Queueing Models and
Service Management
Vol. 2, No. 1, page 083-111, 2019

QMSM
© PU 2018

Numerical Analysis and Long Run Total Cost Optimization of Perishable Queuing Inventory Systems with Delayed Feedback

Agassi Melikov[1,*], Achyutha Krishnamoorthy[2] and Mammad Shahmaliyev[3]

[1]Institute of Control Systems, National Academy of Sciences, B.Vahabzade street 9, Baku, Azerbaijan

[2]Department of Mathematics, Cochin University of Science and Technology, Cochin 682022, India

[3]Deparment of Computer Sciences, National Aviation Academy of Azerbaijan, Baku, Mardakan pr. 30, Azerbaijan

(Received June 2018 ; Accepted September 2018)

Abstract: Three-dimensional Perishable Queueing-Inventory System (PQIS) models with positive service time and delayed feedback are studied in this paper. We assume that the customers either leave the system with/without purchasing an item or join the orbit for the decision making. We apply the replenishment policy with the exponentially distributed positive order lead time. Approximate formulas are developed to calculate the joint distributions and performance measures of the system. High accuracy of the approximate formulas is illustrated by the numerical experiments. perishable queueing-inventory systems, positive service time, delayed feedback, order replenishment policy, finite and infinite 3D Markov Chains, calculation methods.

Keywords: Calculation methods, delayed feedback, finite and infinite 3D Markov chains, perishable queueing-inventory systems, positive service time, (s; S) order replenishment policy.

1. Introduction

The different models of Queuing-Inventory Systems (QIS) has been widely investigated due to its real-world applications in different sectors and industries. The detailed summary of Perisable and Non-Perishable QIS models is given in Karaesmen *et al.* [5] and Krishnamoorthy *et al.* [8].

Classical QIS models are based on the several fundamental assumptions. The first and important one is that after the customer service completion inventory level decreases. But in reality this condition does not always hold, because some customers may refuse to purchase the item after being served. The model with such type of customers was first studied in Krishnamoorthy *et al.* [9, 10]. Later the similar models were analyzed in Melikov and Shahmaliyev [13] and Melikov *et al.* [12] as well.

The second assumption in the studies of QIS models is the absence of feedback. In other words, served customers are not considered for the repeated service call. But in some

* Corresponding author
Email : agassi@science.az

systems the served customers may return to the system instantly (Instantaneous Feedback, IFB) or after some random period of time (Delayed Feedback, DFB) because of the qualitative service. Takacs [16, 17] was the first to study queueing models with unlimited inventory. The summary of the classical queuing models with feedback is provided in Koroliuk *et al.* [7] and Melikov *et al.* [11]. After analysis of the existing literature we found out that the QIS with Feedback (Queuing-Inventory Systems with Feedback, QISwFB) models were less studied and the only papers we could found are Amirthakodi *et al.* [1, 2]. Let's consider these papers in more detail below.

Single-channel QISwDFB with non-perishable inventory and finite queue of the primary customers (p -customers) with Poisson arrival was investigated in Amirthakodi *et al.* [2]. If the queue is full at the arrival of a p -customer then it leaves the system without being served. The p -customer according to Bernoulli trial either joins the orbit for future call or leaves the system after the service completion. The orbit has finite length and every customer after some exponentially distributed random time period recalls for the service independently. The system serves the repeated customers (r -customers) if there are no p -customers and/or inventory level is zero. The r -customer requires only service, that is after the service completion of r -customers the inventory level remains unchanged. Service times of both types of customers are exponentially distributed but with different parameters. The non-preemptive service policy is assumed, so that if the r -customer is being served at the moment of arrival of p -customer, the ongoing service is not interrupted. Repetitive orbit re-joining is considered as well, that is after the service completion of r -customer according to Bernoulli trial it either re-joins the orbit or leaves the system. The (s,S) inventory replenishment policy with positive exponential lead time is applied. The three dimensional Markov Chain (3D MC) is used to describe the mathematical model of the system. The algorithm based on the matrix algebra (see Neuts [14]) was developed for the calculation of steady-state distributions. Additionally, the formulas for the calculation of the average performance measures and the total cost were developed. Laplace-Stieltjes transform of waiting time was derived for both types of customers.

The QISwIFB model with perishable inventory (PQISwIFB) was studied in Amirthakodi *et al.* [1]. This paper investigates a single-channel PQISwIFB with finite queue of p -customers that forms the MAP flow. The inventory item lifetime is finite with exponentially distributed random time period. After the service completion the p -customer either instantly joins the second queue of infinite length for the repeated service or leaves the system according to Bernoulli trial. Then the next p -customer or the r -customer from the second queue is taken for the service. The r -customer after being served either instantly re-joins the second queue or leaves the system. System accepts the p -customers only if the inventory level is positive after the r -customer is served. Otherwise, if there are no p - customers and inventory level is zero the channel becomes idle for an exponentially distributed period of time. If during the idle period the p-customer arrives and inventory level becomes positive the channel starts to serve the customer. If after the idle period no p-customer arrives and the inventory is still empty, the channel begins to serve the r-customers. Likewise in [2] the r-customers requires only the service and the inventory level remains unchanged after service completion. The system uses hybrid replenishment policy, so that if inventory level drops to s then the order of size $S-s$ is placed. The order of size $S-i$ is placed when the inventory level is equal to $i, i <= s$ after the service completion of r-customer. The ordered items are received after a random time which is distributed as phase-type. The system is modeled by 6D-MC and the algorithm based on matrix algebra is developed to calculate the steady-state probabilities. Additionally, the formulas for the performance measures were derived and the total cost minimization problem was considered. It should be noted that the developed algorithm is very complex for the practical implementation and becomes less effective for the models of larger dimension.

The analysis of PQISwDFB models is motivated by its real-world applications in areas like food industry, chemicals, pharmaceuticals, blood bank management and many other related sectors where the perishability needs to be considered. The detailed summary of the PQIS's real-world applications is provided in Goyal and Giri [4] and Shah and Shah [15].

In our paper we present new single-channel PQISwDFB model. It is similar to the model studied in [2] but with the following differences:

- We study the model with perishable inventory.
- There are three options after the service completion for the customer:
  1. Customer leaves the system without purchasing an item.
  2. Customer purchases the item and leaves the system.
  3. Customer does not purchase the item and joins the orbit for "decision making".
- r-customer may purchase inventory item as well.
- Both finite and infinite queues of customers are considered.
- Customers in the queue become impatient when there are no items left in the inventory.

These differences improve the model's correspondence to the real systems. Additionally, we present the efficient method for the calculation of steady-state probabilities. Also we derive the formulas for the performance measures that contains tabulated functions.

The paper is organized as follows. First, we provide the general model description and introduce the problem statement. In the next section, we develop the mathematical model of the system using 3D MC, construct the corresponding transition matrix (Q-matrix) and derive the exact formulas for the system performance measures. Then we analyze the finite and infinite models with respect to the queue length and orbit size. Finally, we provide the numerical results with the illustrations and conclude the article.

2. Model Description and Problem Statement

The system continuously monitors the inventory so that every inventory item becomes unusable (perishes) after some finite exponentially distributed random time. Also, we assume that the reserved item while serving the customer cannot perish.

The p-customers arrive into the system according to Poisson scheme. For the simplicity, all the inventory items are considered identical and after the service completion the inventory level decreases by a single unit if the customer purchases the item.

If at the moment of the customer arrival there are items in the inventory and the channel is idle then the customer is taken to the service. When channel is busy the arrived customer joins the queue and waits for service. The customer either joins the queue according to Bernoulli trial or leaves the system if the inventory level is zero at the moment of arrival. The customers in the queue become impatient when the inventory level drops to zero and they independently leave the system after waiting some exponentially distributed period of time.

We consider the models both with finite and infinite queues. The customer is lost when the finite queue is full. When the queue is infinite all p-customers join the system.

There are three options after the service completion of the p-customer:

1. Customer leaves the system without purchasing an inventory item.
2. Customer purchases the item and leaves the system.
3. Customer does not purchase the item and joins the orbit for "decision making".

We assume that the customers in orbit do not have any information about the queue state or inventory level. After some random period of time every r-customer in orbit recalls for the service independently, while the system does not differentiate between p-customers and r-customers. Impatience rates and service times for both types of customers are the same. The served r-customer may re-join the orbit, that is the repetitive orbit joins are possible.

The r-customers in orbit are assumed to be insistent. If the queue is full or the inventory level is zero at the moment of arrival the r-customer returns to the orbit.

The service time depends on whether the customer purchases the item or not, but it has an exponential distribution with different parameters for each case. This assumption corresponds to reality because the service time needed for the customer that purchases the item is greater than for the one which does not.

For the simplicity, we use the 2-level inventory replenishment policy in our model where the order lead time is an exponentially distributed random variable with finite mean.

The problem is to find the steady-state distribution of the system, calculate the average performance measures. Also we derive the formulas for the performance measures and perform the cost analysis of the system.

3. Calculation Methods

We identify the system parameters as follows:

- S - the maximum inventory size
- s - the replenishment order threshold, $s < S/2$
- N - the maximum queue size for the model with finite queue
- R - the maximum orbit size for the model with finite orbit
- γ^{-1} - the average inventory item lifetime
- λ - the arrival rate of p-customers
- τ^{-1} - the average waiting time in queue when the inventory level is zero
- ϕ_1 - the probability of joining queue when the inventory level is zero
- ϕ_2 - the leaving probability when the inventory level is zero, $\phi_2 = 1 - \phi_1$
- σ_1 - the probability of leaving the system without purchasing an item after the service completion
- σ_2 - the probability of purchasing an item and leaving the system after the service completion
- σ_3 - the probability of joining the orbit for "decision making" without purchasing an item after the service completion
- μ_1^{-1} - the average service time of the customer that does not purchase the item after service completion
- μ_2^{-1} - the average service time of the customer that purchases the item
- ν^{-1} - the average lead time of the order
- η^{-1} - the average dwelling time in the orbit

Remark 1. *Later the term customer will refer to both types of customers (r and p customers), unless indicated explicitly.*

Based on the model description and parameters' definition, the process life cycle is visualized in Figure 1.

