Uludağ University Journal of The Faculty of Engineering, Vol. 28, No. 1, 2023 RESEARCH

DOI: 10.17482/uumfd.1171281

A MAP/PH/1 PRODUCTION INVENTORY MODEL

Serife OZKAR [*]

Received: 05.09.2022; revised: 06.02.2023; accepted: 11.03.2023

Abstract: In this study, a production inventory model with phase type service times where customers join the system occur according to a Markovian arrival process is discussed. When the inventory level is positive, if an arriving customer finds the server idle gets into service immediately. Served customer leaves the system and the on-hand inventory is decreased one unit of item at service completion epoch. Otherwise, the customer enters into a waiting space (queue) of infinite capacity and waits for get served. The production facility produces items according to an (s, S) policy. The production is switched *on* when the inventory level depletes to s and the production remains *on* until the inventory level reaches to the maximum level S. Once the inventory level becomes S, the production process is switched *off*. Applying the matrix-geometric method, we carry out the steady-state analysis of the production inventory model and perform a few illustrative numerical examples includes the effect of parameters on the system performance measures and an optimization study for the inventory policy.

Keywords: Production inventory, Phase-type distribution, Markovian arrival process, Matrix-geometric method, Local purchase, Optimization

MAP/PH/1 Üretim Envanter Modeli

Öz: Bu çalışmada, müşterilerin Markovian varış sürecine göre sisteme katıldığı faz-tipi hizmet sürelerine sahip bir üretim envanter modeli tartışılmaktadır. Envanter seviyesi pozitif olduğunda, gelen bir müşteri hizmet biriminin boş olduğunu tespit ederse hemen hizmete girer. Hizmet verilen müşteri sistemden ayrılır ve eldeki stok, hizmet tamamlanma anında bir birim azalır. Aksi takdirde müşteri sonsuz büyüklükte bir bekleme alanına (kuyruğa) girer ve hizmet almayı bekler. Üretim tesisi, ürünleri (s, S) politikasına göre üretir. Envanter seviyesi s'ye düştüğünde üretim *açılır* ve envanter seviyesi maksimum S seviyesine ulaşana kadar üretim *açık* kalır. Envanter seviyesi S olduğu anda, üretim süreci *kapatılır*. Matris-geometrik yöntemi uygulayarak, üretim envanter modelinin kararlı durum analizini gerçekleştiriyoruz ve parametrelerin sistem performans ölçüleri üzerindeki etkisini ve envanter politikası için bir optimizasyon çalışmasını içeren birkaç açıklayıcı sayısal örnek gerçekleştiriyoruz.

Anahtar Kelimeler: Üretim envanteri, Faz-tipi dağılım, Markovian varış süreci, Matris-geometrik yöntem, Yerel satın alma, Optimizasyon

1. INTRODUCTION

In classical inventory systems, demanded items are directly delivered from stock and the amount of time required to service is negligible. Demand occurred during stock out periods either result in lost sales or is satisfied only after the arrival of the replenishments. In contrast, in most real-life situations, a positive amount of time to serve the inventory is needed, for example, items in inventory require time for retrieval, preparation, packing, and loading. Inventory systems where servicing time for demanded items is a positive value are denominated

[*] Department of International Trade and Logistics, Balikesir University, 10145 Balikesir, Turkey.

Corresponding Author: Serife Ozkar (serife.ozkar@balikesir.edu.tr)

by queueing-inventory systems. These systems must take into account inventory problems in addition to the queueing concepts for the system to work in optimum conditions. Sigman and Simchi-Levi (1992) firstly studied on inventory with positive service time where involve $M/G/1$ queue. The study in Schwarz et al. (2006) was the *first* to produce product form solution in an $M/M/1$ queueing inventory with positive lead time. Saffari et al. (2013) considered an $M/M/1$ queue with associated inventory where the lead time for replenishment is arbitrarily distributed; they produced product form solution for the system state distribution. Queuing-inventory systems can be handled according to many features such as arrival process and service process, inventory policy implemented, shortage or lost sale assumption, if there is queue capacity or not, service interruption or vacation assumption, or perishability of items in inventory. These systems have also been mentioned as inventory models with positive service time. A detailed survey of the literature for queueing-inventory systems can be found in Krishnamoorthy et al. (2011a) and Karthikeyan and Sudhesh (2016).

Filling the inventory by an internal production rather than an external supplier is usually relevant since the manufacturer himself meets the demands. It could be a large enterprise that plans its production based on demand forecasts. Krishnamoorthy et al. (2011b) *firstly* extended queueing inventory problems to production inventory systems. (s, S) inventory policy was implemented; that is when the inventory level drops to s, the production process starts and production continues until the inventory level reaches S. They called the production process \Markovian Production Process" (MPP). The customer arrival process is governed by MAP and the service times are assumed to follow the phase-type distribution. Problem was formulated by quasi birth-death process (QBD) and stationary probabilities were derived and performance measures were presented. Baek and Moon (2014) studied $M/M/1$ queuing system where inventory is replenished by an external supplier and also by an internal production. Customers join the system according to the Poisson process and service time follows an exponential distribution. (r, Q) policy is implemented. A cost optimization was performed and as a conclusion, they claimed that by controlling the production rate, both holding and stock out costs can be controlled. Hence compared to the external supplier internal production makes the system conditions more stable. Baek and Moon (2015) provided an extension to the capacitated single server system and also a multi-server queueing system. Krishnamoorthy and Narayanan (2013) studied an (s, S) production inventory system where the processing of inventory requires a random amount of time. Demands occur according to the Poisson process and service time is exponential. They provided a stochastic decomposition of the steady-state probability vector. They combined the stationary probability vector for the classical $M/M/1$ queue and the stationary vector of a production inventory system. Also, they provided an analysis of the system separately when production is on and they propose a novel approach for the computation of the expected length of a production run. Krishnamoorthy et al (2015) discussed a production inventory system with positive service time where the server and the production process are subject to interruptions. In their study, customers arrive to the system according to Poisson process; the service times and the production times are follows Erlang distributions. The interruptions of the server and the production occur according to Poisson process and the recovery of these follow an exponential distribution. The production inventory system where the arrivals of the demands occur according to Markovian arrival process was studied by He and Zhang (2013). In the system, the production time of a product follows a phase-type distribution. A shipment consolidation policy was used for the shipment of finished products. Namely, as soon as the number of finished products reaches a fixed quantity defined, all the cumulated finished products are shipped together to the inventory.

The loss of customers in queueing-inventory systems can happen in the two ways. First, the customer does not enter the system (called lost) if there is no inventory at the time he arrives in the system. Second, if the inventory level drops to zero as soon as the service of the customer in the service is completed, the customers waiting in the queue will leave the system (lost). The two queueing-inventory models have been discussed in Chakravarthy (2020). In Model 1, if any arriving customer finds the inventory level to be zero, he is being lost. In Model 2, the loss of customers is performed in the two ways. First, if an arriving customer in the idle system finds the inventory level to be zero, the customer is being lost, and secondly, the customers present at time of a service completion with zero inventory are all be lost. In Chakravarthy and Rumyantsev (2020), the two models in Chakravarthy (2020) have been also analyzed by considering MAP demands in batches and phase-type service times.

