Coordinating push and pull flows in a lost sales stochastic supply chain

Georgios Varlas and Michael Vidalis

Department of Business Administration, University of the Aegean
Michalon 8, Chios 82100, Greece
g.varlas@aegean.gr, m.vidalis@aegean.gr

Abstract

In this paper a serial, three echelon, push-pull supply chain is investigated. The supply chain consists of a provider, a distribution centre (buffer) and a retailer. The material flow between upstream stages is push type, while between downstream stages it is driven by continuous review, reorder point/order quantity inventory control policy. Exponentially distributed lead times between stages are assumed. External demand occurs according to pure Poisson, while the demand that cannot be met is lost. The system is modelled using matrix analytic methods as a Markov birth-and-death process. An algorithm is developed to generate the transition matrix for different parameters of the system. Then, the corresponding system of stationary linear equations is generated and the solution of the stationary probabilities is provided. Key performance metrics such as average inventories and customer service levels at each echelon of the system can be computed. The algorithm is programmed in Matlab© and its validity is tested using simulation, with the two approaches giving practically identical results. The contribution of our work is an exact algorithm for a lost sales push-pull supply network. This algorithm can be used to evaluate different scenarios for supply chain design, to explore the dynamics of a push-pull system, or as an optimization tool.

1998 ACM Subject Classification G.3 Probability and Statistics

Keywords and phrases Supply Chain Management, Push-Pull systems, Markov Processes

Digital Object Identifier 10.4230/OASIcs.SCOR.2014.52

1 Introduction and Literature Review

The supply chain (SC) consists of all the parties involved in manufacturing, distribution, and delivery of the product to customers. Members of various echelons in the SC are related to each other, either directly or through intermediaries. Decisions made in any echelon of the SC can affect the costs of other members and vice versa. Supply chain management (SCM) involves the management of flows between the stages of a supply chain, so as to maximize total expected profitability. Six tool drivers may be used to improve supply chain performance: Inventory, Transportation, Facilities, Information, Sourcing, and Pricing [4].

Inventory control plays an important role in supply chain management. Properly controlled inventory can satisfy customers' demand, smooth the production plans, and reduce the operational costs. In practice, inventory control systems usually operate in dynamic environments. Calculating the exact ordering quantity, deciding the proper reordering point, choosing the right inventory reviewing policy, and managing the safety stock are key factors for the SC profitability.

The reordering process is characterized by the review interval, the determination of the order size, the order costs and the objective function. Two types of review systems are widely used in business and industry. Either inventory is continuously monitored (continuous

4th Student Conference on Operational Research (SCOR'14).
Editors: Pedro Crespo Del Granado, Martim Joyce-Moniz, and Stefan Ravizza; pp. 52–62

OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

review), or inventory is reviewed at regular periodic intervals of length R (periodic review). Whether or not to order at a review instant is usually determined by a reorder level denoted by s. This is the inventory position at which a vendor is triggered to place a replenishment order so as to maintain an adequate supply of items to accommodate current and new customers.

Common assumptions of the inventory model to represent the system concern the demand distribution, the lead time (deterministic or stochastic), and the maximum number of outstanding orders. Another important characteristic concerns the demand that cannot be met by the inventory on hand. Such demand can be back-ordered to be met in the future, or can be lost (lost sales assumption). Latter cases appear to be more difficult to analyse and such models have received less attention in the literature [2].

Usually, a Markov model, based on the characteristics of the system and the assumptions made, is developed to represent the on-hand inventory level and the individual outstanding orders. The decision (or ordering) points in such a model are the time instants at which either a demand occurs and no order is outstanding, or a replenishment order is delivered. Based on the transition probabilities and steady-state behaviour of the system, the long-run behaviour of the inventory model is analysed. The stationary distribution function of the on-hand inventory level is used to analyse the inventory system in terms of expected average cost and service level. The pioneer work in this area was that of Clark and Scarf [5] who considered an inventory system with periodic review using echelon stock policy.

In the final phase, the inventory control variables, such as the reorder level and order quantities, are set. Either an exact procedure or an approximation procedure can be used to find these values. Two types of exact procedures are commonly used in literature, namely a policy iteration algorithm and an extensive numerical search procedure.