The model is described by 3D MC with the states (m,n,k), where m is inventory level, n is queue size and k is the orbit size. The state space (SS) of the model is defined as follows:

$$E = \bigcup_{k=0}^{R} E_k,\ E_k \bigcap E_{k'} = \varnothing,\ k \neq k'. \tag{3.1}$$

where $E_k = \{(m,n,k) : m = 0,1,\ldots,S;\ n = 0,1,\ldots,N\},\ k = 0,1,2,\ldots,R$.

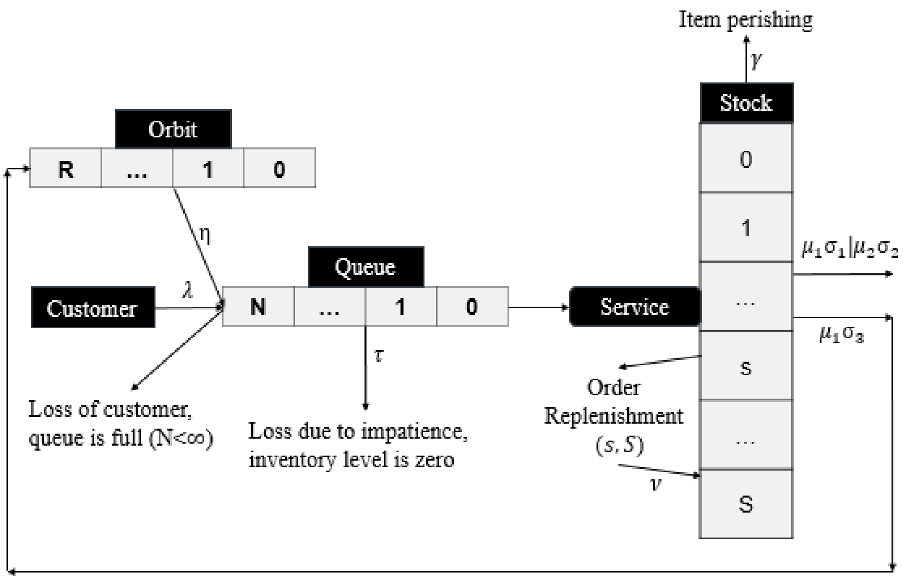


Figure 1. The structure of the investigated Markov model.

We conclude from (3.1) that SS is a set of points with integer coordinates inside the parallelepiped with height $R+1$ and rectangle base $(S+1)\times(N+1)$. The parallelepiped is illustrated in Figure 2 for the model with finite orbit and queue length.

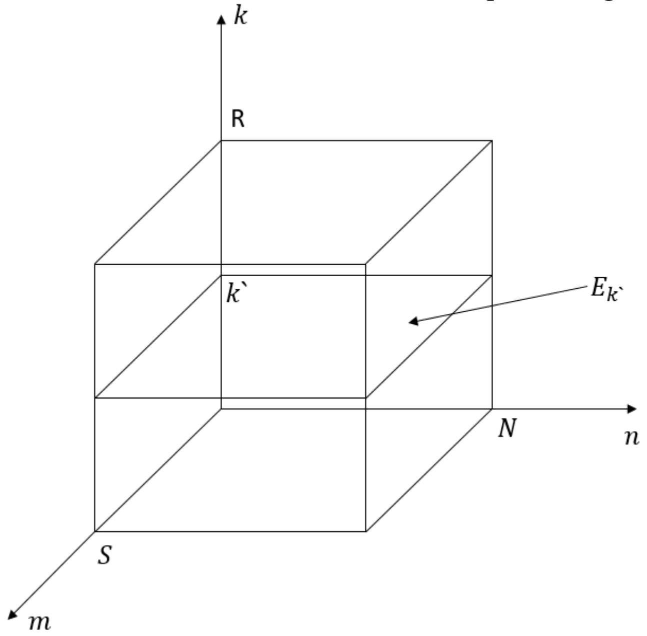


Figure 2. The state space of the 3D Markov model.

The transitions between the states inside the class E_k occur after the following events:

- arrival of p-customer
- inventory replenishment

• service completion
• inventory perishing
• leaving the system due to impatience

The transitions between the classes E_k and $E_{k'}$ are associated with the following events:

• joining the orbit
• r-customer arrival from the orbit

The transition rate from the state $(m_1,n_1,k_1)\in E_{k_1}$ to the state $(m_2,n_2,k_2)\in E_{k_2}$ is denoted by $q((m_1,n_1,k_1),(m_2,n_2,k_2))$. The set of all these rates forms the generator matrix (Q-matrix) of the 3D MC.

According to the accepted service scheme and inventory replenishment policy of the model, we get the following formulas for the transition rates inside the class E_k (see Algorithm 1):

$$q((m_1,n_1,k),(m_2,n_2,k))=\begin{cases}\lambda, & if\ m_2=m_1,\ n_2=n_1+1\\ \mu_1\sigma_1, & if\ m_2=m_1,\ n_2=n_1-1\\ \mu_2\sigma_2, & if\ m_2=m_1-1,\ n_2=n_1-1\\ m_1\gamma, & if\ m_2=m_1-1,\ n_2=n_1=0\\ (m_1-1)\gamma, & if\ m_2=m_1-1,\ n_2=n_1>0\\ \nu, & if\ m_2=m_1+S-s,\ n_2=n_1\\ 0, & otherwise\end{cases}. \tag{3.2}$$

when $m_1>0$,

$$q((0,n_1,k),(m_2,n_2,k))=\begin{cases}\lambda\phi_1, & if\ m_2=0,\ n_2=n_1+1\\ n_1\tau, & if\ m_2=0,\ n_2=n_1-1\\ S\nu, & if\ m_2=S-s,\ n_2=n_1\\ 0, & otherwise\end{cases}. \tag{3.3}$$

when $m_1=0$.

The state transition diagram inside the merged class E_k is illustrated in Figure 3. The rows and columns in Figure 3 describe the changes in inventory level and queue length correspondingly.

Algorithm 1. The calculation of Q-matrix element

function QELEM$(m_1, n_1, k_1, m_2, n_2, k_2)$ $\vartriangle q((m_1,n_1,k_1),(m_2,n_2,k_2))$
  define $q := 0$
  if $k_2 = k_1$ **and** $m_1 > 0$ **then**
    if $m_2 = m_1$ **and** $n_2 = n_1 + 1$ **then** $q := \lambda$
    else if $m_2 = m_1$ **and** $n_2 = n_1 - 1$ **then** $q := \mu_1\sigma_1$
    else if $m_2 = m_1 - 1$ **and** $n_2 = n_1 - 1$ **then** $q := \mu_2\sigma_2$
    else if $m_2 = m_1 - 1$ **and** $n_2 = n_1 = 0$ **then** $q := m_1\gamma$
    else if $m_2 = m_1 - 1$ **and** $n_2 = n_1 > 0$ **then** $q := (m_1 - 1)\gamma$
    else if $m_1 \le s$ **and** $m_2 = m_1 + S - s$ **and** $n_2 = n_1$ **then** $q := \nu$
  else if $k_2 = k_1$ **and** $m_1 = 0$ **then**
    if $m_2 = 0$ **and** $n_2 = n_1 + 1$ **then** $q := \lambda\phi_1$
    else if $m_2 = 0$ **and** $n_2 = n_1 - 1$ **then** $q := n_1\tau$
    else if $m_2 = S - s$ **and** $n_2 = n_1$ **then** $q := \nu$
  else if $k_2 \ne k_1$ **and** $m_2 = m_1 > 0$ **then**
    if $n_2 = n_1 - 1$ **and** $k_2 = k_1 + 1$ **then** $q := \mu_1\sigma_3$
    else if $n_2 = n_1 + 1$ **and** $k_2 = k_1 - 1$ **then** $q := k_1\eta$
  return q

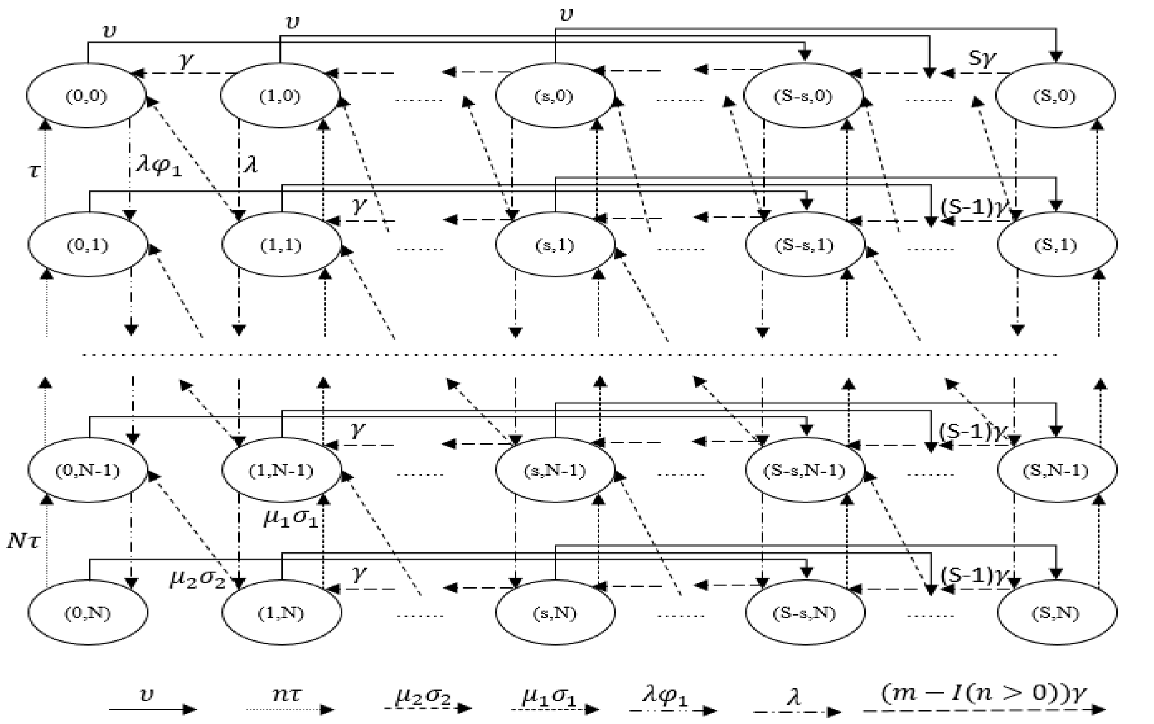


Figure 3. The state transition diagram within the class E_k .

