Vol. 34, No. 4 (2024) | DOI: 10.37190/ord240413

OPEN ACCESS

Operations Research and Decisions

www.ord.pwr.edu.pl

OPERATIONS RESEARCH AND DECISIONS QUARTERLY

A queueing model for an automatic manufacturing system with disasters, breakdowns, and vacations. Optimal design and analysis

Nada Riheb Yatim[1] Amina Angelika Bouchentouf[2*] Pikkala Vijaya Laxmi[3]

[1] *Laboratory of Stochastic Models, Statistic and Applications, University of Saida-Dr. Moulay Tahar, Saida, Algeria*
[2] *Laboratory of Mathematics, Djillali Liabes University of Sidi Bel Abbes, 22000 Sidi Bel Abbes, Algeria*
[3] *Department of Applied Mathematics, Andhra University, Visakhapatnam, India*
[*] *Corresponding author, email address: bouchentouf_amina@yahoo.fr*

Abstract

We study a queueing model with disasters, working breakdowns, balking, reneging, and vacations. This is a novel and realistic queueing model that captures the complex dynamics and behaviors of an automatic manufacturing system (AMS) with various uncertainties and disruptions. The system loses all customers when a disaster occurs and repairs start immediately. New customers get slower service during breakdowns. We use matrix methods to find the system's steady state along with performance measures like the expected number of customers lost, the expected waiting time, and system reliability. We also optimize the system parameters (system capacity, number of servers, service rates) to minimize the cost function using a combined direct search method and quasi-Newton method. Our results can enhance the AMS's performance, profit, and customer satisfaction.

Keywords: *queueing models, disasters, customers' impatienc, vacation, Q-matrix method, optimization*

1. Introduction

This paper focuses on the study of an unreliable machining system modeled as a continuous-time $M/M/c/N$ queueing model with single and multiple vacation policies, waiting servers, disasters, repair, working breakdowns, reneging, and balking. Continuous-time queues have proven to be powerful tools for analyzing the performance of various systems. Their broad range of applications and versatility has made them a key area of research in fields such as computer science, communication systems, and manufacturing [11, 12, 17, 26].

Disasters, also called queue flushing [31] or mass exodus [9] break down the server and clear the system of all work. However, they have no effect when the system is empty. These models can be

Received 8 July 2023, accepted 26 September 2024, published online 19 December 2024
ISSN 2391-6060 (Online)/© 2024 Authors

The costs of publishing this issue have been co-financed by the Department of Operations Research and Business Intelligence at the Faculty of Management, Wrocław University of Science and Technology, Wrocław, Poland

applied to computer networks or industrial systems that are vulnerable to virus infections or reset failures. Extensive literature exists on this subject (e.g., [4, 14, 19, 20, 24, 32]).

Another aspect of queueing models with disasters is the possibility of working breakdowns. These are situations where the service does not stop completely but slows down due to a server malfunction. For example, a computer virus may reduce the system's performance without shutting it down. Many service sectors, such as transportation, telecommunications, and healthcare, encounter this problem when a backup server takes over the main server until it is repaired. Kalidass ans Kastouri [21] pioneered this concept in queueing systems, and [22] applied it to a $M/G/1$ queue with disasters and working breakdown services. Later [3], they studied an $M/G/1$ preemptive priority retrial queue with the same features. Recently, Kim ans Lee [13] explored state-dependent arrival and optional re-service in an $M^X/G/1$ queue with disasters and working breakdown service.

Vacation queues are an important class in which the server may take a break once the system gets empty. These queues are widely studied in queueing theory, as they can be used to improve the efficiency of many real-world systems, such as call centers, healthcare facilities, industry, and transportation. Some of the early notable works on this topic were presented in [16, 28–30].

A significant amount of papers have been devoted to the study of customers' impatience in different contexts of vacation queueing models. Customers may balk, meaning they do not join the queue if it is too long or for other reasons. They may also renege, meaning they leave the queue before being served if they wait too long. These phenomena have been well discussed in the case of vacation, where the server does not serve any customers during the vacation period (cf. [1, 2, 8, 15, 25, 27]) and in the case of working vacation; the server acts at a slow rate of service during the vacation of the server (cf. [5–7, 10, 18, 23, 33]).

This paper proposes a new $M/M/c/N$ queueing model that incorporates various features and adaptations to capture the dynamics of real-world systems, such as multiple and single vacation policies, disasters, working breakdowns, and impatience (balking and reneging). The queueing system has potential applications in different industrial settings, such as manufacturing plants and assembly lines. It combines several characteristics that have not been studied together in the literature. This makes the model complex and challenging to analyze.

We employ the Q-matrix method, a powerful numerical technique for analyzing the steady-state behavior of Markov chains, to conduct the steady-state analysis of the model. We calculate the steady-state probabilities and derive closed-form expressions for various performance measures of the queueing system. It is worth pointing out that there are different techniques for solving mathematical problems, such as analytical approaches that use probability-generating functions or maximum entropy approaches, and matrix-analytic approaches that compute the stationary distribution of Markov chains. In our paper, we use the Q-matrix method because it is suitable for complicated queueing problems that lack straightforward analytical solutions. This method simplifies the construction of system transition probability matrices and expresses the steady-state probabilities as functions of specific state probabilities.

We also construct a cost model and formulate an optimization problem to determine the optimal operating values of different system parameters such as the system capacity, the number of servers, and the service rates during repair and regular busy periods, that minimize the total expected cost of the system. We use the direct search method and the quasi-Newton method to solve the optimization problem.

The remainder of the paper is organized as follows. Section 2 describes the queueing model. Section 3 represents the practical example of the proposed queueing model. Section 4 derives the steady-state distribution of the queueing system. Section 5 determines different performance measures. Section 6, on the other hand, develops a cost model and offers numerical results. In Section 7, we conclude the paper.

2. Mathematical description of the model

An $M/M/c/N$ unreliable multi-server queueing system has been considered. Different assumptions required for the formulation of the model are as:

- The system has c servers that provide service to incoming customers who arrive independently according to Poisson processes with rate λ. The service times follow i.i.d. exponential distributions with parameter $1\mu_B$ and are independent of the arrivals. When a customer arrives and sees that one of the c servers is free, he is immediately served according to the first-come-first-served (FCFS) discipline. Otherwise, the customer has to wait for his turn to receive service.
- After serving all customers, the servers enter a waiting period before going on vacation. The waiting period follows an exponential distribution with parameter ϖ. Once this period ends, the servers go on vacation for a random duration following an exponential distribution with rate ϕ. The system operates in two vacation modes: multiple vacation (MV) and single vacation (SV). In MV mode, if there are no waiting customers at the end of the vacation period, another vacation period begins. Otherwise, the system enters a service phase. On the other hand, in SV mode, when the vacation period ends, the servers switch to the busy period and remain idle until a new arrival occurs. The Kronecker δ is given as:

$$\delta = \begin{cases} 1 & \text{for the single vacation model} \\ 0 & \text{for the multiple vacation model} \end{cases}$$

- While the system is operational, there is a possibility of experiencing a catastrophic failure at any moment, resulting in the loss of all current jobs within the system. This catastrophic event is modeled as a Poisson process with an occurrence rate of γ. If such a disaster occurs, the system promptly enters a repair process to recover from the damage. The repair time of the server has an exponential distribution with rate η.
- During the repair period, the main servers are replaced by the substitute ones which serve new failed machines at a slow rate. The service times during this period are exponentially distributed with rate μ_R, where $\mu_R < \mu_B$.
- Upon arrival, if a customer finds some servers working and others free, he is served immediately. However, during vacation, regular busy, or repair periods, customers have the option to join the queue with a probability of θ_n or choose not to join (balk) with a probability of $\theta'_n = 1-\theta_n$. Here, $0 \leq \theta_{n+1} \leq \theta_n \leq 1$ holds for $c \leq n \leq N-1$ in the case of repair and regular busy periods, and $1 \leq n \leq N-1$ during the vacation period. It is important to note that $\theta_0 = 1, \ldots, \theta_{c-1} = 1$; for repair and regular busy periods, and $\theta_0 = 1$ for the vacation period. Additionally, it should be noted that $\theta_N = 0$ for both cases. Then, in the case of the vacation period, we can express the arrival rate as:

$$\lambda_n = \theta_n \lambda, \quad \text{for } 1 \leq n \leq N$$

For the system in the down or regular operative mode, the arrival rate can be defined as:

$$\lambda_n = \begin{cases} \lambda, & \text{for } n < c \\ \theta_n \lambda, & \text{for } c \leq n \leq N \end{cases}$$

- Upon arrival, if the servers are in the regular working or working breakdown period, the customer activates an impatience timer, either T_B or T_R depending on the period. If the customer's service is not completed before the timer expires, the customer has the option to abandon the system. During the vacation period, a new arrival activates its own timer, denoted as T_V. Customers may choose to give up if the service is unavailable before their impatience timer expires. The impatience times, denoted as T_V, T_B, and T_R are random variables that follow exponential distributions with rates $\chi_V, \chi_B, \chi_R > 0$, respectively. The customers' timers are i.i.d. random variables and independent of the number of waiting customers.
- Different stochastic processes involved in the system are supposed to be independent of each other.

3. Motivation and illustration of the queueing model for an AMS

In this section, we motivate and illustrate the practical application of the proposed queueing model by considering an Automatic Manufacturing System (AMS) as an example. An AMS consists of several machines or workstations that process the arrival of jobs or products according to FCFS policy. These machines can be modeled as servers in the queueing system, and the jobs or products can be modeled as customers. The AMS is subject to various modes and customer behaviors that affect its performance and cost.