In general, production/inventory systems can be classified as push, pull, or hybrid push/pull-type systems. In a push-type system the parts are released to the next station as quickly as possible to avoid starvation of the downstream stations. On the other hand, the pull-type system drives production based upon customer demand. Such systems are widely used and different modelling approaches have been proposed. Chen [3] generalizes the Clark and Scarf model by allowing batch transfers of inventories in a serial network with n stages and backorders. Badinelli [1] constructs a model of the steady-state values of on-hand inventory and backorders for each facility of a serial inventory system, where each facility follows a (Q, R) policy based on installation stock. The descriptive model he presents is intended for optimizing the parameters of such a policy and for obtaining theoretical results about the behaviour of the system. Finally, Gupta and Selvaraju [9] study the effect of stock allocation among different stages, when the total amount of stock in the supply system is fixed at the optimal level. They develop an approximation scheme for performance evaluation of serial supply systems when each stage manages its planned inventories according to a base-stock policy.

In the hybrid push/pull system the production at the earlier upstream stations is push-type, while the production of the later downstream stations is controlled by pull-type policies. The push-pull boundary or junction point is defined as the last push station and determines which stations are push systems and which stations are pull systems. In most cases hybrid systems perform better than pure push, or pure pull systems, while they are more flexible to address growing product variety and shorter product life cycles. However, their analysis is more complicated. Cochran and Kim [6] study with Simulated Annealing a horizontally integrated hybrid production system (HIHPS) with a movable junction point. Their proposed solutions include the location of the junction point, the safety stock level, and the number of

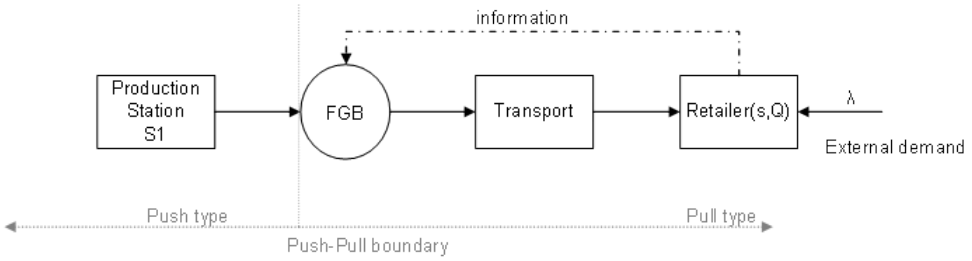

Figure 1 System layout.

kanbans needed in the pull system. Ghrayeb et. al [8] investigate a hybrid push/pull system of an assemble-to order manufacturing environment. They use discreet event simulation along with a genetic algorithm and the objective function for their model is to minimize the sum of inventory holding cost and delivery lead time cost. Finally, Cuypere et. al [7] introduce a Markovian model for push-pull systems with backlogged orders, basing their analysis on quasi birth-and-death processes.

The main goal of our work is to provide an algorithm for the exact evaluation of a push-pull supply network with lost sales. The resulting descriptive model can be used as a design tool or as a tool for the optimization of the parameters of the system.

2 Description of the System

In this article a single product, linear, push-pull supply chain is investigated. The system under consideration is shown in Figure 1. A reliable station S_1 produces (or administers in the system) product units at a rate μ_1 and exponentially distributed inter-arrival times. Finished products are stored in a finite Finished Goods Buffer (FGB). Inventory at buffer is denoted by B_t. In the case where S_1 completes processing, but on completion FGB is full, station S_1 blocks (blocking after processing). Station S_1 consists the push section of the system. Downstream, the retailer R holds inventory I_t and faces external demand with pure Poisson characteristics (customers' inter-arrival times are exponentially distributed and every customer asks for exactly one unit). When the retailer is out of stock, occurring demand is lost. The retailer follows continuous review inventory control policy with parameters (s,Q). When inventory I_t reaches the reorder point s, a replenishment order of Q units is placed on the buffer. The actual level of the sent order depends on the available inventory at buffer. If $B_t \geq Q$, a full order is dispatched to the retailer. Otherwise, an incomplete order is dispatched. In the case where FGB is empty, dispatching is suspended until one unit finishes processing at S_1, upon which it is immediately forwarded for transportation to the retailer. Transportation is modelled as a virtual station T. Inventory in transit is denoted by T_t. In the model, transportation is considered independent from both FGB and retailer. On transportation initiation inventory T_t is subtracted from the buffer and remains in the virtual station T until on transportation completion it is added to the inventory of the retailer I_t. Exponentially distributed times for the transportation are assumed.