The transition rates between the classes E_{k_1} and E_{k_2}, $k_1 \neq k_2$ are defined as follows ($m > 0$):

$$q((m,n_1,k_1),(m,n_2,k_2)) = \begin{cases} \mu_1\sigma_3, & n_2 = n_1 - 1,\ k_2 = k_1 + 1,\ k_1 < R \\ k_1\eta, & n_2 = n_1 + 1,\ k_2 = k_1 - 1,\ n_1 < N \\ 0, & otherwise \end{cases}. \quad (3.4)$$

The stationary probability of the state $(m,n,k) \in E$ is denoted as $p(m,n,k)$. We conclude from the formulas (3.2), (3.3), (3.4) that the Q-matrix of the model is irreducible, therefore there exists the stationary distribution.

The performance measures of the system are calculated via stationary distributions. We will derive the formulas for the following performance measures: S_{av} -average inventory level, Γ_{av} - average inventory perishing intensity, RR - average reorder rate, L_s - average queue length, L_o - average number of the r-customers in the orbit, RL - average customer loss intensity.

The average inventory level, average queue length and average orbit size are defined as the mathematical expectations of the corresponding random variables:

$$S_{av} = \sum_{(m,n,k)\in E} mp(m,n,k). \quad (3.5)$$

$$L_s = \sum_{(m,n,k)\in E} np(m,n,k). \quad (3.6)$$

$$L_o = \sum_{(m,n,k)\in E} kp(m,n,k). \quad (3.7)$$

The average perishing rate, assuming that the reserved item for the service cannot perish, is calculated as follows:

$$\Gamma_{av} = \gamma\left(\sum_{m=1}^{S} m \sum_{(m,0,k)\in E} p(m,0,k) + \sum_{m=2}^{S}(m-1)\sum_{(m,n,k)\in E} p(m,n,k)I(n>0)\right). \quad (3.8)$$

where $I(A)$ is the indicator function of A.

The replenishment order of the inventory is placed independently whenever the inventory level drops to the threshold s:

$$RR = \gamma(s+1)\sum_{(s+1,0,k)\in E} p(s+1,0,k) + (\mu_2\sigma_2 + s\gamma)\sum_{(s+1,n,k)\in E} p(s+1,n,k)I(n>0). \quad (3.9)$$

The customer loss intensity RL consists of three components:

1. the loss intensity of p-customers (RL_p).
2. the loss intensity because of orbit overflow (RL_o).
3. the loss intensity because of impatience of both types of customers (RL_s).

$$RL_p = \lambda \sum_{(m,N,k)\in E} p(m,N,k) + \lambda\phi_2 \sum_{(0,n,k)\in E} p(0,n,k) I(n < N). \quad (3.10)$$

$$RL_o = \mu_1\sigma_3 \sum_{(m,n,R)\in E} np(m,n,R) I(mn > 0). \quad (3.11)$$

$$RL_s = \tau \sum_{(0,n,k)\in E} np(0,n,k). \quad (3.12)$$

In order to calculate the above performance measures we need to obtain the steady-state probability distributions from the balance equations corresponding to the Q-matrix. These are the system of $(S+1)\times(N+1)\times(R+1)$ linear equations that cannot be solved numerically in a reasonable time for larger or infinite values of the parameters ($\mathcal{O}((S+1)(N+1)(R+1))$. Therefore, we apply the method of phase integration of the states of stochastic systems from Korolyuk and Korolyuk [6], where the hierarchical phase integration algorithm is proposed to calculate the stationary distribution of the three-dimensional MC when the certain asymptotic conditions are satisfied. Following the terminology therein, we call it the hierarchical space merging algorithm (SMA).

The idea behind SMA is to divide the original State Space into the $R+1$ parallel sub-planes where each plane denotes the merged state. Later the original stationary distribution is expressed approximately through the probability of the merged state according to the conditional probability formula. This process is repeated hierarchically until the dimension of the system is reduced to one. Therefore, the SMA is computationally efficient method as it diminishes the dimension of the system of equations. It should be noted that SMA produces the approximate solution and currently there are no any formulas for the estimation of the accuracy. The accuracy could be estimated experimentally and it produces highly accurate results when the transitions between the merged states are sufficiently small. The more detailed information about SMA could be found in [6].

For the correct application of the SMA method to our model the transition rates between the states of different classes E_k should be very small compared to the transitions inside the class. This assumption holds for the systems where the probability of joining the orbit is far smaller than the total probability of leaving the system: $\sigma_3 \ll \sigma_2 + \sigma_1$.

Assuming the above condition we will consider four models:

1. Both the queue length N and orbit size R are finite. We will provide the details of SMA and its application for this model, but provide only the final results for other cases.
2. The queue length N is finite and orbit size R is infinite.
3. Both queue length N and orbit size R are infinite.
4. The queue length N is infinite and orbit size R is finite.

3.1. Analysis of the model with finite queue length and orbit size

In this section we will consider the step by step application of SMA for the finite model, $N<\infty$ and $R<\infty$. In the first step of the hierarchy we construct the merge function $U_1(m,n,k)=\langle k\rangle$ based on (3.1), where the merged state $\langle k\rangle$ represents the set of all the states inside the class E_k. The set of the all merged states is denoted by $\Omega_1=\{\langle k\rangle : k=0,1,\dots,R\}$. Then we get the following approximate formula for the steady-state distributions:

$$\tilde{p}(m,n,k)\approx \rho^k(m,n)\pi_1(\langle k\rangle). \tag{3.13}$$

where $\rho^k(m,n)$ is the probability of the state (m,n) inside the class E_k and $\pi_1(\langle k\rangle)$ is the probability of the merged state $\langle k\rangle$, $\langle k\rangle\in\Omega_1$.

Further, based on (3.13) our problem is reduced to finding the probability distributions of the $R+1$ number 2D MC-s and a single 1D MC accordingly.

Now we re-apply SMA to the obtained 2D MC-s with state spaces $E_k, k=0,1,\dots,R$ in order to find the corresponding $\rho^k(m,n)$ probabilities. All the 2D MC-s are identical, therefore we will consider the model with fixed k :

$$E=\bigcup_{m=0}^{S}E_k^m,\ E_k^{m_1}\bigcap E_k^{m_2}=\varnothing,\ m_1\neq m_2. \tag{3.14}$$

where $E_k^m=\{(m,n,k)\in E_k:\ n=0,1,\dots,N\}, m=0,\dots,S.$ Similarly, we construct the merge function $U_2(m,n,k)=\langle m\rangle$ based on (3.14), where the merged state $\langle m\rangle$ represents the set of all the states inside the class E_k^m. The set of the all merged states is denoted by $\Omega_2=\{\langle m\rangle : m=0,1,\dots,S\}.$ Consequently, according to SMA:

$$p^k(m,n)\approx \rho_m^k(n)\pi_2^k(\langle m\rangle). \tag{3.15}$$

where $\rho_m^k(n)$ is the probability of the state (m,n) inside the class E_k^m and $\pi_2^k(\langle m\rangle)$ is the probability of the merged state $\langle m\rangle$, $\langle m\rangle\in\Omega_2$.

Further, we consider the problem of finding the probabilities $\rho_m^k(n)$ of the split models. We conclude from the formulas (3.2), (3.3), (3.4) that the transition rates between the states of the split model with state space E_k^m do not depend on the index k, therefore this index is omitted in $\rho_m^k(n)$ and $\pi_2^k(\langle m\rangle)$ onward. According to the formulas (3.2), the probability distributions inside the all split models with the state space E_k^m, $m=1,\dots,S$ are the same as in the classical model $M/M/1/N$ with load $a=\lambda/(\mu_1\sigma_1)$:

$$\rho_m(n)=a^n(1-a)/(1-a^{N+1}),\ m=1,\dots,S. \tag{3.16}$$

For the model with the infinite queue length we will assume that $a<1$ in order to ensure the system stability.

Similarly, we conclude from the formulas (3.3) that the probability distribution inside the split model with the state space E_k^0 are the same as in the Erlang model $M/M/N/N$ with load $b = \lambda\phi_1/\tau$:

$$\rho_0(n) = \frac{\theta(b,n)}{\sum_{j=0}^{N}\theta(b,j)},\ n = 0,1,\ldots,N. \tag{3.17}$$

where $\theta(i,j) = \frac{i^j}{j!}$.

After performing the mathematical transformations over the formulas (3.2), (3.3), (3.16), (3.17) we derive the following for the transition rates between the merged states $(\langle m_1\rangle),(\langle m_2\rangle) \in \Omega_2$ (see Figure 4):

$$q(\langle m_1\rangle,\langle m_2\rangle) = \begin{cases} \Lambda_1(m_1), & if\ m_2 = m_1 - 1 \\ \nu, & if\ m_1 <= s,\ m_2 = m_1 + S - s. \\ 0, & otherwise \end{cases} \tag{3.18}$$

where $\Lambda_1(m_1) = m_1\gamma\rho(0) + (1-\rho(0))(\mu_2\sigma_2 + (m_1 - 1)\gamma),\ m_1 = 1,2,\ldots,S.$

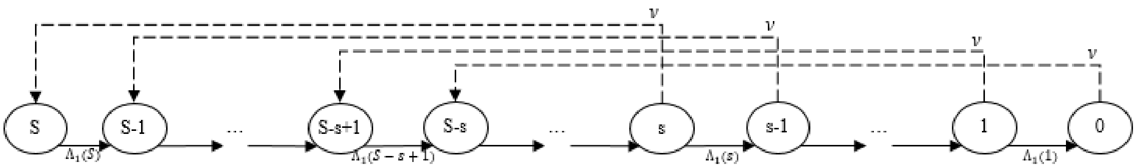


Figure 4. The transition diagram between the merged states E_k^m.

