

On the use of phase-type distributions for inventory management with supply disruptions

Bariş Balcıoğlu^{a*†} and Ülkü Gürler^b

Maintaining the continuity of operations becomes increasingly important for systems that are subject to disruptions due to various reasons. In this paper, we study an inventory system operating under a (q, r) policy, where the supply can become inaccessible for random durations. The availability of the supply is modeled by assuming a single supplier that goes through ON and OFF periods of stochastic duration, both of which are modeled by phase-type distributions (PTD). We provide two alternative representations of the state transition probabilities of the system, one with integral and the other employing Kolmogorov differential equations. We then use an efficient formulation for the analytical model that gives the optimal policy parameters and the long-run average cost. An extensive numerical study is conducted, which shows that OFF time characteristics have a bigger impact on optimal policy parameters. The ON time characteristics are also important for critical goods if disasters can happen. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: phase-type distributions; stochastic inventory; supply disruptions

1. Introduction

Hurricanes in the U.S.A., tsunami in Asia, and terrorist attacks are some examples from recent years which have caused serious disruptions in trade transactions and supply of materials. When Silver [1] and Nahmias [2] first pointed out the need of studying supplier unavailability, disruptions in the context of inventory models were usually considered to be arising due to strikes, shut down plants under preventive maintenance, or vendors committed fully to the contracts of other patrons. While these factors still deserve full attention, the unavoidable natural disasters or unwanted social tragedies have worsened the severity of problems when supply is inaccessible. When the problem involves non-essential consumer goods, the demand of which can display certain levels of variation due to prices set and existence of competitors, managerial tools become crucial to safeguard the continuity of high service levels until the supply chain is retained. In cases of disasters, the humanitarian concerns about providing affected people effectively with basic needs, such as water, food, heating energy, and medicine come forward. Unlike the former case, the demand for most of the items in this category hardly shows variation due to their role in our lives. Under this scenario, it is reasonable to assume that as soon as physical connections are set up, the demanded basic items can be transferred, which implies that an additional lead time may not be observed.

In this paper, we study an inventory problem with deterministic demand where supply can be disrupted at random times for random durations. In particular, we consider a single product, single location inventory system operating under the (q, r) policy. Unlike the common assumption that the inventory can be replenished any time an order is placed, we now assume that the supplier goes through ON (or is accessible) and OFF periods (or is inaccessible) of stochastic durations. Hence, when the inventory level drops to the reorder point r , if the supplier is not available or cannot be reached, the supply chain is interrupted for random durations.

The fundamental difficulty in this problem is obviously due to the uncertainty about when the supply chain will be broken and for how long the disconnect will continue. In some cases, including natural disasters like recurring hurricanes, historical data could be used to fit underlying distributions for the ON and OFF periods of the supplier. Otherwise, hypothetical distributions could be considered for examining worst-case scenarios. When general distributions are fit

^aDepartment of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Rd., Toronto, ON, Canada M5S 3G8

^bDepartment of Industrial Engineering, Bilkent University, Ankara 06800, Turkey

*Correspondence to: Bariş Balcıoğlu, Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Rd., Toronto, ON, Canada M5S 3G8.

†E-mail: baris@mie.utoronto.ca

for modeling ON and OFF periods, resulting models become analytically intractable. As a remedy, in our study we model the ON and OFF periods of the supplier by phase-type distributions (PTD) to obtain the optimal inventory policy parameters. The rationale behind this assumption is the fact that the family of PTD is sufficiently rich to cover most of the commonly used distributions in stochastic systems, such as Exponential, Erlang, and Coxian. Their flexibility and the Markovian structure also allow the use of these distributions for approximating general distributions, resulting in analytically more tractable models. Although the use of PTD is detailed in the context of an inventory problem in this study, the methods offer a more general application in other areas, such as production planning, scheduling, and reliability.

Supplier uncertainty problems of this nature have been studied by several authors. Parlar and Berkin [3] and Berk and Arreola-Risa [4] analyze the supplier uncertainty problem for the classical deterministic (EOQ) model with a single supplier whose ON and OFF periods follow an Exponential distribution. Further extensions are considered in Parlar and Perry [5], Parlar [6], Özekici and Parlar [7], Parlar [8] and Gürler and Parlar [9]. In a more recent work, Xia *et al.* [10] consider disruption management in a production and inventory problem and minimize the deviations from the original plan. Mohebbi [11] analyzes a supply uncertainty model with compound Poisson demand. The basic difference of his work from ours is that Mohebbi considers compound Poisson demands that will always have a non-zero positive variance. Hence, his model cannot simply reduce to our model of deterministic demand, which has a variance of zero. Mohebbi [11] assumes that if there is an outstanding order accepted before the supplier enters its OFF period, this order will continue being processed. Thus, during OFF periods only newly placed orders will be put on hold until the supplier becomes available again. Mohebbi and Hao [12] revise this assumption such that processing outstanding orders stops as soon as the OFF period starts and resumes from the point of interruption when the supplier enters its ON period. Mohebbi and Hao [13], on the other hand, consider that when the supplier becomes available again if processing an order was stopped during the OFF period, the production for that order has to start from scratch. Ross *et al.* [14] assume that the probability of disruption and demand intensity are time dependent. They model the ON and OFF periods by PTD, however, since the demand is assumed to be Poisson and the unmet demand is lost instead of being backordered, their model cannot be transformed to handle our problem. On top of supplier disruption, Schmitt *et al.* [15] consider yield uncertainty as well in a system operating under periodic inventory review and approximate base-stock levels. Among models considering constant demand rate, Parlar [6] studies a supplier unavailability model where the ON periods of the supplier follow an Erlang distribution. The Erlang family can be used to model random variables with coefficient of variations (variance to the squared-mean ratio) less than one. Hence, a significant subset of the distribution functions is excluded from his analysis.

In previous research [5, 6, 9], analysis was focused on modeling the supplier ON periods, where the OFF periods were either taken to be exponential, or even if it were considered to be general, the remaining OFF time was approximated by the limiting forward recurrence time at the instance when the reorder point was crossed. Hence, the real impact of the OFF period distribution was not elaborated, which by nature is a crucial component of the disruption models. Another contribution in our paper is, therefore, to incorporate the exact distribution of the remaining OFF period, if we face a disruption at the instance when we aimed at placing an order. When the OFF time is arbitrary, it is not analytically possible to establish exact results, and time-consuming simulation–optimization approaches should be resorted to. In our study, to circumvent this difficulty, we assume that the OFF periods follow the rich class of PTD and demonstrate how one can make use of the special structure of PTD to yield analytical results.

As discussed in Section 3, our analytical model uses the state-transition probabilities of the continuous-time Markov chain (CTMC) representing the ON/OFF states of the supplier. In the literature (see for instance Gürler and Parlar [9]), these probabilities are obtained by employing techniques such as the trapezoidal rule that approximate the integral equations relating the transition probabilities to one another. However, this results in big sparse matrices, which need to be inverted for each candidate inventory policy. Since the search procedure described in Section 4 evaluates the costs of a large number of candidates until the optimal policy and cost are obtained, i.e. inverts huge sparse matrices for many times, the traditional approach is highly time-consuming. Although it is not one of our primary goals in this study, we provide an alternative, yet more efficient numerical approach to obtain the state-transition probabilities. Instead of approximating the integrals, we take their Laplace transforms and then numerically invert them via an efficient approach by Jagerman [16]. This helps us by-pass inverting big matrices and obtain the state-transition probabilities consuming less time.

Our numerical findings indicate that modeling OFF periods correctly results in significant cost reductions. In particular, if suboptimal EOQ parameters are used for replenishment, the resulting costs are unacceptably high compared to the optimal policy parameters. Similarly, if suboptimal policy parameters, such as the ones obtained from a model assuming Exponential OFF period, are used for problems with non-Exponential OFF periods, costs again increase significantly, especially for moderately reliable suppliers. As the supplier availability increases or decreases, the impact of suboptimal parameters becomes less notable, which is discussed in further detail in Section 4. Yet, computing the corresponding suboptimal cost still remains as a problem if the actual OFF period distribution is not incorporated into the model.

The paper is organized as follows: Section 2 presents a brief review of the PTD. In Section 3, the model is introduced and the long run average cost function is derived and alternative methods to compute the transition probabilities of the

supplier status are presented. Numerical results are presented in Section 4. Finally, in Section 5 concluding remarks are stated.

2. PTD

A PTD of order k is considered to be the distribution of the time until absorption in a CTMC with states $\{1, \dots, k+1\}$, $(k+1)$ st (or state 0) being the single absorbing state. The general structure of PTD is depicted in Figure 1, where $c_{0,i}$ denotes the initial branching probabilities, μ_i is the rate of the sojourn time in state i , and $c_{i,j}$ denotes the transition probability from state i to state j satisfying $c_{i,1} + \dots + c_{i,k} + c_{i,k+1} = 1$ for $i = 1, \dots, k$.