Local purchase is introduced in the inventory systems mainly to maintain customer goodwill. (s, S) inventory system with negligible service time in the concept of local purchase has been studied first by Krishnamoorthy and Raju (1998). Some recent important studies on local purchasing (or emergency purchasing) in the literature can be listed as follows. Shajin et al. (2022) considered a multi-server production inventory system in which an emergency replenishment of one item with zero lead time takes place when the inventory level drops to zero. A production inventory system under a base stock policy for inventory control has been discussed by De la Cruz and Daduna (2022). In their study, when there are no items on the inventory, arriving demand is lost. Barron (2022) examined inventory systems with dual sourcing and emergency orders. He considered two types of supply in the systems: i) a regular supply that follows an (s, S) inventory policy with positive lead time under a lost-sales and ii) an additional emergency supply that brings the inventory up to level $0 \leq S_e \leq S$ when stockout becomes. The emergency supply has a zero lead time but incurs an extra cost has been assumed. Two models of double-source queueing-inventory systems have been studied where instant destruction of inventory is possible in Melikov et al. (2023). Replenishment of stocks from various sources occurs as following. If the inventory level drops to s, a regular order for the supply of inventory to a slow source is generated; if the inventory level falls below a certain threshold value r, the system instantly cancels the regular order and generates an emergency order to the fast source.

In the area of queueing models, Poisson processes and exponential distributions have very nice mathematical properties for the arrival processes and service time distributions. However, the assumptions of the two tools are highly restrictive in applications. Thus, to remove these limitations, Neuts first developed the theory of *phase-type distributions (PH-distributions)* and *related point processes*. The PH-distributions can approximate any probability distribution given by nonnegative random variables. For details on PH-distributions and their properties, we refer the reader to the book in Neuts (1981). The arrival process can be modeled by a Markovian arrival process (MAP) or a batch Markovian arrival process ($BMAP$) in which fairly general tools for modeling stochastic arrival processes. Both of arrival processes can capture the possible correlation and burstiness in the arrival process. We refer to Artalejo et al. (2010), Chakravarthy (2001, 2010) for detailed information. In this study, we use a MAP and PH-distributions for the arrival process and the service times, respectively, and the matrix-analytic method for analyzing stochastic model. One can find the details of the method in the books by Neuts (1981) and Latouche and Ramaswami (1999).

We consider a production inventory system with positive service time. In this study, filling the inventory is considered by an internal production rather than an external supplier. It is assumed that when a customer arrives in the system, there is the least one inventory. In other words, an arriving customer is not being lost at time of an arrival. We analyze a production inventory system with positive service time under local purchase. That is, if there is at least a customer in the queue and the inventory level is reached to zero at time of a service completion, one item is immediately purchased from a nearby retailer for no losing the waiting customer. The main interest is to observe how the system have different arrival processes and service distributions works under various scenarios such as different arrival rate, service rate and production rate and to find an optimum inventory policy. The inventory policy answers two questions, that is, when should I start production and how much should I produce so that I have the minimum total cost. We believe that the model described in this study and the analysis presented will be beneficial to various industrial companies as MAP arrivals and phase-type service times can be observed frequently in practice and also take into account customer satisfaction (not allow loss of customers).

The paper is structured as follows. The assumptions of the production inventory model are elaborated in Section 2. The steady-state analysis of the model including the stability condition and some performance measures of the system are discussed in Section 3. In Section 4, the total expected cost function is structured and presented sensitivity analysis with numerical examples. Finally, some concluding remarks are given in Section 5.

2. MODEL DESCRIPTION

We consider a production inventory system with positive service time where has one production facility. Customers arrive to the system according to a Markovian arrival process (MAP) with representation (D_0, D_1) of order m. The matrices D_0 and D_1 denote, respectively, the transition rates without arrival and the transition rates with arrival. The underlying Markov chain of the MAP is governed by the generator matrix $D = D_0 + D_1$. Hence, the arrival rate of customers is given by $\lambda = \delta D_1 e$ where δ is the stationary probability vector of the matrix D and the vector satisfies as in (1).

$$\delta D = 0, \qquad \delta e = 1. \tag{1}$$

The service time follows a phase-type distribution with representation (β, T) of order n where β is the initial probability vector, $\beta e = 1$, and T is the generator matrix holding the transition rates among the n transient states. It is clear that $Te + T^0 = 0$ in which T^0 is the column vector contains the absorption rates into state 0 from the transient states. The phase-type distribution has the service rate $\mu = [\beta(-T)^{-1}e]^{-1}$.

Items are put into the inventory of the system by production and each arriving customer demands a single item from the inventory. When the inventory level is positive, an arriving customer finding the server idle gets into service immediately. Otherwise, the customer enters into a waiting space (queue) of infinite capacity to be served under the first-come first-served (FCFS) discipline. Served customer leaves the system and the on-hand inventory is decreased by one at service completion epoch.

The queueing-inventory system studied has a single production facility that produce one type of item. The production time of a product is exponentially distributed with rate η. The inventory level in the system is governed by the (s, S)-policy. The production is switched ON when the inventory level depletes to s and the production remains ON until the inventory level reaches to the maximum level S. Once the level becomes S, the production process is switched OFF. The process continues in this fashion. We want to note that if there is at least a customer in the queue and the inventory level is reached to zero, one item is immediately purchased for no losing the waiting customer. Local purchase is introduced in the inventory systems mainly to maintain customer goodwill. For example, when a customer goes to a system for taking an item which is not currently available there, the inventory keeper purchases the same from a nearby retailer and supplies to the customer. The production inventory system described is illustrated in Figure 1.

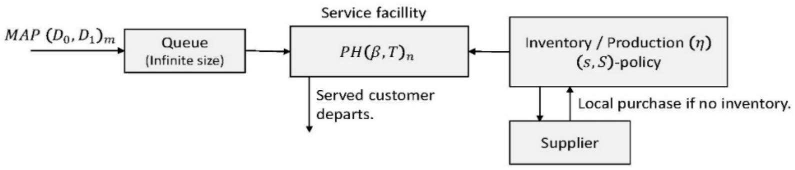

Figure 1:
Production inventory system with local purchase

We define $N(t)$ to be the number of customers in the system; $I(t)$ to be the inventory level in the system; $K(t)$ to be the status of the production process where 0 if the production is OFF and 1 if the production is ON; $J_1(t)$ to be the phase of the service process; and $J_2(t)$ to be the phase of the arrival process at time t. The process $\{(N(t), I(t), K(t), J_1(t), J_2(t)): t \geq 0\}$ is a continuous-time Markov chain $(CTMC)$ on the state space is given by

$$\Omega = \bigcup_{i=0}^{\infty} r(i), \quad (2)$$

where

$$\begin{aligned} r(0) = \; & \{(j, 1, j_2), 1 \leq j \leq s, 1 \leq j_2 \leq m\} \\ & \cup \{(j, k, j_2), s+1 \leq j \leq S-1, k = 0,1, 1 \leq j_2 \leq m\} \\ & \cup \{(S, 0, j_2), 1 \leq j_2 \leq m\}, \text{and} \end{aligned}$$

$$\begin{aligned} r(i) = \; & \{(i, j, 1, j_1, j_2), 1 \leq j \leq s, 1 \leq j_1 \leq n, 1 \leq j_2 \leq m\} \\ & \cup \{(i, j, k, j_1, j_2), s+1 \leq j \leq S-1, k = 0,1, 1 \leq j_1 \leq n, 1 \leq j_2 \leq m\} \\ & \cup \{(i, S, 0, j_1, j_2), 1 \leq j_1 \leq n, 1 \leq j_2 \leq m\}, \;\; i \geq 1. \end{aligned}$$

$r(0)$ of dimension $m(2S - s - 1)$ denotes the set of states corresponding to the system in which the system is idle and the arrival process is in phase j_2 (in one of m phases). The level $(j, 1, j_2)$ denotes the case when the inventory level is j, $1 \leq j \leq s$, the production process is ON; the level (j, k, j_2) denotes the case when the inventory level is j, $s+1 \leq j \leq S-1$, the production process is OFF and ON for $k = 0$ and $k = 1$, respectively; and the level $(S, 0, j_2)$ is related with the case the inventory is the maximum level S and the production process is OFF.