Further from (3.18) we derive (see [6]):

$$\pi_2(\langle m\rangle) = \begin{cases} \alpha_m\pi_2(\langle s+1\rangle), & if\ 0 \le m \le s \\ \beta_m\pi_2(\langle s+1\rangle), & if\ s+1 \le m \le S-s \\ \chi_m\pi_2(\langle s+1\rangle), & if\ S-s+1 \le m \le S \end{cases}. \tag{3.19}$$

where $\alpha_m = \prod_{i=m+1}^{s+1}\frac{\Lambda_1(i)}{\nu + \Lambda_1(i-1)}$, $\beta_m = \frac{\Lambda_1(s+1)}{\Lambda_1(m)}$, $\chi_m = \frac{\nu}{\Lambda_1(m)}\sum_{i=m-S+s}^{S}\alpha_i$, $\Lambda_1(0) = 0$.

The probability $\pi_2(\langle s+1\rangle)$ is found from the normalizing condition:

$$\pi_2(\langle s+1\rangle) = \left(\sum_{m=0}^{s}\alpha_m + \sum_{m=s+1}^{S-s}\beta_m + \sum_{m=S-s+1}^{S}\chi_m\right)^{-1}$$

Consequently, after applying some mathematical transformations we derive the following formula for the transition rates between the classes $\langle k_1\rangle,\langle k_2\rangle \in \Omega_1$:

$$q(\langle k_1\rangle,\langle k_2\rangle)=\begin{cases}\Lambda_2, & if\ k_2=k_1+1\\ k_1 M_2, & if\ k_2=k_1-1.\\ 0, & otherwise\end{cases} \quad (3.20)$$

where $\Lambda_2=\mu_1\sigma_3(1-\rho(0))(1-\pi_2(\langle 0\rangle))$, $M_2=\eta(1-\rho(N))(1-\pi_2(\langle 0\rangle))$.

We conclude from (3.20) that the probabilities of the merged states $\pi_1(\langle k\rangle), \langle k\rangle\in\Omega_1$ are the same as in the model $M/M/R/R$ with load $c=\Lambda_2/M_2$:

$$\pi_1(\langle k\rangle)=\frac{\theta(c,k)}{\sum_{j=0}^{R}\theta(c,j)},\ k=0,1,\ldots,R. \quad (3.21)$$

Finally, according to the formulas (3.13) and (3.15) the approximate steady-state probabilities of the initial 3D model is calculated as follows:

$$\tilde{p}(m,n,k)\approx\rho_m(n)\pi_2(\langle m\rangle)\pi_1(\langle k\rangle). \quad (3.22)$$

After substituting (3.22) in the formulas (3.5)-(3.12) we derive the following approximate formulas for the calculation of the system performance measures:

$$S_{av}\approx\sum_{m=1}^{S}m\pi_2(\langle m\rangle). \quad (3.23)$$

$$\begin{aligned}L_s&\approx\pi_2(\langle 0\rangle)\sum_{n=1}^{N}n\rho_0(n)+(1-\pi_2(\langle 0\rangle))\sum_{n=1}^{N}n\rho(n)\\&=b\pi_2(\langle 0\rangle)(1-E_B(b,N))+(1-\pi_2(\langle 0\rangle))\left(\frac{a}{1-a}-\frac{N+1}{1-a^{N+1}}a^{N+1}\right).\end{aligned} \quad (3.24)$$

$$L_o\approx c(1-E_B(c,R)). \quad (3.25)$$

$$\begin{aligned}\Gamma_{av}&\approx\gamma\sum_{m=1}^{S}\pi_2(\langle m\rangle)(m\rho(0)+(m-1)(1-\rho(0)))\\&=\gamma\sum_{m=1}^{S}\pi_2(\langle m\rangle)\left(m\frac{1-a}{1-a^{N+1}}+(m-1)\frac{a-a^{N+1}}{1-a^{N+1}}\right)\end{aligned}. \quad (3.26)$$

$$\begin{aligned}RR&\approx\pi_2(\langle s+1\rangle)[(s+1)\gamma\rho(0)+(s\gamma+\mu_2\sigma_2)(1-\rho(0))]\\&=\pi_2(\langle s+1\rangle)[(s+1)\gamma\frac{1-a}{1-a^{N+1}}+(s\gamma+\mu_2\sigma_2)\frac{a-a^{N+1}}{1-a^{N+1}}]\end{aligned}. \quad (3.27)$$

$$\begin{aligned}RL_p&\approx\lambda[\rho(N)(1-\pi_2(\langle 0\rangle))+\rho_0(N)\pi_2(\langle 0\rangle)+\phi_2(1-\rho_0(N))\pi_2(\langle 0\rangle)]\\&=\lambda[a^N\rho(N)(1-\pi_2(\langle 0\rangle))+\pi_2(\langle 0\rangle)(E_B(b,N)+\phi_2(1-E_B(b,N)))]\end{aligned}. \quad (3.28)$$

$$\begin{aligned} RL_o &\approx \mu_1\sigma_3\pi_1(\langle R\rangle)(1-\rho(0))(1-\pi_2(0)) \\ &= \mu_1\sigma_3 E_B(c,R)(1-\rho(0))(1-\pi_2(0)) \end{aligned}. \quad (3.29)$$

$$RL_s \approx \tau\pi_2(\langle 0\rangle)\sum_{n=1}^{N} n\rho_0(n) = b\tau\pi_2(\langle 0\rangle)(1-E_B(b,N)). \quad (3.30)$$

Remark 2. *$E_B(x,K)$ quantities are the Erlang B-formulas for the calculation of the customer loss probability for the model $M/M/K/K$ with the load x. We provide the formulas for the case $a \neq 1$, because when $a = 1$ the formulas become even simpler: $\rho(n) = 1/(N+1),\ n = 0,\ldots,N$.*

Remark 3. *We conclude from the formulas (3.15)-(3.22) that the stationary distributions depend on all the load parameters of the system. At the same time, according to the formulas (3.23)-(3.30) only L_o depends explicitly on the arrival intensity of the r-customers. The reason is that according to our assumption, the probability of joining the orbit is far smaller than the total probability of leaving the system, in other words, the arrival intensity of the r-customers are far smaller than of the p-customers. Additionally, the arrival intensity of the p-customers influences the population of r-customers in the orbit, consequently, all the performance measures depend on the arrival of r-customers implicitly.*

The presented methodology could be applied to PQISwDFB with the infinite queue and orbit size, $N = \infty$ and/or $R = \infty$. Below we skip intermediary mathematical transformations and present the resulting formulas for the steady state probabilities and the system performance measures for each case.

3.2. Analysis of the model with finite queue length and infinite orbit size

Now we consider the key points and differences for the case where $N < \infty$ and $R = \infty$:

- $\rho_m(n)$ and $\rho_0(n)$ are calculated by the formulas (3.16) and (3.17) accordingly.
- The probabilities of the merged states $\pi_1(\langle k\rangle), \langle k\rangle \in \Omega_1$ are the same as in the model $M(\Lambda_2)/M(M_2)/\infty$:

$$\pi_1(\langle k\rangle) \approx \frac{c^k}{k!}e^{-c}, k = 0,1,\ldots. \quad (3.31)$$

- Approximate formulas of the performance measures are calculated with the formulas (3.23)-(3.30), except RL_o and L_o. $RL_o = 0$ as the orbit size is infinite and loss probability due to orbit overflow is impossible. The average orbit size is calculated as follows:

$$L_o \approx c. \quad (3.32)$$

3.3. Analysis of the model with infinite queue length and infinite orbit size

In this section we consider the key points and differences for the case where $N=\infty$ and $R=\infty$:

- The probabilities of all states within the split models with the state space $E_k^m, m=1,\dots,S$ are the same as in the classical model $M/M/1/\infty$ with the load $a=\lambda/(\mu_1\sigma_1)$: $\rho_m(n)=(1-a)a^n$, $m=1,\dots,S$. We assume that the ergodicity condition $a<1$ holds true to ensure the system stability.
- The probabilities of all states within the split model with the state space E_k^0 are the same as in the Erlang model $M/M/\infty$ with the load $b=\lambda\phi_1/\tau$: $\rho_0(n)=\theta(b,n)e^{-b}$, $n=1,2,\dots$.
- The probabilities of the states of merged models are calculated by the formulas (3.19) and (3.31), where $\rho(0)=1-a$ and $\rho(N)=0$.
- The approximate values of S_{av} and L_o are calculated by the formulas (3.23) and (3.32) accordingly. The other performance measures are calculated as follows:

$$L_s \approx b\pi_2(\langle 0\rangle)+\frac{a}{1-a}(1-\pi_2(\langle 0\rangle)). \tag{3.33}$$

$$\Gamma_{av} \approx \gamma\sum_{m=1}^{S}\pi_2(\langle m\rangle)(m-a). \tag{3.34}$$

$$RR \approx \pi_2(\langle s+1\rangle)((s+1)\gamma(1-a)+(s\gamma+\mu_2\sigma_2)a). \tag{3.35}$$

$$RL_p \approx \lambda\phi_2\pi_2(\langle 0\rangle). \tag{3.36}$$

$$RL_s \approx \tau b\pi_2(\langle 0\rangle). \tag{3.37}$$

3.4. Analysis of the model with infinite queue length and finite orbit size

Finally, we analyze the case where $N=\infty$ and $R<\infty$:

- The state probabilities within the split models with the state space E_k^m, $m=1,\dots,S$ and E_k^0 are the same as in $N=\infty$ and $R=\infty$ model. We assume that, the ergodicity condition $a=\lambda/(\mu_1\sigma_1)<1$ holds true to ensure the system stability.
- The steady-state probabilities of the states of merged models are calculated by the formulas (3.19) and (3.21), where $\rho(0)=1-a$ and $\rho(N)=0$.
- The approximate values of S_{av} and L_o are calculated by the formulas (3.23) and (3.25) accordingly.