A PTD r.v. can be expressed in matrix notation as well. Let $\hat{C} = (C, 0)$ with $C = (c_{0,1}, c_{0,2}, \dots, c_{0,k})$ be the initial probability vector satisfying $CE = 1$ with E being a $k \times 1$ column vector of 1. The $(k+1) \times (k+1)$ infinitesimal generator \hat{G} of this CTMC chain is then

$$\hat{G} = \begin{pmatrix} \mathbf{G} & \mathbf{G}_0 \\ \mathbf{0} & 0 \end{pmatrix}, \tag{1}$$

where \mathbf{G} is a $k \times k$ matrix with entries as $G_{ii} = -\mu_i$ for $i = 1, \dots, k$ and $G_{ij} = c_{i,j}\mu_i$ for $i \neq j$ and the $k \times 1$ \mathbf{G}_0 vector satisfies $\mathbf{G}\mathbf{E} + \mathbf{G}_0 = \mathbf{0}$ (Neuts [17, Section 2.2]). The pair (C, \mathbf{G}) is called a *representation* of the PTD.

Two special PTDs, namely the Coxian and the Hyper-exponential distributions will be considered in this study due to their popularity in modeling underlying distributions. In a Coxian distribution, once the sojourn time in phase i with exponential rate μ_i is over, the only possible transition types are either to the next state $i+1$ with probability $c_{i,i+1} = a_i$ or to the absorbing state $k+1$ with probability $c_{i,k+1} = 1 - a_i$ for $i = 1, \dots, k-1$ and $c_{k,k+1} = 1$. The graphical representation of a Coxian distribution is provided in Figure 2. In many cases, the probability a_0 is usually assumed to be 1.

Two-stage Coxian distributions with $a_1 \leq 1$ are practical for modeling or approximating distributions with squared-coefficient of variation (variance to the squared-mean ratio) $c^2 \geq 0.5$. Another special case of the Coxian distribution is the generalized Erlang distribution with $a_i = 1$ for $i = 1, \dots, k$. In case $\mu_i = \mu$, the generalized Erlang reduces to the well-known k -stage Erlang distribution. For the two-stage Coxian distribution to be focused on in Sections 3.3 and 4, we have $C = (1, 0)$ and from Equation (1),

$$\hat{G} = \begin{pmatrix} -\mu_1 & a_1\mu_1 & (1-a_1)\mu_1 \\ 0 & -\mu_2 & \mu_2 \\ 0 & 0 & 0 \end{pmatrix}.$$

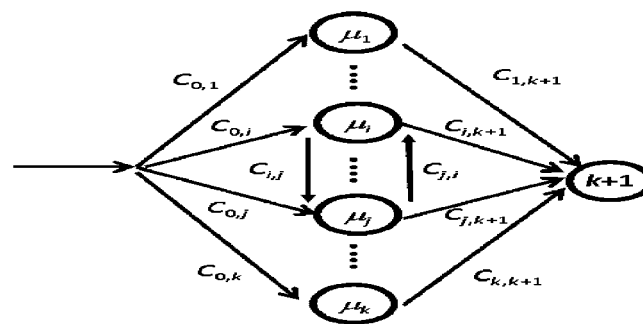


Figure 1. A k -stage Phase-type distribution.

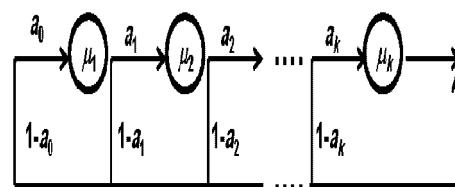


Figure 2. A k -stage Coxian distribution.

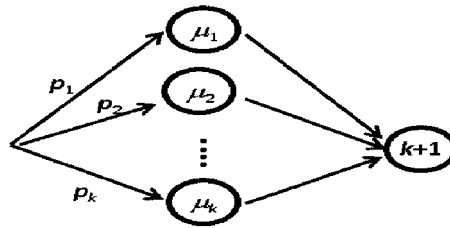


Figure 3. A k -stage Hyper-exponential distribution.

A k -stage Hyper-exponential random variable, on the other hand, is a proper mixture of Exponential random variables (see Figure 3). With an initial branching probability of $c_{0,i} = p_i$, it enters state i to stay for an exponential duration with rate μ_i , where $p_1 + p_2 + \dots + p_k = 1$. Once the sojourn time in that state is over, it enters the absorbing state. The Hyper-exponential r.v. is popularly used to model or approximate distributions with squared-coefficient of variation $c^2 \geq 1$. In many practical settings, the special case of two-stage Hyper-exponential is deemed sufficient for accurate modeling.

3. The model and the cost function

We consider a continuous review inventory system with constant demand and negligible delivery lead times, where the demand per unit time is denoted by D . The holding cost per unit per unit time is assumed to be h and an ordering cost of K is incurred per order. Furthermore, we assume that supply disruptions of random durations can occur due to the unavailability or inaccessibility of the supplier, which goes through ON and OFF periods of stochastic durations. Although, the delivery lead times are assumed negligible, due to the possibility of disruptions with random durations, a modified (q, r) ordering policy is employed as follows: (i) when the inventory position (IP) hits the reorder point r , and the supplier is ON, q units are ordered to bring the IP immediately to $R = q + r$ and (ii) if the supplier is OFF when the IP drops to reorder point r , an amount is ordered to bring the IP to the target level R as soon as the supplier becomes ON again. Obviously, the order quantity in the second scenario becomes a random variable. In this case, unsatisfied demands can occur if the inventory level drops below zero before the supplier becomes available again and such orders are fully backordered at a cost of b per unit.

The objective of the study is to minimize the average cost rate function in reference to the renewal reward theorem.

3.1. The CTMC and the average cost function

The ON/OFF status of the supplier is an alternating renewal process and from the point of view of the retailer, which cannot observe at which ON or OFF phase the system is, the supplier status appears as a semi-Markov process. However, using PTD, we are able to model the ON/OFF status of the supplier as a CTMC. Before introducing it, let us assume that the representation of the ON PTD is (\mathbf{C}, \mathbf{G}) as given in Equation (1). We will assume that the OFF period follows a d -stage PTD such that the sojourn time in any OFF state $0n$ is exponentially distributed with rate λ_n , $n = 1, 2, \dots, d$. The representation of the OFF PTD is denoted by (\mathbf{B}, \mathbf{H}) , where $\mathbf{B} = (c_{0,01}, c_{0,02}, \dots, c_{0,0d})$ is the initial probability vector and \mathbf{H} is a $d \times d$ matrix with entries as $H_{nn} = -\lambda_n$ for $n = 1, \dots, d$ and for $n \neq m$, $H_{nm} = c_{0n,0m} \lambda_n$ with $c_{0n,0m}$ denoting the transition probability from state $0n$ to state $0m$ satisfying $c_{0n,01} + \dots + c_{0n,0d} + c_{0n,0} = 1$, for $n, m = 1, \dots, d$. And finally, $\mathbf{B}\mathbf{E}' = \mathbf{1}$ and the $d \times 1$ \mathbf{H}_0 vector satisfies $\mathbf{H}\mathbf{E}' + \mathbf{H}_0 = \mathbf{0}$, where \mathbf{E}' is a $d \times 1$ column vector of 1. Accordingly, the absorbing state $0(k+1)$ for the OFF PTD (ON PTD) is an ON (OFF) phase and from state $0n$ (i), with probability $c_{0n,0}$ ($c_{i,k+1}$) the system enters an arbitrary ON (OFF) phase.

Now we can introduce $\{\zeta(t), t \geq 0\}$ representing the supplier's status, where $\zeta(t) = 0n$ corresponds to the n th stage of the OFF period and $\zeta(t) = j$ indicates that the supplier is at the j th stage of the ON period. Thus, $P_{xw}(t) = P\{\zeta(t) = w | \zeta(0) = x\}$, $x, w = 1, 2, \dots, k, 01, 02, \dots, 0d$ denotes the transition probability of CTMC being in state w at time t given that it was in state x at time 0.

Considering a typical realization (see Figure 4) of the system under study, one would see that every time the IP reaches R right after the completion of an OFF period, the system regenerates itself. In other words, the time between instances of IP reaching R at the end of OFF periods are considered to be a *regenerative cycle*. Then, the renewal reward theorem [18, p. 78] can be exploited to construct the long run average cost rate function, which would be the ratio of the expected cost of a regenerative cycle to its expected length.

We identify two types of subcycles within a cycle. The first type, referred to as the *standard* subcycle, starts each time the IP is raised to R during an ON period, and has a length of q/D . Standard subcycles repeat themselves until the IP drops down to r during an OFF period. This leads to the *last* subcycle, which will continue until the supplier becomes ON again. A regenerative cycle is composed of a random number, $\tilde{N}(q) - 1$, of standard subcycles and a last one of length $q/D + \tilde{T}_0$,

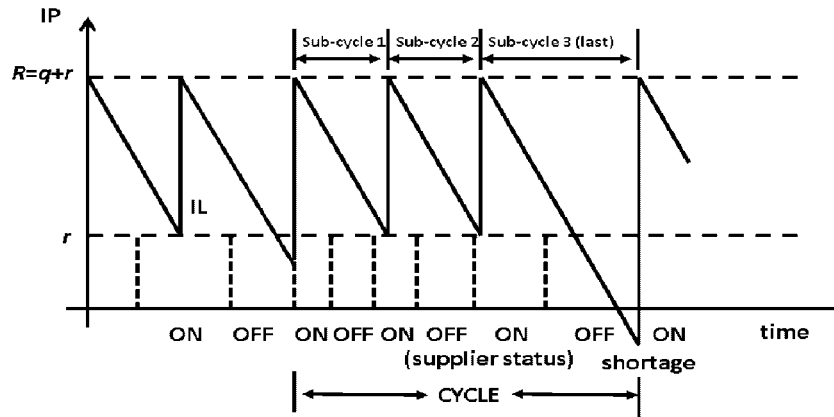


Figure 4. Regenerative cycles of the inventory position.

where \tilde{T}_0 refers to the *waiting time* until the supplier becomes ON again. Let the cumulative distribution function (CDF) and the probability density function (PDF) of \tilde{T}_0 be denoted by $G(y)$ ($G(y) = 1 - G(y)$) and $g(y)$, respectively.

