STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–17.
Published online in International Academic Press (www.IAPress.org)

A Multilayer Perceptron and Cost Optimization of a Single-Server Queue with Bernoulli Feedback and Customer Impatience under a Hybrid Vacation Policy

Houssam Eddine Hamache [1], Louiza Berdjoudj [2,*], Aimen Dehimi [3]

[1,3] *University of Bejaia, Faculty of Exact Sciences, Applied Mathematics Laboratory, 06000 Bejaia, Algeria*
[2] *University of Bejaia, Faculty of Exact Sciences, Research Unit LaMOS, 06000 Bejaia, Algeria*

Abstract This paper deals with a single server queueing system, aiming to handle a hybrid vacation, operating within a finite space, and taking account of Bernoulli feedback and balking, alongside reneging and retention. In case of queue emptiness, coming after a normal busy period, the single server shifts to a working vacation. The server proceeds to take a vacation in case no customers are queued upon the server's return from a working vacation. For analysis purposes, we employed a recursive method to derive the system's steady-state probabilities, thereby facilitating the evaluation of key performance metrics. The numerical results are compared with analytical results and those obtained using a soft computing technique based on a Multilayer Perceptron (MLP) system. Lastly, the Grey Wolf Optimizer is applied to identify the optimal service rates that minimize costs.

Keywords Single-server queue, hybrid vacation, Multilayer Perceptrons (MLP), GWO algorithm, Cost optimization

AMS 2010 subject classifications 60K25, 68M20 ,90B22

DOI: 10.19139/soic-2310-5070-2623

1. Introduction

Queueing systems with server vacation has been the subject of focused study because of the broad range of applications they offer in computer, manufacturing, service, and communication networks as well as production systems. Excellent surveys of the earlier vacation model works were provided in papers by [14, 16], and the book by [39].

Altman and Yechiali [4] examined the $M/M/1$ queue with vacation periods and impatient customers, laying foundational work in this area. Later, Adan et al. [1] studied synchronized reneging in queueing systems with vacations. Five years later, Ibe and Isijola [18] analyzed a single-server system subject to multiple differentiated vacations. Ammar [2] proposed a novel solution framework for the $M/M/1$ queue, incorporating vacations, customer impatience, and a waiting server mechanism. The topic has continued to receive significant attention in the queueing theory literature [3, 8, 10, 13, 22, 36], reflecting its relevance to real-world service systems.

In traditional queueing models with server vacations, it is often assumed that the server ceases all operations. However, in many practical scenarios, a server may remain partially operational during these periods, serving customers at a reduced rate. This concept, known as a working vacation, was pioneered by Servi and Finn [37] in their foundational work.

*Correspondence to: Louiza Berdjoudj (Email: louiza.berdjoudj@univ-bejaia.dz).University of Bejaia, Faculty of Exact Sciences, Research Unit LaMOS, 06000 Bejaia, Algeria.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Building on this foundational concept, the research community quickly began to explore its variations and apply it to more complex scenarios. Wu and Takagi [40] extended the model to an $M/G/1$ queue with multiple working vacations, while Li and Tian [27] investigated an $M/M/1$ queue where vacations could be interrupted, adding a new layer of realism.

The analytical tools used to study these systems also evolved. Li et al. [28] successfully applied the supplementary variable method to analyze a $GI/M/1$ working vacation model. For more complex environments, Jain and Upadhyaya [19] utilized the matrix-geometric method to examine an unreliable multi-server system with synchronous working vacations. A crucial aspect was introduced by Yue et al. [42], who were among the first to integrate customer impatience with the working vacation framework, a theme central to our work.

More recent advancements and comprehensive discussions on Markovian queueing systems with working vacations can be found in the works of Selvaraju and Goswami [38] and Bouchentouf et al. [9], with further relevant studies available in [6, 7, 11].

The characteristic of customer impatience is of significant importance in queueing theory [23, 24, 25]. This behavior primarily manifests as balking (when customers refuse to join a queue) and reneging (when they leave after joining), along with the related concept of customer retention, where the system persuades a reneging customer to stay. It has been established that this research area is both intriguing and challenging, particularly in contexts such as customer service operations for online retailers, critical patient handling in hospital emergency rooms, as well as inventory systems, and various pertinent domains.

In the literature on queueing systems, significant attention has been devoted to models that incorporate both working vacations and customer impatience. For example, Laxmi et al. [26] conducted a comprehensive analysis of a finite-capacity Markovian queueing system incorporating customer impatience behaviors, balking, and reneging, alongside a working vacation policy. Building on this, Majid and Manoharan [29] studied an $M/M/c$ queueing model incorporating customer impatience and synchronous working vacations, underlining the impact of collective server unavailability and time-dependent reneging on the system's stability and performance metrics. Subsequently, Majid [30] further examined an $M/M/1$ queueing model that also integrated customer impatience and working vacations, offering valuable insights into how reneging behavior and reduced service efficiency during vacation periods influence overall system performance.

Research has also explored systems with added layers of operational complexity. For instance, Bouchentouf et al. [12] studied a multi-station unreliable machine model incorporating working vacations, customer impatience, and retention mechanisms, offering analytical and computational insights into system reliability and performance under complex service dynamics. Similarly, Dehimi et al. [15] analyzed a multi-server queueing system with customer impatience under differentiated working vacation policies, providing both analytical and computational insights. More specialized contexts have also been addressed, such as the finite-capacity $M/M/2$ machine repair model investigated by Kadi et al. [21], which featured a triadic service discipline and dual working vacation policies.

Addressing more dynamic customer behaviors, Narmadha and Rajendran [32] analyzed an $M/M/1/K$ queueing model incorporating encouraged arrivals, single working vacations, and customer impatience, demonstrating how dynamic arrival patterns and reneging behavior influence system performance. Further extending the study of complex interactions, the work of Kadi et al. [20] is of particular relevance as it incorporates feedback mechanisms. They developed an $M/M/1/K$ queueing model incorporating working vacations, customer feedback, and impatience behavior under an N-policy, offering an integrated analytical and optimization framework for managing finite-capacity service systems.

Various aspects of queueing systems have been explored in previous research, stating vacation policies analysis, and impatience. Different vacation policies have been inspected for the system's main performance determination, and specifically the influence on customer waiting times. The goal of studying vacations is to collect insights about their impact on service quality. Whereas the investigation of impatience in queueing systems serves the analysis of how customer behavior abandonment implies on system's design. Nonetheless, to our information scope, literature on the holistic consideration of various features, such as vacation, working vacation, balking, reneging, retention, and feedback, within the context of different vacation queueing models incorporating vacation policies remains with gaps. Even in recent studies, the mentioned features lack collective exploration, therefore, the rationale for this research is to treat this specific lack by inspecting a comprehensive system with a simultaneous combination of all the cited features. The study of this integrated system contributes to queueing theory improvement within this context.

In addressing the complexities of queueing systems, artificial neural networks, particularly Multi-Layer Perceptrons (MLPs), provide a robust approach for capturing nonlinear patterns. As a class of deep learning models, MLPs consist of multiple interconnected layers that enable them to model complex relationships within data. By leveraging hidden layers and non-linear activation functions, they excel in tasks such as classification, regression, and pattern recognition. The evolution of neural networks began with Rosenblatt's [34] perceptron model and was significantly enhanced by Rumelhart et al. [35] through the development of the backpropagation algorithm, which allowed for efficient training of MLPs. Given their adaptability and learning efficiency, MLPs have been widely applied across various fields, including science, healthcare, finance, and engineering [5, 33, 41], where data-driven decision-making is crucial.

Our modest contribution to this topic extends the model discussed in [17] by incorporating hybrid vacations. In this approach, the server takes a working vacation upon emptying the queue and transitions to a vacation period when returning from the working vacation to find an empty system. By employing recursive methods, we calculate steady-state probabilities and derive various performance measures. Graphical representations illustrate how system parameters impact performance metrics. Furthermore, MLP-based results are also obtained for comparing the numerical results obtained with the analytical results. Recognizing the importance of optimization, we develop a cost function and formulate an optimization problem to determine the optimal service rates. This is achieved through the application of swarm intelligence, specifically the Grey Wolf Optimizer (GWO) algorithm. The numerical results further enhance the expected cost function. To situate our contribution within the existing literature, Table 1 provides a comparative overview of the characteristics of our proposed model against several relevant recent works.