The other performance measures are calculated using the formulas (3.33)-(3.37), except RL_o:

$$RL_o \approx \mu_1\sigma_3 E_B(c,R)(1-\rho(0))(1-\pi_2(0)).$$

4. Numerical Results

Now we provide the results of the numerical experiments. First, we demonstrate the accuracy of the SMA for the finite model using comparison tables with exact solution. Additionally, dependence graphs for the performance measures are presented. Next, the performance measures' dependence graphs of the partly-infinite (infinite queue or orbit size) models are compared and explained. Lastly, the optimization problem is solved for the full infinite (both queue and orbit size are infinite) model with respect to the reorder level and different replenishment services.

4.1. Analysis of the model with finite queue length and orbit size

First, we estimate the accuracy of the SMA algorithm for the model with finite queue and orbit. We will provide comparison of steady-state probabilities and performance measures. The accuracy will be estimated using the following norms:

- Cosine similarity: $N_1 = \dfrac{\sum_{(m,n,k)\in E} p(m,n,k)\tilde{p}(m,n,k)}{\sqrt{\sum_{(m,n,k)\in E}(p(m,n,k))^2}\sqrt{\sum_{(m,n,k)\in E}(\tilde{p}(m,n,k))^2}}$.

- Maximum absolute difference: $N_2 = \max_{(m,n,k)\in E} |p(m,n,k) - \tilde{p}(m,n,k)|$.

- Root mean square deviation (RMSE):

$N_3 = \left[\dfrac{1}{|E|}\sum_{(m,n,k)\in E}(p(m,n,k) - \tilde{p}(m,n,k))^2\right]^{\frac{1}{2}}$, where $|E|$ is the cardinality of the state space E.

The exact values of steady-state probabilities are calculated from the linear system of balance equations corresponding to the Q-matrix. The system parameters for numerical experiments are accepted as follows:

$$\mu_1 = 55, \mu_2 = 5, \sigma_1 = 0.3, \sigma_2 = 0.5, \phi_1 = 0.3, \nu = 1, \tau = 1.5.$$

The comparison results of the steady-state probabilities and performance measures are given in Tables 1 and 2,3 correspondingly. We conclude from these tables that the accuracy of the approximate approach is very accurate.

Table 1. Estimation of the accuracy of the steady-state probabilities versus various norms.

(S,N)	(s,R)	(λ,η)	Norms		
			$\|\|N\|\|_1$	$\|\|N\|\|_2$	$\|\|N\|\|_3$
(10,10)	(1,2)	(55,5)	0.98964	0.01834	0.00201
	(2,3)	(60,10)	0.98955	0.02042	0.00200
	(4,4)	(65,15)	0.98989	0.01731	0.00194
(10,15)	(1,2)	(55,5)	0.98373	0.01826	0.00154
	(2,3)	(60,10)	0.98456	0.02037	0.00149
	(4,4)	(65,15)	0.98595	0.01726	0.00141
(15,10)	(2,2)	(55,5)	0.95934	0.01823	0.00173
	(5,3)	(60,10)	0.96858	0.02034	0.00154
	(7,4)	(65,15)	0.97482	0.01721	0.00138
(15,15)	(2,2)	(55,5)	0.98686	0.01312	0.00164
	(5,3)	(60,10)	0.98900	0.01207	0.00148
	(7,4)	(65,15)	0.98996	0.01194	0.00141
(20,5)	(2,2)	(55,5)	0.98019	0.01306	0.00124
	(5,3)	(60,10)	0.98423	0.01203	0.00111
	(9,4)	(65,15)	0.98629	0.01190	0.00103
(20,10)	(2,2)	(55,5)	0.95863	0.01303	0.00129
	(5,3)	(60,10)	0.97068	0.01308	0.00110
	(9,4)	(65,15)	0.97645	0.01297	0.00099

Table 2. Estimation of the accuracy of the performance measures. EV-exact value, AV-approximate value.

(S,N)	(s,R,λ,η)	S_{av}		RR		Γ_{av}	
		EV	AV	EV	AV	EV	AV
(10,10)	(1,2,55,5)	2.257	2.641	0.536	0.509	3.279	4.010
	(2,3,60,10)	2.329	2.850	0.622	0.577	3.385	4.375
	(4,4,65,15)	2.236	2.823	0.798	0.722	3.196	4.300
(10,15)	(1,2,55,5)	2.257	2.641	0.536	0.509	3.279	4.010
	(2,3,60,10)	2.329	2.850	0.622	0.577	3.385	4.375
	(4,4,65,15)	2.236	2.823	0.798	0.722	3.196	4.300
(15,10)	(2,2,55,5)	3.396	4.165	0.549	0.511	5.434	6.926
	(5,3,60,10)	3.459	4.397	0.727	0.654	5.520	7.335
	(7,4,65,15)	3.265	4.181	0.860	0.766	5.146	6.912
(15,15)	(2,2,55,5)	3.396	4.165	0.549	0.511	5.434	6.926
	(5,3,60,10)	3.459	4.397	0.727	0.654	5.520	7.335
	(7,4,65,15)	3.265	4.181	0.860	0.766	5.146	6.912
(20,5)	(2,2,55,5)	4.387	5.390	0.507	0.473	7.371	9.336
	(5,3,60,10)	4.661	5.937	0.646	0.583	7.859	10.359
	(9,4,65,15)	4.425	5.689	0.838	0.743	7.389	9.856
(20,10)	(2,2,55,5)	4.387	5.390	0.507	0.473	7.366	9.333
	(5,3,60,10)	4.660	5.937	0.646	0.584	7.855	10.357
	(9,4,65,15)	4.425	5.689	0.838	0.743	7.387	9.854

Table 3. Estimation of the accuracy of the performance measures. EV-exact value, AV-approximate value

(S,N)	(s,R,λ,η)	L_S		L_O		RL	
		EV	AV	EV	AV	EV	AV
(10,10)	(1,2,55,5)	8.993	9.053	0.524	0.583	43.979	43.716
	(2,3,60,10)	9.133	9.202	0.317	0.301	47.957	46.855
	(4,4,65,15)	9.228	9.301	0.230	0.186	52.891	51.684
(10,15)	(1,2,55,5)	13.18	13.03	0.506	0.583	43.918	41.681
	(2,3,60,10)	13.45	13.39	0.307	0.301	47.934	44.876
	(4,4,65,15)	13.63	13.62	0.224	0.186	52.876	49.696
(15,10)	(2,2,55,5)	9.071	9.146	0.524	0.583	43.872	40.477
	(5,3,60,10)	9.203	9.284	0.316	0.301	47.932	43.564
	(7,4,65,15)	9.281	9.359	0.230	0.186	52.876	48.258
(15,15)	(2,2,55,5)	13.39	13.31	0.510	0.583	42.865	43.239
	(5,3,60,10)	13.64	13.63	0.309	0.301	46.770	46.213
	(7,4,65,15)	13.77	13.79	0.225	0.186	51.871	51.174
(20,5)	(2,2,55,5)	4.450	4.535	0.532	0.582	42.806	41.568
	(5,3,60,10)	4.514	4.590	0.320	0.301	46.743	44.620
	(9,4,65,15)	4.562	4.630	0.232	0.186	51.854	49.503
(20,10)	(2,2,55,5)	9.101	9.177	0.524	0.583	42.766	40.580
	(5,3,60,10)	9.240	9.320	0.316	0.301	46.741	43.566
	(9,4,65,15)	9.320	9.399	0.230	0.186	51.854	48.297

Now we consider the dependence of the performance measures on the reorder level for the following system parameters.

$$\lambda = 60, \eta = 10, \mu_1 = 55, \mu_2 = 15, \sigma_1 = 0.2, \sigma_2 = 0.5, \phi_1 = 0.3,$$
$$\nu = 2, \gamma = 2, \tau = 1.5, S = 20, N = 10, R = 1.$$

We conclude from Figure 5 that the average inventory level S_{av} and perishability rate Γ_{av} increases proportionally with the increase of reorder level s but becomes stable for the larger values. Additionally, the intuitive relation $\Gamma_{av} \approx S_{av} * \gamma$ holds true. As expected, reorder rate RR increases with the increase of reorder level s, because the inventory level reaches the threshold value more frequently for the larger values of s. Average queue size L_s as depicted in Figure 6, increases very slowly with the increase of s and remains almost full due to the high arrival intensity. Average orbit size L_o on the other hand, does not depend nor on s neither on maximum queue length N.

