Advances in Industrial Engineering, Autumn 2021, 55(4): 367-401
DOI: 10.22059/aie.2021.332961.1812
RESEARCH PAPER

An Integrated Markovian Queueing-Inventory Model in a Single Retailer- Single Supplier Problem with Imperfect Quality and Destructive Testing Acceptance Sampling

Amir Aghsami *, **Yaser Samimi, Abdollah Aghaie**

School of Industrial Engineering, K. N. Toosi University of Technology (KNTU), Tehran, Iran.

Received: 26 October 2021, Revised: 01 December 2021, Accepted: 03 December 2021

Abstract

This paper proposes a retailer-supplier queueing-inventory problem (RSQIP) in which the imperfect lots are investigated using a single sampling inspection plan. We integrate an M/M/m response queueing system for handling and responding to customers' demands with a classical retailer-supplier inventory model considering defective items and inspection process for the first time. Customers whose demand is met leave the retailer system with exactly one item unless the inventory shortage occurs. The retailer places an order once the inventory level reaches an economic reorder point. The lead time is assumed exponential, and due to the imperfect incoming items, the retailer conducts a destructive acceptance sampling plan. The rate of inspection depends on the sample size. Provided that a lot is rejected, the supplier is required to provide a defect-free shipment. We present the stationary distribution of the number of demands in the response system. Then the joint stationary distribution of the order status and inventory level of the retailer are derived. Several performance measures and the expected total cost are presented steady-state, and a non-linear integer programming model is proposed to minimize the expected total cost. The results are numerically illustrated and reveal the convexity of the expected total cost. The optimal reorder point, order quantity, and the number of servers is computed for some numerical examples. A comprehensive sensitivity analysis is conducted to examine the effect of defective items, response system, and some important parameters on the entire developed model. Finally, useful managerial insights are presented.

Keywords:
M/M/m Queueing-Inventory System;
Retailer-Supplier Inventory Model;
RSQIP;
Defective Items;
Destructive Testing Acceptance Sampling;
Stationary Distribution

Introduction and literature review

In the classical inventory models, the incoming demands are satisfied immediately provided that there is enough inventory in the system while it may take some time to meet the demands or deliver items to the customers [1] . For example, the retailer may need some time to give the customers information about a purchased item to learn how to utilize it [2] Also, he may prepare, pack and load items and then deliver them to customers [1] . From another perspective, the retailer's arrival demands consult with sales experts or inquire them about the availability of items, price, brand, dimension, material, etc., before buying. Afterward, if the requested item is available, they can receive the one. Therefore, in this case, after processing and handling the incoming demands, each of them is met if the inventory level of the retailer isn't zero, otherwise

* Corresponding author: (A. Aghsami)
Email: a.aghsami@ut.ac.ir

lost sales occur. In both cases, a queue of customers or demands is formed so that the queueing-inventory models can be applied to study such systems. However, to the best of our knowledge, the latter case never has been discussed in the related literature.

The queueing-inventory models have been considered by many researchers worldwide in recent decades. These models aim to determine the optimal values of the inventory system's parameters so that the inventory costs, i.e., holding cost, ordering cost, etc., are minimized, and the quality of customer service is improved. On the one hand, the performance measures of the queues of customers (demands), i.e., the average waiting time in the system, the average number of customers in the queue, etc., are affected by the inventory systems characteristics. On the other hand, these performance measures would affect the inventory systems parameters [3] . These factors have motivated researchers to simultaneously study inventory and queueing systems to measure the two systems' impact on each other. Moreover, the close relation between queueing-inventory and the integrated supply chain models has caused researchers to pay more attention to this field [4] . The supply chain's goal is to move products and materials from suppliers to customers [5] , in which they buy items from the retailers. Besides inventory management can enhance the performance and responsiveness of the supply chain [7] , it is one of the practical tools that led to increasing the supply chain's profit; for example, it specifies when and how much should be ordered by the retailers to satisfy demands [8] . Provided the queues of demands are considered in an inventory model, the retailer can obtain the optimal inventory parameters based on the formed queues. As a result, he can manage the inventory system as comprehensively as possible making more profit for the supply chain. Therefore, the queueing-inventory models are more practical, and in addition to the characteristics of inventory models, they also consider the quality of customer service.

Gavish and Graves [9] studied an M/G/1 queueing system in an inventory–production system in which the arriving demands are customers, and the production facility is server. Sigman and Simchi-Levi [10] were pioneers in integrating the service time in the classical inventory models and presented a queueing-inventory model for the mobile server location problem. Berman et al. [11] proposed a queueing system in a service facility to satisfy Poisson-distributed demands and obtain optimal inventory policies. Berman and Sapna [12] derived the stationary distribution of inventory level and queue length in a service facility and developed a cost model for specifying the optimal order quantity. A limited and unlimited M/M/1 queue with inventory and lost sales are modeled by Schwarz et al. [3] . They examined the (r, S) and (r, Q) policies and obtained the joint probabilities of queue length and on-hand inventory level to derive a cost function. Schwarz and Daduna [13] studied an M/M/1 system with inventory regarding (r, Q) policy and backorder. A production system with Poisson demands and two types of items is developed as a queueing-inventory model by Chang and Lu [14] . Saffari et al. [1] extended a queueing-inventory system with multiple suppliers that each one has an exponential lead time. They computed the optimal reorder point and order quantity utilizing some performance measures of the model. Zhao and Lian [15] examined a queueing-inventory model with two classes of customers and different priorities. Krishnamoorthy and Viswanath [16] derived joint stationary distribution of queue length of customers, inventory level, and production status for an inventory-production problem under (s, S) policy. Saffari et al. [17] developed an attached M/M/1 queueing-inventory model with a supplier replenishing items under (r, Q) policy. They assumed the customers are lost when the on-hand inventory isn't available in the system. They extended the long-run average cost concerning r and Q after deriving the item level's limited probability and the number of customers in the system. Sivashankari and Panayappan [18] presented a production inventory system with defective items incorporating a multi-delivery policy. Baek and Moon [19] developed the prior paper to replenish items via an internal production besides an exterior supplier. Also, the lead time was assumed to have an exponential distribution. Krishnamoorthy et al. [20] proposed a queueing-inventory system in which the items are delivered to the customers with a specified probability upon completing service. An inventory-production model with a Markovian queuing model is discussed by Baek and Moon [21] . This article considered a production facility that produces items according to a Poisson process. Manikandan and Nair [22] considered an M/M/1/1 queue with inventory where the unmet customers can retry to get service. An M/M/1 queueing system with an (s, Q) inventory and lost sales were modeled by Baek et al. [23] . Furthermore, Yue et al. [24] studied an M/M/1 queueing model with an attached inventory. Yue and Qin [4] examined an inventory-production system when the return of products is possible. They presented the stationary distributions of queue length and level of inventory to derive performance measures. A queueing-inventory model with contesting suppliers was analyzed by Saffari et al. [25] . A retrial queueing-inventory model with an (s, S) inventory policy was proposed by Shajin and Krishnamoorthy [26] . In this paper, when the server is idle, the arrival customers enter an orbit. Shajin et al. [27] studied a correlated queueing-inventory model where customer arrival and lead-time are Markovian and the service time has phase-type distribution. A queueing-inventory model with two vendors replenishing inventory has been discussed by Chakravarthy and Hayat [28] . They analyzed the model in steady-state utilizing the matrix analytical method. Manikandan and Nair [29] developed an M/M/1 queueing system with inventory considering lost sales, server working vacation, and vacation interruption. Shajin et al. [30] presented the system state probability distribution and cost function for a queueing-inventory model with necessary and optional inventories. Ozkar and Kocer [31] analyzed a two-commodity queueing-inventory model with two classes of customers and an individual ordering policy. Jeganathan et al. [32] developed a queue-dependent service rate in a queueing-inventory model. For more reading, refer to Graves [33] , Manuel et al. [34] , Karthick et al. [35] , Albrecher et al. [36] , Chakravarthy et al. [37] , Marand and Thorstenson [38] , Melikov et al. [39] , Melikov and Shahmaliyev [40] , Gowsalya et al. [41] , Hanukov et al. [42] , Anilkumar and Jose [43] , Keerthana et al. [44] , Krishnamoorthy et al. [45] , Rasmi and Jacob [46] .

In the above papers, the authors assumed that all items are non-defective while the suppliers' production system or internal production may be imperfect. Aghsami et al. [47] have recently studied a queueing-inventory model with an imperfect internal production system and a 100% screening process along with inspection errors. Hence, it is more realistic to consider that the received lots or batches contain a fraction of defective items. Also, it is rational to apply an inspection process to reduce the risk of delivering defective products to the customers if the received lots have faulty items. In the literature of inventory and production models, many studies have discussed defective items. Salameh and Jaber [48] developed the traditional EPQ/EOQ model by considering the fact that the received or produced items are not of perfect quality. Khan et al. [49] extended previous work for the case of learning in the inspection. Al-Salamah [50] studied an economic order quantity (EOQ) model where the buyer places orders from a supplier with perfect and imperfect products. Also, an acceptance sampling plan with destructive testing and inspection errors was applied in this paper. Moussawi-Haidar et al. [51] extended an integrated inventory-quality model where the received orders contain a random fraction of defective items; hence an acceptance sampling plan is performed to the lots instead of a 100% screening process. They showed that using the acceptance sampling plan remarkably increases the profit of the retailer. Hsu and Hsu [52] formulated an EOQ problem with defective items and shortage as well as inspection errors. It is presumed that the inspector may incorrectly classify a defective item as non-defective or vice-versa. Moussawi-Haidar et al. [53] considered an EOQ model with the deteriorating items. It is assumed the orders may contain defective items due to breakage in transportation or imperfect production. Also, the retailer conducts a 100% screening process to eliminate non-conforming items from the received lots. Priyan and Uthayakumar [54] examined a two-level supply chain problem with a vendor and a buyer in which the buyer performs a screening process with errors when a lot is received. To obtain an optimal ordering policy, Chang et al. [55] developed an EOQ model with defective items, inspection errors, and permissible delays in payments. Hasanpour and Sharifi [56] studied an EOQ model with perishable items and destructive acceptance sampling plan and inspection errors. An EOQ model for sampling and sample quality inspection has been proposed by Cheikhrouhou et al. [57] . Upon the lot is received, the retailer performs a sampling process. If the lot is rejected, it is returned to the supplier. Mokhtari and Asadkhani [8] studied an inventory model with defective items, imperfect inspection, and batch replacement. Maleki Vishkaei et al. [58] presented a multi-product multi-supplier single-retailer inventory problem where the retailer purchases products from diff erent suppliers with different buying costs and defective rates. The retailer conducts a destructive testing acceptance-sampling plan in which he returns the rejected lots to the suppliers; another lot is sent to the retailer at no cost. Wangsa et al. [59] studied a single vendor-buyer inventory model where the buyer has an imperfect production system, and the vendor uses a 100% quality check. Safarnezhad et al. [60] focused on a single vendor and single buyer supply chain where if the inventory level reaches a reorder point, the buyer places an order to the vendor. The authors assumed that each lot might have a proportion of defective items that leads to conducting some inspection types, i.e., no inspection, sampling inspection, and 100% inspection to assess the quality of receiving lots. Taleizadeh [61] gives a comprehensive survey of imperfect inventory systems. Asadkhani et al. [62] have recently proposed four EOQ models with various types of imperfect quality items and inspection errors. Additional articles such as Wu et al. [63] , Jolai et al. [64] , Datta [65] , Taleizadeh and Hasani [66] , Jauhari et al. [67] , Pal and Mahapatra [68] , Amirhosseini et al., [69] , Manna et al. [70] , Taheri-Tolgari et al. [71] , Sarkar and Giri [72] , Karakatsoulis and Skouri [73] and Adak and Mahapatra [74] studied various aspects of inventory models with defective items and inspection process.

According to the above literature, only one paper [47] has studied a queueing-inventory model with defective items, while no one of the other queueing-inventory models has considered defective items in the inventory. All of them assumed the internal production system is perfect or the suppliers' lots have no defective items, while due to weak process control, deterioration, deficient planned maintenance, wear and tear of machinery, operator problems, and so on, a production process, whether for the internal productions or suppliers, may produce defective items, which is more realistic and practical [68] . Besides, receiving defective items by customers decreases loyalty, customer satisfaction, and profitability. Hence, conducting an inspection process in such conditions would be rational. Furthermore, the arriving customers to the retailer need to talk with sales experts and get information about the availability of items, function, price, brand, material, after-sales service, etc. Before buying. Therefore, the customers don't have information about products and the status of the retailer's inventory, and as a result, the retailer's system is unobservable for them in such situations. In contrast, the customers or demands in the previous queueing-inventory or imperfect production/inventory models had enough information before entering the system. For example, in [17] [24] , [21] , [47] , etc., it has been assumed that the information of inventory level is observable for the arriving customers, and they don't join the system during the stockout period. Consequently, analyzing the queueing-inventory model with a response system would be more practical.

Table 1 demonstrates some related literature reviews of queueing-inventory and imperfect production/inventory models. We implicitly split this table into three parts that have been separated with black lines. The first two parts indicate queueing-inventory and imperfect production/inventory models, respectively. The third part shows only one article that has recently discussed a queueing-inventory model with defective items and the inspection process [47] . Based on the first part, no queueing-inventory model has discussed defective items except Aghsami et al. [47]]. Also, the response queueing system hasn't been considered in the models of this part because they have been developed in the observable case. Besides, the customers' queue and response system, which are more practical, have been ignored in the classical inventory models, such as articles in the second part of Table 1. As mentioned, the unobservable case in which the arriving customers need to get some information from sellers before buying an item hasn't been studied in the literature of queueing-inventory and imperfect production/inventory models to the best of our knowledge. Aghsami et al. [47] studied a queueing-inventory model with defective items in the observable case. Hence, they didn't discuss the response queueing system. In addition, they assumed an internal production replenishes the inventory while in the retailer-supplier inventory problems, the retailer replenishes inventory through placing orders to suppliers. Moreover, they proposed a 100% screening process to detect defective items, whereas it isn't an appropriate method for destructive tests. On top of that, the 100% screening process can be costly and time-consuming rather than acceptance sampling plans. Therefore, the mentioned gaps and importance of considering the response queueing system and defective items in the retailer-supplier inventory models motivated us to analyze a retailer-supplier inventory problem with imperfect quality and destructive testing acceptance sampling plan integrating with a Markovian response queueing system in order to consider the whole system in the unobservable case. The contributions and novelties of this article can be expressed as follows:

- Integrating a retailer-supplier inventory model with a novel queueing-inventory system.
- Considering defective items in a retailer-supplier queueing- inventory model.
- Considering an M/M/m response queueing model for responding and handling arriving customers in a retailer system.
- Developing a queueing-inventory model with defective items and destructive testing acceptance sampling plan in the unobservable case.
- Modeling the retailer's inventory system with defective items and inspection process as a continuous-time Markov chain (CTMC).

Presenting a nonlinear integer programming model for a new retailer-supplier queueing-inventory problem considering supplier's risk and retailer's risk.

