Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

EVALUATING MIXED INTEGER PROGRAMMING MODELS FOR SOLVING STOCHASTIC INVENTORY PROBLEMS

Bas Bluemink
A. G. de Kok

Eindhoven University of Technology
School of Industrial Engineering
Eindhoven, 5600 MB, THE NETHERLANDS

Balan Srinivasan
Reha Uzsoy

North Carolina State University
E. P. Fitts Dept. of Industrial & Systems Engineering
Raleigh, NC 27695-7906, USA

ABSTRACT

We formulate mixed integer programming (MIP) models to obtain approximate solutions to finite horizon stochastic inventory models. These deterministic formulations of necessity make a number of simplifying assumptions, but their special structure permits very short model solution times under a range of experimental conditions. We evaluate the performance of these models using simulation optimization to estimate the true optimal solutions. Computational experiments identify several demand and cost scenarios in which the MIP models yield near-optimal solutions, and other cases where they fail, suggesting directions for future research.

1 INTRODUCTION

Stochastic inventory models have been the subject of extensive study since the early work of Arrow et al. (1951), and remain one of the most active areas of investigation in the fields of operations management and production systems. However, the computation of optimal inventory policies remains challenging even for apparently simple inventory systems such as single-stage periodic review models with independent demands. While it is well known that a base stock policy is optimal (Clark and Scarf 1960), the only general approach to computing the optimal base stock levels is stochastic dynamic programming, which suffers from the curse of dimensionality to a degree that renders even the solution of quite small instances extremely challenging. Non-stationary demand poses particular challenges for computing optimal policies. As a result, a natural direction for research is the development of heuristics that can consistently provide near-optimal solutions.

The problems of production planning, particularly those of planning the releases of work into capacitated production systems over time, have also been addressed extensively in the literature. The most prevalent mathematical approach to these problems has been deterministic mathematical programming models, which have evolved to a widely accepted common structure. A finite planning horizon is divided into discrete time periods, and decision variables (generally release quantities) are associated with each period. The decision variables are used to compute the values of state variables such as inventory and backorder levels associated with each period, which allow costs to be calculated. The vast majority of this work treats all problem parameters as deterministic. It is interesting to note that developments in stochastic optimization, such as multistage stochastic programming (Birge and Louveaux 1997), simulation optimization (Fu 2015), and robust optimization (Bertsimas and Sim 1994) have as yet had relatively little impact on the production planning domain.

It is interesting that despite their apparent common origin in the problems of supply chain planning and management, research in the areas of production planning and inventory management has evolved almost independently of each other. Inventory theory has focused on modeling the stochastic aspects of demand, largely ignoring problems of capacity. Production planning, on the other hand, has largely avoided problems of stochastic demand to focus on modeling material flows and capacity constraints, as well as lot-sizing issues. While safety stocks are a primary concern in the domain of inventory models, production planning models generally treat them as an exogenous parameter. Attempts to integrate the planning of safety stocks into production planning models are relatively recent. de Kok (2018) provides a simulation-based methodology to determine optimal end-item safety stocks for mathematical-programming-based production planning models.

Due to its heavy reliance on mathematical programming models, the domain of production planning has benefited immensely from the developments in commercial mathematical programming solvers over the last four decades. This suggests the possibility of exploiting these developments to build mathematical programming models that obtain approximate solutions to finite-horizon stochastic inventory problems. In this paper we present the results of an initial effort in this direction. The mixed integer programming (MIP) models we propose share some characteristics with stochastic dynamic programming, in that the demand distribution is discretized, yielding a very large number of binary variables. However, the special structure of the resulting integer programs, together with improvements in solvers and computing power, allow them to be solved in reasonable computation times.

The following section briefly reviews previous related work. Section 3 presents the formal problem statement and the mathematical programming formulation, followed by the design of the computational experiments in Section 4. Section 5 presents the results of our experiments, and Section 6 concludes the paper with a discussion of the primary findings and their implications for future research.