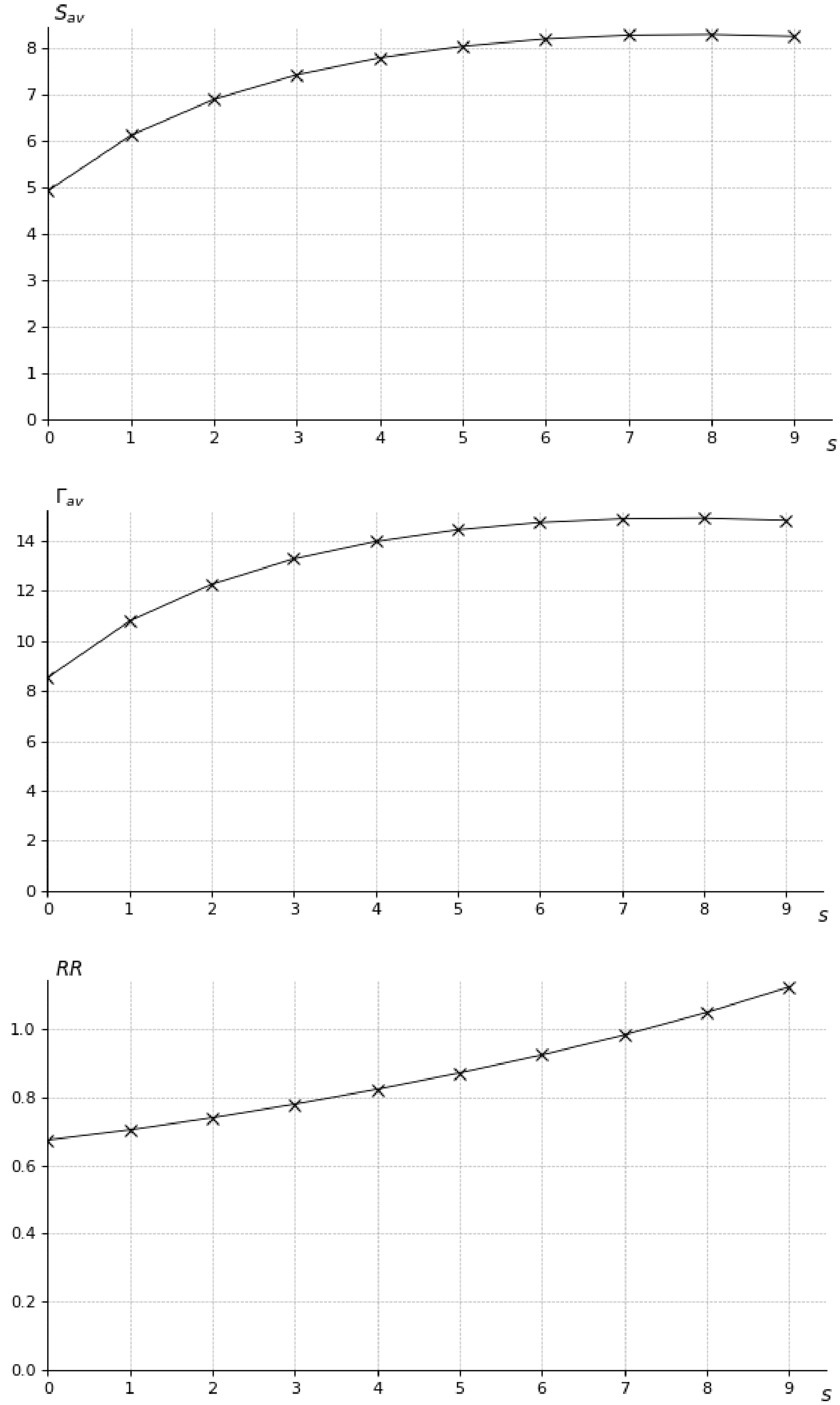


Figure 5. Dependence of S_{av}, Γ_{av} and RR on the reorder level s.

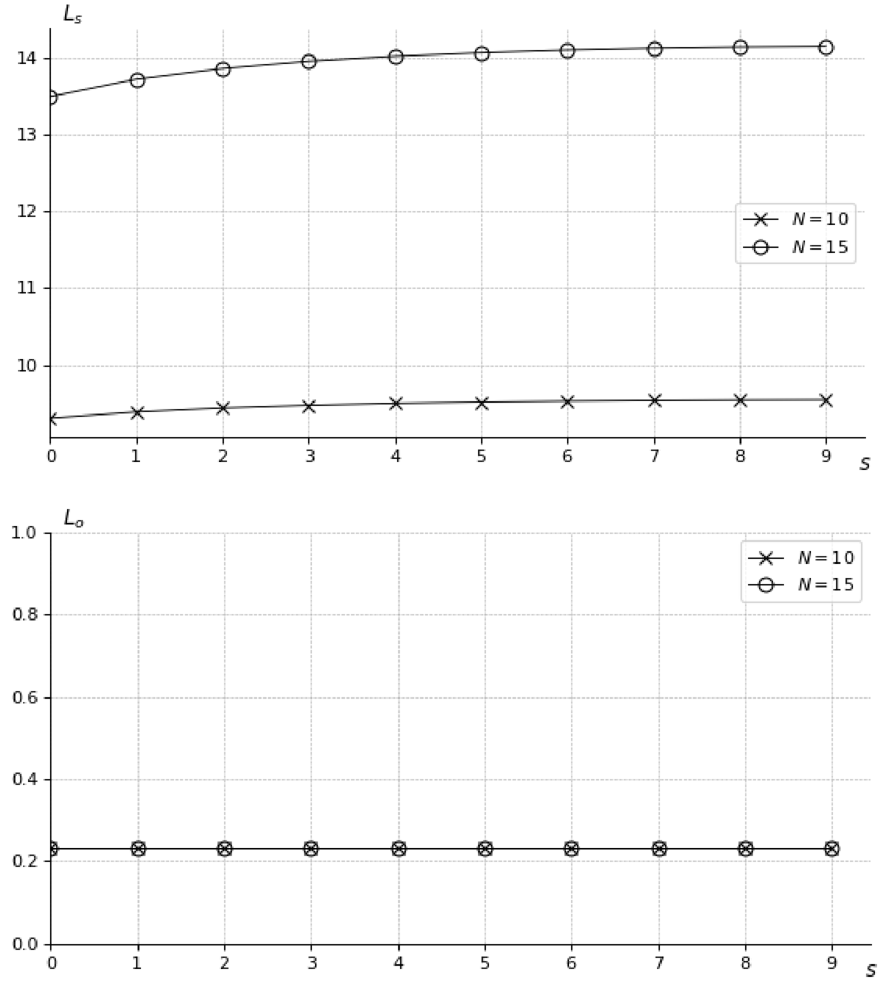


Figure 6. Dependence of L_s and L_o on the reorder level s.

Finally, we conclude from Figure 7 that loss rates due to queue and orbit overflow RL_p and RL_o increase very slowly and almost do not depend on s, while the loss rate due to impatience decreases with the increase of the reorder level s. The increase of RL_p is explained by the growth of the average queue length, while the decrease of RL_s is subject to the growth of the average inventory level, because with the higher S_{av} zero inventory level becomes less probable. Loss intensity due to orbit overflow RL_o is directly proportional to the average queue length S_{av} and L_{av} so that it increases with the growth of s.

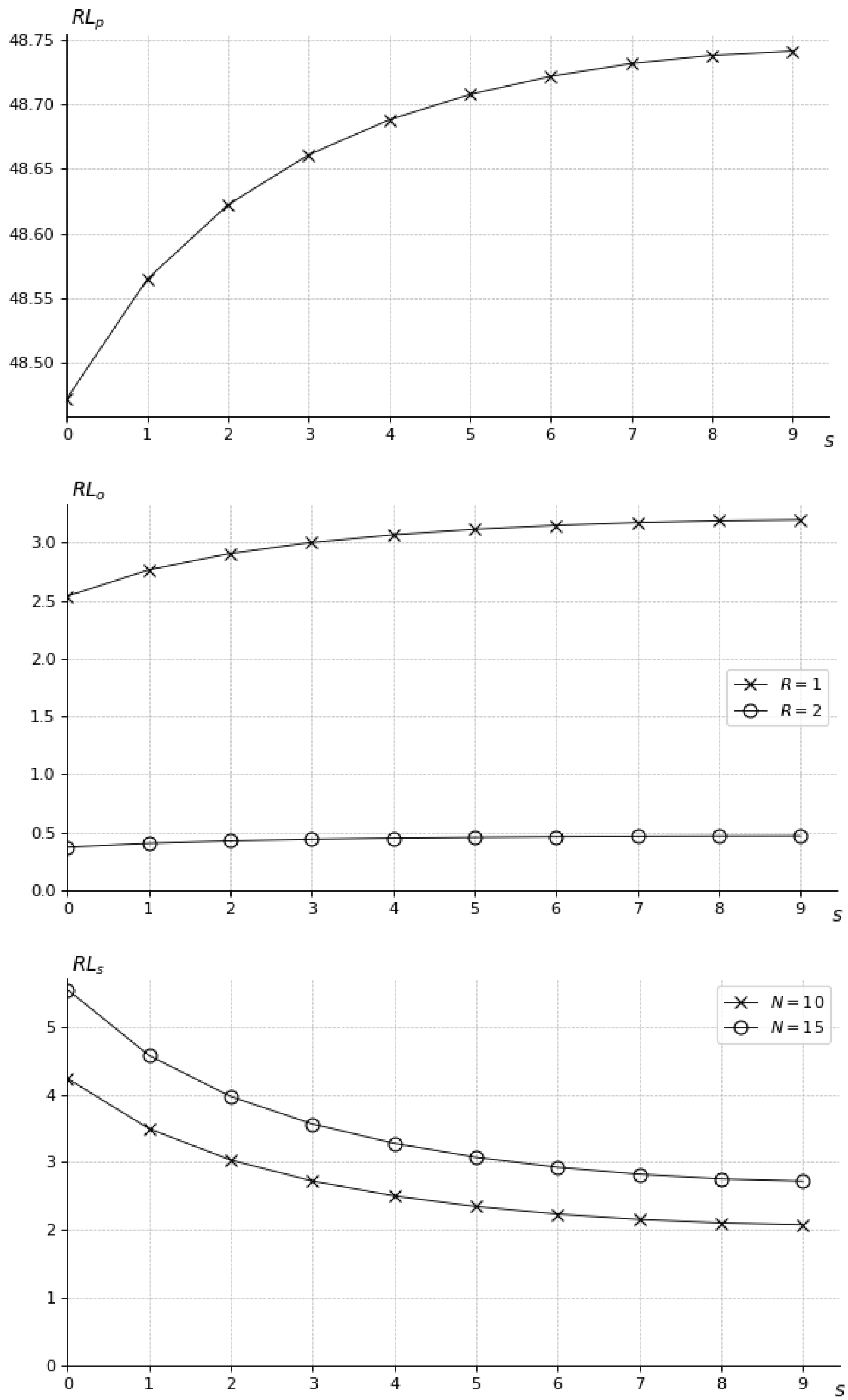


Figure 7. Dependence of RL_p , RL_o and RL_s on the reorder level s.

