2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

On-line Inventory Control in Single-Echelon Systems with Time-Dependent Environment over Multiple Periods

Jianfeng Mao
Division of Systems Engineering and Management
School of Mechanical Aerospace Engineering
Nanyang Technological University, Singapore 639798
jfmao@ntu.edu.sg

***Abstract*— We consider an on-line inventory control in single-echelon systems with time-dependent environment over multiple periods, in which full backlogging and average cost function are adopted. We seek to control order quantity for each period so as to minimize a cost function involving maintenance cost and setup cost for each order. Most of state-of-art inventory control methods require the assumption that demands are independent and identically distributed over multiple periods which could result in a stationary optimal (s, S) or (Q, R) or base stock policy throughout the whole process. The assumption is not quite realistic in practice, especially when we face those disruptive events in supply chain systems. To consider this time-dependent demands, we develop a novel on-line inventory control algorithm to directly obtain the optimal order quantity for the current stage instead of the computation of those optimal policies like (s, S). A "best solution in probability" is introduced, which leads to a high-performance inventory system and can be efficiently obtained by solving a series of corresponding off-line inventory control problems. Numerical examples are included to illustrate our results and show substantial performance improvements over off-line analysis.**
Keywords: **Inventory Control, Single-Echelon, Time-Dependent Demand, Multiple Periods, Average Cost, Full Backlog.**

I. INTRODUCTION

Supply chain managers and academic researchers have been taking huge efforts to improve the efficiency of supply chain systems over years. Many solutions have been developed to increase overall customer value, such as the lean production techniques, which can minimize operating cost by eliminating several types of wastes. However, the effectiveness issues have been ignored somehow (the fast growth of the world economics until several years ago might be a reason). Although those lean techniques could efficiently minimize the operating cost by eliminating several types of wastes, they may tremendously increases the vulnerability of whole supply chain system because they largely reduce the slackness in the supply chain system.

Many disruptive events have occurred in this decade such as 911 Attacks in 2001, SARS Pandemic in 2003, Fuel Crisis in 2008 and H1N1 Pandemic in 2009, which dramatically affected the international business environments. Currently, we are still experiencing the historical financial crisis that results in stock price markdown and demands plummeting over the world. Supply chain managers and researchers gradually realize the importance of supply chain risk management, but the related research works are quite scarce [6]. When facing those catastrophes, managers urge quick-fix solutions in a fire-fighter mode to maintain profitability. Due to the paucity of the literatures about these effectiveness issues, managers can only make snap decisions based on a rule of thumb. They may overreact to time-dependent environment, which may cause unnecessary waste and an opportunity loss in quickly catching up when the economic recovery comes.

Inventory control is one of core issues in supply chain management. To improve the effectiveness of supply chain, time-dependent environment must be taken into account when applying inventory control. Many optimal policies have been developed and studied over decades such as (s, S), (Q, R) and base stock policies [1], [14]. The (s, S) policies is one of the most popular policies, which has been proved to be optimal for some inventory systems since [5] and [8]. More recent optimality results are obtained in [13] and [3]. The computation of the optimal policy is much more complicated than the optimality proofs. Many efforts have been made for many years [9], [12] and [4]. Most of state-of-art inventory control methods require the assumption that demands are independent and identically distributed over multiple periods, whereas the assumption of i.i.d demand only works for the case when products are in a mature stage of a product life cycle and are used regularly [1]. As mentioned above, the assumption is not quite realistic in practice, especially when we face disruptive events in supply chain systems. Even if without those disruptive events, while products are in an initial growth stage or a phase-out stage at the end of the life cycle, the demand will commonly follow a trend model, that is, the demand can be assumed to increase or decrease systematically. Thus, the assumption of i.i.d demand is a little bit strong and not quite matched to many real practices.

Although the (s, S) type policy may be still optimal, the corresponding optimal (s, S) is not stationary, that is, (s, S) differs over periods, which largely increases the computational complexity for policy iteration. The algorithm in [12] is only valid for the i.i.d discrete demand case. The smooth perturbation analysis method proposed in [4] can get an unbias estimator of the expected derivative with respect to a specific control variable based on the sample path analysis,

The author's work is supported in part by NTU under startup grant M58050030 and by AcRF under Tier 1 grant RG 33/10 M52050117.

which has a potential to extend to the time-dependent case. However, the dimensionality of the corresponding optimization problem will be changed from 2 to $2N$ (where N is the number of periods) because of the non-stationary (s,S) policy over each period, which makes it intractable for those perturbation analysis based methods.

