Project 4

Multi-Echelon Control Systems of TPTS spare part network

Alexander Larsson
Erik Lundin
Jakob Bengtsson
Fabian Gunnarson

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Purpose

TPTS is a company that sells spare parts. The distribution network of TPTS can be described as a two-echelon inventory system, in which each item is held in stock by one central warehouse and multiple retailers. These stock points are controlled by separate (R, Q) policies. While the order quantities Q are fixed, TPTS wants help developing a multi-echelon based method for finding near-optimal reorder points R for their items at the different stock locations. In other words, the system cost should be kept as low as possible while also ensuring that service level goals are met. The long-term goal is to implement this method for the entire inventory system of TPTS, however as a first step, the method will be tested on five items with a few stock locations each.

1.2 Data

TPTS has provided data regarding the test items, which will be used as input to develop an analytic optimization method and to conduct the analysis. The types of data that were provided and utilized in the analysis are shown in Table 1.

Table 1: The input data for the analysis.

Q_0	*Batch quantity for orders from warehouse*
Q_i	*Batch quantity for orders from retailer i*
R_0	*Reorder point at warehouse*
R_i	*Reorder point at retailer i*
L_0	*Transportation time for orders from warehouse*
l_i	*Transportation time for orders from retailer i*
$S_{2,i}^{target}$	*Target fill rate for retailer i*
μ_i*	*Mean daily customer demand at retailer i*
σ_i*	*Standard deviation of customer demand at retailer i*
$f_{d,i}$*	*Probability that a customer at retailer i demands d units*

**Empirically observed number*

The numerical values of the input data presented in Table 1 can be seen in the attached file "MION01_assignment4_group2". Moreover, there is also a simulation model of the inventory system at hand, which will be used to evaluate the performance of the analytic model.

1.3 Report structure

The report structure is as follows. First, in section 2 *Method*, a description is given of how the multi-echelon based method was developed. In section 3 *Results*, the solutions that the method generated for the items provided by TPTS are presented and compared to the current solution. Section 4 *Analysis* discusses the performance of the obtained solutions, the model limitations and drawbacks as well as the challenges and opportunities for scaling the model. Lastly, section 5 *Conclusion* attempts to consolidate the findings and provide final remarks. The tools used in the analysis are MatLab and Microsoft Excel. The programming solutions are presented in the files in the attached folder "Assignment4_group2_MatLab".

2 METHOD

2.1 General procedure

Entering this task, it was discussed what would be a reasonable procedure. The so-called METRIC approach has proved to be rather competent in dealing with multi-echelon systems (Axsäter 2006, p. 261) – and so, pursuing a METRIC-type approach served as a starting point for the project. At its core, the METRIC approach entails an approximation of the lead time using an average delay at the warehouse, as seen in equation (1). The transportation time is assumed to be deterministic while the waiting time at the warehouse is treated as stochastic. The lead time approximation, $\overline{L_i}$, enables a breakdown of the multi-echelon problem into single-echelon optimization at each retailer with the approximated lead times being treated as deterministic. As the connection is maintained between reorder points at both the warehouse and retailers, this allows for a joint optimization of the entire system.

$$\overline{L_i} = l_i + E[W_0] \tag{1}$$

$l_i = transportation\ time\ from\ the\ warehouse\ to\ retailer\ i$

$E[W_0] = average\ waiting\ time\ at\ the\ warehouse$

However, when assessing the project data, several nuisances arose. Firstly, it was observed that most retailers were non-identical, with different batch-order quantities and shipment times from the central warehouse. As a result, the average delays for deliveries from warehouse to retailer are likely not equal, something that can be cumbersome to deal with. Here, as is proposed by Axsäter (2006, p. 265), it was reasoned that the approximation of assuming the same average delay $E[W_0]$ for all retailers would be sufficiently accurate in order to describe the system. How such an average was computed will be presented in section 2.3 *Modelling warehouse behaviour*. The method assumes that demand at the retailers and at the warehouse can be accurately approximated by a normal distribution. One important precondition for this procedure is that the customer demand at the retailers follows a compound Poisson process – entailing exponentially distributed interarrival times between customers, as well as considering demanded quantities to be discreetly stochastic entities. Although the actual customer data that was provided by TPTS was not very detailed, revealing nothing about customer arrival times, this assumption was made.

The coefficients of variation for the different retailer demands were all >1, which according to Axsäter (2006, p. 80) implies that this assumption is reasonable.

While implementing the model, various tests were performed, comparing the results regarding fill rates, stock-on-hand levels and average waiting times produced by the analytic model, to what was produced by the simulation model. This approach was helpful both in terms of validating the accuracy of the approximations, as well as discovering coding and mathematical errors in the implementation. The assumption of using a compound Poisson distribution to describe customer demand at retailers was challenged by setting the waiting time equal to zero in the analytic model and observing the metrics at the retailers, see Appendix A. In the simulation model, such a system was produced by assigning the warehouse a very high reorder point, which practically removed the waiting time. This way, potential errors in the waiting time estimations were removed from the equation. Furthermore, the complete model, when including warehouse delays, was validated by using the current set of reorder points. It was found that the analytic model and the simulation model produced a similar set of values, see Appendix B, with the exception of some larger deviations of the fill rates regarding item 4. It was however decided that the accuracy was satisfactory enough to proceed with optimizing the system when modeled under the stated assumptions.

2.2 Calculating the cost of the system

The goal of the reorder point optimization was to minimize the total cost of the system, while still meeting the target fill rates $S_{2,i}$ of each retailer. It was assumed that there was no dependency between the optimal set of reorder points for the different items. Consequently, the costs could be minimized for each item separately. The total costs for a single item could be calculated as in equation (2) (Axsäter 2006, p. 271-272). In the special case where the holding costs are equal regardless of the stock location, the problem boils down to minimizing the stock-on-hand level.

$$C = h_0 E[IL_0^+] + \sum_I h_i E[IL_i^+] \tag{2}$$

C = cost

h_0 =*holding cost at warehouse*

h_i =*holding cost at retailer i*

$E[IL_0^+]$ =*average stock-on-hand level at warehouse*

$E[IL_i^+]$ =*average stock-on-hand level at retailer i*

2.3 Modelling retailer behaviour

The average stock-on-hand for the retailers, $E[IL_i^+]$, was calculated according to equation (3) (Axsäter 2006, p. 264), while the retailer fill rate, $S_{2,i}$, was determined using equation (4) (Axsäter 2006, p. 98).