4.2. Analysis of the models with infinite queue length or orbit size

In this section, we compare the performance measures of the following two models:

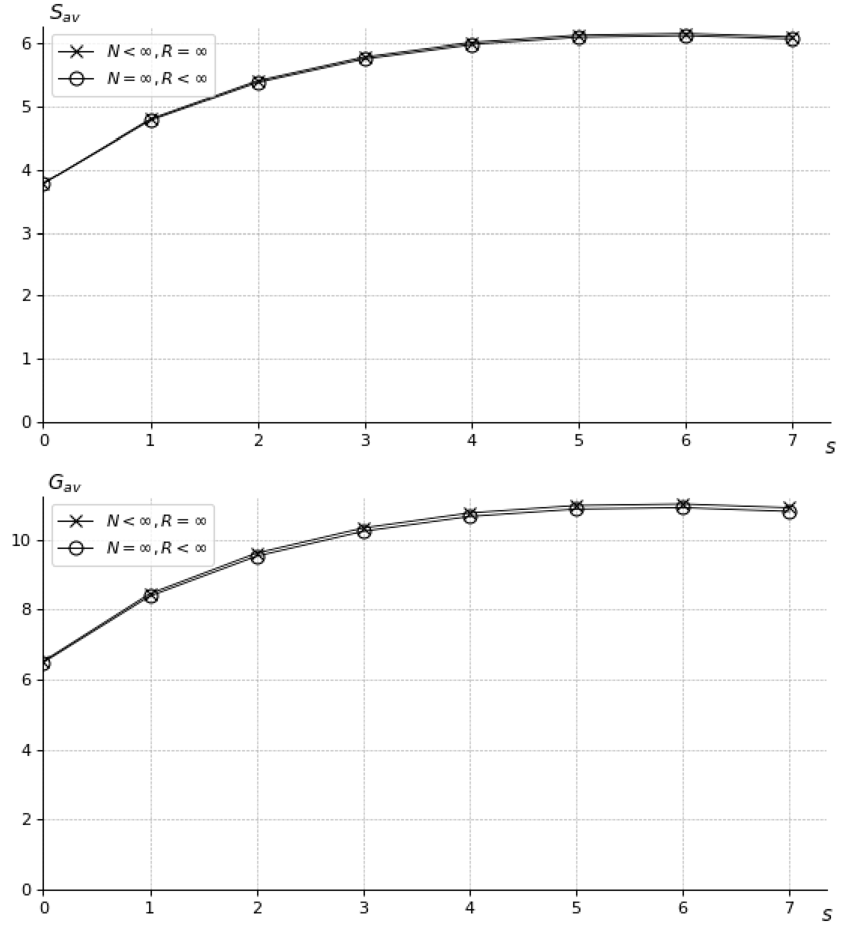
1. The model with finite queue L and infinite orbit R.
2. The model with finite queue L and finite orbit R.

We use the following parameter values in our calculations:

$$\lambda = 10, \eta = 5, \mu_1 = 60, \mu_2 = 15, \sigma_1 = 0.2, \sigma_2 = 0.5, \phi_1 = 0.3,$$
$$\nu = 2, \gamma = 2, \tau = 1.5, S = 15.$$

Additionally, we assume $N = 10, R = \infty$ for the former model and $N = \infty, R = 2$ for the latter one.

First, we consider the inventory related performance measures. We conclude from Figure 8 that S_{av}, G_{av} and RR are almost the same for both models. Such behavior is explained by the fact that these performance measures depend on the state of the inventory and are not related to the queue and orbit length.



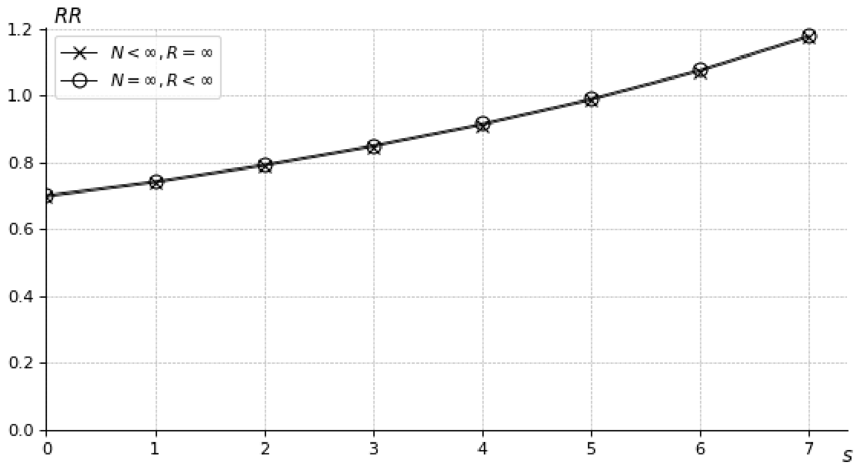
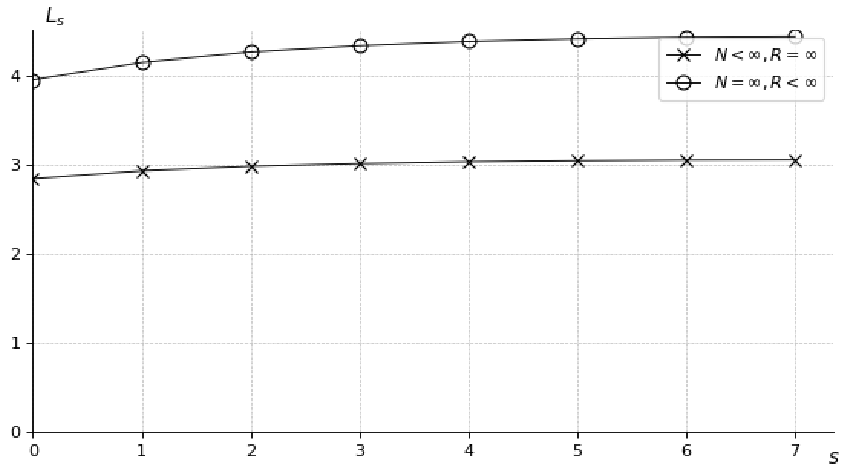
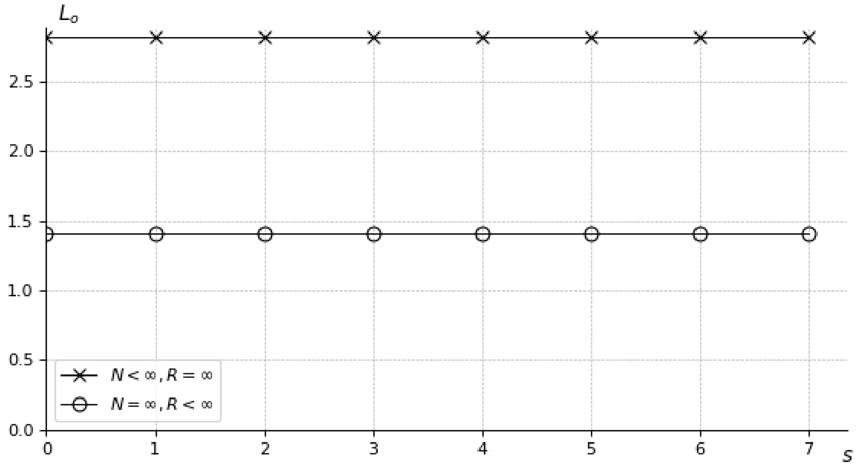


Figure 8. Comparison of the inventory related performance measure with respect to s for partly infinite models

Secondly, we analyze the customer related performance measures. As illustrated in Figure 9 average queue length L_s and average orbit length L_o are higher in the corresponding models with N and R being infinite. Average loss rate is higher in the model with finite orbit because of $RL_o > 0$ as opposed to the model with infinite orbit size where $RL_o = 0$.

Finally, we conclude from the comparisons that inventory related performance measures S_{av}, Γ_{av} and RR are the same for both partly infinite models, but the customer related performance measures differ while preserving their change behavior.





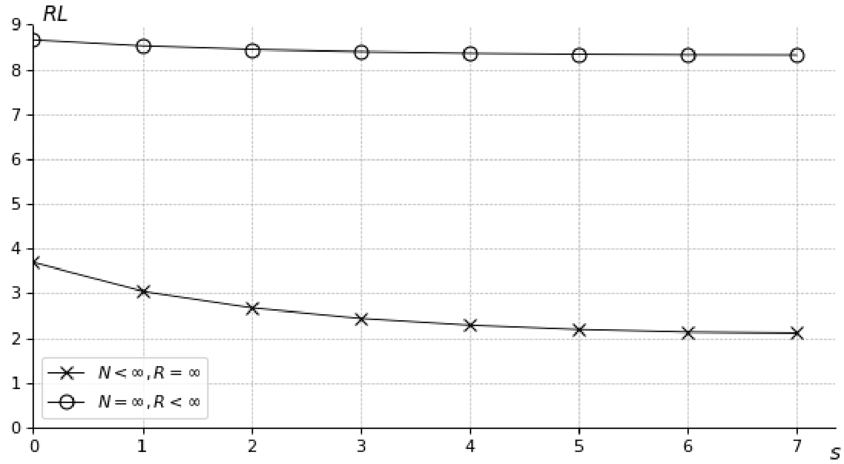


Figure 9. Comparison of the inventory related performance measure with respect to s for partly finite models

4.3. Optimization problem for model with infinite queue and infinite orbit size

In this section we consider the long run total cost optimization problem for the model with $N=\infty$ and $R=\infty$ and apply SMA method to solve it. We will use the stochastic simulation to prove the accuracy of the results.