One of the failure modes is a disaster, which may occur suddenly in the form of an electrical failure, fire, cyber attack, or other event causing major damage to the production facility. Such a disaster results in all jobs or products waiting or currently being processed in the system being lost. This is where our queueing model comes into play. It models this disaster as a Poisson process, allowing to predict and plan for such events. Then, it is necessary to develop a repair and recovery plan that minimizes the downtime and loss of work-in-progress inventory. During the repair period following a disaster, new arrivals may be served at a lower rate (working breakdown). Our model captures this through the use of substitute servers, which serve new failed machines at a slow rate during the repair process. This ensures that production continues as smoothly as possible, maintaining the efficiency of the system and avoiding the loss of potential customers.

Another factor that affects the AMS is vacation periods, which may occur when the machines need downtime for preventive maintenance, upgrades, or periodic breaks after completing all current jobs. During vacations, the machines take a temporary leave from processing new jobs and return after a random period. A vacation can be either single or multiple. When the servers go, simulatively, on break, they leave the system temporarily and return after a random time. Before going on vacation, the servers may wait for a while if there are no customers in the system. A vacation can have both positive and negative effects on the AMS. On one hand, it can improve the performance of the servers. On the other hand, it can increase the workload and waiting time of the remaining servers and customers. Therefore, it is essential to schedule vacations appropriately and balance their benefits and costs. Our queueing model can help with this task by providing useful performance measures such as the expected number of customers in the system during a break period, the expected waiting time of a customer during a break period, and the expected number of impatient customers during a break period.

Customer behavior also affects the AMS. Customers may choose not to enter the system or leave the system before being served due to various reasons. These behaviors are known as balking and reneging, respectively. Balking occurs when a customers decides not to join the queue because he perceives it to be too long or too slow. Reneging occurs when a customers abandons his place in the queue because he becomes impatient or dissatisfied with their waiting time. Both balking and reneging can cause financial losses for the AMS, as there may be costs associated with preparing the order or reserving the place in the queue. Moreover, balking and reneging can reduce the demand and reputation of the AMS, as well as the satisfaction and loyalty of the customers. Therefore, it is crucial to prevent or reduce balking and reneging by improving the service quality and efficiency of the AMS, as well as the satisfaction and loyalty of the customers.

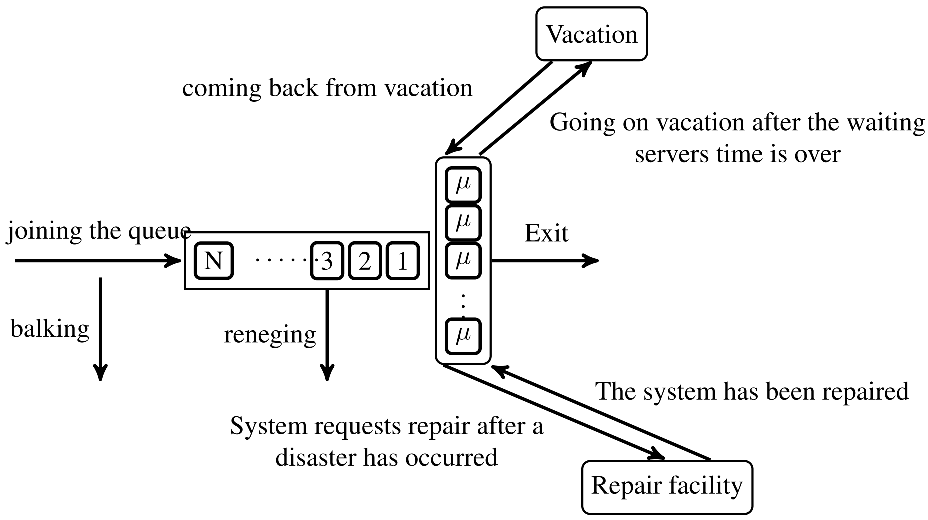

Figure 1. An AMS modeled as $M/M/c/N$ queue with single and multiple vacations, waiting servers, disasters, and working breakdowns

Therefore, it is well noted that our model can provide valuable insights and guidance for practitioners and researchers who are interested in improving the performance and cost-effectiveness of such systems. Figure 1 exemplifies the automated manufacturing system.

4. Derivation of the steady-state distribution

We formulate the process as a special quasi-birth-and-death (QBD) process. The state space of the queueing model at time t is defined by $\{J(t), N(t); t \geq 0\}$. $J(t)$ denotes the server state at time t such that

$$J(t) = \begin{cases} 0, & \text{if the system is on vacation (V)} \\ 1, & \text{if the system is in a regular busy period} \\ 2, & \text{if the system is in a reparation period} \end{cases}$$

and $N(t)$ is the number of customers in the system at that time.

Thus $\{J(t), N(t); t \geq 0\}$ is a bivariate Markov process, and the corresponding state space is

$$\Omega = \begin{Bmatrix} (j,n) : j = 0,1,2,\dots \\ \\ n = 0,1,\dots,N \end{Bmatrix}$$

To study the steady-state behavior of the queueing system, let $\pi_{j,n} = \lim_{t\to\infty} \mathbb{P}\{J(t) = j; N(t) = n\}$ be the steady-state system probabilities of our system, where $(j;n) \in \Omega$. Next, we depict the transition rate diagram of the proposed queueing system (see Figure 2). For this, the following notion is needed:

$$\psi_{0,n} = n\chi_V, \psi_{1,n} = \begin{cases} n(\mu_B + \chi_B), & 1 \leq n \leq c-1 \\ \\ c\mu_B + n\chi_B, & c \leq n \leq N \end{cases}$$

$$\psi_{2,n} = \begin{cases} n(\mu_R + \chi_R), & 1 \leq n \leq c-1 \\ \\ c\mu_R + n\chi_R, & c \leq n \leq N \end{cases}$$

Then, we set the Chapman–Kolmogorov equations specifying the Markov model:

$$(\lambda + \delta\phi)\pi_{0,0} = \varpi\pi_{1,0} + \chi_V\pi_{0,1}, \quad n = 0$$

$$(\theta_1\lambda + \phi + \chi_V)\pi_{0,1} = \lambda\pi_{0,0} + 2\chi_V\pi_{0,2}, \quad n = 1$$

$$(\theta_n\lambda + \phi + n\chi_V)\pi_{0,n} = \theta_{n-1}\lambda\pi_{n-1,0} + (n+1)\chi_V\pi_{0,n+1}, \quad 2 \leq n \leq N-1$$

$$(\phi + N\chi_V)\pi_{0,n} = \theta_{N-1}\lambda\pi_{0,N-1}, \quad n = N$$

$$(\lambda + \varpi)\pi_{1,0} = \eta\pi_{2,0} + \chi_B\pi_{1,1} + \delta\phi\pi_{0,0}, \quad n = 0$$

$$(\lambda + \gamma + n(\mu_B + \chi_B))\pi_{1,n} = \lambda\pi_{1,n-1} + (n+1)(\mu_B + \chi_B)\pi_{1,n+1} + \eta\pi_{2,n} + \phi\pi_{0,n}, \quad 1 \leq n \leq c-1$$

$$(\lambda\theta_c + \gamma + n(\mu_B + \chi_B))\pi_{1,n} = \lambda\pi_{1,n-1} + (c\mu_B + (n+1)\chi_B)\pi_{1,n+1} + \eta\pi_{2,n} + \phi\pi_{0,n}, \quad n = c$$

$$(\lambda\theta_n + \gamma + c\mu_B + n\chi_B)\pi_{1,n} = \theta_{n-1}\lambda\pi_{1,n-1} + (c\mu_B + (n+1)\chi_B)\pi_{1,n+1} + \eta\pi_{2,n} + \phi\pi_{0,n}, \quad c+1 \leq n \leq N-1$$

$$(\gamma + c\mu_B + N\chi_B)\pi_{1,n} = \theta_{N-1}\lambda\pi_{1,N-1} + \eta\pi_{2,n} + \phi\pi_{0,n}, \quad n = N$$

$$(\lambda + \eta)\pi_{2,0} = \gamma\sum_{n=1}^{N}\pi_{1,n} + (\mu_R + \chi_R)\pi_{2,1}, \quad n = 0$$

$$(\lambda + \eta + n(\mu_R + \chi_R))\pi_{2,n} = \lambda\pi_{2,n-1} + (n+1)(\mu_R + \chi_R)\pi_{2,n+1}, \quad 2 \leq n \leq c-1$$

$$(\lambda\theta_c + \eta + n(\mu_R + \chi_R))\pi_{2,n} = \lambda\pi_{2,n-1} + (c\mu_R + (n+1)\chi_R)\pi_{2,n+1}, \quad n = c$$

$$(\lambda\theta_c + \eta + (c\mu_R + n\chi_R))\pi_{2,n} = \theta_{n-1}\lambda\pi_{2,n-1} + (c\mu_R + (n+1)\chi_R)\pi_{2,n+1}, \quad c+1 \leq n \leq N-1$$

$$(\eta + c\mu_R + N\chi_R)\pi_{2,n} = \theta_{N-1}\lambda\pi_{2,N-1}, \quad n = N$$

The normalizing condition is as: $\sum_{n=0}^{N}\pi_{0,n} + \sum_{n=0}^{N}\pi_{1,n} + \sum_{n=0}^{N}\pi_{2,n} = 1.$