To model the system, the following assumptions are made:

1. Both customer demand and lead time are stochastic.
2. There are no back-orders. Demand that cannot be met from inventory on hand is lost both at the retailer and the buffer.
3. At any given time only one order can be in transit from FGB to the retailer.

4. The retailer follows continuous review inventory control policy with parameters (s,Q). Decision variables of the retailer are reorder point s and order quantity Q. In other words, the retailer's problem is the optimization of s and Q simultaneously.
5. Order quantity Q is constant.
6. Transportation is modelled as an independent station and inventory in transit depends on B_t at the time of transportation initiation.
7. Station S_1 never starves.
8. Station S_1 blocks when on completing the processing of a unit, finds the FGB full (blocking after processing). In the case where Q is greater than buffer capacity, the blocked unit is considered available for transportation to the retailer along with the inventory of the buffer (The blocked unit is considered part of the buffer). Otherwise, the blocked unit is transferred to the buffer immediately after there is available space and at the same time station S_1 resumes production.
9. There are no loading/unloading times.
10. All stations are reliable.
11. All times are exponentially distributed.
12. For methodology reasons, it is assumed that no two events can occur at exactly the same time.

3 Description of the Model

3.1 States definition

The system cannot be modelled as a quasi birth-and-death (QBD) process since the assumption of (s,Q) policy allows for transitions between non-adjacent levels. However, the system can be modelled as a continuous time-discreet space Markov process using matrix analytic methods. Taking advantage of repeating structures, an algorithm is developed to generate the transition matrix for different parameters of the system. Then, the corresponding system of stationary linear equations is generated and the solution of the stationary probabilities is provided. Using stationary probabilities, key performance metrics, such as average inventories and service levels at each echelon of the system can be computed.

The design variables that determine the dimension and structure of the transition matrix are:
B: The capacity of the finished goods buffer (FGB).
s: The reorder point at the retailer.
Q: The quantity of the orders requested by the retailer.
All three variables are assumed to be positive integers or zero, with the exception of Q which obviously cannot be zero. Although some scenarios lack physical meaning (for example when $Q > B + 1$), for the development of the algorithm no assumptions about the variable values are made. The other parameters of the model are:
μ_1: The production rate of the Station 1, or rate of units admission in the system (exponential inter-arrival times)
μ_2: The transfer rate of a replenishment order from the buffer to the retailer (exponential times)
λ: The rate of external customers' arrivals (pure Poisson demand)

We will illustrate the methodology using a simple example. We assume buffer capacity $B = 2$, reorder point $s = 1$ and order quantity $Q = 2$. At any moment t, the state of the system can be defined by a three dimensional vector (B_t, T_t, I_t), where:
B_t: The level of inventory on hand at the FGB. $0 \leq B_t \leq B + 1$, where the case $B_t = B + 1$ corresponds to blocking (see assumption 8). In our example $0 \leq B_t \leq 3$.

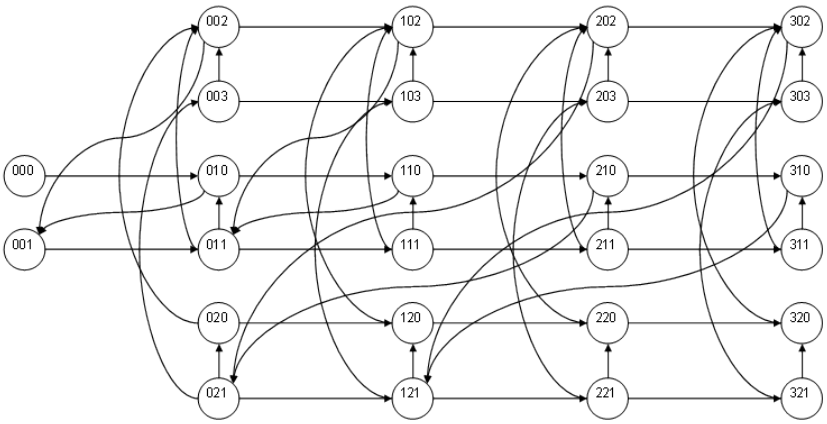

Figure 2 State transition diagram for $B = 2$, $s = 1$, $Q = 2$.