Table 1. Summary of the most pertinent studies vs. the article

There are a lot of cases in the real world for such models with a response system. For example, consider a pharmacy that the demands (customers) arrive at for receiving requirement medicines. The pharmacy assistant takes in a prescription to prepare medicines. It is clear that he/she needs time to read the prescription and check whether the requested medicines are available in stock or not through a computer system. Hence, a queue of arriving demands may be formed. Provided that the requested medicines are available, the demands are satisfied; otherwise, lost sales occur. Because the customers have no information about the pharmacy inventory level, such systems exactly correspond to mentioned response system. The pharmacy would place an order to a supplier when the medicine inventory level reaches a reorder point. After delivering the order, an acceptance sampling plan is conducted by pharmaceutical inspectors due to possible deterioration, adulteration or counterfeiting, contamination, and so forth [77] . Also, there are a lot of products that the customers need to consult with sales experts and get some information about price, function, brand, material, guarantee, after-sales service, etc., Before purchasing. Unquestionably, the sales experts need time to respond to customers' inquires. Hence, the sales experts are the same as the servers in the retailer's response system. As another example, the customers who arrive at a rebar retailer must consult with him about the rebar size, grade, price, type, weight, etc. Besides, after receiving an order from the supplier, the retailer does some tests such as the tensile test, bending test, fatigue test, etc. It is well known that these tests are destructive engineering and materials science tests, and the retailer should apply an acceptance sampling plan. Therefore, such models correspond to a queueing-inventory problem integrated with a retailer-supplier inventory model considering defective

References	Queueing inventory models	Classical inventory model	Response queueing system	Stock replenishment policies		Type of items		Inspection policies			Information level for arrival customers	
				Ordering policy	Production policy	Non-defective	Defective	No inspection	100% inspection	Sampling inspection	Observable case	Unobservable case
Saffari et al. [1]	+			+		+		+			+	
Krishnamoorthy and Viswanath [16]	+				+	+		+			+	
Saffari et al. [17]	+			+		+		+			+	
Baek and Moon [19]	+			+	+	+		+			+	
Baek and Moon [21]	+				+	+		+			+	
Manikandan and Nair [22]	+			+		+		+			+	
Baek et al. [23]	+			+		+		+			+	
Saffari et al. [25]	+			+		+		+			+	
Yue and Qin [4]	+				+	+		+			+	
Shajin and Krishnamoorthy [26]	+			+		+		+			+	
Ozkar and Kocer [31]	+			+		+		+			+	
Rasmi and Jacob [46]	+			+		+		+			+	
Al-Salamah [50]		+		+			+			+	+	
Moussawi-Haidar et al. [51]		+		+			+			+	+	
Hsu and Hsu [52]		+		+			+		+		+	
Moussawi-Haidar et al. [53]		+		+			+		+		+	
Sivashankari and Panayappan [18]		+			+		+	+			+	
Chang et al. [55]		+		+			+		+		+	
Hasanpour and Sharifi [56]		+		+			+			+	+	
Manna et al. [70]		+			+		+		+		+	
Cheikhrouhou et al. [57]		+		+			+			+	+	
Maleki Vishkaei et al. [58]		+		+			+			+	+	
Safarnezhad et al. [60]		+		+			+	+	+	+	+	
Adak and Mahapatra [74]		+		+			+		+		+	
Asadkhani et al. [62]		+		+			+		+		+	
Aghsami et al. [47]	+				+		+		+		+	
This study	+		+	+			+			+		+

items and the inspection process.

Therefore, this paper analyzes defective items' effect on a queueing-inventory system of the retailer from a new point of view. It contributes to the literature in some directions. First, this study proposes a retailer-supplier inventory model with the attached queueing system in which the incoming lots are imperfect so that there are no queueing-inventory models with defective items and ordering policy in the literature. Second, a response system for responding and handling arriving demands has been developed as an M/M/m queueing model that had not ever been considered before. Third, based on the above literature, it is reasonable that the retailer applies an inspection process when the incoming lots are imperfect. Moreover, a 100% screening process is costly and time-consuming [51] and inappropriate for many products such as electrical wires, food, etc. [58] . In such situations, an acceptance sampling plan is conducted. If the testing procedure causes destruction or failure of items, the retailer must use the destructive testing acceptance sampling plan. Hence, we have considered a destructive acceptance-sampling plan in the proposed queueing-inventory system of the retailer. Fourth, the level of the retailer's inventory and its status change over time. Consequently, we look at the proposed retailer's system as a stochastic process for the first time.

Our paper is looking forward to answering the following questions:

- How do the stationary distributions, performance measures, and expected total cost of the retailer-supplier queueing-inventory system with imperfect incoming lots and destructive acceptance-sampling plan compute?
- How do the imperfect incoming lots from the supplier affect the retailer's queueing-inventory system?
- How does an acceptance-sampling plan affect the retailer's queueing-inventory system?
- How does a demand response system that performs as a queueing model affect the retailer-supplier queueing-inventory system?
- How do various parameters of the proposed model affect the optimal reorder point, order quantity, expected total cost, and performance measures?

To answer the questions mentioned above, this paper studies an M/M/m queueing-inventory model for a single retailer-single supplier system in which the supplier sends the imperfect lots to the retailer. The retailer has several servers to respond to the arrival demands that form a queueing system. Also, the retailer conducts a destructive acceptance-sampling plan to inspect the receiving lots due to existing defective items. The retailer continuously checks the inventory level and orders to the supplier when inventory reaches a reorder point. The paper's main objective is to minimize the long-run expected total cost with respect to the reorder point, order quantity, number of servers, acceptance number, and sample size using the joint stationary distributions and performance measures of the entire system.

The description of the model is presented in the next section. We analyze the model and derive stationary distributions in Section 3. In Section 4, we obtain some important performance measures. We develop the cost model and present a nonlinear integer programming model in Section 5. Section 6 discusses comprehensive numerical examples and sensitivity analysis. The discussion and some managerial insights are expressed in Section 7. Finally, we provide the conclusion and give opportunities for future directions in Section 8.

Model description

This paper studies an integrated queueing-inventory model with a single item, a single retailer, and a single supplier. The incoming lots may have defective items that lead to performing the inspection by the retailer. The retailer faces demands where arrive at according to a Poisson process with rate λ. We assume the retailer has been equipped with a response system with m servers to handle and respond to the requests. Each server needs time to respond and consult with arrival demands that is distributed exponentially with rate μ. Hence, a classical M/M/m queueing model is formed in the response system. We suppose $\lambda < m\mu$ to guarantee the ergodicity of the response system queue. After processing and departing from the response system, the demands leave the system with precisely one item if the retailer's inventory is not zero; otherwise, lost sales occur. According to [78] , the departure process of an M/M/m queue in steady-state is a Poisson process with the same arrival rate that here is λ. Therefore, if the inventory level is not zero, it decreases according to a Poisson process with rate λ. It is clear that the number of demands in the response system is independent of the retailer's inventory system. The retailer continuously checks the inventory level. Whenever it reaches the reorder point r, he places an order of size Q to the supplier. It takes an exponential time with parameter ν_1 for delivering an order to the retailer. Due to existing defective items in the received lots, the retailer uses a destructive acceptance-sampling plan to inspect them. Upon receiving a lot, the retailer takes a random sample of size n and screens it. If the number of defective items found in the sample is less than or equal to the acceptance number c, the entire lot is accepted, and the inspected items are discarded. As a result, $Q - n$ items are added to inventory. We assume $Q - n > r$ to avoid the removal of periods where no order is placed. We assume the inspection time of the sample is exponentially distributed with mean $\frac{1}{\omega}$. Because of the sampling inspection plan, it is rational that the inspection time is dependent on the sampling size. Accordingly, we define parameter ω as a linear function of n as $\omega = \frac{a}{n}$ in which a is the screening rate of an item. We also assume that the incoming lots contain a constant fraction of defective items p. In other words, we can presume that an in-control process produces all items. Although each sample's number of defective items follows a hypergeometric distribution when the lot size is finite, for simplicity and similar to [58] and [50] , we compute the acceptance probability of the incoming lots using the Binomial distribution with parameters n and p. Let x be the number of defective items in a sample. Therefore, the acceptance probability of the incoming lots is given by:

$$p_a = p(x \le c) = \sum_{x=0}^{c} \binom{n}{x} p^x (1-p)^{n-x} \tag{1}$$

Provided that the number of defective items found in the sample is greater than the acceptance number c, the entire lot is rejected and returned to the supplier. In this case, the retailer asks the supplier for a new lot without defective items [60] . According to [79] , the supplier should perform the screening and rework activity on the returned lots, which is the best approach to face these situations. Hence, he sends a new lot without defective items. Also, the retailer doesn't pay for ordering and the cost of destructed items in this case [58] . The time it takes a non-defective lot is sent to the retailer follows an exponential distribution with parameter ν_2. Afterward, Q items are added to the retailer's inventory without inspection. Moreover, we would like to design a sampling plan to protect the supplier and retailer. Therefore, we consider the supplier's risk (α), acceptable quality level (AQL), retailer's risk (β), and lot tolerance percent defective (LTPD). The AQL is the poorest level of quality for the supplier's lots that the retailer rejects with the probability α, and the LTPD is the poorest level of quality for the supplier's lots that the retailer accepts with the probability [79] .

Note that we call pre-inspection and after lot rejection orders as regular and special orders in this paper, respectively. Fig. 1 represents a graphical illustration of the model. It is assumed that the inter-arrival times, response times, lead times, and inspection times are mutually independent. We characterize the retailer's inventory system with $S(t)$ and $I(t)$ that is its state at time t, $t \ge 0$. The former indicates the order status, and the latter denotes the inventory level at time t. $S(t)$ takes the values of 0, 1, 2, and 3 if there is no outstanding order, there is a regular outstanding order, the regular order has been received, and it is under inspection, and there is a special outstanding order, respectively. Therefore, we can model the retailer's inventory system as a two-dimensional stochastic process as follows:

$$Z = \{(S(t), I(t)) : t \ge 0\} \tag{2}$$

It is evident that the sojourn time in each state is exponentially distributed, and the process Z is a CTMC with state-space and non-zero transition rates as follows:

$$\Omega = \{(s,i) : s = 1,2,3, 0 \le i \le r\} \cup \{(s,i) : s = 0, r+1 \le i \le r+Q\} \tag{3}$$

$$\begin{array}{ll}
q_{(s,i+1)(s,i)} = \lambda & s = 1,2,3 \quad i = 0,1,\dots,r-1 \\
q_{(0,i+1)(0,i)} = \lambda & i = r+1, r+2, \dots, r+Q-1 \\
q_{(0,r+1)(1,r)} = \lambda & \\
q_{(1,i)(2,i)} = \nu_1 & i = 0,1,\dots,r \\
q_{(2,i)(0,i+Q-n)} = \omega p_a & i = 0,1,\dots,r
\end{array}$$

$$q_{(2,i)(3,i)} = \omega(1-p_a) \qquad i = 0,1,\dots,r$$
$$q_{(3,i)(0,i+Q)} = \nu_2 \qquad i = 0,1,\dots,r$$

Note that we have two transition rate diagrams corresponding to the model. One is related to case $r-n \geq 0$, which is depicted in Fig. 2. The other is related to $r-n<0$. In this case, state $(0, r+Q-n)$ places before the state $(0,Q)$, And Fig. 2 changes slightly accordingly.

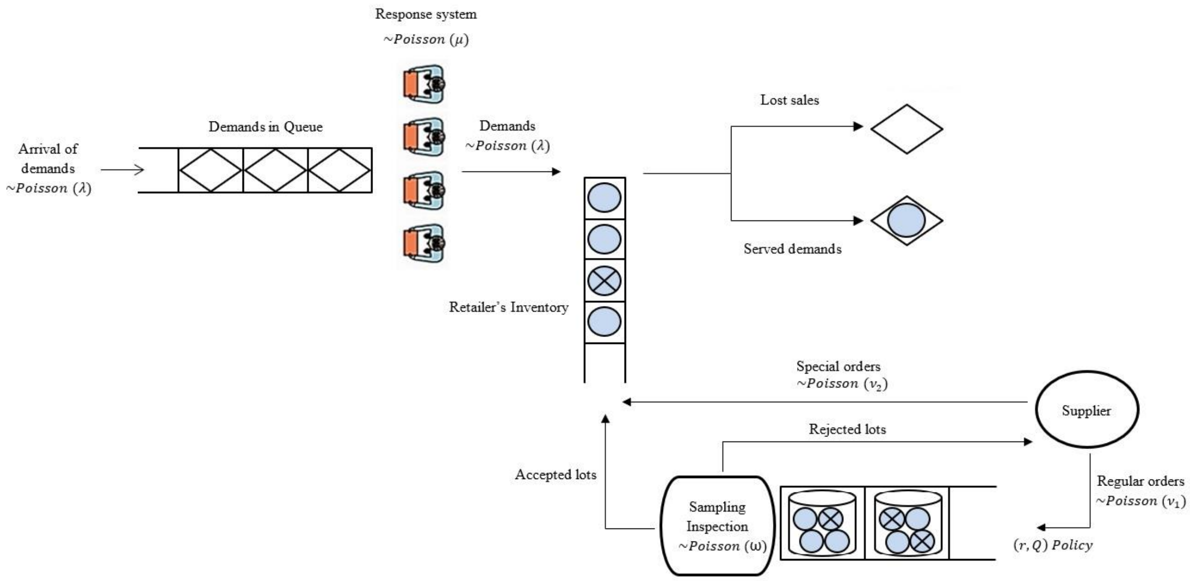


Fig. 1. The integrated queuing-inventory system of the retailer with a sampling inspection plan

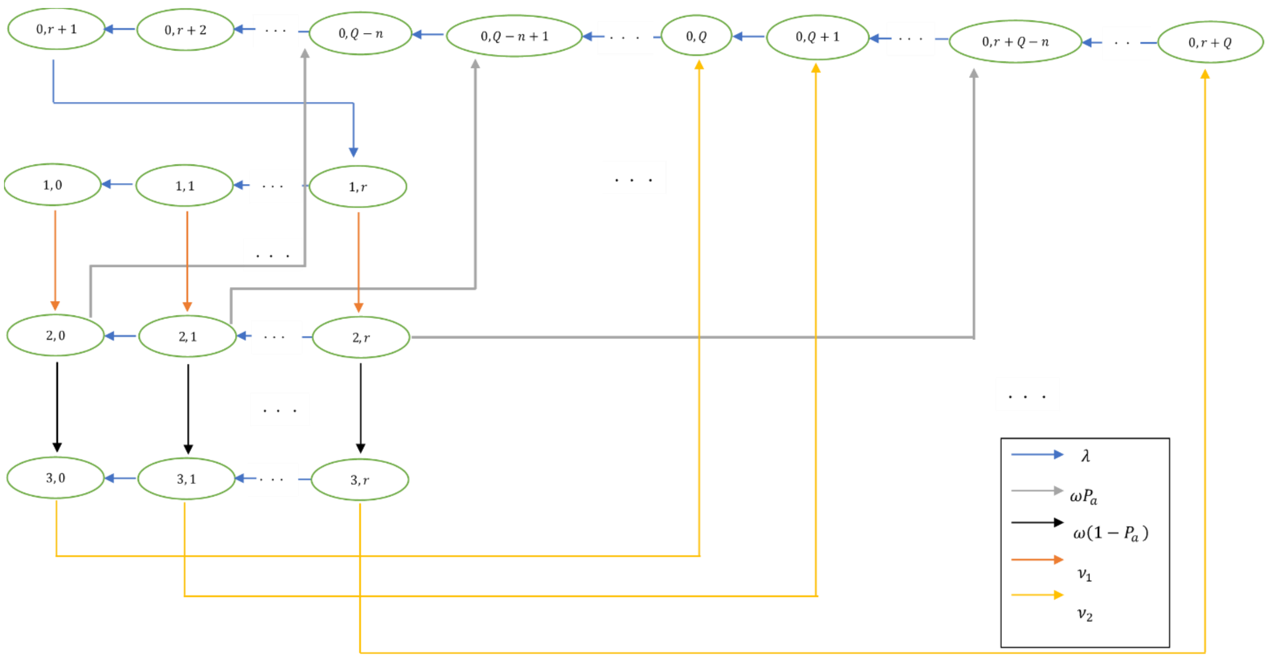


Fig. 2. Transition rate diagram of the model

Analysis of the model

In this section, we first derive the stationary distribution for the demand response queueing system. Afterward, the joint stationary distribution of the order status and inventory level of the retailer is computed.