Table 1. Comparison of queueing models based on system characteristics

	Models			
System characteristics	**Bouchentouf et al. [9]**	**Gupta [17]**	**Keerthana et al. [22]**	**Proposed model**
Working vacation	✓	✓	×	✓
Complete vacation	×	×	✓	✓
Balking	✓	✓	×	✓
Reneging	✓	✓	×	✓
Retenetion	×	×	×	✓
Feedback	×	✓	×	✓
Optimization	Quadratic fit search method	×	PSO	GWO
AI method	×	×	×	MLP

This paper is structured as follows: Section 2 introduces the mathematical description of the suggested queueing model. Section 3 presents the steady-state distribution of the system. In Section 4, key performance metrics are explored alongside the Multilayer Perceptron (MLP) approach. Section 5 provides graphical illustrations to analyze the impact of various system parameters on performance metrics and further discusses the MLP method. Section 6 presents the development of a cost optimization, where the Grey Wolf Optimizer (GWO) method is employed to minimize the cost function. Finally, Section 7 offers concluding insights and discusses the implications for future studies

2. Mathematical model

Consider a $M/M/1/K$ queue with hybrid vacation, Bernoulli feedback, balking, reneging, and retention. This queueing system is founded on the following basic assumptions:

The Arrival Pattern

- ▷ Customers arrive at the system according to a Poisson process at an average rate of λ.

- ▷ The system's assumed finite capacity is denoted as K.
- ▷ Upon a customer's arrival, a decision is made based on the following probabilities: The customer either joins the queue with probability β_k or balks (refuses to join) with the complementary probability $\beta_k^{'} = 1 - \beta_k$. The probabilities β_k adhere to the conditions $0 \leq \beta_{k+1} \leq \beta_k \leq 1$, $1 \leq k \leq K - 1$. Specifically, $\beta_0 = 1$ and $\beta_K = 0$.

The Service Pattern

- ▷ During the normal busy period, the service time follows an exponential distribution with rate μ_b. In the working vacation period, the service time also follows an exponential distribution, but with a lower rate μ_w ($\mu_w < \mu_b$).
- ▷ Customers are served respectively by the FCFS (First-Come-First-Served) principle.

The hybrid vacation mechanism

- ▷ A hybrid vacation combines both a working vacation (WV) and a complete vacation (CV). During the vacation period, the server first enters a working vacation, where it continues to serve customers but at a reduced service rate. The duration of the working vacation follows an exponential distribution with parameter γ. When the working vacation ends, if no customers are waiting, the server takes a complete vacation, with its duration following an exponential distribution with parameter η. If customers are present during the working vacation, the server remains in the working vacation and continues serving them. Once the complete vacation is over, the server resumes normal service for customers waiting in the queue.

Customer behavior and impatience

- ▷ Customers show impatience while waiting for service. Whether during a busy period, a working vacation, or a complete vacation, each customer is associated with an internal patience timer (T_0, T_1, and T_2), respectively, which follows an exponential distribution. The impatience rates (ζ_0, ζ_1, and ζ_2) depend on the server's current state. If a customer's patience runs out, they may abandon the queue (renege). In such cases, the system gets a final chance to convince the customer to stay: the customer leaves permanently with probability α, or may be retained in the queue with probability $\alpha^{'} = 1 - \alpha$.

The Feedback Rule

- ▷ If a customer is dissatisfied with the service, they may choose to leave the system with probability θ, or return with probability $\theta^{'} = 1 - \theta$. Customers who return are treated as new arrivals.

The variables introduced are mutually independent.

2.1. Practical example

This study delves into a $M/M/1/K$ queueing system tailored for a call center environment, where the dynamics of customer arrivals, service times, and operational states are meticulously examined. Incoming calls follow a Poisson process, with service durations exponentially distributed; these times are shorter during normal operations μ_b and longer during reduced-capacity periods ($\mu_w < \mu_b$). The call center can manage up to K simultaneous calls, with customers having the option to balk if all lines are occupied, governed by a probability β_k. Operational modes include hybrid vacations, alternating between working (at reduced capacity) and complete vacation (no service) depending on queue congestion and staffing levels. Dissatisfied customers can either exit the queue θ or return later $\theta^{'} = 1 - \theta$, treated as new arrivals upon re-engagement. Additionally, during various operational phases, customers are subject to impatience timers (T_0, T_1, T_2), abandoning the queue with probability α if service does not commence before timer expiration (reneging), or remaining in the queue with probability $\alpha^{'} = 1 - \alpha$ (retention). This model offers insights into optimizing call center efficiency by balancing service capacity, customer satisfaction, and operational flexibility under varying demand conditions.

3. Steady-state Solution

We will now examine the bi-variate process $\{(L(t),\ S(t)),\ t \geq 0\}$, where $L(t)$ represents customers roster in the system at time t, and $S(t)$ is the server's status at time t, which can take one of three values, such as $S(t) = 0$: when the servers are in normal busy period at time t, $S(t) = 1$ when the server is in a working vacation period at time t, and $S(t) = 2$: when the server is on vacation at time t. The combined probability $P_{k,j} = \lim_{t\to\infty} P\left(L(t) = k,\ S(t) = j\right)$, denotes the steady-state probabilities of the system.
Where
$(k, j) \in \{\{(k, 0) :\ k = 1, ..., K\} \cup \{(k, 1) :\ k = 0, ..., K\} \cup \{(k, 2) :\ k = 0, ..., K\}\}$. Figure 1 illustrates the transitions in the model represented by a diagram. Using the principle of balance equations

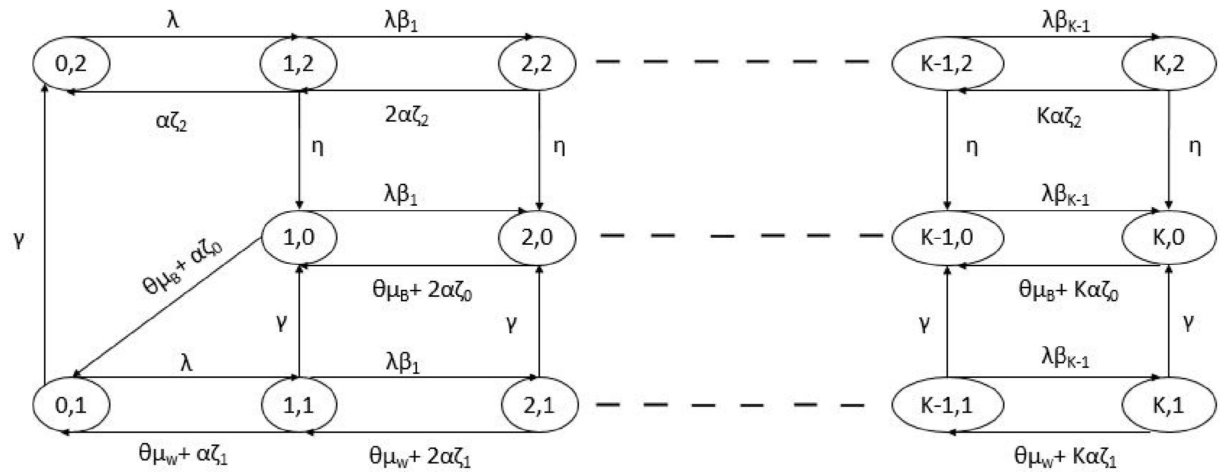


Figure 1. State transition diagram.