$$E[IL_i^+] = \sum_{j=1}^{R_i+Q_i} j \cdot P(IL_i = j) \tag{3}$$

$R_i = reorder\ point\ for\ retailer\ i$

$Q_i = order\ quantity\ for\ retailer\ i$

$$S_{2,i} = \frac{\sum_{d=1}^{x_i} \sum_{j=1}^{R_i+Q_i} min(j,d) \cdot f_d \cdot P(IL_i = j)}{\sum_{d=1}^{x_i} d \cdot f_d} \tag{4}$$

$f_d = probability\ that\ one\ customer\ demands\ d\ units$

$x_i = the\ maximum\ amount\ of\ units\ demanded\ by\ a\ single\ customer\ at\ retailer\ i$

In order to determine $E[IL_i^+]$ and $S_{2,i}$, $P(IL_i = j)$ first needed to be calculated for all relevant values of j, which was done using equation (5) (Axsäter 2006, p. 91). As only the probability for positive inventory levels was needed, the probabilities were calculated for $1 \le j \le R_i + Q_i$. As previously mentioned, the lead time demand at retailers was assumed to be compound Poisson distributed, with empirical compounding distribution as provided in the data material. Consequently, the probability of lead time demand, under the METRIC approximation of the lead time, could be calculated using equation (6) (Axsäter 2006, p. 79). The sum is bounded by the truncation of arriving customers k, elaborated upon below, which implicitly assumes $P(K = k) = 0$for k-values above a certain upper limit.

$$P(IL = j \mid R = r) = \frac{1}{Q} \sum_{k=max(r+1,j)}^{r+Q} P(D(\overline{L_i}) = k - j) \tag{5}$$

$$P(D(\overline{L_i}) = d) = \sum_{k=0}^{\infty} P(K = k) \cdot f_d^k \tag{6}$$

The compounding distribution f_d^k, that k customers demand a total number of d units, was determined for all combinations of d and k. Furthermore, to calculate $P(K = k)$ according to the probability function for the Poisson distribution, found in equation (7) (Axsäter 2006, p. 78), the customer arrival rate at the retailer, λ_i, was required, and was determined from equation (8) (Axsäter 2006, p. 80). The average amount of units purchased by each customer, $E[J]$, was computed using the definition of expected value. The distribution of units purchased per customer, J, (earlier denoted f_d) for each retailer was provided in the data. Lastly the range for k needed to be decided upon. Theoretically, the Poisson distribution allows for all positive integers, but as it is impossible to calculate the probabilities for an infinite amount of k-values, an approximation was needed. The probabilities were calculated for increasing values of k, starting from zero, until the total probability mass was satisfyingly close to 1 so that equation (9) is satisfied for some tolerance threshold, ϵ. The tolerance threshold for this truncation was set to $\alpha = 10^{-6}$.

$$P(K = k) = \frac{(\lambda_i \bar{L}_i)^k}{k!} e^{-\lambda_i \bar{L}_i} \tag{7}$$

$$\mu_i = E[K]E[J] = \lambda_i E[J] \Leftrightarrow \frac{\mu_i}{E[J]} = \lambda_i \tag{8}$$

$$1 - \sum_{k=0}^{y_i} P(K = k) < \alpha \tag{9}$$

$y_i = smallest\ value\ satisfying\ the\ inequality\ in\ equation$ 9

2.4 Modelling warehouse behaviour

One METRIC type approach that can be used to approximate the average warehouse delay, $E[W_0]$, and the average stock-on-hand level, $E[IL_0^+]$, for a two-echelon distribution system with (R, Q) ordering policies and stochastic demand, is described in Axsäter (2003). The general idea with the approach is to arrive at $E[W_0]$and $E[IL_0^+]$by approximating both the retailer demand and the warehouse demand with a normal distribution. This approach seemed particularly suitable for the inventory system under scrutiny in this report, as Axsäter argues that: (i) the technique has potential to optimize quite large systems in an efficient way; (ii) numerical tests indicate that the optimal reorder points are determined very accurately; and (iii) it builds on the assumption that the retailers face a discrete compound Poisson demand, which, as mentioned earlier in section 2.1, was observed to accurately describe demand at the TPTS retailers. Hence, the method was deemed as applicable.

The first steps in the approach was to determine the mean and the standard deviation of the lead time demand, $\mu^{w}(L_0)$ and $\sigma^{w}(L_0)$. The former was obtained through the expression shown in equation (10) (Axsäter, 2003).

$$\mu^{w}(L_0) = \sum_{i=1}^{n} \mu_i^{w}(L_0) = \sum_{i=1}^{n} \mu_i(L_0) = L_0 \sum_{i=1}^{n} \mu_i \quad (10)$$

$\mu_i^{w}(L_0) = average\ demand\ from\ retailer\ i\ at\ the\ warehouse\ during\ the\ warehouse\ lea$

$\mu_i(L_0) = average\ customer\ demand\ at\ retailer\ i\ during\ the\ warehouse\ lead\ time$

$n = number\ of\ retailers$

$\sigma^{w}(L_0)$ was more cumbersome to obtain. First, the probability for k orders at retailer i during the warehouse lead time, $p_{i,k}(L_0)$, needed to be approximated, see equation (11) (Axsäter, 2003). Continuous normally distributed demand is assumed during L_0, and so negative numbers of k were also included, even if k in reality will be a non-negative integer. It was difficult to say beforehand how many values of k this probability needed to be calculated for. It was found by experimentation that k-values <-20 had a probability close to zero, and therefore this was set as a lower bound for k. The probabilities, $p_{i,k}(L_0)$, were then calculated for increasing values of k until the total probability mass was satisfyingly close to 1, fulfilling the inequality in equation (12). The tolerance threshold for this truncation was set to $\alpha = 10^{-6}$.

$$p_{i,k}(L_0) = \frac{\sigma_i(L_0)}{Q_i}\left[G\left(\frac{(k-1)Q_i - \mu_i(L_0)}{\sigma_i(L_0)}\right) + G\left(\frac{(k+1)Q_i - \mu_i(L_0)}{\sigma_i(L_0)}\right) - 2G\left(\frac{kQ_i - \mu_i(L_0)}{\sigma_i(L_0)}\right)\right] \quad (11)$$

$$1 - \sum_{k=-20}^{z_i} p_{i,k}(L_0) < \alpha \quad (12)$$

$\sigma_i(L_0) = standard\ deviation\ for\ customer\ demand\ at\ retailer\ i\ during\ the\ warehouse$