T_t : The number of product units in transit from FGB to the retailer. $0 \leq T_t \leq Q$. $T_t = 0$ means that there is no inventory in transit, while when $T_t = Q$ we have a complete order in transit to the retailer. $0 < T_t < Q$ corresponds to incomplete order. In our example $0 \leq T_t \leq 2$

I_t : The inventory on hand at the retailer. In general $0 \leq I_t \leq s + Q$. In our example $0 \leq I_t \leq 3$

The state space S of the Markov process is comprised of all the possible triplets (B_t, T_t, I_t) and its dimension depends on B, s and Q. For the example under consideration there are 26 possible states. It can be easily proved that for any value of the given parameters, the dimension of the state space is given by

$$N_B^{s,Q} = (s+1) + (s+2) \cdot Q \cdot (B+2)$$

3.2 State transitions

The state of the system can be altered instantaneously by three kinds of events.

1. The completion of processing of one product unit at station S_1. In this case B_t increases one unit. In infinitesimal time dt, the possibility of the event occurring is $\mu_1 \cdot dt + o(dt)$. $o(dt)$ is an unspecified function such that $\lim_{dt \to 0} \frac{o(dt)}{dt} = 0$.
2. The arrival of an outstanding order at the retailer. In this case the inventory on hand of the retailer I_t increases by T_t units. If the new value of I_t is above the reorder point, then T_t resets to zero. Otherwise, a new transfer from FGB is initiated. T_t takes the value of the new order and B_t decreases correspondingly. In infinitesimal time dt, the possibility of the event occurring is $\mu_2 \cdot dt + o(dt)$.
3. The occurrence of external demand. In this case the inventory on hand of the retailer decreases by one unit. If the new inventory equals the reorder point s, a replenishment order is given to the FGB. T_t takes the value of inventory in transit and B_t decreases correspondingly. In infinitesimal time dt, the possibility of the event occurring is $\lambda \cdot dt + o(dt)$.

In Figure 2 is given the state transition diagram for the example under consideration. There are certain symmetries in the diagram. Such symmetries are also present in the transition matrix and form the basis for the development of the algorithm which is the target of our analysis.

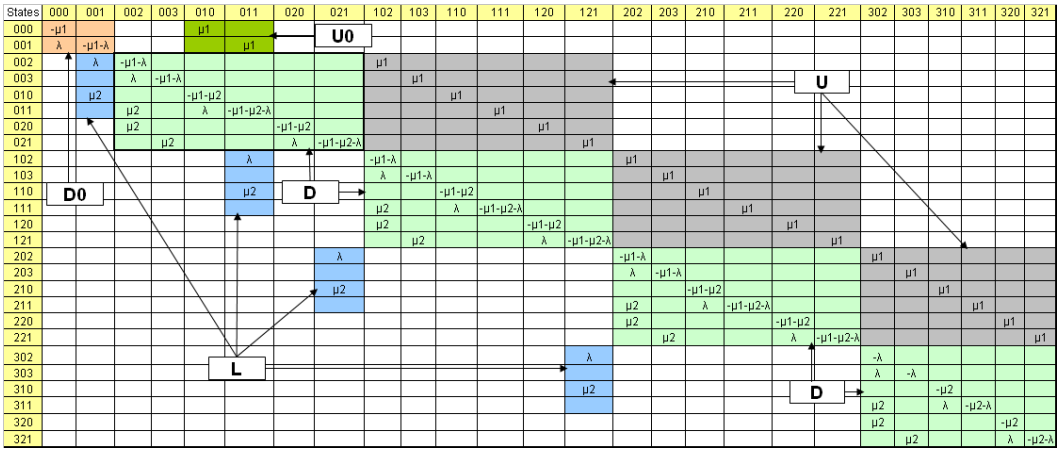

Figure 3 Transition Matrix for $B = 2$, $s = 1$, $Q = 2$.

3.3 The Transition Matrix

Before displaying the transition matrix, we must define a linear ordering of the states. We use the lexicographical ordering [10]. We take as basic level the subset of all states corresponding to a fixed buffer inventory B_t. Such levels correspond to columns in the state transition diagram. Within each level the states are grouped according to the inventory in transit T_t. For fixed level and fixed inventory in transit, the states are ordered by inventory at retailer I_t. To summarize: State (x, y, z) precedes state (x', y', z') if x<x'; State (x, y, z) precedes state (x, y', z') if y<y'; State (x, y, z) precedes state (x, y, z') if z<z'. The transition matrix for the example under consideration is given in Figure 3.

