Теория вероятностей и математическая статистика

Probability theory and mathematical statistics

УДК 519.872

АНАЛИЗ СИСТЕМЫ ОБСЛУЖИВАНИЯ-ЗАПАСАНИЯ С УЧЕТОМ ВРЕМЕНИ ОБСЛУЖИВАНИЯ ФАЗОВОГО ТИПА В УСЛОВИЯХ ГИБРИДНОЙ ПОЛИТИКИ ПОПОЛНЕНИЯ

А. МЕЛИКОВ[1)]***, Ш. ОЗКАР***[2)]

[1)]*Бакинский инженерный университет,*
ул. Гасана Алиева, 120, AZ0101, г. Хырдалан, Азербайджан
[2)]*Балыкесирский университет, кампус Чагыш, 10145, г. Балыкесир, Турция*

Аннотация. Исследуется одноканальная система обслуживания-запасания с гибридной политикой пополнения запасов. Поступление заявок подчиняется пуассоновскому процессу, а время обслуживания имеет фазовое распределение. Гибридная политика пополнения запасов определяется следующим образом: когда уровень запасов падает до точки s, применяется политика (s, Q) при условии, что количество заявок в системе меньше порогового

Образец цитирования:
Меликов А, Озкар Ш. Анализ системы обслуживания-запасания с учетом времени обслуживания фазового типа в условиях гибридной политики пополнения. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;2:77–88 (на англ.).
EDN: ONFKIK

For citation:
Melikov A, Ozkar S. Analysis of queuing-inventory system with phase-type service time under hybrid replenishment policy. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;2:77–88.
EDN: ONFKIK

Авторы:
Агаси Меликов – доктор технических наук, член-корреспондент НАН Азербайджана, профессор; проректор по научной работе, профессор кафедры математики и информатики факультета информационных и компьютерных технологий.
Шерифе Озкар – доктор философии по статистике, доцент; доцент кафедры международной торговли и логистики факультета экономики и менеджмента.

Authors:
Agassi Melikov, doctor of science (engineering), corresponding member of the National Academy of Sciences of Azerbaijan, full professor; vice-rector for scientific affairs and professor at the department of mathematics and computer science, faculty of information and computer technologies.
amelikov@beu.edu.az
https://orcid.org/0000-0003-3721-7820
Serife Ozkar, PhD (statistics), docent; associate professor at the department of international trade and logistics, faculty of economics and administrative sciences.
serife.ozkar@balikesir.edu.tr
https://orcid.org/0000-0003-3475-5666

CC BY NC

значения *r*; если в этот момент количество заявок больше или равно *r*, применяется политика (s, S). Математическая модель исследуемой системы построена как многомерная непрерывная цепь Маркова. Установлено условие устойчивости, и стационарное распределение системы рассчитано с использованием матрично-геометрического метода. Представлены численные результаты.

Ключевые слова: система обслуживания-запасания; гибридная политика пополнения; распределения фазового типа; матрично-геометрический метод.

ANALYSIS OF QUEUING-INVENTORY SYSTEM WITH PHASE-TYPE SERVICE TIME UNDER HYBRID REPLENISHMENT POLICY

A. MELIKOV[a]***, S. OZKAR***[b]

[a]*Baku Engineering University, 120 Hasan Aliyev Street, Khirdalan AZ0101, Azerbaijan*
[b]*Balikesir University, Çağış Campus, Balikesir 10145, Turkey*

Corresponding author: A. Melikov (amelikov@beu.edu.az)

Abstract. We investigate a single-server queuing-inventory system with a hybrid replenishment policy. Customer arrivals follow a Poisson process and service times have a phase-type distribution. A hybrid inventory replenishment policy is defined as follows: when the inventory level falls to *s*, the (s, Q) policy is applied, provided that the number of customers in the system is less than a threshold *r*; if at this point the customers quantity is greater than or equal to *r*, the (s, S) policy is applied. The mathematical model of the investigated queuing-inventory system is constructed as a multi-dimensional continuous-time Markov chain. The stability condition is established, and steady-state distribution is calculated using the matrix-geometric method. Numerical results are presented.

Keywords: queuing-inventory system; hybrid replenishment policy; phase-type service distribution; matrix-geometric method.

Introduction

Many real-world systems can be modelled using queuing-inventory systems (QIS), which combine the properties of both queuing and inventory systems. These systems have been intensively studied by various authors in the last three decades (see the review papers [1–5]). One of the main topics encountered in the QIS models is to determine the replenishment policy (RP). In practice, policies (s, S) and (s, Q) are usually used as a RP. In (s, S) policy, when the inventory level drops to the reorder point *s*, $0 \le s < S$, an order is placed and upon replenishment the inventory level becomes *S*, where *S* denotes the maximum warehouse (storage) capacity of the system. The policy states that the replenishment quantity varies in order to fill the maximum capacity of the inventory when the reorder is placed. This RP is also commonly referred to as «Up to *S*» policy. In (s, Q) policy, when the inventory level drops to *s*, $s < \frac{S}{2}$, an order quantity of $Q = S - s$ is placed and upon replenishment the inventory level becomes a sum of the current items in the inventory and order quantity, i. e. (s, Q) policy states that the replenishment quantity is always fixed.

Above-mentioned policies do not take into account the current number of customers in the system. Therefore, these policies can lead to increasing the reorder rate or unreasonably high levels of inventory that have no buyers. In both cases, the key performance indicators of the QIS under study deteriorate. QIS models with queue-independent RPs are discussed in detail in a large amount of available literature (see, for example, recent works [6–12] and their bibliography).