$Q_i = order\ batch\ quantity\ from\ retailer\ i$

$z_i = smallest\ value\ satisfying\ the\ inequality\ in\ equation\ 12$

The G-function used above is shown in equation (13) (Axsäter, 2003).

$$G(x) = \varphi(x) - x(1 - \Phi(x)) \quad (13)$$

$\varphi(x) = standardized\ normal\ density\ function.$

$\Phi(x) = standardized\ normal\ cumulative\ distribution\ function.$

$\sigma^{w}(L_0)$ could now be obtained using equation (14) (Axsäter, 2003).

$$\sigma^{w}(L_0) = \sqrt{\sum_{i=1}^{n} Var_i^{w}(L_0)} = \sqrt{\sum_{i=1}^{n}\sum_{k=-\infty}^{\infty} (kQ_i - \mu_i^{w}(L_0))^2 p_{i,k}(L_0)} \quad (14)$$

$Var_i^{w}(L_0) = variance\ of\ the\ demand\ from\ retailer\ i\ at\ the\ warehouse\ during\ the\ ware$
$lead\ time$

In equation (14), the sum ranges from negative to positive infinity for k. However, since $p_{i,k}(L_0)$ was truncated as described above, and therefore implicitly assumed to be zero for every other value of k, the sum was bounded accordingly.

Having obtained both $\mu^{w}(L_0)$ and $Var^{w}(L_0)$, the average number of warehouse backorders, $E[IL_0^{-}]$, could be obtained through equation (15) (Axsäter, 2003).

$$E[IL_0^{-}] = \frac{(\sigma^{w}(L_0))^2}{Q_0 - q}\left[H\left(\frac{R_0 + q - \mu^{w}(L_0)}{\sigma^{w}(L_0)}\right) - H\left(\frac{R_0 + Q_0 - \mu^{w}(L_0)}{\sigma^{w}(L_0)}\right)\right] \quad (15)$$

$Q_0 = warehouse\ order\ batch\ quantity$
$q = largest\ common\ factor\ of\ Q_0\ and\ Q_i,\ \forall i$

For all items investigated in this report, q is 1. The H-function used in equation (15) is shown in equation (16).

$$H(x) = \frac{1}{2}[(x^2 + 1)(1 - \Phi(x)) - x\varphi(x)] \quad (16)$$

Using the obtained average warehouse backorder level $E[IL_0^{-}]$, the average warehouse stock on hand $E[IL_0^{+}]$and the average warehouse delay $E[W_0]$ could be obtained using equation (17) and (18), respectively (Axsäter, 2003). The latter of the two equations is an application of Little's Law.

$$E[IL_0^{+}] = R_0 + \frac{Q_0 + q}{2} - \mu^{w}(L_0) + E[IL_0^{-}] \quad (17)$$

$$E[W_0] = \frac{E[IL_0^{-}]}{\sum_{i=1}^{n} \mu_i} \quad (18)$$

2.5 Combined optimization of reorder points

With a model at hand for determining the average waiting time, as well as a way of describing the performance of the retailers under the METRIC approach, it was now possible to optimize the reorder point at the warehouse and at the retailers in combination for a specific item. The optimization algorithm used is described in the pseudo code below:

1. Find a suitable search interval for the optimal warehouse reorder point.
2. Find the average waiting time $E[W_0]$ produced by each R_0 in the interval.
3. For each R_0 in the interval, find the associated reorder point R_i and stock-on-hand level $E[IL_i^+]$ for each retailer.
4. Calculate the total system cost using the stock-on-hand levels for the warehouse and retailers and choose the combination of R_0 and R_i, $\forall i$ that produces the lowest total cost.

Each step in the optimization algorithm is elaborated on in the following subsections. The procedure was, naturally, repeated for each item.

2.5.1 R_0 interval

When choosing an interval, two things were important. Firstly, the interval of course needed to include the optimal R_0 value for certain. Secondly, as a larger interval implies more calculations, and the number of calculations increase rapidly with an increased number of retailers and items, it was desirable to choose as small an interval as possible.

The lower limit was set to $-Q_0$, as this produces a fill rate of 0 and the optimal point is sure to be higher than this value (Axsäter 2006, p. 103). The upper limit needed to be set at a level which would keep the number of iterations at an acceptable level while also, for certain, include the optimal R_0. The effect of incrementing R_0 is two-fold; the central warehouse holding cost will increase, while the waiting time for the retailers will decrease. Decreasing the waiting time in turn enables decrements of the retailer reorder points R_i, which successively leads to lower holding costs at the retailers. Since the shortest possible waiting time is zero time units, further increments of R_0 beyond this point would only increase warehouse costs while retailer costs remain the same. Hence, it was decided to set the upper limits of R_0 to the lowest value which ensured that the waiting time was satisfactory close to zero. The chosen acceptance threshold for this was $W_0 < 10^{-3}$.

2.5.2 $E[W_0]$

The average waiting time was found using the method explained earlier in section 2.4.

2.5.3 R_i and $E[IL_i^+]$

To find the optimal retailer reorder points for a given waiting time, a starting reorder point R_i was chosen for each retailer. Then, the reorder points were optimized one at a time for each retailer. As the cost function for a specific retailer is convex (Axsäter, 2003), the simple algorithm of increasing R_i, until the smallest possible value that satisfied the target fill rate was found, could be used. As a result, the smallest possible number of average stock-on-hand was found according to equation (3) in section 2.3. It was observed that the retailer reorder points strictly increased with longer waiting times. Due to this, the starting point for each R_i was set to the optimal reorder point with $E[W_0] = 0$. Setting the lower limit for R_i this way is also suggested by literature when using the METRIC approach with Poisson distributed lead time demand (Axsäter 2006, p. 272). An upper limit was not needed as the algorithm stopped iterating as soon as the appropriate fill rate was reached.

2.5.4 Optimal R_0 and R_i combination

When the combination of R_0 and R_i, $\forall i$ was found for each value in the R_0 search interval, the total system cost was calculated using equation (2) in section 2.1 *Calculating the cost of the system*. It is also mentioned in that section that the cost-minimization problem boils down to minimizing the total average stock-on-hand level if the holding costs are equal at all locations. As the provided data does not include information about the holding costs, it is assumed that this is the case. Therefore, the combination producing the lowest stock-on-hand was chosen as the new (R, Q) policies for the system.

