

A FLUID EOQ MODEL WITH MARKOVIAN ENVIRONMENT

YONIT BARRON,* *Ariel University*

Abstract

We consider a production-inventory model operating in a stochastic environment that is modulated by a finite state continuous-time Markov chain. When the inventory level reaches zero, an order is placed from an external supplier. The costs (purchasing and holding costs) are modulated by the state at the order epoch time. Applying a matrix analytic approach, fluid flow techniques, and martingales, we develop methods to obtain explicit equations for these cost functionals in the discounted case and under the long-run average criterion. Finally, we extend the model to allow backlogging.

Keywords: Production-inventory model; backlog; cost functional; fluid flow process; martingale; EOQ

2010 Mathematics Subject Classification: Primary 60K10

Secondary 90B05; 60J28

1. Introduction

An important problem in inventory planning is how to effectively manage the inventory control in a dynamic and stochastic environment. The simple economic order quantity (EOQ) model (see, e.g. Nahmias (1997)) is the most fundamental of all inventory models. It describes the trade-off between fixed ordering costs and variable holding costs. In this paper we study a stochastic fluid EOQ model of a single infinite capacity buffer. The buffer content level increases or decreases according to a fluid-flow rate modulated by an n -state continuous-time Markov chain (CTMC). Whenever the buffer becomes empty, the environment state jumps to another state instantaneously with a given probability (or it may stay unchanged). At the same time the buffer is refilled to a state-dependent level instantaneously. In Figure 1, we illustrate a sample path of the process (environment state and inventory level).

Our primary motivation for considering this model is to provide contributions to the study of inventory systems modulated by a Markovian environment. The fluid process is the inventory position or inventory level under continuous review where the environment process represents the background state, for example, the production or sales season. A jump in the fluid level represents an external order placement or an order arrival, and the transition at the background state at the jump point can be the result of repairs or production facilities, etc.

In order to manage such an inventory model, one needs information about the holding cost for the stock, the fixed cost of an order, and the purchasing cost. The main objective of this paper is to develop techniques to determine explicit equations for all these costs under the discounted as well as under the long-run average cost criterion.

Dynamic control of stochastic inventory systems are classified according to whether the demand and production are discrete or continuous. In the case of continuous demand and

Received 10 July 2013; revision received 8 April 2014.

* Postal address: Department of Industrial Engineering and Management, Ariel University, Ariel, 40700, Israel.

Email address: ybarron@ariel.ac.il

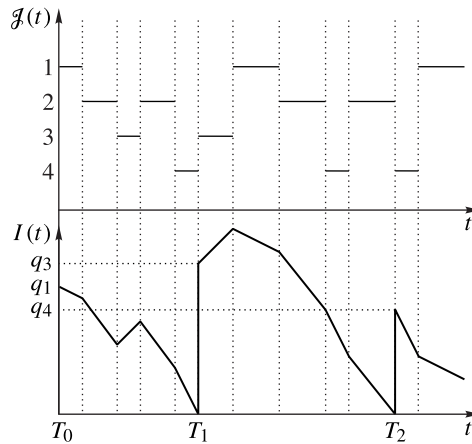


FIGURE 1: A typical sample path of the background environment ($J(t)$) and the inventory level process ($I(t)$).

production, fluid versions of these models were studied by Kulkarni *et al.* (2005) and Kulkarni (2010). The closest papers to our analysis are Kulkarni and Yan (2007) and Yan and Kulkarni (2008), where in both papers the authors consider a continuous fluid-flow system with jumps at the boundary. More specifically, Kulkarni and Yan (2007) derived a system of first-order nonhomogeneous linear differential equations for the limiting distribution of the bivariate (buffer level, environment state) process and Yan and Kulkarni (2008) derived an EOQ policy that minimized the long-run average cost; they also extended the state-space of the model by allowing backlogging. Berman *et al.* (2006) studied a fluid EOQ-type model operating in a Markovian random environment of alternating good and bad periods determining the demand rates. They derived the steady-state mean of the content level and the expected cycle length. Berman and Perry (2006) presented an EOQ-type model in which the demand rate is a function of the inventory level. Later, Berman *et al.* (2008) assumed an EOQ model in which the content level is modelled by a Brownian demand. They derived cost functionals for the discounted case and for the long-run average case. Another related model is the so-called clearing system (see Kella *et al.* (2003), Berman *et al.* (2005), and Perry *et al.* (2005)), which can be regarded as a dual EOQ stochastic model. In a clearing system, the fluid process jumps back to zero when it reaches a certain positive level.

Our analysis is based on a combination of a martingale technique and an application of fluid flow theory. The martingale approach was introduced by Asmussen and Kella (2000) and was frequently used in the study of inventory models, see, e.g. Perry *et al.* (2001) and Kella *et al.* (2003) and the references therein. Fluid flows have been an active area of research in recent years. Bean and O'Reilly (2008) studied the behavior of the fluid in a buffer with threshold controls with a wide range of behaviors possible at the boundaries. Ramaswami (1999) initiated a unified matrix-analytic algorithmic approach to fluid flows, and this was followed by a series of papers by the author, Ahn, and others (see Ramaswami (2006), Ahn and Ramaswami (2003)–(2006), and Ahn *et al.* (2007)). To the best of the author's knowledge, fluid flow methods have not been applied to inventory models and the associated cost-related quantities.

In this paper we use \mathbb{E} and \mathbb{E}_i to represent expectation and conditional expectation operators, respectively. We denote \mathbf{E} to be a matrix (or a vector) of expectations. We denote by \mathbf{e} a column vector with ones, i.e. $\mathbf{e} = (1, 1, \dots, 1)^\top$, by \mathbf{e}_i a row vector with the i th component equal to 1 and all the other components 0, by \mathbf{I} the identity matrix and by $\mathbf{1}_{\{A\}}$ the indicator of an event A .

2. Mathematical description of the model

Let $I(t)$ be the fluid level in the buffer at time t . The rate of change of the fluid level is modulated by a CTMC $\{\mathcal{J}(t) : t \geq 0\}$ on a finite state-space $S = \{1, 2, \dots, n\}$ with a generator matrix $\mathbf{Q} = [Q_{ij}]$. As long as $\mathcal{J}(t)$ is in state i , the production occurs continuously at rate p_i , and there is a demand at rate d_i . The net production rate is, thus, $u_i = p_i - d_i$. Note that u_i may be either negative or positive. Accordingly, we have two disjoint sets (S_1, S_2) , $S = S_1 \cup S_2$, where S_1 is a nonempty set of increasing rates $S_1 = \{i \in S : u_i > 0\}$ and S_2 is a nonempty set of the decreasing rates $S_2 = \{i \in S : u_i < 0\}$. Let $|S_1| = n_1$ and $|S_2| = n_2$; thus, $n_1 + n_2 = n$. Let $\pi = [\pi_1, \dots, \pi_n]$ be the limiting distribution of the $\mathcal{J}(t)$ process, i.e. π is the unique solution to

$$\pi \mathbf{Q} = 0, \quad \pi \mathbf{e} = 1.$$

The system is stable if and only if the expected input rate is negative, i.e.

$$\sum_{i=1}^n \pi_i u_i < 0$$

(see Kulkarni and Yan (2007)). We assume that the stability condition holds throughout the paper.

Let $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_n]$ be the initial probability vector of $\mathcal{J}(t)$. When $I(t)$ down-crosses level 0 in state $i \in S_2$, it is switched instantaneously to state $j \in S$ with probability α_{ij} . Then after changing to state j , the process places a state-dependent order of size $q_j > 0$ from an external supplier who delivers instantaneously. Thus, immediately after the down-crossing to emptiness by the fluid in state i , the fluid process restarts in state q_j with probability α_{ij} .

Let T_k be the time of the k th jump ($T_0 = 0$) and $\mathcal{J}_k = \mathcal{J}(T_k)$ be the environmental state at T_k (just after the jump). We assume that over $[T_{k-1}, T_k)$ $k = 1, 2, \dots$, the process $\{\mathcal{J}(t), t \in [T_{k-1}, T_k)\}$ is an irreducible CTMC on S . We call the points where the process jumps up (the replenishment times) *order points*. They form a semi-renewal process where the T_k s are the semi-regenerative points of the process. Thus, α_{ij} is defined as

$$\alpha_{ij} = \mathbb{P}\{\mathcal{J}_k = j \mid \mathcal{J}(T_k^-) = i\}, \quad i \in S_2, j \in S.$$

The jump probabilities form a matrix $\mathbf{A} = [\alpha_{ij}]$. Note that the case in which the environment is not switched at the moment of an order arrival is a special case of the latter assumption by setting $\mathbf{A} = \mathbf{I}$. Define the k th cycle as the time elapsed between T_{k-1} and T_k , $k = 1, 2, \dots$. Denote by $L_k = T_k - T_{k-1}$, $k = 1, 2, \dots$ ($L_0 = 0$) the inter-replenishment times and let $L = L_1$.