$r(i), i \geq 1$, of dimension $mn(2S - s - 1)$ denotes the set of states corresponding to the system in which the number of customers in the system is i, the service process is in phase j_1 (in one of n phases) and the arrival process is in phase j_2 (in one of m phases). The level $(i, j, 1, j_1, j_2)$ denotes the case when the inventory level is j, $1 \leq j \leq s$, the production process is ON; the level (i, j, k, j_1, j_2) denotes the case when the inventory level is j, $s+1 \leq j \leq S-1$, the production process is OFF for $k = 0$ and is ON for $k = 1$; and the level $(i, S, 0, j_1, j_2)$ gives the case when the inventory level is the maximum level S and the production process is OFF.

The infinitesimal generator matrix Q has a block-tridiagonal matrix structure is given by

$$Q = \begin{pmatrix} B_0 & A_0 & & & \\ C_0 & B & A & & \\ & C & B & A & \\ & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (3)$$

For use in sequel we need to set up the following notations. $e(j)$ is a unit column vector is of dimension j; e_i is a column vector with 1 in the i^{th} position and 0 elsewhere; I is an identity matrix. The Kronecker product and Kronecker sum are given with the symbols $\otimes$ and $\oplus$, respectively. Finally, we define $d_1 = m(2S - s - 1)$ and $d_2 = mn(2S - s - 1)$ to show the dimensions of the matrices in the generator Q.

The matrices A_0 and A in the generator Q have dimensions $(d_1 \times d_2)$ and $(d_2 \times d_2)$, respectively.

$$A_0 = \begin{pmatrix} \beta \otimes D_1 & & & & & & \\ & \ddots & & & & & \\ & & \beta \otimes D_1 & & & & \\ & & & A_{01} & & & \\ & & & & \ddots & & \\ & & & & & A_{01} & \\ & & & & & & \beta \otimes D_1 \end{pmatrix} with\ A_{01} = I_2 \otimes (\beta \otimes D_1). \tag{4}$$

$$A = \begin{pmatrix} I_n \otimes D_1 & & & & & & \\ & \ddots & & & & & \\ & & I_n \otimes D_1 & & & & \\ & & & A_1 & & & \\ & & & & \ddots & & \\ & & & & & A_1 & \\ & & & & & & I_n \otimes D_1 \end{pmatrix} with\ A_1 = I_2 \otimes (I_n \otimes D_1). \tag{5}$$

The matrices B_0 and B in the generator Q have dimensions $(d_1 \times d_1)$ and $(d_2 \times d_2)$, respectively.

$$B_0 = \begin{pmatrix} D_0 - \eta I_m & \eta I_m & & & & & & \\ & \ddots & \ddots & & & & & \\ & & D_0 - \eta I_m & \eta I_m & & & & \\ & & & D_0 - \eta I_m & B_{02} & & & \\ & & & & B_{01} & B_{03} & & \\ & & & & & \ddots & \ddots & \\ & & & & & B_{01} & B_{03} & \\ & & & & & & B_{01} & B_{04} \\ & & & & & & & D_0 \end{pmatrix} with \tag{6}$$

$$B_{01} = \begin{bmatrix} D_0 & \\ & D_0 - \eta I_m \end{bmatrix},\ B_{02} = e_2'(2) \otimes \eta I_m,\ B_{03} = \begin{bmatrix} & \\ & \eta I_m \end{bmatrix},\ B_{04} = e_2(2) \otimes \eta I_m.$$

$$B = \begin{pmatrix} T \oplus D_0 - \eta I_{mn} & \eta I_{mn} & & & & & & \\ & T \oplus D_0 - \eta I_{mn} & \ddots & & & & & \\ & & \ddots & \eta I_{mn} & & & & \\ & & & T \oplus D_0 - \eta I_{mn} & B_2 & & & \\ & & & & B_1 & B_3 & & \\ & & & & & \ddots & \ddots & \\ & & & & & B_1 & B_3 & \\ & & & & & & B_1 & B_4 \\ & & & & & & & T \oplus D_0 \end{pmatrix} \tag{7}$$

$$with\ B_1 = \begin{bmatrix} T \oplus D_0 & \\ & T \oplus D_0 - \eta I_{mn} \end{bmatrix},\ B_2 = e_2'(2) \otimes \eta I_{mn},\ B_3 = \begin{bmatrix} & \\ & \eta I \end{bmatrix},\ B_4 = e_2(2) \otimes \eta I_{mn}.$$

The matrices C_0 and C in the generator Q have dimensions $(d_2 \times d_1)$ and $(d_2 \times d_2)$, respectively.

$$C_0 = \begin{pmatrix} T^0 \otimes I_m & & & & & & & \\ T^0 \otimes I_m & & & & & & & \\ & \ddots & & & & & & \\ & & T^0 \otimes I_m & & & & & \\ & & & C_{01} & & & & \\ & & & & C_{02} & & & \\ & & & & & \ddots & & \\ & & & & & & C_{02} & \\ & & & & & & & C_{03} \quad 0 \end{pmatrix} with \tag{8}$$

$$C_{01} = e(2) \otimes (T^0 \otimes I_m), \quad C_{02} = I_2 \otimes (T^0 \otimes I_m), \quad C_{03} = e_1'(2) \otimes (T^0 \otimes I_m).$$

$$C = \begin{pmatrix} T^0\beta \otimes I_m & & & & & & & \\ T^0\beta \otimes I_m & & & & & & & \\ & \ddots & & & & & & \\ & & T^0\beta \otimes I_m & & & & & \\ & & & C_1 & & & & \\ & & & & C_2 & & & \\ & & & & & \ddots & & \\ & & & & & & C_2 & \\ & & & & & & & C_3 \quad 0 \end{pmatrix} with \tag{9}$$

$$C_1 = e(2) \otimes (T^0\beta \otimes I_m), \quad C_2 = I_2 \otimes (T^0\beta \otimes I_m), \quad C_3 = e_1'(2) \otimes (T^0\beta \otimes I_m).$$

3. THE STEADY-STATE ANALYSIS

3.1. Stability Condition

Let $\pi = (\pi_1, \cdots, \pi_s, \pi_{s+1}, \cdots, \pi_{S-1}, \pi_S)$ be the steady-state probability vector of the generator matrix $F = A + B + C$. The probability vector satisfies

$$\pi F = 0, \qquad \pi e = 1. \tag{10}$$

Lemma: We have

$$\left[\sum_{i=1}^{s} \pi_i + \sum_{i=s+1}^{S-1} \left(\pi_{i,0} + \pi_{i,1}\right) + \pi_S \right] = \mu\beta(-T)^{-1} \otimes \delta. \tag{11}$$

where μ is the service rate and δ is the probability vector of the generator D.