2 PREVIOUS RELATED WORK

The first paper on determination of optimal base-stock levels in a serial supply chain was by Clark and Scarf (1960). They developed an algorithm to solve for the exact base-stock levels in a finite-horizon system and proved that this policy is indeed optimal. Federgruen and Zipkin (1984) proved the optimality of the base-stock policy for infinite-time serial supply chains. Chen and Zheng (1994) extended the method of Clark and Scarf and developed a recursive solution approach. Diks and De Kok (1998; 1999) developed an algorithm to compute near-optimal order-up-to policies for divergent, multi-echelon systems. Shang and Song (2003) developed a heuristic based on the solution of a number of newsvendor subproblems to calculate approximately optimal base-stock levels for infinite-horizon serial systems by solving multiple instances of the classic newsvendor model Snyder and Shen (2011) developed by Arrow, Harris, and Marschak (1951). The heuristic is fast and easy to implement but does not guarantee optimality, although extensive computational experiments demonstrate excellent performance. An overview of extensions to the newsvendor model related to influenced customer demand, supplier pricing policies, and the risk profile of buyers is given by Qin et al. (2011). Other extensions include multi-product systems, extensively treated by Choi (2012), and multi-product systems under constraints, treated by Niederhoff (2007), Zhang (2010), and Abdel-Malek and Otegbeye (2013). Of these, Niederhoff (2007) and Abdel-Malek and Otegbeye (2013) use a separable programming approach that is particularly relevant to the work in this paper.

Parker and Kapuscinski (2004) show that a modified echelon base-stock policy is optimal for a 2-stage serial supply chain that assumes a lower capacity at the most downstream stage and is applicable to both infinite- and finite-horizon problems. Glasserman and Liu (1997) analyzed the multi-stage serial system with capacity constraints under stochastic demand, for which they developed approximations of average inventory, backorders, unfilled demand, and shortfalls.

In summary, several solution methods exist to obtain both exact and approximate base-stock levels in serial supply chains with and without capacity constraints. However, time-dependent capacity constraints and non-stationary demand have not been considered. Non-stationary demand can occur in many cases, e.g., seasonal demand, phase out of a product, or new product introduction. Fluctuations in capacity can occur, e.g., when repairs must be planned or available storage space varies. In this paper, we investigate whether it is possible to incorporate time-dependent capacity constraints and non-stationary demand into the optimization using mathematical programming models.

3 MIXED INTEGER PROGRAMMING FORMULATION

We consider a single-echelon single-item capacitated inventory system under periodic review over a finite planning horizon. We define the following notation, noting that period t runs from time $t-1$ to time t.

T	The set of discrete planning periods $t = 1, ..., \|T\|$
S_t	Base-stock level in period t
h_t	Holding cost at time t
p_t	Penalty cost for backorders at time t
$C_T(S_t)$	Total cost of system for a given set of echelon base stock values S_t
Δ	Width of each discretization interval
L	Number of discretization intervals
D_t	Demand at time t
$f_t(.)$	Probability density function of demand in period t
X_t	Number of units ordered at the end of period $t-1$ which arrive at the end of period t, right after the incoming demand D_t has been fulfilled. Thus, X_t cannot be used to meet demand D_t.
μ_t	Mean demand in period t
I_t	On-hand inventory at the end of period t
B_t	Backorders at the end of period t
Z_t	Net inventory at the end of period t, given by $I_t - B_t$; note that X_t is not included in Z_t
Y_t	Inventory order position at time t, given by Z_t + all outstanding orders
Q_t	Shortfall in period t, defined as $Q_t = max\{0, S_t - Y_t\}$
C_t	Capacity in period t

Demand in each period is independent, but not identically distributed. In each period t the system can replenish its inventory position by ordering up to a maximum of C_t units. There are no fixed ordering costs. At the end of each period, holding costs are incurred equal to h_t times the number of units on hand, while penalty costs incurred equal to p_t times the number of units backlogged. Federgruen and Zipkin (1986) have shown that the optimal policy for this problem has the form of a modified base stock policy, under which a base stock level S_t is determined for each period, and orders placed to bring the inventory position Y_t up to this level. However, due to the capacity constraint, it may not be possible to bring the inventory position up to the base stock level S_t. The amount by which the base stock level exceeds the inventory position, given by $Q_t = max\{0, S_t - Y_t\}$, is referred to as the shortfall for period t.

Our mathematical programming formulation follows previous stochastic dynamic programming and separable convex programming approaches in discretizing the demand distribution in each period using L discretization intervals of with Δ. Binary decision variables are then used to emulate numerical integration in the evaluation of the objective function. Auxiliary binary variables and constraints are implemented to ensure that a base stock policy is correctly implemented. We begin by defining an objective function suitable for the finite-horizon problem, followed by material balance equations and related constraints. Additional constraints then ensure that order releases follow a modified base stock policy. The starting point of our formulation is to express the base stock level S_t associated with period t in terms of the discretization intervals, as

$$S_t = \sum_{l=1}^{L} \Delta y_{l,t}, \tag{1}$$

where Δ denotes the width of the discretization interval and $y_{l,t}$ is a binary decision variable such that

$$y_{lt} = \begin{cases} 1 \ \textit{if } S_t \geq l\Delta \\ 0 \ \textit{otherwise} \ . \end{cases}$$

To ensure the correct definition of the base stock level S_t, we enforce the constraint

$$y_{l-1,t} \geq y_{l,t} \qquad l = 2, ...L, \forall t \in T. \tag{2}$$

This constraint set is particularly important to the computational tractability of the MIP formulation, and captures the order up to structure of the base stock policy. We shall now use this definition of S_t to formulate the objective function and constraints of our MIP model. For clarity of exposition we shall not substitute the expression (1) for the base stock into each expression as we describe the model in more detail, allowing the reader to see the parallels with the stochastic inventory model more clearly.