Next, we consider the cost functionals of the model. The total cost consists of three parts: a fixed ordering cost, a purchasing cost, and a holding cost. The cost components are determined by the environment at the *order points*. Thus, we assume that if the state at an *order point* is $\mathcal{J}_k = i \in S$, the fixed ordering cost is K_i for an order (typically, cost of ordering, shipping, and handling), the cost to purchase one item from an external supplier is c_i and the cost to hold one item in inventory during a time interval of length dt is $h_i dt$ during all the cycle k

in which $\mathcal{J}_k = i$ (i.e. the holding cost for one unit is constant between two consecutive order replacements, i.e. $h(t) = h_i$ for $t \in [T_k, T_{k+1})$ and $\mathcal{J}_k = i$).

We derive in a closed form these cost functionals for the discounted case as well as for the long-run average case. The crucial tools of our analysis are introduced in the next section.

3. Preliminaries

For the determination of the cost functionals, we use two tools: (a) the matrix-analytic approach and the theory of Markov-modulated fluid flows (MMFFs), initiated by a series of papers by Ahn and Ramaswami (2003)–(2007) and Ramaswami (2006) and (b) an application of the optional sampling theorem to the multi-dimensional martingale of Asmussen and Kella (2000).

3.1. The fluid inventory model

Introduce a modulating CTMC $\{\mathcal{J}(t); t > 0\}$ with state-space $S = S_1 \cup S_2$. ($|S_1| = n_1, |S_2| = n_2$). Its infinitesimal generator Q is given in a block form according to transitions between the sets S_i ($i = 1, 2$),

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}.$$

Now let $\mathcal{F}(t)$ be the content level of the fluid at time $t \geq 0$ that is modulated as follows: whenever the Markov chain is $i \in S_1$, the fluid flow increases linearly at rate $u_i > 0$ and whenever it is in $j \in S_2$, the fluid flow decreases linearly at rate $u_j > 0$. The two-dimensional stochastic process $\{\mathcal{F}(t), \mathcal{J}(t), t \geq 0\}$ is called a Markov-modulated fluid flow (MMFF) process. Let U_1, U_2 , and U be diagonal matrices,

$$U_j = \text{diag}\{u_i, i \in S_j\}, \quad U = \text{diag}(U_1, U_2), \quad j = 1, 2.$$

Let $\tau(x) = \inf\{t > 0, \mathcal{F}(t) = x\}$ be the first passage time to level x . Let $\Psi(s)$ be an $(n_1 \times n_2)$ matrix whose ij th component is

$$[\Psi(s)]_{ij} = \mathbb{E}\{e^{-s\tau(0)}, \mathcal{J}(\tau(0)) = j \mid \mathcal{F}(0) = 0, \mathcal{J}(0) = i\}, \quad i \in S_1, j \in S_2,$$

which is the LST (Laplace–Stieltjes transform) of $\tau(0)$ restricted to the event that the inventory process hits level 0 in phase $j \in S_2$ and given that $\mathcal{F}(0) = 0, \mathcal{J}(0) = i \in S_1$. Ramaswami (2006) shows how to compute $\Psi(s)$ and provides a good algorithm for this (note that there are other algorithms including some quadratically convergent ones for computing that matrix; see, e.g. Bean *et al.* (2008)). The following are some interesting first passage times of the process \mathcal{F} :

- $\tau(x, y)$ = first passage time of \mathcal{F} from level x to level y ,
- $^a\tau(x, y)$ = first passage time of \mathcal{F} from level x to level y avoiding a visit to levels in $[a, \infty)$ en route,
- ${}_a\tau(x, y)$ = first passage time of \mathcal{F} from level x to level y avoiding a visit to levels in $[0, a]$ en route,
- $^b_a\tau(x, y)$ = first passage time of \mathcal{F} from level x to level y avoiding a visit to levels in $[0, a] \cup [b, \infty)$ en route.

We use the notation $\hat{f}(x, y, s), {}^a\hat{f}(x, y, s), {}_a\hat{f}(x, y, s)$, and $^b_a\hat{f}(x, y, s)$ to denote, respectively, the matrices of LSTs of the joint distribution of the first passage times $\tau(x, y), {}^a\tau(x, y), {}_a\tau(x, y), {}^b_a\tau(x, y)$, and the state of the phase process at each first passage time.

TABLE 1: Transform matrices.

Quantity	Matrix size
$K(s) = U_1^{-1}(Q_{11} - sI) + \Psi(s)U_2^{-1}Q_{21}$	$n_1 \times n_1$
$H(s) = U_2^{-1}(Q_{22} - sI + Q_{21}\Psi(s))$	$n_2 \times n_2$
$\Xi(s, x) = \Psi(s) \int_{(0,x)} e^{H(s)y} (U_2^{-1}Q_{21}) e^{K(s)y} dy$	$n_1 \times n_1$

TABLE 2: Transform matrices of the reversed process.

Quantity	Matrix size
$K^r(s) = U_2^{-1}Q_{22} + \Psi^r(s)U_1^{-1}Q_{12}$	$n_2 \times n_2$
$H^r(s) = U_1^{-1}(Q_{11} - sI + Q_{12}\Psi^r(s))$	$n_1 \times n_1$
$\Xi^r(s, x) = \Psi^r(s) \int_{(0,x)} e^{H^r(s)y} (U_1^{-1}Q_{12}) e^{K^r(s)y} dy$	$n_2 \times n_2$

TABLE 3: LST of first passage times.

LST	First passage time	Matrix size
${}_0\hat{f}_{12}(x, 0, s) = \Psi(s)e^{H(s)x}$	from (x, S_1) to $(0, S_2)$ avoiding 0 in F	$n_1 \times n_2$
${}_0\hat{f}_{22}(x, 0, s) = e^{H(s)x}$	from (x, S_2) to $(0, S_2)$ avoiding 0 in F	$n_2 \times n_2$
${}_0\hat{f}_{11}(0, x, s) = e^{K(s)x}(I + \Xi(s, x))^{-1}$	from $(0, S_1)$ to (x, S_1) avoiding 0 in F	$n_1 \times n_1$
${}_x\Psi^r(s) = \Psi^r(s) - {}_0\hat{f}_{22}^r(0, x, s)\Psi^r(s)e^{H^r(s)x}$	from $(0, S_2)$ to $(0, S_1)$ avoiding x in F^r	$n_2 \times n_1$
${}_0\hat{f}_{22}^r(0, x, s) = e^{K^r(s)x}(I + \Xi^r(s, x))^{-1}$	from $(0, S_2)$ to (x, S_2) avoiding 0 in F^r	$n_2 \times n_2$

An important variant of the fluid flow \mathcal{F} , a *reflected fluid flow*, is particularly useful in the analysis of our inventory level process. The reflected fluid flow \mathcal{F}^r is obtained by reversing the roles of the up and down environment states. Analogously, $\Psi^r(s)$ is the matrix (of order $(n_2 \times n_1)$) whose (i, j) component is the LST of the time to reach level 0 for the process \mathcal{F}^r restricted to $\mathcal{J}^r(\tau(0)) = j \in S_1$, given that $\mathcal{F}^r(0) = 0$ and $\mathcal{J}^r(0) = i \in S_2$, where $\mathcal{J}^r(t)$ is the modulated state process for \mathcal{F}^r (we use notation $\hat{f}^r(x, y, s)$, ${}^a\hat{f}^r(x, y, s)$, ${}_a\hat{f}^r(x, y, s)$, and ${}_b\hat{f}^r(x, y, s)$ to denote quantities similar to those defined above for \mathcal{F}^r).

For the hitting times that we will use all these matrices are straightforward to evaluate once we have computed $\Psi(s)$. We list these matrices and their sizes in Tables 1, 2, and 3. All matrices have nice probabilistic interpretations. For more details; see Ahn and Ramaswami (2005), Ramaswami (2006), and Ahn *et al.* (2007).