Stationary distribution of demand response system

As mentioned before, the demand response system is the classical M/M/m queueing system. According to [80] , the stationary distribution of the number of demands in the system is given by:

$$\pi_n^d = \begin{cases} \dfrac{\lambda^n}{n!\,\mu^n}\pi_0^d & 0 \le n < \mathrm{m} \\ \dfrac{\lambda^n}{m^{n-\mathrm{m}}\mathrm{m}!\,\mu^n}\pi_0^d & n \ge \mathrm{m} \end{cases} \tag{4}$$

In which π_0^d can be obtained from the following equation:

$$\pi_0^d = \left(\sum_{n=0}^{m-1}\frac{\lambda^n}{n!\,\mu^n} + \sum_{n=m}^{\infty}\frac{\lambda^n}{m^{n-\mathrm{m}}\mathrm{m}!\,\mu^n}\right)^{-1} \tag{5}$$

Stationary distribution of the process Z

In this section, we derive the joint stationary distribution $\pi(s,i)$, which s and i respectively, denote the order status and the inventory level of the retailer in stationary. In other words, we have:

$$\lim_{t\to\infty}\pi(S(t),I(t)) = \pi(s,i) \tag{6}$$

Recalling the state space Ω expressed by Eq. 3. We can rewrite it as $\Omega = \bigcup_{s=0}^{3}\Psi(s)$, in which $\Psi(s)$ is called sth level. We define level s as follows:

$$\Psi(s) = \begin{cases}\{(s,r+1),(s,r+2),\dots,(s,r+Q)\} & s=0\\ \{(s,0),(s,1),\dots,(s,r)\} & s=1,2,3\end{cases} \tag{7}$$

In fact, in level 0, the number of items in the retailer's inventory gets the integer value from $r+1$ to $r+Q$, and in other levels, the number of items has the integer value from 0 to r. The infinitesimal generator matrix of the process Z is as follows:

$$\Phi = \begin{bmatrix} B_{00} & B_{01} & & \\ & B_{11} & B_{12} & \\ B_{20} & & B_{22} & B_{23} \\ B_{30} & & & B_{33}\end{bmatrix} \tag{8}$$

Where the block matrix B_{jk} for j and $k \in s$ indicates the transition from level j to level k, and the block matrix B_{jj} includes transition rates within $\Psi(j)$. Also, the elements of each block matrix indicate the transition rates among states.

Now, we describe the block matrices as follows:

$$B_{00} = \begin{bmatrix} -\lambda & & & & & \\ \lambda & -\lambda & & & & \\ & \lambda & -\lambda & & & \\ & & \cdots & \cdots & & \\ & & & \lambda & -\lambda & \\ & & & & \lambda & -\lambda \end{bmatrix} \tag{9}$$

$$B_{01} = \begin{array}{c} \\ r+1 \\ r+2 \\ \vdots \\ r+Q \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & \cdots & r-1 & r \end{array} \\ \begin{bmatrix} & & & & \lambda \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \end{array} \tag{10}$$

$$B_{11} = \begin{bmatrix} -\nu_1 & & & & & \\ \lambda & -(\lambda+\nu_1) & & & & \\ & \lambda & -(\lambda+\nu_1) & & & \\ & & \cdots & \cdots & & \\ & & & \lambda & -(\lambda+\nu_1) & \\ & & & & \lambda & -(\lambda+\nu_1) \end{bmatrix} \tag{11}$$

$$B_{12} = \begin{bmatrix} \nu_1 & & & & & \\ & \nu_1 & & & & \\ & & \nu_1 & & & \\ & & & \cdots & & \\ & & & & \nu_1 & \\ & & & & & \nu_1 \end{bmatrix} \tag{12}$$

$$B_{20} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ r \end{array} \begin{array}{c} \begin{array}{cccccccc} r+1 & \cdots & Q-n & Q-n+1 & \cdots & r+Q-n & \cdots & r+Q \end{array} \\ \begin{bmatrix} & & \omega p_a & & & & & \\ & & & \omega p_a & & & & \\ & & & & \cdots & & & \\ & & & & & \omega p_a & & \end{bmatrix} \end{array} \tag{13}$$

$$B_{22} = \begin{bmatrix} -\omega & & & & & \\ \lambda & -(\lambda+\omega) & & & & \\ & \lambda & -(\lambda+\omega) & & & \\ & & \cdots & \cdots & & \\ & & & \lambda & -(\lambda+\omega) & \\ & & & & \lambda & -(\lambda+\omega) \end{bmatrix} \tag{14}$$

$$B_{23} = \begin{bmatrix} \omega(1-p_a) & & & & \\ & \omega(1-p_a) & & & \\ & & \cdots & & \\ & & & \omega(1-p_a) & \\ & & & & \omega(1-p_a) \end{bmatrix} \tag{15}$$

$$\begin{array}{cccccc} r+1 & \cdots & Q & Q+1 & \cdots & r+Q \end{array}$$

$$B_{30} = \begin{array}{c} 0 \\ 1 \\ \vdots \\ r \end{array} \begin{bmatrix} & \nu_2 & & & \\ & & \nu_2 & & \\ & & & \cdots & \\ & & & & \nu_2 \end{bmatrix} \quad (16)$$

$$B_{33} = \begin{bmatrix} -\nu_2 & & & & & \\ \lambda & -(\lambda+\nu_2) & & & & \\ & \lambda & -(\lambda+\nu_2) & & & \\ & & \cdots & \cdots & & \\ & & & \lambda & -(\lambda+\nu_2) & \\ & & & & \lambda & -(\lambda+\nu_2) \end{bmatrix} \quad (17)$$

Let $\pi = [\pi_0, \pi_1, \pi_2, \pi_3]$ denotes the stationary distribution vector of the process Z with infinitesimal generator matrix Φ where π_s is defined as:

$$\pi_s = \begin{cases} (\pi(0, \mathrm{r}+1), \pi(0, \mathrm{r}+2), \dots, \pi(0, \mathrm{r}+\mathrm{Q})) & s = 0 \\ (\pi(s, 0), \pi(s, 1), \dots, \pi(s, \mathrm{r})) & s = 1, 2, 3 \end{cases} \quad (18)$$

Accordingly, π satisfies

$$\pi\Phi = 0$$
$$\pi e = 1 \quad (19)$$

In which e is a column vector of 1s of order $(Q + 3(r + 1)) \times 1$.

Taking Eq. 8 into account, we have the extended form of the above equations as follows:

$$\pi_0 B_{00} + \pi_2 B_{20} + \pi_3 B_{30} = 0 \quad (20)$$
$$\pi_0 B_{01} + \pi_1 B_{11} = 0 \quad (21)$$
$$\pi_1 B_{12} + \pi_2 B_{22} = 0 \quad (22)$$
$$\pi_2 B_{23} + \pi_3 B_{33} = 0 \quad (23)$$
$$\pi_0 e_0 + \pi_1 e_1 + \pi_2 e_2 + \pi_3 e_3 = 1 \quad (24)$$

Where e_s is a column vector of 1s of order $Q \times 1$ for $s = 0$ and order $(r + 1) \times 1$ for $s = 1, 2, 3$.

We compute all stationary probabilities by solving the balance Eqs. 20-24 as follows:

Taking Eq. 21 into account, we have:

$$\pi_1 = -\pi_0 B_{01} {B_{11}}^{-1} \quad (25)$$

With the help of Eq. 22, we obtain

$$\pi_2 = -\pi_1 B_{12} {B_{22}}^{-1} \quad (26)$$

By substituting Eq. 25 into Eq. 26, we will have

$$\pi_2 = \pi_0 B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} \quad (27)$$

Based on Eq. 23, π_3 is computed as

$$\pi_3 = -\pi_2 B_{23} {B_{33}}^{-1} \quad (28)$$

By substituting Eq. 27 into Eq. 28, we get

$$\pi_3 = -\pi_0 B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{23} {B_{33}}^{-1} \tag{29}$$

It is clear that the matrices B_{11}, B_{22}, and B_{33} have non-zero determinant and are invertible. By substituting Eqs. 27 and 29 into Eq. 20, we obtain:

$$\pi_0 B_{00} + \pi_0 B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{20} - \pi_0 B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{23} {B_{33}}^{-1} B_{30} = 0 \tag{30}$$

We can rewrite the above equation as follows:

$$\pi_0 \left(B_{00} + B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{20} - B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{23} {B_{33}}^{-1} B_{30} \right) = 0 \tag{31}$$

The Eq. 31 is not enough to compute stationary probabilities of $\Psi(0)$ that are $\left(\pi(0, \mathrm{r}+1), \pi(0, \mathrm{r}+2), \dots, \pi(0, \mathrm{r}+\mathrm{Q})\right)$. Hence, we need the normalized Eq. 24. By substituting Eqs. 25, 27, and 29 into Eq. 24, we will have:

$$\pi_0 e_0 - \pi_0 B_{01} {B_{11}}^{-1} e_1 + \pi_0 B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} e_2 - \pi_0 B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{23} {B_{33}}^{-1} e_3 = 1 \tag{32}$$

After factoring π_0We have the following equation:

$$\pi_0 \left(e_0 - B_{01} {B_{11}}^{-1} e_1 + B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} e_2 - B_{01} {B_{11}}^{-1} B_{12} {B_{22}}^{-1} B_{23} {B_{33}}^{-1} e_3 \right) = 1 \tag{33}$$

Consequently, the system of Eqs. 31 and 33 give the stationary probabilities of $\Psi(0)$. Finally, we can obtain π_1, π_2, and π_3 using Eqs. 25, 27, and 29, respectively.

System performance measure

In this section, some entire system's performance measures in the steady-state are presented. We'll apply them to develop the cost function in the subsequent section. Note that we derive the clauses (I) based on the classical M/M/m queueing system formulas (see [80] and [78]).

(I) The average time that a demand spends in the response system (W^d) is given by

$$W^d = \frac{1}{\mu} + \left(\left(\frac{\lambda}{\mu} \right)^m \frac{1}{m!\,(m\mu)(1-\rho)^2} \right) \pi_0^d$$

(II) The average number of demands in the response system (L^d) per unit time is computed as

$$L^d = \sum_{n=0}^{\infty} n \pi_n^d = \frac{\lambda}{\mu} + \left(\left(\frac{\lambda}{\mu} \right)^m \frac{1}{m!} \frac{\rho}{(1-\rho)^2} \right) \pi_0^d$$

Where $= \frac{\lambda}{m\mu}$.

(III) The expected on-hand inventory level of the retailers per unit time (L_{inv}) is obtained as

$$L_{inv} = \sum_{i=r+1}^{r+Q} i\pi(0, i) + \sum_{s=1}^{3} \sum_{i=0}^{r} i\pi(s, i)$$

(IV) The mean number of regular orders per unit time (L_{ro}) is given by

$$L_{ro} = \pi(1, r)$$

(V) The mean number of regular outstanding orders per unit time (L_{roo}) is given by

$$L_{roo} = \sum_{i=0}^{r} \pi(1, i)$$

(VI) The mean number of special outstanding orders per unit time (L_{soo}) is computed as

$$L_{soo} = \sum_{i=0}^{r} \pi(3, i)$$

(VII) The average number of lost sales per unit time (L_{loss}) is obtained as

$$L_{loss} = \lambda \left(\sum_{s=1}^{3} \pi(s, 0) \right)$$

(VIII) The average number of defective items added to the retailer's inventory (L_{def}) is given by

$$L_{def} = \left(\sum_{i=0}^{r} \pi(2, i) \right) p_a (Q - n) p$$

(IX) The mean number of inspected items (L_{ins}) is computed as

$$L_{ins} = \left(\sum_{i=0}^{r} \pi(2, i) \right) n$$

(X) The average outgoing quality (AOQ) is derived as

$$AOQ = \frac{\left(\sum_{i=0}^{r} \pi(2, i)\right) p_a (Q - n) p}{\left(\sum_{i=0}^{r} \pi(2, i)\right) p_a (Q - n) + \left(\sum_{i=0}^{r} \pi(3, i)\right) Q}$$

(XI) The probability that there is at least one demand in the response system and the on-hand inventory level of the retailer is zero (P_{roiz}) is obtained

$$P_{roiz} = \left(1 - \pi_0^d\right) \left(\sum_{s=1}^{3} \pi(s, 0) \right)$$

(XII) The probability that the response system is empty and the on-hand inventory level of the retailer isn't zero (P_{rzio}) is given by

$$P_{rzio} = \pi_0^d \left(1 - \sum_{s=1}^{3} \pi(s, 0) \right)$$

(XIII) The percentage of time the retailer is doing inspect (P_{rii}) is computed as

$$P_{rii} = \sum_{i=0}^{r} \pi(2, i)$$

(XIV) The percentage of time the retailer doesn't have any outstanding order (P_{rdo}) is obtained as

$$P_{rdo} = \sum_{i=r+1}^{r+Q} \pi(0, i)$$

(XV) The percentage of time the retailer has a regular outstanding order (P_{rhro}) is given by

$$P_{rhro} = \sum_{i=0}^{r} \pi(1, i)$$

(XVI) The percentage of time the retailer has a special outstanding order (P_{rhso}) is given by

$$P_{rhso} = \sum_{i=0}^{r} \pi(3,i)$$

Cost model

This section develops the expected total cost of the system per unit time in the steady-state and presents a nonlinear integer programming model. We define the following cost factors:

C_h: inventory holding cost per item unit per unit time.
C_W: Waiting cost of a demand that leaves the system without receiving the item per unit time.
C_{or}: Cost of ordering a regular replenishment
C_l: Cost of losing a demand
C_p: Cost of purchasing an item
C_{ins}: Cost of inspecting an item
C_{des}: Cost of destructing an item
C_{pdi}: Post-sale defective item cost of the retailer per item
C_{msr}: Marginal cost of a server in the response system per unit time

Note that when a lot is rejected, the process Z moves to $\Psi(3)$ and may enter the state $(3,0)$. In this case, the system faces lost sales, which is an outcome of rejecting a lot. Therefore, the cost of lot rejection has been integrated into lost sale cost. It should be noted that each defective item that delivers to the customers causes penalty and charge for the retailer due to decreasing customers satisfaction and reputation of one. Therefore, we considered C_{pdi} as defective item cost after selling for the retailer.

The long-run average total cost per unit time $ETC(r,Q,n,c,m)$ using the system performance measures derived in Section 4 is as follows:

$$\begin{aligned} ETC(r,Q,n,c,m) &= INVC + WC + OC + LC + PC + INSC + DC + PSC + SC \\ &= C_h L_{inv} + C_W L^d W^d \left(\sum_{s=1}^{3} \pi(s,0)\right) + C_{or} L_{ro} + C_l L_{loss} + C_p Q L_{ro} + C_{ins} L_{ins} \\ &+ C_{des} L_{ins} p_a + C_{pdi} L_{def} + C_{msr} m \end{aligned} \tag{34}$$

Taking into account the ergodicity of the response system queue, the condition for avoiding the removal of periods where no order is placed, and protecting the supplier and retailer, we have a nonlinear integer programming model as:

Minimize

$$\begin{aligned} ETC(r,Q,n,c,m) \\ &= C_h L_{inv} + C_W L^d W^d \left(\sum_{s=1}^{3} \pi(s,0)\right) + C_{or} L_{ro} + C_l L_{loss} + C_p Q L_{ro} + C_{ins} L_{ins} \\ &+ C_{des} L_{ins} p_a + C_{pdi} L_{def} + C_{msr} m \end{aligned} \tag{35}$$

Subject to

$$\lambda < m\mu \tag{36}$$

$$Q - n > r \tag{37}$$

$$1 - \alpha = \sum_{x=0}^{c} \binom{n}{x} AQL^x (1 - AQL)^{n-x} \tag{38}$$

$$\beta = \sum_{x=0}^{c} \binom{n}{x} LTPD^{x}(1-LTPD)^{n-x} \tag{39}$$

$$n, m \in N \tag{40}$$

$$r, Q, c \in Z^{+} \tag{41}$$

Inequality (36) guarantees the ergodicity of the response system queue. Inequality (37) avoids the removal of periods where no order is placed. Eqs. 38 and 39 give a solution to n and c considering suppliers' and retailers' risks. The constraint (40) indicates that the variables n and m are natural numbers. The constraint (41) indicates that the other variables are natural numbers.