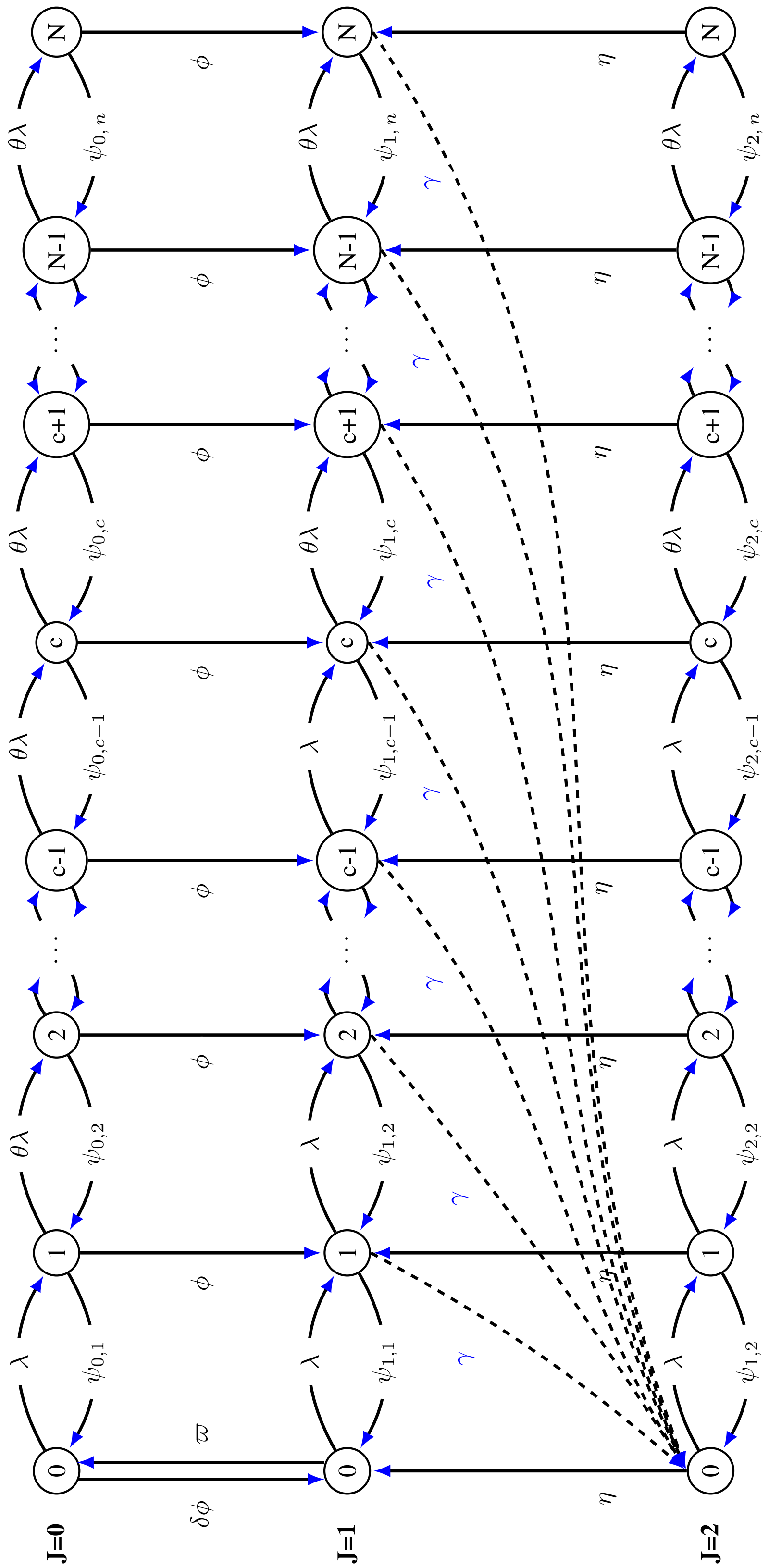

Figure 2. Transition diagram

The following notions are needed for the sequel of analysis:

$$\alpha_n = \begin{cases} -(\lambda + \delta\phi), & n = 0 \\ -(\theta_n\lambda + \psi_{0,n} + \phi), & 1 \le n \le N-1 \\ -(\psi_{0,n} + \phi), & n = N \end{cases}$$

$$\beta_n = \begin{cases} -(\lambda + \varpi), & n = 0, \\ -(\lambda + \psi_{1,n} + \gamma), & 1 \le n \le c-1 \\ -(\theta_c\lambda + \psi_{1,n} + \gamma), & n = c, \\ -(\theta_n\lambda + \psi_{1,n} + \gamma), & c+1 \le n \le N-1 \\ -(\psi_{1,n} + \gamma), & n = N, \end{cases}$$

and

$$\varepsilon_n = \begin{cases} -(\lambda + \eta), & n = 0, \\ -(\lambda + \psi_{2,n} + \eta), & 1 \le n \le c-1, \\ -(\theta_c\lambda + \psi_{2,n} + \eta), & n = c \\ -(\theta_n\lambda + \psi_{2,n} + \eta), & c+1 \le n \le N-1 \\ -(\psi_{2,n} + \eta), & n = N. \end{cases}$$

After having arranged the system's states, we construct the infinitesimal generator matrix denoted by Q

$$Q = \begin{pmatrix} \mathit{\Lambda}_1 & \mathit{\Lambda}_2 & \mathit{\Lambda}_3 \\ \mathit{\Theta}_1 & \mathit{\Theta}_2 & \mathit{\Theta}_3 \\ \mathit{\Omega}_1 & \mathit{\Omega}_2 & \mathit{\Omega}_3 \end{pmatrix}$$

with

$$\mathit{\Lambda}_1 = \begin{pmatrix} \alpha_0 & \lambda_0 & & & & & \\ \psi_{0,1} & \alpha_1 & \lambda_1 & & & & \\ & \psi_{0,2} & \alpha_2 & \lambda_2 & & & \\ & & \ddots & & \ddots & & \ddots \\ & & & & \psi_{0,N-1} & \alpha_{N-1} & \lambda_{N-1} \\ & & & & & \psi_{0,n} & \alpha_N \end{pmatrix}_{N+1 \times N+1}$$

$$\mathit{\Lambda}_2 = \begin{pmatrix} \delta\phi & 0 & \dots & \dots & 0 \\ 0 & \phi & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & & & 0 \\ 0 & 0 & \dots & \dots & \phi \end{pmatrix}_{N+1 \times N+1},$$

$$\mathit{\Theta}_1 = \begin{pmatrix} \varpi & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}_{N+1 \times N+1}$$

$$\Theta_2 = \begin{pmatrix} \beta_0 & \lambda & & & & & & & \\ \psi_{1,1} & \beta_1 & \lambda & & & & & & \\ & \psi_{1,2} & \beta_2 & \lambda & & & & & \\ & & \ddots & \ddots & \ddots & & & & \\ & & & \psi_{1,c-1} & \beta_{c-1} & \lambda & & & \\ & & & & \psi_{1,c} & \beta_c & \lambda_c & & \\ & & & & \ddots & \ddots & \ddots & & \\ & & & & & \psi_{1,N-1} & \beta_{N-1} & \lambda_{N-1} \\ & & & & & & \psi_{1,n} & \beta_N \end{pmatrix}_{N+1\times N+1}$$

$$\Theta_3 = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ \gamma & 0 & \dots & \dots & 0 \\ \gamma & & & & 0 \\ \vdots & & & & 0 \\ \gamma & & \dots & \dots & 0 \end{pmatrix}_{N+1\times N+1}$$

$$\Omega_3 = \begin{pmatrix} \varepsilon_0 & \lambda & & & & & & & \\ \psi_{2,1} & \varepsilon_1 & \lambda & & & & & & \\ & \psi_{2,2} & \varepsilon_2 & \lambda & & & & & \\ & & \ddots & \ddots & \ddots & & & & \\ & & & \psi_{2,c-1} & \varepsilon_{c-1} & \lambda & & & \\ & & & & \psi_{2,c} & \varepsilon_c & \lambda_c & & \\ & & & & \ddots & \ddots & \ddots & & \\ & & & & & \psi_{2,N-1} & \varepsilon_{N-1} & \lambda_{N-1} \\ & & & & & & \psi_{2,n} & \varepsilon_N \end{pmatrix}_{N+1\times N+1}$$

and

$$\Omega_2 = \begin{pmatrix} \eta & & \\ & \ddots & \\ & & \eta \end{pmatrix}_{N+1\times N+1}$$

The matrices Λ_3 and Ω_1 are zero matrices of order $N+1\times N+1$. Next, we concentrate on the steady-state distribution of the process $\{J(t),\ N(t), t \geq 0\}$. Let Π be the steady-state probability vector of the matrix Q, where $\Pi = (\Pi_0,\ \Pi_1,\ \Pi_2)$ such that $\Pi_0 = (\pi_{0,0}, \pi_{0,1}, \dots \pi_{0,n})$, $\Pi_1 = (\pi_{1,0}, \pi_{1,1}, \dots \pi_{1,n})$, $\Pi_2 = (\pi_{2,0}, \pi_{2,1}, \dots \pi_{2,n})$. The steady-state equations

$$\Pi Q = 0 \tag{1}$$

must be verified by the vector $\varPi$ and the normalisation condition:

$$\varPi e = 1 \tag{2}$$

with 0 a zero row vector, $e = (e_1, e_2, e_3)$ is an $(3N+3)$ column vector with ones, and $e_j,\ \ j = 0, 1, 2$ are an (N+1) columns vectors.