Proof: The steady-state equations in (10) are rewritten as follows

$$\pi_1[(I_n \otimes D_1) + (T \oplus D_0) - \eta I_{mn} + (T^0\beta \otimes I_m)] + \pi_2(T^0\beta \otimes I_m) = 0,$$

$$\begin{gathered}\pi_{j-1}\eta I_{mn}+\pi_j[(I_n\otimes D_1)+(T\oplus D_0)-\eta I_{mn}]+\pi_{j+1}(T^0\beta\otimes I_m)=0,\\ 2\le j\le s-1\\ \pi_{s-1}\eta I_{mn}+\pi_s[(I_n\otimes D_1)+(T\oplus D_0)-\eta I_{mn}]+\pi_{s+1}C_1=0,\\ \pi_sB_2+\pi_{s+1}[A_1+B_1]+\pi_{s+2}C_2=0,\\ \pi_{j-1}B_3+\pi_j[A_1+B_1]+\pi_{j+1}C_2=0,\qquad s+2\le j\le S-2\\ \pi_{S-2}B_3+\pi_{S-1}[A_1+B_1]+\pi_SC_3=0,\\ \pi_{S-1}B_4+\pi_S[(I_n\otimes D_1)+(T\oplus D_0)]=0,\end{gathered}\tag{12}$$

with the normalizing condition

$$\left[\sum_{i=1}^{s}\pi_i+\sum_{i=s+1}^{S-1}\left(\pi_{i,0}+\pi_{i,1}\right)+\pi_S\right]e=1.\tag{13}$$

Now adding the equations given in (12) we get

$$\left[\sum_{i=1}^{s}\pi_i+\sum_{i=s+1}^{S-1}\left(\pi_{i,0}+\pi_{i,1}\right)+\pi_S\right][(T+T^0\beta)\oplus D]=0.\tag{14}$$

The last equation yields the result in the Lemma from the uniqueness of the steady-state vector of the generators $(T+T^0\beta)$ and D along with the normalizing condition in (13).

Theorem: The production inventory system with MAP arrival and phase-type services under study is stable *if and only if* the following condition is satisfied.

$$\lambda<\mu\tag{15}$$

where μ and λ are the service rate and arrival rate, respectively.

Proof: The production inventory model under study is stable if and only if $\pi Ae<\pi Ce$ in Neuts (1981). The stability condition is given as

$$\left[\sum_{i=1}^{s}\pi_i+\sum_{i=s+1}^{S-1}\left(\pi_{i,0}+\pi_{i,1}\right)+\pi_S\right](I\otimes D_1)e<\left[\sum_{i=1}^{s}\pi_i+\sum_{i=s+1}^{S-1}\left(\pi_{i,0}+\pi_{i,1}\right)+\pi_S\right](T^0\beta\otimes I)\,e.\tag{16}$$

The proof is completed as in (15).

3.2. Steady State Probability Vector

Let $x=(x^*,x(1),x(2),\cdots)$ denote the steady-state probability vector of the generator matrix Q in (3). The vector satisfies

$$xQ=0,\qquad xe=1.\tag{17}$$

The vector x^* of dimension $m(2S-s-1)$ is partitioned into vectors of dimension m as $x^* = [x^*(1,1), \cdots, x^*(s,1), x^*(s+1,0), x^*(s+1,1), \cdots, x^*(S-1,0), x^*(S-1,1), x^*(S,0)]$. The vector $x^*(j,1), 1 \le j \le s$, denotes the probability that the system is idle, the inventory level is j, the production process is ON and the arrival process is in one of m phases. The vector $x^*(j,k), s+1 \le j \le S-1$, gives the probability that the system is idle, the inventory level is j, the production process is OFF for $k=0$ (is ON for $k=1$) and the arrival process is in one of m phases. The vector $x^*(S,0)$ is the probability that the system is idle, the inventory level is the maximum level S and the production process is OFF and the arrival process is in one of m phases.

The vector $x(i), i \ge 1$, of dimension $mn(2S-s-1)$ is partitioned into vectors of dimension mn as $x(i) = [x(i,1,1), \cdots, x(i,s,1), x(i,s+1,0), x(i,s+1,1), \cdots, x(i,S-1,0), x(i,S-1,1), x(i,S,0)]$. The vector $x(i,j,1), 1 \le j \le s$, denotes the probability that the number of customers in the system i, the inventory level is j, the production process is ON, the service process is in one of n phases and the arrival process is in one of m phases. The vector $x(i,j,k), s+1 \le j \le S-1$, gives the probability that the number of customers in the system i, the inventory level is j, the production process is OFF for $k=0$ (is ON for $k=1$), the service process is in one of n phases and the arrival process is in one of m phases. The vector $x(i,S,0)$ is the probability that the number of customers in the system i, the inventory level is the maximum level S, the production process is OFF, the service process is in one of n phases and the arrival process is in one of m phases.

Under the stability condition given in (15) the steady-state probability vector x is obtained (see Neuts (1981)) as

$$x(i) = x(1)R^{i-1}, \qquad i \ge 2, \tag{18}$$

where the rate matrix R is the minimal nonnegative solution to the matrix quadratic equation

$$R^2C + RB + A = 0, \tag{19}$$

and the vectors x^* and $x(1)$ are obtained by solving

$$\begin{aligned} x^*B_0 + x(1)C_0 &= 0, \\ x^*A_0 + x(1)(B+RC) &= 0, \end{aligned} \tag{20}$$

subject to the normalizing condition

$$x^*e + x(1)(I-R)^{-1}e = 1. \tag{21}$$

3.3. Performance Measures

Some performance measures of the system under study are listed in this section.

The mean number of customers in the system

$$E_N = \sum_{i=1}^{\infty} i\,x(i)e = x(1)(I-R)^{-2}e. \tag{22}$$

The mean inventory level

$$E_I = \left[\sum_{j=1}^{s} jx^*(j,1)\, e + \sum_{j=s+1}^{S-1} j[x^*(j,0) + x^*(j,1)]e + Sx^*(S,0)e\right] + \sum_{i=1}^{\infty}\left[\sum_{j=1}^{s} jx(i,j,1)\, e + \sum_{j=s+1}^{S-1} j[x(i,j,0) + x(i,j,1)]e + Sx(i,S,0)e\right]. \quad (23)$$

The mean production rate

$$E_{PR} = \eta \sum_{j=1}^{S-1} x^*(j,1)e + \eta \sum_{i=1}^{\infty}\sum_{j=1}^{S-1} x(i,j,1)e. \quad (24)$$

The mean rate at which production process is switched ON

$$ER_{ON} = \mu \sum_{i=1}^{\infty} x(i, s+1, 0)\, e. \quad (25)$$

The mean local purchase rate

$$ER_{LP} = \mu \sum_{i=1}^{\infty} x(i,1,1)\, e. \quad (26)$$

4. NUMERICAL STUDY

We perform the numerical examples to represent the effects of various system parameters on the performance measures in the Section 4.1 and to discuss optimum inventory policies by using a constructed cost in the Section 4.2.

