Chapter 5

Vacation Queueing Model with F-Policy and Vacation Interruption

5.1 Introduction

Any framework in which the accessible resources are not sufficient at the point of time to fulfill the demands placed randomly upon them is a contender for queueing investigation. Queues occur broadly in all spheres of life and many useful technological systems. They earned sovereignty in machine shops with a requirement for a quick repair, in banks with customers being served by raconteur, calls appearing at telephone exchanges and many more.

In many day-to-day real-life scenarios, the controllable service models are designed to discover the optimal conditions for turning the server on/off, which results in the lowest expected cost. Controllable queue-based service systems could be divided into two parts; first is arrival control policy, which includes F-policy (*cf.* [92]) whereas the second is service control policy which includes the vacation policy like the N-policy, the D-policy and the T-policy, which are proposed by [297], [25], and [99] respectively. Gupta [92] first established the interrelationship between F-policy and N-policy to give significant insight into the queueing problem. Tadj and Choudhury [246] gave a comprehensive review to establish the optimal control and optimal designs of waiting-line models. A recursive approach for the F-policy finite capacity $G/M/1$ queue was established by [278] using the supplementary variable technique with the assumption of the remaining inter-arrival time as a supplementary variable. Chang et al. [39] studied the controllable F-policy $M/H_2/1$ queueing model with two types of service policies and obtained the steady-state analytic solutions besides the optimal capacity of the system K and threshold value F simultaneously to minimize the expected total cost. More recently, [108] dealt with the optimal threshold admission control policies in the machining system besides the concept of working vacation for the optimum maintainability cost. Shekhar et al. [227] discussed maintainability issues and standby workstations provisioning for a manufacturing system and computed the transient-state probabilities via the spectral method.

In most of the investigations on queue-based service systems, customers always wait in the service system until his job is completed, which is the common assumption. But in many practical waiting line scenarios, such as network communication, emergency rooms in a hospital, and a maintenance department, the customers may become impatient and depart the service system without receiving the service in search of alternative service when the expected waiting time is extremely long. Thus, the impatience behavior of the customer with which he may not join the queue on arrival is known as balking. While in the reneging behavior, customer joins the queue but leaves impatiently without being served after waiting for some time into the service system. Haight [95] first introduced the balking behavior of customers for a single server Markovian queueing model. Later, [96] again investigated a $M/M/1$ queueing model in which the customers leave (renege) the service system without getting service. Queueing problems with balking or/and reneging behavior of the customers have motivated many researchers to investigate (*cf.* [13], [14], [75], [119], [216], [1]). Liou [179] employed the matrix-analytic approach to computing the steady-state probabilities and stability condition for an infinite capacity Markovian queue with a single unreliable service station and impatient customer. Recently, [228] extended the reneging behavior for the point process in the finite population queueing model and provided a profitable insight to the industrial engineers through the sensitivity analysis for various performance indices.

In this chapter, we also consider real-time system-oriented service attitude of the server that it continues to work with the slower service rate rather than complete termination of his service during the vacation period. Such vacation policy has been described as a working vacation, which had been first introduced by [222]. Liu et al. [185] determined the stochastic decomposition structures of the expected queue length and waiting time in the system for a single server Markovian queueing model with working vacation. Lin and Ke [177] studied an infinite capacity multi-server queue with single working vacation and employed the matrix-geometric approach to derive the explicit closed-form of the rate matrix of the quasi-birth and death process. Using the supplementary variable technique, [318] depicted the queue length distribution and service condition at the arbitrary point for the $M/G/1$ queueing model with single working vacation. The N-policy $M/M/1$ queue with working vacation and server breakdowns was investigated by [302]. By obtaining the stability condition and stationary probability distribution of the number of customers in the system, they proposed the closed matrix form expressions of various performance characteristics. Recently, [312] considered a $M/M/1$ queueing model with single working vacation and multiple vacations and by using R-matrix method, G-matrix method they obtained the stationary distribution of queue length. Shekhar et al. [232] surveyed the recent literature in detail for modeling and application of a finite population queue-based system with different types of impatience behavior among customers and various type of vacation policies adopted by service provider.

The assumption of being in the working vacation for the server until the end of the vacation period is unrealistic and restrictive. The server resumes the normal working attribute at the end of the random vacation period only. To overcome this limitation, [170] proposed the vacation interruption policy in a single server Markovian queue with working vacations. According to this policy, the server can join back from the vacation to the normal working level once some performance characteristics, such as the number of customers, achieve a certain value in the vacation period. The server may come back from the vacation without completing the vacation. The model proposed by [170] was generalized to the $M/G/1$ queueing model by [320], to $M/PH/1$ queueing model by [23] and to $MAP/G/1$ model again by [321]. Gao and Liu [78] assumed a single working vacation and Bernoulli vacation interruption policy for the $M/G/1$ queueing model. Recently, discrete-time queues with working vacations and vacation interruptions were investigated by [81], [80], [79].

Metaheuristic algorithms are emergent nature-inspired optimization techniques which are employed when traditional techniques fail for optimal analysis of the real-time complex service system. These algorithms may be population-based (genetic algorithm) or agent-based (particle swarm optimization (PSO), cuckoo search (CS), firefly algorithm (FA) (*cf.* [178], [36], [309])). The particle swarm optimization technique was first conceptualized by [67] in 1995 based on swarm behavior of bird and fish schooling in nature. From that point of view, PSO has been used in almost every area in optimization, computational intelligence, communication systems, and design applications. There are more than two dozen PSO variants and hybrid algorithms by combining PSO with other existing algorithms.

The prime objective of the present study is to investigate the controllable arrival policy of service systems with realistic design assumptions. This chapter establishes and narrows the following research gaps simultaneously (i) the finite capacity Markovian queue-based service system with controllable arrival policy, *i.e.* F-policy; (ii) the single working vacation with vacation interruption; (iii) impatience behavior of the customers. The purpose of the present chapter is threefold (i) to establish mathematical model of finite capacity Markovian queue with F-policy, single working vacation, vacation interruption and impatience behavior of the customers; (ii) to propose matrix analytical method for computing the steady-state probabilities and various service system performance characteristics; (3) to employ an efficient nature-inspired optimization technique, particle swarm optimization for optimal analysis of expected total cost; and classical differential calculus for the sensitivity analysis.