Clearly, RPs that take into account both the current queue length and the inventory level are more efficient than state-independent policies. In recent years, some authors have begun to study QIS models, in which the type of RP changes only depending on the current inventory level, i. e. they do not take into account the number of customers in the queue (see [13–17] and their bibliography). To the best of our knowledge, QIS models under queue-dependent RPs have received little research. The first work devoted to the study of the QIS model, in which

the RP type changes depending on both the current inventory level and the queue length, is paper [18]. In that paper, the QIS model was considered under simple Poisson/exponential assumptions. This work is an extension of the model considered in paper [18] for the case of phase-type distribution (PH-distribution) of service time.

Model description

Customers arrive to the system according to Poisson process with parameter λ. Upon arrival, customers line up in one infinite queue. Customer service is carried out in the order of their arrival, and after servicing a customer the inventory level decreases by one item. We define the threshold parameter r, $r > 0$, for the number of customers in the system. The hybrid RP is defined as follows: (i) if the inventory level drops to s and the number of customers in the system is less than r, then $(s,\ Q)$ policy is applied; (ii) when the inventory level drops to s and the number of customers in the system is more or equal to r, then $(s,\ S)$ policy is applied.

The service times follow PH-distribution with parameters $(\beta,\ T)$. The service rate is given by $\mu = \left[\beta(-T)^{-1}e\right]^{-1}$ and T^0 is a column vector satisfying $Te + T^0 = 0$. Formulas for computation of other characteristics of the PH-distribution can be found in papers [19–21].

The customer arriving at the moment when the warehouse is empty is lost and permanently leaves the system. When inventory levels drop to zero, customers in the queue begin to lose patience, i. e., at these moments, only one customer (at any point in the queue) becomes impatient, regardless of the length of the queue. The impatience time is assumed to have an exponential distribution with parameter τ.

Lead times follow exponential distribution with parameter θ_1 for $(s,\ Q)$ policy (with parameter θ_2 for $(s,\ S)$ policy).

Our goal is to determine the key performance measures of the QIS under study.

Steady state analysis

Let $N(t)$, $I(t)$ and $J(t)$ denote the number of customers in the system, the inventory level and the phase of the service process at time t, respectively. The process $\{(N(t),\ I(t),\ J(t)),\ t \geq 0\}$ is a continuous-time Markov chain and the state space is given as follows:

$$\Omega = \{(0,\ i) : 0 \leq i \leq S\} \cup \{(n,\ i,\ j) : n \geq 1,\ 0 \leq i \leq S,\ 1 \leq j \leq m\}.$$

The level $\{(0,\ i) : 0 \leq i \leq S\}$ of dimension $(S+1)$ corresponds to the case when there are no customers in the system and the inventory level is i. The level $\{(n,\ i,\ j) : n \geq 1,\ 0 \leq i \leq S,\ 1 \leq j \leq m\}$ of dimension $m(S+1)$ corresponds to the case when there are n customers in the system, the inventory level is i and the service process is in one of m phases.

The generator matrix of the system has a block-tridiagonal matrix structure, which is given by

$$Q = \begin{bmatrix} B_0 & A_0 & & & & & & \\ C_0 & B_1 & A & & & & & \\ & C & B_1 & A & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & C & B_1 & A & & \\ & & & & C & B_2 & A & \\ & & & & & C & B_2 & A \\ & & & & & & \ddots & \ddots & \ddots \end{bmatrix}. \quad (1)$$

The upper-diagonal matrices in (1) are A_0 with dimension $(S+1)\times m(S+1)$ and A with dimension $m(S+1)\times m(S+1)$ are given by

$$A_0 = \begin{pmatrix} 0 & & & \\ & \lambda\beta & & \\ & & \ddots & \\ & & & \lambda\beta \end{pmatrix},\ A = \begin{pmatrix} 0 & & & \\ & \lambda I & & \\ & & \ddots & \\ & & & \lambda I \end{pmatrix}.$$

The lower-diagonal matrices in (1) are C_0 with dimension $m(S+1)\times(S+1)$ and C with dimension $m(S+1)\times m(S+1)$ are given by

$$C_0 = \begin{pmatrix} \tau e_m & & & \\ T^0 & 0 & & \\ & \ddots & \ddots & \\ & & T^0 & 0 \end{pmatrix}, \; C = \begin{pmatrix} \tau I & & & \\ T^0\beta & 0 & & \\ & \ddots & \ddots & \\ & & T^0\beta & 0 \end{pmatrix}.$$

The main-diagonal matrices in (1) are B_0 with dimension $(S+1)\times(S+1)$, B_1 and B_2 with dimension $m(S+1)\times m(S+1)$ are given by

$$B_0 = \begin{pmatrix} -\theta_1 & & \theta_1 & & & \\ & -(\theta_1+\lambda) & & \theta_1 & & \\ & & \ddots & & \ddots & \\ & & -(\theta_1+\lambda) & & & \theta_1 \\ & & & -\lambda & & \\ & & & & \ddots & \\ & & & & & -\lambda \end{pmatrix},$$

$$B_1 = \begin{pmatrix} -(\theta_1+\tau)I & & \theta_1 I & & & \\ & d_1 & & \theta_1 I & & \\ & & \ddots & & \ddots & \\ & & d_1 & & & \theta_1 I \\ & & & T-\lambda I & & \\ & & & & \ddots & \\ & & & & & T-\lambda I \end{pmatrix}$$

with $d_1 = T-(\theta_1+\lambda)I$,

$$B_2 = \begin{pmatrix} -(\theta_2+\tau)I & & & & \theta_2 I \\ & d_2 & & & \theta_2 I \\ & & \ddots & & \vdots \\ & & d_2 & & \theta_2 I \\ & & & T-\lambda I & \\ & & & & \ddots \\ & & & & T-\lambda I \end{pmatrix}$$

with $d_2 = T-(\theta_2+\lambda)I$.