We consider the following five sets of values for the arrival process. The arrival processes have the same mean of 1 but each one of them is qualitatively different. The values of the standard deviation of the interarrival times of the arrival processes with respect to ***ERLA*** are, respectively, 1, 1.41421, 3.17451, 1.99336, and 1.99336. The MAP processes are normalized to have a specific arrival rate λ as given in Chakravarthy (2010). The arrival processes labeled ***MNCA*** and ***MPCA*** have negative and positive correlation for two successive interarrival times with values -0.4889 and 0.4889, respectively, whereas the first three arrival processes have zero correlation for two successive interarrival times.

Erlang distribution (***ERLA***):

$$D_0 = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}, D_1 = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}. \quad (27)$$

Exponential distribution (***EXPA***):

$$D_0 = (-1), D_1 = (1). \quad (28)$$

Hyper-exponential distribution (***HEXA***):

$$D_0 = \begin{pmatrix} -1.9 & 0 \\ 0 & -0.19 \end{pmatrix}, D_1 = \begin{pmatrix} 1.71 & 0.19 \\ 0.171 & 0.019 \end{pmatrix}. \tag{29}$$

MAP with negative correlation (***MNCA***):

$$D_0 = \begin{pmatrix} -1.00222 & 1.00222 & 0 \\ 0 & -1.00222 & 0 \\ 0 & 0 & -225.75 \end{pmatrix}, D_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0.01002 & 0 & 0.9922 \\ 223.4925 & 0 & 2.2575 \end{pmatrix}. \tag{30}$$

MAP with positive correlation (***MPCA***):

$$D_0 = \begin{pmatrix} -1.00222 & 1.00222 & 0 \\ 0 & -1.00222 & 0 \\ 0 & 0 & -225.75 \end{pmatrix}, D_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0.9922 & 0 & 0.01002 \\ 2.2575 & 0 & 223.4925 \end{pmatrix}. \tag{31}$$

For the service times, we consider the following three phase-type distributions. The phase-type distributions have the same mean of 1 but each one of them is qualitatively different. The values of the standard deviation of the distributions are, respectively, 0.70711, 1, and 2.24472. The distributions are normalized at a specific value for the service rate μ.

Erlang distribution (***ERLS***):

$$\beta = (1 \quad 0), T = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}. \tag{32}$$

Exponential distribution (***EXPS***):

$$\beta = (1), T = (-1). \tag{33}$$

Hyper-exponential distribution (***HEXS***):

$$\beta = (0.9 \quad 0.1), T = \begin{pmatrix} -1.9 & 0 \\ 0 & -0.19 \end{pmatrix}. \tag{34}$$

4.1. Effects of the System Parameters

Example 1: We discuss the behavior of the performance measures, ER_{LP}, E_{PR} and E_I under the various service time distributions and the arrival processes in Figure 2. Towards this end, we fix the inventory policy $(s, S) = (3,8)$, the arrival rate $\lambda = 0.5$, the service rate $\mu = 1.1$ and vary the production rate $\eta = 0.5, 0.7, \dots, 2.5$.

The mean local purchase rate ER_{LP} appears to decrease with increasing values of η. In the cases both of ERLS and HEXS, the effect of variability in the arrival processes on ER_{LP} is very significant at the low production rate. The MAP processes with positive correlation labeled MPCA and with high variability labeled HEXA are significantly separated from the other MAP processes, especially for the systems where the production rate is low. The values of ER_{LP} are the lower in the case of HEXS where is phase-type distribution with high variability. We can say that the effect of variability in the distributions of service times on ER_{LP} is important by depending on the MAP processes (i.e., see MPCA and HEXA).

When the production rate increases, it is expected to increasing the mean production rate E_{PR} as in Figure 2. Similar to previous comments, the values of E_{PR} in the cases MPCA and

HEXA are significantly separated from the values in the other MAP processes, especially for the systems where the production rate is low. The values of ER_{LP} are slightly lower in the case of HEXS.

The mean inventory level E_I increases as the production rate η increases, especially when lower production rates. The variabilities of the MAP processes and PH-distributions have slightly effects on the values of E_I except to MPCA with positive correlation.