Numerical examples and sensitivity analysis

This section presents some numerical examples to reveal the possible convexity of the expected total cost developed in the previous section in the feasible region. Due to the complexity of the expected total cost, it is extremely hard to demonstrate its convexity analytically. Hence, we employ a numerical procedure to show the cost function's convexity corresponds to the variables authorizing to change. Also, we compute the (local) optimum values of the variables r, Q, and m for fixed values of n and c obtained from solving the system of Eqs. 38 and 39. We consider values set for r, Q and m and fix any one of the variables by these values. Afterward, we allow the other two variables to change in an interval to determine their (local) optimum values. Therefore, by the procedure, we can study the behavior of the expected total cost $ETC(r, Q, n, c, m)$. Moreover, we assess various parameters' effect on the optimal decision variables, corresponding expected total cost, and some performance measures. It should be noted that our goal in this article is to analyze the behavior of expected total cost and illustrate the characteristic of one because this model has been developed for the first time. Regarding that we cannot depict a four-dimensional plot of the expected total cost, we fix any of the variables r, Q and m and draw a three-dimensional plot of the expected total cost for the other two variables to reveal the characteristics of it. Hence, we don't focus on obtaining the optimal global solutions because it can be done directly by search methods such as grid search in a pre-determined interval. Moreover, we're going to perform a comprehensive sensitivity analysis to understand better the model's behavior and present valuable managerial insight to managers and practitioners.

We assume that $\alpha = 0.05$, $\beta = 0.1$, $AQL = 0.01$ and $LTPD = 0.06$. It is possible to use the Binomial nomograph instead of solving Eqs. 38 and 39 and construct the sampling plan n and c. According to [79] , the intersection of the line connecting $(AQL, 1-\alpha)$ and $(LTPD, \beta)$ on the Binomial nomograph locates a region that includes desired sampling plan. Fig. 3 shows the Binomial nomograph. Although this procedure may give several sampling plans with Operating Characteristics (OC) curves that pass close to the mentioned points, we consider the sampling plan $n = 89$ and $c = 2$ that is too close to passing through points (0.01, 0.95) and (0.06, 0.1) on the OC curve [79] . Also, We let $\lambda = 40$ customers/hour, $\mu = 50$ customers/hour, $a = 50$ items/hour, $\nu_1 = 1.2$ orders/hour, $\nu_2 = 0.7$ orders/hour, $p = 0.02$, $C_h = 8$ \$/item, $C_W = 60$ \$/customer/hour, $C_{or} = 200$ \$/order, $C_l = 100$ \$/customer, $C_p = 40$ \$/item, $C_{ins} = 0.8$ \$/item, $C_{des} = 40$ \$/item, $C_{pdi} = 500$ \$/item and $C_{msr} = 2$ \$/server/hour. We assume the time unit is an hour.

First, we fix $m = 2$, and obtain the expected total cost per hour under the function of $ETC(r, Q, 89, 2, 2)$, that its values are shown in Table 2. In this table, bold form and underlined values indicate the row and column minimum, respectively. Provided that a value is in the bold and underlined form simultaneously, it is the optimal value of the expected total cost correspondence to the table. Also, it may be the minimum value of the expected total cost. Fig. 4 depicts a three-dimensional plot of the expected total cost $ETC(r, Q, 89, 2, 2)$ that reveals the cost function's convexity for various combinations r and Q.

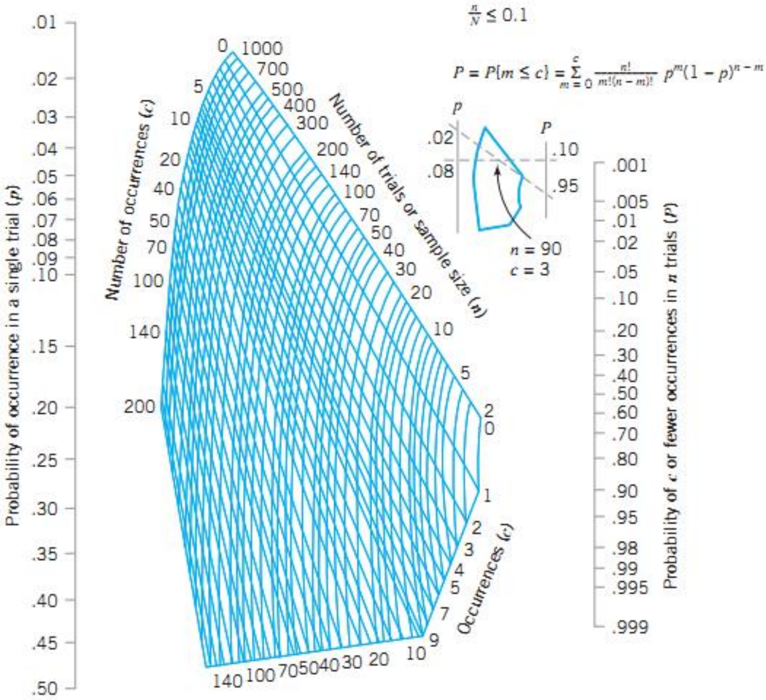


Fig. 3. Binomial nomograph (Montgomery [79])

Based on Table 2 and Fig. 4, the (local) optimal reorder point and order quantity values are $r^* = 81$, $Q^* = 318$, respectively. Moreover, the (local) minimum expected total cost is $ETC(r^*, Q^*, 89, 2, 2) = 2703.77839$. Table 3 represents the optimal performance measures corresponding to this point. The stationary distributions of all levels are shown in Fig. 5.

Table 2. The expected total cost as a function of r and Q

Q / r	316	317	318	319	320	321
78	2704.12855	2704.05360	2704.00521	**2703.98313**	2703.98709	2704.01685
79	2703.97013	2703.90579	2703.86798	**2703.85646**	2703.87097	2703.91126
80	2703.87333	2703.81957	2703.79234	**2703.79138**	2703.81643	2703.86724
81	2703.83823	2703.79506	**2703.77839**	2703.78798	2703.82356	2703.88489
82	2703.86493	2703.83232	**2703.82621**	2703.84634	2703.89249	2703.96426
83	2703.95348	**2703.93144**	2703.93587	2703.96652	2704.02312	2704.10543
84	2704.103946	**2704.09246**	2704.10741	2704.14857	2704.21567	2704.30844

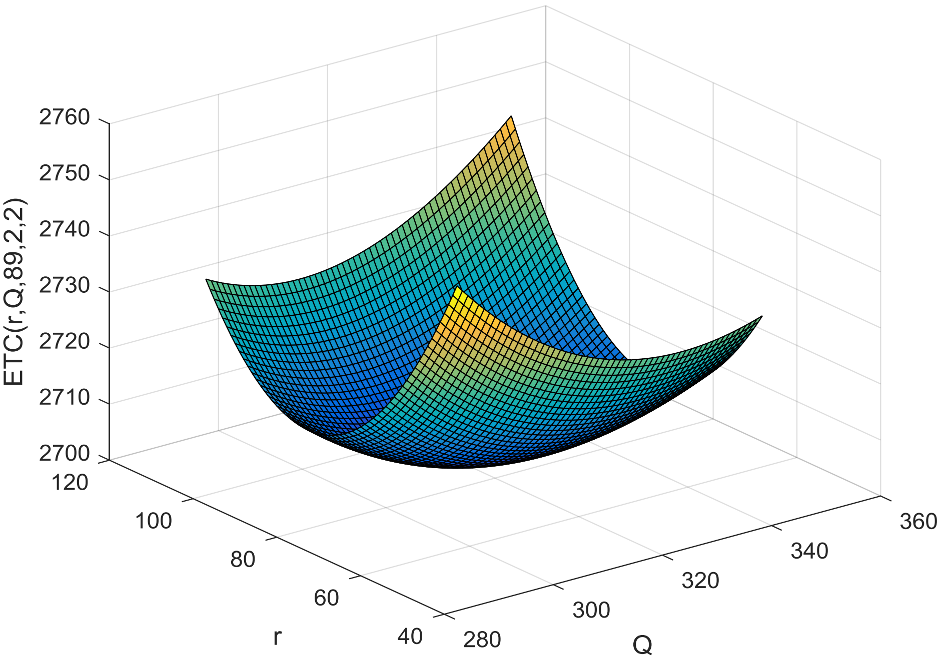
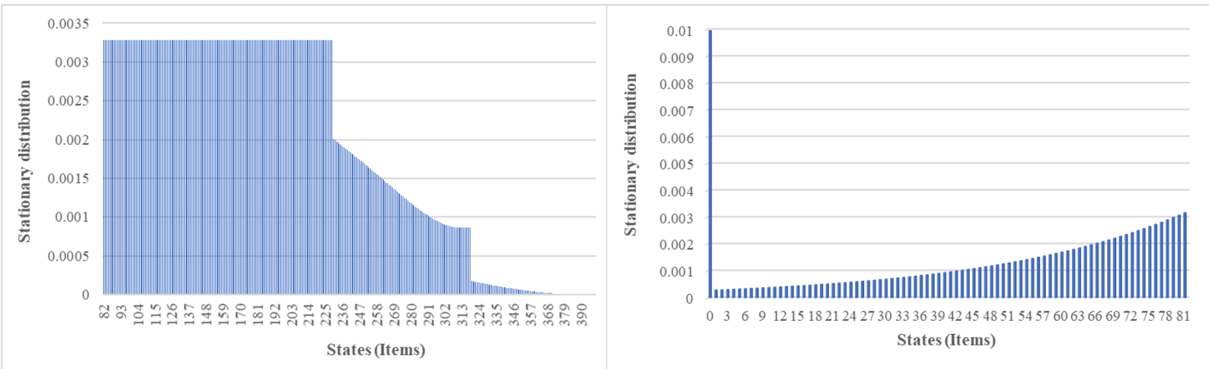


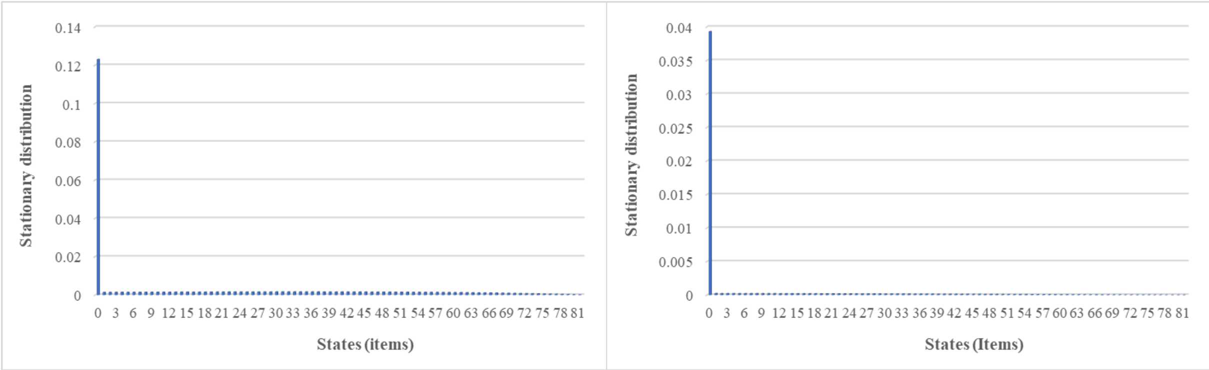
Fig. 4. A three-dimensional plot of the expected total cost $ETC(r, Q, 89, 2, 2)$

Table 3. The optimal system performance measures corresponding to $(r^* = 81, Q^* = 318, 89, 2, 2)$

Performance measure	Optimal values	Performance measure	Optimal values
W^d (hours)	0.02381	L_{ins} (Items)	20.77535
L^d (Customers)	0.95238	AOQ	0.01430
L_{inv} (Items)	118.31659	P_{roiz}	0.09849
L_{ro} (Orders)	0.00318	P_{rzio}	0.35471
L_{roo} (orders)	0.10928	P_{rii}	0.23343
L_{soo} (orders)	0.04935	P_{rdo}	0.60793
L_{loss} (Customers)	6.89417	P_{rhro}	0.10928
L_{def} (Items)	0.78749	P_{rhso}	0.04935



(a) (b)



(c) (d)

Fig. 5. The stationary distributions of levels of the process Z corresponding to $(81, 318, 89, 2, 2)$: (a) π_0 ; (b) π_1 ; (c) π_2 ; (d) π_3

Now, we fix $r = 81$, and study the expected total cost per unit time under the function of $ETC(81, Q, 89, 2, m)$, that its values are shown in Table 4. Moreover, Fig. 6 demonstrates a three-dimensional plot of the expected total cost $ETC(81, Q, 89, 2, m)$ that reveals the cost function's convexity for various combinations Q and m. In a similar manner in the previous case and based on Table 4 and Fig. 5, the (local) optimal order quantity and the number of servers is $Q^* = 81, m^* = 2$, respectively. Furthermore, the (local) minimum expected total cost is $ETC(81, Q^*, 89, 2, m^*) = 2703.77839$. Finally, we fix $Q = 318$, and analyze the expected total cost per unit time under the function of $ETC(r, 318, 89, 2, m)$, that its values are shown in Table 5. In addition, Fig. 7 represents a three-dimensional plot of the expected total cost $ETC(r, 318, 89, 2, m)$ that reveals the cost function's convexity for various combinations r and m. In a similar manner in the previous cases and based on Table 5 and Fig. 7, the (local) optimal reorder point and the number of servers is $r^* = 81, m^* = 2$, respectively. Besides, the (local) minimum expected total cost is $ETC(r^*, 318\ , 89, 2, m^*) = 2703.77839$.

Table 4. The expected total cost as a function of m and Q

Q / m	316	317	318	319	320	321
1	2705.76599	2705.70990	2705.68040	**2705.67724**	2705.70015	2705.74889
2	2703.83823	2703.79506	**2703.77839**	2703.78798	2703.82356	2703.88488
3	2705.77671	2705.73374	**2705.71728**	2705.72706	2705.76284	2705.82436
4	2707.76974	2707.72679	**2707.71035**	2707.72016	2707.75596	2707.81750
5	2709.76886	2709.72591	**2709.70947**	2709.71928	2709.75509	2709.81664

Table 5. The expected total cost as a function of m and Q

r / m	78	79	80	81	82	83
1	2706.017866	2705.84354	2705.73102	**2705.68040**	2705.69176	2705.76516
2	2704.00521	2703.86798	2703.79234	**2703.77839**	2703.82621	2703.93587
3	2705.94237	2705.80571	2705.73065	**2705.71728**	2705.76567	2705.87589
4	2707.93524	2707.79865	2707.72366	**2707.71035**	2707.75881	2707.86909
5	2709.93434	2709.79776	2709.72277	**2709.70947**	2709.75794	2709.86823
6	2711.93423	2711.79765	2711.72267	**2711.70937**	2711.75784	2711.86813

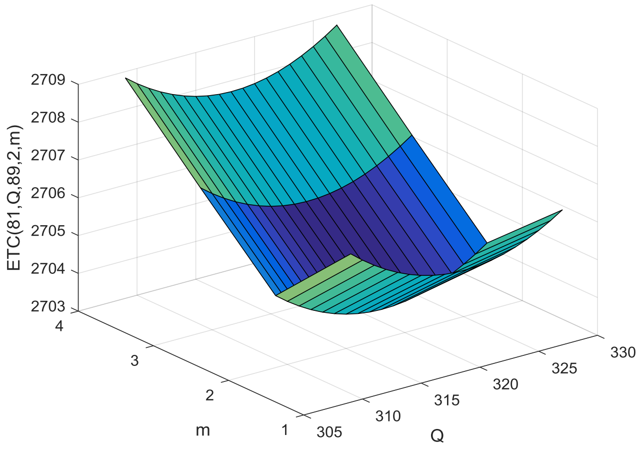


Fig. 6: A three-dimensional plot of the expected total cost $ETC(81, Q, 89, 2, m)$