3 RESULTS

3.1 Comments on the presentation of the results

The five items subject to analysis in this report are held in stock by at least two out of eight retailers. As interesting as it would have been to analyze how much the inventory system could be improved from the perspective of each individual inventory location, a fundamental perk with modeling the system as a multi-echelon one, is that it enables a more holistic system view. Thus, instead of presenting the results from an item-aggregated viewpoint of each retailer and the central warehouse, each item is treated as its own separate system, and the results are also presented as such. Treating the results in such a way also matches well with the way the results were obtained. Consequently, in this section, all the retailers are “anonymized”, and every item is therefore treated as if they have their own unique retail locations. Throughout the remainder of this report, the retailers will be numbered from 1-5, with no correlation between the retailer numbers for different items.

3.2 Reorder points

The optimization method renders a completely new set of reorder points for both the central warehouse and the different retailers. These are shown in Table 2. Precisely the same information for the current inventory system setup is shown in Table 3 for the sake of comparison. A table of the reorder points for the items at the *real* inventory locations can be found in Appendix C.

Tables 2 and 3 reveal that a common factor for all items is that their central warehouse reorder points drastically decrease for the new (R, Q) policies compared to the current one.

Table 2: The reorder points (R) for the new (R,Q) policies.

	Item 1	Item 2	Item 3	Item 4	Item 5

	R	R	R	R	R
CW	16	-76	8	-14	-6
Retailer 1	63	20	23	8	6
Retailer 2	10	31	5	14	1
Retailer 3	63		20	1	1
Retailer 4			6	1	
Retailer 5			18		

Table 3: The reorder points (R) for the current (R,Q) policies.

	Item 1	Item 2	Item 3	Item 4	Item 5
	R	R	R	R	R
CW	47	11	37	9	4
Retailer 1	32	1	1	1	2
Retailer 2	1	3	2	1	1
Retailer 3	34		10	1	1
Retailer 4			1	1	
Retailer 5			8		

3.3 Fillrates & Lead times

In this section the fill rates and lead times for each item and retailer are plotted. The target fill rates, depicted in black, are provided by TPTS. The “simulated current” values, depicted in yellow, are generated by the simulation model when provided with the reorder points currently in use in the system, as stated in Table 3, and can be interpreted as the performance of the system as it is running today. The “simulated new” values, depicted in orange, are generated by the simulation model when provided with the new R-values in Table 2 and can be interpreted as the actual performance of the system given the new reorder points. Lastly, the “estimated” values, depicted in blue, are the values generated from the analytic model when estimating the system for the newly proposed R-values in Table 2.

3.3.1 Item 1

For Item 1, the analytic model accurately estimates the performance, as the estimated and simulated values are very similar regarding both fill rate and lead time, as seen in Figure 1. There is also a significant increase in the fill rate at retailer 2 as all retailers now meet the target fill rate.

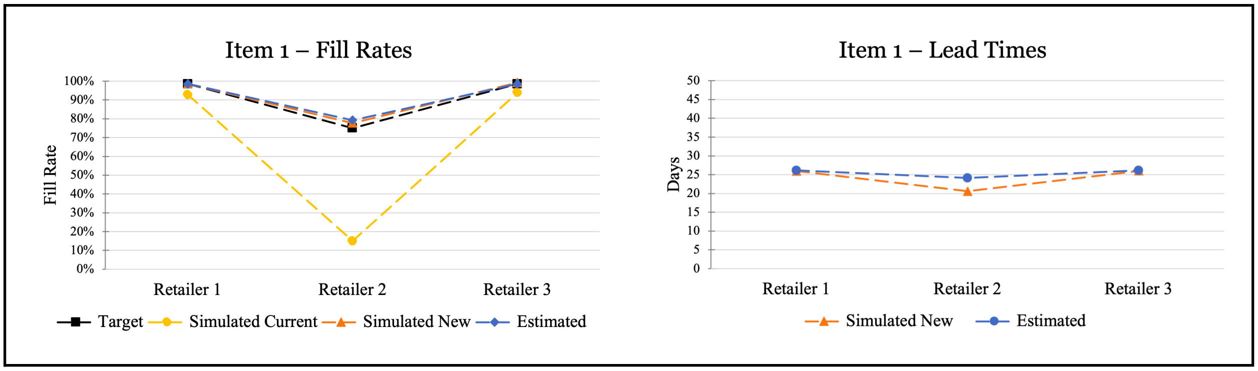

Figure 1: Fill rates and lead time for item 1.

3.3.2 Item 2

As seen in Figure 2 regarding Item 2, the analytic model overestimates the fill rate slightly compared to the simulated values. However, the actual fill rates have improved significantly in terms of meeting the target. The lead times are slightly underestimated compared to the simulated values.

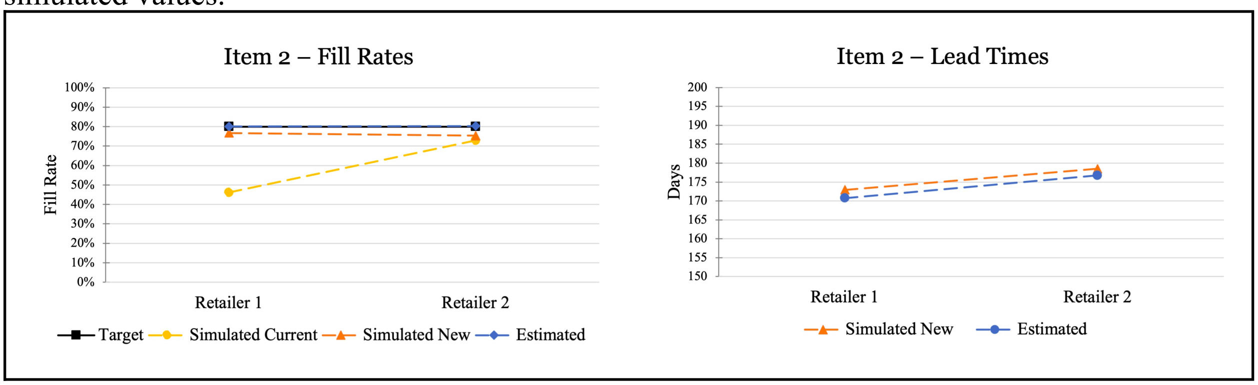

Figure 2: Fill rates and lead times for item 2.

3.3.3 Item 3

For Item 3, as seen in Figure 3, the fill rates have improved significantly in terms of meeting the target. The estimated values are also very accurate in comparison with the simulated new fill rates, with the only exception of a slight underestimation at retailer 2. Regarding the lead times there is a slight overestimation for all retailers except retailer 1.