Analogous to the matrices considered above for F , we also introduce the matrices $K^r(s)$, $H^r(s)$, and $\Xi^r(s)$ associated to the rate-reverse flow \mathcal{F}^r , by changing the index from 1 to 2, and 2 to 1.

In our model the bivariate process $(I(t), \mathcal{J}(t))$, between two sequential *order points*, has the same distribution as the MMFF process $(\mathcal{F}(t), \mathcal{J}(t))$; thus, we will use the quantities in Table 3 for relevant functionals.

3.2. The multi-dimensional martingale

Let $\{X(t), t \geq 0\}$ be a right-continuous Markov modulated Lévy process with modulating process $\{J(t), t \geq 0\}$ which is a right-continuous irreducible finite state-space CTMC. Let $\{Y(t), t \geq 0\}$ be an adapted continuous process with a finite expected variation on finite intervals and let $Z(t) = X(t) + Y(t)$. Asmussen and Kella (2000) have shown that for such a process the matrix with elements $\mathbb{E}_i\{e^{\alpha X(t)}; J(t) = j\}$ has the form of $e^{tK(\alpha)}$ for some matrix $K(\alpha)$. Theorem 2.1 of Asmussen and Kella (2000) yields that under certain conditions on $\{Z(t), t \geq 0\}$, the multi-dimensional process

$$M(\alpha, t) = \int_0^t e^{\alpha Z(s)} \mathbf{1}_{\{J(s)\}} ds K(\alpha) + e^{\alpha Z(0)} \mathbf{1}_{\{J(0)\}} - e^{\alpha Z(t)} \mathbf{1}_{\{J(t)\}} + \alpha \int_0^t e^{\alpha Z(s)} \mathbf{1}_{\{J(s)\}} dY(s) \tag{3.1}$$

is a (row) vector-valued zero mean martingale. Some of the relevant functionals in this paper will be obtained by applying the optional stopping theorem (OST) (or Doob’s optional sampling theorem, see Doob (1953)) to appropriate special cases of (3.1). For our model $X(t)$ has piecewise linear sample paths with slope u_j on intervals where $\mathcal{J}(t) = j$.

In this paper for any matrix B , we shall denote its elements by $(B)_{ij}$ or by $[B]_{ij}$ and reserve the notation B_{ij} for the sub-matrix of B with row indices in S_i and column indices in S_j . Moreover, we use Δ_q for a diagonal matrix as follows: $\Delta_q = \text{diag}(q_1, q_2, \dots, q_n)$.

4. The discounted model

Let us now introduce the functionals indicating the expected discounted costs in our model using a discount factor $\beta > 0$: (a) the order cost, includes a fixed set up cost whenever an order is placed and a purchasing cost, and (b) the holding cost for the stock.

4.1. Order cost

Assume that $\mathcal{J}_k = i \in S$. Then an order of size q_i is placed and $I(T_k) = q_i$. The order cost is $K_i + c_i q_i$. Let $OC(\beta)$ be the expected discounted order cost and let $\hat{O}(\beta)$ be an $(n \times 1)$ vector whose i th component $\hat{O}_i(\beta)$ is given by

$$\hat{O}_i(\beta) = \mathbb{E}_i \left\{ \sum_{k=0}^{\infty} e^{-\beta T_k} (K_{\mathcal{J}_k} + c_{\mathcal{J}_k} q_{\mathcal{J}_k}) \right\},$$

i.e. the expected discounted order cost, given $\mathcal{J}(0) = i \in S, I(0) = q_i$. Then, we have

$$OC(\beta) = \gamma \cdot \hat{O}(\beta).$$

Regarding Table 2, let ${}_0\hat{f}_{12}(q, 0, s)$ be an $(n_1 \times n_2)$ matrix whose ij th component is

$$({}_0\hat{f}_{12}(q, 0, s))_{ij} = (\Psi(s)e^{H(s)q_i})_{ij}, \quad i \in S_1, j \in S_2.$$

The ij th component $({}_0\hat{f}_{12}(q, 0, s))_{ij}$ represents the LST of the time until the content level process hits level 0 in state $j \in S_2$, given $\mathcal{J}(0) = i \in S_1, I(0) = q_i$ (for a proof; see Theorem 5

of Ramaswami (2006)). Similarly, let ${}_0\hat{f}_{22}(q, 0, s)$ be a matrix of order $(n_2 \times n_2)$ whose ij th component is given by

$$({}_0\hat{f}_{22}(q, 0, s))_{ij} = e^{H(s)q_i}, \quad i \in S_2, j \in S_2.$$

The ij th component $({}_0\hat{f}_{22}(q, 0, s))_{ij}$ represents the LST of the time until the process hits level 0 in state $j \in S_2$, given $\mathcal{J}(0) = i \in S_2, I(0) = q_i$. Denote by $\hat{f}(q, s)$ an $(n \times n_2)$ matrix,

$$\hat{f}(q, s) = \begin{bmatrix} {}_0\hat{f}_{12}(q, 0, s) \\ {}_0\hat{f}_{22}(q, 0, s) \end{bmatrix}.$$

Let A_2 be an $(n_2 \times n)$ matrix given by $A_2 = [A_{21}, A_{22}]$ (from last paragraph of Section 3.2, A_{ij} means the sub-matrix of A with row indicates in S_i and column indicates in S_j).

Lemma 4.1. *The total expected discounted order cost vector $\hat{O}(\beta)$ of order $(n \times 1)$ satisfies the following equation:*

$$\hat{O}(\beta) = (I - \hat{f}(q, \beta)A_2)^{-1} \Delta_{K+cq} e. \tag{4.1}$$

Proof. Recall that L is the time of the next order. We can write $\hat{O}(\beta)$ as

$$\hat{O}(\beta) = \Delta_{K+cq} e + E(e^{-\beta L}) A_2 \hat{O}(\beta). \tag{4.2}$$

We use $E(e^{-\beta L})$ as a shorthand notation for the $(n \times n_2)$ matrix whose ij th component is given by

$$\mathbb{E}\{e^{-\beta L}\}_{ij} = \mathbb{E}\{e^{-\beta L} \mathbf{1}_{\{\text{level 0 hit at time } L \text{ in phase } j\}} \mid \mathcal{J}(0) = i, I(0) = q_i\}$$

Applying the fluid model $E(e^{-\beta L}) = \hat{f}(q, \beta)$ and solving (4.2) for $\hat{O}(\beta)$, we obtain (4.1).

4.2. Holding cost

The expected discounted holding cost can be expressed as

$$HC(\beta) = \mathbb{E}\left\{ \int_0^\infty h(t) e^{-\beta t} I(t) dt \right\}.$$

Let $\hat{h}(\beta)$ be an $(n \times 1)$ vector whose i th component is given by

$$\hat{h}_i(\beta) = \mathbb{E}_i\left\{ \int_0^\infty h(t) e^{-\beta t} I(t) dt \right\}$$

i.e. the expected discounted holding cost, given $\mathcal{J}(0) = i \in S$. Thus, we have $HC(\beta) = \gamma \cdot \hat{h}(\beta)$.

Lemma 4.2. *The vector $\hat{h}(\beta)$ of order $(n \times 1)$ satisfies the following equation:*

$$\hat{h}(\beta) = (I - \hat{f}(q, \beta)A_2)^{-1} \Delta_h E\left(\int_{t=0}^L e^{-\beta t} I(t) dt \right).$$

Proof. In a similar manner to that of Lemma 4.1, we can write $\hat{h}(\beta)$ as

$$\hat{h}(\beta) = \Delta_h E\left(\int_{t=0}^L e^{-\beta t} I(t) dt \right) + E(e^{-\beta L}) A_2 \hat{h}(\beta).$$

The first vector $E(\int_{t=0}^L e^{-\beta t} I(t) dt)$ is the expected discounted inventory level of the first cycle. From that point, the proof is similar to (4.2).