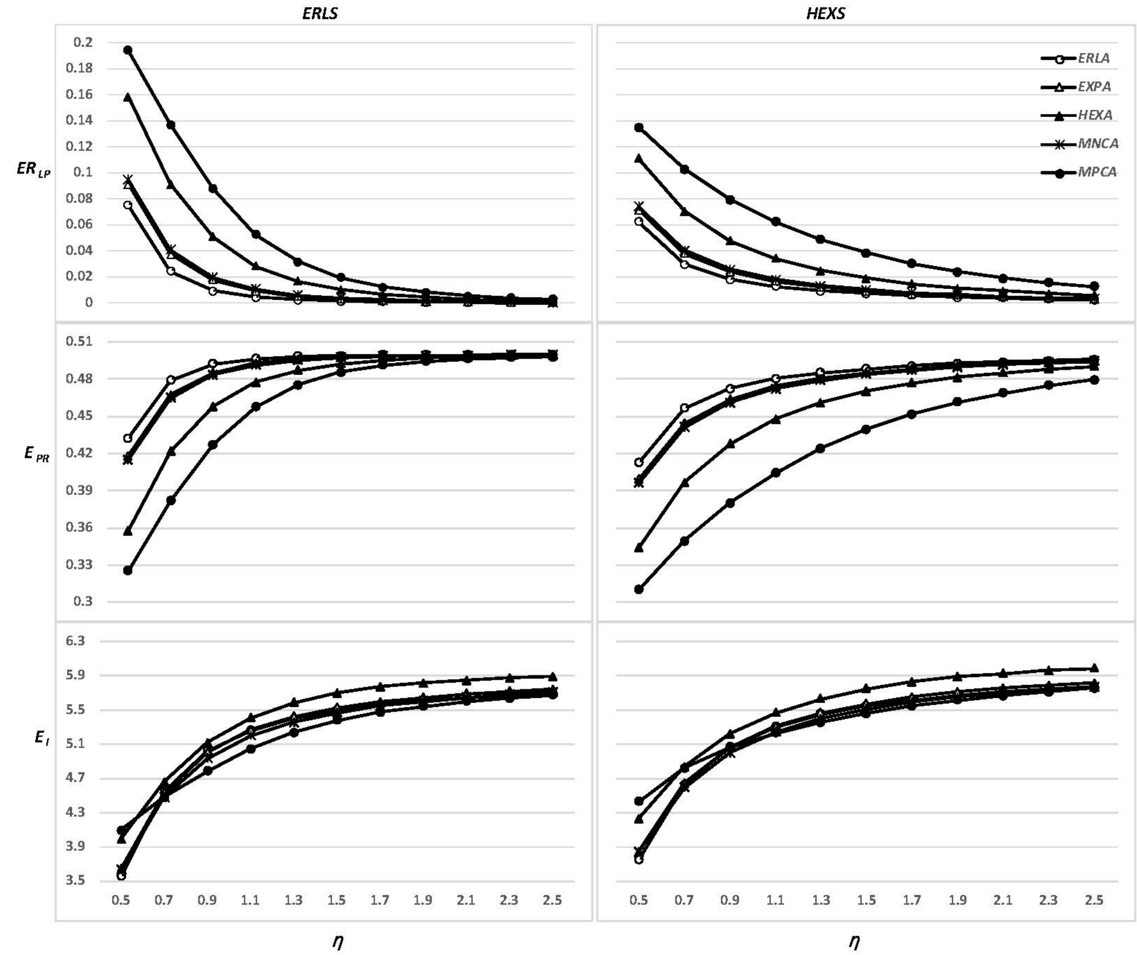

Figure 2:
Effect of the production rate η on various performance measures

Example 2: We investigate the effects of the arrival rate and the service rate on the mean number of customers in the system E_N under the various scenarios by displaying the plots in Figure 3. For this purpose, we fix the inventory policy $(s, S) = (3,8)$ and the production rate $\eta = 0.3$ on some service time distributions and arrival processes. Moreover, we fix the arrival rate $\lambda = 1$ (the service rate $\mu = 1.1$) and vary the service rate μ (the arrival rate λ) for the left-side of Figure 3 (for the right-side of Figure 3).

As expected, the mean number of customers in the system E_N decreases with increasing values of μ in Figure 3a and increases with increasing values of λ in Figure 3b.

The values of E_N dramatically increases by depending on the MAP arrivals at lower service rates or higher arrival rates, in other words, when the system has high traffic intensity. For example, the values are around 25, 45 and 450 for ERLA, HEXA and MPCA, respectively, in the case of $\mu = 1.1$ and HEXS in Figure 3a.

When looking at the effect of the variability of service distribution, at lower μ values (when the system is on high traffic intensity), the effect of service variability on E_N is significant, except for the case MPCA. When there is a positive correlation in arrivals (MPCA), the effect of service distribution disappears. The all comments can be seen in Figure 3a.

Similar comments about the effect of the variability of service distribution can be said for the higher λ values (when the system is on high traffic intensity) in Figure 3b. The effect of variability on E_N is significant. The effect of service distribution disappears in the case of MPCA.

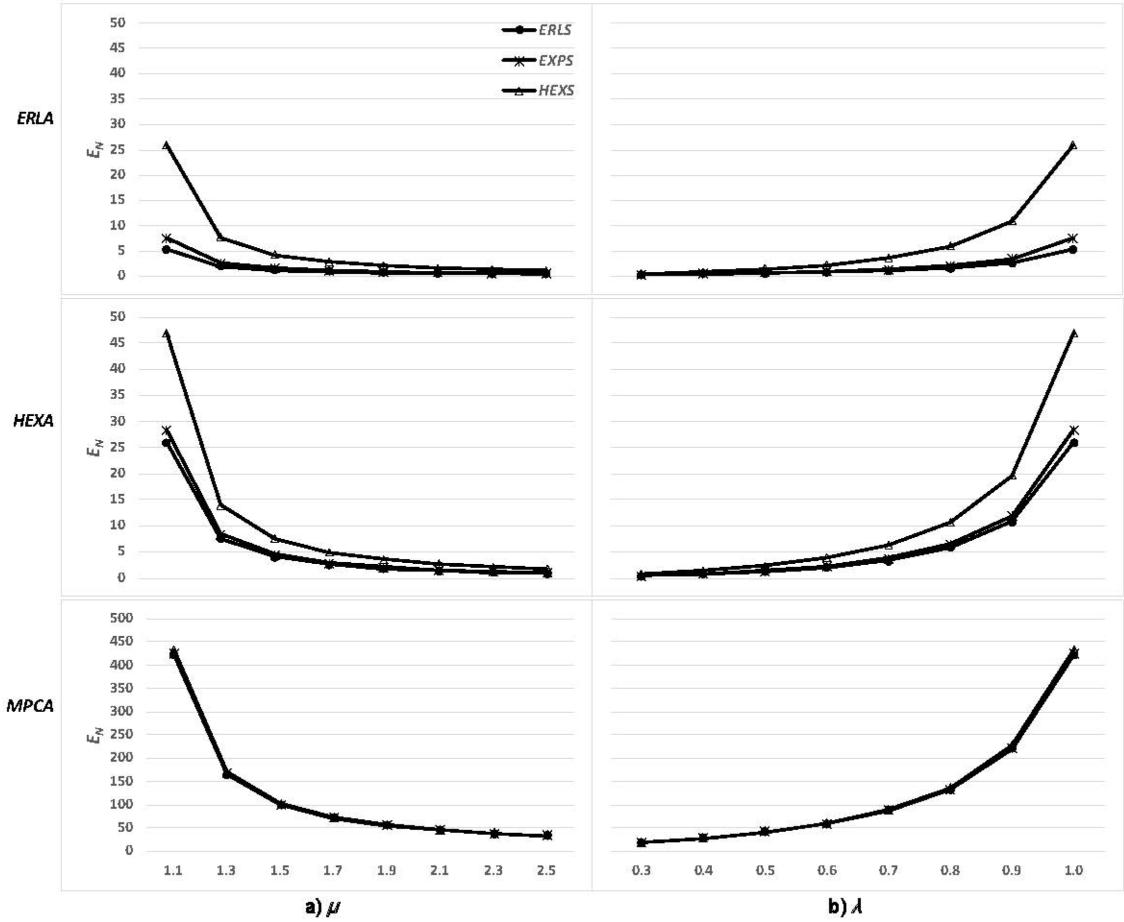

Figure 3:
Behavior of the performance measure E_N under the various scenarios;
***a)** for the service rate, **b)** for the arrival rate*

Example 3: We look at the effects of the arrival rate and the service rate on the mean inventory level E_I considering various the distributions of service times and the arrival processes. For this purpose, we fix the inventory policy $(s, S) = (3,8)$ and vary the production rate $\eta = 0.3, 1$. Moreover, we fix the arrival rate $\lambda = 1$ (the service rate $\mu = 1.1$) and vary the service rate μ from 1.1 to 2.5 (the arrival rate λ from 0.3 to 1) in Figure 4a (in Figure 4b).

As we showed earlier in Figure 2, the values of E_I increases as the production rate increases (recall the Example 1 has been performed for the constant values of both λ and μ). We can also see this result in all plots in Figure 4 (note that the y-scales in all plots). In this example, we observe the behavior of E_I for two different production rates (consider two different systems producing with low rate and producing with high rate) under various scenarios. We purpose to see if the behavior of variability on E_I is different at different production rates.