The transition matrix can be divided into sub-matrices with well defined and predictable characteristics. Some examples of constituent sub-matrices are given in Figure 4.

The first diagonal sub-matrix D_0 corresponds to the boundary states where there is no inventory in transit, $I_t < s$ and $B_t = 0$. In our example it is a 2x2 block at the top left. In general it is a $(s+1) \times (s+1)$ sub-matrix. On the diagonal, the basic repeating block D is a $(s+2) \cdot Q \times (s+2) \cdot Q$ sub-matrix. It corresponds to analogous transitions for different values of buffer inventory (levels), and it is repeated B+2 times. D can be further analysed into constituent sub-sub-matrices.

- D_1 is a $Q \times Q$ block on the diagonal of D. It corresponds to the diagonal of transition matrix P, where no event occurs, and to the occurrence of external demand when no replenishment order is initiated ($I_t > s$)
- D_2 is a $(s+1) \times (s+1)$ block also on the diagonal of D. Within each D block, D_2 is repeated Q times. It corresponds to the diagonal of transition matrix P and to the occurrence of external demand when no replenishment order is initiated ($T_t > 0$).
- D_3 is a $k \cdot (s+1) \times Q$ block, where k=min(s,Q). It is located just below D_1 and corresponds to the arrival of replenishment orders at the retailer when the new I_t exceeds s. D_3 consists of k blocks of s+1 lines. ith block consists of s+1-i zero lines (corresponding to the arrival of a replenishment order when the new $I_t \leq s$) and a left aligned $i \times i$ diagonal matrix of μ_2.
- D_4 occupies the left down corner of D. It occurs only when $Q > s$ and corresponds to replenishment orders where $T_t > s$. D_4 is a $(s+1) \cdot f \times Q$ sub-matrix, where f= max(Q-s,0). It can be divided to Q-s, left aligned diagonal blocks of μ_2. Each block has dimension $(s+1) \times (s+1)$ and each subsequent block is located one column to the right.

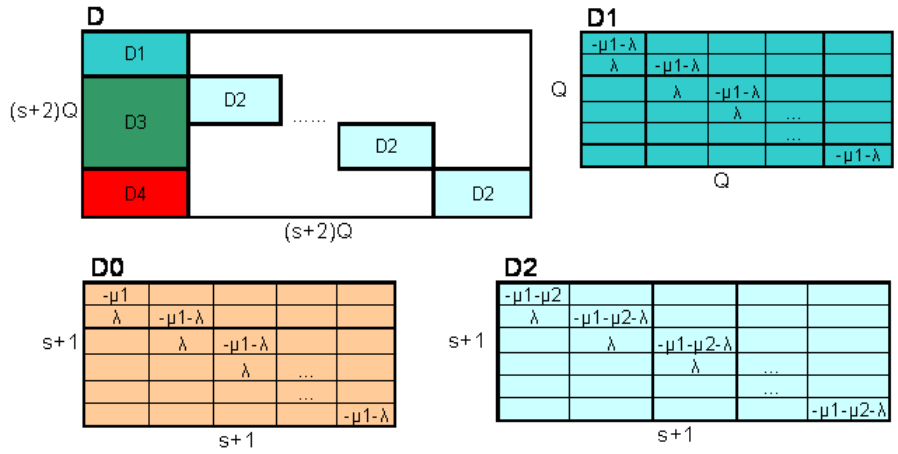

Figure 4 Examples of sub-matrices.

The last sub-matrix D corresponds to the boundary states where $B_t = B + 1$. In these states station S_1 is blocked and therefore there can be no arrivals at buffer. Consequently $-\mu_1$ is subtracted from the diagonal elements of the last D sub-matrix.

Upper diagonal sub-matrices correspond to arrivals from S_1 to buffer. Since S_1 processes one unit at a time, only transitions to adjacent levels occur. The first upper diagonal sub-matrix U_0 corresponds to arrivals at buffer, while the system is at the boundary states where $T_t = 0$, $I_t < s$ and $B_t = 0$. In such cases, the arriving unit is immediately forwarded for transportation to the retailer. U_0 is a $(s + 1) \times (s + 1)$ diagonal block of μ_1. U is the repeating upper diagonal block. It is a $(s + 2) \cdot Q \times (s + 2) \cdot Q$ diagonal matrix of μ_1. It is repeated B+1 times, corresponding to different levels.