In this paper, we develop a novel on-line optimal inventory control algorithm to the single-echelon problem with time-dependent environment, in which the optimal order quantity is directly obtained instead of computing the optimal policy (s,S) for each period. The corresponding on-line optimization problem could be formulated as a stochastic programming problem that is very hard to solve for the following two reason. First, the expected value of the cost function cannot be analytically calculated. Thus, it is necessary to invoke simulation-based methods to estimate quantities of interests. Second, the evaluation of the cost function itself is an off-line optimization problem. More details will be provided in the next section. We propose and analyze a new idea that can reduce the complexity of such simulation-based approaches by orders of magnitude.

In a standard stochastic programming method, we first estimate the expected value of the cost function and then optimize based on the estimated cost. Since the cost function itself is an off-line optimization problem, $O(IM)$ off-line problems are required to solve to derive the optimal solution, where M is the number of sample paths simulated for estimation purposes, and I is the total number of solutions evaluated. Since I and M are usually very large, standard stochastic programming methods may not be applicable to on-line control. The idea we develop in this paper is to reverse the estimation-optimization order above: we first optimize over sample paths and then use the cumulative distribution function of the solutions to estimate an optimal solution termed "*best solution in probability*"(BSIP), which was first defined in [7]. We show that the BSIP can be obtained by only solving $O(M)$ off-line problems, leading to a much faster process amenable to on-line control.

In Section II, we first formulate the off-line inventory control problem. Based on it, the on-line problem is defined afterwards, In Section III, we study the structural property for the off-line problem and an efficient algorithm is derived to solve the off-line problem. In Section IV, we introduce the "best solution in probability", leading to an algorithm for deriving a complete solution of the on-line problem. Simulation results are given in Section V illustrating the on-line capability of the proposed approach and we close with conclusions presented in Section VI.

II. Problem Formulation

In this section, we will consider a periodic review inventory system with time-dependent environment, which includes random demand, lead time and yield ratio. The demand is continuous and full backlogging. In what follows, to avoid the distraction, we will start with the case without lead time and random yield where only the time-dependent demand is considered, whose solution method can be easily extended to the case with random lead time and yield ratio by using the framework proposed in this paper.

A. *Off-line Problem Formulation*

Before proceeding to the on-line problem, we need first to formulate its corresponding off-line inventory control problem. Let

- x_i: the inventory level in period i;
- $d_i \geq 0$: the demand in period i;
- $u_i \geq 0$: the order quantity in period i;
- $h \geq 0$: the holding cost for the inventory;
- $p \geq 0$: the penalty cost for the backlog;
- K: the set-up cost;
- c: the purchase cost;
- s: the selling price.

A typical inventory control process could be depicted by Fig. 1, in which the order u_i is set at the beginning of each period and it is followed by the demand d_i. The inventory level x_i for the ith period is the inventory level on hand after the demand d_i.

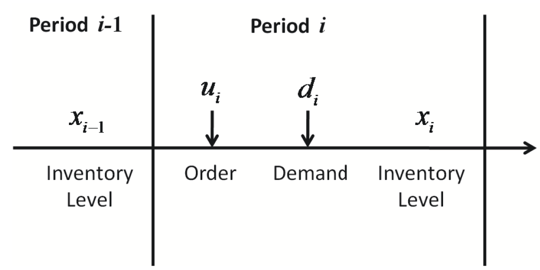

Fig. 1. Inventory Control Process

The off-line problem in this paper is an optimization problem defined over some specific sample path. Although the demand d_i for each period is a random variable, it becomes a deterministic parameter over a specific sample path. Therefore, the off-line problem is actually a deterministic optimization problem. Its objective is to maximize the profit by setting optimal order quantities for each period.