Now, we have to find the $(n \times 1)$ vector $E(\int_{t=0}^L e^{-\beta t} I(t) dt)$ by applying the OST to the multi-dimensional martingale. Assume that $\mathcal{J}(0) = i \in S, I(0) = q_i$. Consider a Lévy process $\{X_i(t)\}$,

$$X_i(t) = X_i(0) - \int_{s=0}^t u_{\mathcal{J}(s)} ds, \quad X_i(0) = -q_i, \quad 0 \leq t < L.$$

Note that u_i may be either negative or positive. It is not difficult to see that the latter process up to time L , i.e. $(X_i(t))_{0 \leq t < L}$, has the same distribution as $(-I_i(t))_{0 \leq t < L}$. From Chapter XI, p. 311 of Asmussen (2003) it follows that

$$\mathbb{E}_i\{e^{\alpha X_i(t)}; \mathcal{J}(t) = j\} = (e^{tK(\alpha)})_{ij},$$

where

$$K(\alpha) = Q - \alpha U.$$

Let $Y(t) = -(\beta/\alpha)t$ (for an arbitrary $\alpha > 0$) and let $Z_i(t) = X_i(t) + Y(t)$. Since $Y(t)$ is adapted and has paths of a finite expected variation, the process

$$\begin{aligned} M_i(\alpha, t) &= \int_0^t e^{\alpha Z_i(s)} \mathbf{1}_{\{\mathcal{J}(s)\}} ds K(\alpha) + e^{\alpha Z_i(0)} \mathbf{1}_{\{\mathcal{J}(0)\}} \\ &\quad - e^{\alpha Z_i(t)} \mathbf{1}_{\{\mathcal{J}(t)\}} + \alpha \int_0^t e^{\alpha Z_i(s)} \mathbf{1}_{\{\mathcal{J}(s)\}} dY(s) \\ &= \int_0^t e^{\alpha X_i(s) - \beta s} \mathbf{1}_{\{\mathcal{J}(s)\}} ds (K(\alpha) - \beta I) \\ &\quad + e^{\alpha X_i(0)} \mathbf{1}_{\{\mathcal{J}(0)\}} - e^{\alpha X_i(t) - \beta t} \mathbf{1}_{\{\mathcal{J}(t)\}} \end{aligned} \tag{4.3}$$

is an n -dimensional row vector-valued zero mean martingale. The OST yields $\mathbb{E}M_i\{\alpha, 0\} = \mathbb{E}M_i\{\alpha, L\} = 0$, i.e.

$$\begin{aligned} \mathbb{E}_i \left\{ \int_0^L e^{\alpha X_i(s) - \beta s} ds \right\} \\ = [E_i(e^{\alpha X_i(L) - \beta L} \mathbf{1}_{\{\mathcal{J}(L)\}}) - E_i(e^{\alpha X_i(0)} \mathbf{1}_{\{\mathcal{J}(0)\}})](K(\alpha) - \beta I)^{-1} \mathbf{e}. \end{aligned} \tag{4.4}$$

Obviously,

$$E_i(e^{\alpha X_i(0)} \mathbf{1}_{\{\mathcal{J}(0)\}}) = e^{-\alpha q_i} \mathbf{e}_i,$$

or in an $(n \times n)$ matrix form

$$E(e^{\alpha X(0)} \mathbf{1}_{\{\mathcal{J}(0)\}}) = e^{-\alpha q} I = \Delta_{-\alpha q} \tag{4.5}$$

(note that we use $X(t)$ as a shorthand notation for the vector $(X_1(t), \dots, X_n(t))^T$). Next, we have to derive $E_i(e^{\alpha X_i(L) - \beta L} \mathbf{1}_{\{\mathcal{J}(L)\}})$. Note that $X_i(L) = 0$ and applying the fluid method leads to the $(n \times n)$ matrix form

$$E(e^{-\beta L} \mathbf{1}_{\{\mathcal{J}(L)\}}) = (\mathbf{0}_{n \times n_1} \quad \hat{f}(q, \beta)), \tag{4.6}$$

where $\mathbf{0}_{a \times b}$ is a shorthand notation for an $(a \times b)$ matrix with all entries equal to 0. For convenience, we will drop the size $(a \times b)$ and apply $\mathbf{0}$ in the cases that were previously mentioned (instead of $\mathbf{0}_{a \times b}$). Substituting (4.5) and (4.6) into (4.4), we obtain the $(n \times 1)$ vector

$$E\left(\int_0^L e^{\alpha X(s) - \beta s} ds\right) = \left((\mathbf{0} \quad \hat{f}(q, \beta)) - \Delta_{e^{-\alpha q}}\right)(K(\alpha) - \beta I)^{-1} \mathbf{e}. \tag{4.7}$$

Now, take the derivative of both sides of (4.7) with respect to α and let $\alpha = 0$. This leads to

$$E\left(\int_0^L e^{-\beta s} X(s) ds\right) = -E\left(\int_0^L e^{-\beta s} I(s) ds\right). \tag{4.8}$$

Accordingly, the expected discounted total cost (TC) is

$$TC(\beta) = OC(\beta) + HC(\beta).$$

5. Long run average analysis

Let $TC(t)$ be the total cost until time t . We are interested in the long-run average total cost per time unit:

$$\overline{TC} = \lim_{t \rightarrow \infty} \frac{TC(t)}{t}.$$

Note that (\mathcal{J}_k, T_k) is a Markov renewal process, and the content level process $I(t)$ is a semi-regenerative process with respect to (\mathcal{J}_k, T_k) . The T_k s are semi-regenerative points of the process. Thus, the process $\mathcal{J}_k, k = 0, 1, 2, \dots$ is an irreducible positive recurrent Markov chain with transition probability matrix \mathbb{P}^* and stationary probability vector π^* where

$$\pi^* \mathbb{P}^* = \pi^*, \quad \pi^* \mathbf{e} = 1.$$

Recall that L denotes the first cycle length and let $\mathbb{E}\{L\} = \pi^* \mathbf{E}(L)$ where $\mathbf{E}(L)$ is an $(n \times 1)$ vector whose i th component is $\mathbb{E}_i\{L\} = \mathbb{E}\{L \mid \mathcal{J}_0 = i \in S\}$. Denote by $N(t) = \sup\{n \geq 0 \mid L_0 + \dots + L_n \leq t\}, t \geq 0$, the corresponding counting process. Let

$$OC(t) = \sum_{k=0}^{N(t)} (K_{\mathcal{J}_k} + c_{\mathcal{J}_k} q_{\mathcal{J}_k})$$

be the order cost up to time t and let

$$HC(t) = \int_0^t h(s) I(s) ds$$

be the inventory cost up to time t . Thus, the long-run average total cost per time unit

$$\lim_{t \rightarrow \infty} \frac{TC(t)}{t} = \lim_{t \rightarrow \infty} \left(\frac{OC(t)}{t} + \frac{HC(t)}{t} \right).$$

Denote by $\overline{\mathbf{O}}$ an $(n \times 1)$ vector whose i th component is $\overline{O}_i = K_i + c_i q_i, i \in S$. Similarly, denote by \overline{hc} an $(n \times 1)$ vector whose i th component is $\overline{hc}_i = \mathbb{E}_i\{\int_0^L h_i I(t) dt\}$. From Theorem 3.1 and Proposition 5.2 of Asmussen (2003) (see also Theorem 3.6.1 of Ross (1996)) it follows that:

$$\lim_{t \rightarrow \infty} \frac{OC(t)}{t} = \frac{\sum_i \pi_i^* (K_i + c_i q_i)}{\mathbb{E}\{L\}} = \frac{\pi^* \overline{\mathbf{O}}}{\mathbb{E}\{L\}}, \tag{5.1}$$

$$\lim_{t \rightarrow \infty} \frac{HC(t)}{t} = \frac{\sum_i \pi_i^* \mathbb{E}_i\{h_i \int_0^L I(t) dt\}}{\mathbb{E}\{L\}} = \frac{\pi^* \overline{hc}}{\mathbb{E}\{L\}}. \tag{5.2}$$

To complete the derivation of the long-run analysis, we have to find $\pi^*, \mathbb{E}\{L\}$, and \overline{hc} .

5.1. Determination of π^*

Lemma 5.1. *The stationary probability vector π^* satisfies the following equations:*

$$\pi^* = \pi^* \hat{f}(q, 0)A_2, \quad \pi^* e = 1. \tag{5.3}$$

Proof. Assume that $\mathcal{J}(0) = r \in S, I(0) = q_r$. The probability that at order point the state is $j \in S_2$ is $\hat{f}(q, 0)_{rj}$. From that point the process switches to state i with probability α_{ji} and restarts from level q_i . Hence, we obtain

$$\pi_i^* = \sum_{r \in S} \pi_r^* \sum_{j \in S_2} (\hat{f}(q, 0)_{rj} \alpha_{ji}), \quad \sum_{i \in S} \pi_i^* = 1. \tag{5.4}$$

Applying (5.4) in a matrix form leads to (5.3).