Stability condition

Let $\pi = (\pi_0, \pi_1, \ldots, \pi_s, \pi_{s+1}, \ldots, \pi_S)$ be the steady-state probability vector of the finite generator $D = A + B_2 + C$. This matrix is given by

$$D = \begin{pmatrix} -\theta_2 I & & & & & & \theta_2 I \\ T^0\beta & T-\theta_2 I & & & & & \theta_2 I \\ & \ddots & \ddots & & & & \vdots \\ & & T^0\beta & T-\theta_2 I & & & \theta_2 I \\ & & & T^0\beta & T & & \\ & & & & \ddots & \ddots & \\ & & & & & T^0\beta & T \end{pmatrix}.$$

The probability vector π satisfies the following balance equations:

$$\pi D = 0 \text{ and } \pi e = 1. \tag{2}$$

We can rewrite the equations in (2) as follows:

$$-\pi_0 \theta_2 I + \pi_1 T^0 \beta = 0,$$

$$\pi_i \left(T - \theta_2 I\right) + \pi_{i+1} T^0 \beta = 0,\ 1 \le i \le s,$$

$$\pi_i T + \pi_{i+1} T^0 \beta = 0,\ s+1 \le i \le S-1, \tag{3}$$

$$\left(\pi_0 + \ldots + \pi_s\right) \theta_2 I + \pi_S T = 0,$$

with the normalising condition

$$\sum_{i=0}^{S} \pi_i e = 1. \tag{4}$$

Adding the first three equations in (3), we obtain the following equality:

$$\sum_{i=1}^{S-1} \pi_i T - \sum_{i=0}^{s} \pi_i \theta_2 I - \sum_{i=1}^{S} \pi_i T^0 \beta = 0. \tag{5}$$

The equality given in (5) is first multiplied by the matrix T^{-1} and then by the column vector e_m on the right to obtain the following equation:

$$\sum_{i=1}^{S-1} \pi_i e_m - \sum_{i=0}^{s} \pi_i \theta_2 \left(-T\right)^{-1} e_m = \sum_{i=1}^{S} \pi_i T^0 \beta \left(-T\right)^{-1} e_m. \tag{6}$$

We obtain

$$\pi_S e_m = \sum_{i=0}^{s} \pi_i \theta_2 \left(-T\right)^{-1} e_m$$

by multiplying the last equality in (3) by the matrix T^{-1} and the vector e_m, respectively. Using the obtained equation and the service rate $\mu = \left[\beta \left(-T\right)^{-1} e\right]^{-1}$, the equation in (6) becomes as following:

$$\mu \sum_{i=1}^{S} \pi_i e_m = \sum_{i=1}^{S} \pi_i T^0. \tag{7}$$

Finally, by using the normalising condition in (4), we can rewrite the equality in (7) as follows:

$$\mu \left(1 - \pi_0 e_m\right) = \sum_{i=1}^{S} \pi_i T^0. \tag{8}$$

According to paper [22], the system is stable if and only if $\pi A e_m < \pi C e_m$. That is,

$$\lambda \left(1 - \pi_0 e_m\right) < \tau \pi_0 e_m + \sum_{i=1}^{S} \pi_i T^0. \tag{9}$$

From (8) and (9) we conclude that the following theorem is true.

Theorem 1. *The queuing-inventory system under study is stable if and only if the following condition is satisfied:*

$$\lambda < \mu + \frac{\tau \pi_0 e_m}{1 - \pi_0 e_m}.$$

By solving the system of balance equations in (3), (4), the probability vector π_0 can be obtained.

Let $x = \left(x_0,\ x_1,\ x_2,\ \ldots,\ x_r,\ x_{r+1},\ \ldots\right)$ be the vector of steady-state probabilities of the generator matrix Q in (1). That is, the vector x satisfies the following balance equations:

$$xQ = 0 \text{ and } xe = 1.$$

$(S+1)$-dimensional row vector x_0 is partitioned into components represented as

$$x_0 = \left[x_0(0),\ x_0(1),\ \ldots,\ x_0(S)\right].$$

Here $x_0(i)$ gives the probability that there are zero customers in the system and the inventory level is $i,\ 0 \le i \le S$. The row vector x_n of dimension $m(S+1)$ is partitioned into components represented as

$$x_n = \left[x_n(0),\ x_n(1),\ \ldots,\ x_n(S)\right],\ n \ge 1.$$

Here $x_n(i)$ of dimension m gives the steady-state probabilities that there are n customers in the system, the inventory level is i, $0 \le i \le S$, and the service process is in one of m phases.

The following theorem can be used to calculate the probabilities of the system (see [22]).

Theorem 2. *Under the stability condition given in theorem 1 the steady-state probability vector x is obtained by solving the following system of linear equations:*

$$x_0 B_0 + x_1 C_0 = 0,$$

$$x_0 A_0 + x_1 B_1 + x_2 C = 0,$$

$$x_{n-1} A + x_n B_1 + x_{n+1} C = 0,\ 2 \le n \le r-2,$$

$$x_{r-2} A + x_{r-1}\left[B_1 + RC\right] = 0$$

with

$$x_0 e + \sum_{n=1}^{r-2} x_0 e + x_{r-1}\left(I - R\right)^{-1} e = 1,$$

where the rate matrix R is the minimal non-negative solution to the matrix quadratic equation

$$R^2 C + R B_2 + A = 0.$$

The QBD structure of the generator given in (1), under the stability condition, yields a modified matrix-geometric solution. So, the probabilities of non-boundary states are given by

$$x_{n+r} = x_{r-1} R^{n+1},\ n \ge 0.$$

The performance measures

Below there are the key performance measures of the system:

- the expected loss rate of customers due to no inventory

$$ELR_1 = \lambda x_0(0) + \lambda \sum_{n=1}^{\infty} x_n(0) e;$$

- the expected loss rate of customers due to impatience

$$ELR_2 = \tau \sum_{n=1}^{\infty} x_n(0) e;$$

- the expected number of customers in the system

$$EN = \sum_{n=1}^{r-1} n x_n e + x_{r-1} R\left[r\left(I - R\right)^{-1} + R\left(I - R\right)^{-2} \right] e;$$

- the expected inventory level

$$EI = EI_1 + EI_2,$$

$$EI_1 = \sum_{n=0}^{r-1} \sum_{i=1}^{S} i x_n(i) e,\ \ EI_2 = \sum_{n=r}^{\infty} \sum_{i=1}^{S} i x_n(i) e;$$

- the expected order size

$$ES = ES_1 + ES_2,$$

$$ES_1 = (S-s) \sum_{n=0}^{r-1} \sum_{i=0}^{s} x_n(i) e,\ \ ES_2 = \sum_{n=r}^{\infty} \sum_{i=S-s}^{S} i x_n(S-i) e;$$

• the expected replenishment rate

$$ER = \theta_1 \sum_{n=0}^{r-1} \sum_{i=0}^{s} x_n(i)\, e + \theta_2 \sum_{n=r}^{\infty} \sum_{i=0}^{s} x_n(i)\, e.$$

Numerical study

Here we examine how some performance measures behave in different scenarios. For the service times the Erlang distribution (**ERLS**), exponential distribution (**EXPS**) and hyper-exponential distribution (**HEXS**) are considered. Each of the distributions has the same mean of 1, but qualitatively different. The values of the standard deviation of **ERLS**, **EXPS** and **HEXS** are 0.707 11, 1 and 2.244 72, respectively. The distributions are normalised at a specific value for the service rate μ.

ERLS:

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

EXPS:

$$\beta = (1),\ T = (-1).$$

HEXS:

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

The tables and figures in this section give the behaviour of some performance measures when the arrival rate λ changes and when the service rate μ changes. For this purpose, we fix some parameter values in all tables and figures. That is, the replenishment rate when the system operates under (s, Q) policy is $\theta_1 = 1$, the replenishment rate when the system operates under (s, S) policy is $\theta_2 = 1.5$, the impatient rate is $\tau = 2$, and the maximum inventory level is $S = 20$. In the examples given in tables and figures, when the service rate changes, the arrival rate is fixed as $\lambda = 2$ and when the arrival rate changes, the service rate is fixed as $\mu = 4$.

The way the expected number of customers in the system *EN* changes when there are different service time distributions by varying the threshold value as $r = 4, 8$ and the reorder point as $s = 3, 7$ is given in table 1 when the values of λ are increased and in table 2 when the values of μ are increased. When the arrival rate of customers increases, the traffic intensity increases, i. e. more customers are expected to be in the system. This observation is represented in table 1 for all scenarios. It is seen more clearly in the dense system that the coefficients of variation (CV) in service time distributions have an effect on *EN* values. The value of CV is 1 in **EXPS**, less than 1 in **ERLS** and greater than 1 in **HEXS**. Table 1 shows that when the value of CV increases, the values of *EN* also increase. This increase is less in the system with lower dense (i. e., see the case of $\lambda = 2.2$) and more in the system with higher dense (i. e., see the case of $\lambda = 3.8$).

Table 1

The expected number of customers in the system for arrival rate

s	λ	*r* = 4			*r* = 8		
		ERLS	**EXPS**	**HEXS**	**ERLS**	**EXPS**	**HEXS**
3	2.2	1.015	1.174	2.455	1.013	1.171	2.443
	2.6	1.471	1.752	3.990	1.462	1.738	3.958
	3.0	2.247	2.754	6.756	2.211	2.709	6.681
	3.4	3.908	4.921	12.785	3.791	4.791	12.629
	3.8	10.051	12.842	33.803	9.696	12.503	33.502
7	2.2	1.044	1.209	2.539	1.042	1.207	2.531
	2.6	1.529	1.824	4.180	1.522	1.815	4.156
	3.0	2.373	2.916	7.213	2.347	2.884	7.155
	3.4	4.271	5.400	14.189	4.186	5.307	14.064
	3.8	12.756	16.399	43.311	12.496	16.149	43.058

As the speed of customer service increases, the traffic intensity decreases, meaning that fewer customers are expected to be in the system. This behaviour of *EN* is given in table 2 for all scenarios. The effect of the CV on the values of *EN* is seen more clearly in the dense system. Table 2 shows that when the value of CV increases, the values of *EN* also increase. This increase is less in the system with lower traffic intensity (i. e., see the case of $\mu = 3.8$) and more in the system with higher traffic intensity (i. e., see the case of $\mu = 2.2$). In all scenarios, comparing with other service time distributions, it is clearly seen that the expected number of customers in the system is the largest in the **HEXS** case, especially in a dense system.

Table 2

The expected number of customers in the system for service rate

s	μ	$r = 4$			$r = 8$		
		ERLS	**EXPS**	**HEXS**	**ERLS**	**EXPS**	**HEXS**
3	2.2	6.969	8.831	23.143	6.815	8.678	22.991
	2.6	2.565	3.158	7.816	2.526	3.112	7.749
	3.0	1.603	1.918	4.419	1.591	1.903	4.386
	3.4	1.177	1.377	2.977	1.173	1.371	2.959
	3.8	0.935	1.075	2.197	0.933	1.072	2.188
7	2.2	7.630	9.814	26.827	7.583	9.763	26.754
	2.6	2.670	3.301	8.314	2.656	3.283	8.279
	3.0	1.654	1.983	4.614	1.648	1.976	4.595
	3.4	1.209	1.417	3.082	1.207	1.414	3.072
	3.8	0.958	1.102	2.264	0.957	1.101	2.258

In this model a customer loss is considered to occur in two different scenarios: (i) if there are no items in the inventory when a customer arrives, the customer does not enter the system (the customer is lost) – ELR_1; (ii) if the number of items in the inventory is zero when the customer in the service receives service and leaves, a customer in the queue shows impatience and leaves the queue at a certain rate (the customer is lost) – ELR_2. Changing on both of the expected loss rates of customers when the values of λ are increased and also when the values of μ are increased are given in table 3. It illustrates two service time distributions, the threshold value $r = 4$ and the reorder point $s = 3$.