Let $H(x)$ denote the maintenance cost involving both the holding cost and shortage cost, which could be defined by tradition as

$$H(x) = h \cdot \max(x, 0) + p \cdot \max(-x, 0) \tag{1}$$

Then the off-line problem over a finite horizons (N periods) can be formulated as:

$$\begin{aligned} \min_{u_1,\ldots,u_N} & \sum_{i}^{N} \Big\{ H(x_i) + K \cdot \delta(u_i) + c \cdot u_i \Big\} \\ & - s \cdot \min\left(\sum_{i=1}^{N} d_i,\ x_0 + \sum_{i=1}^{N} u_i \right) \\ & s.t.\ \ x_i = x_{i-1} - d_i + u_i,\ i = 1, \ldots, N. \end{aligned}$$

where

$$\delta(u_i) = \begin{cases} 1 & u_i > 0 \\ 0 & u_i = 0 \end{cases}$$

and x_0 is the initial inventory level. The cost function is the inverse profit function. Its first term is the operating cost including maintenance cost, set up cost and purchase cost. Its second term is the total revenue over the N periods.

Since demands are deterministic over a specific sample path, it is possible to match the total demands by ordering exact same amount of orders. It is common to assume that more sales should result in a higher profit when the selling price s is reasonably larger than the purchase cost c, that is, the optimal solution of the off-line problem above must satisfy the equality below:

$$\sum_{i=1}^{N} u_i + x_0 = \sum_{i=1}^{N} d_i$$

Based on this property, the off-line problem above could be equivalently reduced to the following one:

$$\begin{aligned} &\min_{u_1,\dots,u_N} \sum_{i=1}^{N} \Big\{ H(x_i) + K \cdot \delta(u_i) \Big\} \\ &s.t.\ \ x_i = x_{i-1} - d_i + u_i,\ i = 1, ..., N \\ &\qquad \sum_{i=1}^{N} u_i + x_0 = \sum_{i=1}^{N} d_i. \end{aligned} \tag{2}$$

Its cost function only involves the maintenance cost and set up cost, which is actually a common practice in most inventory control models defined in literatures.

B. On-line Problem

After defining the off-line problem, we now proceed to formulate the on-line problem. The off-line problem is based on an open-loop control scheme and its optimal controls are all obtained at the very beginning. While in the on-line problem, a closed-loop control policy is adopted, *i.e.*, only the optimal control for the current period is computed and applied at the beginning of the corresponding period. Since the exact amount of the demand cannot be known in advance, a closed-loop scheme can derive a much better solution than the off-line solution based on an open-loop scheme because the closed-loop scheme could take opportunity to improve its control based on most updated information observed in the uncertain environment. However, this advantage of the on-line framework does not come at free, which causes a much higher computational complexity. It is very crucial to develop an efficient on-line algorithm to fully utilize this advantage. The detail analysis will be revealed after the on-line problem formulation.

Since only the optimal control of the current period needs to be determined in the on-line framework, we will begin with an auxiliary problem below, which is adapted from the off-line problem (2) and only focusing on the control u_1,

$$\begin{aligned} &\min_{u_1} J(u_1; x_0, \omega_N) = H(x_1) + K \cdot \delta(u_1) + L(x_1) \\ &\quad s.t.\ \ x_1 = x_0 + u_1 - d_1. \end{aligned} \tag{3}$$

where ω_N represents a specific sample path over a finite horizons of N periods, *i.e.*, a sequence of $d_1, ..., d_N$ in this case and the function $L(x_1)$ is defined as follows,

$$\begin{aligned} &L(x_1) \equiv \min_{u_2,\dots,u_N} \sum_{i=2}^{N} \Big\{ H(x_i) + K \cdot \delta(u_i) \Big\} \\ &s.t.\ \ x_i = x_{i-1} - d_i + u_i,\ i = 2, ..., N \\ &\qquad \sum_{i=2}^{N} u_i + x_1 = \sum_{i=1}^{N} d_i. \end{aligned} \tag{4}$$