The rest content of this chapter is organized in different sections as follows. In section 5.2, the assumptions and notations required for mathematical modeling of the service system are described along with the steady-state solution procedure through the matrix analytic method. In section 5.3, various service system performance characteristics for analysis purposes are established. The expected total cost function is proposed for the governing model in the next section 5.4 and employs the particle swarm optimization technique (PSO) for optimal analysis in section 5.5. In section 5.6, the sensitivity analysis for expected total cost function (TC) in terms of various parameters involved in service system is performed. The results of numerical simulation, sensitivity, and optimal analysis are depicted in tables and graphs in section 5.7 to give a quick insight into the studied queue-based service system. Finally, in section 5.8, some conclusions are drawn by highlighting the noble inference and future scopes are provided.

5.2 Model Description

In the present chapter, the controllable arrival policy, namely F-policy and working vacation policy in finite capacity service system are considered with vacation interruption. The threshold parameter F is used to establish a controllable arrival in the queueing model, which becomes more flexible and realistic dealing with congestion conditions experienced in many areas such as computer & communication systems, transportation systems, and production systems, etc. From the literature survey, it is also perceived that the service models with F-policy and vacation interruption have not been studied previously. The incorporation of these sensible behaviors makes our service model more adaptable and closer to realistic queueing situations. Following notations and assumptions are considered for the present work.

- The customers arrive according to Poisson fashion in the service system with a mean arrival rate of λ. The arrived customers form a single waiting queue.
- When the number of customers in the service system reaches its finite capacity of K, it forbids any customer to join the system until the length of the customers in the service system decreases to a pre-specified threshold value F.
- As per controllable arrival policy, F-policy, at the allowable epoch, *i.e.* epoch at which the customers are allowed to join the service system for service, the

system requires a startup time which is exponentially distributed with parameter γ.

- After completion of all his services, when the service system becomes empty, $i.e.$ the server finds no waiting customer queueing up in the system; the server adopts a working vacation mode of the random duration. The duration of working vacation is exponentially distributed with mean time $\frac{1}{\theta}$.
- On resuming from vacation to normal busy period, if the server finds no customer waiting in the service system, the server is not permitted to adopt another vacation, and he waits idly for the newly arriving customer whom he immediately starts to provide the service, $i.e.$ the server follows single working vacation policy.
- During working vacation, the server continues to facilitate the service to the waiting customers with the slower service rate, and the prospective customer continues to join the waiting line of the service system.
- A single server follows $First\,Come\,First\,Serve\,(FCFS)$ service discipline. The service times during working vacation mode and normal working attribute follow an independently and identically distributed exponential distribution with meantime $\frac{1}{\mu_v}$ and $\frac{1}{\mu_b}$ $(\mu_v < \mu_b)$ respectively.
- If the server finds at least F waiting customers in the service system after completion of service for any customer in the working vacation attribute then the server's vacation is interrupted immediately, and the server resumes the normal working attribute.
- If an arriving customer finds waiting customers queueing up into the service system, then either he decides to balk with probability ξ or remains into the service system with the complementary probability $\bar{\xi}$.
- When the server is busy, due to unpredictable long waiting times, the customer exhibits an impatience and renege after the time which follows an exponential distribution with parameter η.
- All processes and events are repeated all over again and independent to the state of the other.

Fig. 5.1 depicts the transition states of the governing finite capacity Markovian service model with working vacation, vacation interruption and randomized arrival control policy.

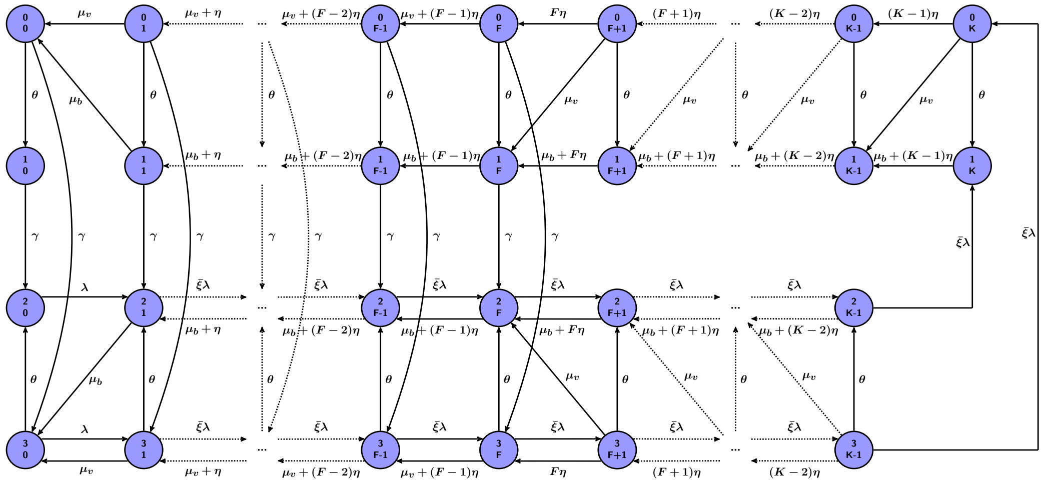

Figure 5.1: State transition diagram

5.2.1 Notations

In the present study, for the brevity of mathematical modeling, the states of the governing model at any instant t are defined as

$$J(t) \equiv \text{The state of the server at instant } t.$$
$$N(t) \equiv \text{Number of the customers in the service system at instant } t.$$

where

$$J(t) \equiv \begin{cases} 0; & \text{Customers are not allowed to join the service system} \\ & \text{and server is in working vacation attribute.} \\ 1; & \text{Customers are not allowed to join the service system} \\ & \text{and server is in normal busy mode.} \\ 2; & \text{Customers are allowed to join the service system} \\ & \text{and server is in normal busy mode.} \\ 3; & \text{Customers are allowed to join the service system} \\ & \text{and server is in working vacation attribute.} \end{cases}$$

Then, $\{J(t), N(t); t \geq 0\}$ represents a two dimensional continuous-time Markov chain (CTMC) with the state space

$$\Theta \equiv \{\{(j,n); j = 0,1 \;\&\; n = 0,1,2,\ldots,K\} \cup \{(j,n); j = 2,3 \;\&\; n = 0,1,2,\ldots,K-1\}\}$$

Now, let us define the state probabilities at instant t as

$$P_{j,n}(t) = \text{Prob}[J(t) = j \,\&\, N(t) = n];\; j = 0,1;\; n = 0,1,\ldots,K$$
$$P_{j,n}(t) = \text{Prob}[J(t) = j \,\&\, N(t) = n];\; j = 2,3;\; n = 0,1,\ldots,K-1$$