3.1 Objective Function

Since we consider only linear holding and penalty costs, we approximate the objective function by assuming that at the start of each period the inventory position Y_t is raised to its (time-dependent) base stock level S_t. This formulation neglects the possibility that it may not be possible to raise the inventory position to the desired base stock level due to the limited capacity C_t, i.e., the possibility of positive shortfall $Q_t > 0$. However, the constraints of the mathematical program explicitly recognize that this may not be possible due to capacity constraints (inducing positive shortfall), or to excess inventory carried from previous periods. By discretizing the time-dependent demand probability functions and summing over all time periods the objective function can be expressed as follows:

$$C_T(S_t) = \sum_{t=0}^{T} \left\{ h_t \sum_{l=1}^{L} (S_t - l\Delta) P_{l,t} + (p_t + h_t) \sum_{l=1}^{L} \{ (l\Delta - S_t) P_{l,t} \} (1 - y_{lt}) \right\}, \tag{3}$$

where

$$P_{l,t} = \frac{\Delta}{2} (f_t((l-0.5)\Delta) + f_t((l+0.5)\Delta)). \tag{4}$$

This cost function essentially implements the classical newsvendor cost function in each period t (Snyder and Shen 2011) using Equation (4) to implement the trapezoidal rule for numerical integration. The first summation inside the braces computes the expected holding cost of the net inventory, and the second the expected number of missed demands.

3.2 Material Balance Equations

The net-inventory at the end of each period t is the sum of the net-inventory of the previous period plus the amount ordered minus the demand in this period. Therefore, the material balance equation (5) must always hold at the end of every period.

$$Z_t = Z_{t-1} + X_{t-1} - D_t \tag{5}$$

However, since the period demand D_t is a random variable, this constraint cannot be implemented in this form in a mathematical programming model. In a scenario-based stochastic programming approach, this would be enforced using a number of demand realizations (scenarios), but this results in exponential growth in the size of the formulation with the number of realizations and planning periods. Hence, we take the approach of enforcing Equation (5) in expectation, yielding the deterministic material balance constraint

$$Z_t = Z_{t-1} + X_{t-1} - \mu_t \tag{6}$$

Finally, since the net inventory Z_t can take negative values and the mathematical programming model requires non-negative variables, we use the relation $Z_t = I_t - B_t$ to obtain the final form of the material balance constraint given by

$$I_t - B_t = I_{t-1} - B_{t-1} + X_{t-1} - \mu_t \tag{7}$$

If demand rises over time, the inventory can only be built up to the maximum that the capacity allows. Therefore, the base-stock level cannot increase by more than the capacity C_t from one period to the next:

$$S_t \leq S_{t-1} + C_t \tag{8}$$

We also assume that the sum of the previous period's net inventory and the current period's order will always equal the base-stock level (cf. our discussion on the objective function above):

$$Z_{t-1} + X_t + X_{t-1} = S_t \tag{9}$$

A limitation of Equation (9) is that if demand drops after time t; the base-stock level at time t must be at least Z_{t-1}. To allow the base-stock level at time t to go below this, Equation (9) is relaxed if $X_t = 0$ and $S_t < S_{t-1} - \mu_{t-1}$. This is accomplished by introducing three new auxiliary binary decision variables defined as follows:

$$z_t = \begin{cases} 1 \ \textit{if } X_t = 0 \\ 0 \ \textit{if } X_t > 0 \end{cases} \tag{10}$$

$$u_t = \begin{cases} 1 \ \textit{if } S_t < S_{t-1} - \mu_{t-1} \\ 0 \ \textit{if } S_t \geq S_{t-1} - \mu_{t-1} \end{cases} \tag{11}$$

$$v_t = \begin{cases} 1 \ \textit{if } z_t u_t = 1 \\ 0 \ \textit{if } z_t u_t = 0 \end{cases} \tag{12}$$