Let's consider the overall expenses of the system. We have the maintenance expenses per inventory item, per customer in the queue and orbit. Additionally, extra penalty is charged for the every lost customer. There are expenses for the delivery services that deliver the inventory replenishment orders. The delivery service charges for the shipping and per item carriage.

The problem is to minimize the total cost by choosing the optimal delivery service d and the reorder threshold s. We introduce following long run total cost function TC:

$$TC(s,d)=(K+c_r(S-s))RR+c_sS_{av}+c_p\Gamma_{av}+c_lRL+c_{ws}L_s+c_{wo}L_o. \quad (4.1)$$

The parameters in (4.1) have following meanings:

- d - Designates the delivery service. The number of delivery services to choose from is $D, d : d = 0,\ldots,D$. The delivery service d is defined by the triple $(\nu, K, c_r)_d$.
- K - predefined fixed shipping cost of the chosen delivery service.
- c_r - per item carriage cost of the delivery service.
- c_s - maintenance cost per inventory unit.
- c_p - per unit perishing cost.
- c_l - penalty per customer loss.
- c_{ws} - queue maintenance cost per customer.
- c_{wo} - orbit maintenance cost per customer.

We have the following delivery services ($D = 8$):

$$\{(0.5,1,0.5),(0.5,1,1),(0.5,2,0.5),(0.5,2,1),(1,1,0.5),(1,1,1),(1,2,0.5),(1,2,1)\}$$

The list of the services is sorted by the lead intensity ν and its cost.

The system parameters and the cost function constants are chosen as follows:

$$\lambda = 5, \eta = 10, \mu_1 = 50, \mu_2 = 30, \sigma_1 = 0.2, \sigma_2 = 0.7, \phi_1 = 0.4, \gamma = 2, \tau = 0.5, S = 15,$$

$$c_s = 0.5, c_p = 0.5, c_l = 0.5, c_{ws} = 0.2, c_{wo} = 0.1.$$

We use the brute force method to find the optimal (s,d) pair that minimizes the cost function TC. The SMA algorithm is used to calculate the system performance measures. To ensure the system stability we assume that the ergodicity condition is hold: $\lambda / (\mu_1 \sigma_1) < 1$. The result of experiment is depicted in Figure 10. TC gets its minimal value at $s = 0, d = 7$ and $TC(0,7) = 6.56$.

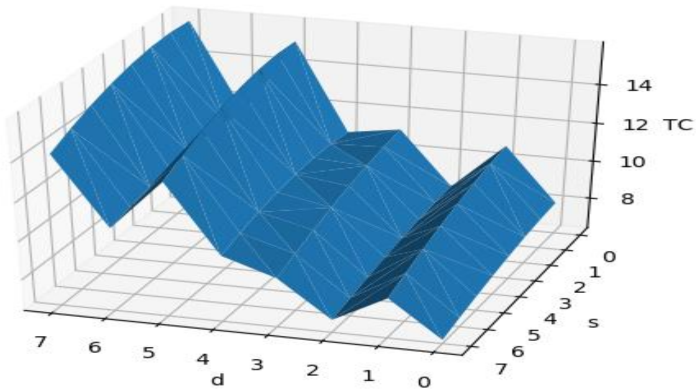


Figure 10. Long run total cost TC dependence on the reorder level s and delivery services d.

Although the optimal pair depends on the chosen parameters, in general, zero reorder level s lowers the average inventory level S_{av} and minimizes the inventory cost and the reorder rate RR, while the higher delivery rate decreases the probability of the zero inventory and prevents the customer loss that in turn decreases the customer loss penalty cost.

To prove the accuracy of the SMA we have run the same experiment, but applied Gillespie's Direct algorithm for the stochastic simulation (see Gillespie [3]) to calculate the system performance measures. As expected, the optimal pair was the same $s=0, d=7$, while there was the minor difference in the value of the cost function $TC(0,7)=5.85$.

In conclusion, we have solved the cost optimization problem for the system using the SMA and demonstrated the dependence of TC on the delivery service and reorder threshold selection. We checked the accuracy of the results obtained by SMA using Gillespie's stochastic simulation algorithm.

4.4. Pros and Cons

The main disadvantage of SMA is that it produces approximate results, while its accuracy is very high as confirmed by the numerical experiments. Additionally, due to its low complexity and computational efficiency SMA could be used in the calculation of the system performance measures, as well as, in optimization problems for the models with the higher parameter values and dimension.

4.5. Discussions

Let's consider the behavior of the model based on the numerical results above.

First, we consider the inventory related performance measures. When we increase the reorder level s the replenishment orders are made more frequently that results in the growth of the reorder intensity RR. If the inventory is replenished frequently the probability of the zero inventory decreases and the average inventory level S_{av} goes up. Therefore, the average perishing rate Γ_{av} increases as well according to the formula $\Gamma_{av} \approx \gamma * S_{av}$. This logic is illustrated in the Figures 5 and 8.

Secondly, we discuss the customer related performance measures. We observe the slight increase of the average queue length L_{av} for the higher values of the reorder level s in Figures 6 and 9. This is explained with the increase of the average inventory level S_{av}, because the higher S_{av} means the lower probability of zero inventory and less customer loss due to impatience RL_s. The average orbit length L_o is not affected by the reorder level s as it mostly depends on the parameter σ_3 as concluded from the formula (3.20).

5. Conclusion

The finite and infinite 3D PQIS models with positive service time and delayed feedback are studied in this paper. It is assumed that the customers either leave the system with/without purchasing an item or join the orbit for the decision making. When the inventory level is zero, customers join the system according to Bernoulli trial, while customers in the queue become impatient. The inventory replenishment policy belongs to (s,S) class.

The exact and approximate formulas are given for the calculation of the steady-state probabilities and performance measures of the system. Exact method is based on the solving of balance equations and is suitable only for the finite models. The approximate approach is based on the State Merging Algorithm (SMA) of Markov Chains and is applicable for both finite and infinite systems. The high accuracy of the given formulas is demonstrated using numerical experiments and the corresponding comparison tables are provided. The limitations and advantages of the SMA algorithm are shortly explained.

The dependence of performance measures on the reorder level s is presented for the finite model. The performance measures of the partly-infinite ($N=\infty$ or $R=\infty$) models are compared and described using graphical illustrations.

Finally, optimization problem is solved for the fully infinite model ($N=\infty$ and $R=\infty$). The dependence of long run total cost function TC on the reorder level s and different delivery services d is shown using 3D graph.

5.1. Future work

We still have some work left on the SMA algorithm for the future papers. First, we are working on the accuracy estimation of the SMA and we would like to develop the analytic formulas for this purpose. Currently, we estimate the SMA using experiments or compare with the results of the available methods and algorithms. Secondly, we think on the modification of the SMA to return the exact results. Thirdly, we apply the SMA on the different systems and perform experiments to find out the models the SMA works the best. We investigate the models with variable incoming intensity rate, MMPP (Markov modified Poissen process) that is the generalization of the studied model. We will present the results of our researches in the next papers.

References

[1] Amirthakodi, M., Radhamami, V., & Sivakumar, B. (2015). A perishable inventory system with service facility and feedback customers. *Annals of Operations Research*, 233, 25-55.

[2] Amirthakodi, M., & Sivakumar, B. (2015). An inventory system with service facility and finite orbit for feedback customers. *OPSEARCH*, 52, 225-255.

[3] Gillespie, D. T. (1976). A General method for numerically simulating the stochastic time evolution of of coupled chemical reactions. J*ournal of Computational Physics*, 22, 403-434.

[4] Goyal, S. K., & Giri, B. C. (2001). Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*, 134, 1-16.

[5] Karaesmen, I., Scheller-Wolf, A., & Deniz, B. (2011). Managing perishable and aging inventories: Review and future research directions. Planning production and inventories in the extended enterprise. *A state of the art handbook.* (Eds. Kempf K., Keskinocak P, Uzsoy P.), 1, 393-438.

[6] Korolyuk, V. S., & Korolyuk, V. V. (1999). Stochastic Models of Systems. Kluwer, Boston.

[7] Koroliuk, V. S., Melikov, A. Z., Ponomarenko, L. A., & Rustamov, A. M. (2016). Methods for analysis of multi-channel queuing models with instantaneous and delayed feedbacks. *Cybernetics and System Analysis*, 52, 58-70.

[8] Krishnamoorthy, A., Lakshmy, B., & Manikandan, R. (2011). A survey on inventory models with positive service time. *OPSEARCH*, 48, 153-169.

[9] Krishnamoorthy, A., Manikandan, R., & Lakshmy, B. (2015). Revisit to queuing-inventory system with positive service time. *Annals of Operations Research*, 233, 221-236.

[10] Krishnamoorthy, A., Manikandan, R., & Shajin, D. (2015). Analysis of a multi-server queuing-inventory system. *Advances in Operations Research*, Volume 2015, Article ID 747328, 16 pages.

[11] Melikov, A. Z., Ponomarenko, L. A., & Rustamov, A. M. (2015). Methods for analysis of queuing models with instantaneous and delayed feedbacks. *Communications in Computer and Information Sciences*, 564, 185-199.

[12] Melikov, A. Z., Ponomarenko, L. A., & Shahmaliyev, M. O. (2017). Analysis of perishable queuing-inventory systems with different types of requests. *Journal of Automation and Information Sciences*, 49, 42-60.

[13] Melikov, A. Z., & Shahmaliyev, M. O. (2017). A perishable queuing-inventory system with positive service time and (S-1, S) replenishment policy. *Communications in Computer and Information Sciences*, 800, 83-96.

[14] Neuts, M. F. (1981). Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach. The John Hopkins University Press, Baltimore, MD.

[15] Shah, N. H., & Shah, Y. K. (2000). Literature survey on inventory models for deteriorating items, *Ekonomski Anali*, 44, 221-237.

[16] Takacs, L. (1963). A single-server queue with feedback. *Bell System Technical Journal*, 42, 505-519.

[17] Takacs, L. (1977). A queuing model with feedback. *Operations Research*, 11, 345-354.