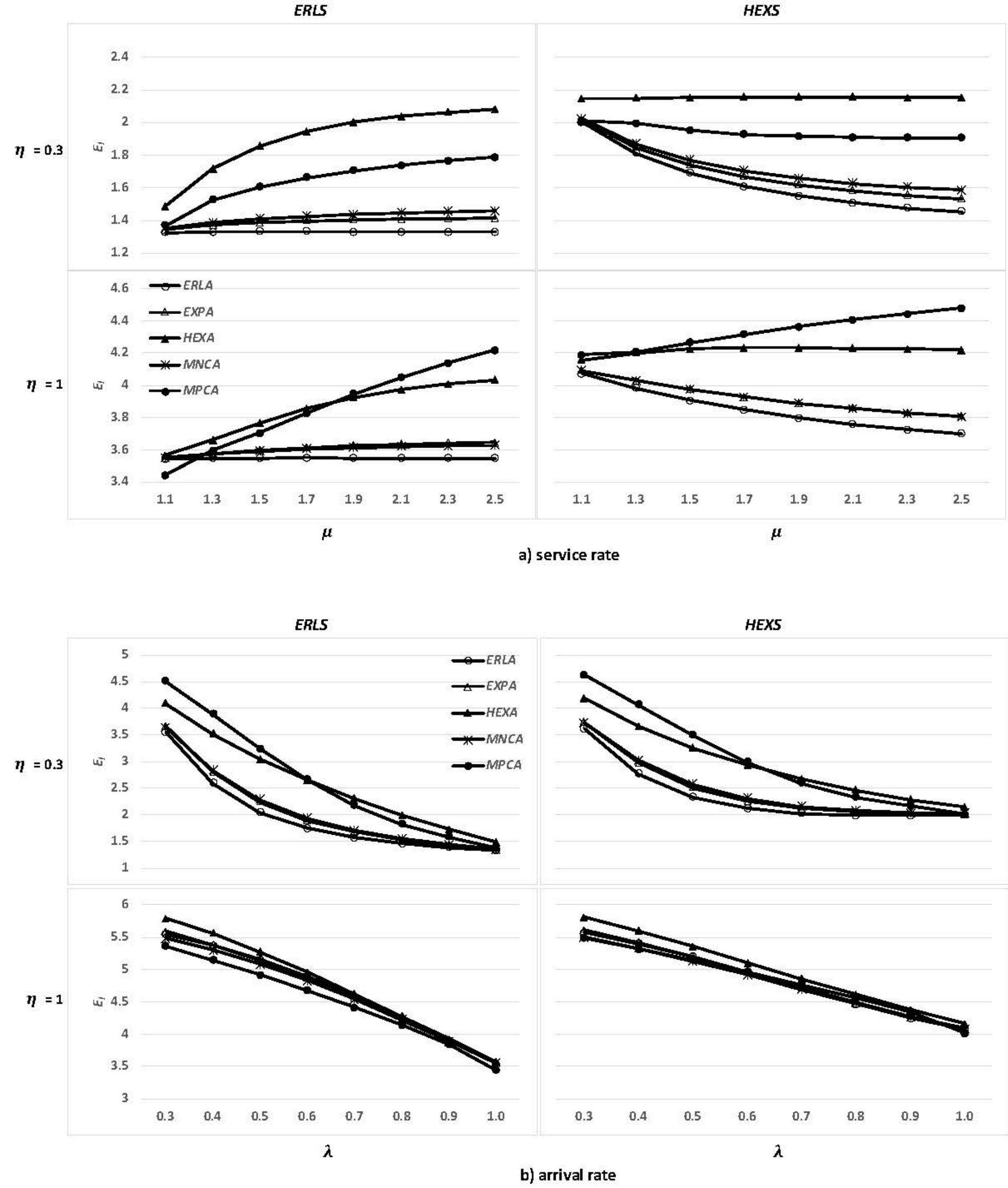

Figure 4:
Behavior of the performance measure E_I under the various scenarios

In Figure 4a, looking when the production occurs with rate $\eta = 0.3$, in the case of ERLS, as the service rate μ increases, the values of E_I are almost constant for the three MAP processes (ERLA, EXPA and MNCA) and the values of E_I increase for the HEXA and MPCA. On the other hand, in the case of HEXS, as the service rate μ increases, the values of E_I are almost constant for the two MAP processes (HEXA and MPCA) and the values of E_I decrease for the

ERLA, EXPA and MNCA. In summary, we can say the following results. If the service distribution has low variability (ERLS), the MAP arrivals have high variability (HEXA and MPCA) make a difference. In contrast, if the service distribution has high variability (HEXS), the MAP arrivals have low variability (ERLA, EXPA and MNCA) make a difference. Looking when the production occurs with rate $\eta = 1$ in Figure 4a, in the case of ERLS, we can say comments similar to the case $\eta = 0.3$. One observation can be added that the increment of the values of E_I occur dramatically for the case MPCA comparing with the case $\eta = 0.3$. In the case of HEXS, we observe results similar to the case $\eta = 0.3$ except to the case of MPCA. In this case, the values of E_I also increase. By the second result, we can say that positive correlation is additionally important.

As expected, it can be seen from all plots in Figure 4b that while the arrival rate λ increased, the values of E_I decreased. Looking at both production rates $\eta = 0.3, 1$, it is seen that MAP has an effect on E_I when the arrival rate is low, especially in the case of $\eta = 0.3$.

4.2. Optimization

In this section, we establish an objective function, ETC, giving the expected total cost per unit of time for discussing optimum inventory policy as follows. Towards finding the optimum values of inventory policy, we fix the unit values of the costs by $c_h = 5$, $c_p = 20$, $c_s = 200$, $c_{lp} = 30$, and $c_w = 250$.

$$ETC = c_h E_I + c_p E_{PR} + c_s ER_{ON} + c_{lp} ER_{LP} + c_w E_W \tag{35}$$

where

c_h : Holding cost of each inventory per unit time,
c_p : Producing cost of each inventory per unit time,
c_s : Cost for starting production,
c_{lp} : Local purchase cost for each inventory,
c_w : Waiting cost of a customer in the queue per unit time.

Example 4: Under various distributions of the service times and the arrival processes, we give the optimum values of the maximum inventory level S^* in this example. We fix, respectively, the arrival rate and the service rate by $\lambda = 1$ and $\mu = 1.1$ and vary the production rates $\eta = 1, 4, 7$.

Table 1. Optimum S^* for $s = 2$

		$\eta = 1$		$\eta = 4$		$\eta = 7$	
Arrival	Service	ETC	S^*	ETC	S^*	ETC	S^*
ERLA	***ERLS***	1376.179	8	1401.591	10	1404.860	11
	EXPS	1940.224	8	1964.812	10	1968.135	10
	HEXS	6526.170	7	6548.225	9	6552.458	10
EXPA	***ERLS***	1976.701	8	2001.869	10	2005.177	11
	EXPS	2545.455	8	2569.795	10	2573.144	10
	HEXS	7135.765	7	7157.598	9	7161.877	10
HEXA	***ERLS***	6544.177	8	6569.160	10	6572.542	10
	EXPS	7145.407	8	7169.472	10	7172.878	10
	HEXS	11820.561	7	11842.016	9	11846.453	9
MNCA	***ERLS***	2039.504	8	2064.653	10	2067.983	10
	EXPS	2601.209	8	2625.527	10	2628.875	10
	HEXS	7180.445	7	7202.190	9	7206.549	9
MPCA	***ERLS***	105691.759	8	105716.006	10	105719.324	11
	EXPS	105980.845	8	106004.136	10	106007.494	10
	HEXS	108257.731	7	108278.205	9	108282.543	10

As the variability in the distribution of the service times increase, the optimum value of S^* decreases and the optimum total cost increases. The result can be seen, specifically, in the cases of HEXS. On the other hand, we observe that the variability in the arrival process has not effect on the optimum inventory policy while it causes the increment of the optimum total cost. Finally, the optimum values both ETC and S^* increase as the production rate η increases as expected.

