{2} and E.H.M. van der Schoot\textsuperscript{3}

\textsuperscript{1}Faculty of Economics and Business Administration, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
\textsuperscript{2}CentER for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
\textsuperscript{3}CMG Trade, Transport & Industry, SAP Competence Center, P.O. Box 8566, 3009 AN, The Netherlands

\textit{Key words:} Forecasting, inventory, simulation, intermittent demand

\textbf{ABSTRACT}

This paper examines the performance of two different \((s, Q)\) inventory models for spare parts in a production plant of confectionery in The Netherlands, a simple and an advanced model. The simple approach is more or less standard: undershoot of the reorder level is not taken into account, the normal distribution is used as the distribution of demand during lead-time. The advanced one takes undershoots into account, differentiates between zero and nonzero demands during lead-time, and utilizes the gamma distribution as the demand distribution. Both models are fed with parameters estimated by a procedure that forecasts demand sizes and time between demand occurrences separately (intermittent demand). The results show that the advanced approach yields service levels close to the desired one under many circumstances, while the simple approach is not consistent leading to much larger inventories if one wants to be sure that the service level is obtained for all spare parts.

\textbf{INTRODUCTION}

The organization for which this research project was undertaken\textsuperscript{1}, MARS, owns a production plant in Veghel (The Netherlands). MARS produces confectionery for the European market with highly automated machines, 24 hours per day, 365 days per year. As a machine breakdown is very costly, a maintenance organization needs spare parts in order to maintain the equipment. The spare parts distribution center is divided into two sections: an open (for ‘cheep’ and ‘fast moving’ spare parts) and a closed (for ‘expensive’ and/or very critical parts: often seldom used) storage section.

The open storage section (created in 1991) yielded reduction of waiting times for mechanics and savings on total controlling costs. The mechanics can retrieve the spare parts without making any
registration. Two suppliers of spare parts nearby the distribution center perform a daily visual control of the inventory and take care of the replenishments within four working days using a \((s,Q)\)-system.

In the closed storage section mechanics ought to register the retrieval of parts. MARS itself controls the inventory position in that case. In this section also a kind of \((s,Q)\)-inventory control is used, whereby every transaction is registered. Just as in the open section, the reorder point and order size are based on intuition and experience regarding inventory turn over rates, product value and transportation/ordering costs. There is an endeavor to restrict the number of different distributors from 250 to 100. The lead-times from those distributors are quite different in length. Some variation in lead-time is possible; however, this mainly depends on the arrangements on deliverance. To reduce complexity we therefore assume fixed lead-times. After delivery of the service parts, they are checked and registered in the SAP/R3 system.

Demand for spare parts is very intermittent (cf. Figure 1): of all spare parts at MARS 40 percent had no usage during the years 1995-1997 and even 80 percent had no usage in the last half year. Only 1.15 \% spare parts had a mean usage of more than 0.15 per day.

![FIG.1: Relative frequency of daily usage during Sept.'97- Febr.'98 (Logarithmic scale for percentage)](image)

As the company is aware of the intuitive nature of the inventory control procedure, a study was undertaken to design and test a procedure which could outperform the existing one in service and/or costs. Management considers it important to have one well-defined procedure (especially for the closed section),
which could be used for all spare parts, independent of their demand characteristics. Given the information
lack regarding the performance of the existing policy and the intuitive nature of it, we firstly had to invent a
simple and formal control rule approximating closely or even outperforming the existing one. In this way it
would become possible to compare the simple approach adhered by the company with a more advanced
policy involving more computer time, a less intuitive decision rule, and formulas which are difficult to
comprehend by the employees concerned. Furthermore, with formally described inventory policies a
simulation study could be set up to investigate the performance of both (the existing and the proposed)
systems. It is clear that only substantial improvements upon the existing policy could convince
management to pass on to a newly designed more complex system.

The next section discusses the characteristics of the spare parts demand and some possible control
procedures. After introducing the necessary notation we formulate the used models precisely and describe
the main results of a Monte Carlo investigation. The last section contains the conclusions and suggestions
for further research.