Clearly, for a specific sample path ω_N, the problem (3) is equivalent to the off-line problem (2) and its optimal solution u_1^* can be obtained by solving (2). However, the demands $d_1, ..., d_N$ are random variables and they may not be independent and identically distributed. To consider this uncertain effects, it is necessary to introduce a stochastic programming setting to formulate the on-line problem. In this paper, we adopt the average cost over infinite horizons to measure the system performance. Let

$$\bar{J}(u_1; x_0, \omega_N) = \lim_{N \to +\infty} \frac{J(u_1; x_0, \omega_N)}{N} \tag{5}$$

and on-line problem could be formulated as

$$\min_{u_1} \Big\{ E_\omega \big(\bar{J}(u_1; x_0, \omega_N) \big) \Big\} \tag{6}$$

The on-line problem (6) is tremendously hard to solve. It might be able to be approached by dynamic programming method, but the policy iteration is notoriously time consuming. The estimation of the expected value is very costly. Since the demands are not independent and identically distributed, it is impossible to reach a closed form of the expected value of $\bar{J}(u_1; x_0, \omega_N)$. Besides, evaluating the function $J(u_1; x_0, \omega_N)$ for some u_1 is equivalently to solve an off-line problem because the function $L(x_1)$ involved in $J(u_1; x_0, \omega_N)$ is an off-line optimization problem. Now imagine if there are total I solutions are evaluated over M sample paths during the whole optimization process, then there are total MI off-line problems required to solve to find the best solution among these I solutions. To obtain a good enough solution, we have to accurately evaluate a large amount of solutions, which implies M and I should be large. Furthermore, we could not let N go to infinity in practice. Usually, we pick a large enough N to approximate the cost function of (6), which also largely increases the computational burden because of the high dimensionality N. To overcome these difficulties, a new solution framework are required. As mentioned in the introduction, we are going to develop a new idea to solve this problem, in which the estimation-optimization order is reversed.

III. Off-line Solution

To evaluate the function $J(u_1; x_0, \omega_N)$ for some u_1, the problem (4) is required to solve. Since the problem (4) is actually a kind of off-line problem defined in (2), we will first focus on developing an efficient algorithm to solve the off-line problem in this section.

Several methods had been proposed to solve problems similar to the off-line problem defined in this paper. In [10], an efficient algorithm is developed to solve the case without backlogging. In [11], although backlogging is considered, it is not quite efficient to implement because the dominant set is required to generate and its size grows exponentially with respect to N. In the following, we will develop a new algorithm to solve the off-line problem (2) in the computational complexity of $O(N^2 \log N)$.

We will start with solving a special case with zero initial inventory level, that is, $x_0 = 0$. For those cases with non-zero initial inventory, they could be equivalently reduced to the case with $x_0 = 0$.

A. Renewal Cycle Property

To solve the off-line problem (2) with $x_0 = 0$, we need to first reveal an optimality property, termed as "Renewal Cycle Property". Let us first introduce the concept of "Renewal Cycle" as follows.

Definition 1: A renewal cycle is a continuous set of periods $\{k, ..., n\}$ such that $x^*_{k-1} = 0, x^*_n = 0$ and $x^*_i \neq 0$ for $i = k, ..., n-1$.

Since $x_0 = 0$ and $x_N = 0$, there is at least one renewal cycle within the periods $\{1,...,N\}$. Before proceeding to the "Renewal Cycle Property", we need to reveal a lemma below to show that each order should always cover the demands in a number of consecutive periods.

Lemma 1: If $u^*_l > 0$, then there exists some p, q such that $u^*_l = \sum_{j=p+1}^{q} d_j$.

(The proofs in this paper are omitted due to page limits.)

Following this lemma, we could finally prove the "renewal cycle property" addressed in the theorem below.

Theorem 1: Only one order will be set within a renewal cycle.

A typical optimal trajectory could be illustrated as in Fig. 2 based on Theorem 1.