Table 2. Optimum S^* for $s = 5$

		$\eta = 1$		$\eta = 4$		$\eta = 7$	
Arrival	Service	ETC	S^*	ETC	S^*	ETC	S^*
ERLA	***ERLS***	1379.914	10	1416.479	13	1419.839	14
	EXPS	1944.067	10	1979.617	13	1983.090	13
	HEXS	6530.491	9	6562.805	12	6567.401	13
EXPA	***ERLS***	1980.467	10	2016.748	13	2020.155	14
	EXPS	2549.375	10	2584.587	13	2588.095	13
	HEXS	7140.182	9	7172.168	12	7176.818	13
HEXA	***ERLS***	6547.641	10	6584.021	13	6587.512	13
	EXPS	7149.177	10	7184.234	13	7187.823	13
	HEXS	11825.129	9	11856.547	12	11861.376	12
MNCA	***ERLS***	2043.205	10	2079.530	13	2082.877	14
	EXPS	2605.070	10	2640.317	13	2643.826	13
	HEXS	7184.945	9	7216.758	12	7221.477	12
MPCA	***ERLS***	105695.546	10	105730.438	13	105733.863	14
	EXPS	105984.947	10	106018.462	13	106021.999	13
	HEXS	108262.744	9	108292.242	12	108296.986	13

5. CONCLUSIONS

In this study, we considered a production inventory system with MAP arrival and phase-type service time in which one item is purchased when the inventory level fall to zero. There is one production facility in the system and it is governed by (s, S)-policy. We obtained the stability condition of the production inventory system in closed form and then established its steady-state analysis under stability by using the matrix-geometric method. Finally, some numerical examples were performed to see the effects of the parameters on the performance measures, to see the effect of the variabilities of both the MAP process and the service process on the measures, and to define the optimum inventory policy. In all the examples, we observed that the variability in the inter-arrival times (the MAP processes) and the variability in the distributions of the service times (PH-distributions) affect the values of the performance measures and the optimum inventory policy. This observation is very important in the modelling of real systems.

CONFLICT OF INTEREST

The author confirms that there is no known conflict of interest or common interest with any institution/organization or person.

AUTHOR CONTRIBUTION

The author takes all responsibility of the manuscript.

REFERENCES

1. Artalejo, J.R., Gomez-Corral, A. and He, Q.M. (2010) Markovian arrivals in stochastic modelling: a survey and some new results. *SORT: Statistics and Operations Research Transactions*, 34(2), 101-144.

2. Baek, J.W. and Moon, S.K. (2014) The $M/M/1$ queue with a production-inventory system and lost sales, *Applied Mathematics and Computation*, 233, 534-544. doi.org/10.1016/j.amc.2014.02.033

3. Baek, J.W. and Moon, S.K. (2016) A production-inventory system with a Markovian service queue and lost sales, *Journal of the Korean Statistical Society*, 45, 14-24. doi.org/10.1016/j.jkss.2015.05.002

4. Barron, Y. (2022) The continuous (S, s, S_e) inventory model with dual sourcing and emergency orders, *European Journal of Operational Research*, 301, 18-38. doi.org/10.1016/j.ejor.2021.09.021

5. Chakravarthy, S.R. (2001) The batch Markovian arrival process: a review and future work. In: Krishnamoorthy, A., Raju, N. and Ramaswami, V. (eds) Advances in probability and stochastic processes. Notable, New Jersey, 21-49.

6. Chakravarthy, S.R. (2010) Markovian arrival processes, *Wiley Encyclopedia of Operations Research and Management Science*. doi.org/10.1002/9780470400531.eorms0499

7. Chakravarthy, S.R. (2020) Queueing-inventory models with batch demands and positive service times, *Automation and Remote Control*, 81, 713-730. doi.org/10.1134/S0005117920040128

8. Chakravarthy, S.R. and Rumyantsev, A. (2020) Analytical and simulation studies of queueing-inventory models with *MAP* demands in batches and positive phase type services, *Simulation Modelling Practice and Theory*, 103, 1-15. doi.org/10.1016/j.simpat.2020.102092

9. De la Cruz, N.N. and Daduna, H. (2022) Analysis of second order properties of production–inventory systems with lost sales, *Annals of Operations Research*. doi.org/10.1007/s10479-022-05061-z

10. He, Q.M. and Zhang, H. (2013) Performance analysis of an inventory–production system with shipment consolidation in the production facility, *Performance Evaluation*, 70(9), 623-638. doi.org/10.1016/j.peva.2013.05.007

11. Karthikeyan, K. and Sudhesh, R. (2016) Recent review article on queueing inventory systems, *Research Journal of Pharmacy and Technology*, 9(11), 1451-1461. doi:10.5958/0974-360X.2016.00421.2

12. Krishnamoorthy, A. and Raju, N. (1998) (s, S) inventory with lead time-the *N*-policy. *International Journal of Information and Management Sciences*, 9, 45-52.

13. Krishnamoorthy, A., Lakshmy, B. and Manikandan, R. (2011a) A survey on inventory models with positive service time, *OPSEARCH*, 48, 153-169. doi.org/10.1007/s12597-010-0032-z

14. Krishnamoorthy, A., Viswanath, C. and Narayanan, V.C. (2011b) Production inventory with service time and vacation to the server, *IMA Journal of Management Mathematics*, 22, 33-45. doi.org/10.1093/imaman/dpp025

15. Krishnamoorthy, A. and Narayanan, V.C. (2013) Stochastic decomposition in production inventory with service time, *European Journal of Operational Research*, 228, 358-366. doi.org/10.1016/j.ejor.2013.01.041

16. Krishnamoorthy, A., Nair, S.S. and Narayanan, V.C. (2015) Production inventory with service time and interruptions, *International Journal of Systems Science*, 46(10), 1800-1816. doi.org/10.1080/00207721.2013.837538

17. Latouche, G. and Ramaswami, V. (1999) *Introduction to Matrix Analytic Methods in Stochastic Modelling*, ASASIAM, Philadelphia.

18. Melikov, A., Mirzayev, R. and Sztrik, J. (2023) Double-sources queuing-inventory systems with finite waiting room and destructible stocks, *Mathematics*, 11, 226. doi.org/10.3390/math11010226

19. Neuts, M.F. (1981) *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*, The Johns Hopkins University Press, Baltimore, MD. [1994 version is Dover Edition].

20. Saffari, M., Asmussen, S. and Haji, R. (2013) The $M/M/1$ queue with inventory, lost sale, and general lead times, *Queueing Systems*, 75, 65-77. doi.org/10.1007/s11134-012-9337-3

21. Schwarz, M., Sauer, C., Daduna, H., Kulik, R. and Szekli, R. (2006) $M/M/1$ queueing system with inventory, *Queueing Systems: Theory and Applications*, 54, 55-78. doi.org/10.1007/s11134-006-8710-5

22. Sigman, K. and Simchi-Levi, D. (1992) Light traffic heuristic for an $M/G/1$ queue with limited inventory, *Annals of Operations Research*, 40, 371-380. doi.org/10.1007/BF02060488

23. Shajin, D., Krishnamoorthy, A., Melikov, A.Z. and Sztrik, J. (2022) Multi-server queuing production inventory system with emergency replenishment, *Mathematics*, 10(20), 3839. doi.org/10.3390/math10203839