$$(\lambda\beta_1 + \mu_b\theta + \alpha\zeta_0)P_{1,0} = (\theta\mu_b + 2\alpha\zeta_0)P_{2,0} + \eta P_{1,2} + \gamma P_{1,1},\ \ k = 1, \tag{1}$$

$$\begin{aligned}(\lambda\beta_k + \theta\mu_b + k\zeta_0)P_{k,0} = \lambda\beta_{k-1}P_{k-1,0} + (\theta\mu_b + (k+1)\alpha\zeta_0)P_{k+1,0} + \eta P_{k,2} + \gamma P_{k,1},\\ 2 \leq k \leq K-1,\end{aligned} \tag{2}$$

$$(\theta\mu_b + K\alpha\zeta_0)P_{K,0} = \lambda\beta_{k-1}P_{k-1,0} + \eta P_{K,2} + \gamma P_{K,1},\ \ k = K, \tag{3}$$

$$(\lambda + \gamma)P_{0,1} = (\theta\mu_b + \alpha\zeta)P_{1,1} + (\theta\mu_b + \alpha\zeta_0)P_{1,0},\ \ k = 0, \tag{4}$$

$$(\lambda\beta_1 + \gamma + \theta\mu_w + 2\alpha\zeta_1)P_{1,1} = \lambda P_{0,1} + (\theta\mu_w + 2\alpha\zeta_1)P_{2,1},\ \ k = 1, \tag{5}$$

$$\begin{aligned}(\lambda\beta_k + \gamma + \theta\mu_w + k\alpha\zeta_1)P_{k,1} = \lambda\beta_{k-1}P_{k-1,1} + (\theta\mu_w + (n+1)\alpha\zeta_1)P_{k+1,1},\\ 2 \leq k \leq K-1,\end{aligned} \tag{6}$$

$$(\gamma + \theta\mu_w + K\alpha\zeta_1)P_{K,1} = \lambda\beta_{K-1}P_{K-1,1}, \qquad k = K, \tag{7}$$

$$\lambda P_{0,2} = \gamma P_{0,1} + \alpha\zeta_2 P_{1,2},\ \ k = 0, \tag{8}$$

$$(\lambda\beta_1 + \alpha\zeta_2 + \eta)P_{1,2} = \lambda P_{0,2} + 2\alpha\zeta_2 P_{2,2}, \qquad k = 1, \tag{9}$$

$$\begin{aligned}(\lambda\beta_k + k\alpha\zeta_2 + \eta)P_{k,2} = \lambda\beta_{k-1}P_{k-1,2} + (k+1)\alpha\zeta_2 P_{k+1,2},\\ 2 \leq k \leq K-1,\end{aligned} \tag{10}$$

$$(\eta + K\alpha\zeta_2)P_{K,2} = \lambda\beta_{K-1}P_{K-1,2}, \qquad k = K. \tag{11}$$

The normalizing condition is

$$\sum_{k=0}^{K} (P_{k,0} + P_{k,1} + P_{k,2}) = 1. \tag{12}$$

Presented below is the theorem outlining the solution to the above equations.

Theorem 1. The steady-state probabilities representing the system size during various operational phases, specifically the vacation period ($P_{2,k}$), the working vacation period ($P_{1,k}$), and the regular busy period ($P_{0,k}$), are given respectively and expressed as follows:

$$\begin{aligned} P_{k,2} &= \Gamma_k P_{K,2}, \\ &= \Gamma_k \left(\sum_{k=1}^{K} (\varrho_2 \psi_k - \phi_k) + \sum_{k=0}^{K} (\varrho_1 \delta_k + \Gamma_k) \right)^{-1}. \end{aligned} \tag{13}$$

$$P_{k,1} = \varrho_1 \delta_k P_{K,2}. \tag{14}$$

$$P_{k,0} = (\varrho_2 \psi_k - \phi_k) P_{K,2}, \tag{15}$$

where

$$\Gamma_k = \begin{cases} 1, & k = K, \\ \frac{K\alpha\zeta_2 + \eta}{\lambda\beta_{K-1}}, & k = K-1, \\ \frac{\lambda\beta_{k+1} + (k+1)\alpha\zeta_2 + \eta}{\lambda\beta_k} \Gamma_{k+1} - \frac{(k+2)\alpha\zeta_2}{\lambda\beta_k} \Gamma_{k+2}, & 0 \le k < K-2, \end{cases} \tag{16}$$

$$\delta_k = \begin{cases} 1, & k = K, \\ \frac{\gamma + \theta\mu_w + K\alpha\zeta_1}{\lambda\beta_{K-1}}, & k = K-1, \\ \frac{\lambda\beta_{k+1} + \gamma + \theta\mu_w + (k+1)\alpha\zeta_1}{\lambda\beta_k} \delta_{k+1} - \frac{\theta\mu_w + (k+2)\alpha\zeta_1}{\lambda\beta_k} \delta_{k+2}, & 0 \le k < K-2, \end{cases} \tag{17}$$

$$\varrho_1 = \frac{\lambda\Gamma_0 - \alpha\zeta_2\Gamma_1}{\gamma\delta_0}. \tag{18}$$

$$\psi_k = \begin{cases} 1, & k = K, \\ \frac{\theta\mu_b + K\alpha\zeta_0}{\lambda\beta_{K-1}}, & k = K-1, \\ \frac{\lambda\beta_{k+1} + \theta\mu_b + (k+1)\alpha\zeta_0}{\lambda\beta_k} \psi_{k+1} - \frac{\theta\mu_b + (k+2)\alpha\zeta_0}{\lambda\beta_k} \psi_{k+2}, & 1 \le k < K-2, \end{cases}$$

$$\phi_k = \begin{cases} 0, & k = K, \\ \frac{\eta + \gamma\varrho_1}{\lambda\beta_{K-1}}, & k = K-1, \\ \frac{\eta\Gamma_{k+1} + \gamma\varrho_1\delta_{k+1}}{\lambda\beta_k}, & 0 \le k < K-2, \end{cases}$$

$$\varrho_2 = \frac{\delta_0\varrho_1(\lambda + \gamma) - \delta_1(\theta\mu_w + \alpha\zeta_1)\varrho_1 + \phi_1(\theta\mu_b + \alpha\zeta_0)}{\psi_1(\theta\mu_b + \alpha\zeta_0)}, \tag{19}$$

and

$$P_{K,2} = \left(\sum_{k=1}^{K} (\varrho_2 \psi_k - \phi_k) + \sum_{k=0}^{K} (\varrho_1 \delta_k + \Gamma_k) \right)^{-1}. \tag{20}$$

Proof
by solving equations recursively (9) − (11), we find $P_{k,2} = \Gamma_k P_{K,2}$, such that (16) represent Γ_k. By equations (5) − (7), we get $P_{k,1} = \delta_k P_{K,1}$, such that (17) represent δ_k. We use equation (8) and we obtain (14) − (18). Via equations (2) − (3), we obtain $P_{k,0}$ in terms of $P_{K,0}$ and $P_{K,2}$. Using (4), we can obtain $P_{k,0}$ in terms of $P_{K,2}$ that is given by (15). Finally, by applying the normalization condition we derive equation (20).

4. Performance measures

▷ The probabilities representing the different server states–busy normal period, working vacation, and vacation–are specified as follows:

$$P_b = P_{K,2} \sum_{k=1}^{K} (\varrho_2 \psi_k - \phi_k).$$

$$P_{wv} = \varrho_1 P_{K,2} \sum_{k=0}^{K} \delta_k.$$

$$P_v = P_{K,2} \sum_{k=0}^{K} \Gamma_k.$$

▷ The formulas for the expected number of customers in the system (L_s) and in the queue (L_q) are defined as follows:

$$\begin{aligned} L_s &= \sum_{k=0}^{K} k(P_{k,0} + P_{k,1} + P_{k,2}) \\ &= P_{K,2} \left[\sum_{k=1}^{K} (\varrho_2 k \psi_k - k\phi_k + \varrho_1 k \delta_k + k\Gamma_k) \right]. \end{aligned} \tag{21}$$

$$\begin{aligned} L_q &= \sum_{k=1}^{K} (k-1)(P_{k,0} + P_{k,1}) + \sum_{k=1}^{K} k P_{k,2} \\ &= P_{K,2} \left[\sum_{k=1}^{K} (\varrho_2 (k-1)\psi_k - (k-1)\phi_k + \varrho_1 (k-1)\delta_k + k\Gamma_k) \right]. \end{aligned} \tag{22}$$

▷ The expected balking rate:

$$\begin{aligned} B_r &= \lambda \sum_{k=1}^{K} (1-\beta_k)(P_{k,0} + P_{k,1} + P_{k,2}) \\ &= \lambda P_{K,2} \left[\sum_{k=1}^{K} (\varrho_2 \beta_k^{'} \psi_k - \beta_k^{'} \phi_k + \varrho_1 \beta_k^{'} \delta_k + \beta_k^{'} \Gamma_k) \right]. \end{aligned} \tag{23}$$

▷ The formulas for the expected waiting time of customers in the system (W_s) and in the queue (W_q) are given by:

$$W_s = \frac{L_s}{\lambda^{'}}, \ \text{ where } \lambda^{'} = \lambda - B_r, \ \ W_q = \frac{L_q}{\lambda^{'}}. \tag{24}$$