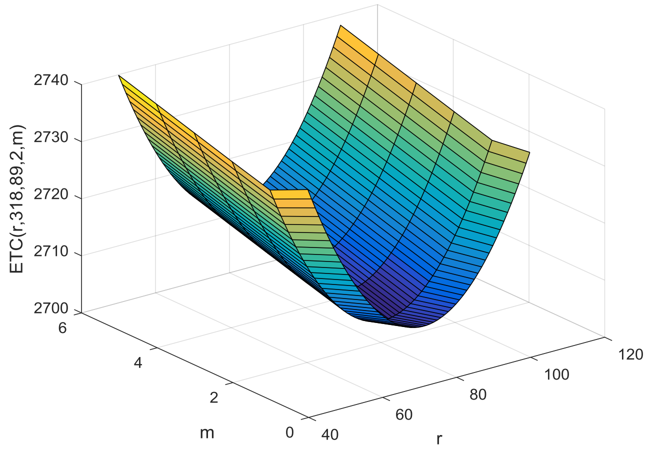


Fig. 7: A three-dimensional plot of the expected total cost $ETC(r, 318, 89, 2, m)$

Example 1. In this example, we study the impact of the proportion of defective items in each lot p on the optimal values of the reorder point and order quantity as well as corresponding expected total cost and some performance measures. For this purpose, the values of variables as $n = 89$, $c = 2$, $m = 2$ and values of parameters as $\lambda = 40$ customers/hour, $\mu = 50$ customers/hour, $a = 50$ items/hour, $\nu_1 = 1.2$ orders/hour, $\nu_2 = 0.7$ orders/hour, $p = 0.02$, $C_h = 8$ \$/item, $C_W = 60$ \$/customer/hour, $C_{or} = 200$ \$/order, $C_l = 100$ \$/customer, $C_p = 40$ \$/item, $C_{ins} = 0.8$ \$/item, $C_{des} = 40$ \$/item, $C_{pdi} = 500$ \$/item and $C_{msr} = 2$ \$/server/hour are fixed. From Table 5 and Figs. 8 to 13, we observe the following results:

1. According to Table 6 and Fig. 8, by increasing the fraction of defective items p, the optimal reorder point r^* and order quantity Q^* considerably increases and decreases, respectively. It decreases the percentage of time that the process Z is in level 0 (the percentage of time that the retailer doesn't have any outstanding order) and increases the percentage of time that it is in the other levels. This behavior continues until the probability of defective items is 0.04. Because the minimum of the total cost function will be in the feasible region if the probability of defective items changes in the interval [0 0.04]. When $p > 0.04$, the reorder point and order quantity related to the minimum of the total cost function doesn't satisfy the inequality (37), and the model searches the optimal values in the feasible region's boundary, and the behavior of the optimal reorder point and order quantity will change. For example, the optimal reorder point is 131 when $p = 0.045$. Based on inequality (37), $Q > 220$ is the feasible interval, and the optimal order quantity does not decrease like before. Hence, we call point $p = 0.04$ **the trending change point**. After this point, the optimal reorder point r^* and order quantity Q^* slightly increase. Also, the percentage of time that the retailer doesn't have any outstanding order mildly increases while the percentage of time that the retailer is waiting for regular orders or doing inspection decreases mildly (see Fig. 13). According to Fig. 13, the incoming lots are more rejected, and the percentage of time that the retailer is waiting for a special order rises when p increases.
2. The optimal expected total cost reduces when p increases (see Table 6 and Fig. 8) because the order quantity decreases, and the inventory cost and cost of destructing items reduce.
3. As shown in Table 6 and Fig. 9, the optimal purchased and destructive testing costs decline when the probability of defective items increases. It is because of the decrease of order quantity and increase of the reorder point up to **the trending change point**. Afterward, this behavior continues because of the rise in the reorder point.

4. Increasing p up to **the trending change point**, the waiting cost of demands that leave the system without receiving any item rises (see Table 6). Besides, the lost sale cost has a behavior same as waiting cost so that based on Fig. 9 and 12, the number of lost sales and associated cost rise until **the trending change point** due to the increase of stockout period. After passing the probability of defective items from **the trending change point**, the waiting and lost sales costs slightly decrease, but their values are still greater than values before this point because the average inventory increases somewhat after the point (see Fig. 12).
5. According to Table 6, the ordering cost increases until **the trending change point** because the optimal order quantity declines up to that. Consequently, the average number of orders rises that leading to increasing the ordering cost. Afterward, the ordering cost mildly decreases because the optimal order quantity mildly rises, and the average number of orders decreases after the point.
6. From Fig. 13, we observe that by increasing p up to **the trending change point**, the percentage of time the process Z is in the inspection state increments that lead to increasing the average inspected items L_{ins}, and the related cost $INSC$. Afterward, their values decline at a lower speed (see Fig. 11 and Table 6).
7. Taking Table 6 and Fig. 12 into account, the average on-hand inventory and its cost reduce when p increases to the trending change point, and afterward, the inventory cost rises. It is because the reorder point and order quantity increase and decrease, respectively up to **the trending change point.** Then, order quantity mildly increases.

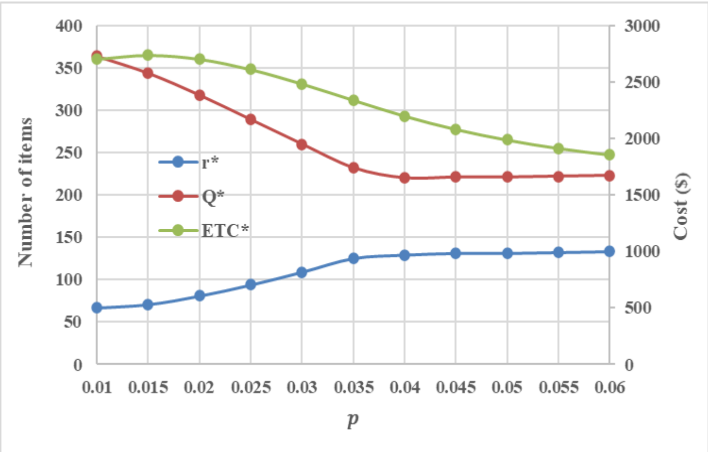


Fig. 8. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus p

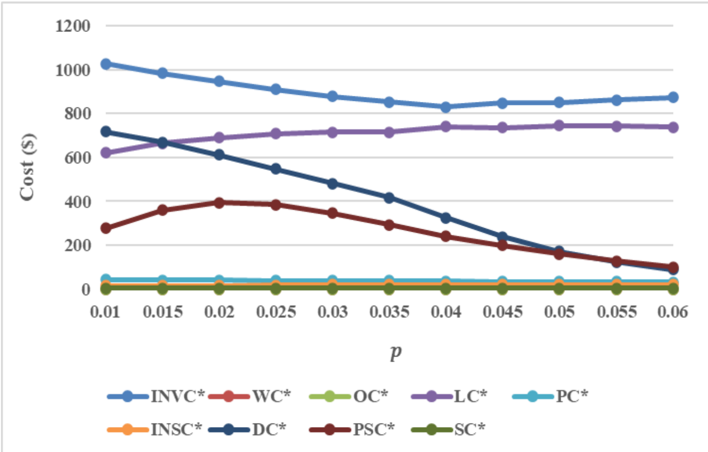


Fig. 9. The optimal values of the different components of the expected total cost versus p

Table 6. The optimal values with respect to changes in p

p	0.01	0.015	0.02	0.025	0.03	0.035	0.04	0.045	0.05	0.055	0.06
r^*	67	71	81	94	109	125	129	131	131	132	133
Q^*	364	344	318	289	260	232	220	221	221	222	223
$INVC^*$ ($)	1026.35	983.77	946.53	909.38	878.22	853.22	830.57	848.28	851.12	862.00	872.28
WC^* ($)	0.211	0.226	0.235	0.241	0.244	0.243	0.252	0.250	0.253	0.252	0.251
OC^* ($)	0.585	0.604	0.637	0.682	0.739	0.810	0.820	0.790	0.768	0.751	0.737
LC^* ($)	620.72	663.35	689.42	708.13	715.75	714.32	740.7258	736.23	744.93	742.00	737.50
PC^* ($)	42.60	41.53	40.49	39.43	38.45	37.57	36.08	34.93	33.96	33.32	32.88
$INSC^*$ ($)	15.28	15.76	16.62	17.81	19.30	21.14	21.41	20.63	20.06	19.59	19.25
DC^* ($)	717.71	669.90	612.11	547.91	481.10	416.11	325.61	237.95	172.56	123.97	88.41
PSC^* ($)	277.21	359.88	393.74	384.77	346.64	292.51	239.63	198.52	159.96	127.36	99.84
SC^* ($)	4	4	4	4	4	4	4	4	4	4	4
ETC^* ($)	2704.66	2739.01	2703.78	2612.35	2484.44	2339.93	2199.10	2081.58	1987.61	1913.25	1855.14

8. Fig. 10 shows the average outgoing quality (AOQ) vs. p. As expected, if the probability of defective items is very low or high, AOQ will be at a lower level. The maximum on the AOQ curve shows the worst possible average quality resulting from the inspection process, and this point is called the average outgoing quality limit (AOQL) (Montgomery [79]). Based on Fig. 10, The AOQL is 0.01449, which occurs in $p = 0.025$.
9. Based on Fig. 11, when p rises to 0.02, L_{def} ascends and then descends. Thus, for $p = .02$, L_{def} is maximum. Fig. 9 depicts the post-sale defective cost has the same behavior.
10. Fig. 13 demonstrates that if p becomes greater than 0.035, the percentage of time the retailer is waiting for special order is more than the percentage of time the retailer is waiting for regular orders and vice versa.
11. We fix the number of servers in the response system; hence the server cost remains unchanged when p rises.

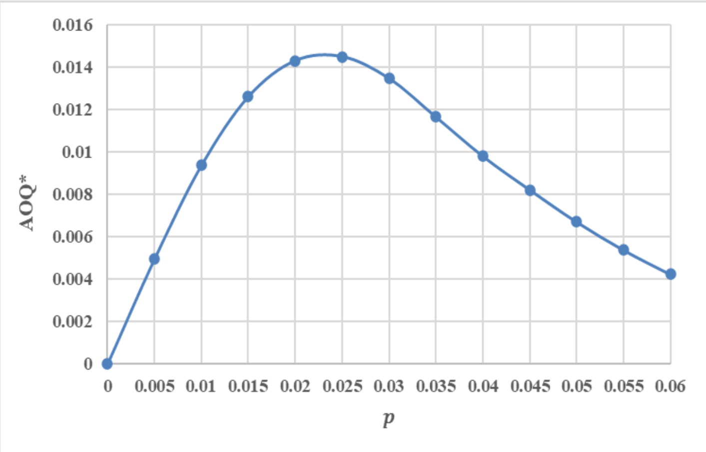


Fig. 10. The optimal average outgoing quality AOQ* versus p; $r = r^*$, $Q = Q^*$, $n = 89$, $c = 2$, $m = 2$

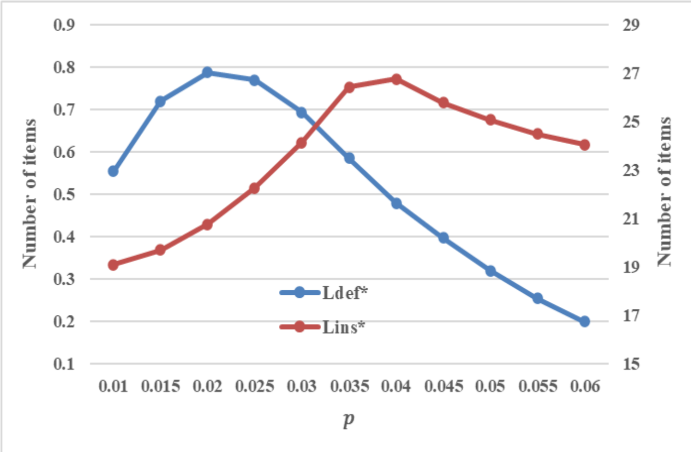


Fig. 11. The optimal average number of defective items is added to the retailer's inventory L^*_{def} and the optimal mean number of inspected items L^*_{ins} versus p; $r = r^*$, $Q = Q^*$, $n = 89$, $c = 2$, $m = 2$

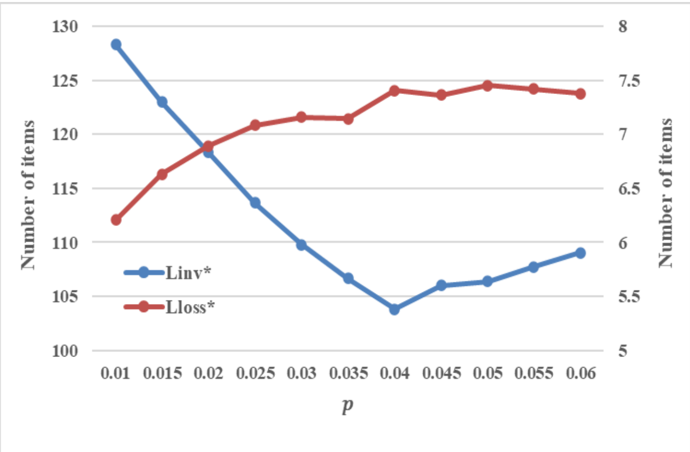


Fig. 12. The optimal expected on-hand inventory level of the retailers per time unit L^*_{inv} and the optimal average number of lost sales per time unit L^*_{loss} versus p; $r = r^*$, $Q = Q^*$, $n = 89$, $c = 2$, $m = 2$

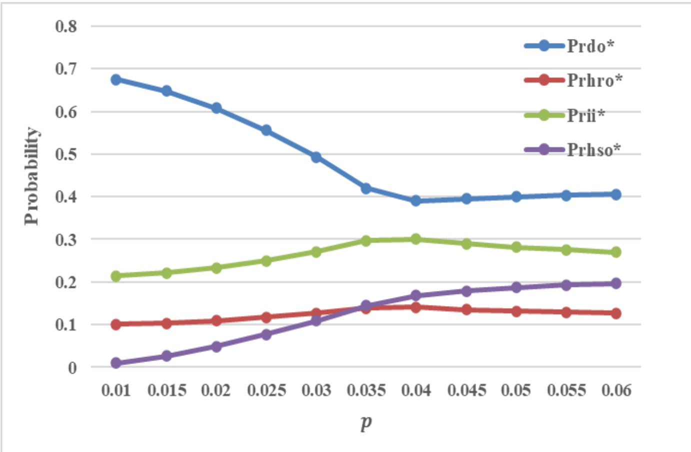


Fig. 13. The percentage of time that the retailer is in the different levels versus p; $r = r^*$, $Q = Q^*$, $n = 89$, $c = 2$, $m = 2$

Example 2. In this example, we reveal the impact of the inventory holding cost C_h, the waiting cost of unsatisfied demand C_W, the lost sale cost C_l and the post-sale defective item cost C_{pdi} on the optimal values of the reorder point and order quantity as well as the corresponding expected total cost. Toward this end, the values of variables as $n = 89$, $c = 2$, $m = 2$ and values of parameters as $\lambda = 40$ customers/hour, $\mu = 50$ customers/hour, $a = 50$ items/hour, $\nu_1 = 1.2$ orders/hour, $\nu_2 = 0.7$ orders/hour, $p = 0.02$, $C_{or} = 200$ \$/order, $C_p = 40$ \$/order, $C_{ins} = 0.8$ \$/item, $C_{des} = 40$ \$/item and $C_{msr} = 2$ \$/server/hour are fixed. From Table 7 and Figs. 14 to 16, we observe the following results:

1. As shown in Fig. 14, the optimal reorder point and order quantity decline as the inventory cost rises. It leads to a decrease in the inventory level, and increases lost sales costs. Finally, the optimal expected total cost will increase sharply. It should be noted that when the optimal reorder point decreases, we place an order later that causes decreasing the inventory level.
2. As the waiting cost for unmet demands increases in the response system, the optimal order quantity, reorder point, and expected total cost increase to reduce the percentage of times that inventory is unavailable. In addition, when the number of servers is less, the number of demands in the response system and total waiting time increase, so changes in order quantity and reorder point will be more sensitive to changes in waiting costs. For example, for m=1, by rising waiting cost, the mentioned variables increase more (see Table 7).
3. Table 7 demonstrates that for constant waiting cost, the optimal order quantity, reorder point, and expected total cost decrease when the number of servers in the response system rises. The higher the waiting cost, the greater the reduction. Because by increasing the number of the server, the waiting time of the demands reduces, and as a result, the higher the waiting cost, the more significant cost function saving.
4. When the lost sale cost increases, the optimal order quantity and reorder point increase; consequently, the stockout period and lost sales values decrease, and the total cost is minimized. The speed of rising the reorder point is greater than the rate of growing the order quantity. Also, the expected total cost has an ascending trend (see Fig. 15).
5. Fig. 16 shows that the optimal order quantity and reorder point slightly decrease as the cost of post-sale defective items increases. It is because by decreasing optimal order quantity and reorder point, the inventory level reduces, and accordingly, the number of defective items delivered to customers decreases. Hence, these values are less sensitive to the post-sale defective item cost. Also, the expected total cost rises.