Constraint (9) can now be implemented using the constraints

$$Z_{t-1} + X_t + X_{t-1} \geq S_t \tag{13}$$

$$Z_{t-1} + X_t + X_{t-1} \leq S_t + v_t M \tag{14}$$

where M represents a very large number. If $v_t = 0$, Equations (13) and (14) enforce Equation (9). However, if $v_t = 1$, the right hand side of Equation (14) becomes sufficiently large that the constraint is no longer binding. The following constraints then ensure the correct system dynamics:

$$(1 - z_t) \leq X_t \qquad \forall t \in T \tag{15}$$

$$(1 - z_t)M \geq X_t \qquad \forall t \in T \tag{16}$$

$$S_t + u_t M \geq S_{t-1} - \mu_{t-1} \qquad \forall t \in T \tag{17}$$

$$S_t < S_{t-1} - \mu_{t-1} + (1 - u_t)M \qquad \forall t \in T \tag{18}$$

$$v_t \leq z_t \qquad \forall t \in T \tag{19}$$

$$v_t \leq u_t \qquad \forall t \in T \tag{20}$$

$$v_t \geq z_t + u_t - 1 \qquad \forall t \in T \tag{21}$$

The first two constraint sets link the values of the order quantity X_t to the values of the binary decision variable z_t. These two constraints together ensure that $X_t \geq 0$. The third and fourth constraint sets allow

u_t to take the correct values based on the relation between $s_t, S_{t=1}$ and μ_t. The final three constraint sets ensure that the binary decision variables z_t take values consistent with the values of u_t and v_t.

Finally, the incoming order can never exceed the available capacity:

$$X_t \leq C_t \tag{22}$$

The final formulation thus minimizes the objective function (3) subject to the constraints

$$(2), (7), (14) - (20) \tag{23}$$

$$I_t \geq 0 \qquad \forall t \in T \tag{24}$$

$$B_t \geq 0 \qquad \forall t \in T \tag{25}$$

$$y_{l,t} \in \{0,1\} \qquad l = 1, ..., L, \forall t \in T \tag{26}$$

$$z_t \in \{0,1\} \qquad \forall t \in T \tag{27}$$

$$u_t \in \{0,1\} \qquad \forall t \in T \tag{28}$$

$$v_t \in \{0,1\} \qquad \forall t \in T \tag{29}$$

The size of this formulation, and hence the computation time required for its solution, evidently depend on several factors, whose values must be determined experimentally. These include the discretization unit Δ and the number L of discretization intervals used. The formulation involves $O(LT)$ decision variables and $O(T)$ constraints. Thus, the size of the MIP instance that must be solved is primarily determined by the number of discretization intervals used to approximate the objective function.

The formulation described above is an approximation to the stochastic inventory problem that we wish to solve. The first approximation takes place in the objective function (3), which assumes that the inventory position can always be brought up to the base stock level S_t; the probability of positive shortfall is neglected. This can be expected to adversely effect the quality of the solution proposed by the formulation when capacity constraints are binding a significant proportion of the time. The second approximation is the assumption in the material balance Equation (6) that demand will always be realized at its mean level. This is necessary to avoid the exponential growth in formulation size that would result from using specific demand realizations, i.e., scenarios, but may result in suboptimal base stock levels. Our computational experiments in the next section examine the degree to which these approximations affect the quality of the solutions obtained from the formulation.

4 EXPERIMENTAL DESIGN

The objective of these exploratory experiments is to assess the potential of the MIP model presented in the previous section to generate near-optimal solutions to the problem, and to develop an understanding of the reasons for poor performance so that they can be enhanced and improved. The finite horizon consists of 10 periods, with $t \in T, T = \{0, 1, ..., 9\}$, and $t = 0$ the time at which the calculation of future base-stock levels is executed. In order to define the cost parameters in an intuitive manner, we set the values of p and h to yield the specified service level (probability of no stockout) α in an uncapacitated newsvendor problem, i.e., $\frac{p}{p+h} = \alpha$. The following parameters are varied:

- Service level: 90% or 95%
- Squared coefficient of variation (CV^2) of period demand: 0.5, 1, or 2
- Mean period demand: constant, rising, decreasing, or constant with one peak. The variations used for the mean demand per period are given in Table 1.
- Capacity per period: low, high, unlimited, or high with sudden decrease. For the unlimited capacity case, a value must be defined as input, for which the value of 50 is chosen. The value 50 represents unlimited capacity in this case, as it is very seldom reached with the demand levels currently proposed. The variations used for the capacity per period are given in Table 2.

Table 1: Mean demand profiles.

Variation \ t	1	2	3	4	5	6	7	8	9	Mean
Constant	9	9	9	9	9	9	9	9	9	9
Rising	5	6	7	8	9	10	11	12	13	9
Decreasing	9	10	11	13	11	9	7	6	5	9
Peak	7	7	7	7	10	12	17	7	7	9

Table 2: Capacity profiles.