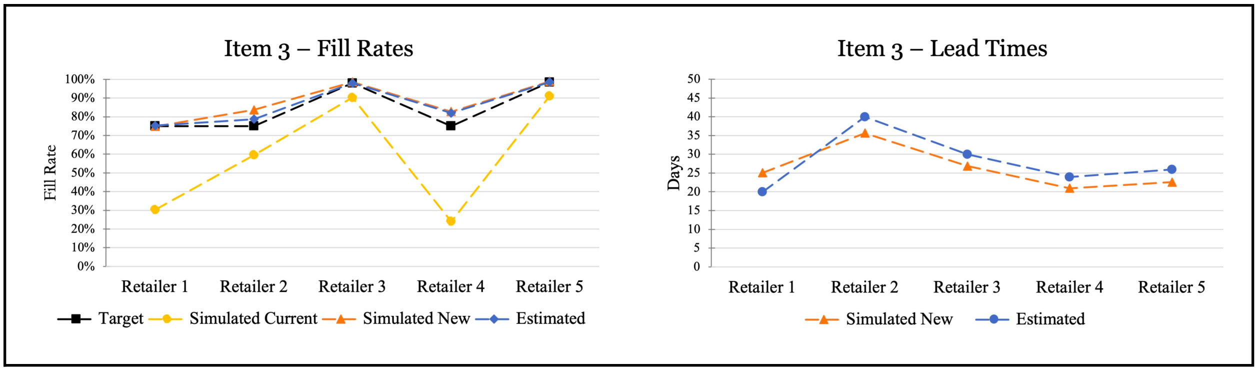

Figure 3: Fill rates and lead times for item 3.

3.3.4 Item 4

For Item 4, as seen in Figure 4, the fill rate for retailer 1 and 2 are slightly underestimated by the analytic model. However, the actual fill rates are significantly improved in comparison with the current fill rates of the system. For this item, the estimations of the lead time are less accurate in comparison with the other items.

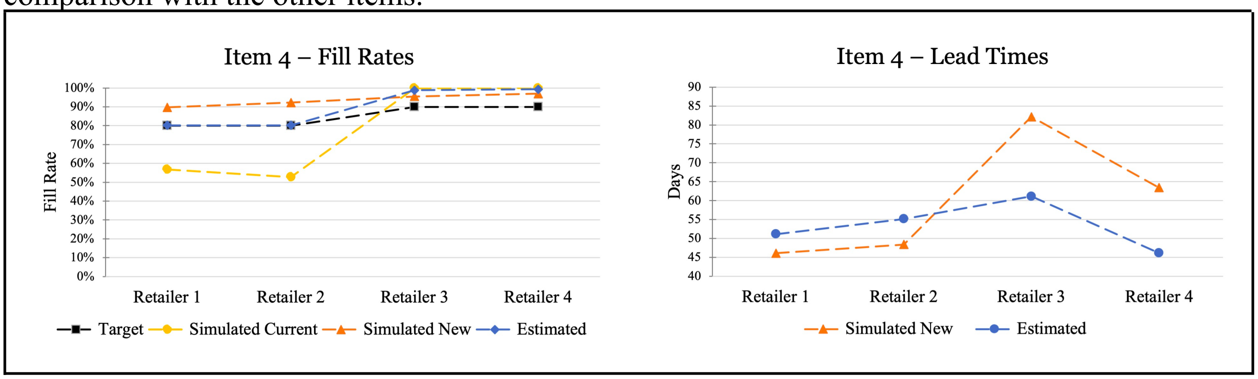

Figure 4: Fill rates and lead times for item 4.

3.3.5 Item 5

Lastly, for Item 5 in figure 5, the estimated values are accurate both regarding fill rates and lead times. As the current reorder points already meet the target fill rates, there is no significant improvement for this item in this regard.

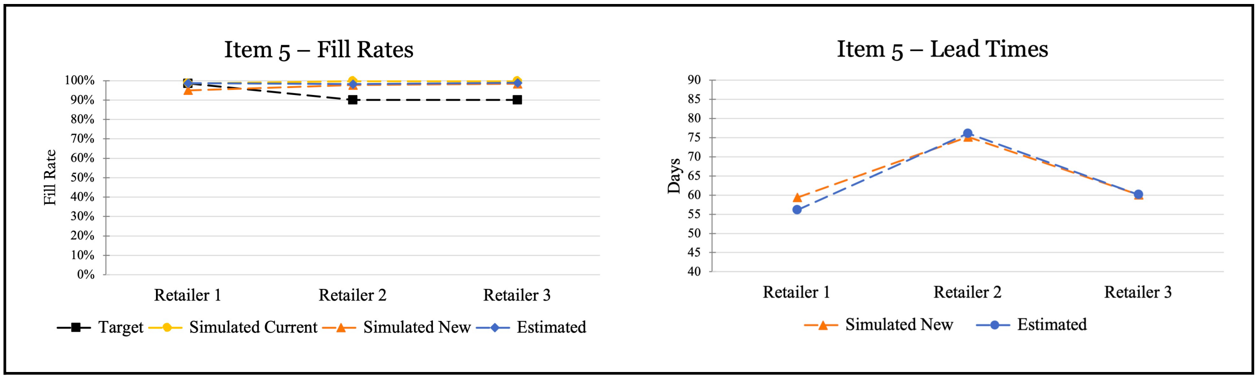

Figure 5: Fill rates and lead times for item 5.

3.4 Stock-on-hand

While the suggested new reorder points improve the performance in terms of meeting fill rates in the system, the total stock on hand did not change significantly, as seen in Figure 6. The stock-on-hand was increased for some item systems, and reduced for others, but the total stock on hand for the system stayed at approximately 350 units. Under the assumption that holding costs are equal for all items, this means that the total cost of the system has stayed at the same level.

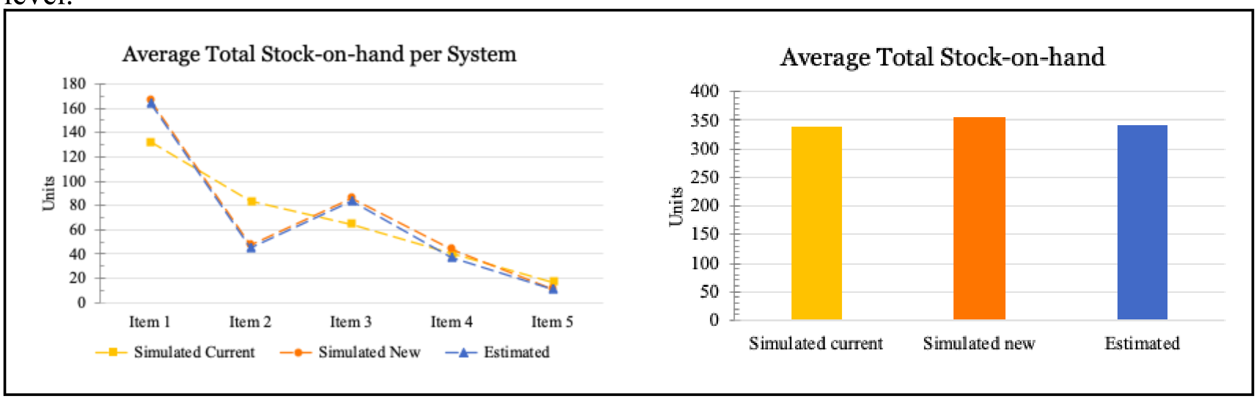

Figure 6: Stock-on-hand levels.