5.2. The expected cycle length

Assume that $i \in S$. In order to obtain $\mathbb{E}_i\{L\}$, we apply the theory of the Markov Additive Process (MAP) as introduced in Section XI, p. 312–313 of Asmussen (2003). Consider a Lévy process

$$\tilde{X}_i(t) = \tilde{X}_i(0) + \int_{s=0}^t u_{\mathcal{J}(s)} ds, \quad \tilde{X}_i(0) = 0. \tag{5.5}$$

Proposition 2.2 of Asmussen (2003) reads:

$$\mathbb{E}_i\{e^{\alpha \tilde{X}_i(t)}; \mathcal{J}(t) = j\} = (e^{t\tilde{K}(\alpha)})_{ij},$$

where

$$\tilde{K}(\alpha) = Q + \alpha U. \tag{5.6}$$

The matrix $\tilde{K}(\alpha)$ has a real eigenvalue $\tilde{k}(\alpha)$ with maximal real part (cf. Chapter XI, p. 312 of Asmussen (2003)). The corresponding left and right eigenvectors v^α and h^α may be chosen with strictly positive components. Moreover, without loss of generality, it can be assumed that $v^\alpha h^\alpha = 1$, and that $\pi h^\alpha = 1$ where $\pi = v^0$ is the stationary distribution of $\{\mathcal{J}(t)\}$ and h^0 is a column vector of 1s. Let h' be the derivative of h^α at $\alpha = 0$. Corollary 2.6 of Asmussen (2003) implies that for any stopping time L with $\mathbb{E}_i\{L\} < \infty$,

$$\mathbb{E}_i\{\tilde{X}_i(L)\} = \tilde{k}'(0)\mathbb{E}_i\{L\} + h'_i - \mathbb{E}_i\{h'_{\mathcal{J}(L)}\}. \tag{5.7}$$

By Corollaries 2.7 and 2.8 of Asmussen (2003):

$$\tilde{k}'(0) = \sum_{i \in S} \pi_i u_i.$$

To find the vector h' , we apply the same technique as in Asmussen and Kella (2000, p. 385). For $\tilde{k}(\alpha)$ the eigenvalue with the largest real part of $\tilde{K}(\alpha)$ and h^α as the corresponding right-normalized eigenvector, we have

$$\tilde{k}(\alpha)h^\alpha = \tilde{K}(\alpha)h^\alpha. \tag{5.8}$$

Taking the derivative of (5.8) with respect to α , let $\alpha = 0$ and noting that $\tilde{k}(0) = 0, h^0 = e, \tilde{K}'(0) = U$, and $\tilde{K}(0) = Q$ yields

$$\tilde{k}'(0)e = Qh' + Ue. \tag{5.9}$$

Note that $\pi \mathbf{h}^\alpha = 1$, thus, $\pi \mathbf{h}' = 0$. Subtracting $e\pi \mathbf{h}' = 0$ from the two sides of (5.9) yields

$$\mathbf{h}' = (\mathbf{Q} - e\pi)^{-1}[k'(0) - \mathbf{U}]e.$$

Now, let \mathbf{h}'_2 be the sub-vector of \mathbf{h}' of order $(n_2 \times 1)$ (with all entries corresponding to states in S_2). Regarding (5.7), we have

$$\mathbb{E}_i\{\tilde{X}_i(L)\} = -q_i, \quad \mathbb{E}_i\{h'_{g(L)}\} = e_{i0}\hat{f}(q, 0)\mathbf{h}'_2. \tag{5.10}$$

Substituting (5.10) into (5.7), we obtain the $(n \times 1)$ vector of expectations

$$\mathbf{E}(L) = \frac{{}_0\hat{f}(q, 0)\mathbf{h}'_2 - \mathbf{h}' - \Delta_q e}{k'(0)}. \tag{5.11}$$

5.3. Holding cost

We use the notation from Section 4.2 to obtain the n -column vector $\overline{\mathbf{h}\mathbf{c}}$ whose i th component is $\mathbb{E}_i\{\int_0^L h_i I(t) dt\}$. Substitute $\beta = 0$ in (4.7) and take the derivative with respect to α . Then set $\alpha = 0$. Since $\mathbf{K}(\alpha)$ is singular at $\alpha = 0$, we conclude that

$$\frac{d}{d\alpha}([\mathbf{0} \quad \hat{f}(q, 0)] - \Delta_{e^{-\alpha q}}]\mathbf{K}(\alpha)^{-1}e \longrightarrow \mathbf{E}\left(\int_0^L X(s) ds\right) \tag{5.12}$$

as $\alpha \rightarrow 0$ through values for which $\mathbf{K}(\alpha)$ is nonsingular. Relation (5.12) amounts to applying twice L'Hôpital's rule to determine the $(n \times 1)$ vector $\mathbb{E}\{\int_0^L X(s) ds\}$ and, we obtain

$$\overline{\mathbf{h}\mathbf{c}} = -\Delta_h \mathbf{E}\left(\int_0^L X(s) ds\right).$$

Example 5.1. We consider a Markov chain with $n = 4$ states with a generator matrix \mathbf{Q} and a probability matrix \mathbf{A} ,

$$\mathbf{Q} = \begin{pmatrix} -7 & 1.5 & 2.5 & 3 \\ 3 & -9 & 2 & 4 \\ 2 & 4 & -9 & 3 \\ 4 & 3 & 2 & -9 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0.7 & 0.1 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.15 & 0.05 \\ 0.02 & 0.08 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.2 & 0.7 \end{pmatrix}.$$

The initial probability vector is $\gamma = (0.2, 0.3, 0.35, 0.15)$. The vector of the net production rates is $u = (1, 1.5, -1.5, -2)$, thus, $S_1 = \{1, 2\}$ and $S_2 = \{3, 4\}$. For simplicity, we assume that the costs and quantities are equal; $k_i = 40$ and $c_i = 5$ for $i \in S$ and a discount factor $\beta = 0.01$. We assume that $h_i = h$ for $i \in S$ and we let h vary in $\{0.25, 0.5, 0.75, 1, 1.5\}$ and q vary in $\{2, 3, 4, \dots, 11, 12\}$. In Figure 2(a), we present the expected discounted total cost, $\text{TC}(\beta)$, as a function of the ordered quantity q for values of h and in Figure 2(b), we present the long-run average cost per time unit, $\overline{\text{TC}}$, as a function of q for several values of h . We see that the behavior of the plots in both figures is similar to the basic EOQ behavior; both costs, $\text{TC}(\beta)$ and $\overline{\text{TC}}$, appear to be convex functions. Moreover, we can conclude that for each holding cost there is an optimal order quantity which increases as the holding cost decreases.

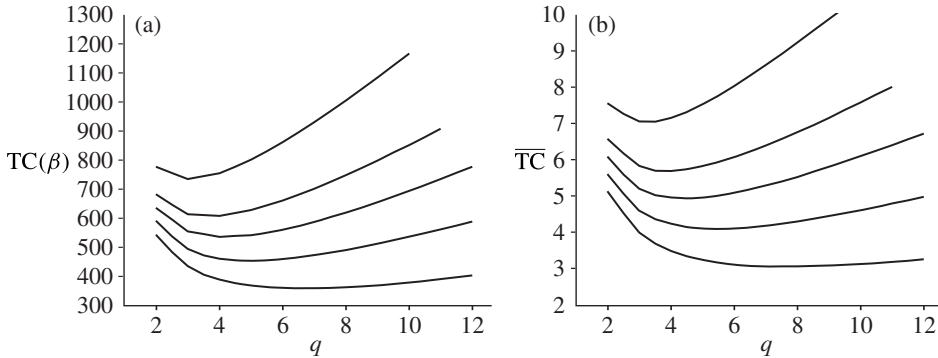


FIGURE 2: (a) The expected discounted total cost, $TC(\beta)$, as a function of the order quantity, q . (b) The long-run average cost per time unit, \overline{TC} , as a function of the order quantity, q . The values for h are, ascending vertically, 0.25, 0.5, 0.75, 1, and 1.5.

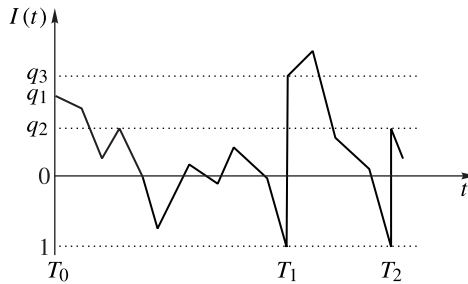


FIGURE 3: A sample path of the inventory process with backlogging.