Variation \ t	1	2	3	4	5	6	7	8	9	Mean
Low	10	10	10	10	10	10	10	10	10	10
High	12	12	12	12	12	12	12	12	12	12
Unlimited	50	50	50	50	50	50	50	50	50	50
Sudden decrease	13	13	13	13	4	4	13	13	13	11

Rather than presenting a full factorial design, we focus on the 15 cases presented in Table 3. Cases 1–4 are generated to examine the behavior of the MIP model under different capacity levels. Cases 5–12 are generated to test the behavior with capacity constraints, while Cases 13–15 examine the effect of the value of CV^2.

The MIP formulation was solved using CPLEX 12; solution times for the largest formulations were of the order of 40 minutes, although times less than 5 minutes were typical. No attempt was made to enhance the formulation using valid inequalities, and CPLEX was run with all settings at their default values. The performance of the base stock levels obtained from the MIP model is evaluated by simulating their execution for 2000 replications using Microsoft Excel.

Due to the difficulty of implementing an exact stochastic dynamic programming approach in the time available to us, we use a simulation optimization approach to generate near-optimal solutions as a benchmark for the MIP model. The search mechanism employed is a Genetic Algorithm (GA) implemented using the Standard Evolutionary Engine in the Frontline Solver add-in to Microsoft Excel. Each chromosome corresponds to a vector of base stock levels $S_t, t = 1, ..., 10$. The algorithm was implemented using a population size of 200, with the fitness of each chromosome being measured by the average cost of 2,000 independent simulation replications. A single Excel worksheet is implemented to simulate the performance of a given vector of base stock levels, which are specified as the decision variables over which the GA searches. Initial experiments sought to minimize the average cost over all 2,000 replications. However, it was noted that in some instances the evolutionary algorithm generated very large base stock levels that were clearly redundant in the presence of the capacity constraints. To prevent this behavior, a penalty term equal to 5% of the sum of the base stock levels in all periods was added to the average cost over the simulation replications, and used as the objective function for the simulation optimization. However, the results reported in the next section do not include the penalty term. A maximum computation time limit of 7,200 seconds was used for all runs, with the algorithm terminating if the current best solution was not improved in 30 minutes. The average run time for the simulation optimization procedure was of the order of one hour on a MacBook Pro with a 2.5 GHz Intel Core i7 processor and 16GB RAM.

We recognize that the solutions provided by the simulation optimization procedure cannot be guaranteed to be optimal. However, there are several results in the literature proving that a GA will, when run for sufficiently long time under the right parameterization, reach a global optimal solution with probability 1 (Goldberg 1989). Thus, we use a very long run of the GA as an admittedly imperfect surrogate for the exact optimal solution to the stochastic optimization problem.

Table 3: Experiment cases.

Case	Service level	CV^2	t:	1	2	3	4	5	6	7	8	9
1	95%	0.5	μ_t:	5	6	7	8	9	10	11	12	13
			C_t:	50	50	50	50	50	50	50	50	50
2	95%	0.5	μ_t:	9	9	9	9	9	9	9	9	9
			C_t:	50	50	50	50	50	50	50	50	50
3	95%	0.5	μ_t:	7	7	7	7	10	12	17	7	7
			C_t:	50	50	50	50	50	50	50	50	50
4	95%	0.5	μ_t:	9	10	11	13	11	9	7	6	5
			C_t:	50	50	50	50	50	50	50	50	50
5	90%	1	μ_t:	9	9	9	9	9	9	9	9	9
			C_t:	10	10	10	10	10	10	10	10	10
6	90%	1	μ_t:	9	9	9	9	9	9	9	9	9
			C_t:	13	13	13	4	4	13	13	13	13
7	90%	1	μ_t:	5	6	7	8	9	10	11	12	13
			C_t:	10	10	10	10	10	10	10	10	10
8	90%	1	μ_t:	5	6	7	8	9	10	11	12	13
			C_t:	12	12	12	12	12	12	12	12	12
9	90%	1	μ_t:	9	10	11	13	11	9	7	6	5
			C_t:	10	10	10	10	10	10	10	10	10
10	90%	1	μ_t:	9	10	11	13	11	9	7	6	5
			C_t:	12	12	12	12	12	12	12	12	12
11	90%	1	μ_t:	7	7	7	7	10	12	17	7	7
			C_t:	10	10	10	10	10	10	10	10	10
12	90%	1	μ_t:	7	7	7	7	10	12	17	7	7
			C_t:	12	12	12	12	12	12	12	12	12
13	95%	0.5	μ_t:	5	6	7	8	9	10	11	12	13
			C_t:	10	10	10	10	10	10	10	10	10
14	95%	1	μ_t:	5	6	7	8	9	10	11	12	13
			C_t:	10	10	10	10	10	10	10	10	10
15	95%	2	μ_t:	5	6	7	8	9	10	11	12	13
			C_t:	10	10	10	10	10	10	10	10	10