▷ The expected reneging rate:

$$
\begin{aligned}
R_r &= \alpha\zeta_0 \sum_{k=1}^{K} kP_{k,0} + \alpha\zeta_1 \sum_{k=1}^{K} kP_{k,1} + \alpha\zeta_2 \sum_{k=1}^{K} kP_{2,k} \\
&= \alpha P_{K,2} \left[\sum_{k=1}^{K} (\zeta_0 \varrho_2 k\psi_k - \zeta_0 k\phi_k + \zeta_1 \Theta_1 k\delta_k) \right] \\
&+ \alpha P_{K,2} \left[\zeta_2 \sum_{k=1}^{K} k\Gamma_k \right].
\end{aligned}
\tag{25}
$$

▷ The expected retention rate:

$$
\begin{aligned}
R_t &= \alpha^{'}\zeta_0 \sum_{k=1}^{K} kP_{k,0} + \alpha\zeta_1 \sum_{k=1}^{K} kP_{k,1} + \alpha\zeta_2 \sum_{k=1}^{K} kP_{2,k} \\
&= \alpha^{'} P_{K,2} \left[\sum_{k=1}^{K} (\zeta_0 \varrho_2 k\psi_k - \zeta_0 k\phi_k + \zeta_1 \Theta_1 k\delta_k) \right] \\
&+ \alpha^{'} P_{K,2} \left[\zeta_2 \sum_{k=1}^{K} k\Gamma_k \right].
\end{aligned}
\tag{26}
$$

4.1. Multilayer Perceptrons (MLP)

Multilayer Perceptrons (MLP) are a class of artificial neural networks designed to model complex relationships in data through multiple layers of interconnected neurons. Each neuron processes inputs using weighted connections and activation functions, enabling nonlinear transformations. The foundation of MLPs traces back to Frank Rosenblatt's Perceptron [34], but their full potential was realized with the development of the backpropagation algorithm by Rumelhart, Hinton, and Williams (1986) [35], which allowed efficient training of deep networks. As a key component of soft computing, MLPs are widely applied in classification, regression, and pattern recognition, playing a crucial role in fields such as finance, healthcare, and engineering. The MLP model was implemented with three hidden layers, each containing 64 neurons, using the ReLU activation function. The output layer consists of a single neuron with a linear activation function for regression. The model was trained for 420 epochs with a batch size of 32, using the Adam optimizer with a learning rate of 0.001. The Mean Squared Error (MSE) was used as the loss function. The dataset for training was generated from our analytical model by creating 4001 data points based on combinations of the arrival rate (λ) and service rate (μ_b). These data points were divided into training (80%) and validation (20%) sets to ensure robust model performance.

The following algorithm describes the training process of a Multilayer Perceptron (MLP):

Algorithm 1 Multilayer Perceptron (MLP) Training Algorithm

Input: λ, μ_b
Output: L_q.
Step 1: Initialize MLP Model
- Define network architecture (input, hidden, and output layers).
- Initialize weights and biases randomly.

Step 2: Training Process
for $t = 1$ *to* $MaxIter$ **do**
  Forward Pass: Compute activations for each layer.
  Compute Loss: Evaluate the difference between predictions and actual values.
  Backward Pass: Compute gradients using backpropagation.
   Update weights and biases using Adam optimizer.

Step 3: Output the Optimized MLP Model
Return L_q.

5. Performance and sensitivity analysis

5.1. Performance analysis

The objective of this subsection of the study is to empirically validate the analytical results presented earlier while also demonstrating the effect of various parameters. A computational experiment was conducted on diversified performance measures of interest governed by different scenarios, and to accomplish this, numerical results are depicted through the graphs. The following cases are being considered to achieve this objective:

- ♦ $\lambda = 0.01 : .01 : 5,\ \mu_b = 2,\ \mu_w = 1,\ \gamma = 1.5,\ \alpha = 0.7,\ \theta = 0.5,\ K = [5; 15; 25],\ \eta = 0.5,\ \zeta_0 = 1,\ \zeta_1 = 1.5,\ \zeta_2 = 2,\ \ \beta_k = 1 - k/K.$

- ♦ $\lambda = 3.5,\ \mu_b = 2,\ \mu_w = 1,\ \gamma = 1.5,\ \alpha = 0.7,\ \theta = 0.5,\ K = 10,\ \eta = [0.01 : .01 : 2],\ \zeta_0 = 1,\ \zeta_1 = 1.5$ $,\ \zeta_2 = 2,\ \ \beta_k = 1 - k/K.$

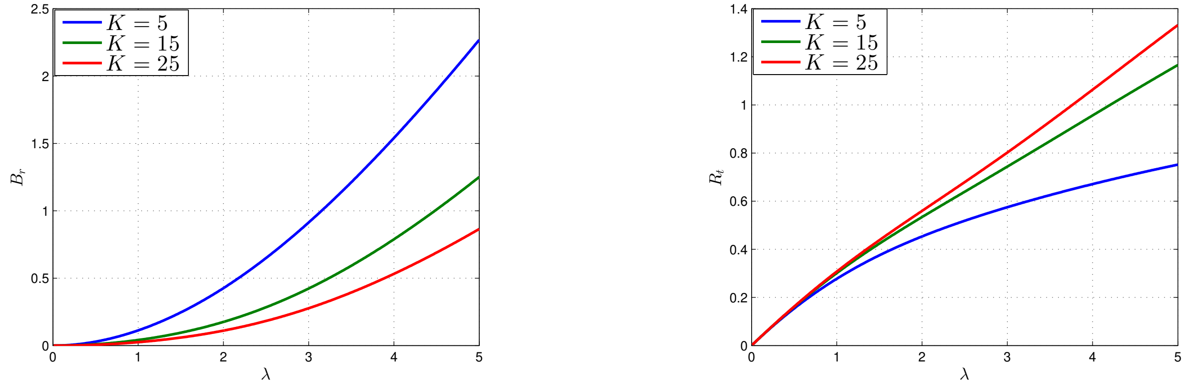
- ♦ $\lambda = 3.5,\ \mu_b = 2,\ \mu_w = 1,\ \gamma = [0.01 : .01 : 3],\ \alpha = 0.7,\ \theta = 0.5,\ K = 10,\ \eta = 0.5,\ \zeta_0 = 1,\ \zeta_1 = 1.5,\ \zeta_2 = 2,\ \ \beta_k = 1 - k/K.$

- ♦ $\lambda = 3.5,\ \mu_b = [0.01 : .01 : 4],\ \mu_w = 1,\ \gamma = 1.5,\ \alpha = [0.1; 0.5; 0.9],\ \theta = 0.5,\ K = 10,\ \ \eta = 0.5,\ \zeta_0 = 1,\ \zeta_1 = 1.5,\ \zeta_2 = 2,\ \ \beta_k = 1 - k/K.$

- ♦ $\lambda = 3.5,\ \mu_b = 2,\ \mu_w = 1,\ \gamma = [0.01 : .01 : 3],\ \alpha = 0.7,\ \theta = 0.5,\ K = 10,\ \eta = 0.5,\ \zeta_0 = [1; 1.3; 1.5],\ \zeta_1 = 1.8,\ \zeta_2 = 2,\ \ \beta_k = 1 - k/K.$

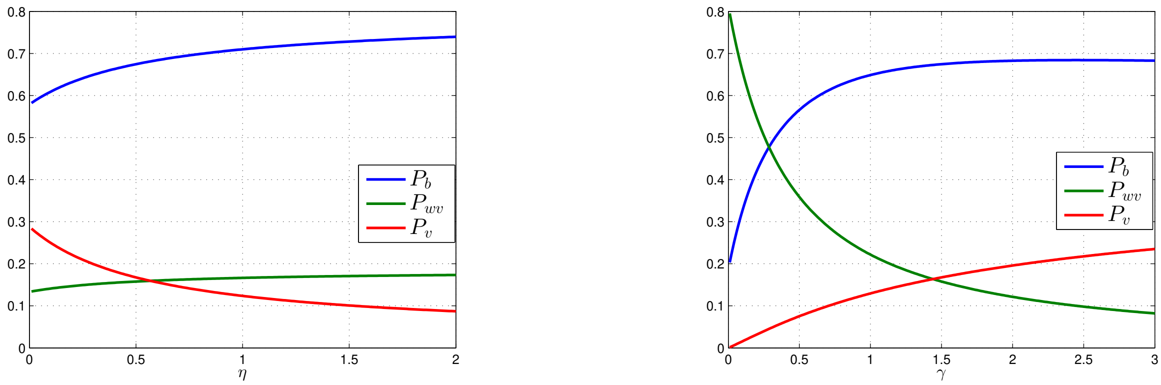
- ♦ $\lambda = 3.5,\ \mu_b = 2,\ \mu_w = 1,\ \gamma = [0.01 : .01 : 3],\ \alpha = 0.7,\ \theta = 0.5,\ K = 10,\ \eta = 0.5,\ \zeta_0 = 1,\ \zeta_1 = [1.5; 1.8; 2.1],\ \zeta_2 = 2,\ \ \beta_k = 1 - k/K.$