Table 7. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, m, 2, 2)$ with respect to changes in C_w and m; $C_h = 8$, $C_l = 100$, $C_{pdi} = 500$

m	C_W	60	800	1540	2280	3020	3760
1	r^*	81	88	94	100	105	110
	Q^*	319	322	325	327	329	331
	ETC^*	2705.67724	2754.61421	2800.14795	2842.72115	2882.70387	2920.38625
2	r^*	81	81	81	82	82	83
	Q^*	318	318	319	319	319	319
	ETC^*	2703.77839	2706.67051	2709.55330	2712.41345	2715.26916	2718.12486

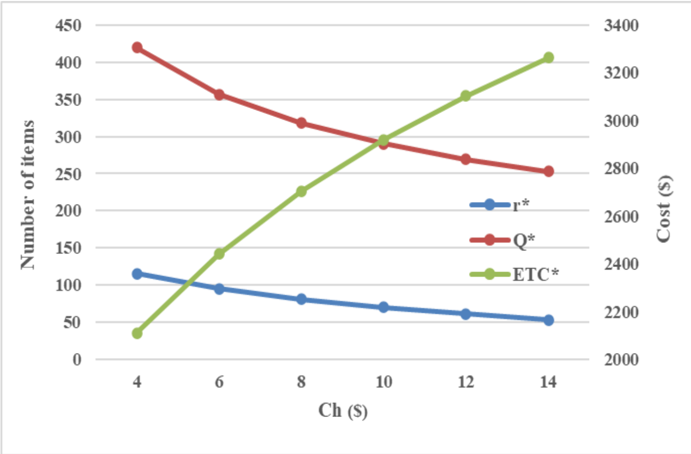


Fig. 14. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus C_h ; $C_W = 60$, $C_l = 100$, $C_{pdi} = 500$

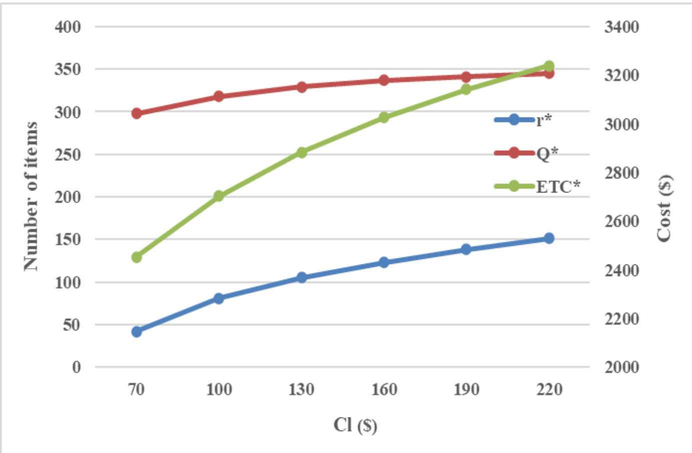


Fig. 15. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus C_l ;$C_h = 8$, $C_W = 60$, $C_{pdi} = 500$

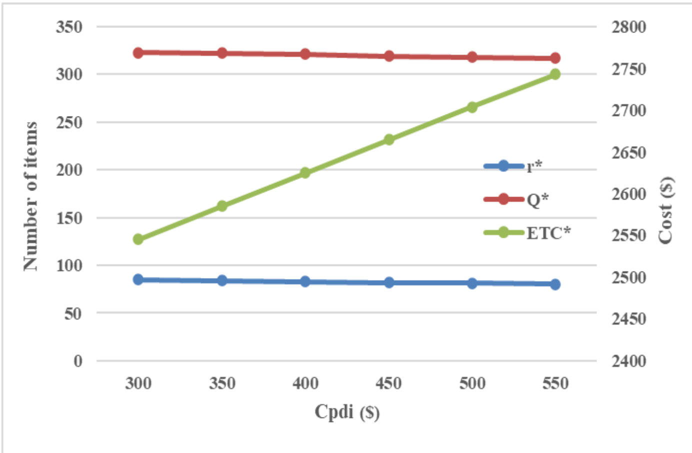


Fig. 16. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus C_{pdi} ;$C_h = 8$, $C_W = 60$, $C_l = 100$

Example 3. In this example, we investigate the impact of the demand arrival rate λ, the response system service rate μ, the regular order lead time rate ν_1, the special order lead time rate ν_2 and the screening rate of an item a on the optimal values of the reorder point and order quantity, as well as the corresponding expected total cost. For this purpose, the values of variables as $n = 89$, $c = 2$, $m = 2$ and values of parameters as $p = 0.02$, $C_{or} = 200$ \$/order, $C_p = 40$ \$/item, $C_{ins} = 0.8$ \$/item, $C_{des} = 40$ \$/item and $C_{msr} = 2$ \$/server/hour are fixed. From Figs. 17 to 20, we observe the following results:

1. According to Fig. 17, increasing the arrival rate of demand to system, the optimal reorder point, order quantity, and expected total cost rise. By increasing the reorder point, the retailer orders sooner, which causes to increase the inventory level and meets more demands. Moreover, clearly, the higher order quantity, the higher satisfied demands.
2. As shown in Fig. 18, whenever the response system's service rate increases up to 30, the optimal reorder point and order quantity decrease very little. After $\mu = 30$, these values almost remain constant. By increasing μ, the average waiting time of demands reduces, leading to a reduction in the expected total cost; consequently, the retailer can reduce the reorder point and order quantity mildly. But after $\mu = 30$, increasing service rate does not affect waiting time due to the fixed arrival rate, and the reorder point and the order quantity remain unchanged. Furthermore, when the service rate is low, increasing

it affects the optimal values considerably because it causes the waiting time to be significantly reduced, but the effect is gradually reduced. Therefore, with the increase of service rate, first, the optimal expected total cost decreases significantly, and the slope of reduction decreases and reaches a fixed value.

3. By increasing the regular order lead time rate ν_1, the optimal order quantity almost doesn't change, but the optimal reorder point declines, and the optimal expected total cost decreases (see Fig. 19). In fact, if the retailer receives the orders sooner, he can order later, which causes to decrease in the inventory level and expected total cost.
4. By increasing the special order lead time rate ν_2, we observe that the optimal order quantity declines, and the reorder point first ascends and then descends. Also, the expected total cost has a descending trend. It is because the orders are delivered sooner, and the lost sales are reduced (see Fig. 20).
5. Considering Fig. 21, if the inspection rate increases, the optimal order quantity decreases, and reorder point increases. Moreover, the expected total cost has a descending behavior. Rising the inspection rate will allow the inventory replenishment process to be sooner, and the stockout period will decrease. Overall, the waiting time of items before entering the retailer inventory system decrease that causes the total cost.

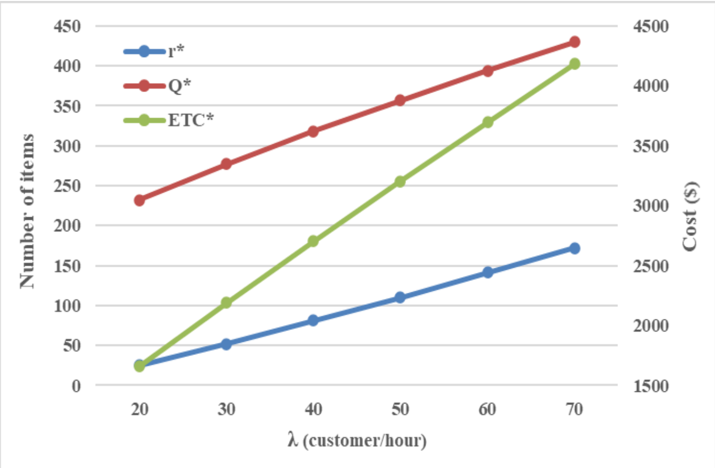


Fig. 17. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus λ ; $\mu = 50$, $a = 50$, $\nu_1 = 1.2$, $\nu_2 = 0.7$

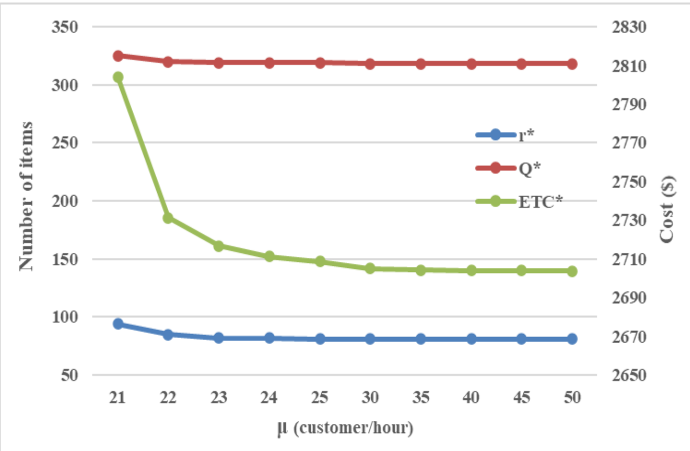


Fig. 18. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus μ ; $\lambda = 40$, $a = 50$, $\nu_1 = 1.2$, $\nu_2 = 0.7$

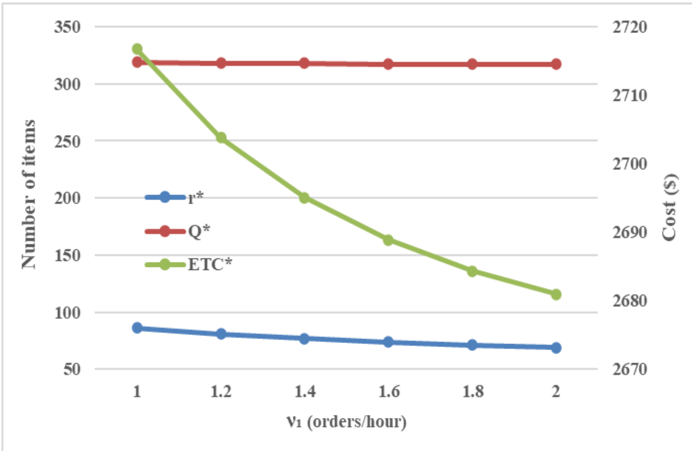


Fig. 19. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus ν_1 ; $\lambda = 40$, $\mu = 50$, $a = 50$, $\nu_2 = 0.7$

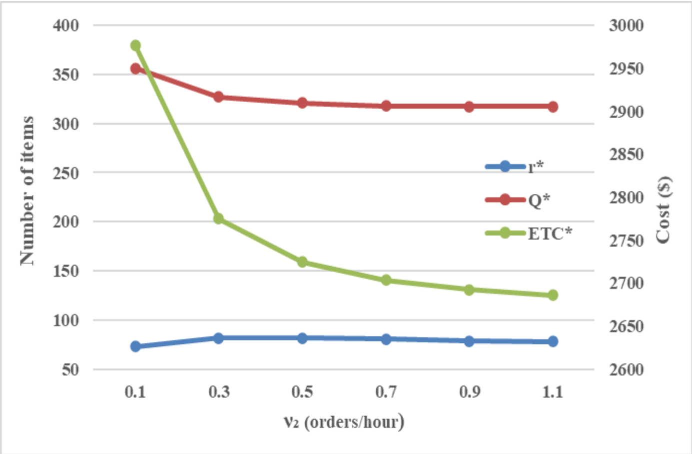


Fig. 20. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus ν_2 ; $\lambda = 40$, $\mu = 50$, $a = 50$, $\nu_1 = 1.2$

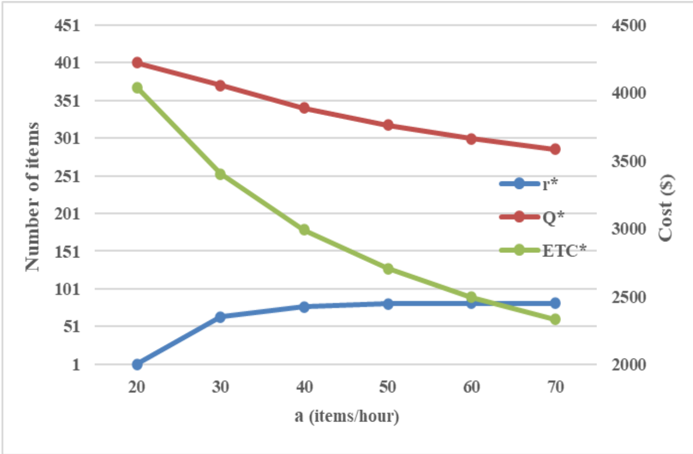


Fig. 21. The optimal reorder point r^*, order quantity Q^* and the expected total cost $ETC(r^*, Q^*, 89, 2, 2)$ versus a ; $\lambda = 40$, $\mu = 50$, $\nu_1 = 1.2$, $\nu_2 = 0.7$

Example 4. In this example, we examine the impact of the demand arrival rate λ, the regular order lead time rate ν_1, the special order lead time rate ν_2 and the screening rate a on the optimal values of the reorder point and order quantity, as well as the corresponding, expected total cost. For this purpose, the values of variables as $n = 89$, $c = 2$, $m = 2$ and values of parameters as $p = 0.02$, $C_{or} = 200$ \$/order, $C_p = 40$ \$/order, $C_{ins} = 0.8$ \$/item, $C_{des} = 40$ \$/item and $C_{msr} = 2$ \$/server/hour are fixed. From Tables 8 to 10 and Figs. 17 to 20, we observe the following results:

1. According to Table 8, the optimal average on-hand inventory, the number of lost sales, the number of defective items added to the retailer's inventory, and the number of inspected items increase when the arrival demand rises. Because based on Fig. 17, the optimal reorder point and order quantity rise. Fig. 22 demonstrates that the percentage of time that the retailer doesn't have any outstanding order reduces by increasing the arrival rate of demand because he orders sooner. As a result, the sojourn time in the other levels increases. As shown in Fig. 22, the percentage of time the retailer process is in the inspection mode is more sensitive to the demand rate than the waiting states for the order to arrive. In addition, lost sales increase as demand increases.
2. We observe in Table 9 that by increasing the delivery rate of regular orders, the optimal inventory level and the number of lost sales decreases slightly, and the optimal expected number of defective items added to the retailer's inventory and the number of inspected items slightly rises. Moreover, the percentage of time that the retailer doesn't have any outstanding order rises, and the percentage of time that the retailer has a regular outstanding order reduces. Besides, the sojourn time in the other levels mildly increases (see Fig. 23).
3. We observe in Table 10 that by increasing the delivery rate of special orders, the optimal inventory level, the number of inspected items, and the number of defective items added to the retailer's inventory rise, and the number of lost sales declines. According to Fig. 24, as expected, the percentage of time that the retailer has a special outstanding order reduces, and the sojourn time in the other levels increases.
4. Based on Table 11, as the inspection rate of items increases, the average inventory increases, and lost sales decrease. Also, the average defective items added to the inventory and the inspected items decrease because, according to Fig. 25, the percentage of time the retailer is inspecting reduces significantly. The percentage of time the system is in other states increases slightly.