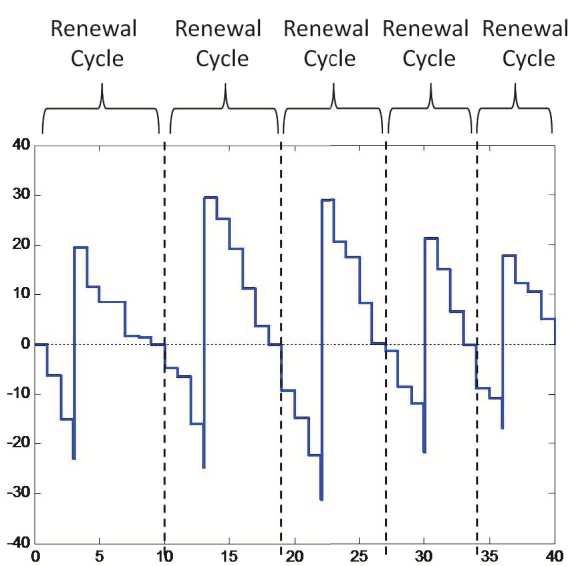

Fig. 2. A Typical Optimal Trajectory

B. Off-line Algorithm

Let J_n denote the optimal cost over period $\{1, 2, ..., n\}$ ending with the inventory level as zero and $G_{k,n}$ denote the optimal cost over a single renewal cycle over period $\{k, ..., n\}$. Then we could have

$$J_n = \min_{1 \leq k \leq n-1} \left\{ J_k + G_{k,n} \right\}$$

Based on the renewal cycle property revealed in Theorem 1, we could define $G_{k,n}$ as the optimal cost of the following problem

$$\begin{aligned} G_{k,n} \equiv & \min_{z_k,...,z_n} \sum\nolimits_{i=k}^{n} H(x_i) \\ \text{s.t.} \quad & x_i = -\sum\nolimits_{j=k}^{i} d_j + u \cdot z_i, \; i = k, ..., n; \\ & \sum\nolimits_{i=k}^{n} z_i = 1; \quad u = \sum\nolimits_{i=k}^{n} d_i; \\ & z_i = \{0, 1\}, \quad i = k, ..., n. \end{aligned} \tag{7}$$

The problem (7) can be easily solved by a binary search algorithm with the computational complexity of $O(\log N)$.

The off-line algorithm is developed in Table I, in which we start from J_1 and gradually increase n from 1 to N. Finally, we reach J_N that corresponds to the off-line problem (2). Its computational complexity is $O(N^2 \log N)$.

TABLE I
OFF-LINE ALGORITHM

Step 1:	Start from $n = 1$ and $J_1 = G_{1,1}$;
Step 2:	Compute $G_{k,n}$ for $k = 1, ..., n$;
Step 3:	Compute $J_n = \min_{1 \leq k \leq n-1} \{J_k + G_{k,n}\}$;
Step 4:	$n = n + 1$; if $n > N$ stop, otherwise goto step 2.

C. Nonzero Initial Inventory

Before proceeding to the case with non-zero initial inventory, we need to first reveal the follow lemma.

Lemma 2: If $x_0 \geq d_1$, then $u^*_1 = 0$.

There are two possible non-zero initial inventory cases: $x_0 < d_1$ and $x_0 \geq d_1$.

Case (I) $x_0 < d_1$. We could reduce it to an equivalent zero initial inventory case where $d'_1 = d_1 - x_0$ and $d'_i = d_i$ for $i = 2, ..., N$.

Case (II) $x_0 \geq d_1$. Without loss of generality, there always exists some k such that

$$x_0 - \sum\nolimits_{i=1}^{k-1} d_i \geq d_k \quad \text{and} \quad x_0 - \sum\nolimits_{i=1}^{k} d_i < d_{k+1}$$

We can have $u^*_i = 0$ for $i = 1, ..., k$ based on Lemma 2. The remain process $\{k+1, ..., N\}$ becomes to Case (I) we discussed above.

D. Lead Time and Yield Ratio

Let L denote the lead time and α denote the yield ratio. The L and α may also be random variables like the demand d_i. However, over a specific sample path, L and α are just deterministic parameters. To solve an off-line problem considering L and α, we could still first solve the off-line problem (2) and obtain its optimal order quantity u^*_i. Then we just need to set of the amount of order as u^*_i/α and place it L periods before the ith period.

IV. ON-LINE SOLUTION

In this section, we will develop an efficient on-line algorithm to solve the stochastic programming problem (6), in which we search for the best solution in probability instead of the one minimizing the expected value.