The repeating block L below the diagonal describes transitions where there is triggering of replenishment orders. The events that can trigger replenishment orders are a) the occurrence of external demand while $I_t = s + 1$ and b) the arrival of a replenishment order when the updated I_t does not exceed s. For s=0, $L = \lambda$. For s>0, L is a $(s + 1) \cdot k + Q \times s$ block, where k=min(s,Q). For such cases, block L has the following structure: The element (1,s)=λ; then follow Q-1 zero lines; then follow k, $(s + 1) \times s$ blocks where the ith block is a $(s + 1 - i) \times (s + 1 - i)$ right aligned diagonal matrix of μ_2 followed by i zero lines.

L is the most complicated sub-matrix since it corresponds to transitions between non-adjacent levels. Sub-matrix L occurs B+2 times and its exact positioning in the transition matrix P depends on the parameters B,s and Q. The first occurrence of L corresponds to the boundary conditions where the buffer is empty. The next h = min(B+1,Q) occurrences correspond to cases where $B_t \leq Q$. Finally, the next B+1-h occurrences correspond to initiation of complete orders when $B_t > Q$.

3.4 Performance Measures

From the transition matrix, the corresponding system of linear equations can be determined and the vector of stationary probabilities can be computed. Using the stationary probabilities, we can calculate performance measures for the system under consideration. Again, we take advantage of the structure of the transition matrix. To illustrate through an example, for $B = 2$, $s = 1$, $Q = 2$, the percentage of external demand that is met from inventory on hand:

$$\begin{aligned} \text{Fill Rate retailer} &= P(I_t > 0) = 1 - P(I_t = 0) = \\ &= 1 - (\pi_{000} + \pi_{010} + \pi_{020} + \pi_{110} + \pi_{120} + \pi_{210} + \pi_{220} + \pi_{310} + \pi_{320}) . \end{aligned}$$

In general, taking advantage of the transition matrix structure,

$$FillRate = 1 - \pi_1 - \sum_{j=0}^{B+1} \sum_{i=0}^{Q-1} \pi_{r+i\cdot(s+1)} ,$$

where $r = s + Q + 2 + j \cdot (s + 2) \cdot Q$, and π_i is the ith element of the stationary probability vector π. It is reminded that the sequence of system states is defined according to certain rules as expounded earlier.

Similarly, for the average inventory at buffer (Work in Process Buffer or WIP Buffer), including blocked units:
Inventory buffer $= c \otimes b$, where:
c: is a line vector with the possible positive values of B_t In our example c= [1, 2, 3], and generally c=[1, 2, . . . ,B+1].
b: is a column vector with the ith element giving the probability that i units belong to the buffer. In our example $b^\tau = [(\pi_{102} + \pi_{103} + \pi_{110} + \pi_{111} + \pi_{120} + \pi_{121})(\pi_{202} + \pi_{203} + \pi_{210} + \pi_{211} + \pi_{220} + \pi_{221})(\pi_{302} + \pi_{303} + \pi_{310} + \pi_{311} + \pi_{320} + \pi_{321})]$

From the matrix structure, and especially from the levels we have defined according to B_t, it can be inferred that for the j+1 element of vector b

$$b_{j+1} = \sum_{i=r}^{r+(s+2)\cdot Q-1} \pi_i ,$$

where $r = (s + 2)(1 + Q + j \cdot Q)$
so that b can be calculated and average WIP buffer can be computed as the product $c \otimes b$.

In a similar way we can also calculate the rest performance measures of concern, namely average inventory at retailer (WIP retailer), percentage blocked for S_1, and average inventory in transit from the buffer to the retailer (WIP in transit).