For the current (R, Q) policies, the warehouse is generally stocked with more items compared to the levels at the retailers. For the estimated new values, and the simulated equivalents, this ratio is reversed, meaning that stock-on-hand levels are generally higher at the retailers compared to levels at the warehouse. This result is visualized in Figure 7, where stock-on-hand levels for retailers and warehouses are calculated as averages across all five systems.

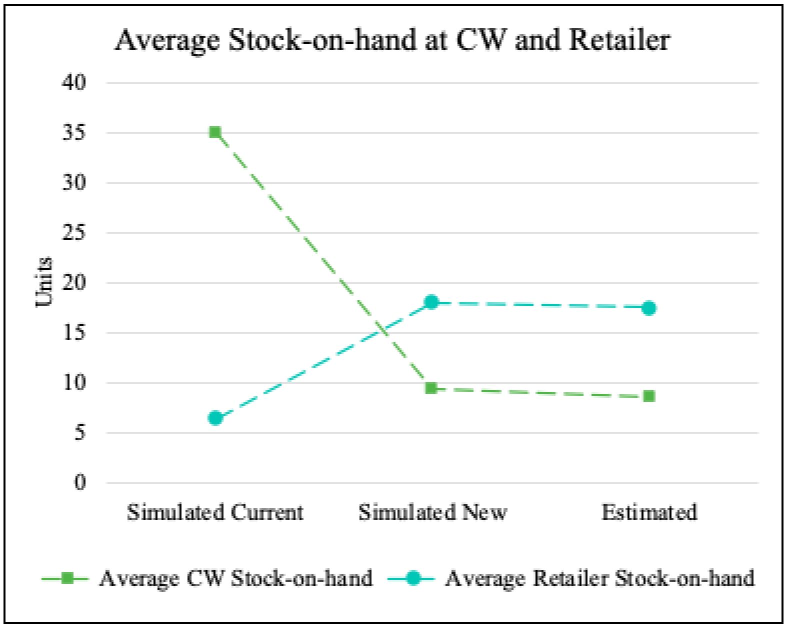

Figure 7: Stock-on-hand levels for different echelons.

3.5 Waiting time

In Table 4 the estimated and simulated average waiting times, $E[W_0]$, for each item, when using the new (R, Q) policies, can be found. The analytic model can be seen to accurately estimate the simulated average waiting times in both absolute and relative terms.

In Table 5 the simulated waiting times with the relative error of the estimation are found for each retailer and each item Remember that the waiting time for every retailer for a specific item is estimated by the average waiting time found in Table 4. In Table 6 the absolute and relative errors of the waiting time estimations are found.

Mostly, the estimations are accurate. However, the estimations for Item 3 and Item 4 stand out as being less so. For Item 3, it is explained by Retailer 1 having a significantly longer waiting time in comparison to the other retailers, which are all close to 6 days. For Item 4, it is explained by Retailer 1 and 2 experiencing waiting times of about 35 days, while retailer 3 and 4 experience waiting times of about 60 days, rendering the average estimation poor for all four retailers.

Table 4: Estimated and simulated average waiting time.

Average waiting time, $E[W]$	Item 1	Item 2	Item 3	Item 4	Item 5
Estimated	10.12	156.75	9.94	41.15	46.11
Simulated	8.85	158.72	8.21	47.77	46.88
Relative difference of estimation	13%	-1%	17%	-16%	-2%

Table 5: Simulated waiting times.

Retailer	Item 1	Item 2	Item 3	Item 4	Item 5
1	9.94	158.92	15.05	36.09	49.37
2	6.60	158.51	5.63	34.42	45.18
3	10.00		6.84	62.15	46.09
4			6.93	58.42	
5			6.58		

Table 6: Absolute and relative error of estimated waiting times.

Retailer	Item 1	Item 2	Item 3	Item 4	Item 5
1	0.18 (+2%)	-2.18 (-1%)	-5.11 (-51%)	5.06 (+12%)	-3.26 (-7%)
2	3.52 (+35%)	-1.76 (-1%)	4.31 (+43%)	6.73 (+16%)	0.93 (+2%)
3	0.12 (+1%)		3.10 (+31%)	-21.00 (-51%)	0.02 (0%)
4			3.01 (+30%)	-17.27 (-42%)	
5			3.36 (+34%)		

4 ANALYSIS

4.1 Model performance

4.1.1 System improvement

Regarding target fill rates, a significant improvement of the inventory system can be noted by applying the proposed analytic multi-echelon method. The slight increase in total stock-on-hand for the simulated model before and after applying the new (R, Q) policies will somewhat increase the tied-up capital. On the other hand, the significant increase in service level will most likely generate more satisfied customers, goodwill for the company and in the long run increase sales. In the best of cases, the total stock-on-hand for a specific item is reduced while the target fill rate is achieved, which was the case for Item 2.

The results obtained when simulating the analytically suggested reorder points should preferably be verified somehow to be able to conclude that these reorder points really do entail a near-optimal solution. While this is difficult to do without actually testing the simulation model for all combinations of reorder points in a trial-and-error fashion – a very extensive task – Axsäter (2006, p. 281) suggests a design characteristic which is oftentimes typical for optimal solutions. Albeit generalizations for optimal designs of inventory systems are difficult to make, the stock-on-hand relationship between echelons is not seldom such as having downstream installments stocked higher compared to the upper echelons. This result was indeed obtained when simulating the suggested near-optimal order points, see Figure 7.