- ♦ $\lambda = 3.5,\ \mu_b = 2,\ \mu_w = 1,\ \gamma = [0.01 : .01 : 3],\ \alpha = 0.7,\ \theta = 0.5,\ K = 10,\ \eta = 0.5,\ \zeta_0 = 1,\ \zeta_1 = 1.5,\ \zeta_2 = [2; 2.3; 2.6],\ \ \beta_k = 1 - k/K.$



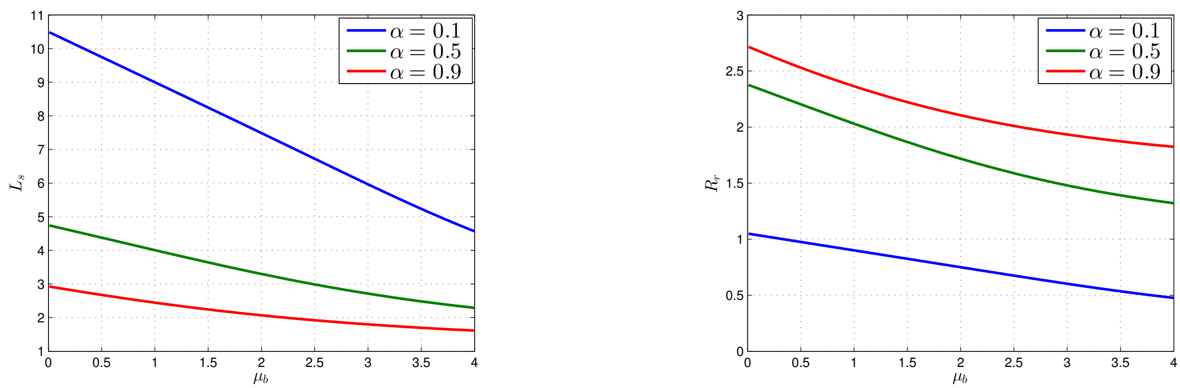
(a) (b)

Figure 2. Effect of λ on B_r and R_t for different values of K.



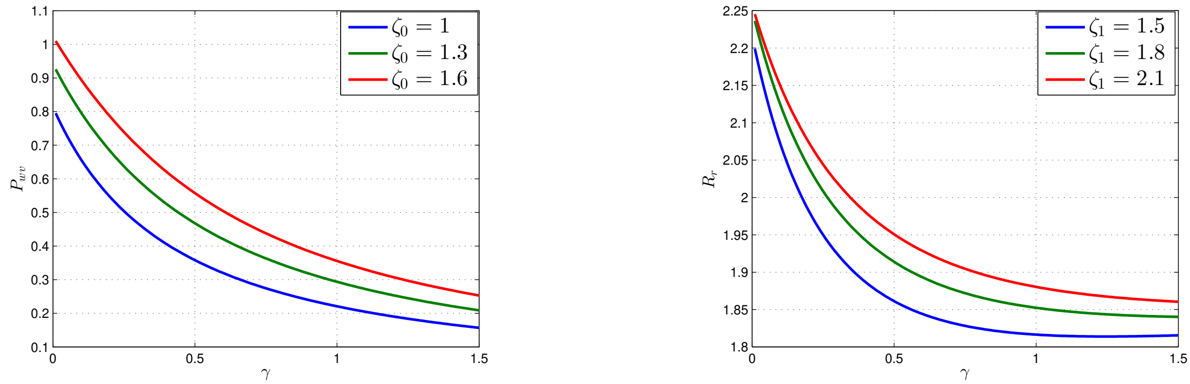
(a) (b)

Figure 3. The probabilities of system in different states according to η and γ.



(a) (b)

Figure 4. Effect of λ on L_s and R_r for different values of α.



(a) (b)

Figure 5. Effect of γ on P_{wv} (for different ζ_0 values) and on R_r (for different ζ_1 values).

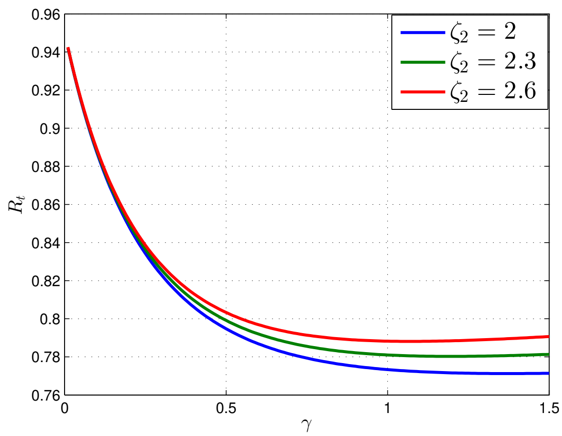


Figure 6. Effect of γ on R_t for different ζ_2 values.

- ▷ Arrival rate and system capacity impact: The system size shows a state of increase when λ is large to a certain degree. This leads to a high probability of normal busy period P_b. Additionally, the average balking and average retention rates are seen to increase (see Figures 2a and 2b). As K escalates, more customers are incentivized to merge into the queue, leading to a fall in the average balking rate (see Figure 2a). However, this scenario also results in a higher average rate of customers abandoning the queue prematurely, Consequently, the system adopts a specific strategy to retain these customers, thereby increasing the average retention rate (see Figure 2b).

- ▷ The effect of working vacation rate(γ): As the working vacation rate γ increases, the system transitions more rapidly between periods of normal business and vacation. This results in a growth in the probabilities of the normal busy period Pb and vacation period P_v, while the probability of the working vacation period P_{wv} drops (see Figures 3b and 5a). Furthermore, an increase in the working vacation rate leads to a shorter average waiting time for customers. Consequently, this reduction in waiting time results in a decrease in both the average reneging rate and the average retention rate (see Figures 5b and 6).

▷ The effect of vacation rate (η): As the vacation rate rises higher, the probability of vacation tends to drop, and the probability of the normal busy period is in a growth state. Additionally, the probability of a working vacation also increases, albeit at a slower rate.

▷ The effect of service rate and (μ_b) and non-retention (α): As the service rate (μ_b) increases to a certain extent, both the mean number of customers and the average reneging tend to decrease (see Figures 4a and 4b). Furthermore, a higher probability of non-retention results in a decrease in L_s (see Figure 4a) while simultaneously high the average reneging (see Figure 4b).

▷ The effect of impatience rates ζ_i: As the impatience rates ζ_i $(i = 0, 1, 2)$ increase, whether during the normal busy period, working vacation, or complete vacation, they become crucial factors that directly influence key system outcomes. These include increasing the average rates of reneging and retention and affecting the probability of the server being in a working vacation state (see Figures 5a, 5b, and 6).

5.2. Results MLP

This part presents the results obtained from both the Multilayer Perceptrons (MLP) model and the analytical queueing approach. The MLP was trained using system parameters such as arrival rate and service time to estimate queue length. After the training phase, the model's predictions were evaluated and compared to the analytical results derived from classical queuing theory formulas. The comparison reveals the degree of accuracy in MLP predictions and highlights any deviations from the theoretical values.