Table 3

The expected loss rate of customers because of no inventory and impatient

λ	μ	ELR_1		ELR_2	
		ERLS	**HEXS**	**ERLS**	**HEXS**
1.4	–	0.021 6	0.024 9	0.004 5	0.006 5
2.2	–	0.081 1	0.088 0	0.018 9	0.025 2
3.0	–	0.168 2	0.175 8	0.045 8	0.058 0
3.8	–	0.220 8	0.251 1	0.082 6	0.104 4
–	2.2	0.035 6	0.049 8	0.021 7	0.036 5
–	3.0	0.055 9	0.063 9	0.018 6	0.026 8
–	3.8	0.061 8	0.068 3	0.014 9	0.020 5
–	4.6	0.064 2	0.069 8	0.012 2	0.016 1

The more customers come to the system, the more likely the system will have zero inventory. So, as expected, as the values of λ increase, both ELR_1 and ELR_2 values increase. As customers arriving at the same constant rate receive faster service (as μ increases), the probability of the number of items in the inventory being zero increases. Consequently, the case at which an incoming customer encounters zero items and therefore does

not join the system, in other words ELR_1, increases as μ increases. The situation where the inventory level is zero when a customer receives service is expected to increase as μ increases. However, since an increase in μ will also cause the decrease of the number of customers in the system, it is seen that the expected loss rate of customers due to impatience ELR_2 also decreases in table 3.

As the arrival rate increases (as the traffic intensity increases), the expected number of items in the inventory *EI* decreases until the system reaches a certain density and then increases (fig. 1). This behaviour is valid for both threshold values of *r*, but at large *r* values the change point is seen when the system is denser.