5 COMPUTATIONAL RESULTS

We present the experimental results in two separate sections. In the first set of experiments, shown in Table 5, we examine the impact of different capacity levels on the performance of the MIP formulation. Each row of the table gives the average cost of the policy suggested by the MIP model (Avg.), its standard deviation (SD), and the lower and upper limits of a 95% confidence interval for the mean cost (LL and UL, respectively). The C_t column gives the capacity, i.e., the maximum order size that can be shipped in any period. The final column gives the ratio of the average cost of the MIP policy to that obtained using the GA, which we use as a surrogate for the optimal value. Ratios below 1 indicate cases where the simulation optimization was not able to improve on the MIP solution; this is probably due to becoming trapped in a local optimum, or the presence of the term penalizing high base stock levels in the GA.

It is immediately apparent that when $C_t = 50$, i.e., capacity constraints are not binding, the MIP produces solutions of essentially the same quality as the simulation optimization procedure. Under this condition, the assumption in the objective function that the base stock can always be achieved is valid, although the assumption of mean demand in the material balance equation (6) is not. When $C_t = 12$, the performance of the MIP again improves, except for Case 3. Case 3 experiences a peak in mean demand in periods 6 through 8, requiring inventory to be accumulated ahead of this temporary bottleneck. In the intermediate capacity cases of $C_t = 11$ or 12, the maximum deviation of the MIP solution from that obtained by simulation optimization is 15%, which occurs, rather surprisingly, in Case 2 where the demand distribution remains constant in all periods. It is interesting to observe that the simulation optimization procedure consistently yields higher base stock levels than the MIP, despite the fact that high base stock levels are explicitly penalized in the fitness measures used in the simulation optimization procedure.

Table 4: Results for Cases 1 through 4.

Case	C_t	Avg.	SD	LL	UL	MIP/GA
1	50	194.23	121.99	189.74	198.72	0.98
	10	343.24	495.63	325	361.49	1.15
	11	278.12	365.71	264.66	291.58	1
	12	257.45	346.72	244.69	270.21	1.03
2	50	206.15	122.24	201.65	210.64	1
	10	311.84	481.07	294.13	329.54	1.08
	11	290.88	456.8	274.07	307.69	1.15
	12	254.98	342.49	242.38	267.59	1.02
3	50	214.21	131.5	209.37	219.05	0.99
	10	392.3	578.44	371.01	413.59	1.09
	11	347.57	493.85	329.4	365.75	1.08
	12	326.79	476.65	309.25	344.34	1.11
4	50	218.06	134.02	213.13	223	1.02
	10	321.8	581.79	300.39	343.21	0.98
	11	292.58	503.1	274.07	311.1	1.07
	12	269.36	419.6	253.91	284.8	1

The results for Cases 5 through 15 are shown in Table 5. The CAP/GA column gives the performance of a policy that orders the capacity in each period, i.e., $X_t = C_t$ for all periods t. In this case, the maximum error, except for Case 10, is 13%, suggesting that overall the MIP formulation produces solutions that are very reasonable given the severity of the assumptions it makes. In Case 10, mean demand is decreasing over time, implying that the variability of demand is also decreasing since the CV is held constant across the planning horizon. Case 9 has the same demand pattern, but has C_9 = 10 while C_{10} = 12. This case is particularly difficult as in the later periods it is frequently not necessary to order, due to the higher probability of inventory remaining on hand from earlier periods. Examining the base stock levels generated, the simulation optimization has higher base stocks in the earlier periods than the MIP, and lower ones in the final periods.

The CAP/GA column also yields some interesting insights. In all cases except Cases 9 and 10, ordering up to the full capacity in each period appears to perform very well. Cases 9 and 10 are, of course, the cases with declining mean demand over time, where ordering up to capacity in the later periods will yield high levels of excess inventory. Comparison to the MIP/GA results suggests that the MIP is systematically setting its base stock levels too low, since due to the capacity constraints no base stock level will yield higher order quantities than those from the CAP policy. This bias towards lower base stock levels is to be expected from the mean demand assumption in the material balance constraints (6), which ignores the variability in demand that creates the need for safety stock. On the other hand, the objective function contains an explicit description of the demand distribution in each period, limited by the accuracy of the discretization used and the assumption that the inventory position can always be brought up to the specified base stock level S_t, which effectively relaxes the capacity constraint in the objective function.