MODEL SELECTION

In order to provide an impression of the characteristics of the demand patterns ten spare parts are selected
from the closed storage section. For these ten parts the mean time (days) between usage (interarrival time),
$E(A)$, the mean demand when the part is required, $E(D^*)$, and the demand variance, $Var(D^*)$ are
obtained empirically. Together with the fixed lead time, $L$, the parts value, $v$, and the ordering cost, these
figures are recorded in Table 1 illustrating the variability of the main characteristics which is representative
for many parts. Some parts are needed seldom (M8, M9), while the required number is 1 or 2 (M9), or
varies between 1 and for example 7 (M8); other parts are needed frequently like M0, M1, M2, M4 also with
strongly varying required numbers. Part M2, a very critical part in spite of the low price, is measured in
centimeters. The coefficient of variation of the interarrival time which is an important characteristic when
designing a new control procedure, appears to be close to one as can be expected from the nature of
demand (failures of machines). Due to the limited period for which historical information was available this
could not be checked for M8 and M9, unfortunately. However, we assume from now on that this coefficient of variation equals \( \sqrt{1-1/E(A)} \), the coefficient of variation of the underlying geometric variable.

Lead-time variability is sufficiently low (see introduction) to assume it to be deterministic. Table 1 demonstrates that demand is very intermittent, in general. A well-defined control policy has to account for these demand patterns. In consultation with the company a service criterion should be used to determine the decision parameters \( s \) and \( Q \). The main reason for this choice is the lack of cost information on the various cost components needed for a cost model. Of the commonly used service criteria the so-called \( P_2 \)-criterion (also called the fill rate) has been chosen, meaning that on the average during a delivery cycle, a shortage of \( 100(1-P_2) \) percent of the mean demand during an arbitrary cycle is allowed.

<table>
<thead>
<tr>
<th>Material</th>
<th>( E(A) )</th>
<th>( E(D^+) )</th>
<th>( Var(D^+) )</th>
<th>Lead-time ( (L) )</th>
<th>Value ( (v) )</th>
<th>Ordering cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>5.9</td>
<td>1.6</td>
<td>1.0</td>
<td>5</td>
<td>14.50</td>
<td>13.62</td>
</tr>
<tr>
<td>M1</td>
<td>7.0</td>
<td>21.0</td>
<td>447.6</td>
<td>5</td>
<td>11.83</td>
<td>13.62</td>
</tr>
<tr>
<td>M2</td>
<td>7.5</td>
<td>122.7</td>
<td>7709.6</td>
<td>14</td>
<td>1.18</td>
<td>38.62</td>
</tr>
<tr>
<td>M3</td>
<td>7.7</td>
<td>3.4</td>
<td>14.2</td>
<td>14</td>
<td>28.00</td>
<td>38.62</td>
</tr>
<tr>
<td>M4</td>
<td>24.0</td>
<td>19.4</td>
<td>48.2</td>
<td>30</td>
<td>58.00</td>
<td>38.62</td>
</tr>
<tr>
<td>M5</td>
<td>24.5</td>
<td>2.8</td>
<td>4.4</td>
<td>21</td>
<td>22.50</td>
<td>38.62</td>
</tr>
<tr>
<td>M6</td>
<td>24.7</td>
<td>7.0</td>
<td>8.0</td>
<td>30</td>
<td>20.02</td>
<td>13.62</td>
</tr>
<tr>
<td>M7</td>
<td>24.8</td>
<td>1.2</td>
<td>0.2</td>
<td>25</td>
<td>93.65</td>
<td>38.62</td>
</tr>
<tr>
<td>M8</td>
<td>119.0</td>
<td>2.5</td>
<td>4.5</td>
<td>31</td>
<td>725.00</td>
<td>38.62</td>
</tr>
<tr>
<td>M9</td>
<td>204.0</td>
<td>1.5</td>
<td>0.5</td>
<td>42</td>
<td>444.00</td>
<td>38.62</td>
</tr>
</tbody>
</table>

Table 1: Some observations on and characteristics of ten selected spare parts.

In the following we shall review some of the approaches found in the literature which could be considered for application in the MARS-case.

**Normal distribution of demand**

The reorder point calculation, for a \((s,Q)\)-system, based on the normal probability distribution states (see e.g. Silver et al.\(^2\)) that the reorder point \( s \) can be obtained from

\[
s = \hat{x}_L + k\hat{\sigma}_L
\]  

(1)
where the safety stock $k\hat{\sigma}_L$ is the product of the forecast error standard deviation of demand during the lead-time and the safety factor $k$, which very much depends on the desired type of service and the normality assumption. Some objections against the normality assumption are:

- non-unit demand can lead to an inventory level under $s$ when the order is placed;
- a large coefficient of variation leads to an unacceptable large chance of negative demands in the model;
- skewness is not taken into account.