Let $\tilde{N}_i(q)$ be the number of subcycles needed to complete the cycle with an expected value of $N_i(q)$, if $\zeta(t)$ is in the ON state i , $i = 1, 2, \dots, k$ when the inventory is raised to $R = q + r$ at the beginning of a cycle. Since the state of $\zeta(t)$ when a new cycle starts will be i with probability $c_{0,i}$ (irrespective of which OFF state was preceding, hence 0 denotes any OFF state), we can write the expected number of subcycles $N(q)$ within a cycle as

$$N(q) = \sum_{i=1}^k c_{0,i} N_i(q) = \mathbf{C}\mathbf{N},$$

where as before $\mathbf{C} = (c_{0,1}, c_{0,2}, \dots, c_{0,k})$ and $\mathbf{N}^T = (N_1(q), N_2(q), \dots, N_k(q))$. Depending on the state of $\zeta(t)$ when the inventory drops to r for the first time after the cycle starts, we can write

$$\tilde{N}_i(q) = I(\zeta(q/D) = 0) + \sum_{j \neq i, j=1}^k (1 + \tilde{N}_j(q)) I(\zeta(q/D) = j) + (1 + \tilde{N}_i^*(q)) I(\zeta(q/D) = i),$$

where $\tilde{N}_i^*(q)$ has the same distribution as $\tilde{N}_i(q)$, $i = 1, 2, \dots, k$ and $I(\cdot)$ is the indicator function, which is used to imply that the reorder point is down-crossed in one of the ON or OFF state. Taking the expectation of both sides and arranging the terms give the expected number of standard subcycles as

$$N(q) = \mathbf{C}(\mathbf{I} - \mathbf{P})^{-1} \mathbf{E},$$

where $\mathbf{P} = \{P_{ij}(q/D)\}_{i,j=1,2,\dots,k}$ is the probability transition matrix between ON states over the interval $(0, q/D)$ and \mathbf{I} is the $k \times k$ identity matrix, and as in Section 2, \mathbf{E} is the $k \times 1$ column vector of 1. Hence, the expected length of a cycle will be

$$T(q) = \frac{q}{D} \mathbf{C}(\mathbf{I} - \mathbf{P})^{-1} \mathbf{E} + T_0,$$

where $T_0 = E[\tilde{T}_0]$.

The expected total cost within a cycle consists of ordering, inventory holding, and backordering costs. Let $c(q, r)$ be the expected cost of a standard subcycle given as

$$c(q, r) = K + h \left[\frac{q^2}{2D} + \frac{rq}{D} \right].$$

The cost incurred during the waiting time in the last subcycle depends on the length of \tilde{T}_0 . The expected holding cost HC during the waiting time is

$$HC = h \left\{ \int_0^{r/D} \left(ru - \frac{u^2 D}{2} \right) g(u) du + \frac{r^2}{2D} \bar{G}(r/D) \right\}, \quad (2)$$

whereas the expected backordering cost is

$$BC = Db \int_{r/D}^{\infty} (u - r/D)g(u) du. \quad (3)$$

The summation of the two gives the expected cost, $C(q, r) = HC + BC$, that the system incurs during the waiting time.

Now, invoking the renewal reward theorem, the long run average cost rate function $\mathcal{K}(q, r)$ is obtained as

$$\mathcal{K}(q, r) = \frac{N(q)c(q, r) + C(q, r)}{T(q)}. \quad (4)$$

The cost rate function given above does not provide closed-form solutions and the optimization should be done by numerical methods. However, to implement the numerical methods, we first need to obtain the state transition probabilities, $P_{xw}(t)$, of the CTMC $\{\zeta(t), t \geq 0\}$, which we will discuss next.

3.2. State-transition probabilities

Let $F_i(t) = 1 - e^{-\mu_i t}$, $t \geq 0$ be the CDF of the sojourn time in ON state i with rate μ_i , ($\bar{F}_i(t) = 1 - F_i(t)$) and $f_i(t)$ be its density function, $i = 1, \dots, k$. Similarly, let $G_n(t) = 1 - e^{-\lambda_n t}$, $t \geq 0$ be the CDF of the sojourn time in OFF state $0n$ with rate λ_n , ($\bar{G}_n(t) = 1 - G_n(t)$) and $g_n(t)$ be its density function, $n = 1, \dots, d$.

One way to obtain $P_{xw}(t)$ $x, w = 1, \dots, k, 01, \dots, 0d$ is exploiting the integral equations as given below.

Theorem 1

The state-transition probabilities $P_{xw}(t)$, $t \geq 0$ of the CTMC representing the supplier availability whose ON and OFF times follow a k -stage and a d -stage PTD, respectively, are the solutions of the following integral equations:

$$P_{ii}(t) = \bar{F}_i(t) + \sum_{m=1, m \neq i}^k c_{i,m} \int_0^t P_{mi}(t-x) f_i(x) dx + c_{i,k+1} \sum_{n=1}^d c_{0,0n} \int_0^t P_{0n,i}(t-x) f_i(x) dx, \quad i = 1, \dots, k, \quad (5)$$

$$P_{ij}(t) = \sum_{m=1, m \neq i}^k c_{i,m} \int_0^t P_{mj}(t-x) f_i(x) dx + c_{i,k+1} \sum_{n=1}^d c_{0,0n} \int_0^t P_{0n,j}(t-x) f_i(x) dx, \quad i, j = 1, \dots, k, \quad i \neq j, \quad (6)$$

$$P_{i,0n}(t) = \sum_{m=1, m \neq i}^k c_{i,m} \int_0^t P_{m,0n}(t-x) f_i(x) dx + c_{i,k+1} \sum_{l=1}^d c_{0,0l} \int_0^t P_{0l,0n}(t-x) f_i(x) dx, \quad i = 1, \dots, k, \quad n = 1, \dots, d, \quad (7)$$

$$P_{0n,0n}(t) = \bar{G}_n(t) + \sum_{l=1, l \neq n}^d c_{0n,0l} \int_0^t P_{0l,0n}(t-x) g_n(x) dx + c_{0n,0} \sum_{m=1}^k c_{0,m} \int_0^t P_{m,0n}(t-x) g_n(x) dx, \quad n = 1, \dots, d, \quad (8)$$

$$P_{0n,0v}(t) = \sum_{l=1, l \neq n}^d c_{0n,0l} \int_0^t P_{0l,0v}(t-x) g_n(x) dx + c_{0n,0} \sum_{m=1}^k c_{0,m} \int_0^t P_{m,0v}(t-x) g_n(x) dx, \quad n, v = 1, \dots, d, \quad n \neq v, \quad (9)$$

$$P_{0n,j}(t) = \sum_{l=1, l \neq n}^d c_{0n,0l} \int_0^t P_{0l,j}(t-x) g_n(x) dx + c_{0n,0} \sum_{m=1}^k c_{0,m} \int_0^t P_{m,j}(t-x) g_n(x) dx, \quad n = 1, \dots, d, \quad j = 1, \dots, k. \quad (10)$$

Proof

We will only derive Equation (6), since the others are similar. Note that the probability of making a transition from ON state i in the interval $(x, x + dx)$ is $f_i(x)dx$. With probability $c_{i,m}$ (see Figure 1) the transition from state i will be into ON state m . $P_{mj}(t-x)$ gives the probability of transition into ON state j during the remaining $t-x$ time units from ON state m . Similarly, with probability $c_{i,k+1} \times c_{0,0n}$, upon leaving ON state i , CTMC enters OFF state $0n$. $P_{0n,j}(t-x)$ gives the probability of transition into ON state j during the remaining $t-x$ time units from OFF state $0n$. The final equation is obtained by considering all possible values of m and n via the summation. Equation (7) is very similar except this time we find the probability of starting from an ON state i and ending at an OFF state. Equations (8)–(10) are the probabilities of starting from an OFF state. \square

The set of integral equations such as the ones in Theorem 1 has been used, e.g. by Gürler and Parlar [9], Parlar [6] to obtain the transition probabilities. This set of integral equations can be approximated by employing techniques such as

the trapezoidal rule before inverting, but this usually yields huge sparse matrices, which are difficult to invert (Gürler and Parlar [9]).

Another approach is to exploit the Kolmogorov equations as presented below.