Table 5: Computational results for Cases 5 through 15

Case	Avg.	SD	LL	UL	MIP/GA	CAP/GA
5	395.16	545.16	375.09	415.22	1.1	0.99
6	329.90	346.98	317.13	342.67	1.15	1.09
7	364.59	446.56	348.15	381.02	1.09	1.00
8	324.3	410.74	309.18	339.41	1.12	1.04
9	407.82	599.97	385.73	429.9	1.03	1.06
10	419.23	649.89	395.31	443.15	1.22	1.26
11	398.97	509.33	380.22	417.72	1.04	1.02
12	342.03	424.29	326.41	357.65	1.01	1.07
13	334.58	312.53	323.08	346.08	1.14	1.02
14	474.95	661.29	450.61	499.29	1.07	0.94
15	747.21	1133.52	705.49	788.93	1.13	0.96

6 CONCLUSIONS AND FUTURE DIRECTIONS

Taken as a whole, the results from our experiments are quite encouraging given the potentially severe approximations made by the MIP formulation. The formulation ignores the capacity constraint in the objective function, but enforces it in expectation through the material balance constraints linking the periods. In contrast, the objective function has a complete description of the demand distribution, but the material balance constraints use only the mean demand in each period. The generally good results obtained suggest that the trade-offs made in modeling the constraints and objective function compensate for each other to some degree, although not enough to avoid poor performance under some experimental conditions, notably the decreasing demand scenarios of Cases 9 and 10. However, we are well aware that the results reported here are exploratory in nature. A larger computational experiment with a better benchmark for the optimal solution, ideally an exact optimal solution, is necessary to draw generalizable conclusions.

The MIP models have proven to be computationally tractable, allowing relatively large instances to be solved in reasonable, although not always very short, CPU times. The computational efficiency of the formulation is in large part due to the consecutive ones structure implied by constraints (2), but we have made no other effort to enhance the formulation. Improving the computational efficiency of the formulation by developing valid inequalities and formulating tighter constraints remains an interesting direction for future research.

The exploratory results presented here raise a number of interesting questions for future work. The simulation optimization procedure seems to suggest that there are a large number of alternative policies that have very similar average cost values in the neighborhood of the optimal solution. While this makes it difficult to obtain the global optimum using simulation optimization without using a large number of replications for each evaluation of the fitness measure, it also suggests that obtaining a near-optimal solution using an approximate approach may be quite practical. There is also the possibility of enhancing the MIP formulation in several ways. One of these may be to include a limited number of carefully selected demand realizations to guide the MIP model, e.g., by assuming a two-point distribution with the same mean and variance as the original demand distribution. Another enhancement might be the inclusion of chance constraints to enable a better description of demand uncertainty in the material balance constraints.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Karthick Gopalswamy of North Carolina State University for his help in preparing this paper.

REFERENCES

Abdel-Malek, L. L. and M. Otegbeye. 2013. "Separable Programming/Duality Approach to Solving the Multi-Product Newsboy/Gardener Problem with Linear Constraints". *Applied Mathematical Modelling* 37(6):4497–4508.

Arrow, K. J., T. Harris, and J. Marschak. 1951. "Optimal Inventory Policy". *Econometrica* 19:250–272.

Bertsimas, D. and M. Sim. 1994. "The Price of Robustness". *Operations Research* 52(1):35–53.

Birge, J. R. and F. Louveaux. 1997. *Introduction to Stochastic Programming*. New York: Springer.

Chen, F. and Y.-S. Zheng. 1994. "Lower Bounds for Multi-Echelon Stochastic Inventory Systems". *Management Science* 40(11):1426–1443.

Choi, T.-M. 2012. *Handbook of Newsvendor Problems: Models, Extensions and Applications*. New York, NY: Springer,

Clark, A. J. and H. Scarf. 1960. "Optimal Policies for a Multi-Echelon Inventory Problem". *Management Science* 6(4):475–490.

de Kok, T. 2018. "Inventory Management: Modeling Real-life Supply Chains and Empirical Validity". *Foundations and Trends in Technology, Information and Operations Management* 11(4):343–437.

Diks, E. B. and A. De Kok. 1998. "Optimal Control of a Divergent Multi-Echelon Inventory System". *European Journal of Operational Research* 111(1):75–97.

Diks, E. B. and A. De Kok. 1999. "Computational Results for the Control of a Divergent N-echelon Inventory System". *International Journal of Production Economics* 59(1–3):327–336.