Making use of equations (1) and (2) and the fact that $\varLambda_3$ and $\varOmega_1$ are zero matrices, we obtain

$$\varPi_0 \varLambda_1 + \varPi_1 \varTheta_1 = 0 \tag{3}$$

$$\varPi_0 \varLambda_2 + \varPi_1 \varTheta_2 + \varPi_2 \varOmega_2 = 0 \tag{4}$$

$$\varPi_1 \varTheta_3 + \varPi_2 \varOmega_3 = 0 \tag{5}$$

$$\varPi_0 e_1 + \varPi_1 e_2 + \varPi_2 e_3 = 1 \tag{6}$$

The matrices $\varLambda_2$, $\varTheta_1$ and $\varTheta_3$ can be written as follows:

$$\varLambda_2 = \begin{pmatrix} \delta\phi & O_1 \\ O_2 & \phi I_N \end{pmatrix}, \ \varTheta_1 = \begin{pmatrix} \varpi & O_1 \\ O_2 & O_3 \end{pmatrix}, \ \varTheta_3 = \begin{pmatrix} O_4 & \\ \gamma J_N & O_3 \end{pmatrix}$$

where $O_i, i = 1, 2, 3, 4$ are zero matrices and I_N an identity matrix, where O_1, O_2, O_3, O_4 of order $1 \times N$, $N \times 1$, $N \times N$, $1 \times N + 1$, respectively, and J_N is N-dimensional column vector with ones.

From equation (3), we get

$$\varPi_0 = -\varPi_1 \begin{pmatrix} \varpi o \\ O_5 \end{pmatrix} = -\pi_{1,0} \varpi o \tag{7}$$

where O_5 is an $N \times N + 1$ matrix and $o = (o_0, \tilde{o})$ such that $\tilde{o} = (o_1, \ldots, o_N)$ is an N row vector of the matrix $\varLambda_1^{-1}$.

From equation (5), we have

$$\varPi_2 = -\varPi_1 \varTheta_3 \varOmega_3^{-1} \tag{8}$$

By replacing equation (7) and equation (8) into equation (4), we find $-\pi_{1,0} \varpi o \varLambda_2 + \varPi_1 \tilde{\varTheta} = 0$, where $\tilde{\varTheta} = (\varTheta_2 - \eta \varTheta_3 \varOmega_3^{-1})$.

Therefore,

$$\varPi_1 = \pi_{1,0} \varpi \varLambda_2 o \tilde{\varTheta}^{-1} = \begin{cases} \pi_{1,0} \varpi \phi o \tilde{\varTheta}^{-1}, & \delta = 1 \\ \pi_{1,0} \varpi \phi \dot{\tilde{o}} \tilde{\varTheta}^{-1}, & \delta = 0 \end{cases} \tag{9}$$

and

$$\varPi_2 = -\varPi_1 \gamma \dot{\varOmega}^{-1} = \begin{cases} -\pi_{1,0} \varpi \phi o \Upsilon \tilde{\varTheta}^{-1}, & \delta = 1 \\ -\pi_{1,0} \varpi \phi \dot{\tilde{o}} \Upsilon \tilde{\varTheta}^{-1}, & \delta = 0 \end{cases} \tag{10}$$

where $\dot{\tilde{o}} = (0, o_1, \ldots, o_N)$ and $\Upsilon = \varTheta_3 \varOmega_3^{-1}$.

Finally, based on the above analysis, from (7), (9), and (10), we obtain the expressions for $\pi_{j,n}$ in terms of $\pi_{1,0}$, for single and multiple vacation policies, then use normalization condition to derive this latter. We summarize our results in the following theorem.

Theorem 1. For a finite source-multiserver Markovian queueing system with single and multiple vacation policies, waiting servers, disasters, working breakdowns, and impatience, the stationary probabilities $\pi_{j,n}$ can be expressed as:

1. For single vacation policy $(\delta = 1)$

$$\pi_{j,n} = \begin{cases} -\varpi o_n \pi_{1,0}, & j = 0,\ 0 \le n \le N \\ \varpi\phi \sum_{i=0}^{N} o_i \tilde{\theta}_{in} \pi_{1,0}, & j = 1,\ 0 \le n \le N \\ -\varpi\phi \sum_{j=0}^{N}\sum_{i=0}^{N} o_i \tilde{\theta}_{ij} \omega_{jn} \pi_{1,0}, & j = 2,\ 0 \le n \le N \end{cases}$$

where

$$\pi_{1,0} = \big(-\varpi \sum_{n=0}^{N} o_n + \varpi\phi \sum_{n=0}^{N}\sum_{i=0}^{N} o_i \tilde{\theta}_{in} - \varpi\phi \sum_{n=0}^{N}\sum_{j=0}^{N}\sum_{i=0}^{N} o_i \tilde{\theta}_{ij}\omega_{jn} \big)^{-1}$$

2. For multiple vacation policy case $(\delta = 0)$

$$\pi_{j,n} = \begin{cases} -\varpi o_n \pi_{1,0}, & j = 0, 0 \le n \le N, \\ \varpi\phi \sum_{0=1}^{N} \dot{\tilde{o}}_i \tilde{\theta}_{in} \pi_{1,0}, & j = 1, 0 \le n \le N \\ -\varpi\phi \sum_{j=0}^{N}\sum_{i=0}^{N} \dot{\tilde{o}}_i \tilde{\theta}_{ij} \omega_{jn} \pi_{1,0}, & j = 2, 0 \le n \le N \end{cases}$$

where

$$\pi_{1,0} = \big(-\varpi \sum_{n=0}^{N} o_n + \varpi\phi \sum_{n=1}^{N}\sum_{i=0}^{N} \dot{\tilde{o}}_i \tilde{\theta}_{in} - \varpi\phi \sum_{n=1}^{N}\sum_{j=0}^{N}\sum_{i=0}^{N} \dot{\tilde{o}}_i \tilde{\theta}_{ij}\omega_{jn} \big)^{-1}$$

such that ω_{jn} are the elements of the matrix $\Upsilon = \Theta_3 \Omega_3^{-1}$ and $\tilde{\theta}_{ij}$ are the elements of matrix $\tilde{\Theta} = (\Theta_2 - \eta\Theta_3\Omega_3^{-1})$, $o = (o_0, \tilde{o})$, where $\tilde{o} = (o_1, \ldots, o_N)$ is an N row vector of the matrix Λ_1^{-1}, and $\dot{\tilde{o}} = (0, o_1, \ldots, o_N)$

Proof. The theorem has been proved in the above. □

5. Performance measures

In this section, based on the steady-state probabilities, we obtain different performance measures. The results are presented as follows

- The probability that the service is available

$$P_B = \begin{cases} \sum_{n=0}^{N} \pi_{1,n} = \varpi\phi \sum_{n=0}^{N}\sum_{i=0}^{N} o_i\tilde{\theta}_{in}\pi_{1,0}, & \delta = 1 \\ \sum_{n=0}^{N} \pi_{1,n} = \varpi\phi \sum_{n=0}^{N}\sum_{i=0}^{N} \dot{\tilde{o}}_i\tilde{\theta}_{in}\pi_{1,0}, & \delta = 0 \end{cases}$$

- The probability that the service is unavailable due to vacation

$$P_V = \sum_{n=0}^{N} \pi_{0,n} = -\varpi \sum_{n=0}^{N} o_n\pi_{1,0}$$

- The probability that the service is under repair

$$P_R = \sum_{n=0}^{N} \pi_{2,n} = \begin{cases} -\varpi\phi \left(\sum_{n=0}^{N}\sum_{j=0}^{N}\sum_{i=0}^{N} o_i\tilde{\theta}_{ij}\omega_{jn} \right) \pi_{1,0}, & \delta = 1 \\ -\varpi\phi \left(\sum_{n=0}^{N}\sum_{j=0}^{N}\sum_{i=0}^{N} \dot{\tilde{o}}_i\tilde{\theta}_{ij}\omega_{jn} \right) \pi_{1,0}, & \delta = 0 \end{cases}$$

- The mean system size

$$\begin{aligned} L_s &= \sum_{n=1}^{N} n(\pi_{0,n} + \pi_{1,n} + \pi_{2,n}) \\ &= \begin{cases} -\varpi \left(\sum_{n=1}^{N} no_n - \phi \sum_{n=1}^{N}\sum_{i=0}^{N} no_i\tilde{\theta}_{in} + \phi \sum_{n=1}^{N}\sum_{j=1}^{N}\sum_{i=0}^{N} no_i\tilde{\theta}_{ij}\omega_{jn} \right) \pi_{1,0}, & \delta = 1 \\ -\varpi \left(\sum_{n=1}^{N} no_n - \phi \sum_{n=1}^{N}\sum_{i=0}^{N} n\dot{\tilde{o}}_i\tilde{\theta}_{in} + \phi \sum_{n=1}^{N}\sum_{j=0}^{N}\sum_{i=0}^{N} n\dot{\tilde{o}}_i\tilde{\theta}_{ij}\omega_{jn} \right) \pi_{1,0}, & \delta = 0 \end{cases} \end{aligned}$$

- The effective arrival rate

$$\dot{\lambda} = \sum_{n=0}^{N} \lambda_n(\pi_{0,n} + \pi_{1,n} + \pi_{2,n})$$

- The mean waiting time of a customer in the system

$$W_s = \frac{L_s}{\dot{\lambda}}$$

- The average balking rate

$$R_b = \lambda - \dot{\lambda}$$

- The system reliability

$$P_{re} = 1 - P_R = 1 - \sum_{n=0}^{N} \pi_{2,n} = \begin{cases} 1 + \varpi\phi \left(\sum_{n=0}^{N} \sum_{j=0}^{N} \sum_{i=0}^{N} o_i \tilde{\theta}_{ij} \omega_{jn} \right) \pi_{1,0}, & \delta = 1 \\ 1 + \varpi\phi \left(\sum_{n=0}^{N} \sum_{j=0}^{N} \sum_{i=0}^{N} \dot{\tilde{o}}_i \tilde{\theta}_{ij} \omega_{jn} \right) \pi_{1,0}, & \delta = 0 \end{cases}$$