Theorem 2

The transient probabilities of the CTMC representing the supplier's ON/OFF status are obtained as the solution of the following Kolmogorov equations:

$$\mathbf{P}'(t) = \hat{\mathbf{A}}\mathbf{P}(t), \quad \mathbf{P}(0) = \mathbf{I}', \tag{11}$$

where $\mathbf{P}(t) = [P_{xw}(t)]$, $t \geq 0$, $x, w = 1, \dots, k, 01, \dots, 0d$ is the $(k+d) \times (k+d)$ matrix of finite time transition probabilities, \mathbf{I}' is the $(k+d) \times (k+d)$ identity matrix, and $\hat{\mathbf{A}}$ is the $(k+d) \times (k+d)$ infinitesimal generator of the CTMC, given as

$$\hat{\mathbf{A}} = \begin{pmatrix} \mathbf{G} & \mathbf{G}_0\mathbf{B} \\ \mathbf{H}_0\mathbf{C} & \mathbf{H} \end{pmatrix}, \tag{12}$$

the rows and columns of which are indexed as $1, \dots, k, 01, \dots, 0d$.

Proof

By definition, the matrix \mathbf{G} gives the transition rates of flow balance equations among ON states. Similarly, the matrix \mathbf{H} has the same rates among OFF states. Since the vector \mathbf{G}_0 contains the rates of entering an arbitrary OFF state from each ON state, the product $\mathbf{G}_0\mathbf{B}$ yields the matrix of transition rates of entering a given OFF state from a given ON state. Similarly, the product $\mathbf{H}_0\mathbf{C}$ is the matrix of transition rates of entering a given ON state from a given OFF state. \square

The solution of Equation (11) is

$$\mathbf{P}(t) = e^{t\hat{\mathbf{A}}}\mathbf{P}(0), \tag{13}$$

and one can approximately compute $e^{t\hat{\mathbf{A}}}$ by following techniques in the literature [19].

From $\mathbf{P}(q/D)$ one can obtain $P_{xw}(q/D)$ for $x, w = 1, \dots, k$, namely, transition probabilities between ON states and form \mathbf{P} to compute $N(q)$ and $T(q)$, which are finally used in Equation (4). Note that the expected wait time T_0 that is needed for $T(q)$ and $\bar{G}(t)$ with $g(t)$ that are needed in Equation (2) ($g(t)$ for Equation (3) as well) have to be obtained considering the OFF PTD. To do this, we define $\bar{P}_{i,0n}(m, q/D)$ as the steady-state probability that a cycle starting with ON state i ends after m subcycles when the IP down-crosses r in state $0n$. Note that $\bar{P}_{i,0n}(m, q/D)$ goes rapidly to zero when m gets large. Hence, the steady-state probability of down-crossing r in state $0n$ of the OFF period in an arbitrary cycle, \bar{P}_{0n} , is given as follows by adding over m , for $n = 1, \dots, d$,

$$\bar{P}_{0n} = \sum_{i=1}^k c_{0,i} \sum_{m=1}^{\infty} \bar{P}_{i,0n}(m, q/D), \tag{14}$$

where for $m > 1$ we find the required probabilities by conditioning on the ON state where the CTMC is found at the end of the first subcycle:

$$\bar{P}_{i,0n}(m, q/D) = \sum_{j=1}^k P_{ij}(q/D) \bar{P}_{j,0n}(m-1, q/D) \tag{15}$$

and $\bar{P}_{i,0n}(1, q/D) = P_{i,0n}(q/D)$ for $i = 1, \dots, k$ and $n = 01, \dots, d$, which are readily available from either the solution of Theorem 1 or Equation (13).

Letting $R_n(t)$ denote the CDF of the waiting time of the last subcycle if r is down-crossed in state $0n$ of the OFF period, we have

$$G(t) = \sum_{n=1}^d \bar{P}_{0n} R_n(t)$$

from which its derivative $g(t)$ and its expected value, namely, T_0 can be obtained. Obviously, $R_n(t)$ depends on the underlying OFF PTD. For instance, if a k -stage Erlang is the underlying OFF period distribution, the waiting time will be a $k-n-1$ -stage Erlang r.v. if r is down-crossed in state $0n$ of the OFF period.

Although we summarized two alternative approaches by Theorems 1 and 2, obtaining $P_{xw}(t)$ becomes difficult especially when the number of PTD phases increases. Thus, similar to earlier work in the literature we will restrict ourselves to simpler yet highly versatile PTDs in Section 3.3. Due to its flexibility in modeling distributions with $c^2 \geq 0.5$ while providing a small

state space with only two different stages, we have chosen two-stage Coxian distribution for modeling the supplier ON time. In similar vein, a two-stage Hyper-exponential OFF period is considered, with 01 and 02 as the possible OFF period states. The model studied in Section 3.3 will be employed in our numerical examples presented in Section 4 while testing the impact of different ON/OFF time distributions.

3.3. The case with two-stage Coxian on times and two-stage Hyper-exponential off times

In this special case, the supplier ON period is modeled by Coxian (a_1, μ_1, μ_2) and the supplier OFF period follows $H2(p, \lambda_1, \lambda_2)$ distribution with $p(1-p)$ being the probability of entering OFF state 01 (02). Observe that when $a_1 = 1$ and $\mu_1 = \mu_2$, the ON period distribution becomes a two-stage Erlang distribution and when $p = 1$, the OFF period distribution is Exponential with rate λ_1 . We will consider both these distributions in numerical examples in Section 4 as well.

The state-transition probabilities $P_{xw}(t), t \geq 0$ of the CTMC modeling the supplier status are the solutions of the integral equations from Theorem 1. In particular, we need the transition probabilities among ON states for constructing the matrix $\mathbf{P} = \{P_{ij}(q/D)\} i, j = 1, 2$, which are

$$\begin{aligned} P_{11}(t) &= e^{-\mu_1 t} + a_1 \int_0^t P_{21}(t-x)\mu_1 e^{-\mu_1 x} dx + (1-a_1)p \int_0^t P_{01,1}(t-x)\mu_1 e^{-\mu_1 x} dx \\ &\quad + (1-a_1)(1-p) \int_0^t P_{02,1}(t-x)\mu_1 e^{-\mu_1 x} dx, \\ P_{12}(t) &= a_1 \int_0^t P_{22}(t-x)\mu_1 e^{-\mu_1 x} dx + (1-a_1)p \int_0^t P_{01,2}(t-x)\mu_1 e^{-\mu_1 x} dx \\ &\quad + (1-a_1)(1-p) \int_0^t P_{02,2}(t-x)\mu_1 e^{-\mu_1 x} dx, \\ P_{21}(t) &= p \int_0^t P_{01,1}(t-x)\mu_2 e^{-\mu_2 x} dx + (1-p) \int_0^t P_{02,1}(t-x)\mu_2 e^{-\mu_2 x} dx, \\ P_{22}(t) &= e^{-\mu_2 t} + p \int_0^t P_{01,2}(t-x)\mu_2 e^{-\mu_2 x} dx + (1-p) \int_0^t P_{02,2}(t-x)\mu_2 e^{-\mu_2 x} dx. \end{aligned} \tag{16}$$

Direct method of solving the integral equations such as the ones above is based on approximating an integral using one of many classical formulae, such as the trapezoidal rule, the Simpson's rule, or the Bode's rule. For instance, Gürler and Parlal [9] employ the trapezoidal rule towards this end. However, this approach results in large sparse matrices to be inverted for different (q, r) values during the search process described in Section 4.

Alternatively, one can exploit the Kolmogorov equations of this system following Theorem 2, for which

$$\hat{\mathbf{A}} = \begin{pmatrix} -\mu_1 & a_1\mu_1 & p(1-a_1)\mu_1 & (1-p)(1-a_1)\mu_1 \\ 0 & -\mu_2 & p\mu_2 & (1-p)\mu_2 \\ \lambda_1 & 0 & -\lambda_1 & 0 \\ \lambda_2 & 0 & 0 & -\lambda_2 \end{pmatrix}$$

to be used in Equation (13) to obtain the required transition probabilities.

In this paper, we have chosen to exploit the integral equations from Theorem 1 to obtain $P_{xw}(t), t \geq 0$ but employed a different approach as explained in Appendix A from approximations such as the Trapezoidal rule. Our approach requires having the Laplace transforms of the integral equations, which are numerically inverted. This approach turns out to give quick and accurate computations. With these probabilities found in Appendix A, the matrix $\mathbf{P} = \{P_{ij}(q/D)\} i, j = 1, 2$ and then $N(q)$ in Section 3.1 can be obtained. In order to obtain the expected cycle length $T(q)$, one must keep track of the OFF state that has initiated the waiting time in the last subcycle. Noting that each cycle starts with ON state 1, for $m > 1$ we need $\bar{P}_{1,0n}(m, q/D), n = 1, 2$ from Equation (15) to be substituted in Equation (14) to compute \bar{P}_{01} and $\bar{P}_{02} (= 1 - \bar{P}_{01})$. For $m = 1$, as noted before, $\bar{P}_{1,0n}(1, q/D) = P_{1,0n}(q/D)$. In Appendix A, we also explain how $P_{1,0n}(q/D)$ can be computed.