For steady-state analysis, on equilibrium condition (as $t \to \infty$), the steady-state probabilities are defined as limiting form as follows

$$P_{j,n} = \lim_{t\to\infty} P_{j,n}(t);\; \forall j = 0,1;\; n = 0,1,2,\ldots,K$$
$$P_{j,n} = \lim_{t\to\infty} P_{j,n}(t);\; \forall j = 2,3;\; n = 0,1,2,\ldots,K-1.$$

5.2.2 Matrix Analytic Solutions

The matrix-analytic method is a very convenient approach for understanding and formulating the stochastic models. It is widely used to deal with numerous waiting line problems that desire the exact analysis. Neuts [200] first time proposed the matrix-analytic method to investigate the embedded Markov chain of many realistic waiting line problems. In the present research, the corresponding block tridiagonal rate matrix $\mathbf{Q}$ of this Markov chain having dimension $(4K+2)\times(4K+2)$ has the structure

$$\mathbf{Q} = \begin{bmatrix} \mathbf{A}_0 & \mathbf{B}_0 & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{C}_1 & \mathbf{A}_1 & \mathbf{B}_1 & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \mathbf{A}_2 & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_{K-2} & \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{C}_{K-1} & \mathbf{A}_{K-1} & \mathbf{B}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{C}_K & \mathbf{A}_K \end{bmatrix}$$

where, the block matrices $\mathbf{A}_K$, $\mathbf{B}_2$ and $\mathbf{C}_K$ on the lower level of the rate matrix $\mathbf{Q}$ have dimensions (2×2), (4×2) and (2×4) respectively. Other block matrices are square matrices of dimension (4×4). Each block entry of the rate matrix $\mathbf{Q}$ is indexed as follows

$$\mathbf{A}_0 = \begin{bmatrix} -(\gamma+\theta) & \theta & 0 & \gamma \\ 0 & -\gamma & \gamma & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & \theta & -(\lambda+\theta) \end{bmatrix}$$

and, for $1 \leq i \leq F$

$$\mathbf{A}_i = \begin{bmatrix} x_i & \theta & 0 & \gamma \\ 0 & y_i & \gamma & 0 \\ 0 & 0 & z_i & 0 \\ 0 & 0 & \theta & w_i \end{bmatrix}$$

where,

$$\begin{aligned} x_i &= -[\mu_v + (i-1)\eta + \gamma + \theta]; \\ y_i &= -[\mu_b + (i-1)\eta + \gamma] \\ z_i &= -\left[\bar{\xi}\lambda + \mu_b + (i-1)\eta\right]; \\ w_i &= -\left[\bar{\xi}\lambda + \mu_v + (i-1)\eta + \theta\right]; \end{aligned}$$

Similarly, for $F+1 \leq i \leq K-1$

$$\mathbf{A}_i = \begin{bmatrix} x_i & \theta & 0 & 0 \\ 0 & y_i & 0 & 0 \\ 0 & 0 & z_i & 0 \\ 0 & 0 & \theta & w_i \end{bmatrix}$$

where,

$$\begin{aligned} x_i &= -[\mu_v + (i-1)\eta + \theta]; \\ y_i &= -[\mu_b + (i-1)\eta]; \\ z_i &= -\left[\bar{\xi}\lambda + \mu_b + (i-1)\eta\right]; \\ w_i &= -\left[\bar{\xi}\lambda + \mu_v + (i-1)\eta + \theta\right]; \end{aligned}$$

and

$$\mathbf{A}_K = \begin{bmatrix} -[\mu_v + \theta + (K-1)\eta] & \theta \\ 0 & -[\mu_b + (K-1)\eta] \end{bmatrix}$$

$$\mathbf{B}_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix}; \quad \mathbf{B}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \xi\lambda & 0 \\ 0 & 0 & 0 & \bar{\xi}\lambda \end{bmatrix}; \mathbf{B}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \bar{\xi}\lambda \\ \bar{\xi}\lambda & 0 \end{bmatrix}$$

and

$$\mathbf{C}_1 = \begin{bmatrix} \mu_v & 0 & 0 & 0 \\ 0 & \mu_b & 0 & 0 \\ 0 & 0 & 0 & \mu_b \\ 0 & 0 & 0 & \mu_v \end{bmatrix}$$

for $2 \leq i \leq F$

$$\mathbf{C}_i = \begin{bmatrix} [\mu_v + (i-1)\eta] & 0 & 0 & 0 \\ 0 & [\mu_b + (i-1)\eta] & 0 & 0 \\ 0 & 0 & [\mu_b + (i-1)\eta] & 0 \\ 0 & 0 & [\mu_v + (i-1)\eta] & 0 \end{bmatrix}$$

Similarly, for $F+1 \leq i \leq K-1$

$$\mathbf{C}_i = \begin{bmatrix} (i-1)\eta & \mu_v & 0 & 0 \\ 0 & [\mu_b + (i-1)\eta] & 0 & 0 \\ 0 & 0 & [\mu_b + (i-1)\eta] & 0 \\ 0 & 0 & \mu_v & (i-1)\eta \end{bmatrix}$$

and

$$\mathbf{C}_K = \begin{bmatrix} (K-1)\eta & \mu_v & 0 & 0 \\ 0 & [\mu_b + (K-1)\eta] & 0 & 0 \end{bmatrix}$$

5.2.3 Steady-State Probabilities

Let Π represent the steady-state probability vector corresponding to rate matrix $\mathbf{Q}$ which is partitioned as $\Pi = [\Pi_0, \Pi_1, \ldots, \Pi_{K-1}, \Pi_K]$, where the sub-vectors $\Pi_n = [P_{0,n}, P_{1,n}, P_{2,n}, P_{3,n}]$ for $n = 0, 1, 2, ..., K-1$ and $\Pi_K = [P_{0,K}, P_{1,K}]$ are of dimensions (1×4) and (1×2) respectively. Let $\mathbf{e}$ represents a column vector of dimension $(4K+2)$ having all elements 1. Now, by solving the homogeneous system of linear equations $\Pi\mathbf{Q} = 0$ along with the normalizing condition $\Pi e = 1$, we obtain steady-state probability equations in matrix form as

$$\Pi_0 \mathbf{A}_0 + \Pi_1 \mathbf{C}_1 = \mathbf{0} \tag{5.1}$$

$$\Pi_0 \mathbf{B}_0 + \Pi_1 \mathbf{A}_1 + \Pi_2 \mathbf{C}_2 = \mathbf{0} \tag{5.2}$$