However, the simplicity of the reorder point formula and its relatively easy implementation explains its widely usage in many inventory management systems and commercial software. This standard procedure is a useful candidate to approach the existing control procedure. It is simple to understand, applicable in many situations and -what is most important- closely related to what management considers appropriate. As management is inclined to use equal safety factors over different spare parts, individually determined safety factors should outperform the system adhered by the company. Then, showing that an advanced system is even better only requiring the implementation of some software and a little more computer time, better values for the decision parameters could be determined on a regular basis, yielding a closer approximation of the desired service level. In the section on model formulation the more or less standard method based on the normal distribution is presented.

*Poisson or compound Poisson distribution of demand*

Silver et al.² suggest the following: “If the average demand during lead time is below 10 units, then a discrete demand distribution such as the Poisson distribution is more appropriate”. However, when the mean is not approximately equal to the variance of the lead time demand, then a compounding process might be more appropriate. Among others the following models are found in literature:

(a) Friend³ uses a Poisson distribution for demand occurrence, combined with demands of constant size.

(b) Hadley and Whitin⁴ and Adelson⁵ use a stuttering Poisson distribution: a Poisson distribution for demand occurrence with a geometric distribution for the demand size.
Nahmias and Demmy\textsuperscript{6} use demand occurrences according to a Poisson process, logarithmic distributed demand sizes, and gamma distributed lead-times.

Care must be taken in choosing the right compounding, as the reorder point calculation might be seriously affected by false assumptions. For the case considered, (a) and (c) do not apply: we do not have constant demands and lead times are assumed deterministic. A stuttering Poisson distribution (b) is close to what is required. However, the values for $E(D^*)$ and $Var(D^*)$ in Table 1 suggest that a geometric distribution for the demands is not appropriate.

For slow moving expensive items and for the case of Poisson demand and various types of shortage costs, there are efficient methods\textsuperscript{2,7,8} for finding $s$ and $Q$ simultaneously. However, as stated earlier, cost models are difficult to implement in practice.

\textit{Compound Bernoulli distribution of demand}

Using a compound Poisson process assumes that empirical information is available of each individual demand and its demand size. In practice this is often not a very realistic starting point. Rather one has information on discrete time units, such as days or weeks, where demand is aggregated over time units. Also in the present case one works with demand per day information. Janssen et al.\textsuperscript{9} describe an $(R,s,Q)$ inventory model with a service level restriction, and where demand is modelled as a compound Bernoulli process. It is shown that this kind of modelling is especially suitable for intermittent demand. Further it is shown that incorporating the undershoot of the reorder level yields a substantial improvement on the attained performance levels when demand is intermittent. However, this procedure is only tested with known parameter values of the demand distribution. The described procedure seems a very good candidate for using it as the basis of a newly designed more advanced system for application in the present case.

\textit{Unknown demand parameters}

A problem of many methods, suggested in the literature, is the limited reported experience of applying these methods in practice. It is often unknown to what extent an a priori chosen performance level is attained. The performance of these methods is not known when using parameter estimates in stead of the true values of the parameters. Silver and Rahnama\textsuperscript{10} reported on a $(s,Q)$-model with a cost criterion, where underestimating the safety factor tends to lead to a higher cost penalty than equivalently
overestimating it. Therefore they suggest to bias the safety factor deliberately upwards so as to reduce the expected cost penalty associated with statistically estimating the reorder point. An important conclusion of this study should be that all methods, which are advocated in literature, should be tested, not with complete information on those parameters, but with an estimation procedure included. The common way to produce estimates of parameters of demand distributions is by forecasting. Especially exponential smoothing is applied, for its ability to incorporate non-stationary behaviour of demand. Table 1 suggests that simple exponential smoothing is not appropriate. Croston\textsuperscript{11} showed that when demand is intermittent, which is obviously the case for spare parts, the forecast error can be reduced by smoothing the time between demands and the demand sizes separately, and using these for the forecasting of lead-time demand. Many authors after 1972 showed the relevance of this Croston method.