With \bar{P}_{01} and \bar{P}_{02} , which are the steady-state probabilities of starting the waiting time of an arbitrary cycle in OFF state 01 or 02, respectively, we can write the expected length of a cycle as

$$T(q) = \frac{q}{D} \mathbf{C}(\mathbf{I} - \mathbf{P})^{-1} \mathbf{E} + T_0 = \frac{q}{D} \mathbf{C}(\mathbf{I} - \mathbf{P})^{-1} \mathbf{E} + \frac{\bar{P}_{01}}{\lambda_1} + \frac{\bar{P}_{02}}{\lambda_2}$$

due to the fact that in either of the OFF states, the remaining time will be exponential with the appropriate sojourn rate.

Similarly, we can also obtain the respective expected holding and backordering costs during the waiting time using Equations (2) and (3),

$$HC = h \left\{ \bar{P}_{01} \frac{\lambda_1 r - D + D e^{-\lambda_1 r / D}}{\lambda_1^2} + \bar{P}_{02} \frac{\lambda_2 r - D + D e^{-\lambda_2 r / D}}{\lambda_2^2} \right\},$$

$$BC = Db \left\{ \bar{P}_{01} \frac{e^{-\lambda_1 r / D}}{\lambda_1} + \bar{P}_{02} \frac{e^{-\lambda_2 r / D}}{\lambda_2} \right\}.$$

4. Numerical results

In this section, we present the results of our numerical study, which aims to elaborate (i) the significance of exact modeling of the OFF time distribution and (ii) the impact of different ON and OFF distributions on the cost figures.

Defining the unavailability rate as $E(OFF)/[E(ON) + E(OFF)]$, we considered three cases: (i) a reliable supplier with 11% overall unavailability, (ii) a highly reliable supplier with 5% overall unavailability, and (iii) an unreliable supplier with 40% overall unavailability. In each case, we conducted tests using four pairs of ON/OFF time distributions: (a) two-stage Erlang (Erlang(2, μ)) ON times with a $c_{ON}^2 = 0.5$ and Exponential(λ) OFF times with a $c_{OFF}^2 = 1$, (b) two-stage Coxian (Coxian (a_1, μ_1, μ_2)) ON times with a $c_{ON}^2 = 10$ and Exponential(λ) OFF times, (c) Erlang (2, μ) ON times and two-stage Hyper-exponential (H2(p, λ_1, λ_2)) OFF times with a $c_{OFF}^2 = 32.87$, and finally (d) Coxian(a_1, μ_1, μ_2) ON times and H2(p, λ_1, λ_2) OFF times. In the numerical analysis, the demand is taken as $D = 100$ per unit time, holding cost is set to $h = 1$ and several ordering and back order costs are considered.

As stated earlier, the expected cost rate function does not yield a closed-form expression, hence numerical optimization methods are used to find the optimal (q^*, r^*). We applied adaptive random search (ARS) [20] method, which is proven to be effective and simple. The algorithm to compute the optimal (q^*, r^*) values is summarized next, where \mathcal{C} is used for $\mathcal{H}(q, r)$:

- Step 0. Set \mathcal{C} to a high positive initial value.
- Step 1. Start with a feasible (q, r) in a set that is known to contain the optimal pair.
- Step 2. Evaluate $P_{ij}(q/D)$
- Step 3. Generate the \mathbf{P} matrix, compute the expected cycle length and the expected cycle cost using Equation (4).
- Step 4. Evaluate \mathcal{C} . If it is smaller than the previous one, keep the corresponding (q, r). If the improvement is negligible, stop. Otherwise go to the next step.
- Step 5. Generate a new feasible q value as follows: If q^* is known to be in the interval $[q_l, q_h]$ and if q_{old} refers to the previous point, generate the new point $q_{new} = q_{old} + (q_h - q_l)(2\theta - 1)^v$, where θ is a random number between 0 and 1 and v is an odd integer. Similarly find a new feasible r value. Go to Step 2.

We next summarize our results. In Tables I–VIII, the first two columns show the K and b parameters assumed for the numerical examples. In all tables, except Table II, we present the economic order quantity $q_{EOQ} = \sqrt{2KD/h}$ in the third

K	b	$q_{EOQ} (= C_{EOQ})$	C_{EOQ}^A (Actualized)	q^*	r^*	$\mathcal{H}(q^*, r^*)$	$\% \frac{(C_{EOQ}^A - \mathcal{H}(q^*, r^*))}{\mathcal{H}(q^*, r^*)}$
50	25	100	165.04	124.07	13.68	145.41	13.5%
	50		232.79	103.28	45.49	163.67	42.2%
	150		503.78	141.48	75.85	193.4	160.5%
	300		910.27	141.56	76.47	207.33	339%
	500		1452.3	135.74	95.53	219.73	560.9%
100	25	141.42	187.14	187.06	1.21	180.99	3.4%
	50		235.6	148.53	43.93	198.45	18.7%
	150		429.44	143.3	64.59	224.54	91.3%
	300		720.2	143.3	64.59	246.21	192.5%
	500		1107.9	244.57	67.27	266.47	315.8%
400	25	282.84	304.41	282.84	0	304.41	0%
	50		328.74	285.59	12.44	312.77	2.2%
	150		426.03	282.87	50.18	349.67	21.8%
	300		571.96	282.87	50.18	369.27	54.9%
	500		766.54	264.75	107.77	395.01	94.1%

Table II. Coxian(0.05,1,0.05) ON, Exponential(4) OFF times with 11% overall supplier unavailability.

K	b	q^*	r^*	$\mathcal{H}(q^*, r^*)$
50	25	124.07	13.68	144.71
	50	103.28	45.49	163.17
	150	141.48	75.85	193.17
	300	141.56	76.47	206.82
	500	135.74	95.53	219.3
100	25	187.06	1.21	180.55
	50	148.53	43.93	198.21
	150	143.3	64.59	224.17
	300	143.3	64.59	245.41
	500	244.57	67.27	266.16
400	25	282.84	0	304.27
	50	285.59	12.44	321.59
	150	282.87	50.18	349.55
	300	282.87	50.18	369.03
	500	264.75	107.77	394.98

Table III. Erlang(2,1) ON, H2(0.95, 47.5, 0.2174) OFF times with 11% overall supplier unavailability.

K	b	$q_{EOQ} (= C_{EOQ})$	$C_{Er/Exp}^A$ (Actualized)	q^*	r^*	$\mathcal{H}(q^*, r^*)$	$\% \frac{(C_{Er/Exp}^A - \mathcal{H}(q^*, r^*))}{\mathcal{H}(q^*, r^*)}$
50	25	100	320.17	143.47	1.18	315.06	1.6%
	50		540.07	215.91	15.29	526.61	2.6%
	150		1272.86	385.95	334.3	1009.29	26.1%
	300		2376.58	895.53	230.63	1464.52	62.5%
	500		3735.19	963.46	300.86	1856.93	101.1%
100	25	141.42	343.76	187.14	0.33	343.43	0.1%
	50		562.34	269.35	13.55	544.9	3.2%
	150		1320.27	577.09	256.06	1030.43	28.1%
	300		2452.06	800.94	282.53	1443.81	69.8%
	500		3619.1	895.3	241.37	1988.53	82%
400	25	282.84	451.51	282.84	0	451.51	0%
	50		641.82	390.91	1.93	620.68	3.4%
	150		1330.12	813.77	176.93	1088.89	22.2%
	300		2352.34	918.75	208.61	1515.71	55.2%
	500		3721.17	915.48	141.48	2227.08	67.1%

Table IV. Coxian(0.05,1,0.05) ON, H2(0.95, 47.5, 0.2174) OFF times with 11% overall supplier unavailability.

K	b	$q_{EOQ} (= C_{EOQ})$	$C_{Er/H2}^A$ (Actualized)	q^*	r^*	$\mathcal{H}(q^*, r^*)$	$\% \frac{(C_{Er/H2}^A - \mathcal{H}(q^*, r^*))}{\mathcal{H}(q^*, r^*)}$
50	25	100	318.26	143.47	1.18	318.26	0%
	50		529.95	237.89	7.91	528.99	0.2%
	150		1009.77	385.98	384.48	1006.67	0.3%
	300		1457.73	895.56	280.81	1423.77	2.4%
	500		1846.99	640.44	461.43	1754.9	5.2%
100	25	141.42	345.56	187.14	0.33	345.56	0%
	50		546.93	313.59	6.22	545.82	0.2%
	150		1028.57	577.12	306.24	1025.96	0.31%
	300		1438.07	800.97	332.7	1407.21	2.2%
	500		1977.3	958.6	331.2	1813.46	9.0%
400	25	282.84	452.34	282.84	0	452.34	0%
	50		620.94	458.93	11.04	620.46	0.1%
	150		1085.33	809.11	203.45	1083.75	0.1%
	300		1508.53	643.56	335.68	1471.66	2.5%
	500		2212.98	915.51	191.66	2096.56	5.6%

Table V. Erlang(2,1) ON, Exponential(9.5) OFF times with 5% overall supplier unavailability.