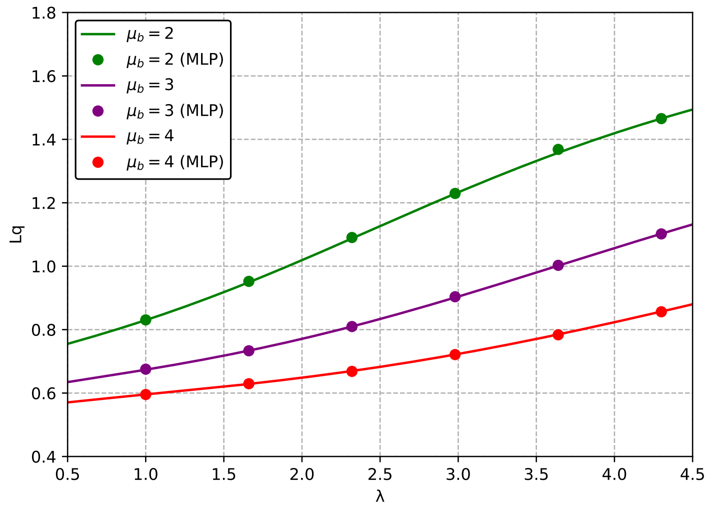


Figure 7. Effect on L_q of λ by varying μ_b

– From Figures 7-8, it can be observed that as the arrival rate increases, the average number of customers in the queue also rises, aligning with real-world scenarios. Additionally, Figure 7 illustrates that an increase in the service rate μ_b leads to a decrease in L_q. Furthermore, Figure 8 demonstrates that the average number of customers in the queue decreases as the probability of non-feedback (θ) increases. Figure 9 indicates that an increase in the service rate during busy periods leads to a reduction in L_q, whereas a rise in arrival rate corresponds to an increase in L_q. Table 2 summarizes the performance of the Multilayer Perceptron (MLP) model for key hyperparameters (θ, λ, and μ_b). The results demonstrate that the model achieves high accuracy, as indicated by the low Root Mean Square Error (RMSE) values and R^2 coefficients consistently

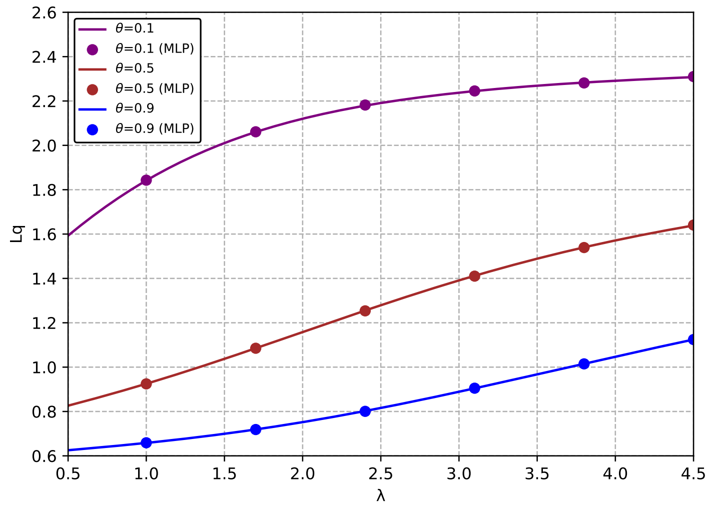


Figure 8. Effect on L_q of λ by varying θ

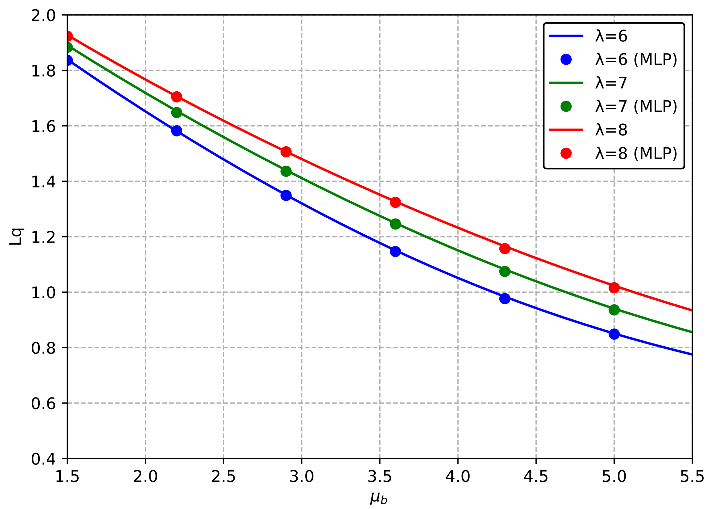


Figure 9. Effect on L_q of μ_b by varying λ

close to unity across the training, validation, and full datasets. These findings are consistent with the graphical results in Figures 7, 8, and 9, where MLP predictions closely match the analytical solutions across various parameter settings. The minimal discrepancies between predicted and theoretical values confirm the MLP model's robustness in capturing the nonlinear dynamics of the system and providing accurate queue length estimations, underscoring its value as a reliable alternative to traditional analytical methods for complex queueing dynamics.

Table 2. Performance Metrics for MLP Regression Under Different Hyperparameters

Hyperparameter	Value	Train RMSE	Train R^2	Val RMSE	Val R^2	Full RMSE	Full R^2
μ_b	2	0.0030493	0.999855	0.0022239	0.999904	0.0029017	0.999864
	3	0.0031152	0.999683	0.0033828	0.999537	0.0031710	0.999659
	4	0.0013290	0.999852	0.0013043	0.999821	0.0013240	0.999847
θ	0.1	0.0026849	0.999815	0.0027139	0.999725	0.0026908	0.999802
	0.5	0.0044138	0.999697	0.0040089	0.999696	0.0043350	0.999697
	0.9	0.0012186	0.999936	0.0012772	0.999921	0.0012307	0.999934
λ	6	0.0056883	0.999680	0.0058633	0.999570	0.0057241	0.999662
	7	0.0038856	0.999840	0.0038714	0.999802	0.0038828	0.999834
	8	0.0035342	0.999855	0.0034197	0.999831	0.0035114	0.999851

6. Cost optimization

We present a model to quantify the costs incurred in our model. In this context, the first step is to define the total pricing that the system costs per unit of time:

$$\begin{aligned}\Lambda(\mu_b,\mu_w) &= C_bP_b + C_vP_v + C_{wv}P_{wv} + C_qL_q + C_r(R_r+B_r)\\ &\quad + C_tR_t + \mu_bC_{\mu_b} + \mu_wC_{\mu_w} + \theta^{'}(\mu_b+\mu_w)C_f + C_a.\end{aligned}$$

- C_b (resp. C_v): Denotes the cost per time unit when the server is in work state (resp. vacation) during normal busy periods.
- C_{wv} (resp.C_q): Presents the cost per time unit when the server is in the work state during the working vacation period (resp. in case of customer waiting in the queue).
- C_r (resp. C_t): Denotes the cost per time unit during the loss of the customer (resp. retains),
- C_{μ_b} (resp. C_{μ_w}): Determines the servicing cost per time unit in a normal busy period (resp. during the working vacation period).
- C_f: The servicing cost per time unit for a feedback customer.
- C_a The purchase cost of the server is fixed per unit.

6.1. *Grey Wolf Optimizer*

The Grey Wolf Optimizer (GWO) algorithm was first introduced by [31], inspired by the social hierarchy and hunting behavior of grey wolves in nature. Its simplicity, adaptability, and proven effectiveness make it a valuable tool for researchers across various disciplines. The GWO algorithm demonstrates strong potential in solving complex optimization problems by efficiently exploring the search space while ensuring rapid convergence. This novel technique was employed to explore (μ_b,μ_w) in the purpose of searching the global minimum of $\Lambda(\mu_b,\mu_w)$.

6.2. *Numerical cost optimum*

This subsection seeks the minimization of the total cost expected to be incurred by the system. To solve the optimization problem, the Grey Wolf Optimizer (GWO) was employed. The GWO algorithm was configured with a population size of 30 wolves and run for a maximum of 100 iterations. The convergence criterion was set to a stagnation of the best solution for 30 consecutive iterations. Concretely, we perform an evaluation of the cost function Λ based on the parameters μ_b and μ_w.