6. Inventory model with backlogging

In the previous sections we considered a model where we place an order as soon as the inventory on hand down-crosses level zero. Many businesses find it practical to operate with planned backlogging. In this section we consider the same model as in the sections above, but allow backlogging, and assume that unsatisfied demands are fully backlogged. Let $I(t)$ be the net inventory level at time t (i.e. the inventory on hand at time t minus backorders at time t). We always use any inventory on hand to fill demands; backorders accumulate only when we run out of stock entirely. Thus, if $I(t)$ is positive, it represents the amount of inventory on hand. If it is negative, it represents the negative amount of backorders at time t . Now, besides all the previous costs, there is also a backlogging cost. We consider a policy under which we place an order of size $(q_i - l)$ whenever the inventory level decreases to a reorder point l ($l < 0$). Thus, we start a new cycle with state $i \in S$ and with an inventory level q_i . We assume zero lead times, so the orders arrive instantaneously and the inventory level jumps to q_i . In Figure 3, we illustrate a typical sample path of the inventory process.

Note that the time of the k th jump is $T_k = \inf\{t > T_{k-1}; I(t) = l\}$ and $T_0 = 0$. Obviously, $\mathcal{J}_k = \mathcal{J}(T_k)$ is the environmental state at T_k (just after the jump), (\mathcal{J}_k, T_k) is a Markov renewal

process with T_k s as the semi-regenerative points. Let $L_k^l = T_k - T_{k-1}$ and $L_0^l = 0$ be the inter-replenishment times, and let $L^l = L_1^l$. We assume that during a cycle k in which $\mathcal{J}_k = i$, the shortage cost for one item during a time interval of length dt equals $b_i dt$ (i.e. $b(t) = b_i$ for $t \in [T_k, T_{k+1})$ and $\mathcal{J}_k = i$).

Without loss of generality, we assume that the inventory level at time 0 has a shortage of l items ($I(0^-) = l$). Given $\mathcal{J}(0) = i \in S$ we order $(q_i - l)$ items and, hence, $I(0) = q_i$.

6.1. The expected discounted analysis

Let $OC(\beta, l)$, $HC(\beta, l)$, and $BC(\beta, l)$ be the expected discounted order cost, holding cost, and shortage cost as a function of the reorder point l , respectively. The expected discounted total cost is given by

$$TC(\beta, l) = OC(\beta, l) + HC(\beta, l) + BC(\beta, l).$$

Similar to Section 4, let $\hat{O}(\beta, l)$, $\hat{h}(\beta, l)$, and $\hat{B}(\beta, l)$ be $(n \times 1)$ vectors whose i th component is the expected discounted cost given $\mathcal{J}(0) = i, I(0) = q_i$. Clearly, $OC(\beta, l) = \gamma \hat{O}(\beta, l)$, $HC(\beta, l) = \gamma \hat{h}(\beta, l)$, and $BC(\beta, l) = \gamma \hat{B}(\beta, l)$. Note that the order cost $\hat{O}(\beta, l)$ vector is obtained directly from Lemma 4.1 by setting $(q_i - l)$ instead of q_i in (4.1). The computation of the holding and backlogging costs are carried out in the next two subsections.

6.1.1. *Holding cost.* We modify the results of Lemma 4.2 to obtain

$$\hat{h}(\beta, l) = (I - \hat{f}(q - l, \beta)A_2)^{-1} \Delta_h E \left(\int_{t=0}^{L^l} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0\}} dt \right).$$

Regarding L , the first time that the process down-crosses level 0 (see Section 2), we obtain

$$\begin{aligned} & E \left(\int_{t=0}^{L^l} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0\}} dt \right) \\ &= E \left(\int_{t=0}^L e^{-\beta t} I(t) dt \right) + E(e^{-\beta L}) E \left(\int_{t=0}^{L^l - L} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0\}} dt \right). \end{aligned} \tag{6.1}$$

Note that $E(e^{-\beta L}) = \hat{f}(q, \beta)$. The first term of (6.1) is given by (4.7) and (4.8). In order to derive the right side term of (6.1), we shift the time origin to L . Denote by $\tilde{L} = L^l - L$ and let \tilde{E} be an $(n_2 \times 1)$ vector

$$\tilde{E} = E \left(\int_{t=0}^{\tilde{L}} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0\}} dt \right).$$

We define two sets of stopping times $\{\tilde{L}_k\}_{k=0,1,\dots}$ and $\{\tilde{\tau}_k\}_{k=1,2,\dots}$ during the interval $[0, \tilde{L}]$ by

$$\begin{aligned} \tilde{L}_0 &= 0, & \tilde{L}_k &= \inf\{t : t > \tilde{L}_{k-1}, I(t) = 0, \mathcal{J}(t) \in S_2\}, \\ \tilde{\tau}_k &= \inf\{t : t > \tilde{L}_{k-1}, I(t) = 0, \mathcal{J}(t) \in S_1 \text{ or } I(t) = l, \mathcal{J}(t) \in S_2\} \end{aligned}$$

i.e. \tilde{L}_k is the k th time that the process down-crosses level 0 and $\tilde{\tau}_k$ is the k th time that the process up-crosses level 0 or hits level l . Note that $\{\tilde{L}_k\}_{k=0,1,\dots}$ are semi-regenerative points during the interval $[0, \tilde{L}]$. Obviously, $\tilde{\tau}_k \leq \tilde{L}_k$ for $k > 0$.

Proposition 6.1. *The $(n_2 \times 1)$ vector \tilde{E} satisfies the following matrix equation:*

$$\tilde{E} = [I - {}^{-l}\Psi^r(\beta)\Psi(\beta)]^{-1} {}^{-l}\Psi^r(\beta) \mathbf{E} \left(\int_0^{\tilde{L}_1 - \tilde{\tau}_1} e^{-\beta t} I(t) dt \right). \tag{6.2}$$

Proof. We reveal two cases after time $\tilde{L}_0 = 0$.

- (i) In the first case, the process up-crosses level 0 before it hits level l (thus, $I(\tilde{\tau}_1) = 0$). In this case, the interval $[0, \tilde{L}]$ can be divided into $[0, \tilde{\tau}_1) \cup [\tilde{\tau}_1, \tilde{L}_1) \cup [\tilde{L}_1, \tilde{L}]$. From time \tilde{L}_1 , the process starts again (as \tilde{L}_1 is a semi-regenerative point).
- (ii) In the second case, the process hits level l (thus, $I(\tilde{\tau}_1) = l$). In this case $\tilde{\tau}_1 = \tilde{L}$ and $[0, \tilde{L}) = [0, \tilde{\tau}_1)$.

Note that

$$\mathbf{E} \left(\int_{t=0}^{\tilde{\tau}_1} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0, \tilde{\tau}_1 < \tilde{L}\}} dt \right) = 0$$

and

$$\mathbf{E} \left(\int_{t=0}^{\tilde{\tau}_1} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0, \tilde{\tau}_1 = \tilde{L}\}} dt \right) = 0.$$

Applying the fluid parameters $({}^{-l}\Psi^r(\beta)$ and $\Psi(\beta)$, see Table 3), we obtain

$$\begin{aligned} \tilde{E} &= \mathbf{E} \left(\int_{t=\tilde{\tau}_1}^{\tilde{L}_1} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) > 0, \tilde{\tau}_1 < \tilde{L}\}} dt \right) + \mathbf{E}(e^{-\beta \tilde{L}_1} \mathbf{1}_{\{\tilde{\tau}_1 < \tilde{L}\}}) \tilde{E} \\ &= {}^{-l}\Psi^r(\beta) \mathbf{E} \left(\int_0^{\tilde{L}_1 - \tilde{\tau}_1} e^{-\beta t} I(t) dt \right) + {}^{-l}\Psi^r(\beta) \Psi(\beta) \tilde{E}. \end{aligned} \tag{6.3}$$

Solving (6.3), we obtain (6.2).