K	b	$q_{\text{EOQ}} (= C_{\text{EOQ}})$	C_{EOQ}^A (Actualized)	q^*	r^*	$\mathcal{H}(q^*, r^*)$	$\% \frac{(C_{\text{EOQ}}^A - \mathcal{H}(q^*, r^*))}{\mathcal{H}(q^*, r^*)}$
50	25	100	112.51	100	0	112.51	0%
	50		125.72	124.07	13.68	121.32	3.6%
	150		178.2	91.35	21.14	132.49	34.5%
	300		256.9	91.31	37.26	142.04	80.9%
	500		361.89	91.31	37.26	145.38	148.9%
100	25	141.42	150.18	141.42	0	150.18	0%
	50		159.45	150.25	4.49	157.06	1.5%
	150		196.57	224.47	12.33	179.6	9.4%
	300		252.24	217.67	24.68	185.99	35.6%
	500		326.47	148.53	43.93	187.7	73.9%
400	25	282.84	286.97	282.84	0	286.97	0%
	50		291.6	282.84	0	291.6	0%
	150		310.18	285.59	12.44	303.23	2.3%
	300		338.04	285.59	12.44	311.69	8.5%
	500		375.18	285.59	12.44	322.97	16.2%

Table VI. Erlang(2,1) ON, H2(0.9954, 13, 0.1603) OFF times with 5% overall supplier unavailability.

K	b	$q_{\text{EOQ}} (= C_{\text{EOQ}})$	$C_{\text{Er/Exp}}^A$ (Actualized)	q^*	r^*	$\mathcal{H}(q^*, r^*)$	$\% \frac{(C_{\text{Er/Exp}}^A - \mathcal{H}(q^*, r^*))}{\mathcal{H}(q^*, r^*)}$
50	25	100	135.32	100	0	135.32	0%
	50		173.86	143.49	3.42	172.66	0.7%
	150		296.29	115.16	22.93	294.96	0.5%
	300		474.73	110.44	24.2	468.18	1.4%
	500		700.74	200.15	33.74	691.54	1.3%
100	25	141.42	173.58	141.42	0	173.58	0%
	50		206.98	150.25	4.49	206.98	0%
	150		334.14	185.37	8.13	328.97	1.6%
	300		500.41	217.67	24.68	500.41	0%
	500		724.36	219.95	50.39	717.08	1%
400	25	282.84	308.78	282.84	0	308.78	0%
	50		338.02	282.84	0	338.02	0%
	150		451.94	329.83	5.11	449.11	0.6%
	300		612.03	391.23	7.91	605.62	1.1%
	500		825.48	456.41	12.85	798.89	3.3%

Table VII. Erlang(2,1) ON, Exponential(0.75) OFF times with 40% overall supplier unavailability.

K	b	$q_{\text{EOQ}} (= C_{\text{EOQ}})$	C_{EOQ}^A (Actualized)	q^*	r^*	$\mathcal{H}(q^*, r^*)$	$\% \frac{(C_{\text{EOQ}}^A - \mathcal{H}(q^*, r^*))}{\mathcal{H}(q^*, r^*)}$
50	25	100	773.78	134.49	223.04	391.69	97.5%
	50		1475.6	175.55	282.07	479.28	207.9%
	150		4283.1	184.49	434.17	623.91	586.5%
	300		8494.2	187.07	503.89	716	1086.3%
	500		14109	169.84	592.95	783.09	1701.7%
100	25	141.42	718.6	257.07	169.36	411.67	74.6%
	50		1330.4	210.97	276.49	499.81	166.2%
	150		3777.5	206.55	429.88	644.62	486%
	300		7448.3	189.16	529.57	737.25	910.3%
	500		12343	228.58	556.38	803.82	1435.5%
400	25	282.84	634.42	368.53	109.88	495.22	28.1%
	50		1030.8	364.76	256.85	590.19	74.7%
	150		2616.6	370.96	353.96	727.74	259.6%
	300		4995.1	391.4	446.47	819.91	509.2%
	500		8166.6	409.14	505.84	888.34	819.3%

Table VIII. Erlang(2,1) ON, H2(0.015, 0.0225, 1.4775) OFF times with 40% overall supplier unavailability.

K	b	$q_{EOQ} (= C_{EOQ})$	$C_{Er/Exp}^A$ (Actualized)	q^*	r^*	$\mathcal{K}(q^*, r^*)$	$\% \frac{(C_{Er/Exp}^A - \mathcal{K}(q^*, r^*))}{\mathcal{K}(q^*, r^*)}$
50	25	100	721.78	201.94	81.06	678.07	6.4%
	50		1225.98	148.44	168.44	1180.1	3.9%
	150		3076.58	201.27	262.59	3057.81	0.6%
	300		5688.19	959.72	481.69	5636.04	0.9%
	500		9061.7	942.49	481.69	8834.09	2.6%
100	25	141.42	734.4	206.89	90.96	697.04	5.4%
	50		1248.12	183.42	133.97	1199.78	4.0%
	150		3094.65	223.33	258.29	3074.46	0.7%
	300		5698.22	961.81	418.31	5632.78	1.2%
	500		9083.42	879.31	553.06	8796.9	3.3%
400	25	282.84	795.79	371.64	32.27	779.09	2.1%
	50		1335.96	409.52	140.4	1286.59	3.8%
	150		3153.89	567.52	184.96	3138.43	0.5%
	300		5751.73	931.43	463.5	5649.79	1.8%
	500		9113.28	949.17	522.87	8821	3.3%

column. Note that because of the cost parameters we have chosen, the cost rate C_{EOQ} satisfies $C_{EOQ} = q_{EOQ}$, hence the same column also gives the EOQ cost that would be incurred if the supplier were always available.

4.1. Sensitivity and the benefit of modeling the OFF time distribution

We first discuss a more realistic case of 11% overall unavailability, for which the results are presented in Tables I–IV.

Table I shows the results when ON time is Erlang(2,1), with $E[ON]=2$ and OFF time is Exponential(4) with $E[OFF]=0.25$, where columns 5–7 display the optimal inventory policy parameters q^* , r^* , and the resulting cost $\mathcal{K}(q^*, r^*)$. As expected, we observe that as the backorder costs increase, both the (q^*, r^*) values and the corresponding costs increase, however, the rate of these increases become less for higher ordering costs. We also observe that as the ordering costs increase, the order quantity approaches EOQ. If supplier unavailability were ignored, one would have to resort to the suboptimal EOQ parameters $(q, r) = (q_{EOQ}, 0)$ (see column 3 for q_{EOQ}), incurring a cost of C_{EOQ}^A given in column 4. The percentage increase in the cost by using suboptimal parameters are given in the last column. These error terms indicate that severity of using EOQ policy instead of the optimal policy is felt more for higher backordering costs, which indicate the criticality of the goods under study. On the other hand, higher ordering costs slightly mitigate the cost discrepancy, which implies that if less frequent but larger orders are placed, the impact of missing the optimal policy is lessened to an extent.

In Table II, ON times are taken as Coxian(0.05,1,0.05) with the same $E[ON]=2$, whereas the OFF times are the same as given in Table I. We see that the optimal policy parameters presented in columns 3 and 4 are the same as those of the parallel cases in Table I. The optimal costs in the last column of Table II are slightly less than those with Erlang(2,1) ON times. At first glance, this looks surprising since Coxian(0.05,1,0.05) ON times with a $c_{ON}^2 = 10$ is obviously more variable than Erlang(2,1) ON times. On the other hand, we should consider the failure (or hazard) rate function of a variable defined as $r(t) = f(t)/\bar{F}(t)$, where $f(t)$ and $\bar{F}(t)$ are the probability density and complementary distribution functions, respectively. Coxian(0.05,1,0.05) has a decreasing failure rate (DFR) function, which makes the supplier with such ON time distributions more reliable than the one with Erlang(2,1) ON times, which has an increasing failure rate (IFR) function.

In Table III, the OFF time is taken as H2(0.95, 47.5, 0.2174) which preserves $E[OFF]=0.25$ of Table I, but results in a $c_{OFF}^2 = 32.87$, and the ON time is the same Exponential distribution of Table I. The optimal policy parameters and the resulting cost are given in columns 5–7. Except the case with $K = 400, b = 25$, policies from Table I are no more optimal (in fact, for this case EOQ is still the optimal policy). We observe in general that the costs have increased significantly when the OFF time has Hyper-exponential (H2) distribution and optimal parameters no longer get close to the EOQ parameters as K increases. This can be explained by the fact that H2(0.95, 47.5, 0.2174) has a DFR function, as opposed to the constant failure rate of the Exponential distribution. Note that the specific H2 OFF time structure provides a good model to account for extremely rare disaster cases. Here, only with 5% probability, the system will experience a long recovery time with a mean of $1/0.2174 = 4.6$ more than twice as much as an average ON period.

If it is assumed that OFF times are Exponential when in fact they are H2(0.95, 47.5, 0.2174), the optimal policy parameters of Table I would be obtained which would result in costs as given in Column 4 of Table III. We observe that not using the optimal policy costs us more as backordering cost increases, whereas higher ordering costs slightly decrease the penalty.

This can be traced from the last column, which lists the percent error of $C_{Er/Exp}^A$ with respect to $\mathcal{K}(q^*, r^*)$. We observe that the loss due to wrongly assuming an Exponential OFF distribution may result in as much as 101.1%.