Table 8. Some optimal performance measures with respect to changes in λ; $r = r^*$, $Q = Q^*$, $\mu = 50$, $a = 50$, $\nu_1 = 1.2$, $\nu_2 = 0.7$

λ (customers/hour)	20	30	40	50	60	70
L_{inv}^* (items)	73.90268	96.1659	118.3166	139.8944	161.8818	183.5445
L_{loss}^* (customers)	3.639022	5.262993	6.894172	8.608867	10.27031	11.97983
L_{def}^* (items)	0.36859	0.576736	0.787485	0.998055	1.210938	1.423542
L_{ins}^* (items)	15.57219	18.53363	20.77535	22.4989	23.98632	25.22072

Table 9. Some optimal performance measures with respect to changes in ν_1; $r = r^*$, $Q = Q^*$, $\lambda = 40$, $\mu = 50$, $a = 50$, $\nu_2 = 0.7$

ν_1 (orders/hour)	1	1.2	1.4	1.6	1.8	2
L_{inv}^* (items)	118.385	118.3166	118.4093	118.01	117.7206	117.7196
L_{loss}^* (customers)	7.10832	6.894172	6.75788	6.681852	6.651982	6.604202
L_{def}^* (items)	0.782707	0.787485	0.790727	0.792213	0.792923	0.794059
L_{ins}^* (items)	20.55952	20.77535	20.86088	20.99175	21.01057	21.04067

Table 10. Some optimal performance measures with respect to changes in ν_1; $r = r^*$, $Q = Q^*$, $\lambda = 40$, $\mu = 50$, $a = 50$, $\nu_1 = 1.2$

ν_2 (orders/hour)	0.1	0.3	0.5	0.7	0.9	1.1
L_{inv}^* (items)	104.7063	115.3782	117.8038	118.3166	118.2018	118.6237
L_{loss}^* (customers)	13.44092	8.572536	7.366611	6.894172	6.710434	6.565109
L_{def}^* (items)	0.64022	0.750197	0.777181	0.787485	0.791533	0.794989
L_{ins}^* (items)	14.48638	19.04319	20.23837	20.77535	20.97374	21.0653

Table 11. Some optimal performance measures with respect to changes in a; $r = r^*$, $Q = Q^*$, $\lambda = 40$, $\mu = 50$, $\nu_1 = 1.2$, $\nu_1 = 0.7$

a (items/hour)	20	30	40	50	60	70
L_{inv}^* (items)	101.9876	118.6316	120.0355	118.3166	115.627	113.0521
L_{loss}^* (customers)	16.07516	10.468	8.182998	6.894172	6.090088	5.536525
L_{def}^* (items)	1.458785	1.19159	0.954122	0.787485	0.666893	0.576855
L_{ins}^* (items)	28.24738	25.52814	22.87415	20.77535	19.0948	17.69056

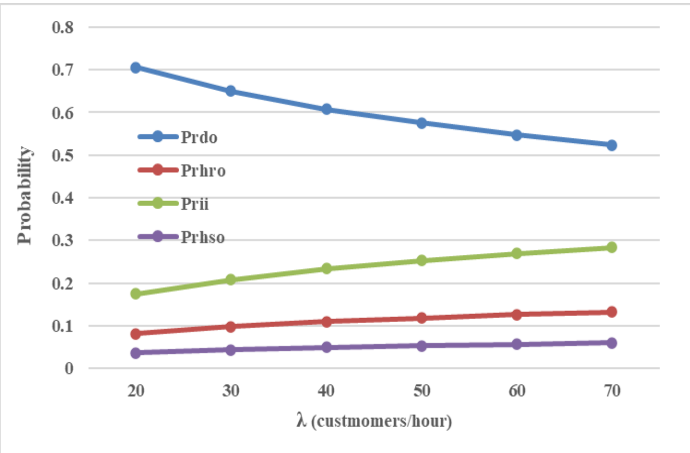


Fig. 22. The percentage of time that the retailer is in the different levels versus λ; $r = r^*$, $Q = Q^*$, $\mu = 50$, $a = 50$, $\nu_1 = 1.2$, $\nu_2 = 0.7$

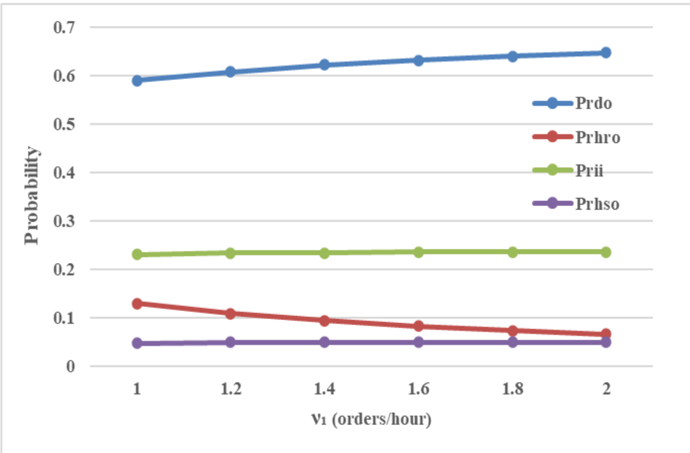


Fig. 23. The percentage of time that the retailer is in the different levels versus ν_1; $r = r^*$, $Q = Q^*$, $\lambda = 40$, $\mu = 50$, $a = 50$, $\nu_2 = 0.7$

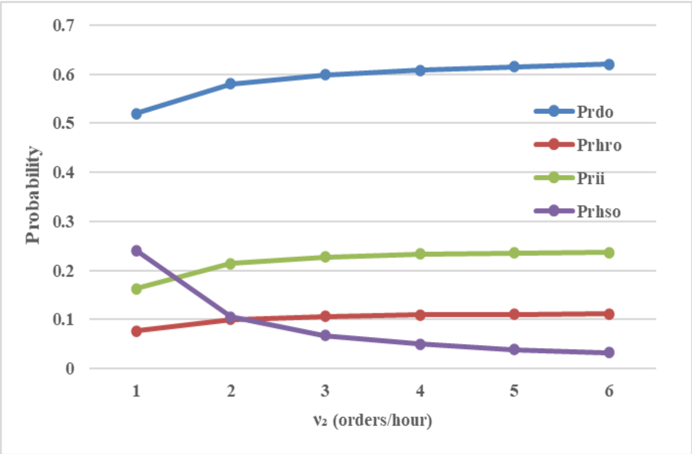


Fig. 24. The percentage of time that the retailer is in the different levels versus ν_1; $r = r^*$, $Q = Q^*$, $\lambda = 40$, $\mu = 50$, $a = 50$, $\nu_1 = 1.2$

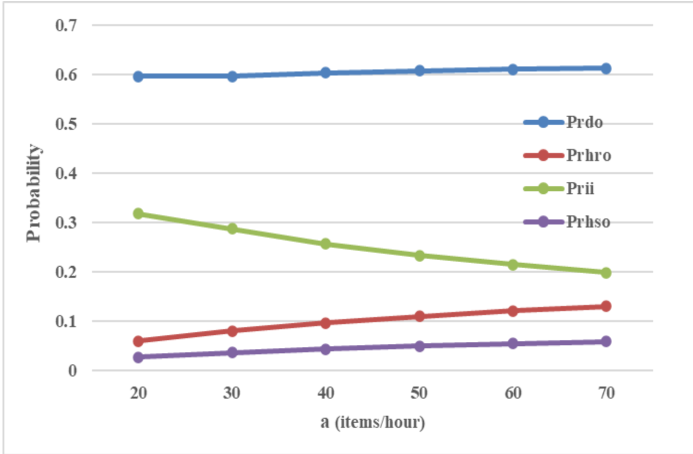


Fig. 25. The percentage of time that the retailer is in the different levels versus a; $r = r^*$, $Q = Q^*$, $\lambda = 40$, $\mu = 50$, $\nu_1 = 1.2$, $\nu_2 = 0.7$

Discussion and managerial insights

This paper studies an M/M/m queueing-inventory model for a retailer-supplier system with defective items and destructive testing acceptance sampling. Considering defective items in the queueing-inventory model is more practical and realistic, which has not been included in the previous researches in this area, such as [17] , [19] , [22] , etc., except to [47] , but they proposed an M/M/1 queueing-inventory model with defective items in the observable case and not for a retailer-supplier problem. Hence, in such circumstances, conducting an inspection process for incoming items to discover defective items is rational. The 100% screening process can be costly and time-consuming rather than acceptance sampling plans and isn’t an appropriate method for destructive tests. So, we considered a destructive testing acceptance sampling plan while Aghsami et al. [47] applied a 100% inspection process to detect defective items. Moreover, it is more realistic to assume the retailer's demands first arrive at a response system for inquiring, processing, etc. Hence, we proposed a response system for the retailer equipped with multiple servers. It works as a queueing system that has not been considered in previous studies related to retailer-supplier problems. Also, such a system with this new point of view isn’t addressed in the queueing-inventory literature.

On the other hand, although numerous researches have discussed defective items in the literature of retailer-supplier inventory models, no one integrates these models with the arrival demands queueing system. Considering that the retailer's order status and inventory level can be in each possible state at any point in time, it is more practical to look at the retailer's inventory system as a stochastic process. Accordingly, this paper modeled a retailer-supplier system with the properties mentioned above, considering destructive testing acceptance-sampling as CTMC model, which has not been addressed in [50] , [51] , [58] , etc. Therefore, our model has been presented to fulfill these research gaps.

The impact of different parameters on the optimal reorder point, order quantity, expected total cost, and performance measures was assessed to achieve a thorough perception of the proposed model behavior in the previous section. Consequently, based on the result of the numerical examples and sensitive analysis, some managerial implications can be presented as follows:

1. According to Table 5 and Fig. 8, if the percentage of defective items in the incoming lots from the supplier increases up to the trending change point, the retailer should

decrease order quantity and increase reorder point to minimize the expected total cost. After that, he should slightly raise both of them.

2. Considering Table 5, when the probability of defective items rises, the waiting time of demands grows, leading to increase demands dissatisfaction level and lost sales. Also, the purchased and destructive testing costs reduce. Hence, the retailer can spend this savings on advertising and even supplying items from other sources.
3. Taking Fig. 13 into account, the percentage of time that the retailer is doing inspection reduces mildly after the trending change point. So, in this case, the retailer can employ several laborers from the other departments' inspection section.
4. Based on Table 6, by increasing the probability of defective items, the lost sale has an Incremental trend to reach a high level; therefore, our model suggests that the retailer should supply some demands from other suppliers.
5. Based on Figs. 10 and 11, the retailer should focus on the inspection system for $0.02 \leq p \leq 0.025$ and even use a tightened inspection in this case.
6. According to Fig. 13, in case $0.035 \leq p \leq 0.04$, The percentage of times the process is in the inspection level is at a maximum. The retailer is advised to allocate as much manpower as possible to the inspection department.
7. If the holding inventory cost increases, the retailer should order later with less quantity to minimize the expected total cost (see Fig. 14).
8. According to Table 7, if the waiting cost of unsatisfied demand increases, the retailer should increase the reorder point and order quantity. This increase must be based on the number of servers in the response system. For example, if the number of servers is low so that the formed queue is long in the response system, the retailer should increase the order quantity and reorder point further when the waiting cost rises.
9. According to Table 7, the retailer can reduce the expected total cost by increasing the number of servers to a sufficient extent. The higher the waiting cost, the more influential the number of servers in reducing the total cost. Provided that the retailer has limited storage space or the holding cost is high, he can decrease the reorder point and order quantity by increasing the number of servers if possible.
10. If the lost sales cost increases, the retailer should increase the reorder point and the order quantity. According to Fig. 15, it is better to increase the reorder point more than the order quantity.
11. Based on Fig. 16, our model suggests that the retailer should mildly decline the reorder point and order quantity whenever the post-sale defective item cost rises.
12. According to Fig. 17, the retailer should order sooner and more by increasing the system's demand.
13. According to Fig. 18, The retailer is advised to keep the service rate optimal to minimize the expected total cost. If he increases the service rate more than a certain level, it will not significantly affect his cost. Hence, the retailer can decrease laborers in the response system until the service rate reaches an optimal level and employ them in the other section. On the other hand, if it is not possible to increase the service rate, he can compensate for the lack of laborers in the response system by increasing the reorder point and order quantity.
14. Provided the regular order lead time rate increases, based on Fig. 19, our model recommends that the retailer decline the reorder point to minimize the total cost. Furthermore, he can reduce a bit the order quantity.
15. According to Fig. 20, it is recommended to the retailer that if the delivery time of a special order is reduced, he should order later in a smaller quantity, reducing the cost of inventory and purchase.

16. If the retailer can increase the inspection speed, it can reduce the order amount, but he must raise the re-order point, which reduces the average total cost. Therefore, if he wants to reduce the order quantity, he must employ more laborers in the inspection system, ultimately reducing the average total cost (see Fig. 21).
17. Our model recommends to the retailer that if the arrival rate of demand increases, he should concentrate on the inspection system more and allocate more storage space for holding inventories (see Table 8).
18. Table 9 and Fig. 23 show that if the retailer chooses a supplier with less lead time, he can order later. The percentage of time that doesn't have any outstanding order increases. Consequently, he stores less inventory.
19. If the retailer selects a supplier with less lead time, he can have fewer lost sales (see Table 10).
20. According to Table 11, one-way retailers can increase average inventory and reduce lost sales is to increase inspection speed.

Conclusion and future direct

This article studied the RSQIP where the supplier sends the imperfect lots to the retailer. He continually checks the inventory and places an order Q when the inventory level reaches the reorder point r. The order delivers to the retailer after an exponential time that we called regular order. Also, he uses a destructive testing acceptance sampling to accept or reject the lots that the inspection time follows an exponential distribution and is dependent on sampling size. If a lot is rejected, the retailer asks the supplier to send a lot without defective items. This order delivers to the retailer after an exponential time that we called special order. The demands arrive at the retailer's response system according to a Poisson process. There are several servers in the response system to handle the arrival demands, and each one needs an exponential time to respond to them. Hence, a queue of requests is formed in the response system that acts as an M/M/m queueing system. After completing service in the response system, the demands leave the system with precisely one item provided the inventory is available; otherwise, lost sales occur. We derived the stationary distributions of the number of demands in the response system and the joint stationary distribution of the order status and the retailer's inventory level. Afterward, some key performance measures have been presented, and a cost model has been developed in steady-state. We proposed a non-linear integer programming model to minimize the long-run expected total cost concerning the reorder point, order quantity, number of servers, acceptance number, and sampling size. We have presented comprehensive numerical examples to demonstrate the expected total cost behavior and the impact of various key parameters on optimal order quantity, reorder point, and the expected total cost. Finally, based on the sensitivity analysis, we recommended some practical managerial insights. To the best of our knowledge, this is the first time a retailer-supplier inventory system considering defective items and destructive testing acceptance sampling is integrated with a queueing system and modeled as CTMC, which is more practical and realistic.

As future researches, the presented model can be extended for other ordering policies, e.g., (R, T) policy. Studying an imperfect internal production besides the imperfect supplier could be an interesting issue for future researches. Moreover, developing the model to multi-retailer and multi-supplier is another suggestion for the future. In addition, this model could be extended for general arrival, service rate, or lead time. Also, the 100% screening process can be performed in the inspection system instead of the acceptance sampling plan. Moreover, in this paper, we assumed the servers are reliable in the retailer's system while they may fail at any point in time, which may cease the service entirely or lead to the service continuing at a slower rate [81] . Therefore, studying the server breakdowns in the response system would be an interesting problem for future research.

References

[1] Saffari, M., Haji, R. and Hassanzadeh, F., 2011. A queueing system with inventory and mixed exponentially distributed lead times. *The International Journal of Advanced Manufacturing Technology*, 53(9-12), pp.1231-1237.

[2] Jeganathan, K., Reiyas, M.A., Padmasekaran, S. and Lakshmanan, K., 2017. An $$ M/E_{K}/1/N $$ M/EK/1/N Queueing-Inventory System with Two Service Rates Based on Queue Lengths. *International Journal of Applied and Computational Mathematics*, 3(1), pp.357-386.

[3] Schwarz, M., Sauer, C., Daduna, H., Kulik, R. and Szekli, R., 2006. M/M/1 queueing systems with inventory. *Queueing Systems*, 54(1), pp.55-78.

[4] Yue, D. and Qin, Y., 2019. A production inventory system with service time and production vacations. *Journal of Systems Science and Systems Engineering*, *28*(2), pp.168-180.