$$\Pi_{n-1} \mathbf{B}_1 + \Pi_n \mathbf{A}_n + \Pi_{n+1} \mathbf{C}_{n+1} = \mathbf{0};\ 2 \leq n \leq K-1 \tag{5.3}$$

$$\Pi_{K-1} \mathbf{B}_2 + \Pi_K \mathbf{A}_K = \mathbf{0} \tag{5.4}$$

Now, after recursively substitution,

$$\Pi_0 = \Pi_1 \mathbf{C}_1 \left(-\mathbf{A}_0^{-1}\right) = \Pi_1 \mathbf{X}_0 \tag{5.5}$$

$$\Pi_1 = \Pi_2 \mathbf{C}_2 \left[-\{\mathbf{X}_0 \mathbf{B}_0 + \mathbf{A}_1\}^{-1}\right] = \Pi_2 \mathbf{X}_1 \tag{5.6}$$

$$\Pi_n = \Pi_{n+1} \mathbf{C}_{n+1} \left[-\{\mathbf{X}_{n-1} \mathbf{B}_1 + \mathbf{A}_n\}\right]^{-1}$$

$$\Rightarrow \quad \Pi_n = \Pi_{n+1} \mathbf{X}_n;\ n = 2,3,\ldots,K-1 \tag{5.7}$$

and again by suitable substitution, we get

$$\Pi_n = \Pi_K \{\mathbf{X}_{K-1} \mathbf{X}_{K-2} \mathbf{X}_{K-3} \ldots \mathbf{X}_{n+2} \mathbf{X}_{n+1} \mathbf{X}_n\};\ n = 0,1,2,\ldots,K-1$$

$$\Rightarrow \Pi_n = \Pi_K \prod_{j=n}^{K-1} \mathbf{X}_j = \Pi_K \Phi_n;\ n = 0,1,2,\ldots,K-1 \tag{5.8}$$

Let define $\mathbf{e}_1 = [1,1,1,1]^T$ and $\mathbf{e}_2 = [1,1]^T$. Now, by using the normalization condition $\Pi \mathbf{e} = 1$, we obtain

$$\sum_{n=0}^{K-1} \Pi_n \mathbf{e}_1 + \Pi_K \mathbf{e}_2 = 1$$

$$[\Pi_0 + \Pi_1 + \Pi_2 + \cdots + \Pi_{K-1}]\, \mathbf{e}_1 + \Pi_K \mathbf{e}_2 = 1$$

$$\Pi_K [\Phi_0 + \Phi_1 + \cdots + \Phi_{K-1}]\, \mathbf{e}_1 + \Pi_K \mathbf{e}_2 = 1$$

$$\Rightarrow \Pi_K \left[\sum_{n=0}^{K-1} \Phi_n \mathbf{e}_1 + \mathbf{e}_2\right] = 1 \tag{5.9}$$

Therefore, the probability vector Π_K can be obtained explicitly by solving the above eq^n(5.4) and eq^n(5.9). After obtaining the probability vector Π_K, all the other steady-state probability vectors are easily determined from eq^n(5.8).

5.3 System Performance Measures

For the comparative study, the performance modeling of the service system is mandatory and important. There are certain universal performance indices with which the performance of the designed finite service system with randomized arrival control policy (F-policy) and vacation interruption can be characterized for the decision purpose. These performance indices are quite interrelated, and each presumes its importance in a particular background. In this section, some service system performance indices by using the steady-state probabilities which are computed in the previous section, are established as follows

- The expected number of customers in the service system

$$L_S = \Pi_K \left[\sum_{n=1}^{K-1} n\Phi_n \mathbf{e}_1 + K\mathbf{e}_2 \right] \tag{5.10}$$

- The probability that the server is in working vacation attribute

$$P_{WV} = \Pi_K \left[\sum_{n=0}^{K-1} \Phi_n \mathbf{u}_1 + \mathbf{u}_2 \right] \tag{5.11}$$

- The probability that the server is on regular busy mode

$$P_B = \Pi_K \left[\sum_{n=0}^{K-1} \Phi_n \mathbf{w}_1 + \mathbf{w}_2 \right] \tag{5.12}$$

- The probability that the server's vacation is interrupted

$$P_{VI} = \Pi_K \left[\sum_{n=F+1}^{K-1} \Phi_n \mathbf{u}_1 + \mathbf{u}_2 \right] \tag{5.13}$$

- Throughput of the service system

$$\tau_p = \Pi_K \left[\sum_{n=1}^{K-1} \mu_v \Phi_n \mathbf{u}_1 + \mu_v \mathbf{u}_2 + \sum_{n=1}^{K-1} \mu_b \Phi_n \mathbf{w}_1 + \mu_b \mathbf{w}_2 \right] \tag{5.14}$$

- The probability that the server starts to allow customers to join the service system

$$P_S = \Pi_K \sum_{n=0}^{F} \Phi_n \mathbf{b}_1 \tag{5.15}$$

- The probability that customers are not allowed to join the service system *i.e.* the system is blocked

$$P_L = \Pi_K \left[\sum_{n=0}^{K-1} \Phi_n \mathbf{b}_1 + \mathbf{e}_2 \right] \tag{5.16}$$

- Average balking rate

$$ABR = \Pi_K \sum_{n=1}^{K-1} \left(\bar{\xi} \lambda \right) \Phi_n \mathbf{a}_1 \tag{5.17}$$

- Average reneging rate

$$ARR = \Pi_K \left[\sum_{n=2}^{K-1} (n-1)\eta\, \Phi_n \mathbf{e}_1 + (K-1)\eta\, \mathbf{e}_2 \right] \tag{5.18}$$

- Effective arrival rate

$$\lambda_{\text{eff}} = \Pi_K \left[\lambda\, \Phi_0\, \mathbf{a}_1 + \sum_{n=1}^{K-1} (\bar{\xi}\lambda)\, \Phi_n\, \mathbf{a}_1 \right] \tag{5.19}$$

- The expected waiting time of the customers in the service system

$$W_S = \frac{L_S}{\lambda_{\text{eff}}} \tag{5.20}$$

where $\mathbf{e}_1$, $\mathbf{u}_1$, $\mathbf{w}_1$, $\mathbf{a}_1$ and $\mathbf{b}_1$ are column vectors of order four having vector forms $[1,1,1,1]^T$, $[1,0,0,1]^T$, $[0,1,1,0]^T$, $[0,0,1,1]^T$ and $[1,1,0,0]^T$ respectively. Similarly $\mathbf{e}_2$, $\mathbf{u}_2$ and $\mathbf{w}_2$ are column vectors of order two having vector forms $[1,1]^T$, $[1,0]^T$ and $[0,1]^T$ respectively.