In Table IV, ON times are taken as Coxian(0.05,1,0.05), which is the only difference from Table III. We observe that using Coxian ON times as opposed to Erlang does not have a monotone impact on costs and in most of the cases the optimal cost difference in the two models are not significant. However, using the suboptimal parameters of the Erlang ON times still may cause significant increases in the cost, especially with high backorder costs. Note for example in the last column that for the cases where $b=500$, using suboptimal parameters increases the costs by 5.2, 9.0 and 5.6% for $K=50, 100, 400$, respectively, which may be highly significant in nominal terms. As a summary of the reliable supplier case with 11% overall unavailability, we see that OFF time characteristics are more crucial on the optimal policy than those of ON time, although the latter is also important for critical goods.

We next examine the highly reliable supplier case with a 5% overall unavailability:

Table V presents the results with Erlang(2,1) ON times and Exponential(9.5) OFF times with $E[OFF]=0.105$. Increase in availability obviously has reduced the optimal costs significantly, and in comparison to Table I the cost increase as backorder increases is less emphasized in this case. We observe that EOQ policy $(q, r)=(q_{EOQ}, 0)$ is optimal when backordering cost is low and ordering cost is high. When backordering cost increases, on the other hand, EOQ policy is not optimal anymore. Note that, using the suboptimal EOQ policy becomes unacceptably costly for higher backorder costs, even with a supplier which is available 95% of the time (see column 8).

Assuming Coxian(0.05,1,0.05) distribution for modeling ON times instead of Erlang(2,1) distribution has not changed the optimal policies. Similar to our comparison of Tables II and I, we have detected a slight decrease in optimal cost when Coxian(0.05,1,0.05) with DFR function is considered. Hence, the table for this scenario is not presented.

The impact of changing the OFF time distribution is given in Table VI, which differs from Table V in its H2(0.9954, 13, 0.1603) OFF times with a $c_{OFF}^2=32.37$. Comparing these two tables, we observe that H2 OFF distribution results in much higher costs and increase in backorder costs result in steep increase in the optimal costs, whereas increase in ordering costs has relatively less impact. The costs with suboptimal parameters of Exponential OFF times are given in Column 4 and the percentage difference with the optimal costs are provided again in the last column. When the supplier is highly available, these cost differences appear to be tolerable. However, the impact of wrong modeling of the OFF time distribution on cost estimation is further discussed below.

Our numerical results with Coxian(0.05,1,0.05) ON times and H2(0.9954, 13, 0.1603) OFF times were almost the same as the results in Table VI. Hence, we have chosen not to present that table here.

We close our discussion by presenting the analysis of the unreliable supplier with a 40% overall unavailability:

Table VII shows the results when ON times follow Erlang(2,1) distribution and OFF times are Exponential(0.75) with $E[OFF]=1.334$. As expected, when the supplier is highly unavailable, in general the costs increase significantly and become more notable as higher backorder costs are incurred. We observe that EOQ policy $(q, r)=(q_{EOQ}, 0)$ is not acceptable since its implementation costs, C_{EOQ}^A , provided in column 4 are drastically higher than the optimal costs.

Assuming Coxian(0.05,1,0.05) distribution for modeling ON times instead of Erlang(2,1) distribution has changed the optimal policies in a few cases. However, the optimal costs were almost the same as the costs presented in column 7 of Table VII, which leads us to the conclusion that ON time characteristics do not play a crucial role when supplier is 40% unavailable, hence results are not presented.

In Table VIII, results with H2(0.015, 0.0225, 1.4775) OFF times with a $c_{OFF}^2=32.84$ are presented, which differ from Table VII in this OFF time distribution. In this case also, the OFF time distribution makes a drastic difference in the optimal costs. If the suboptimal parameters of the Exponential OFF time model were used, the cost differences get as high as 6.4% and the minimum differences are observed for the cases with moderate backorder costs.

As in the previous cases, our numerical results with Coxian(0.05,1,0.05) ON times and H2(0.015, 0.0225, 1.4775) OFF times were almost the same as the results in Table VIII. Hence, they are not presented.

In light of the above discussions, we observed that modeling the OFF time distribution correctly results in significant savings and its importance is more notable when the supplier is unavailable about 11% of the time. This result is intuitive, since when the supplier is highly reliable, the impact of OFF times is naturally not very important, hence its modeling is also not a crucial issue. When unavailability is very high on the other hand, the costs increase so drastically that correct modeling of the OFF time period falls short of improving the costs.

4.2. Impact of correctly modeling the off period in estimating costs

Next we would like to discuss the practical significance of modeling the ON and OFF periods correctly, when estimation of the cost figures is needed in advance to allocate funds for operations and investments. In our numerical studies, we have observed that in some cases, the optimal parameters of different models turn out to be close, which might lead to the wrong conclusion that using the right model is not indispensable when it comes to computing inventory cost correctly. To be more specific, let us examine the case with $K=50, b=25$ in Tables V and VI. We see that the optimal policy parameters are the

same as those of the EOQ model. Based on this observation, if we conclude that we can use the EOQ model when the supplier is highly reliable and the backorder cost is low, we would be in fact making a gross error since we would report the EOQ cost (which is 100 in this case) although the optimal policy parameters are set correctly. In the case of Exponential(9.5) OFF times, this corresponds to a 12.51% error, and in the H2(0.9954, 13, 0.1603) OFF times scenario, we would underestimate the cost by 35.32%.

Similar caution should apply if it is claimed that it will be sufficiently accurate to model the ON and OFF times by mean preserving Exponential r.v. even when the actual distributions are non-Exponential. In particular, by ignoring the fact that the actual OFF times are two-stage Hyper-exponential and instead assuming a mean preserving Exponential r.v. may lead to significant cost estimation errors. In comparison of parallel cases in Tables V and VI, this corresponds to using the suboptimal policy parameters presented in columns 5 and 6 of Table V when in fact OFF times follow H2(0.9954,13,0.1603). This implies that we would have to report the costs displayed in column 7 of Table V irrespective of the original OFF time distribution. Our mistake in reporting (computing) the costs would be realized only when the suggested policy were implemented and the cost figures $C_{Er/Exp}^A$ in column 4 of Table VI were witnessed. In summary, reported costs of policies from Table V would be on the average 114.4% less than their actualization costs when OFF times follow H2(0.9954,13,0.1603). A similar comparison of Tables I and III dictates the same precaution. If one chooses to consider the Exponential OFF times in modeling, i.e. running the system with parameters presented in columns 5 and 6 of Table I when the actual OFF times follow H2(0.95, 47.5, 0.2174), the average error of reporting wrong costs will be 551.8%.

Hence, we would like to emphasize once more that, even when they might not have a serious impact on the optimal policy parameters, the ON and OFF time distributions should be correctly reflected into the model in order to compute the system costs correctly.

5. Conclusion

In this study, an inventory problem is considered where the supplier may become unavailable at random times for random durations, which is modeled as a stochastic process going through ON and OFF periods. Both the ON and OFF periods are assumed to follow PTD. Assuming constant demand and negligible lead time, the exact long run average cost function is derived, the optimization of which is done by numerical methods.

Unlike the earlier research works that concentrate on ON time modeling, we point out the difficulty of incorporating the OFF time distribution into the analysis, yet successfully complete the model. In obtaining the state-transition probabilities, we propose using a more efficient numerical approach.

Extensive numerical studies demonstrate that on the extremes of the range of the supplier availability (highly reliable or unreliable), OFF time characteristics are important to obtain the costs correctly. When the supplier is reliable, both ON and OFF time distributions should be included in the analysis in order to obtain the correct optimal policy. Otherwise, much higher system costs would be incurred.

Appendix A: The solution of the integral equations

In this section, starting with the integral equations of the system studied in Section 3.3, we will obtain the Laplace transforms of the transition probabilities and eventually invert them numerically. To solve Equation (16), we need the following probabilities that appear on the right-hand side of Equation 16

$$P_{01,1}(t) = \int_0^t P_{11}(t-x)\lambda_1 e^{-\lambda_1 x} dx,$$

$$P_{01,2}(t) = \int_0^t P_{12}(t-x)\lambda_1 e^{-\lambda_1 x} dx,$$

$$P_{02,1}(t) = \int_0^t P_{11}(t-x)\lambda_2 e^{-\lambda_2 x} dx,$$

$$P_{02,2}(t) = \int_0^t P_{12}(t-x)\lambda_2 e^{-\lambda_2 x} dx.$$

Defining the Laplace transform (LT) of a function $f(t)$ as $\tilde{f}(s) = \int_0^\infty e^{-st} f(t) dt$ and observing that they are convolutions, we take the LT of Equation (16) and the equations above