**NOTATION**

The notation to be used includes

$L$ = the deterministic lead-time in days

$D_t$ = demand on day $t$

$Z_t(L) = \sum_{r=1}^{L} D_{t+r}^{\tau} \,$, total demand during the days $t+1,\ldots, t+L$

$\hat{D}_t(\tau) = \text{forecast of demand on day } t \text{ for day } t+\tau$

$\hat{Z}_t(L) = \sum_{r=1}^{L} \hat{D}_t(\tau) \,$, forecast of demand during the days $t+1,\ldots, t+L$

$e_t(\tau) = D_{t+\tau} - \hat{D}_t(\tau) \,$, forecast error for $D_{t+\tau}$

$E_t(L) = \sum_{r=1}^{L} e_t(\tau) \,$, forecast error for $Z_t(L)$

$A_z$ = the $z$-th interarrival time between demands

$D_z^*$ = the $z$-th demand size (always positive)

$\sigma^2 = \text{Var}(D_z^*)$

$a = \text{E}(D_z^*)$

$p$ = the probability of a positive demand on a day
MODEL FORMULATION

This section will be split up into three subsections; one on the forecasting approach, the next on the simple inventory model which will be used to simulate the company’s preferred model and the last on the advanced model which should outperform the simple model. Both models apply the forecasting method explained in the next subsection.

*Forecasting the intermittent demand structure of the spare parts*

For the demand of spare parts which can be characterized by one (or more) days of usage and a large number of days of non-usage, Croston\(^{11}\) introduced a point forecasting procedure, which separates interarrival times of non-usage and order sizes at a demand occurrence. Johnston and Boylan\(^{12}\) and Willemain *et al.*\(^{13}\) show that this idea of Croston is superior to applying simple exponential smoothing to all the demand data, zero demand or not. In the \((s, Q)\) inventory models described in the next two subsections it is of interest to calculate forecasts, not for a single period, but rather the sum of forecasts for the next \(L\) periods, the length of the lead-time. Then the point forecast of this aggregated or cumulative
demand is given by \( \hat{Z}_t(L) \) with forecast error \( E_t(L) \). For safety stock determination, we are interested in \( \text{Var}(E_t(L)) \). The model assumptions for the process \( \{D^*_z\} \) are
\[
D^*_z = a + \varepsilon_z \quad \text{where} \quad E(\varepsilon_z) = 0, \quad \text{Var}(\varepsilon_z) = \sigma^2 \quad \text{and} \quad \varepsilon_z \text{ are i.i.d.} \quad (2)
\]
Since a positive demand on day \( t \) occurs with probability \( p \), the distribution of \( A_z \) is:
\[
P(A_z = k) = (1 - p)^{k-1} p, \quad k \geq 1 \quad (3)
\]
with
\[
E(A_z) = 1/p, \quad \text{Var}(A_z) = (1 - p)/p^2 \quad (4)
\]
In the sequel we assume that the processes \( \{D^*_z\} \) and \( \{A_z\} \) are independent. Using a single exponential smoothing procedure for these processes, gives
\[
\hat{D}^*_z(1) = \alpha D^*_z + (1 - \alpha)\hat{D}^*_{z-1}(1) \quad (5)
\]
\[
\hat{A}_z(1) = \beta A_z + (1 - \beta)\hat{A}_{z-1}(1) \quad (6)
\]
The smoothing parameters commonly should satisfy \( 0 < \alpha \leq 0.3 \) and \( 0 < \beta \leq 0.3 \). Estimates of the standard deviation of the forecast error of \( D^*_z \) are obtained by smoothing the mean absolute deviation of forecast error (and multiply it by 1.25):
\[
MAD^*_z = \omega |D^*_z - \hat{D}^*_{z-1}| + (1-\omega)\text{MAD}^*_{z-1} \quad (7)
\]
The point forecast of demand on day \( t \) for day \( t + \tau \) is
\[
\hat{D}_t(\tau) = \frac{\hat{D}^*_z(1)}{\hat{A}_z(1)}, \quad \tau = 1, 2, \ldots \quad (8)
\]
Note that indices \( z \) and \( t \) are related in an obvious way: \( D^*_z \) is the last observed demand previous to day \( t \).

The point forecast for the aggregated demand during the lead time \( L \), is
\[
\hat{Z}_t(L) = \sum_{\tau=1}^{L} \hat{D}_t(\tau) = L \hat{D}^*_z(1)/\hat{A}_z(1) \quad (9)
\]
The variance of the sum of the forecast errors during \( L \), is
\[
\text{Var}(E_t(L)) = \text{Var}\left[\sum_{\tau=1}^{L} D_{t+\tau} - \sum_{\tau=1}^{L} \hat{D}_t(\tau)\right] = \text{Var}\left[\sum_{\tau=1}^{L} D_{t+\tau}\right] + \text{Var}\left[\sum_{\tau=1}^{L} \hat{D}_t(\tau)\right] \quad (10)
\]
The second equal sign is allowed since the realizations of demand after \( t \) and its forecasts are independent of each other. For the expectations and variances of the forecasts we obtain

\[
E\{\hat{D}_z(1)\} = E\{\alpha \sum_{k=0}^{\infty} (1-\alpha)^k D_{z-k}^*\} = a
\]