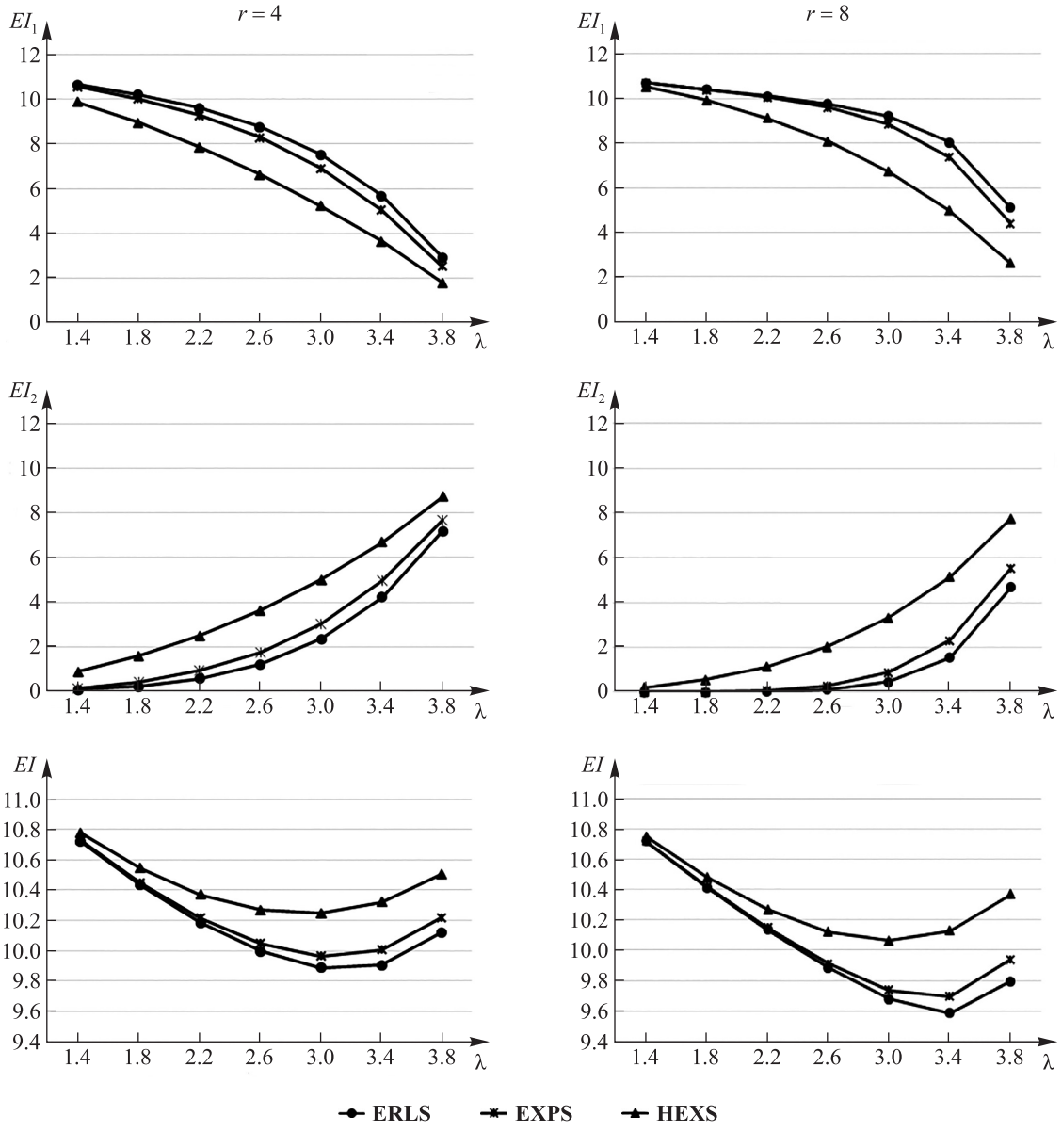


Fig. 1. The expected number of items in the inventory versus the arrival rate

(s, Q) policy is followed until the value of *EN* reaches *r*. In other words, the size of orders arriving in the inventory is constant up to the threshold value *r* (as much as $S - s = Q$). While the system operates in this policy the expected number of items in the inventory is represented by EI_1. (s, S) policy is followed when the value of *EN* becomes *r* and more. Namely, after the threshold value is reached, the number of items in the inventory will always be at the maximum level *S* with the orders coming. As the system operates in this policy the expected number of items in the inventory is represented by EI_2. As the arrival rate of customers increases, the expected number of customers in the system also increases. Therefore, fig. 1 shows that the increase in the arrival rate causes the values of EI_1 to decrease and the values of EI_2 to increase. In other words, when the traffic intensity

is low in the system, the system operates more in (s, Q) policy, and as the intensity increases, the frequency of operation in (s, S) policy increases. It is clearly seen from fig. 1 that the CV has an effect on the values of *EI*. In the cases where the variation in service times is 1 and less than 1 (in **ERLS** and **EXPS** cases), the values of *EI* are almost the same, especially at low arrival rates. Compared to **ERLS** and **EXPS** service times, the values of *EI* are higher in **HEXS** service time with high variation. When looking at EI_1 and EI_2 parts in fig. 1, the values of these performance measures in the case of **HEXS** are apart from the others.

As the service rate μ increases, the expected number of customers in the system decreases and so the value of *EI* decreases (fig. 2). The behaviours of EI_1 and EI_2 values are seen to be exactly the opposite of that in fig. 1.