A. Best Solution in Probability

Although we find an efficient way to evaluate the function $J(u_1; x_0, \omega_N)$ in (3), we still face the difficulty of evaluating its expected value. In the following, we provide an alternative to standard stochastic programming based on a new idea which can bypass the difficulty of evaluating this expected value and fully utilizes the efficiency of the algorithm which solves the off-line problem(2) as discussed earlier.

Clearly, a closed-form expression for $E_\omega\{\bar{J}(u_1; x_0, \omega_N)\}$ in (6) cannot be derived and has to be estimated. If we proceed via Monte Carlo simulation as in standard stochastic programming, there are several notable difficulties: **(i)** it is costly to evaluate $J(u_1; x_0, \omega_N)$ for each u_1. Assume we randomly generate M sample paths (i.e., realizations of demands $d_i, i = 1, ..., N$, the lead time L and the yield ratio α) and solve problem (4) for each sample path. Since problem (4) can be solved in $O(N^2 \log N)$ by using the off-line algorithm, the complexity of evaluating a solution u_i is $O(MN^2 \log N)$. Due to the central limit theorem, the accuracy of the estimation of this expected value is improved in the rate of $O(1/\sqrt{M})$, which implies a large number, M, of simulations are required to get a good estimation; **(ii)** if one is to use derivatives in an optimization algorithm, both $dE_\omega\big(\bar{J}(u_1; x_0, \omega_N)\big)/du_k$ and $d\bar{J}(u_1; x_0, \omega_N)/du_k$ are hard to compute because $J(u_1; x_0, \omega_N)$ involves $L(x_1)$ defined in (4) which has no closed form. Only finite differences can be obtained, which requires two time-consuming evaluations; **(iii)** assuming the total number of iterations is I, the total complexity of solving the on-line problem is $O(IMN^2 \log N)$. It usually may take many iterations to converge to the optimal solution of (6), which implies that I is also very large; **(iv)** the cost function in (6) is defined over infinite horizons. Although we could use a finite-horizon setting to approximate it, we need to choose a large enough N to reach a better approximation of the cost function. For the analysis above, the computational complexity of the standard stochastic programming method is $O(IMN^2 \log N)$ where I, M and N are all very large. Such huge complexity is not suitable for on-line control.

In the following, we will bypass much of this complexity by developing a probabilistic comparison algorithm, in which we search for the "Best Solution in Probability" defined below instead of the on-line optimal solution, i.e., the one minimizing $E_\omega\{\bar{J}(u_1; x_0, \omega_N)\}$.

Definition 2: u_1^* is the *Best Solution in Probability (BSIP)* if and only if u_1^* satisfies,

$$\Pr\left[\bar{J}(u_1^*; x_0, \omega_N) \leq \bar{J}(u_1; x_0, \omega_N))\right] \geq 0.5 \qquad (8)$$

By this definition, when comparing to any arbitrary solution u_1, the BSIP u_1^* is always more likely to be better than the arbitrarily selected solution u_1.

Remark 1: Generally, the BSIP is not the on-line optimal solution in the sense of minimizing an expected value as in (6). One can interpret it as an alternative definition of optimality to the usual "optimality in expectation" (which may in fact not be the best choice in some applications).

Remark 2: The BSIP coincides with the on-line optimal solution under the *Non-singularity Condition (NSC)* introduced in [7].

Remark 3: It remains an open problem whether the NSC is satisfied in this particular problem. If so, the BSIP is an estimate of the optimal solution of the on-line problem (6) which as we will see converges to the true solution very fast. Otherwise, the BSIP is a sub-optimal solution which we expect to be quite close to optimal, assuming "optimality" is still in the traditional sense of minimizing an expectation.

Based on the analysis above, even if NSC cannot be satisfied, the BSIP can still serve as a near optimal or good enough solution to approximate the on-line optimal solution. In the following, we will prove the existence of the BSIP for this particular problem by using a construction method. Then, an algorithm is developed to determine the BSIP in $O(MN^2 \log N + M \log M)$ complexity. To begin with, we exploit a property of $J(u_1; x_0, \omega_N)$ as shown next.