4 Verification

The validity of the algorithm is verified using simulation. A simulation model of the system described in section 2 is developed using Arena© simulation package. An approach similar to the cycle view of supply chains is adopted. The system is modelled using three cycles corresponding to the interfaces between Buffer, Transportation and Retailer. The results of the algorithm described in section 3 are collated with simulation results for the same system parameter values and the two approaches are found to give practically identical results. A simulation time of 1000000 time units was selected as it was deemed long enough to provide statistically vigorous results. Moreover, a warm-up period of 20000 time units was selected, so as to eliminate the effect of the initial conditions. Figure 5 gives the comparison of analytic and simulation solutions for performance measure Fill Rate across various scenarios. The parameters of the system are: $\mu_1 = 1$, $\mu_2 = 0.5$, $\lambda = 1$, $0 \leq B \leq 10$, $0 \leq s \leq 10$, and $0 \leq Q \leq 11$. % Fill Rate Difference $= \frac{FR_{Simulation} - FR_{Analytic}}{FR_{Analytic}}$. The difference does not exceed 0.5 %, well within the limits of the expected variability due to the statistical nature of simulation results.

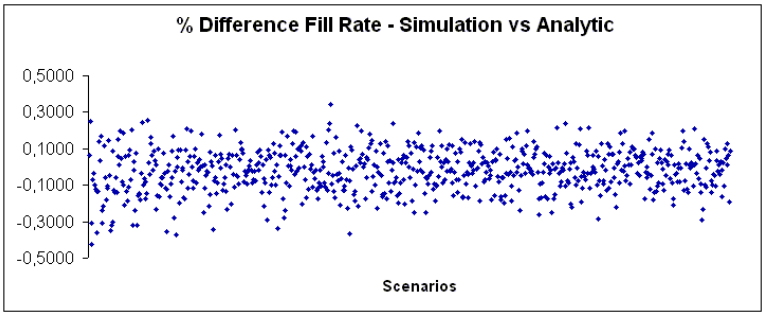

Figure 5 Simulation vs Analytic results for various scenarios.

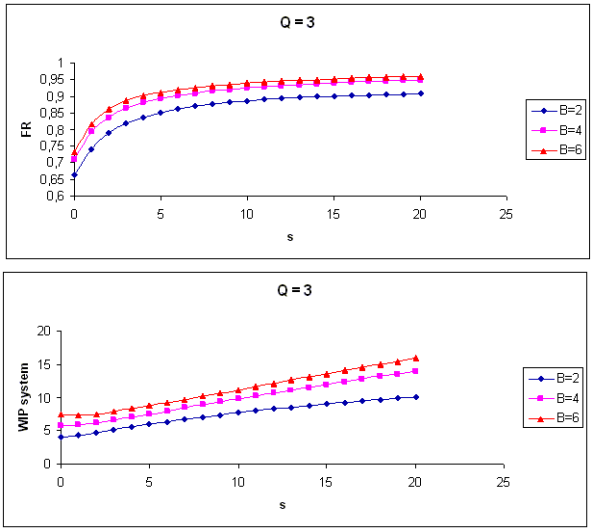

Figure 6 The effect of s on Fill rate and WIP system for different B levels.

5 Results

We use the analytic model to investigate the effect of each parameter B, s and Q on the performance measures of the system. We choose a balanced system, where $\mu_1 = \mu_2 = \lambda = 1$. Costs in a supply chain are mainly associated with holding inventories and lost sales, so we focus our analysis on the performance measures Fill Rate and average inventory in the system (WIP System). In push-type systems costs are also associated with blocking, in terms of lowered utilization and production disruptions, but at this stage blocking is not investigated. Figure 6 gives the effect of s on the performance measures for different levels of B and constant Q. For a given level of B, increasing s causes an almost linear increase in the average system inventory (Work in Process, or WIP system). At the same time, the improving effect on fill rate diminishes with increasing s. From the model it is inferred that beyond a point, changing the value of s would not be an advisable strategy for the improvement of supply chain performance.

Figure 7 gives the effect of B on the performance measures of concern for different levels of Q and constant s. We can see that for a range of values of B there are different (B,s,Q) policies yielding different levels of Fill Rate for approximately the same level of average system inventory (WIP system). In such cases there is a potential of enhancing supply chain