5.4 Cost Analysis

In this section, a steady-state expected total cost function TC for the finite capacity Markovian service system with working vacation, vacation interruption, and randomized arrival control policy is formulated. In this function, four variables, K, F, μ_v, and μ_b, are assumed as decision variables for the optimal design of the queue-based service system. The discrete variables K and F are required to be a positive integer (Z^+), and μ_v and μ_b are continuous variables having non-negative real value. Our main aim is to analyze the optimal capacity of the service system, say K^*, the optimum threshold value, say F^*, the optimal values of the service rates, say μ_v^* and μ_b^* during a working vacation attribute and regular busy mode respectively.

Having a significant objective in view of the service system design, the expected total cost function is developed using various cost elements associated with numerous performance indices and states of the system. Now, let define the following unit cost elements as

$C_h \equiv$ Cost for each customer present in the service system

$C_b \equiv$ Cost when the server is busy

$C_l \equiv$ Cost associated with each lost customer when the service system is blocked

$C_s \equiv$ Cost associated to each allowed customer to enter the service system

$C_w \equiv$ Cost per customer when customers are waiting for service

$C_k \equiv$ Cost for the capacity of the service system

$C_1 \equiv$ Cost for service when the server is in a regular busy mode

$C_2 \equiv$ Cost for service when the server is in a working vacation attribute

$C_3 \equiv$ Lost cost when one customer balks or reneges

$C_{vi} \equiv$ Cost for initiating vacation interruption

Thus, on the basis of the definitions of cost elements listed above, the expected total cost function is framed as

$$\begin{aligned} TC(K,F,\mu_v,\mu_b) = & C_h L_S + C_b P_B + C_l \lambda P_L + C_s P_S + C_w W_S + C_k K + C_1 \mu_b \\ & + C_2 \mu_v + C_3 (ABR + ARR) + C_{vi} P_{VI} \end{aligned} \tag{5.21}$$

where, K, $F \in Z^+$ and $\mu_v, \mu_b \in R^+ \cup \{0\}$. The cost minimization problem of the designed service model can be illustrated mathematically as a constraint problem

$$\begin{aligned} & TC(K^*,F^*,\mu_v^*,\mu_b^*) = \underset{(K,F,\mu_v,\mu_b)}{\text{minimize}} \; TC(K,F,\mu_v,\mu_b) \\ & \text{subject to} \quad 0 < \mu_v < \mu_b \\ & K,F \in Z^+ \text{ and } \mu_v, \mu_b \in R^+ \cup \{0\} \end{aligned} \tag{5.22}$$

The cost elements listed above are considered to be linear in nature. Due to extremely high non-linearity and complexity of the expected total cost function, it would have been an exhausting work to establish analytic analysis for finding optimal values of the decision variables $(K^*,F^*,\mu_v^*,\mu_b^*)$. Metaheuristic optimizing techniques for analyzing the optimal value of decision variables are a suitable alternative. Among known metaheuristic algorithm, the use of the particle swarm optimization technique (PSO) for global optimization of the expected total cost function is computationally efficient. In support, the numerical illustration is also presented to show that the expected cost function is highly convex in nature and PSO is converging algorithm.

5.5 Particle Swarm Optimization

For important information and details about the PSO algorithm, refer the section 1.10.3 and its pseudo-code therein. For the numerical simulation of the cost optimization problem, here we use the service system design parameters, K, F, μ_v, μ_b

and cost function TC in place of the variables x_1, x_2, x_3, x_4, objective function f in the cited subsection, respectively.

5.6 Sensitivity Analysis

In this section, a sensitivity analysis is performed for the expected total cost function (5.21) for the studied service system with respect to the changes in specific values of various governing parameters λ, μ_v, μ_b, θ, γ, η, and ξ. For the analysis with respect to arrival rate λ of the customers, on differentiating eq.^{n} $\Pi\mathbf{Q} = \mathbf{0}$ with respect to λ, the following partial-differential expression are reducing

$$\frac{\partial \Pi}{\partial \lambda}\mathbf{Q} + \Pi\frac{\partial \mathbf{Q}}{\partial \lambda} = \mathbf{0}$$

or equivalently

$$\frac{\partial \Pi}{\partial \lambda} = -\Pi\frac{\partial \mathbf{Q}}{\partial \lambda}\mathbf{Q}^{-1}$$

Now, the first order partial derivatives $\frac{\partial P_{i,n}}{\partial \lambda}; 0 \leq i \leq 3, 0 \leq n \leq K-1$ and $\frac{\partial P_{j,K}}{\partial \lambda}; j = 0,1$ can easily be obtained. In the same fashion the first order partial derivation with respect to other system parameters $\frac{\partial \Pi}{\partial \mu_v}$, $\frac{\partial \Pi}{\partial \mu_b}$, $\frac{\partial \Pi}{\partial \theta}$, $\frac{\partial \Pi}{\partial \gamma}$, $\frac{\partial \Pi}{\partial \eta}$, and $\frac{\partial \Pi}{\partial \xi}$ can also be determined.

Next, on differentiating the expected total cost function (5.21) with respect to λ, eq^n(5.21) deduce as

$$\frac{\partial TC}{\partial \lambda} = \Bigg[C_h \times \frac{\partial L_S}{\partial \lambda} + C_b \times \frac{\partial P_B}{\partial \lambda} + C_l \times \left\{\lambda\frac{\partial P_L}{\partial \lambda} + P_L\right\} + C_s \times \frac{\partial P_S}{\partial \lambda} + C_w \times \frac{\partial W_S}{\partial \lambda} \\ + C_{vi} \times \frac{\partial P_{vi}}{\partial \lambda} + C_3 \times \left\{\frac{\partial (ABR)}{\partial \lambda} + \frac{\partial (ARR)}{\partial \lambda}\right\}\Bigg]$$