Lemma 3: $L(x_1)$ is K-convex with respect to x_1, that is, for any $x_1 < x_1' < x_1''$, it satisfies that

$$K + L(x_1'') \geq L(x_1') + (x_1'' - x_1')\left(\frac{L(x_1') - L(x_1)}{x_1' - x_1}\right).$$

$L(x_1)$ is actually a cost-to-go function and Lemma 3 can be similarly proved in the context of dynamic programming as shown in Section 4.2 in [2].

Let t denote the optimal solution of the problem below

$$t = \arg\min_{u_1} \frac{J(u_1; x_0, \omega_N)}{N}$$

and the function $J_1(u_1'')$ possesses the following property.

Lemma 4: If $t < u_1' < u_1''$, then

$$2K + J_1(u_1'') \geq J_1(u_1')$$

Since $J(u_1''; x_0, \omega_N)$ depends on the sample path ω_N, t is also a random variable. Let us define $G(u_1) = \Pr[t \leq u_1]$ and

$$u_1^* = \inf_u\{u : u = G^{-1}(0.5)\} \qquad (9)$$

where $G^{-1}(\cdot)$ is the inverse function of $G(\cdot)$. Generally, the BSIP may not always exist. As we proved in [7], if $J(u_1; x_0, \omega_N)$ is unimodal in u_1, then the BSIP is guaranteed to exist. Although a K-convex function might not be unimodal, the parameter K can be regard as the degree of the unimodality. When K approaches 0, the function approaches unimodal. Based on this insight, we have the following lemma.

Lemma 5: If $t < u_1' < u_1''$, then

$$\bar{J}(u_1''; x_0, \omega_N) \geq \bar{J}_1(u_1'; x_0, \omega_N)$$

This lemma finally leads us to the following theorem that determines the BSIP.

Theorem 2: The u_1^* defined in (9) is a BSIP satisfying (8).

B. On-line Algorithm

From Theorem 2, we can obtain u_1^* through the cumulative distribution function of t, $G(u_1)$. Since $G(u_1)$ is not available in closed form, we need to estimate u_1^* by estimating $G(u_1)$. We can resort to a Monte Carlo simulation method, in which we generate M sample paths where a sample path is generated based on the time-dependent demands $d_1, ..., d_N$. Let ω_N^j denote the jth sample path for $j = 1, ..., M$ and let t^j denote the solution of minimizing $J(u_1; x_0, \omega_N^j)$ in the jth sample path ω_N^j. Then, $G(u_1)$ can be estimated by

$$\hat{G}(u_1) = \frac{1}{M} \sum_{j=1}^{M} \mathbf{1}[t^j \le u_1]$$

where $\mathbf{1}[\cdot]$ is an indicator function. Let $\hat{u}_1^* = \inf_u\{u : u = \hat{G}^{-1}(0.5)\}$. Based on the strong law of large number, $\hat{G}(u_1)$ converges to $G(u_1)$ w.p.1 as M approaches to infinity. Combining this with $\hat{u}_1^* = \hat{G}^{-1}(0.5)$ and $u_1^* = G^{-1}(0.5)$, $\hat{u}_1$ also converges to u_1^* w.p.1 as $M \to +\infty$. Furthermore, using the Chernoff bound, we can show that $\hat{u}_1^*$ approaches u_1^* exponentially fast as M increases, that is,

Lemma 6: For any $\epsilon > 0$, there always exists $C > 0$ such that $\Pr\left[\, |\hat{u}_1^* - u_1^*| \ge \epsilon \right] \le 2e^{-CM}$.

The remaining question is how to obtain t^j for its corresponding sample path ω_N^j, which can be fulfilled by apply the off-line algorithm in Table I. Finally, we could develop the on-line algorithm in Tab II to efficiently obtain an estimate of BSIP u_1^*. In Step 2 of this algorithm, we only solve M off-line problems, which can be accomplished in $O(MN^2 \log N)$. Moreover, although a good approximation of the cost function $\bar{J}(u_1; x_0, \omega_N)$ requires a large N, the optimal solution t could quickly converge with respect to N. Since we only need to compute the optimal solution t^j for each sample path ω_N^j in the on-line algorithm, we can just pick $n << N$, a much smaller number of total periods to approximate the optimal solution t^j. This effect can be observed in the numerical results. The complexity of the sorting procedure in Step 3 is $O(M \log M)$. Thus, deriving u_1^* is accomplished in $O(Mn^2 \log n + M \log M)$ complexity, which is clearly a vast improvement over $O(IMN^2 \log N)$.