[5] Naimi Sadigh, A., Chaharsooghi, S.K. and Sheikhmohammady, M., 2016. Game-theoretic analysis of coordinating pricing and marketing decisions in a multi-product multi-echelon supply chain. *Scientia Iranica*, *23*(3), pp.1459-1473.

[6] Mohtashami, Z., Aghsami, A. and Jolai, F., 2020. A green closed loop supply chain design using queuing system for reducing environmental impact and energy consumption. *Journal of cleaner production*, *242*, p.118452.

[7] Zokaee, M., Nazari, A., Aghsami, A. and Jolai, F., 2021. An inventory system with coordination among manufacturers and retailers under buyback contract, vertical integration, retailer's effort and carbon footprint constraint. *International Journal of Sustainable Engineering*, pp.1-21.

[8] Mokhtari, H., Asadkhani, J. (2019). 'Economic Order Quantity for Imperfect Quality Items Under Inspection Errors, Batch Replacement and Multiple Sales of Returned Items', *Scientia Iranica*, doi: 10.24200/sci.2019.52075.2520

[9] Gavish, B. and Graves, S.C., 1981. Production/inventory systems with a stochastic production rate under a continuous review policy. *Computers & Operations Research*, *8*(3), pp.169-183.

[10] Sigman, K. and Simchi-Levi, D., 1992. Light traffic heuristic for an M/G/1 queue with limited inventory. *Annals of Operations Research*, *40*(1), pp.371-380.

[11] Berman, O., Kaplan, E.H. and Shevishak, D.G., 1993. Deterministic approximations for inventory management at service facilities. *IIE transactions*, *25*(5), pp.98-104.

[12] Berman, O. and Sapna, K.P., 2000. Inventory management at service facilities for systems with arbitrarily distributed service times. *Stochastic Models*, 16(3-4), pp.343-360.

[13] Schwarz, M. and Daduna, H., 2006. Queueing systems with inventory management with random lead times and with backordering. *Mathematical Methods of Operations Research*, *64*(3), pp.383-414.

[14] Chang, K.H. and Lu, Y.S., 2010. Queueing analysis on a single-station make-to-stock/make-to-order inventory-production system. *Applied Mathematical Modelling*, *34*(4), pp.978-991.

[15] Zhao, N. and Lian, Z., 2011. A queueing-inventory system with two classes of customers. *International Journal of Production Economics*, *129*(1), pp.225-231.

[16] Krishnamoorthy, A. and Viswanath, N.C., 2013. Stochastic decomposition in production inventory with service time. *European Journal of Operational Research*, *228*(2), pp.358-366.

[17] Saffari, M., Asmussen, S. and Haji, R., 2013. The M/M/1 queue with inventory, lost sale, and general lead times. Queueing Systems, 75(1), pp.65-77.

[18] Sivashankari, C.K. and Panayappan, S., 2014. Production inventory model for two levels production with defective items and incorporating multi-delivery policy. *International Journal of Operational Research*, *19*(3), pp.259-279.

[19] 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*, pp.534-544.

[20] Krishnamoorthy, A., Manikandan, R. and Lakshmy, B., 2015. A revisit to queueing-inventory system with positive service time. *Annals of Operations Research*, *233*(1), pp.221-236.
[21] 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(1), pp.14-24.
[22] Manikandan, R. and NAIR, S.S., 2017. M/M/1/1 queueing-inventory system with retrial of unsatisfied customers. *Communications in Applied Analysis, 21*(2), pp.217-236.
[23] Baek, J.W., Bae, Y.H., Lee, H.W. and Ahn, S., 2018. Continuous-type (s, Q)-inventory model with an attached M/M/1 queue and lost sales. *Performance Evaluation*, *125*, pp.68-79.
[24] Yue, D., Zhao, G. and Qin, Y., 2018. An M/M/1 Queueing-Inventory System with Geometric Batch Demands and Lost Sales. *Journal of Systems Science and Complexity*, 31(4), pp.1024-1041.
[25] Saffari, M., Sajadieh, M.S. and Hassanzadeh, F., 2019. A queuing system with inventory and competing suppliers. *European Journal of Industrial Engineering*, 13(3), pp.420-433.
[26] Shajin, D. and Krishnamoorthy, A., 2020. Stochastic decomposition in retrial queueing-inventory system. *RAIRO-Operations Research*, 54(1), pp.81-99.
[27] Shajin, D., Krishnamoorthy, A. and Manikandan, R., 2020. On a queueing-inventory system with common life time and Markovian lead time process. *Operational Research*, pp.1-34.
[28] Chakravarthy, S. and Hayat, K., 2020. Queueing-Inventory Models for a Two-Vendor System with Positive Service Times. *Queueing Models and Service Management*, *3*(1), p.1.
[29] M Manikandan, R. and Nair, S.S., 2020. An M/M/1 Queueing-Inventory System with Working Vacations, Vacation Interruptions and Lost Sales. *Automation and Remote Control*, 81, pp.746-759.
[30] Shajin, D., Jacob, J. and Krishnamoorthy, A., 2021. On a queueing inventory problem with necessary and optional inventories. *Annals of Operations Research*, pp.1-26.
[31] Ozkar, S. and Kocer, U.U., 2021. Two-commodity queueing-inventory system with two classes of customers. *OPSEARCH*, *58*(1), pp.234-256.
[32] Jeganathan, K., Selvakumar, S., Anbazhagan, N., Amutha, S. and Hammachukiattikul, P., 2021. Stochastic modeling on M/M/1/N inventory system with queue-dependent service rate and retrial facility. *AIMS Mathematics*, *6*(7), pp.7386-7420.
[33] Graves, S.C., 1982. The application of queueing theory to continuous perishable inventory systems. *Management Science*, *28*(4), pp.400-406.
[34] Manuel, P., Sivakumar, B. and Arivarignan, G., 2008. A perishable inventory system with service facilities and retrial customers. Computers & Industrial Engineering, 54(3), pp.484-501.
[35] Karthick, T., Sivakumar, B. and Arivarignan, G., 2015. An inventory system with two types of customers and retrial demands. *International Journal of Systems Science: Operations & Logistics*, *2*(2), pp.90-112.
[36] Albrecher, H., Boxma, O.J., Essifi, R. and Kuijstermans, R., 2017. A queueing model with randomized depletion of inventory. *Probability in the Engineering and Informational Sciences*, *31*(1), pp.43-59.
[37] Chakravarthy, S.R., Maity, A. and Gupta, U.C., 2017. An '(s, S)'inventory in a queueing system with batch service facility. *Annals of Operations Research*, *258*(2), pp.263-283.
[38] Marand, A.J., Li, H. and Thorstenson, A., 2019. Joint inventory control and pricing in a service-inventory system. *International Journal of Production Economics*, *209*, pp.78-91.
[39] Melikov, A., Krishnamoorthy, A. and Shahmaliyev, M., 2018. Perishable queuing inventory systems with delayed feedback. In *Information Technologies and Mathematical Modelling. Queueing Theory and Applications* (pp. 55-70). Springer, Cham.
[40] Melikov, A.Z. and Shahmaliyev, M.O., 2019. Queueing System M/M/1/∞ with Perishable Inventory and Repeated Customers.*Automation and Remote Control*, 80(1), pp.53-65.
[41] Gowsalya, V., Selvakumar, C. and Elango, C., 2019. Finite Source Retrial Queue with Inventory Management: *Semi MDP.Journal of Computer and Mathematical Sciences*, 10(5), pp.1032-1042.
[42] Hanukov, G., Avinadav, T., Chernonog, T. and Yechiali, U., 2020. A multi-server system with inventory of preliminary services and stock-dependent demand. *International Journal of Production Research*, pp.1-19.

[43] Anilkumar, M.P. and Jose, K.P., 2020. A Geo/Geo/1 inventory priority queue with self-induced interruption. *International Journal of Applied and Computational Mathematics*, *6*(4), pp.1-14.

[44] Keerthana, M., Saranya, N. and Sivakumar, B., 2020. A stochastic queueing-inventory system with renewal demands and positive lead time. *European Journal of Industrial Engineering*, *14*(4), pp.443-484.

[45] Krishnamoorthy, A., Joshua, A.N. and Kozyrev, D., 2021. Analysis of a Batch Arrival, Batch Service Queuing-Inventory System with Processing of Inventory While on Vacation. *Mathematics*, *9*(4), p.419.

[46] Rasmi, K. and Jacob, M.J., 2021. Analysis of a multiserver queueing inventory model with self-service. *International Journal of Mathematical Modelling and Numerical Optimisation*, *11*(3), pp.275-291.

[47] Aghsami, A., Samimi, Y. and Aghaei, A., 2021. A novel Markovian queueing-inventory model with imperfect production and inspection processes: a hospital case study. *Computers & Industrial Engineering*, p.107772.

[48] Salameh, M.K. and Jaber, M.Y., 2000. Economic production quantity model for items with imperfect quality. *International journal of production economics*, *64*(1-3), pp.59-64.

[49] Khan, M., Jaber, M.Y. and Wahab, M.I.M., 2010. Economic order quantity model for items with imperfect quality with learning in inspection. *International journal of production economics*, *124*(1), pp.87-96.

[50] Al-Salamah, M., 2011. Economic order quantity with imperfect quality, destructive testing acceptance sampling, and inspection errors. *Advances in Management and Applied Economics*, *1*(2), p.59.

[51] Moussawi-Haidar, L., Salameh, M. and Nasr, W., 2013. An instantaneous replenishment model under the effect of a sampling policy for defective items. *Applied Mathematical Modelling*, *37*(3), pp.719-727.

[52] Hsu, J.T. and Hsu, L.F., 2013. Two EPQ models with imperfect production processes, inspection errors, planned backorders, and sales returns. Computers & Industrial Engineering, 64(1), pp.389-402.

[53] Moussawi-Haidar, L., Salameh, M. and Nasr, W., 2014. Effect of deterioration on the instantaneous replenishment model with imperfect quality items. *Applied Mathematical Modelling*, *38*(24), pp.5956-5966.

[54] Priyan, S. and Uthayakumar, R., 2015. Mathematical modeling and computational algorithm to solve multi-echelon multi-constraint inventory problem with errors in quality inspection. *Journal of Mathematical Modelling and Algorithms in Operations Research*, *14*(1), pp.67-89.

[55] Chang, C.T., Cheng, M.C. and Soong, P.Y., 2016. Impacts of inspection errors and trade credits on the economic order quantity model for items with imperfect quality. *International Journal of Systems Science: Operations & Logistics*, 3(1), pp.34-48.

[56] Hasanpour Rodbaraki, J. and Sharifi, E., 2016. Economic Order Quantity for Deteriorating Items with Imperfect Quality, Destructive Testing Acceptance Sampling, and Inspection Errors. *Advances in Industrial Engineering*, *50*(2), pp.235-246.

[57] Cheikhrouhou, N., Sarkar, B., Ganguly, B., Malik, A.I., Batista, R. and Lee, Y.H., 2018. Optimization of sample size and order size in an inventory model with quality inspection and return of defective items. *Annals of Operations Research*, *271*(2), pp.445-467.

[58] Maleki Vishkaei, B., Niaki, S.T.A., Farhangi, M. and Mahdavi, I., 2019. A single-retailer multi-supplier multi-product inventory model with destructive testing acceptance sampling and inflation. *Journal of Industrial and Production Engineering*, *36*(6), pp.351-361.

[59] Wangsa, I.D. and Wee, H.M., 2019. A vendor-buyer inventory model for defective items with errors in inspection, stochastic lead time and freight cost. *INFOR: Information Systems and Operational Research*, 57(4), pp.597-622.

[60] Safarnezhad, M., Aminnayeri, M. and Ghasemy Yaghin, R., 2020. Joint pricing and lot sizing model with statistical inspection and stochastic lead time. *INFOR: Information Systems and Operational Research*, pp.1-34.

[61] Taleizadeh, A.A., 2021. Imperfect Inventory Systems: Inventory and Production Management. *Springer Nature*.
[62] Asadkhani, J., Mokhtari, H. and Tahmasebpoor, S., 2021. Optimal lot-sizing under learning effect in inspection errors with different types of imperfect quality items. *Operational Research*, pp.1-35.
[63] Wu, K.S., Ouyang, L.Y. and Ho, C.H., 2007. Integrated vendor-buyer inventory system with sublot sampling inspection policy and controllable lead time. *International Journal of Systems Science*, *38*(4), pp.339-350.
[64] Jolai, F., Gheisariha, E. and Nojavan, F., 2011. Inventory Control of Perishable Items in a Two-Echelon Supply Chain. *Advances in Industrial Engineering*, *45*(Special Issue), pp.69-77.
[65] Datta, T.K., 2013. An inventory model with price and quality dependent demand where some items produced are defective. *Advances in Operations Research*, *2013*.
[66] Taleizadeh, A. and Hasani, M., 2015. A multi-product inventory control model with imperfect production process and rework under delayed payment policy. *Advances in Industrial Engineering*, *49*(2), pp.223-235.
[67] Jauhari, W.A., Sofiana, A., Kurdhi, N.A. and Laksono, P.W., 2016. An integrated inventory model for supplier-manufacturer-retailer system with imperfect quality and inspection errors. *International Journal of Logistics Systems and Management*, *24*(3), pp.383-407.
[68] Pal, S. and Mahapatra, G.S., 2017. A manufacturing-oriented supply chain model for imperfect quality with inspection errors, stochastic demand under rework and shortages. *Computers & Industrial Engineering*, *106*, pp.299-314.
[69] Amirhosseini, K., Mohammadi, M. and Pasandideh, S.H., 2017. Modeling and Solving a Multi-product Economic Order Quantity Problem with Imperfect Items and Emergency Buy or Repair Options. *Advances in Industrial Engineering*, *51*(4), pp.405-416.
[70] Manna, A.K., Dey, J.K. and Mondal, S.K., 2017. Imperfect production inventory model with production rate dependent defective rate and advertisement dependent demand. *Computers & Industrial Engineering*, 104, pp.9-22.
[71] Taheri-Tolgari, J., Mohammadi, M., Naderi, B., Arshadi-Khamseh, A. and Mirzazadeh, A., 2019. An inventory model with imperfect item, inspection errors, preventive maintenance and partial backlogging in uncertainty environment. *Journal of Industrial & Management Optimization*, 15(3), pp.1317-1344.
[72] Sarkar, S. and Giri, B.C., 2020. Stochastic supply chain model with imperfect production and controllable defective rate. *International Journal of Systems Science: Operations & Logistics*, *7*(2), pp.133-146.
[73] Karakatsoulis, G. and Skouri, K., 2021. Optimal reorder level and lot size decisions for an inventory system with defective items. *Applied Mathematical Modelling*, *92*, pp.651-668.
[74] Adak, S. and Mahapatra, G.S., 2021. Two-echelon imperfect production supply chain with probabilistic deterioration rework and reliability under fuzziness. *Journal of Management Analytics*, pp.1-25.
[75] Khan, M., Jaber, M.Y. and Bonney, M., 2011. An economic order quantity (EOQ) for items with imperfect quality and inspection errors. *International Journal of Production Economics*, 133(1), pp.113-118.
[76] Ghosh, P.K., Manna, A.K. and Dey, J.K., 2017. Deteriorating manufacturing system with selling price discount under random machine breakdown. *Int. J. Comput. Eng. Manage*, 20, pp.8-17.
[77] World Health Organization, 2005. WHO guidelines for sampling of pharmaceutical products and related materials. *WHO Technical Report Series*, *929*, pp.59-93.
[78] Ross, S.M., 2014. *Introduction to probability models*. Academic press.
[79] Montgomery, D.C., 2009. *Introduction to statistical quality control*. John Wiley & Sons.
[80] Shortle, J.F., Thompson, J.M., Gross, D. and Harris, C.M., 2018. Fundamentals of queueing theory.
[81] Aghsami, A. and Jolai, F., 2020. Equilibrium threshold strategies and social benefits in the fully observable Markovian queues with partial breakdowns and interruptible

setup/closedown policy. *Quality Technology & Quantitative Management*, *17*(6), pp.685-722.