(11)

\[
E\{\hat{A}_z(1)\} = E\{\beta \sum_{k=0}^{\infty} (1-\beta)^k A_{z-k}\} = \frac{1}{p}
\]

(12)

\[
Var\{\hat{D}_z(1)\} = Var\{\alpha \sum_{k=0}^{\infty} (1-\alpha)^k D_{z-k}^*\} = \frac{\alpha}{2-\alpha} \sigma^2
\]

(13)

\[
Var\{\hat{A}_z(1)\} = Var\{\beta \sum_{k=0}^{\infty} (1-\beta)^k A_{z-k}\} = \frac{\beta}{2-\beta} \frac{1-p}{p^2}
\]

(14)

These approximations hold when enough historical information is available and the demand data is stationary. Using the following expressions for the variance of a product and the quotient of two independent stochastic variables \( x \) and \( y \):

\[
Var(xy) = \mu_x^2 \sigma_y^2 + \mu_y^2 \sigma_x^2 + \sigma_x^2 \sigma_y^2
\]

(15)

\[
Var(x/y) = \frac{\mu_x^2}{\mu_y^2} (\sigma_x^2/\mu_x^2 + \sigma_y^2/\mu_y^2)
\]

(16)

for fixed integer \( L \) the next expressions follow:

\[
Var\{\hat{D}_z(L)\} = L^2 Var\{\hat{D}_z(1)/\hat{A}_z(1)\} = (pL)^2 \left( \frac{\alpha}{2-\alpha} \sigma^2 + \frac{\beta}{2-\beta} (1-p)a^2 \right)
\]

(17)

\[
Var\{\sum_{t=0}^{L} D_{z+\tau}\} = L Var(D_t) = L \left\{ p \sigma^2 + a^2 p(1-p) \right\}
\]

(18)

Combining (17) and (18) yields:

\[
Var\{E_z(L)\} = pL \left\{ pL \left( \frac{\alpha}{2-\alpha} \sigma^2 + \frac{\beta}{2-\beta} (1-p)a^2 \right) + \sigma^2 + a^2 (1-p) \right\}
\]

(19)

In this expression, which represents the variance of the sum of the forecast errors during the lead-time, the unknown parameters can be estimated as follows:

\[
\hat{p} = \frac{1}{\hat{A}_z(1)}, \quad \hat{a} = \hat{D}_z^*(1), \quad \hat{\sigma} = 1.25 \text{MAD}_z \sqrt{(2-\alpha)/2}
\]

(20)

yielding \( \hat{\text{Var}}(E_z(L)) \) as an estimate of \( \text{Var}(E_z(L)) \). The estimate for \( \sigma \) is obtained according to a well-known relationship\(^{15}\). For the parameter setting of \( \alpha \), \( \beta \) and \( \omega \), we refer to the next section.
Under the regime of an \((s, Q)\) inventory model, the inventory position is monitored every day. When the inventory position is below \(s\), an amount \(Q\) is ordered such that the inventory position is raised to a value between \(s\) and \(Q\). The daily monitoring was considered to be important by the MARS management, as delays in deliverance of spare parts can cause costly interruptions of the production process. The next subsection describes in detail how the simple method can be implemented in practice when the demand parameters are unknown.

**The simple \((s, Q)\) inventory model**

When applying the simple method \((1)\), we need the estimates \(\hat{x}_L\) and \(\hat{\sigma}_L\). These can be obtained with the aforementioned forecasting procedure:

\[
\hat{x}_L = \hat{Z}_L(L)
\]

\[
\hat{\sigma}_L^2 = \text{Var}(E_i(L))
\]

The safety factor \(k\) can be obtained by solving the corresponding service equation \(2\)

\[
G_u(k) = \frac{Q}{\sigma_L}(1 - P_2)
\]

where \(G_u(k)\) is defined as

\[
G_u(k) = \int_k^\infty (u - k)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)du.
\]

**The advanced \((s, Q)\) inventory model for spare parts**

The inventory model described here is based on Janssen et al. \(^9\) and combines their compound Bernoulli model (CBM) with the aforementioned forecasting procedure. Let’s first start with some important modelling assumptions, some of which are based on the MARS situation ((a), (c), (e)):

(a) lead-time will be considered as fixed and integer;

(b) both the demand during lead-time (if positive) and the undershoot of the reorder levels are distributed as mixed Erlang distributions;

(c) the reorder point \(s\) is positive;

(d) daily demand is i.i.d. and considered as a continuous variable;

(e) as service criterion a so called \(P_2\)-service level is used with backordering possibility;
(f) the a priori fixed order quantity $Q$ is sufficiently large compared with the forecasted demand during lead time plus forecasted undershoot at period $t$.