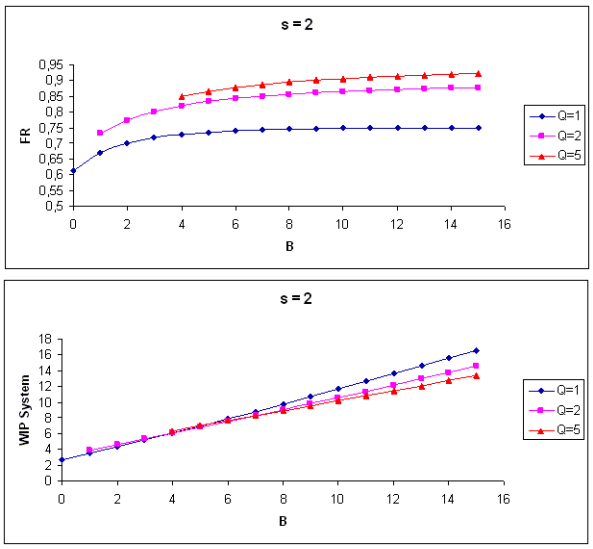

Figure 7 The effect of B on Fill Rate and WIP system for different levels of Q.

performance without incurring further costs. In a real situation, decision makers should further experiment in order to define the optimal policy within the given constraints. It should be noted that a change of policy would have different effects on the different members of the supply chain. For example, for $B = 5$, changing the value of Q from 1 to 5 would increase Fill rate significantly, but the average inventory (and thus the corresponding costs) at the retailer would also increase. However, such an increase would be compensated by the decrease of the average inventory at the buffer so that the average inventory in the system remains the same. On the whole, the performance of the supply chain can be improved, but a "global" viewpoint and centralised decision making would be required.

6 Conclusions – Future research

In our work we developed an exact algorithm for the analysis of a simple, serial, push-pull supply chain. The proposed descriptive model captures relationships between variables, offers insight on key features of the system at hand, and can be used as a design tool for the evaluation of appropriate systems and the determination of their optimal characteristics. By extensive enumeration and evaluation of the possible policies (B,s,Q), the optimal policy that will minimize average system inventory for a given threshold value of Fill Rate, or that will maximize Fill Rate for a given maximum average system inventory, can be determined.

As indicated by the results for a balanced system, all parameters B, s and Q can have an impact on system performance. The relative importance of each depends on the specific range of its values. With regard to the average inventory in the system, Q seems to have a lesser effect since changes in the average inventory at the retailer are counterbalanced by the changes in the average buffer inventory. On the other hand, an increase in s or B causes an almost linear increase in average system inventory. With regard to Fill Rate, increasing any of the parameters B,s,and Q improves system performance. However, changing only one of the parameters, there is a threshold Fill Rate value that cannot be exceeded, and the effect of each parameter diminishes as this value is approached. Due to the dynamic nature of the system, effective decision making should take into account the effect of each policy on the whole supply chain, since a local view may lead to sub-optimal solutions.

In a further step of our research, the algorithm can be expanded to include different demand characteristics (for example compound Poisson) or longer chains. Members may be added either upstream (push segment), or downstream (pull segment). The use of phase type distributions (Erlang, Coxian) instead of exponential distribution could also be a possible object of further research.

References

1 R. D. Badinelli. A model for continuous-review pull policies in serial inventory systems. *Operations Research*, 40(1):142–156, 1992.

2 M. Bijvank and I. Vis. Lost-sales inventory theory: A review. *European Journal of Operational Research*, 215:1–13, 2011.

3 F. Chen. Optimal policies for multi-echelon inventory problems with batch ordering. *Operations Research*, 48(3):376–389, 2000.

4 S. Chopra and P. Meindl. *Supply Chain Management, Strategy, Planning & Operations, 3rd edition*, 44–72. Pearson International, 2007.

5 A. J. Clark and H. Scarf. Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4):475–490, 1960.

6 J. K. Cochran and S. S. Kim. Optimum junction point location and inventory levels in serial hybrid push/pull production systems. *International Journal of Production Research*, 36(4):1141–1155, 1998.

7 E. Cuypere, K. Turck and D. Fiems. A Queueing Theoretic Approach to Decoupling Inventory. Analytical and Stochastic Modeling Techniques and Applications. In proceedings *19th International Conference, ASMTA 2012, Grenoble, France*. 150–164, 2012.

8 O. Ghrayeb, N. Phojanamongkolkij and B. A. Tan. A hybrid push/pull system in assemble-to-order manufacturing environment. *Journal of Intelligent Manufacturing*, 20(4):379–387, 2009.

9 D. Gupta and N. Selvaraju. Performance Evaluation and Stock Allocation in Capacitated Serial Supply Systems. *Manufacturing and Service Operations Management*, 8(2):169–191, 2006.

10 G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM Series on statistics and Applied probability, 1999.