Similarly, differentiating eq^n(5.21) with respect to other service system parameters μ_v, μ_b, θ, γ, η and ξ respectively, we get

$$\frac{\partial TC}{\partial \delta} = \Bigg[C_h \times \frac{\partial L_S}{\partial \delta} + C_b \times \frac{\partial P_B}{\partial \delta} + \lambda C_l \times \frac{\partial P_L}{\partial \delta} + C_s \times \frac{\partial P_S}{\partial \delta} + C_w \times \frac{\partial W_S}{\partial \delta} + C_{vi} \times \frac{\partial P_{vi}}{\partial \delta} \\ + C_3 \times \left\{\frac{\partial (ABR)}{\partial \delta} + \frac{\partial (ARR)}{\partial \delta}\right\} + \Theta\Bigg]$$

where,

$$\delta = \mu_v, \mu_b, \gamma, \theta, \eta, \xi \text{ and } \Theta = \begin{cases} C_1, & \text{if } \delta = \mu_b \\ C_2, & \text{if } \delta = \mu_v \\ 0, & \text{if } \delta = \gamma, \theta, \eta, \xi \end{cases}$$

5.7 Numerical Results

In this section, the numerical illustration is furnished through various graphs and tables to provide the justification for our formulation. All the numerical results are achieved by preparing the mathematical program in MATLAB 2018b (9.5.0.944444), 64bit (Win64), license number 925317 on Intel(R) Core(TM) i5-5200U CPU@2.20 GHz processor, 8MB RAM computing system. The results of the sensitivity and optimal analysis have also been compiled to give a quick insight into the design parameters of the service system.

5.7.1 Numerical Simulation

Because the analytical expressions of performance indices are not enough to interpret the performance of the service system, we compute the values of various performance indices numerically. For that purpose, the default values of the parameters involved in service system are fixed as $F = 7$, $K = 12$, $\lambda = 0.3$, $\mu_v = 0.2$, $\mu_b = 0.5$, $\xi = 0.01$, $\eta = 0.02$, $\gamma = 0.1$, and $\theta = 0.2$, and characterize the results in Figs. 5.2–5.5 and Tables 5.1–5.2.

Figs. 5.2 and 5.3 show the variability in the expected number of customers in the service system (L_S) with respect to varied values of λ and μ_b respectively. From Fig. 5.2, it is clear that the incremental change in λ increases the value of L_S. And in Fig. 5.3, the value of L_S decreases for the increased value of μ_b. Both of these facts are quite obvious. Also, from Figs. 5.2 and 5.3, it is observed the L_S increases with respect to service system capacity (K), the threshold value (F) and parameter γ but decreases with respect to increased values of parameters ξ, η, and θ. For maintaining the high performance of the queue-based service system, system analyst must concentrate on system design K, F & γ. The duration of vacation and decision for vacation interruption also play important role in maintaining the high performance.

Figs. 5.4 and 5.5 describe the trends of the throughput of the service system (τ_p) with incremental change in the value of λ and μ_b respectively. It is observed from Figs. 5.4 and 5.5 that the throughput of the service system increases for the increasing values of λ and μ_b. This is the factual observation that more throughput for

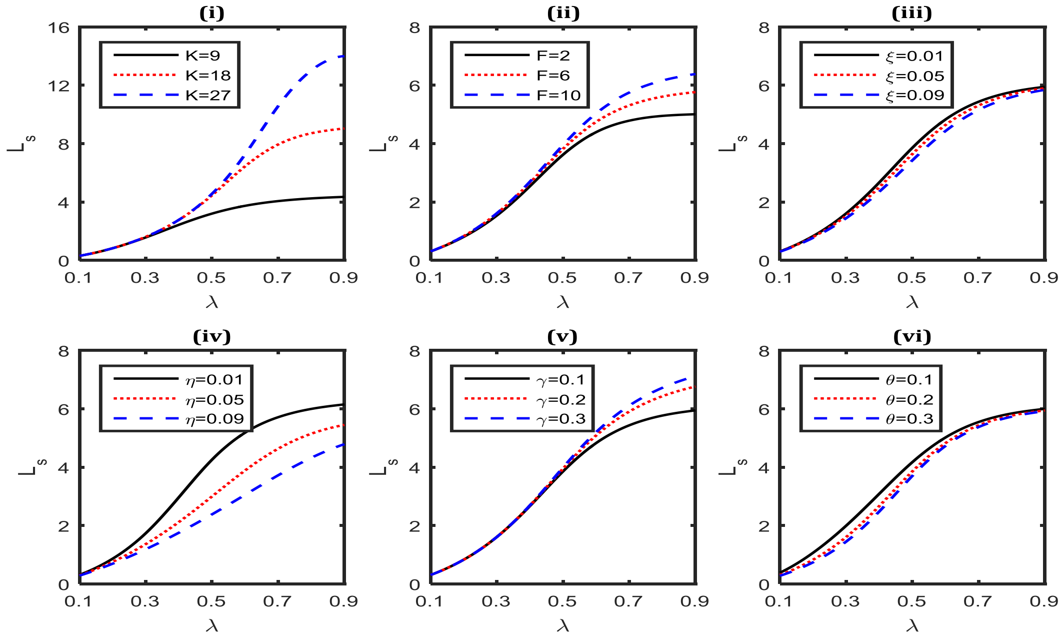

Figure 5.2: Expected number of customer in the service system (L_S) wrt λ for (i) K, (ii) F, (iii) ξ, (iv) η, (v) γ, and (vi) θ.