In the sequel we make use of the same notation as in the previous subsection, however we omit the indices.

The relation between the distribution functions $F_D(\cdot)$ and $F_D^\ast(\cdot)$ is\textsuperscript{9}

$$F_D(x) = \begin{cases} 1 - p, & \text{if } x = 0 \\ 1 - p + p F_D^\ast(x), & \text{if } x > 0 \end{cases}$$  \hspace{1cm} (24)

The moments of $D$ and $D^\ast$ are related according to:

$$E(D) = pa$$  \hspace{1cm} (25)

$$Var(D) = p \sigma^2 + a^2 p (1 - p)$$  \hspace{1cm} (26)

The first two moments of $D^\ast$ (viz. $a$ and $\sigma^2$) and the parameter $p$ can be estimated according to (20).

Using assumption (a) and (d), the demand during the lead time, $Z(L)$ is an $L$-fold convolution, with expectation and variance

$$E(Z(L)) = L E(D)$$  \hspace{1cm} (27)

$$Var(Z(L)) = L Var(D)$$  \hspace{1cm} (28)

The probability of a positive demand during the lead-time, is equal to

$$p_L = 1 - (1 - p)^L$$  \hspace{1cm} (29)

The relation between the distribution functions $F_{Z(L)}(\cdot)$ and $F_{Z^\ast(L)}(\cdot)$ is

$$F_{Z(L)}(x) = \begin{cases} 1 - p_L, & \text{if } x = 0 \\ 1 - p_L + p_L F_{Z^\ast(L)}, & \text{if } x > 0 \end{cases}$$  \hspace{1cm} (30)

with the following moment relation between $Z^\ast(L)$ and $Z(L)$:

$$E(Z^\ast(L)) = E(Z(L))/p_L$$  \hspace{1cm} (31)

$$Var(Z^\ast(L)) = Var(Z(L))/p_L - (1 - p_L)(E(Z(L)))^2 / p_L^2$$  \hspace{1cm} (32)

Janssen et al.\textsuperscript{9} show that the undershoot effect is very important and should be taken into account. As daily demand is assumed to be a continuous variable and $Q$ sufficiently large, then the first two moments of the undershoot $U$ can be approximated\textsuperscript{16} by
\[ E(U) = E(D^2)/(2E(D)) = E(D^2)/(2E(D)) \]  
\[ E(U^2) = E(D^3)/(3E(D)) = E(D^3)/(3E(D)) \]

Thus the variance of \( U \) is

\[ \text{Var}(U) = E(U^2) - (E(U))^2 \]

From experience we observe right skewness in the process \( \{D^*\} \). If we assume that \( D^* \) can be approximated by a gamma distribution, then

\[ E(D^{*2}) = (1 + c_{D^*}^2)(1 + 2c_{D^*}^2)a^3 \]

In this way \( D^* \) is only a function of its first two moments, which can be estimated by the forecasting procedure. Based on the right skewness of \( \{D^*\} \), we may expect that the processes \( \{Z^*(L)\} \) and \( \{U\} \) are right skewed too. The first two moments of \( Z^* = Z^*(L) + U \) can be obtained via the first two moments of \( Z^*(L) \) and \( U \), assuming that these are independent. Approximating \( Z^*(L) \) and \( U \) by mixtures of Erlang distributions has two advantages:

(a) the property of right skewness;
(b) numerical attractiveness.\(^{16}\)

The \( P_2 \)-service equation looks as follows now.\(^9\)

\[ 1 - P_2 = \left\{ p_L (E(Z^*-s)^+ - E(Z^*-s-Q)^+) + (1 - p_L)(E(U-s)^+ - E(U-s-Q)^+) \right\}/Q \]  

In accordance to assumption (f) \( Q \) will be fixed as follows:

\[ Q = EOQ \quad \text{if} \quad EOQ > 1.5\hat{Z}^*(L) \]
\[ = 1.5\hat{Z}^*(L), \quad \text{elsewhere} \]  

Here \( \hat{Z}^*(L) \) is the forecasted demand during lead-time. For given value of \( Q \) and the target service level \( P_2 \), the above service equation can be solved for \( s \) by a local search procedure. The expressions of the type \( E(\cdot)^+ \) are evaluated by using mixed Erlang distributions, which are fitted on the basis of the forecasted moments of the corresponding stochastic variables.