Due to the complexity and significant non-linearity of optimization problems, analytical solutions are often challenging to obtain. However, by utilizing suitable nonlinear optimization techniques, we can derive optimal solutions in the cost model. In this instance, we define the parameters and apply the grey wolf optimizer algorithm to obtain the optimal values (u_b^*,u_w^*) for the service rates. We write the problem designed to optimize:

$$\min_{\mu_b,\mu_w} \Lambda(\mu_b,\mu_w)$$
$$\text{s.t} \begin{cases} \mu_b - \mu_w > 0, \\ \mu_w > 0, \\ (\mu_b,\mu_w) \in \mathbb{R}_+^2. \end{cases}$$

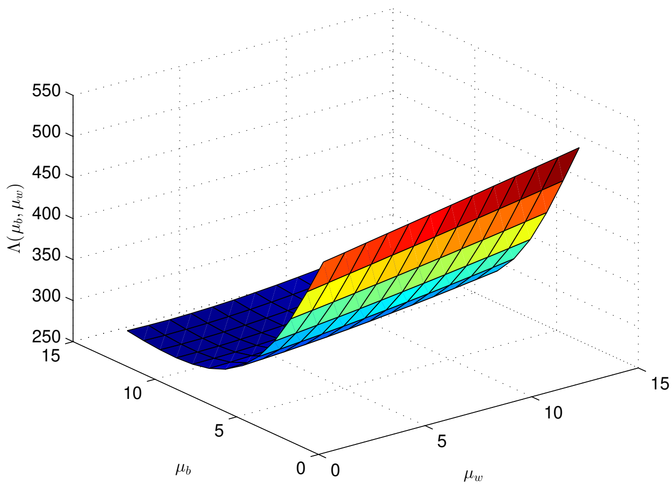


Figure 10. $\Lambda(\mu_b,\mu_w)$ vs. μ_b and μ_w

Figure 10 clearly illustrates the convex nature of the objective function Λ according the service rates μ_b and μ_w.

To proceed with analyzing the cost optimization of the queueing model, we firstly set the parameters regulating the cost: $C_b = 80$ USD/hour, $C_v = 60$ USD/hour, $C_{wv} = 70$ USD/hour, $C_q = 55$ USD/hour, $C_r = 30$ USD/hour, $C_t = 15$ USD/hour, $C_{\mu_b} = 3$ USD/customer, $C_{\mu_w} = 2$ USD/customer, $C_f = 7$ USD/customer, $C_a = 3$ USD/hour.

Table 3. Optimal values of (μ_b^*,μ_w^*) and $\Lambda^*(\mu_b^*,\mu_w^*)$ for various values of arrival rate (λ) and vacation rate (η), when $\lambda = 6:1:8$, $K = 12$, $\alpha = 0.7$, $\theta = 0.6$, $\gamma = 2$, $\eta = [2; 2.5; 3]$, $\xi_0 = 0.9$, $\xi_1 = 1.6$, $\xi_2 = 1.9$.

η	λ	μ_b^* (cust/hr)	μ_w^* (cust/hr)	Λ^* (USD/hr)
2	6	10.5074	2.1155	274.9516
	7	11.8579	3.5971	304.1049
	8	13.1917	5.1614	332.3367
2.5	6	10.7098	2.5624	273.0545
	7	12.0748	4.1477	302.0183
	8	13.4340	5.7542	330.1020
3	6	10.8870	2.9701	271.6902
	7	12.2621	4.6054	300.4876
	8	13.6206	6.1863	328.4362

– From Table 3, the minimum expected cost is seen to increase when λ increases. Nevertheless, when the vacation rate is on the rise the minimum expected cost is dropping. This confirms that reducing the vacation rate is a costly endeavor.

Table 4. Optimal values of (μ_b^*, μ_w^*) and $\Lambda^*(\mu_b^*, \mu_w^*)$ for various value of non-feedback probability (θ), when $\lambda = 6.5$, $K = 12$, $\alpha = 0.7$, $\theta = [0.4; 0.6; 0.8]$, $\gamma = 2$, $\eta = 3$, $\xi_0 = 0.9$, $\xi_1 = 1.6$, $\xi_2 = 1.9$.

θ	μ_b^* (cust/hr)	μ_w^* (cust/hr)	Λ^* (USD/hr)
0.4	13.7709	2.1155	344.1649
0.6	11.5769	3.7784	286.2084
0.8	10.7260	7.7133	238.7107

Table 5. Optimal values of (μ_b^*, μ_w^*) and $\Lambda^*(\mu_b^*, \mu_w^*)$ for various value of working vacation rate (γ), when $\lambda = 6.5$, $K = 12$, $\alpha = 0.7$, $\theta = 0.6$, $\gamma = [1.5; 2; 2.5]$, $\eta = 3$, $\xi_0 = 0.9$, $\xi_1 = 1.6$, $\xi_2 = 1.9$.

γ	μ_b^* (cust/hr)	μ_w^* (cust/hr)	Λ^* (USD/hr)
1.5	11.3372	7.0923	292.0922
2	11.5757	3.7884	286.2084
2.5	11.8271	2.1155	278.8619

Table 6. Optimal values of (μ_b^*, μ_w^*) and $\Lambda^*(\mu_b^*, \mu_w^*)$ for various value of non-retention probability (α), when $\lambda = 6.5$, $K = 12$, $\alpha = [0.3; 0.5; 0.7]$, $\theta = 0.6$, $\gamma = 2$, $\eta = 3$, $\xi_0 = 0.9$, $\xi_1 = 1.6$, $\xi_2 = 1.9$.

α	μ_b^* (cust/hr)	μ_w^* (cust/hr)	Λ^* (USD/hr)
0.3	13.1045	5.5258	297.005
0.5	12.3133	4.6534	291.4246
0.7	11.5754	3.7850	286.2048

- From Tables 4-6, we observe that with the leap of θ, there is a diminution in the minimum expected cost, and it can also be seen that a drop of the optimal anticipated cost with the hike of γ and α. This means that reducing the working vacation time, feedback probability, and retention probability results in additional cost savings.

7. Conclusion

Our modest paper examines a queue of $M/M/1/K$ model including Bernoulli feedback under a hybrid vacation policy scenario with impatient customers. Employing a recursive method, steady-state probabilities were derived, and metrics were formulated to assess the system's performance. Additionally, the results obtained from Multilayer Perceptrons (MLP) are compared with the numerical results. Furthermore, numerical solutions were achieved through the implementation of the Grey Wolf Optimizer to ensure optimizing the rates of the services and minimizing the function that expresses the expected cost. Finally, experimental computation results were used to emphasize the effects of several parameters on (μ_b^*, μ_w^*) and $\Lambda(\mu_b^*, \mu_w^*)$.

This research can be extended in various directions by incorporating concepts such as retrial queues, priority mechanisms, catastrophic events, and heterogeneous customer types into the queuing model in future studies. Furthermore, a limitation of the current model is the assumption of linear balking and constant reneging probabilities. In reality, customer impatience is often nonlinear and time-dependent. A significant direction for future research would be to incorporate more dynamic customer behavior models, such as using exponential growth functions for balking or for reneging rates that increase with waiting time. This would enhance the model's realism and predictive power, though it would likely necessitate a simulation-based approach rather than an analytical one.

It is also important to acknowledge the assumptions underlying our model, such as the Poisson arrival process and exponentially distributed service and vacation times. While these Markovian assumptions are necessary for analytical tractability, they may not perfectly capture the dynamics of all real-world systems. Future research could extend this work by considering more general distributions (e.g., phase-type distributions) or by using simulation to analyze non-Markovian models, thereby broadening its applicability. Finally, a crucial next step for future research will be to validate the entire framework using diverse, real-world datasets. Applying these methods to empirical data from different industries would be essential to confirm the model's generalizability and practical utility.

Acknowledgement

The authors are thankful to the Editor in Chief as well as the anonymous referees for their valuable comments and suggestions, which substantially improved the quality of this paper.