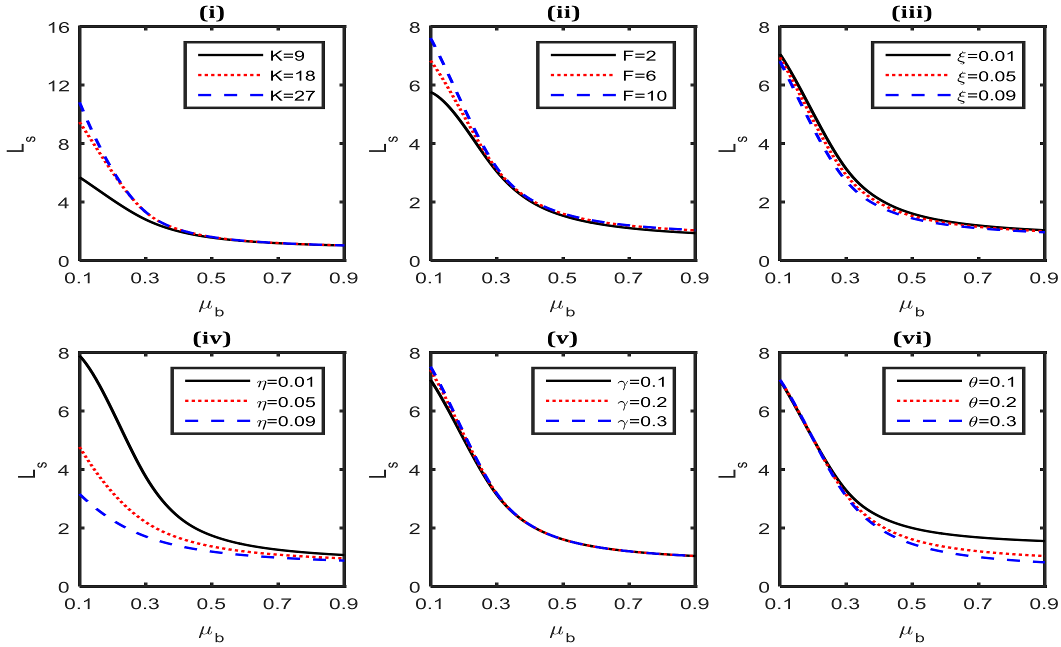

Figure 5.3: Expected number of customer in the service system (L_S) wrt μ_b for (i) K, (ii) F, (iii) ξ, (iv) η, (v) γ, and (vi) θ.

more customers in the service system and becomes saturated after a certain number depending on service system design. For the service rate, this observation is quite obvious. From these figures, it is also observed that the τ_p increases with respect to K, γ, θ but decreases with the incremental change in ξ and η. The results are more

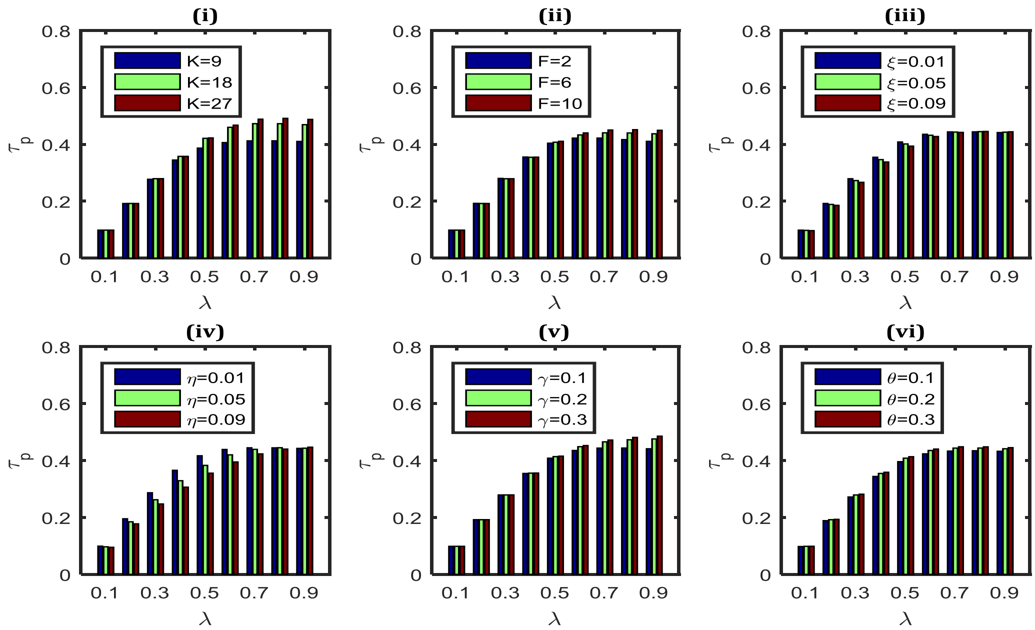

Figure 5.4: Throughput of the service system (τ_p) wrt λ for (i) K, (ii) F, (iii) ξ, (iv) η, (v) γ, and (vi) θ.

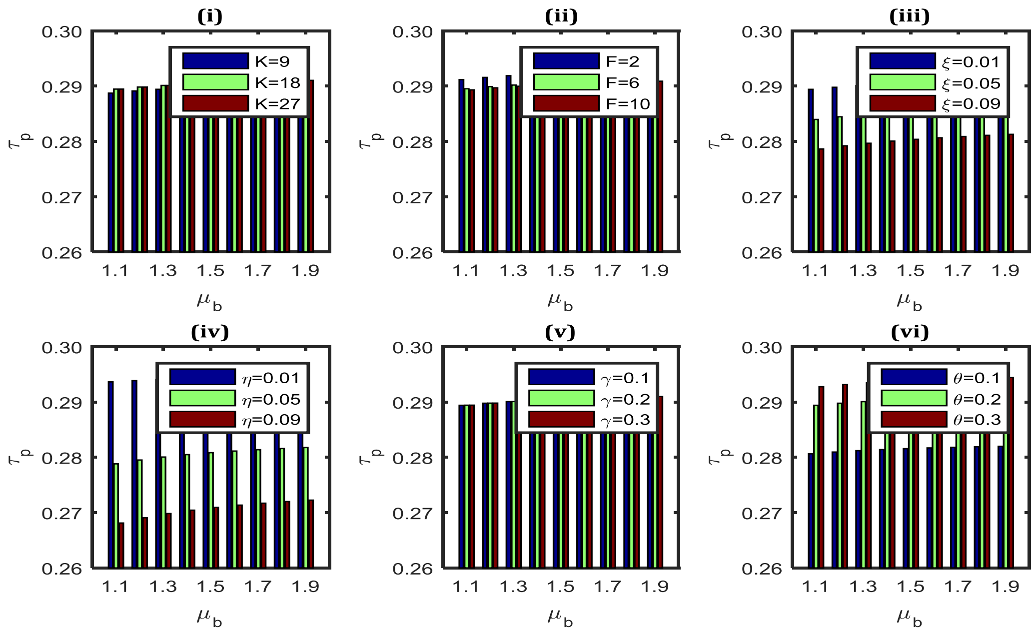

Figure 5.5: Throughput of the service system (τ_p) wrt μ_b for (i) K, (ii) F, (iii) ξ, (iv) η, (v) γ, and (vi) θ.

prominent for the higher value of λ and lower value of μ_b. Some opposite trends are observed for F in Figs. 5.4 and 5.5. These results prompt that increasing the service facility need not always give increased throughput. For more throughput, system design, vacation policy etc are also governing.