$$\begin{aligned} \tilde{P}_{11}(s) &= \frac{1}{\mu_1 + s} + \frac{\mu_1}{\mu_1 + s} \{a_1 \tilde{P}_{21}(s) + (1 - a_1)[p \tilde{P}_{01,1}(s) + (1 - p) \tilde{P}_{02,1}(s)]\}, \\ \tilde{P}_{12}(s) &= \frac{\mu_1}{\mu_1 + s} \{a_1 \tilde{P}_{22}(s) + (1 - a_1)[p \tilde{P}_{01,2}(s) + (1 - p) \tilde{P}_{02,2}(s)]\}, \\ \tilde{P}_{21}(s) &= \frac{\mu_2}{\mu_2 + s} \{p \tilde{P}_{01,1}(s) + (1 - p) \tilde{P}_{02,1}(s)\}, \\ \tilde{P}_{22}(s) &= \frac{1}{\mu_2 + s} + \frac{\mu_2}{\mu_2 + s} \{p \tilde{P}_{01,2}(s) + (1 - p) \tilde{P}_{02,2}(s)\}, \\ \tilde{P}_{01,1}(s) &= \frac{\lambda_1}{\lambda_1 + s} \tilde{P}_{11}(s), \\ \tilde{P}_{01,2}(s) &= \frac{\lambda_1}{\lambda_1 + s} \tilde{P}_{12}(s), \\ \tilde{P}_{02,1}(s) &= \frac{\lambda_2}{\lambda_2 + s} \tilde{P}_{11}(s), \\ \tilde{P}_{02,2}(s) &= \frac{\lambda_2}{\lambda_2 + s} \tilde{P}_{12}(s), \end{aligned}$$

the four equations of which further reduce to

$$\begin{aligned} \tilde{P}_{11}(s) &= \frac{1}{\mu_1 + s} \frac{1}{1 - \frac{\mu_1}{\mu_1 + s} \tilde{g}(s) \left[a_1 \frac{\mu_2}{\mu_2 + s} + (1 - a_1) \right]}, \\ \tilde{P}_{12}(s) &= \frac{a_1}{\mu_2 + s} \frac{\mu_1}{\mu_1 + s} \frac{1}{1 - \frac{\mu_1}{\mu_1 + s} \tilde{g}(s) \left[a_1 \frac{\mu_2}{\mu_2 + s} + (1 - a_1) \right]}, \\ \tilde{P}_{21}(s) &= \frac{\mu_2}{\mu_2 + s} \tilde{g}(s) \tilde{P}_{11}(s), \\ \tilde{P}_{22}(s) &= \frac{1}{\mu_2 + s} + \frac{\mu_2}{\mu_2 + s} \tilde{g}(s) \tilde{P}_{12}(s), \end{aligned}$$

where $\tilde{g}(s) = p\lambda_1/(\lambda_1 + s) + (1 - p)\lambda_2/(\lambda_2 + s)$ is the LT of the two-stage Hyper-exponential OFF time density function. We can now exploit some numerical LT inversion techniques such as the one by Jagerman [16] to obtain $P_{ij}(q/D)$, $i, j = 1, 2$, which are used in the probability transition matrix \mathbf{P} and compute the expected number of subcycles $N(q)$.

We further need the following probabilities to compute $\tilde{P}_{1,0n}(1, q/D) = P_{1,0n}(q/D)$ to be used in Equation (15) to eventually obtain \tilde{P}_{0n} in Equation (14).

$$\begin{aligned} P_{1,01}(t) &= a_1 \int_0^t P_{2,01}(t-x) \mu_1 e^{-\mu_1 x} dx + (1 - a_1) p \int_0^t P_{01,01}(t-x) \mu_1 e^{-\mu_1 x} dx \\ &\quad + (1 - a_1)(1 - p) \int_0^t P_{02,01}(t-x) \mu_1 e^{-\mu_1 x} dx, \\ P_{1,02}(t) &= a_1 \int_0^t P_{2,02}(t-x) \mu_1 e^{-\mu_1 x} dx + (1 - a_1) p \int_0^t P_{01,02}(t-x) \mu_1 e^{-\mu_1 x} dx \\ &\quad + (1 - a_1)(1 - p) \int_0^t P_{02,02}(t-x) \mu_1 e^{-\mu_1 x} dx, \end{aligned}$$

$$P_{2,01}(t) = p \int_0^t P_{01,01}(t-x)\mu_2 e^{-\mu_2 x} dx + (1-p) \int_0^t P_{02,01}(t-x)\mu_2 e^{-\mu_2 x} dx,$$

$$P_{2,02}(t) = p \int_0^t P_{01,02}(t-x)\mu_2 e^{-\mu_2 x} dx + (1-p) \int_0^t P_{02,02}(t-x)\mu_2 e^{-\mu_2 x} dx,$$

$$P_{01,01}(t) = e^{-\lambda_1 t} + \int_0^t P_{1,01}(t-x)\lambda_1 e^{-\lambda_1 x} dx,$$

$$P_{01,02}(t) = \int_0^t P_{1,02}(t-x)\lambda_1 e^{-\lambda_1 x} dx,$$

$$P_{02,01}(t) = \int_0^t P_{1,01}(t-x)\lambda_2 e^{-\lambda_2 x} dx,$$

$$P_{02,02}(t) = e^{-\lambda_2 t} + \int_0^t P_{1,02}(t-x)\lambda_2 e^{-\lambda_2 x} dx.$$

Similar to what we have done before, by taking the LT of these convolutions and carrying out simplifications, we arrive at

$$\tilde{P}_{1,01}(s) = \frac{\mu_1}{\mu_1 + s} \frac{p}{\lambda_1 + s} \left\{ \frac{a_1 \mu_2}{\mu_2 + s} + (1 - a_1) \right\} \frac{1}{1 - \frac{\mu_1}{\mu_1 + s} \tilde{g}(s) \left[a_1 \frac{\mu_2}{\mu_2 + s} + (1 - a_1) \right]},$$

$$\tilde{P}_{1,02}(s) = \frac{\mu_1}{\mu_1 + s} \frac{(1-p)}{\lambda_2 + s} \left\{ \frac{a_1 \mu_2}{\mu_2 + s} + (1 - a_1) \right\} \frac{1}{1 - \frac{\mu_1}{\mu_1 + s} \tilde{g}(s) \left[a_1 \frac{\mu_2}{\mu_2 + s} + (1 - a_1) \right]}.$$

Now we can again use the numerical LT inversion technique by Jagerman [16] to obtain $P_{1,0n}(q/D)$, $n = 1, 2$.

Acknowledgements

We thank Dr Mahmut Parlar for suggesting the use of PTD in supplier uncertainty models. We also thank Dr Emre Berk and Dr Hakan Özaktaş for their valuable comments in several stages of this work. We are also grateful to the Associate Editor and the two anonymous referees whose comments and suggestions helped us to improve the presentation of the paper significantly.

References

1. Silver AE. Operations research in inventory management: a review and critique. *Operations Research* 1981; **29**:628–645.
2. Nahmias S. *Production and Operations Analysis*. Homewood: Illinois, 1989.
3. Parlar M, Berkin D. Future supply uncertainty in EOQ models. *Naval Research Logistics* 1991; **38**:107–121.
4. Berk E, Arreola-Risa A. Note on future uncertainty in EOQ models. *Naval Research Logistics* 1994; **41**:129–132.
5. Parlar M, Perry D. Inventory models of future supply uncertainty with single and multiple suppliers. *Naval Research Logistics* 1996; **43**:191–210.
6. Parlar M. Continuous-review inventory problem with random supply interruptions. *European Journal of Operational Research* 1997; **99**:366–385.
7. Özekici S, Parlar M. Inventory models with unreliable suppliers in a random environment. *Annals of Operations Research* 1999; **91**:123–136.
8. Parlar M. Probabilistic analysis of renewal cycles: an application to a non-Markovian inventory problem with multiple objectives. *Operations Research* 2000; **48**:243–255.
9. Gürler Ü, Parlar M. An inventory problem with two randomly available suppliers. *Operations Research* 1995; **45**:904–918.
10. Xia Y, Yang M-H, Golany B, Gilbert S, Yu G. Real time disruption management in a two-stage production and inventory system. *IIE Transactions* 2004; **36**:111–125.
11. Mohebbi E. A replenishment model for the supply-uncertainty problem. *International Journal of Production Economics* 2004; **87**:25–37.
12. Mohebbi E, Hao D. When supplier's availability affects the replenishment lead time—an extension of the supply-interruption problem. *European Journal of Operational Research* 2006; **175**:992–1008.
13. Mohebbi E, Hao D. An inventory model with non-resuming randomly interruptible lead time. *International Journal of Production Economics* 2008; **114**:755–768.
14. Ross AM, Rong Y, Snyder LV. Supply disruptions with time-dependent parameters. *Computers and Operations Research* 2008; **35**:3504–3529.
15. Schmitt AJ, Snyder LV, Shen ZM. Inventory systems with stochastic demand and supply: properties and approximations. *European Journal of Operational Research* 2010; **206**:313–328.
16. Jagerman DL. An inversion technique for the Laplace transform. *Bell System Technical Journal* 1982; **61**(8):1995–2002.
17. Neuts MF. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Dover: New York, 1981.
18. Ross S. *Stochastic Processes*. Wiley: New York, 1983.
19. Moler C, Van Loan C. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review* 2003; **45**:3–49.
20. Gottfried BS, Weisman J. *Introduction to Optimization Theory*. Prentice-Hall: Englewood Cliffs, NJ, 1973.