REFERENCES

1. I. Adan, A. Economou, and S. Kapodistria, *Synchronized reneging in queueing systems with vacations*, Queueing Systems, vol. 62, pp. 1–33, 2009.
2. S. I. Ammar, *Transient solution of an $M/M/1$ vacation queue with a waiting server and impatient customers*, Journal of the Egyptian Mathematical Society, vol. 25, no. 3, pp. 337–342, 2017.
3. F. Afroun, D. Aïssani, D. Hamadouche, and M. Boualem, *Q-matrix method for the analysis and performance evaluation of unreliable $M/M/1/N$ queueing model*, Mathematical Methods in the Applied Sciences, vol. 41, no. 18, pp. 9152–9163, 2018.
4. E. Altman and U. Yechiali, *Analysis of customers' impatience in queues with server vacations*, Queueing Systems, vol. 52, pp. 261–279, 2006.
5. S. Bagheri, S. Taridashti, H. Farahani, P. Watson, and E. Rezvani, *Multilayer perceptron modeling for social dysfunction prediction based on general health factors in an Iranian women sample*, Frontiers in Psychiatry, vol. 14, article 1283095, 2023.
6. A. A. Bouchentouf, A. Guendouzi, and A. Kandouci, *Performance and economic study of heterogeneous $M/M/2/N$ feedback queue with working vacation and impatient customers*, ProbStat Forum, vol. 12, pp. 15–35, 2019.
7. A. A. Bouchentouf, M. Boualem, L. Yahiaoui, and H. Ahmad, *A multi-station unreliable machine model with working vacation policy and customers impatience*, Quality Technology and Quantitative Management, vol. 19, no. 6, pp. 766–796, 2022.
8. A. A. Bouchentouf, M. Cherfaoui, and M. Boualem, *Performance and economic analysis of a single server feedback queueing model with vacation and impatient customers*, Opsearch, vol. 56, no. 1, pp. 300–323, 2019.
9. A. A. Bouchentouf, A. Guendouzi, and S. Majid, *On impatience in Markovian $M/M/1/N/DWV$ queue with vacation interruption*, Croatian Operational Research Review, vol. 11, pp. 21–37, 2020.
10. A. A. Bouchentouf, M. Cherfaoui, and M. Boualem, *Analysis and performance evaluation of Markovian feedback multi-server queueing model with vacation and impatience*, American Journal of Mathematical and Management Sciences, vol. 40, no. 3, pp. 261–282, 2021.
11. A. A. Bouchentouf, A. Guendouzi, M. Hoalef, and S. Majid, *Analysis of a single server queue in a multi-phase random environment with working vacations and customers' impatience*, Operations Research and Decisions, vol. 32, no. 2, pp. 16–33, 2022.
12. A. A. Bouchentouf, M. Boualem, L. Yahiaoui, and H. Ahmad, *A Multi-Station Unreliable Machine Model With Working Vacation Policy and Customers' Impatience*, Quality Technology and Quantitative Management, vol. 19, no. 6, pp. 766–796, 2022.
13. A. A. Bouchentouf, and M. Boualem, *Modelling and simulation of Bernoulli feedback queue with general customers' impatience under variant vacation policy*, International Journal of Operational Research, vol. 46, no. 4, pp. 451–480, 2023.
14. B. T. Doshi, *Single server queues with vacation: A survey*, Queueing Systems, vol. 1, no. 1, pp. 29–66, 1986.
15. A. Dehimi, M. Boualem, A. A. Bouchentouf, S. Ziani, and L. Berdjoudj, *Analytical and computational aspects of a multi-server queue with impatience under differentiated working vacations policy*, Reliability: Theory and Applications, vol. 19, no. 3 (79), pp. 393–407, 2024.
16. B. T. Doshi, *Single server queues with vacations*, Stochastic Analysis of the Computer and Communication Systems, pp. 217–264, 1990.
17. N. Gupta, *Analysis of an $M/M/1/K$ feedback working vacation queue with reneging*, Reliability: Theory and Applications, vol. 18, no. 4 (76), pp. 178–188, 2023.
18. O. C. Ibe, and O. A. Isijola, *$M/M/1$ multiple vacation queueing systems with differentiated vacations*, Modelling and Simulation in Engineering, pp. 18–18, 2014.
19. M. Jain, and S. Upadhyaya, *Synchronous working vacation policy for finite-buffer multiserver queueing system*, Applied Mathematics and Computation, vol. 217, no. 24, pp. 9916–9932, 2011.
20. A. Kadi, M. Boualem, N. Touche, and A. Dehimi, *Modeling and Optimization of an $M/M/1/K$ queue with single working vacation, feedback, and impatience timers under N-policy*, Discrete and Continuous Models in Applied Computational Science, vol. 33, no. 1, pp. 5–22, 2025.
21. A. Kadi, N. Touche, A. A. Bouchentouf, and M. Boualem, *Finite-Capacity $M/M/2$ machine repair model with impatient customers, Triadic Discipline, and Two Working Vacation Policies*, Journal of Mathematical Modeling, vol. 13, no. 1, pp. 447–460, 2025.
22. S. Keerthana, J. E. A. Bagyam, and R. Remya, *A Cost Analysis of Single-Server Discouraged Arrivals With Differentiated Vacation Queueing Model*, Croatian Operational Research Review, vol. 15, no. 2, pp. 145–158, 2024.
23. R. Kumar, and S. K. Sharma, *An $M/M/1/N$ queueing model with retention of reneged customers and balking*, American Journal of Operational Research, vol. 2, no. 1, pp. 1–5, 2012.

24. R. Kumar, and S. K. Sharma, *A multi-server Markovian feedback queue with balking, reneging and retention of reneged customers*, AMO-Advanced Modeling and Optimization, vol. 16, no. 2, pp. 395–406, 2014.
25. R. Kumar, and S. Sharma, *Transient analysis of an $M/M/c$ queueing system with balking and retention of reneging customers*, Communications in Statistics-Theory and Methods, vol. 47, no. 6, pp. 1318–1327, 2018.
26. P. V. Laxmi, V. Goswami, and K. Jyothsna, *Analysis of finite buffer Markovian queue with balking, reneging and working vacations*, International Journal of Strategic Decision Sciences, vol. 4, no. 1, pp. 1–24, 2013.
27. J. Li, and N. Tian, *The $M/M/1$ queue with working vacations and vacation interruptions*, Journal of Systems Science and Systems Engineering, vol. 16, no. 1, pp. 121–127, 2007.
28. J. H. Li, N. S. Tian, and Z. Y. Ma, *Performance analysis of $GI/M/1$ queue with working vacations and vacation interruption*, Applied Mathematical Modelling, vol. 32, no. 12, pp. 2715–2730, 2008.
29. S. Majid, and P. Manoharan, *Impatient customers in an $M/M/c$ queue with single and multiple synchronous working vacations*, Pakistan Journal of Statistics and Operation Research, vol. 14, no. 3, pp. 571–594, 2018.
30. S. Majid, *Performance Analysis of a Markovian Queue with Impatient Customers and Working Vacation*, Journal of the Operations Research Society of China, vol. 11, no. 1, pp. 133–156, 2023.
31. S. Mirjalili, S. M. Mirjalili, and A. Lewis, *Grey wolf optimizer*, Advances in Engineering Software, vol. 69, pp. 46–61, 2014.
32. V. Narmadha and P. Rajendran, *Analysis of an encouraged arrival Markovian queue with single working vacation, impatience, and reneging of customers*, Reliability: Theory and Applications, vol. 20, no. 1 (82), pp. 897–902, 2025.
33. K. A. Rashedi, M. T. Ismail, S. Al Wadi, A. Serroukh, T. S. Alshammari, and J. J. Jaber, *Multi-layer perceptron-based classification with application to outlier detection in Saudi Arabia stock returns*, Journal of Risk and Financial Management, vol. 17, no. 2, article 69, 2024.
34. F. Rosenblatt, *The perceptron: a probabilistic model for information storage and organization in the brain*, Psychological Review, vol. 65, no. 6, pp. 386–408, 1958.
35. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature, vol. 323, no. 6088, pp. 533–536, 1986
36. C. Shekhar, S. Varshney, and A. Kumar, *Matrix-geometric solution of multiserver queueing systems with Bernoulli scheduled modified vacation and retention of reneged customers: A meta-heuristic approach*, Quality Technology and Quantitative Management, vol. 18, no. 1, pp. 39–66, 2021.
37. L. D. Servi, and S. G. Finn, *$M/M/1$ queues with working vacations ($M/M/1/WV$)*, Performance Evaluation, vol. 50, no. 1, pp. 41–52, 2002.
38. N. Selvaraju, and C. Goswami, *Impatient customers in an $M/M/1$ queue with single and multiple working vacations*, Computers and Industrial Engineering, vol. 65, no. 2, pp. 207–215, 2013.
39. N. Tian, and Z. G. Zhang, *Vacation queueing models: theory and applications*, Springer Science and Business Media, 2006.
40. D. A. Wu, and H. Takagi, *$M/G/1$ queue with multiple working vacations*, Performance Evaluation, vol. 63, no. 7, pp. 654–681, 2006.
41. E. Yüksel, D. Soydaner, and H. Bahtiyar, *Nuclear binding energy predictions using neural networks: Application of the multilayer perceptron*, International Journal of Modern Physics E, vol. 30, no. 03, article 2150017, 2021.
42. D. Yue, W. Yue, and G. Xu, *Analysis of customers impatience in an $M/M/1$ queue with working vacations*, Journal of Industrial and Management Optimization, vol. 8, no. 4, pp. 895–908, 2012.