- The average reneging rate

$$R_{\text{ren}} = \begin{cases} -\varpi \left(\chi_V \sum_{n=1}^{N} n o_n - \chi_B \phi \sum_{n=1}^{N} \sum_{i=0}^{N} n o_i \tilde{\theta}_{in} + \chi_R \phi \sum_{n=1}^{N} \sum_{j=0}^{N} \sum_{i=0}^{N} n o_i \tilde{\theta}_{ij} \omega_{jn} \right) \pi_{1,0}, & \delta = 1 \\ -\varpi \left(\chi_V \sum_{n=1}^{N} n o_n - \chi_B \phi \sum_{n=1}^{N} \sum_{i=0}^{N} n \dot{\tilde{o}}_i \tilde{\theta}_{in} + \chi_R \phi \sum_{n=1}^{N} \sum_{j=0}^{N} \sum_{i=0}^{N} n \dot{\tilde{o}}_i \tilde{\theta}_{ij} \omega_{jn} \right) \pi_{1,0}, & \delta = 0 \end{cases}$$

- The mean number of customers served per unit time

$$E_s = \mu_B \left(\sum_{n=1}^{c-1} n\pi_{1,n} + c \sum_{n=c}^{N} \pi_{1,n} \right) + \mu_R \left(\sum_{n=1}^{c-1} n\pi_{2,n} + c \sum_{n=c}^{N} \pi_{2,n} \right)$$

6. Optimization analysis

Now, we focus on the cost parameter optimization. We aim to find the values of the decision variables (optimum system capacity, minimum number of servers, and optimum service rates during both repair and regular busy periods) that minimize the objective function (the total expected cost of the system) while satisfying the constraints. Combined with the above results, we then investigate the effects of main system parameters on different performance measures given above as well as on the expected cost function per unit time T_c and on the total expected profit per unit time of the system T_{ep} that are given as:

$$T_c = c_B P_B + c_V P_V + c_R P_R + c_h E(L_s) + c_l (R_{\text{ren}} + R_b) + c \times c_f + c \times (\mu_B c_{s_1} + \mu_R c_{s_2})$$

where

c_B – cost per unit time when the system is in a normal busy period,
c_V – cost per unit time when the system is in vacation period,
c_R – cost per unit time when the system is under repair,
c_h – holding cost per unit time when a customer enters the queue, and waits for service,
c_{s_1} – cost per service per unit time in a normal busy period,
c_{s_2} – cost per service per unit time in repair period,
c_l – cost per service per unit time when a customer is lost,
c_f – fixed purchase cost of the server per unit,, *and*

$$T_{ep} = T_{ev} - T_c$$

$T_{ev} = \mathrm{R} \times E_s$ denotes the total expected revenue per unit time of the system and R represents the revenue earned by providing service to a customer.

The objective function (11) is complex and highly non-linear. In this case, it is beneficial to use the direct search method and quasi-Newton method for finding good approximations to the optimal solution. For the numerical analysis, we consider the different cost elements as: $c_V = \$18$, $c_B = \$50$, $c_R = \$60$, $c_h = \$18$, $c_{s1} = \$1$, $c_{s2} = \$1$, $c_l = \$18$, $c_f = \$1$, and $R = 50$.

6.1. Direct search method

In this subsection, we aim to obtain the joint optimal values (c^*, N^*) by employing the direct search method. If the function $T_c(c, N)$ is convex (unimodal), a single relative minimum exists. In order to get $T_c(c^*, N^*)$, we have to prove that the $T_c(c, N)$ is convex. We take different parameters as follows: $\lambda = 3$, $\chi_V = 0.5$, $\chi_B = 0.3$, $\chi_R = 0.7$, $\mu_R = 1$, $\mu_B = 4$, $\phi = 1.2$, $\eta = 0.5$, $\gamma = 0.2$, $\varpi = 0.5$, $\theta = 1 - (n/N)$ and vary the system capacity (N) as well as the number of servers in the system (c). The obtained numerical results have been presented in the following figures and tables.

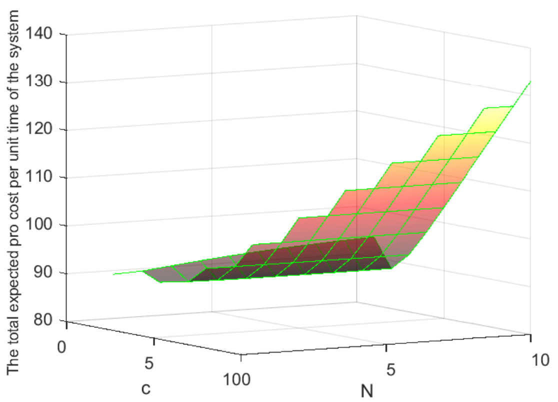

Figure 3. The expected cost vs. c, N for SVP

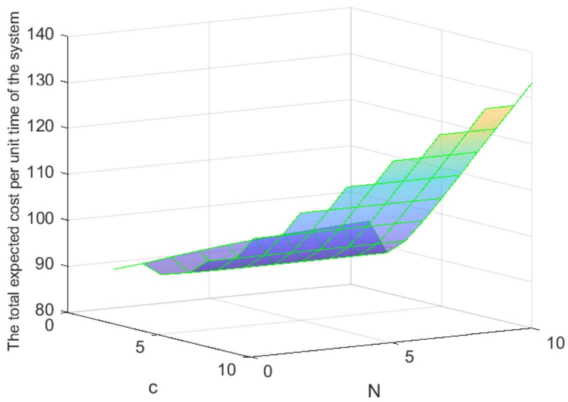

Figure 4. The expected cost vs. c, N for MVP

Figures 3 and 4 clearly show the convexity of the curves for both single and multiple vacation policies; there exist certain values of the number of servers (c) and the system capacity (N) that minimize the total expected cost function for the chosen set of model parameters. More precisely, from Table 1, we have the combined optimal solution is $(c^*, N^*) = (2, 7)$ with $T_c = \$88.2949$, in single vacation policy, and $(c^*, N^*) = (2, 5)$ with $T_c = \$88.6103$, in multiple vacation policy.

6.2. Quasi-Newton method

After determining the optimal values of c and N using the direct search method, we proceed to search for the optimal service rates (μ_B^*, μ_R^*). This is achieved using the quasi-Newton method. Our focus then shifts to examining the effects of different system parameters on optimization results.

Next, we conduct a sensitivity analysis. This analysis evaluates the impacts of various system parameters. These parameters include the arrival rate (λ), vacation (ϕ), waiting servers rate (ω), disaster rate (γ), and repair rate (η). We also consider the reneging rates during vacation, busy, and repair periods, represented by ξ_V, ξ_B, and ξ_R respectively. Additionally, we look at the non-balking probability (θ_n).

Table 1. N/c for various T_c

	N/c	1	2	3	4	5	6	7	8
SVP	1	90.1648	–	–	–	–	–	–	–
	2	90.4144	88.7606	–	–	–	–	–	–
	3	91.0930	88.4810	91.9394	–	–	–	–	–
	4	91.7545	88.3668	91.9066	97.1769	–	–	–	–
	5	92.3352	88.3182	91.8922	97.2018	103.0250	–	–	–
	6	92.8344	88.2993	91.8862	97.2214	103.0509	109.0118	–	–
	7	93.2633	**88.2949**	91.8845	97.2372	103.0712	109.0322	115.0239	–
	8	93.6337	88.2979	91.8849	97.2501	103.0874	109.0484	115.0399	121.0381
MVP	1	89.7226	–	–	–	–	–	–	–
	2	90.3297	88.8882	–	–	–	–	–	–
	3	91.1546	88.6897	92.1756	–	–	–	–	–
	4	91.9022	88.6248	92.1708	97.4343	–	–	–	–
	5	92.5414	**88.6103**	92.1756	97.4727	103.2904	–	–	–
	6	93.0836	88.6167	92.1837	97.5023	103.3246	109.2835	–	–
	7	93.5456	88.6320	92.1927	97.5258	103.3513	109.3097	115.3009	–
	8	93.9423	88.6508	92.2017	97.5449	103.3726	109.3306	115.3213	121.3195

Table 2. ϕ and λ vs. different performance measures and T_c
for $\chi_V = 0.5, \chi_B = 0.3, \chi_R = 0.7, \eta = 0.5, \ \gamma = 0.2, \varpi = 0.5$