Federgruen, A. and P. Zipkin. 1984. "Computational Issues in an Infinite-Horizon, Multiechelon Inventory Model". *Operations Research* 32(4):818–836.

Federgruen, A. and P. Zipkin. 1986. "An Inventory Model with Limited Production Capacity and Uncertain Demands I. The Average-Cost Criterion". *Mathematics of Operations Research* 11(2):193–207.

Fu, M. C. 2015. *Handbook of Simulation Optimization*. New York: Springer.

Glasserman, P. and T.-W. Liu. 1997. "Corrected Diffusion Approximations for a Multistage Production-Inventory System". *Mathematics of Operations Research* 22(1):186–201.

Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley.

Niederhoff, J. A. 2007. "Using Separable Programming to Solve the Multi-Product Multiple Ex-Ante Constraint Newsvendor Problem and Extensions". *European Journal of Operational Research* 176(2):941–955.

Parker, R. P. and R. Kapuscinski. 2004. "Optimal Policies for a Capacitated Two-Echelon Inventory System". *Operations Research* 52(5):739–755.

Qin, Y., R. Wang, A. J. Vakharia, Y. Chen, and M. M. Seref. 2011. "The Newsvendor Problem: Review and Directions for Future Research". *European Journal of Operational Research* 213(2):361–374.

Shang, K. H. and J.-S. Song. 2003. "Newsvendor Bounds and Heuristic for Optimal Policies in Serial Supply Chains". *Management Science* 49(5):618–638.

Snyder, L. V. and Z.-J. M. Shen. 2011. *Fundamentals of Supply Chain Theory*. New York: John Wiley & Sons.

Zhang, G. 2010. "The Multi-Product Newsboy Problem with Supplier Quantity Discounts and a Budget Constraint". *European Journal of Operational Research* 206(2):350–360.

AUTHOR BIOGRAPHIES

BAS BLUEMINK is an M.S. student from the faculty of Industrial Engineering & Innovation Sciences at Eindhoven University of Technology, the Netherlands. At this university, he is pursuing a M.S. degree in Manufacturing Systems Engineering, a joint program of Industrial Engineering and Mechanical Engineering. He holds a B.S. degree in Industrial Engineering and Management from University of Groningen, the Netherlands. Through internships at companies in the semiconductor and electronics industry he has gained experience working in a supply chain environment in both the Netherlands and Malaysia. His email address is b.w.bluemink@student.tue.nl.

TON DE KOK is a Professor of Quantitative Analysis of Operations at the School of Industrial Engineering at Eindhoven University of Technology in The Netherlands. He holds B.S. degrees in Mathematics and Physics, a M.S degree in Mathematics from Leiden University and a Ph.D. degree from Free University in Amsterdam. He is a 2004 Edelman Finalist. He has published over 100 papers in ISI journals, including Management Science, MSOM, POM, and OR. He founded the European Supply Chain Forum. His current research interests include Build-To-Order Supply Chains and generic models for supply chain optimization under stochastic demand and supply. His email address is a.g.d.kok@tue.nl.

BALAN SRINIVASAN is a Ph.D. student in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He holds a B.S. in Industrial Engineering from the College of Engineering, Guindy, Anna University and a Master in Industrial Engineering from the National Institute of Industrial Engineering, India. His research interests are in multi-echelon inventory optimization, supply chain risk management, and supply chain engineering. Before coming to NC State University, he worked in the industry for eight years in the automotive, manufacturing, healthcare, agribusiness, and consulting sectors. He authored five refereed journal papers, and is a certified Six Sigma Green Belt. He received the voice of the future award and excellence award during his work at Murugappa group and Olam international ltd. His email address is sbalan@ncsu.edu.

REHA UZSOY is the Clifton A. Anderson Distinguished Professor in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He holds B.S. degrees in Industrial Engineering and Mathematics and a M.S. in Industrial Engineering from Bogazici University, Istanbul, Turkey. He received his Ph.D. in Industrial and Systems Engineering in 1990 from the University of Florida, and held faculty positions in Industrial Engineering at Purdue University prior to joining North Carolina State University in 2007. His teaching and research interests are in production planning and supply chain management. Before coming to the US, he worked as a production engineer with Arcelik AS, a major appliance manufacturer in Istanbul, Turkey. He has also been a visiting researcher at Intel Corporation and IC Delco. He was named a Fellow of the Institute of Industrial Engineers in 2005, Outstanding Young Industrial Engineer in Education in 1997, and has received awards for both undergraduate and graduate teaching. He is the author of more than 120 refereed journal papers and book chapters, one monograph, and three edited books. His e-mail address is ruzsoy@ncsu.edu.