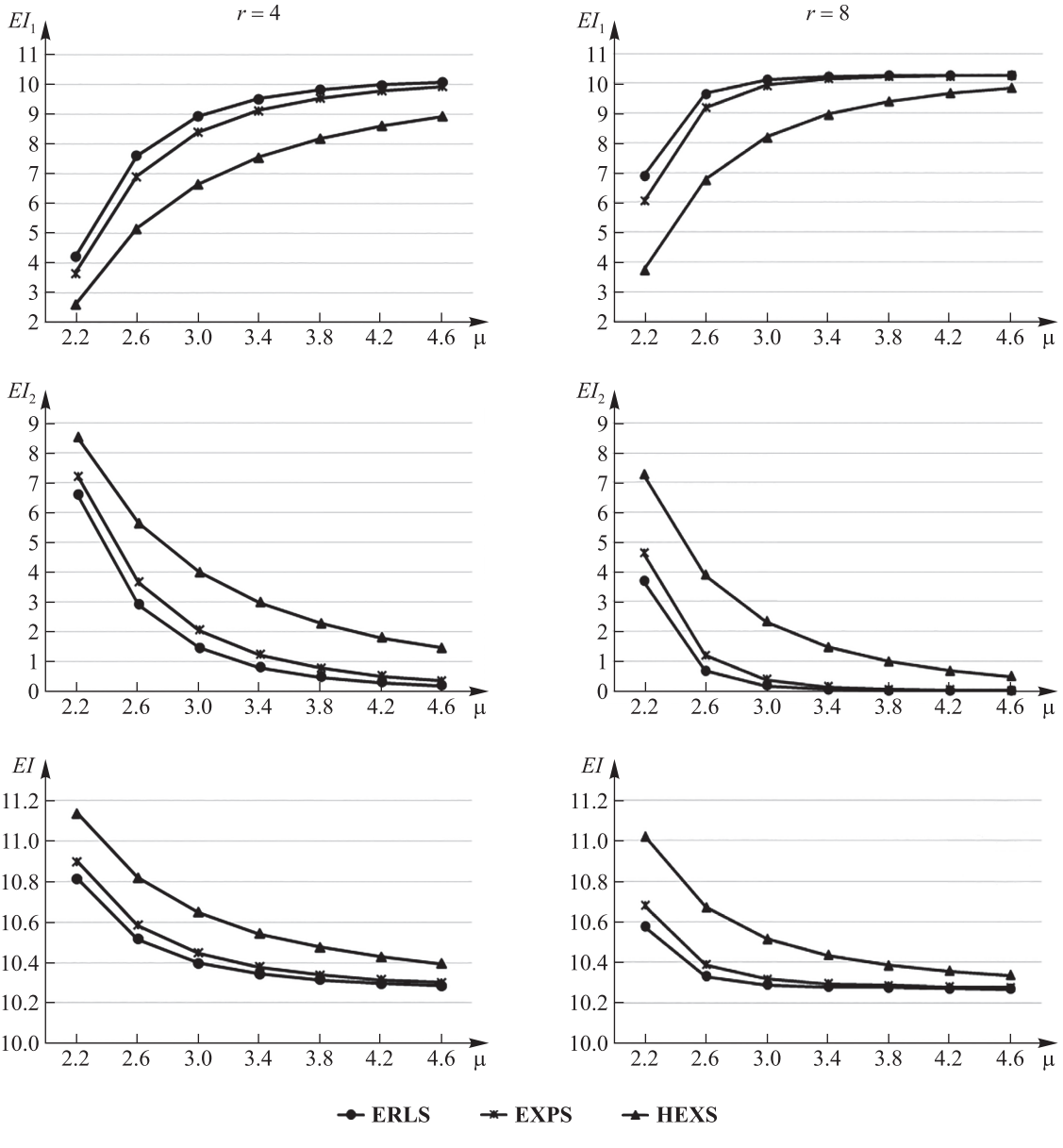


Fig. 2. The expected number of items in the inventory versus the service rate

As the number of customers entering the system increases, the expected order size *ES* is supposed to increase. This increase occurs rapidly until the system reaches a certain density, but after the system reaches a certain density, it remains almost the same (see $\lambda = 3.4,\ 3.8$). This result is observed in fig. 3. As previously commented, the inventory policy applied changes depending on the values of λ. If customers arrive with a low λ rate, the system is not of high density and (s, Q) policy is followed more. So, the values of ES_1 are expected to be larger. For example, when looking at the case of $\lambda = 1.4$, the values of ES_2 are almost zero while the values of ES_1 are around 1.5. On the other hand, if customers arrive with a high λ rate, the system is of high density and (s, S) policy is followed more. For example, when looking at the cases of $\lambda = 3$ and more, it is seen that the values of ES_1 decrease and the values of ES_2 increase.

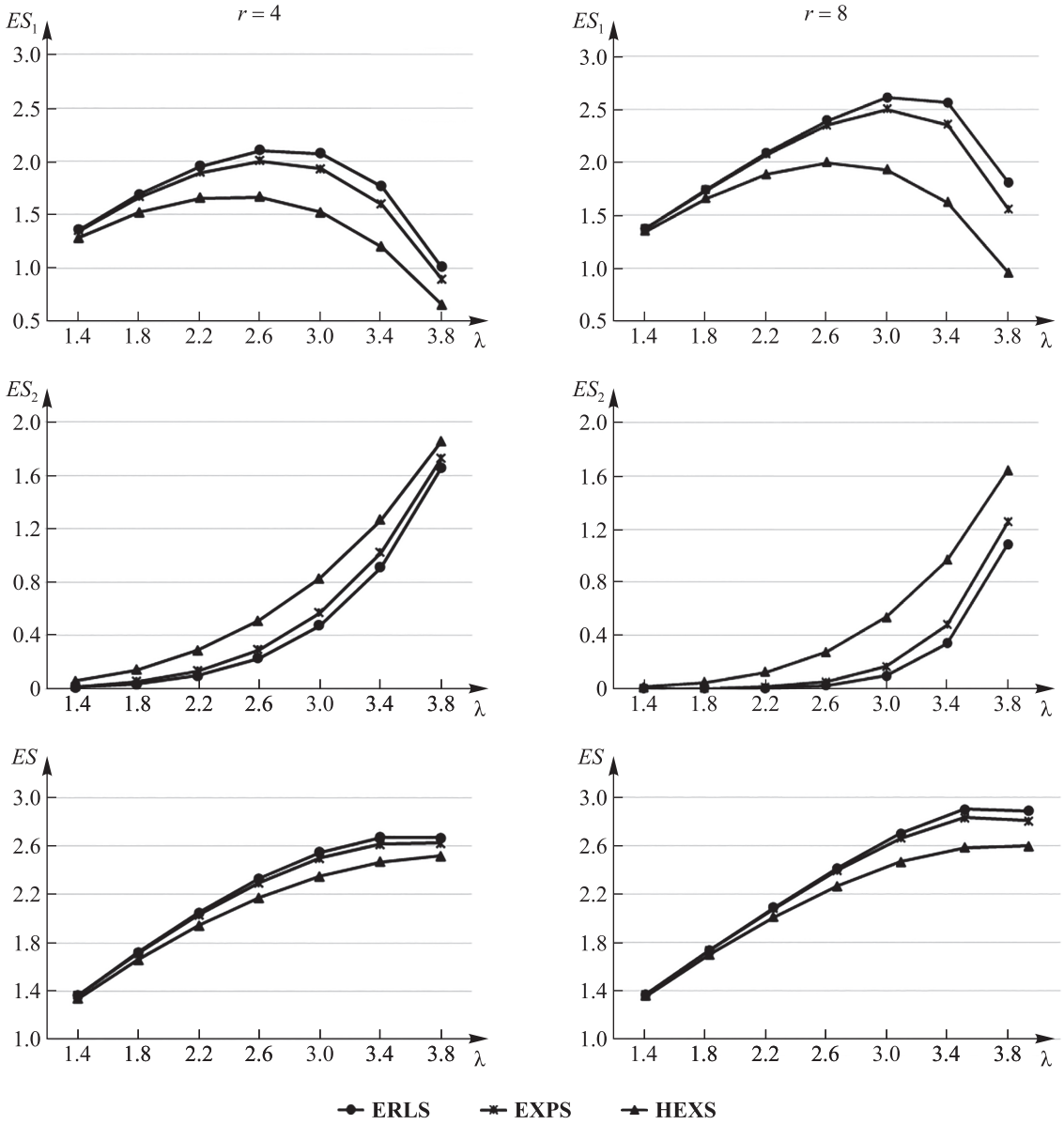


Fig. 3. The expected order size versus the arrival rate

Conclusions

A new queue-dependent hybrid replenishment policy is developed in the QIS model with PH-distribution of service time. It is based on the combined use of known replenishment policies (s, Q) and (s, S). When the inventory level drops to the reorder point s if the number of customers in the system is less than a threshold value r, then (s, Q) policy is applied, and if the number is more or equal to r, then (s, S) policy is applied. To calculate the performance measures of the proposed replenishment policy, a matrix-geometric approach is used. Numerical examples performed in different scenarios demonstrate the behaviour of the system performance measures. As directions for further research, we can point out the study of a similar QIS model with MAP arrivals and solving optimisation problems for such models.