It can be shown \(^9\) that the expected average physical stock is equal to
\[ \mu_1(s, Q) = p_L \frac{E((s + Q - Z^*(L))^2 - E((s - Z^*(L))^2)}{2Q} + (1 - p_L) \frac{(s + Q)^2 + s^2}{2Q} \]

Using a mixed Erlang distribution for \( Z^*(L) \) this expression can also be easily evaluated. We observed in all cases we analyzed, that, for high \( P_2 \)-service and positive reorder point, the complex formula above could be safely approximated by the following simple formula

\[ \mu_2(s, Q) = s + Q / 2 - E Z(L) \]

(Note that (40) is different from the simple formula mentioned by Janssen et al.\(^9\) by the expectation of the undershoot).

**SIMULATION STUDY**

In order to get an idea of the performance of the methods developed in the previous section, a simulation study has been performed. The object of the study is to compare the simple and the advanced approach in a setting, which is relevant for the company. That is, using those regimes, we will investigate by simulation both the attained \( P_2 \)-service and average inventory levels thereby simulating data which resemble the company’s usage data for various spare parts (cf. Table 1). An important aspect of the simulation study is the use of the forecasting procedures to estimate the demand parameters. Janssen et al.\(^9\) tested the CBM method with known demand parameters. However, using estimated parameter values will influence the performance of the system\(^10,17\). This section proceeds as follows. Firstly, we describe the design of the simulation program. Next the results of the simulations will be presented and discussed.

*Simulation design*

The simulation program is set up according to the ‘next event’ principle\(^18\). This principle means that after taking care of all events on a certain day, the programs continues to the day of the following event. The events, which may occur, are:

- A delivery by the supplier, resulting in an increase of the inventory position.
- A usage of the spare part, leading to an update of the forecasts of mean time between demands, mean demand size, and corresponding mean absolute deviation of forecast error.
- The inventory position gets under the reorder level, which evokes an order.
Demand sizes are generated according to a mixed Erlang distribution, whereas the time between demand occurrences is generated according to a geometric distributed variable +1 (the interarrival time is at least 1). Using a run-in period of 100 demands measurements of performance are obtained after 100,000 demand occurrences. The estimations of the demand parameters are updated after each period of 90 days according to the actual values from the forecast procedure. As is common practice a small value for $\omega$ is chosen ($\omega = 0.025$) The nature of demand may normally considered to be reasonable stationary. Replacing parts by better alternatives cause dramatic changes in the demand pattern, for example. These causes cannot be accounted for by forecasting. The events, which influence the demand pattern strongly, should be accounted for by the ‘management-by-exception’ principle. Thus, relatively small values for $\alpha$ and $\beta$ are chosen ($\alpha = \beta = 0.05$). The most important assumptions that are being used throughout the various simulations are:

- Non-integer values for the reorder point and the order size generated by the simulation program are being rounded upwards.
- Demand sizes generated by the program are being rounded to integers with a minimum of one.

The main output of the simulation program is as follows:

- Attained $P_2$-service level, $SL$, defined by

$$SL = 1 - \frac{\text{mean shortage per cycle}}{\text{mean demand per cycle}}$$

where \text{mean shortage per cycle} is defined as the accumulated shortage at the end of all replenishment cycles during the simulation horizon divided by the number of cycles, and \text{mean demand per cycle} is defined by aggregated demand over the simulation horizon divided by the number of cycles. A correction for a possible shortage at the beginning of the replenishment cycle seems not necessary as the service levels used in the simulation and desired by the company are very high (0.95 and 0.99).