TABLE II

ON-LINE ALGORITHM

Step 1:	Randomly generate M Sample paths, $\omega_N^1, ..., \omega_N^M$;
Step 2:	Obtain t^j by using the off-line algorithm for each sample path ω_N^j, for $j = 1, ..., M$;
Step 3:	Sort $t^1, ..., t^M$ to derive $\hat{G}(u_1)$ and then obtain $\hat{u}_1^*$.

(The numerical results are omitted in this paper due to page limits)

V. CONCLUSIONS

In this paper, we have addressed the on-line inventory control problem in single-echelon systems with time-dependent environment, in which a stationary optimal policy like (s, S) does not exist. To consider this time-dependent environment, we introduced a close-loop control scheme and developed a novel on-line inventory control algorithm, in which only the optimal order quantity for the current stage will be computed and applied. To fulfill this task, we first obtained an highly efficient algorithm to solve its corresponding off-line problem and then proposed an approach leading to a "best solution in probability" (BSIP). This solution estimates the on-line optimal control (and converges to it exponentially fast) if the non-singularity condition holds and otherwise provides suboptimal solutions. Obtaining the BSIP entails reversing the order of the usual stochastic programming approach where one first estimates an expected cost and then seeks to minimize it; we first optimize the cost over individual sample paths and then estimate the cumulative distribution function of the solutions from which the BSIP is easily obtained. This has the advantage of drastically reducing computational complexity.

Future work is aiming at studying a more general inventory control system where the inventory depletion is also allowed. Moreover, we assume that information on probability distributions of these time-dependent demands are known beforehand. However, in some applications only rough information of this type may be available. We plan to incorporate a forecasting technique, such as time series analysis, to estimate these probability distributions based on past history and study their convergence properties when stationarity applies.

REFERENCES

[1] S. Axsäter. *Inventory Control (Second Edition)*. Springer, 2006.

[2] D. P. Bertsekas. *Dynamic Programming and Optimal Control Vol.1 (Second Edition)*. Athena Scientific, 2000.

[3] D. Beyer and S. P. Sethi. The classical average-cost inventory models of iglehart (1963) and veinott and wagner (1965) revisited. *Journal of Optimization Theory and Applications*, 101(3):523–555, 1999.

[4] M.C. Fu. Sample path derivatives for (s,S) inventory systems. *Operations Research*, 42(2):351–364, 1994.

[5] Donald L. Iglehart. Optimality of (s, S) policies in the infinite-horizon dynamic inventory problem. *Management Science*, 9(2):259–267, 1963.

[6] U. Jüttner, H. Peck, and M. Christopher. Supply chain risk management: Outlining an agenda for future research. *International Journal of Logistics*, 6(4):197–210, 2003.

[7] J. Mao and C. G. Cassandras. On-line optimal control of a class of discrete event systems with real-time constraints. *Journal of Discrete Event Dynamic Systems*, 20(2):187–213, 2010.

[8] Arthur F. Veinott. On the optimality of (s, S) inventory policies: New conditions and a new proof. *SIAM Journal on Applied Mathematics*, 14(5):1067–1083, 1966.

[9] Artur F. Veinott and Harvey M. Wagner. Computing optimal (s,S) inventory policies. *Management Science*, 11(5):525–552, 1966.

[10] Harvey M. Wagner and Thomson M. Whitin. Dynamic version of the economic lot size model. *Management Science*, 5(1):89–96, 1958.

[11] Willard I. Zangwill. A deterministic multi-period production scheduling model with backlogging. *Management Science*, 13(1):105–119, 1966.

[12] Y. Zheng and A. Federgruen. Finding optimal (s,S) policies is about as simple as evaluating a single policy. *Operations Research*, 39(4):654–665, 1991.

[13] Y. S. Zheng. A simple proof for optimality of (s, s) policies in infinite-horizon inventory systems. *Journal of Applied Probability*, 28:802–810, 1991.

[14] P. H. Zipkin. *Foundations of Inventory Management*. McGraw-Hill, Singapore, 2000.