	ϕ	λ	c^*	N^*	μ_B^*	μ_R^*	P_V	P_B	P_R	L_s	W_s	R_b	R_{ren}	P_{re}	E_s	T_c
SVP	1	3.5	2	4	6.5707	1.3195	0.1962	0.6862	0.1175	0.8085	0.2677	0.4804	0.3550	0.8824	4.0728	92.0711
		4	2	4	6.9933	1.6311	0.1894	0.6872	0.1234	0.8566	0.2511	0.5883	0.3760	0.8766	4.6001	97.2001
		4.5	2	4	7.4020	1.9396	0.1834	0.6881	0.1285	0.8987	0.2366	0.7009	0.3942	0.8715	5.1145	101.9905
	1.5	3.5	2	10	6.9049	1.6298	0.1425	0.7346	0.1229	0.7833	0.2360	0.1813	0.3385	0.8771	4.5158	89.1947
		4	2	7	7.4295	1.8682	0.1386	0.7341	0.1273	0.8058	0.2180	0.3038	0.3484	0.8727	5.0619	93.6769
		4.5	2	6	7.9318	2.1329	0.1350	0.7336	0.1315	0.8325	0.2037	0.4134	0.3600	0.8685	5.6114	98.0306
	2	3.5	2	13	7.1081	1.6889	0.1132	0.7628	0.1240	0.7168	0.2121	0.1202	0.3060	0.8760	4.6930	87.7859
		4	2	9	7.6670	1.9484	0.1100	0.7614	0.1286	0.7468	0.1969	0.2081	0.3192	0.8714	5.2793	91.9320
		4.5	2	7	8.2296	2.1929	0.1074	0.7603	0.1323	0.7686	0.1835	0.3105	0.3288	0.8677	5.8540	96.0738
MVP	1	3.5	2	4	6.3678	1.2905	0.2386	0.6457	0.1157	0.8699	0.2940	0.5413	0.3856	0.8843	3.8935	93.1785
		4	2	4	6.7694	1.6092	0.2248	0.6532	0.1221	0.9141	0.2731	0.6527	0.4043	0.8779	4.4117	98.2658
		4.5	2	4	7.1620	1.9237	0.2134	0.6589	0.1277	0.9527	0.2553	0.7683	0.4205	0.8723	4.9235	103.1658
	1.5	3.5	2	8	6.8076	1.5410	0.1923	0.6880	0.1197	0.8361	0.2576	0.2539	0.3655	0.8803	4.3304	89.9404
		4	2	7	7.2652	1.8421	0.1800	0.6943	0.1256	0.8726	0.2389	0.3468	0.3809	0.8744	4.8878	94.5163
		4.5	2	6	7.7347	2.1122	0.1707	0.6991	0.1302	0.8943	0.2217	0.4653	0.3900	0.8698	5.4270	99.0260
	2	3.5	2	12	7.0610	1.6350	0.1672	0.7118	0.1210	0.7798	0.2329	0.1518	0.3372	0.8790	4.5440	88.0928
		4	2	8	7.5837	1.8809	0.1563	0.7177	0.1260	0.7996	0.2141	0.2652	0.3456	0.8740	5.1072	92.5729
		4.5	2	7	8.0844	2.1661	0.1471	0.7222	0.1306	0.8276	0.1996	0.3538	0.3575	0.8694	5.6821	96.7991

The impacts of these parameters on different performance measures and $T_c(c^*, N^*, \mu_B^*, \mu_R^*)$ are considered. The numerical results of this analysis are provided in Tables 2–5. Furthermore, the impact of these system parameters on $T_{ep}(c^*, N^*, \mu_B^*, \mu_R^*)$ is illustrated in Figures 5–10. The numerical study is presented for both SVP and MVP.

Table 3. γ, ϖ, and η vs. different performance measures and T_c for $\chi_V = 0.5, \chi_B = 0.3, \chi_R = 0.7,\ \lambda = 3, \phi = 1.2$

	γ	ϖ	η	c^*	N^*	μ_B^*	μ_R^*	P_V	P_B	P_R	L_s	W_s	R_b	R_{ren}	P_{re}	E_s	T_c
SVP	0.2	0.6	0.8	2	17	6.2955	0.2096	0.2093	0.7163	0.0744	0.8803	0.3050	0.1142	0.3806	0.9256	3.7719	83.8095
			1	2	35	6.1974	0.0001	0.2101	0.7283	0.0616	0.9063	0.3080	0.0577	0.3842	0.9384	3.8184	82.5542
		0.8	0.8	2	10	6.2265	0.0834	0.2557	0.6721	0.0721	0.9083	0.3250	0.2053	0.3969	0.9279	3.6018	84.3462
			1	2	15	6.0834	0.0001	0.2559	0.6841	0.0600	0.9350	0.3271	0.1419	0.3999	0.9400	3.6523	83.1603
	0.4	0.6	0.8	2	12	5.9993	1.3056	0.1930	0.6667	0.1404	0.8398	0.2945	0.1481	0.3718	0.8596	3.5953	86.3136
			1	2	44	6.0669	0.8200	0.1974	0.6861	0.1165	0.9080	0.3073	0.0456	0.3978	0.8835	3.6570	84.9475
		0.8	0.8	2	8	5.9284	1.2086	0.2364	0.6273	0.1363	0.8655	0.3130	0.2347	0.3863	0.8637	3.4395	86.8294
			1	2	16	6.0019	0.7196	0.2416	0.6453	0.1131	0.9354	0.3262	0.1320	0.4135	0.8869	3.5003	85.5008
MVP	0.2	0.6	0.8	2	9	6.1993	0.0462	0.2718	0.6568	0.0714	0.9201	0.3325	0.2330	0.4033	0.9286	3.5451	84.5228
			1	2	13	6.0446	0.0001	0.2715	0.6690	0.0595	0.9468	0.3342	0.1671	0.4061	0.9405	3.5978	83.3578
		0.8	0.8	2	6	6.0559	0.0001	0.3242	0.6069	0.0689	0.9316	0.3528	0.3592	0.4110	0.9311	3.3392	85.0607
			1	2	8	5.9007	0.0001	0.3233	0.6191	0.0576	0.9653	0.3551	0.2814	0.4174	0.9424	3.3996	83.9861
	0.4	0.6	0.8	2	7	5.9024	1.1697	0.2519	0.6133	0.1348	0.8707	0.3191	0.2715	0.3896	0.8652	3.3808	87.0042
			1	2	13	5.9774	0.6843	0.2568	0.6312	0.1119	0.9426	0.3325	0.1648	0.4179	0.8881	3.4448	85.6799
		0.8	0.8	2	5	5.8095	1.0461	0.3023	0.5679	0.1298	0.8845	0.3394	0.3936	0.3989	0.8702	3.1894	87.5230
			1	2	8	5.8880	0.5675	0.3074	0.5846	0.1079	0.9623	0.3536	0.2786	0.4304	0.8921	3.2578	86.2343

Table 4. χ_R, χ_B, and χ_V vs. different performance measures and T_c for $\lambda = 3, \eta = 0.5, \gamma = 0.2, \phi = 1.2, \varpi = 0.5$

	χ_R	χ_B	χ_V	c^*	N^*	μ_B^*	μ_R^*	P_V	P_B	P_R	L_s	W_s	R_b	R_{ren}	P_{re}	E_s	T_c
SVP	0.1	0.4	0.3	2	3	6.3288	1.2313	0.1810	0.7091	0.1098	0.6926	0.2707	0.4420	0.2152	0.8902	3.6733	86.7194
			0.5	2	3	6.3080	1.2289	0.1813	0.7091	0.1096	0.6747	0.2619	0.4238	0.2499	0.8904	3.6547	86.6412
		0.6	0.3	2	3	6.4353	1.1848	0.1847	0.7090	0.1063	0.6815	0.2659	0.4375	0.2766	0.8937	3.6057	87.5144
			0.5	2	3	6.4134	1.1824	0.1849	0.7090	0.1061	0.6633	0.2570	0.4191	0.3118	0.8939	3.5869	87.4303
	0.3	0.4	0.3	2	3	6.3817	1.1099	0.1817	0.7091	0.1092	0.6857	0.2674	0.4355	0.2396	0.8908	3.6576	86.7539
			0.5	2	4	6.2940	1.2454	0.1801	0.7091	0.1108	0.7071	0.2661	0.3428	0.2913	0.8892	3.6831	86.5679
		0.6	0.3	2	3	6.4879	1.0613	0.1853	0.7090	0.1057	0.6749	0.2628	0.4312	0.3003	0.8943	3.5904	87.5400
			0.5	2	3	6.4658	1.0588	0.1856	0.7090	0.1055	0.6567	0.2538	0.4128	0.3355	0.8945	3.5715	87.4553
MVP	0.1	0.4	0.3	2	3	6.1713	1.1966	0.2346	0.6581	0.1073	0.7518	0.3020	0.5107	0.2329	0.8927	3.4993	87.2180
			0.5	2	3	6.1319	1.1905	0.2390	0.6542	0.1068	0.7327	0.2921	0.4920	0.2801	0.8932	3.4617	87.1493
		0.6	0.3	2	2	6.3540	0.8976	0.2443	0.6550	0.1007	0.6651	0.2823	0.6439	0.2667	0.8993	3.3380	88.0555
			0.5	2	2	6.3175	0.8894	0.2496	0.6502	0.1002	0.6525	0.2749	0.6264	0.3092	0.8998	3.3012	88.0154
	0.3	0.4	0.3	2	3	6.2240	1.0731	0.2355	0.6579	0.1066	0.7452	0.2986	0.5045	0.2571	0.8934	3.4833	87.2463
			0.5	2	3	6.1845	1.0666	0.2400	0.6539	0.1061	0.7261	0.2888	0.4858	0.3044	0.8939	3.4456	87.1763
		0.6	0.3	2	2	6.3994	0.7920	0.2450	0.6548	0.1002	0.6601	0.2793	0.6366	0.2882	0.8998	3.3250	88.0741
			0.5	2	3	6.2854	1.0167	0.2446	0.6527	0.1026	0.7160	0.2844	0.4822	0.3650	0.8974	3.3773	87.9385

Table 5. θ_n vs. different performance measures, T_c, and T_{ep} for $\lambda = 3, \eta = 0.5, \gamma = 0.2, \phi = 1.2, \varpi = 0.5, \chi_V = 0.5,\ \chi_B = 0.3, \chi_R = 0.7$

	c^*	N^*	μ_B^*	μ_R^*	P_V	P_B	P_R	L_s	W_s	R_b	R_{ren}	P_{re}	E_s	T_c	T_{ep}
SVP	1	1	9.1546	0.0001	0.2187	0.7080	0.0733	0.3751	0.1539	0.5626	0.1614	0.9267	6.4818	73.6724	250.4176
	2	7	6.2488	1.1907	0.1766	0.7092	0.1142	0.7623	0.2741	0.2189	0.3325	0.8858	3.7517	86.0172	101.5661
	2	5	6.2026	1.2486	0.1752	0.7093	0.1156	0.7862	0.2779	0.1706	0.3431	0.8844	3.7782	85.8523	103.0591
MVP	1	1	8.8091	0.0001	0.2984	0.6317	0.0698	0.4149	0.1745	0.6224	0.1825	0.9302	5.5651	72.9149	205.3381
	2	5	6.1487	1.0532	0.2337	0.6563	0.1100	0.7957	0.2983	0.3325	0.3512	0.8900	3.5396	86.6557	90.3234
	2	4	6.1121	1.1163	0.2309	0.6576	0.1115	0.8255	0.3026	0.2718	0.3646	0.8885	3.5775	86.4957	92.3810

For the 1st and 4th row of the table $\theta_n = 1/(n+1)$, for the 2nd and 5th row $\theta_n = 1 - (n/N)$, and for the 3rd and 6th row $\theta_n = 1 - (n/N)^2$.