- Average (physical) inventory level $\mu(s, Q)$ as a result of the simulation.
- Average inventory level according to formulae (39) and (40).
Simulation results

The simple inventory control procedure will be denoted by STM (standard method), the advanced one by CBM (compound Bernoulli method). Table 1 has been used to set constant and varying values for the parameters of the control procedures in the simulation. The simulation yields the performances (realized service level, and average inventory level as measured during the simulation) of both procedures for each of the two desired service levels (0.95 and 0.99) as a function of one varying key variable. The constant (varying) values of the parameters are set as follows: \( E(A) = 25 \{5,10,15,20,25,50,75,100,150,200\} \), \( E(D^*) = 3 \), \( Var(D^*) = 9 \{0,2,4,6,8,10,15,20\} \), \( E(L) = 20 \{5,10,20,30,40,50\} \), \( c_A = \sqrt{1-1/E(A)} \{0.4,0.6,0.8,1.0,1.2,1.4,1.6\} \). When, as is the case in the underlying CBM, the interarrival time has a geometric distribution with minimal value 1, the corresponding coefficient of variation \( c_A \) equals \( \sqrt{1-1/E(A)} \), which is close to 1 in many cases. As there is some doubt on the validity of the assumption of the geometric distribution, we have performed a simulation experiment where the interarrival time is generated by a gamma distribution with varying \( c_A \).

Figures 2 and 3 show the simulation results for varying lead-time. As lead-time increases the performance of CBM is consistent. However, the attained service level is 1-2% under the desired one, which is mainly due to the fact that the distribution parameters are unknown and substituted by estimated values. A better approach of the desired service level could simply be obtained in practice by upgrading the \( P_2 \)-service a little. On the contrary, STM yields a service level that is much lower in the first place and is not consistent. The increasing service level for increasing lead-times is due to the fact that the distribution of demand during lead-time will be closer to the normal distribution as lead-time increases. There is no obvious way to correct STM such that the attained service level would be closer to the desired level. Figure 3 shows the consequences for the mean inventory of increasing lead-times. The combinations CBM-0.95 and STM-0.99 at \( L = 40 \) in both figures show that STM needs a little bit more inventory to obtain the same service level as CBM.

Figures 4 and 5 show the results for varying mean interarrival time. Again, as should be, CBM produces a consistent level of the attained service. However, STM turns out to be extremely sensitive to
Increasing $L$ (Figures 2 and 3) has a similar effect as decreasing $E(A)$. The explanation of the behaviour of STM in Figure 4 is comparable with the explanation that is given for Figure 2.

Figures 6 and 7 show the results for varying variance of demand size. The decrease of the attained service level with CBM and increasing $\text{Var}(D^*)$ is mainly due to the fact that the factor $1.25$ used in order to obtain an estimate of the forecast error standard deviation, should be larger\textsuperscript{19}. With STM, again the normality assumption is less appropriate with increasing $\text{Var}(D^*)$, yielding a decreasing service level. Figures 8 and 9 show the extent of robustness of CBM against violating the assumption that interarrival times are geometrically distributed. Obviously, the inventory decisions do not change much with varying $c_A$, yielding approximately the same mean inventory levels. However, the attained service levels diminish, which is mainly due to the fact that $A_2$ and $D_2$ are not i.i.d. (as is used for both STM and CBM, cf. (14) and (18)) when $c_A$ deviates from the corresponding value of a geometric variable\textsuperscript{9}.

The differences between the formulas $\mu_1(s,Q)$ and $\mu_2(s,Q)$ for the average physical inventory are illustrated in Figure 10. For all different simulation situations (62 in number) realizations of (39) and (40) are evaluated. Both formulas apparently give nearly the same results. Only when the coefficient of variation of $A$ is large there is a slight difference. Formula (40) is preferred for its simplicity. Figure 11 shows the results of (40) in relation with the realization of the average physical inventory in the 62 simulated situations. The average inventory is slightly overestimated by formula (40) unless the coefficient of variation is large. It turns out that formula (40) is a very useful formula to obtain an estimate of the average inventory in the described situations.

**CONCLUSIONS**

As spare parts control was considered important by the management of MARS, they wondered whether going from an intuitive control to a more consistent complex combined forecasting/replenishment approach, as indicated in this paper, would pay off. The results of the new approach were such that they decided to implement it in the near future.
LITERATURE


FIG. 2: Attained service level versus the delivery time

FIG. 3: Average physical inventory versus the delivery time.
FIG. 4: Attained service level versus the mean interarrival time.

FIG. 5: Average physical inventory versus the mean interarrival time.
FIG. 6: Attained service level versus the variance of $D^*$

FIG. 7: Average physical inventory versus the variance of $D^*$. 
FIG. 8: Attained service level versus the coefficient of variation of the interarrival time.

FIG. 9: Average physical inventory versus the coefficient of variation of the interarrival time. The coefficient of variation of the interarrival time.
FIG. 10: Comparison of formulae (39) and (40).

FIG. 11: Realized physical inventory compared with formula (40).