4.1.2 Fill rate estimation

For most items, the simulated model renders a fill rate above the target for each retailer when reorder points obtained from the analytic approach are used. However, when observing the results of Item 2, no retailer achieves the target fill rate when the inventory system is modelled in the simulation. For such results, the optimal reorder points obtained from the analytic method could be used as a base in the process of determining optimal reorder points. The simulation model could be used to adjust the obtained reorder points slightly to achieve the target fill rates. On the other hand, implementing simulation models for the full scale system could be complicated and time consuming. Therefore, it could be more cost-effective to accept slight errors in some fill rates. Furthermore, it should be noted that even without a simulation model at hand, the fill rates rendered by the suggested analytic approach is much more satisfying in comparison to the current system. Nevertheless, if the targets are a critical matter, the reorder points generated by the suggested analytic approach could be slightly increased to avoid an unsatisfactory service level. However, as it is very difficult to know whether the fill rate is over- or underestimated based on the estimations, this approach might increase stock-on-hand levels unnecessarily. Therefore, it is recommended to only assign this extra buffer for items where achieving the fill rate is extra important. It should be noted that it is also possible to hedge against underestimations by increasing the fill rate targets slightly.

4.1.3 Lead time estimation

To further evaluate the accuracy of the model, it is also important to compare the generated lead times from the analytic method with the lead times obtained when running the simulation, as the estimations are modeled under the approximation of every retailer facing the same waiting time at the warehouse. For instance, there is a significant difference between estimated and simulated lead time for two retailers for Item 4. The difference can be explained by the fact that there is a significant difference between the simulated and estimated delay at the warehouse, as was noted in section 3.5. Interestingly, the fill rate and stock-on-hand level estimations, while being the most inaccurate of the bunch, are still quite accurate even for this item, which implies robustness in the model. Furthermore, the estimations of fill rate and lead times for Item 3 are very accurate, implying that large errors in absolute values have a larger impact than large errors in relative values as both Item 3 and 4 stood out in this regard.

Lastly, it should be noted that the less accurate estimations of Item 4 could be a result of some other error in the model since similar deficiencies were observed when evaluating the model with a waiting time of 0 (see Appendix A).

4.2 Scaling the model

4.2.1 Analytical considerations

There are several emerging challenges that come as a result of upscaling the analytic model to a level where it can be applied on all items and retailers within the inventory system. Firstly, the simulation model might need to be extended to include more retailers to enable testing and validation of the reorder points obtained from the analytical approach. Furthermore, in the case where one or several retailers have significant errors between real and average delay at the warehouse, the reorder points might need to be tweaked to a larger extent. This risk could potentially be compounded for larger inventory systems with a multitude of retailers keeping the same item in stock. However, as mentioned by Axsäter (2003) it is possible to efficiently optimize quite large systems with the approximation of delay used in this solution.

Another challenge with scaling the analytic model presents itself as a company-strategy issue. As the analytic model is built, it assumes that orders placed at both echelons are delivered according to a first-come, first-served policy. This means that the stochastic lead time to get an order delivered is unaffected by any orders that are placed *after* the order point (Axsäter 2006, p. 119). However, in a real setting, there might be situations when a company like TPTS would momentarily override this model-necessary assumption. This could occur when TPTS receives orders that are deemed especially strategically important, for example orders from a prominent customer. One-off instances like this could very well happen, but a problem arises if the company management agrees to let this happen on a rather frequent basis. Ultimately, a point could be reached where the analytically proposed near-optimal reorder points for the large system, which are based on an estimation of the waiting time enabled by a strict FCFS policy, no longer adequately can describe the system in practice.

4.2.2 Numerical considerations

When modelling a test system for only five items, algorithm and coding inefficiencies are not very prominent nor problematic. However, if the model is to be implemented to handle the entire inventory system, computer processing speed and memory capacity are likely to act as bottlenecks.

A first important step to prepare the system model for upscaling should be to ensure that the source code is free of possible errors or inefficiencies, such as errant iterations or redundant function evaluations. The implementation types themselves are also important to scrutinize, as sometimes it might be more computationally efficient to e.g. substitute a loop for a recursive method, or vice versa.

It is just as important to improve the optimization algorithms described in pseudo code in section 2.5; there are for certain inefficiencies there to be found and eliminated. The algorithm contains several parts where an increment in either items or retailers leads to drastic increases in required iterations. Three of these stand out as especially interesting in this regard: (i) the optimization procedure for optimal retailer reorder points R_i for a given R_0; (ii) the iterative search through the interval of possibly optimal reorder points R_0; and (iii) the calculation of $P(D(L) = d)$ when obtaining$S_{2,i}$.

The reason for concern with (i) is that the way the optimization procedure is currently implemented, if a total of r reorder points are iterated through, an addition of n retailers holding an item in stock would require $n * r$ more iterations by the program. To remedy this problem, a cleverer search algorithm could be implemented. As earlier stated, the cost function of a specific retailer is convex. Hence, a simple line search algorithm like the golden section search, or Fibonacci search, would significantly reduce the number of function evaluations (Böiers 2010, p. 31). A drawback is that these search methods require an upper limit for the search interval. Such a limit could be attempted to find by testing a large value for R_i to make sure that it is producing a fill rate above the target level. Should it fail to do so, that reorder point could then be set as a new lower limit, while another, larger reorder point is tested as an upper limit.

The problem with (ii) is that it is difficult to decide an adequate R_0 interval and subsequently shorten it in an effective manner when searching for the optimum, as the total cost function was found not to be strictly convex. This can be seen when examining the function in the more detailed view in figure 8. It can also be seen from the figure, however, that from a bird's-eye perspective, the function looks quite convex, which suggests that it could be possible to find a shorter interval by first calculating the total cost for a few R_0 values. Such a technique was attempted by implementing a simple algorithm which tried n different R_0 values and then found the value that produced the lowest total costs. The two adjacent values were then compared in order to determine the smallest of the two. Due to the close-to-convexity, the optimal R_0 value was theoreticized to be between these two values. However, the interval was created with its upper and lower limits set as the R_0values adjacent to the two found reorder points, enlarging the interval and hedging against the optimal to be found just outside the bounds of the smaller interval. Lastly, all R_0 values in this shortened interval were evaluated for optimality. The method worked well as it did reduce the computations needed significantly. However, it was problematic to find a suitable value for n, as well as an interval size that worked well for every item. The method did, however, produce at least close-to-optimal values for R_0. Although not yet perfected, this method could be revisited and improved upon when trying to implement the model on a larger scale.