References

1. Bijvank M, Vis IFA. Lost-sales inventory theory: a review. *European Journal of Operational Research*. 2011;215(1):1–13. DOI: 10.1016/j.ejor.2011.02.004.

2. Krishnamoorthy A, Manikandan R, Lakshmy B. A revisit to queueing-inventory system with positive service time. *Annals of Operations Research*. 2015;233:221–236. DOI: 10.1007/s10479-013-1437-x.

3. Krishnamoorthy A, Shajin D, Narayanan VC. Inventory with positive service time: a survey. In: Anisimov V, Limnios N, editors. *Queueing theory 2: advanced trends*. London: ISTE; 2021. p. 201–238 (Sciences. Mathematics: queuing theory and applications). Co-published by the John Wiley & Sons. DOI: 10.1002/9781119755234.ch6.

4. Salini K, Arya PS, Manikandan R. Queueing-inventory systems: a survey. arXiv:2308.06518v3 [Preprint]. 2023 [cited 2025 March 21]: [40 p.]. Available from: https://arxiv.org/abs/2308.06518v3.

5. Wang J, Zhang Y. Research progress on queueing-inventory systems. *Operations Research Transactions*. 2025;29(3):77–92. Chinese. DOI: 10.15960/j.cnki.issn.1007-6093.2025.03.004.

6. Chakravarthy SR, Rumyantsev A. Analytical and simulation studies of queueing-inventory models with MAP demands in batches and positive phase type services. *Simulation Modeling Practice and Theory*. 2020;103:102092. DOI: 10.1016/j.simpat.2020.102092.

7. Otten S, Daduna H. Stability of queueing-inventory systems with customers of different priorities. *Annals of Operations Research*. 2023;331(2):963–983. DOI: 10.1007/s10479-022-05140-1.

8. Shajin D, Jacob J, Krishnamoorthy A. On a queueing inventory problem with necessary and optional inventories. *Annals of Operations Research*. 2022;315(2):2089–2114. DOI: 10.1007/s10479-021-03975-8.

9. Rasmi K, Jacob MJ, Rumyantsev AS, Krishnamoorthy A. A multi-server heterogeneous queuing-inventory system with class-dependent inventory access. *Mathematics*. 2021;9(9):1037. DOI: 10.3390/math9091037.

10. Jeganathan K, Anzen Koffer V, Pavai Madheswari S, Venkatesan T. An investigation on retrial queueing-inventory system with server maintenance based on customer satisfaction. *International Journal of Applied and Computational Mathematics*. 2025;11(5):193. DOI: 10.1007/s40819-025-02023-4.

11. Jeganathan K, Anzen Koffer V, Lakshmanan K, Ammar SI. Performance analysis of synergistic servers in a multi-server queueing-inventory system with encouraged arrivals. *Journal of the Indian Society for Probability and Statistics*. 2026;27(1):159–202. DOI: 10.1007/s41096-025-00255-7.

12. Dudin AN. Admission control in a parcel delivery system with account of the finite capacity of the warehouse at the pick-up point. *Journal of the Belarusian State University. Mathematics and Informatics.* 2025;1:68–78. EDN: ZIAXUB.

13. Melikov A, Mirzayev R, Nair SS. Double sources queuing-inventory system with hybrid replenishment policy. *Mathematics*. 2022;10(14):2423. DOI: 10.3390/math10142423.

14. Melikov AZ, Mirzayev RR, Nair SS. Numerical study of a queuing-inventory system with two supply sources and destructive customers. *Journal of Computer and Systems Sciences International*. 2022;61(4):581–598. DOI: 10.1134/S1064230722030091.

15. Lawrence AS, Melikov A, Sivakumar B. Analysis and optimization of hybrid replenishment policy in a double-sources queueing-inventory system with MAP arrivals. *Annals of Operations Research*. 2023;331(2):1249–1267. DOI: 10.1007/s10479-023-05646-2.

16. Mirzayev RR. A model of a queuing system with a base stock policy and two supply sources. *Informatics and Control Problems*. 2022;42(1):26–32. DOI: 10.54381/icp.2022.1.04.

17. Mirzayev R. Optimal cost analysis in a double-source queuing-inventory system under hybrid replenishment policy. *Journal of Baku Engineering University. Mathematics and Computer Science*. 2025;9(1):19–25. DOI: 10.30546/09090.2025.210.019.

18. Ozkar S, Melikov A. Algorithmic approach to study queueing-inventory systems with queue-dependent hybrid replenishment policy. In: Dudin A, Nazarov A, Moiseev A, editors. *Information technologies and mathematical modelling: queueing theory and related fields. Revised selected papers of the 23rd International conference; 2024 October 20–26; Karshi, Uzbekistan*. Cham: Springer; 2025. p. 288–302 (Communications in computer and information science; volume 2472). DOI: 10.1007/978-3-031-88307-1_23.

19. Chakravarthy SR. *Introduction to matrix-analytic methods in queues 1: analytical and simulation approach – basics*. London: ISTE; 2022. XV, 341 p. (Mathematics and statistics series). Co-published by the John Wiley & Sons. DOI: 10.1002/9781394165421.

20. Chakravarthy SR. *Introduction to matrix-analytic methods in queues 2: analytical and simulation approach – queues and simulation*. London: ISTE; 2022. XV, 415 p. (Mathematics and statistics series). Co-published by the John Wiley & Sons. DOI: 10.1002/9781394174201.

21. Dudin AN, Klimenok VI, Vishnevsky VM. *The theory of queuing systems with correlated flows*. Cham: Springer; 2020. XXII, 410 p. DOI: 10.1007/978-3-030-32072-0.

22. Neuts MF. *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Baltimore: John Hopkins University Press; 1981. XIII, 332 p.

Received 12.01.2026 / revised 15.02.2026 / accepted 25.03.2026.