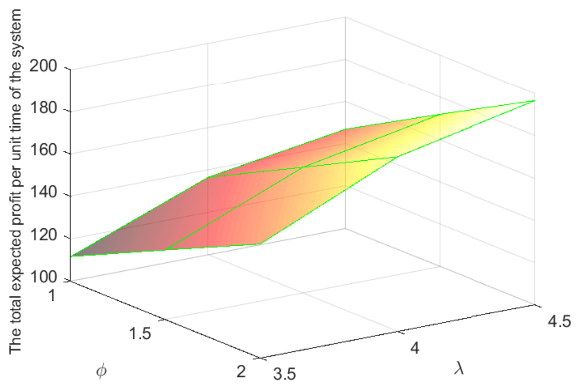

Figure 5. T_{ep} near $(c^*,\, N^*,\, \mu_B^*,\, \mu_R^*)$ for $\delta = 1$ and different values of ϕ and λ

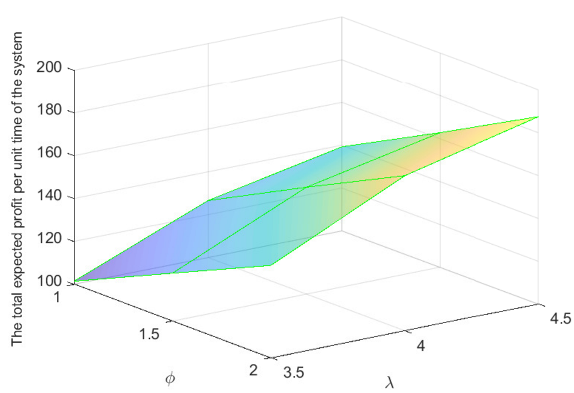

Figure 6. T_{ep} near $(c^*,\, N^*,\, \mu_B^*,\, \mu_R^*)$ for $\delta = 0$ and different values of ϕ and λ

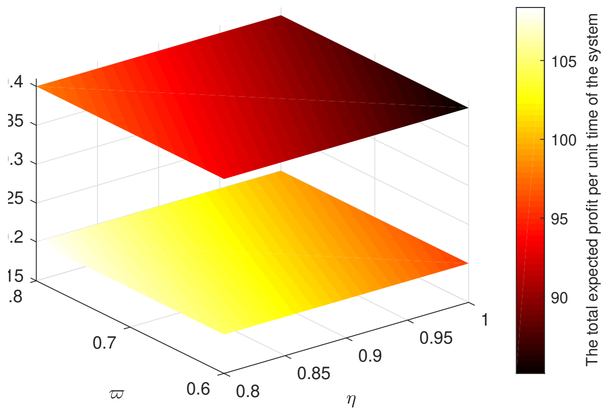

Figure 7. T_{ep} near $(c^*,\, N^*,\, \mu_B^*,\, \mu_R^*)$ for different values of ϖ, γ, and $\delta = 1$

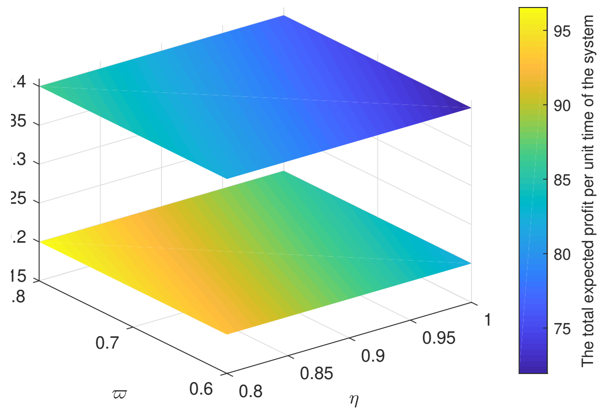

Figure 8. T_{ep} near $(c^*,\, N^*,\, \mu_B^*,\, \mu_R^*)$ for different values of ϖ, γ, and $\delta = 0$

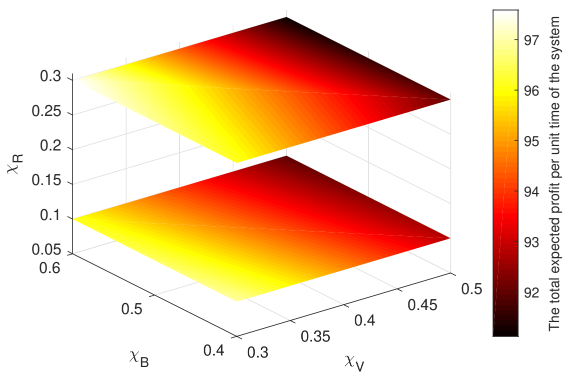

Figure 9. T_{ep} near $(c^*,\, N^*,\, \mu_B^*,\, \mu_R^*)$ for different values of $\chi_V, \chi_B,\, \chi_R$ and $\delta = 1$

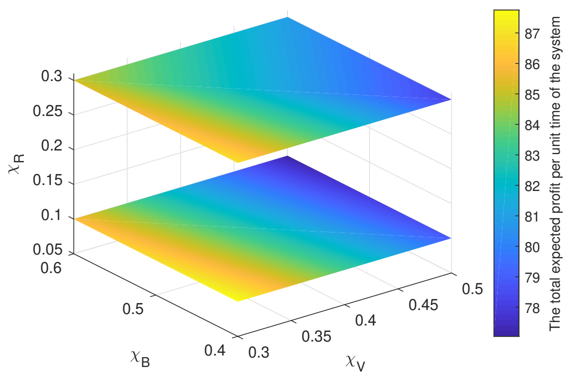

Figure 10. T_{ep} near $(c^*,\, N^*,\, \mu_B^*,\, \mu_R^*)$ for different values of χ_V, χ_B, χ_R and $\delta = 0$

6.2.1. Discussion

The cost optimization analysis of the queueing model can help determine optimal values for the capacity, number of servers, and service rates in the AMS to minimize overall costs. The optimized system capacity N^* guides setting the number of machines/workstations in the production line to balance investment and inventory holding costs. The optimal number of servers c^* helps determine the server capacity that minimizes costs due to disruptions and customer abandonment. The optimized service rates μ_B^*, μ_R^* help define scheduling policies for production and repair operations that maximize throughput while controlling labor expenses.

Taking into account various cost tradeoffs, the optimization framework identifies the ideal combination of capacity, staffing, and service rates for AMS to maintain high quality, throughput, and customer satisfaction while keeping costs in control. The visualization of costs further helps illustrate the impact of different decisions and failure dynamics on the bottom line. This connects the analytical model to tangible planning, configuration, and control policies for real-world flexible manufacturing systems. From the above results, Tables 2–5 and figures 5–10, the following impacts are observed:

- The total cost and the total profit increase whenever λ increases. Since the arrival rate is growing up, there will be a high number of customers arriving in the system $(L_s \nearrow)$. Thus, the total cost will be increased along with the total profit. This is indeed the case as the higher number of jobs entering the service unit will be placed and served resulting in the enhanced total profit.
- (If the vacation completion rate (ϕ) increases, the respective total expected cost (total expected profit) decreases (increases). This is because an increase in the vacation completion rate leads to a decrease in the server's vacation time $(P_V \searrow)$, and thus an increase in its availability to serve customers $(P_B \nearrow)$. Therefore, $(E_s \nearrow)$. This in turn reduces the number of customers in the system$(L_s \searrow)$, resulting in a decrease in T_c and in an increase in T_p. This is obvious, as the machines are available for production and processing, the queueing time will be less and further generated more profits from the system point of view.
- The waiting servers rate (ϖ) impacts the total cost by raising it as it increases: A decrease in the waiting server rate can increase the waiting time for customers $(W_s \nearrow)$, which can augment customer impatience and a decreases in the product value. This in turn increases the total expected cost and decreases the total expected profit.
- The increase of the failure rate (γ) means that fewer customers are being served, which can lead to a decrease in the mean number of customers served $(E_s \searrow)$. This leads to a decrease in revenue and an increase in costs and a decrease in the total expected profit.
- Obviously, high repair rate (η) has a nice impact on the total expected cost function and total expected profit. When the repair rate increases, it means that the failure can be quickly addressed and resolved. This leads to less downtime which results in a reduction of the repair time $(P_R \searrow)$ and an increase in the station's availability to serve customers $(P_B \nearrow)$. As a result, the system can serve more customers and generate more revenue, which can increase the total expected profit. Additionally, a higher repair rate can lead to lower maintenance costs, as regular maintenance can prevent failures and reduce the need for repairs. This can also contribute to a decrease in the total expected cost function.