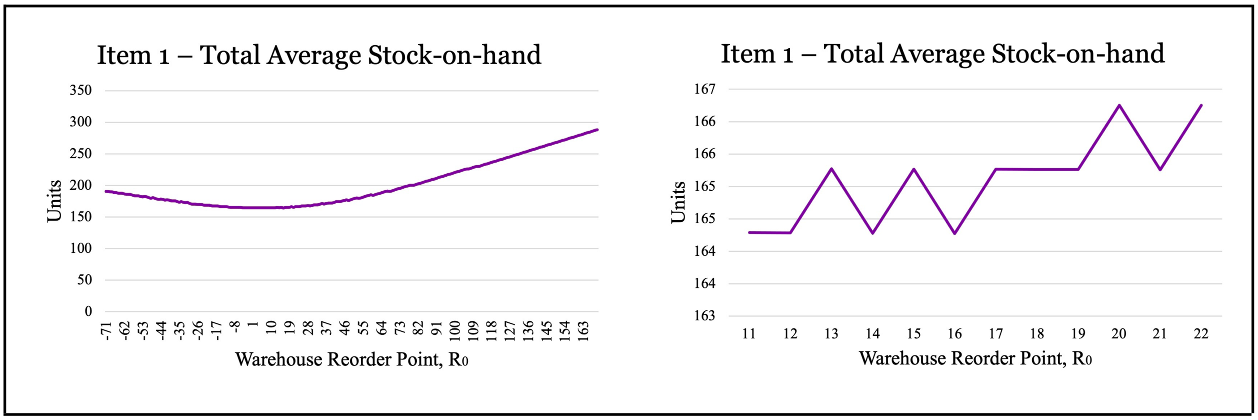

Figure 8: The optimal average stock-on-hand for different warehouse reorder points.

The issue with (iii) is that $P(D(L) = d)$, in theory, could be infinitely large. Consequently, it needs to be truncated with a pre-decided level of tolerance. Relaxing this tolerance might reduce the necessary number of computations drastically, which is likely to have a resounding impact on the total computation time, as this part of the code is executed more times than any other due to its placement at the very bottom of the multiple-loop rabbit hole. It is quite difficult to say what an acceptable tolerance level is when truncating. If the tolerance level is too relaxed by too much, however, the results become unreliable. In order to find a suitable balance, different tolerance levels can be tried and evaluated.

5 CONCLUSION

Introducing a multi-echelon model to control the inventory system at TPTS could improve the performance of the system significantly in terms of reaching fill rates, while only slightly increasing the total average stock-on-hand level. The analytic approach described in the report produced a model which successfully manages to estimate both fill rates, stock-on-hand levels and lead-times for all items, with the exception of the lead time estimation for retailers in one of the item inventory systems.

A pursuit of a large scale system implementation could prove difficult, considering the risk of facing increasingly worse lead time approximations, as well as other numerical considerations stemming from inclusions of additional items and retailers. However, as implied by the successful results of this test run, the full-scale implementation of the model may very well be worth the effort.

6 REFERENCES

Axsäter, S. 2003. Approximate optimization of a two-level distribution inventory system. *Int. J. Production Economics* 81-82: 545-553. doi: https://doi.org/10.1016/S0925-5273(02)00270-0

Axsäter, S. 2006. *Inventory Control, 2nd edition*. Springer.

Böiers, L. 2010. *Mathematical Methods of Optimization.* Studentlitteratur AB.

APPENDIX A

Table A1: Fill rates obtained from the simulation model given $R_0 = 1000$.

Retailer*	Item 1	Item 2	Item 3	Item 4	Item 5
1	0,9588	0,4718	0,3027	0,5699	0,9871
2	0,1525	0,7453	0,5967	0,5315	0,9965
3	0,9547		0,8963	0,9967	1
4			0,23910	1	
5			0,9105		

**Retail numbers represent the chronological order in which they occurred in the given data, and do not denote the number of actual retailers.*

Table A2: Fill rates obtained from the METRIC approach model, given $E(W_0) = 0$.

Retailer*	Item 1	Item 2	Item 3	Item 4	Item 5
1	0,9461	0,4602	0,2829	0,3649	0,9846
2	0,1481	0,7382	0,5393	0,3128	0,9969
3	0,9543		0,9049	0,9986	0,9993
4			0,2406	0,9999	
5			0,9124		

**Retail numbers represent the chronological order in which they occurred in the given data, and do not denote the number of actual retailers.*

Table A3: Percentage difference between fill rates obtained from the METRIC approach model in relation to the simulation model, given $E(W_0) = 0$.

Retailer*	Item 1	Item 2	Item 3	Item 4	Item 5
1	−1%	−2%	−7%	−36%	0%
2	−3%	−1%	−10%	−41%	0%
3	0%		1%	0%	0%
4			1%	0%	
5			0%		

**Retail numbers represent the chronological order in which they occurred in the given data, and do not denote the number of actual retailers.*

APPENDIX B

Table A1: Fill rates obtained from the simulation model given current reorder points.

Retailer*	Item 1	Item 2	Item 3	Item 4	Item 5
1	0,9279	0,4613	0,3028	0,5679	0,9856
2	0,1500	0,7286	0,5949	0,5273	0,9976
3	0,9388		0,9023	0,9991	0,9970
4			0,2396	1,0000	
5			0,9106		

**Retail numbers represent the chronological order in which they occurred in the given data, and do not denote the number of actual retailers.*

Table A2: Fill rates obtained from the analytic model given current reorder points.

Retailer*	Item 1	Item 2	Item 3	Item 4	Item 5
1	0,9457	0,4600	0,2827	0,3648	0,9845
2	0,1480	0,7378	0,5390	0,3127	0,9969
3	0,9539		0,9045	0,9986	0,9993
4			0,2406	0,9999	
5			0,9120		

**Retail numbers represent the chronological order in which they occurred in the given data, and do not denote the number of actual retailers.*

Table A3: Percentage difference between fill rates obtained from the analytical approach in relation to the simulation model, given current reorder points.

Retailer*	Item 1	Item 2	Item 3	Item 4	Item 5
1	2%	0%	-7%	-36%	0%
2	-1%	1%	-9%	-41%	0%
3	2%		0%	0%	0%
4			0%	0%	
5			0%		

**Retail numbers represent the chronological order in which they occurred in the given data, and do not denote the number of actual retailers.*

APPENDIX C

Table C1: The reorder points for the specific inventory locations

Location	CW	Retailer 2	Retailer 5	Retailer 7	Retailer 11	Retailer 12	Retailer 19	Retailer 30	Retailer 32
Item 1	16			63			10	63	
Item 2	-76		20			31			
Item 3	8	23			5	20	6	18	
Item 4	-14	8	14			1			1
Item 5	-6	6			1		1		