To complete our derivation, let $\tilde{\Psi}$ and \tilde{I} be $(n_1 \times n)$ matrices,

$$\tilde{\Psi} = (\mathbf{0}_{n_1 \times n_1} \quad \Psi(\beta)), \quad \tilde{I} = (I \quad \mathbf{0}_{n_1 \times n_2}).$$

Lemma 6.1. *The $(n_1 \times 1)$ vector $\mathbf{E}(\int_0^{\tilde{L}_1 - \tilde{\tau}_1} e^{-\beta t} I(t) dt)$ satisfies the following:*

$$\frac{d}{d\alpha} \left[(\tilde{\Psi} - \tilde{I})(\mathbf{K}(\alpha) - \beta I)^{-1} \mathbf{e} \right] \Big|_{\alpha=0} \longrightarrow -\mathbf{E} \left(\int_0^{\tilde{L}_1 - \tilde{\tau}_1} e^{-\beta t} I(t) dt \right). \tag{6.4}$$

Proof. By shifting the time original to $\tilde{\tau}_1$, it is easy to modify the method of Section 4.2. For this, we apply again the optimal sampling theory to the martingale (4.3), this time to the stopping time $\tilde{L}_1 - \tilde{\tau}_1$ with $q_i = 0$ (see also (4.4)–(4.7)) and obtain

$$\mathbf{E} \left(\int_0^{\tilde{L}_1 - \tilde{\tau}_1} e^{\alpha X(t) - \beta t} I(t) dt \right) = (\tilde{\Psi} - \tilde{I})(\mathbf{K}(\alpha) - \beta I)^{-1} \mathbf{e}. \tag{6.5}$$

Now, derive (6.5) with respect to α and setting $\alpha = 0$ leads to (6.4).

6.1.2. *Backlogging cost.* We modify the renewal theory and the result of Lemma 4.2 yields

$$\hat{\mathbf{B}}(\beta, l) = -(I - \hat{f}(q - l, \beta)A_2)^{-1} \Delta_b \mathbf{E} \left(\int_{t=0}^{L^l} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) < 0\}} dt \right).$$

Obviously, $I(t) > 0$ during the time interval $[0, L]$. Regarding $\tilde{L} = L^l - L$, we introduce the $(n_2 \times 1)$ vector

$$\tilde{\mathbf{B}} = \mathbf{E} \left(\int_{t=0}^{\tilde{L}} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) < 0\}} dt \right).$$

The term $\tilde{\mathbf{B}}$ represents the expected discounted shortage level during $[0, \tilde{L}]$. We obtain

$$\mathbf{E} \left(\int_{t=0}^{L^l} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) < 0\}} dt \right) = \hat{f}(q, \beta) \tilde{\mathbf{B}}. \tag{6.6}$$

Theorem 6.1. *The vector $\tilde{\mathbf{B}}$ satisfies the following matrix equation:*

$$\tilde{\mathbf{B}} = [I - {}^{-l}\Psi^r(\beta)\Psi(\beta)]^{-1} \mathbf{E} \left(\int_{t=0}^{\tilde{\tau}_1} e^{-\beta t} I(t) dt \right). \tag{6.7}$$

Proof. Similar to the proof of Proposition 6.1 (note that $\mathbf{E}(\int_{\tilde{\tau}_1}^{\tilde{L}} e^{-\beta t} I(t) \mathbf{1}_{\{I(t) < 0, \tilde{\tau}_1 < \tilde{L}\}} dt) = 0$).

For the final determination of (6.7), we have to derive $\mathbf{E}(\int_{t=0}^{\tilde{\tau}_1} e^{-\beta t} I(t) dt)$. Applying the Lévy process $\{\tilde{X}_i(t)\}$ for $i \in S_2$ (see (5.5) and (5.6)), it is not difficult to see that the latter process up to time $\tilde{\tau}_1$, i.e. $(\tilde{X}_i(t))_{0 \leq t < \tilde{\tau}_1}$, has the same distribution as $(I(t))_{0 \leq t < \tilde{\tau}_1}$. For $Y(t) = -(\beta/\alpha)t$ let $Z_i(t) = \tilde{X}_i(t) + Y(t)$. Applying Theorem 2.1 of Asmussen and Kella (2000) yields that the process

$$\mathbf{M}_i(\alpha, t) = \int_0^t e^{\alpha \tilde{X}_i(s) - \beta s} \mathbf{1}_{\{\mathcal{J}(s)\}} ds (\tilde{\mathbf{K}}(\alpha) - \beta I) + e^{\alpha \tilde{X}_i(0)} \mathbf{1}_{\{\mathcal{J}(0)\}} - e^{\alpha \tilde{X}_i(t) - \beta t} \mathbf{1}_{\{\mathcal{J}(t)\}}$$

is an n -dimensional row vector-valued zero mean martingale. The OST yields $\mathbb{E}\mathbf{M}_i\{\alpha, 0\} = \mathbb{E}\mathbf{M}_i\{\alpha, \tilde{\tau}_1\} = 0$. So, we obtain

$$\mathbf{E} \left[\int_0^{\tilde{\tau}_1} e^{\alpha \tilde{X}(s) - \beta s} \mathbf{1}_{\{\mathcal{J}(s)\}} ds \right] (\tilde{\mathbf{K}}(\alpha) - \beta I) = \mathbf{E}(e^{\alpha \tilde{X}(\tilde{\tau}_1) - \beta \tilde{\tau}_1} \mathbf{1}_{\{\mathcal{J}(\tilde{\tau}_1)\}}) - \mathbf{E}(e^{\alpha \tilde{X}(0)} \mathbf{1}_{\{\mathcal{J}(0)\}}). \tag{6.8}$$

Since $\tilde{X}_i(0) = 0$, the $(n_2 \times n)$ matrix $\mathbf{E}(e^{\alpha \tilde{X}(0)} \mathbf{1}_{\{\mathcal{J}(0)\}}) = (\mathbf{0}_{n_2 \times n_1} \ l)$. Regarding the two cases ($\tilde{\tau}_1 < \tilde{L}$ and $\tilde{\tau}_1 = \tilde{L}$) and applying the fluid model yields to

$$\begin{aligned} \mathbf{E}(e^{\alpha \tilde{X}(\tilde{\tau}_1) - \beta \tilde{\tau}_1} \mathbf{1}_{\{\mathcal{J}(\tilde{\tau}_1)\}}) &= \mathbf{E}(e^{-\beta \tilde{\tau}_1} \mathbf{1}_{\{\mathcal{J}(\tilde{\tau}_1), \tilde{\tau}_1 < \tilde{L}\}}) + \mathbf{E}(e^{\alpha l} e^{-\beta \tilde{\tau}_1} \mathbf{1}_{\{\mathcal{J}(\tilde{\tau}_1), \tilde{\tau}_1 = \tilde{L}\}}) \\ &= ({}^{-l}\Psi^r(\beta) \ \mathbf{0}_{n_2 \times n_2}) + e^{\alpha l} (\mathbf{0}_{n_2 \times n_1} \ 0 \hat{f}_{22}^r(0, -l, \beta)). \end{aligned} \tag{6.9}$$

(Note that for $\tilde{\tau}_1 < \tilde{L}$, $\tilde{X}(\tilde{\tau}_1) = 0$ and for $\tilde{\tau}_1 = \tilde{L}$, $\tilde{X}(\tilde{\tau}_1) = l$.) Substituting (6.9) into (6.8), multiplying by the vector \mathbf{e} , we arrive at

$$\begin{aligned} \mathbf{E} \left(\int_0^{\tilde{\tau}_1} e^{\alpha \tilde{X}(s) - \beta s} ds \right) \\ = [({}^{-l}\Psi^r(\beta) \ \mathbf{0}) + e^{\alpha l} (\mathbf{0} \ 0 \hat{f}_{22}^r(0, -l, \beta)) - (\mathbf{0} \ I)] (\tilde{\mathbf{K}}(\alpha) - \beta I)^{-1} \mathbf{e}. \end{aligned} \tag{6.10}$$

Now, we can determine the $(n_2 \times 1)$ vector $\mathbf{E}(\int_0^{\tilde{\tau}_1} e^{-\beta t} I(t) dt)$ by taking the derivative of (6.10) with respect to α and setting $\alpha = 0$.

6.2. The long-run average analysis

Throughout this section we apply the same methods as in Section 5. We use the subscript l to denote quantities similar to those above. Similar to (5.1) and (5.2), we obtain the following proposition.