  However, it is important to note that an excessively high repair rate can also lead to higher costs. This is because maintaining a high repair rate may require expensive equipment or highly skilled technicians. Moreover, there may be diminishing returns as the cost of increasing the repair rate further may outweigh the benefits. Therefore, it is important to balance the repair rate with the costs involved in order to optimize the total expected cost function and total expected profit.
- The impatience rates of customers (χ_V, χ_B, χ_R) have several negative impacts on the system's performance and profitability; abandoning the system can result in a decrease in the total number of customers served and revenue generated, which can negatively impact the system's profitability.

- A higher non-balking probability (θ_n) indicates a higher willingness of customers to join the system, which can increase the number of customers served and the revenue generated. This can positively impact the optimal performance measures and increase the total expected profit.
- From the above observations it is important to note that the majority of the results agree with our intuition while others are not so straightforward to interpret. For instance, we remark that when the joining probability increases from $\theta_n = \frac{1}{n+1}$ to $\theta_n = 1 - \frac{n}{N}$, the total expected profit decreases. This contradicts our intuition, and it can be due to the chosen costs and parameters. It is well known that higher join probability would normally increase profitability (more customers being served). But in this case, perhaps the higher loads lead to disproportionately higher holding costs that outweigh the gains.
- Further, the values of P_V, R_{ren}, R_b, L_s, W_s, P_{re}, T_c for single vacation are lower than those for multiple vacation, and the values of P_R, P_B, E_s, T_{ep} for multiple vacation are lower than those for single vacation.

Remark 1. The obtained results align with intuitive expectations—the single vacation model demonstrates better performance measures versus the multiple vacation model.

7. Conclusion

We proposed a queueing model tailored to machining systems within automated manufacturing environments. Our model integrates critical factors, including disasters, working breakdowns, single and multiple vacations, waiting servers, and customer impatience (balking and reneging). Unlike existing literature models that often isolate specific dynamics, our unified framework captures the interplay of multiple uncertainties.

Our contributions extend beyond theory. By directly applying our queueing model to automated manufacturing systems (AMS), production and quality managers gain practical insights. They can optimize critical parameters such as system capacity, the number of servers, and service rates during disruptions to maximize throughput while balancing costs. Moreover, our customer-centric design accounts for behaviors like impatience and service refusal, ensuring AMS facilities meet operational goals while enhancing overall effectiveness.

Acknowledgement

The authors are thankful to the Editor-in-chief and the anonymous referees that their careful reading gave them the opportunity to improve the quality of the paper.

References

[1] Afroun, F., Aïssani, D., Hamadouche, D., and Boualem, M. Q–matrix method for the analysis and performance evaluation of unreliable $M/M/1/N$ queueing model. *Mathematical Methods in the Applied Sciences 41*, 18 (2018), 9152–9163.

[2] Ammar, S. I. Transient solution of an $M/M/1$ vacation queue with a waiting server and impatient customers. *Journal of the Egyptian Mathematical Society 25*, 3 (2017), 337–342.

[3] Ammar, S. I., and Rajadurai, P. Performance analysis of preemptive priority retrial, queueing system with disaster under working breakdown services. *Symmetry 11*, 3 (2019), 419.

[4] Atencia, I., and Moreno, P. The discrete-time *Geo*/*Geo*/1 queue with negative customers and disasters. *Computers & Operations Research 31*, 9 (2004), 1537–1548.

[5] Bouchentouf, A. A., Boualem, M., Yahiaoui, L., and Ahmad, H. A multi-station unreliable machine model with working vacation policy and customers' impatience. *Quality Technology & Quantitative Management 19*, 6 (2022), 766–796.

[6] Bouchentouf, A. A., Cherfaoui, M., and Boualem, M. Performance and economic analysis of a single server feedback queueing model with vacation and impatient customers. *OPSEARCH 56* (2019), 300–323.

[7] Bouchentouf, A. A., Cherfaoui, M., and Boualem, M. Analysis and performance evaluation of Markovian feedback multi–server queueing model with vacation and impatience. *American Journal of Mathematical and Management Sciences 40*, 3 (2021), 261–282.

[8] Bouchentouf, A. A., Medjahri, L., Boualem, M., and Kumar, A. Mathematical analysis of a Markovian multi-server feedback queue with a variant of multiple vacations, balking and reneging. *Discrete and Continuous Models and Applied Computational Science 30*, 1 (2022), 21–38.

[9] Chen, A., and Renshaw, E. The $M/M/1$ queue with mass exodus and mass arrivals when empty. *Journal of Applied Probability 34*, 1 (1997), 192–207.

[10] Cherfaoui, M., Bouchentouf, A. A., and Boualem, M. Modeling and simulation of Bernoulli feedback queue with general customers' impatience under variant vacation policy *International Journal of Operational Research 46*, 4 (2023), 451–480.

[11] Chew, S. Continuous-service M/M/1 queuing systems. *Applied System Innovation 2*, 2 (2019), 16.

[12] Daigle, J. N. Elementary continuous-time Markov chain-based queueing models. In *Queueing Theory with Applications to Packet Telecommunication*, J. N. Daigle, Ed., Springer US, Boston, MA, 2005, pp. 57–105.

[13] Deena Merit, C. K., and Haridass, M. A simulation study on the necessity of working breakdown in a state dependent bulk arrival queue with disaster and optional re-service. *International Journal of Ad Hoc and Ubiquitous Computing 41*, 1 (2022), 1–15.

[14] Demircioglu, M., Bruneel, H., and Wittevrongel, S. Analysis of a discrete-time queueing model with disasters. *Mathematics 9*, 24 (2021), 3283.

[15] Yue, D., Yue, W., and Sun Y. Performance analysis of an M/M/c/N queueing system with balking, reneging and synchronous vacations of partial servers In *Operations Research and Its Applications. The Sixth International Symposium on Operations Research and Its Applications (ISORA'06) Xinjiang, China, August 8–12, 2006*, X.-S. Zhang, D.-G. Liu and L.-Y. Wu, Eds., pp. 128–143.

[16] Doshi, B. T. Queueing systems with vacations – A survey. *Queueing Systems 1*, 1 (1986), 29–66.

[17] Graves, S. C. The application of queueing theory to continuous perishable inventory systems. *Management Science 28*, 4 (1982), 400–406.

[18] Houalef, M., Bouchentouf, A. A., and Yahiaoui, L. A multi-server queue in a multi-phase random environment with waiting servers and customers' impatience under synchronous working vacation policy. *Journal of the Operations Research Society of China 11*, 3 (2023), 459-487.

[19] Jain, G., and Singman, K. A Pollaczek–Khintchine formula for $M/G/1$ queues with disasters. *Journal of Applied Probability 33*, 4 (1996), 1191–1200.

[20] Jiang, T., Liu, L., and Li, J. Analysis of the $M/G/1$ queue in multi-phase random environment with disasters. *Journal of Mathematical Analysis and Applications 430*, 2 (2015), 857–873.

[21] Kalidass, K., and Kasturi, R. A queue with working brekdowns. *Computers & Industrial Engineering 63*, 4 (2012), 779–783.

[22] Kim, B. K., and Lee, D. H. The M/G/1 queue with diasters and working breakdowns. *Applied Mathematical Modelling 38*, 5-6 (2014), 1788–1798.

[23] Kumar, A., Boualem, M., Bouchentouf, A. A., and Savita. Optimal analysis of machine interference problem with standby, random switching failure, vacation interruption and synchronized reneging. In *Applications of Advanced Optimization Techniques in Industrial Engineering* (2022), A. Goel, A. Chauhan and A. K. Malik, Eds., CRC Press, pp. 155–168.

[24] Lee, D. H., Yang, W. S., and Park, H. M. Geo/G/1 queues with disasters and general repair times. *Applied Mathematical Modelling 35*, 4 (2011), 1561–1570.

[25] Padmavathy, R., Kalidass, K., and Ramanath, K. Vacation queues with impatient customers and a waiting server. *International Journal of Software Engineering and Knowledge Engineering 1*,1 (2011), 10–19.

[26] Rashid, R., Hoseini, S. F., Gholamian, M. R., and Feizabandi, M. Application of queuing theory in production-inventory optimization. *Journal of Industrial Engineering International 11*, 4 (2015), 485–494.

[27] Suranga Sampath, M. I. G., and Liu, J. Impact of customers' impatience on an M/M/1 queueing system subject to differentiated vacations with a waiting server. *Quality Technology & Quantitative Management 17*, 2 (2018), 125–148.

[28] Takagi, H. *Queueing Analysis: A Foundation of Performance Analysis, Volume 1: Vacation and Priority Systems*. Elsevier, Amsterdam, 1991.

[29] Tian, N., and Zhang, Z. G. Vacation queueing models: Theory and applications. *Springer Science & Business Media*. 2006.

[30] Tian, N., Li, Q.-L., and Gao, J. Conditional stochastic decompositions in the M/M/c queue with server vacations. *Communications in Statistics. Stochastic Models 15*, 2 (1999), 367–377.

[31] Towsley, D., and Tripathi, S. K. A single server priority queue with server failures and queue flushing. *Operations Research Letters 10*, 6 (1991), 353–362.

[32] Yang, W. S., and Chae, K. C. A note on $GI/M/1$ queue with poisson negative arrival. *Journal of Applied Probability 38*, 4 (2001), 1081–1085.

[33] Ziad, I., Vijaya Laxmi, P., Bhavant, E. G., Bouchentouf, A. A., and Majid, S. A matrix geometric solution of a multi-server queue with waiting servers and customers' impatience under variant working, vacation and vacation interruption. *Yugoslav Journal of Operations Research 33*, 3 (2023), 389-407.