Proposition 6.2. *It holds that*

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{OC_l(t)}{t} &= \frac{\sum_i (\pi_l^*)_i (K_i + c_i (q_i - l))}{\mathbb{E}\{L^l\}} = \frac{\pi_l^* \bar{O}_l}{\mathbb{E}\{L^l\}}, \\ \lim_{t \rightarrow \infty} \frac{HC_l(t)}{t} &= \frac{\sum_i (\pi_l^*)_i \mathbb{E}_i \left\{ \int_0^{L^l} h_i I(t) \mathbf{1}_{\{I(t) > 0\}} dt \right\}}{\mathbb{E}\{L^l\}} = \frac{\pi_l^* \Delta_h \bar{h}c_l}{\mathbb{E}\{L^l\}}, \\ \lim_{t \rightarrow \infty} \frac{BC_l(t)}{t} &= \frac{-\sum_i (\pi_l^*)_i \mathbb{E}_i \left\{ \int_0^{L^l} b_i I(t) \mathbf{1}_{\{I(t) < 0\}} dt \right\}}{\mathbb{E}\{L^l\}} = \frac{-\pi_l^* \Delta_b \bar{b}c_l}{\mathbb{E}\{L^l\}}. \end{aligned}$$

We now have explicit equations for all the components above.

- (a) The stationary probability vector π_l^* is obtained by (5.4) with $\hat{f}(q - l, 0)$ replacing $\hat{f}(q, 0)$.
- (b) The average cycle length $\mathbb{E}\{L^l\}$ is obtained by (5.11) with $q - l$ replacing q .
- (c) To obtain $\bar{h}c_l = \mathbf{E}(\int_0^{L^l} I(t) \mathbf{1}_{\{I(t) > 0\}} dt)$ set $\beta = 0$ in (6.1) (and, thus, in (4.7)), (6.2), and (6.4). Then apply twice L'Hôpital's rule (similar to Section 5.3).
- (d) Similarly, to obtain $\bar{b}c_l = \mathbf{E}(\int_0^{L^l} I(t) \mathbf{1}_{\{I(t) < 0\}} dt)$ set $\beta = 0$ in (6.6), (6.7), and (6.10) and apply twice L'Hôpital's rule.

Example 6.1. We consider the same data as in Example 5.1 with $h_i = 0.5$ and $q_i = 5$ for $i \in S$. We assume that $b_i = b$ for $i \in S$ and let b vary in $\{0.1, 0.25, 0.5, 1, 2, 5\}$ and $(-l)$ vary in $\{0, 1, 2, 3, \dots, 11\}$. In Figure 4(a), we present the expected discounted total cost, $TC(\beta, l)$, as a function of the order point l for values of b and in Figure 4(b), we present the long-run average cost per time unit, \bar{TC}_l , as a function of l for several values of b .

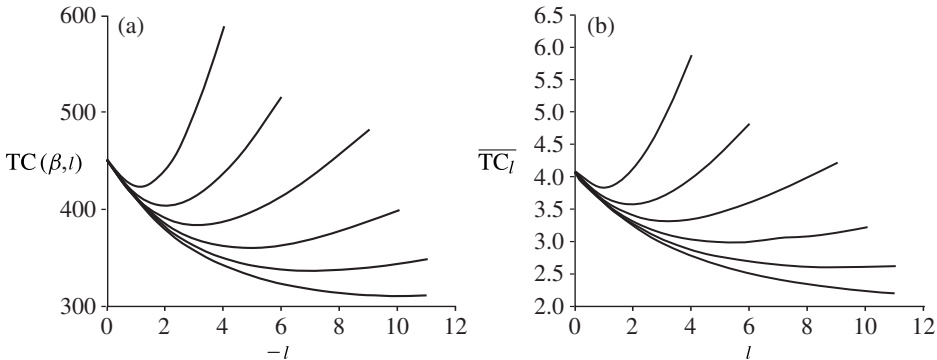


FIGURE 4: (a) The expected discounted total cost, $TC(\beta, l)$, as a function of l . (b) The long-run average cost per time unit, \bar{TC}_l , as a function of l . The values of b are, ascending vertically, 0.1, 0.25, 0.5, 1, 2, and 5.

We see that the behavior of the plots in both figures is similar. Again, both costs, $TC(\beta, l)$ and \overline{TC}_l , appear to be convex functions. Moreover, we can conclude that as b increases, the optimal value of $-l$ decreases (due to the high cost of the shortage). Obviously, the case of the model without backlogging is a special case when $l = 0$.

Acknowledgements

The author is grateful to Professors David Perry, Onno Boxma, and the anonymous referee for helpful comments and suggestions that led to substantial improvements of the manuscript.

References

- AHN, S. AND RAMASWAMI, V. (2003). Fluid flow models and queues—a connection by stochastic coupling. *Stoch. Models* **19**, 325–348.
- AHN, S. AND RAMASWAMI, V. (2005). Efficient algorithms for transient analysis of stochastic fluid flow models. *J. Appl. Prob.* **42**, 531–549.
- AHN, S. AND RAMASWAMI, V. (2006). Transient analysis of fluid models via elementary level-crossing arguments. *Stoch. Models* **22**, 129–147.
- AHN, S., BADESCU, A. L. AND RAMASWAMI, V. (2007). Time dependent analysis of finite buffer fluid flows and risk models with a dividend barrier. *Queueing Systems* **55**, 207–222.
- ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd edn. Springer, New York.
- ASMUSSEN, S. AND KELLA, O. (2000). A multi-dimensional martingale for Markov additive processes and its applications. *Adv. Appl. Prob.* **32**, 376–393.
- BEAN, N. G. AND O'REILLY, M. M. (2008). Performance measure of a multi-layer Markovian fluid model. *Ann. Operat. Res.* **160**, 99–120.
- BEAN, N. G., O'REILLY, M. M. AND TAYLOR, P. G. (2008). Algorithms for the Laplace–Stieltjes transforms of first return times for stochastic fluid flows. *Methodol. Comput. Appl. Prob.* **10**, 381–408.
- BERMAN, O., PARLAR, M., PERRY, D. AND POSNER, M. J. M. (2005). Production/clearing models under continuous and sporadic review. *Methodol. Comput. Appl. Prob.* **7**, 203–224.
- BERMAN, O. AND PERRY, D. (2006). An EOQ model with state-dependent demand rate. *Europ. J. Operat. Res.* **171**, 255–272.
- BERMAN, O., PERRY, D. AND STADJE, W. (2006). A fluid EOQ model with a two-state random environment. *Prob. Eng. Inf. Sci.* **20**, 329–349.
- BERMAN, O., PERRY, D. AND STADJE, W. (2008). Optimal replenishment in a Brownian motion EOQ model with hysteretic parameter changes. *Internat. J. Inventory Res.* **1**, 1–19.
- DOOB, J. L. (1953). *Stochastic Processes*. John Wiley, New York.
- KELLA, O., PERRY, D. AND STADJE, W. (2003). A stochastic clearing model with a Brownian and a compound Poisson component. *Prob. Eng. Inf. Sci.* **17**, 1–22.
- KULKARNI, V. G. (2010). *Modeling and Analysis of Stochastic Systems*, 2nd edn. CRC Press, Boca Raton, FL.
- KULKARNI, V. AND YAN, K. (2007). A fluid model with upward jumps at the boundary. *Queueing Systems* **56**, 103–117.
- NAHMAS, S. (1997). *Production and Operations Analysis*, 3rd edn. Irwin, Chicago, IL.
- PERRY, D., BERG, M. AND POSNER, M. J. M. (2001). Stochastic models for broker inventory in dealership markets with a cash management interpretation. *Insurance Math. Econom.* **29**, 23–34.
- PERRY, D., STADJE, W. AND ZACKS, S. (2005). Sporadic and continuous clearing policies for a production/inventory system under an M/G demand process. *Math. Operat. Res.* **30**, 354–368.
- RAMASWAMI, M. V. (1999). Matrix analytic methods for stochastic fluid flows. *Proceedings of the International Teletraffic Congress, ITC-16*, Edinburgh, Elsevier, pp. 1019–1030.
- RAMASWAMI, V. (2006). Passage times in fluid models with application to risk processes. *Methodol. Comput. Appl. Prob.* **8**, 497–515.
- ROSS, S. M. (1996). *Stochastic Processes*, 2nd edn. John Wiley, New York.
- TZENOVA, E. I., ADAN, I. J. B. F. AND KULKARNI, V. G. (2005). Fluid models with jumps. *Stoch. Models* **21**, 37–55.
- YAN, K. AND KULKARNI, V. G. (2008). Optimal inventory policies under stochastic production and demand rates. *Stoch. Models* **24**, 173–190.