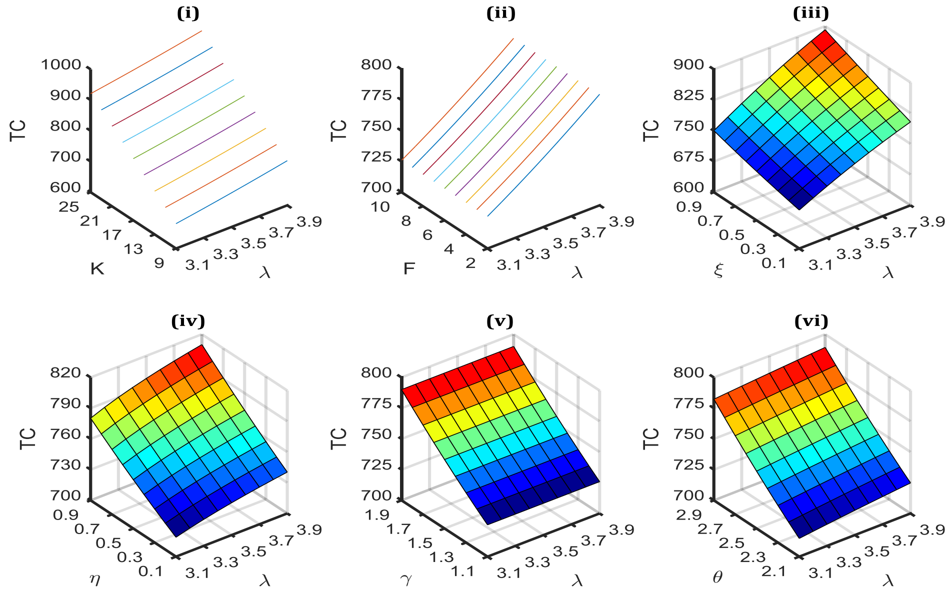

Figure 5.6: Expected total cost of the service system (TC) wrt λ for (i) K, (ii) F, (iii) ξ, (iv) η, (v) γ, and (vi) θ.

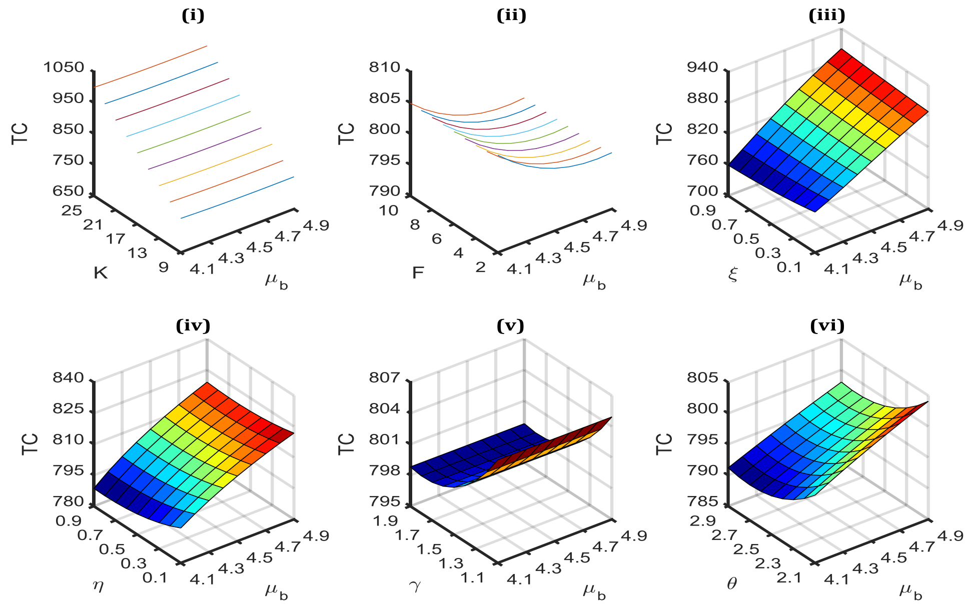

Figure 5.7: Expected total cost of the service system (TC) wrt μ_b for (i) K, (ii) F, (iii) ξ, (iv) η, (v) γ, and (vi) θ.

For various system parameters, the change in the value of the total expected cost (TC) formulated in eq^n(5.21) is depicted in Figs. 5.6 and 5.7 with respect to arrival rate (λ) and service rate in the normal busy mode (μ_b) respectively. For finding the value of expected total cost (TC), we fix various unit cost elements as $C_h = 5$, $C_b = 200$, $C_l = 150$, $C_s = 100$, $C_w = 60$, $C_k = 15$, $C_1 = 50$, $C_2 = 12$, $C_3 = 170$, and

$C_{vi} = 10$ along with default values of service system parameters in Figs. 5.2–5.5 and Tables 5.1–5.2. It is cleared from all sub-figures of Fig. 5.6 that the expected total cost TC increases for all system parameters. Fig. 5.7 shows mix trends of increasing and decreasing of TC with respect to μ_b. It prompts μ_b as a decision variable which is expected to be optimal for the minimum cost.

Table 5.1: Service system characteristics wrt K, λ, μ_b, and ξ.

(K,λ,μ_b,ξ)	L_S	P_B	P_{WV}	P_{VI}	τ_p	P_S	P_L
(10, 4.0, 4.5, 0.2)	1.77352	0.71760	0.28233	7.31E-05	3.08982	2.23E-03	4.17E-03
(10, 4.5, 4.5, 0.2)	2.16296	0.74826	0.25162	1.27E-04	3.34533	5.16E-03	9.65E-03
(10, 5.0, 4.5, 0.2)	2.58159	0.78115	0.21866	1.94E-04	3.56628	1.03E-02	1.92E-02
(10, 4.0, 5.0, 0.2)	1.56275	0.67997	0.31995	8.28E-05	3.17651	1.36E-03	2.48E-03
(10, 4.0, 5.5, 0.2)	1.40401	0.64744	0.35247	9.13E-05	3.24272	8.74E-04	1.54E-03
(10, 4.0, 4.5, 0.3)	1.49195	0.68595	0.31401	3.98E-05	2.92149	8.41E-04	1.57E-03
(10, 4.0, 4.5, 0.4)	1.25634	0.65612	0.34386	1.85E-05	2.75908	2.59E-04	4.84E-04
(15, 4.0, 4.5, 0.2)	1.79671	0.71856	0.28137	7.30E-05	3.09552	2.91E-05	9.05E-05
(15, 4.5, 4.5, 0.2)	2.22599	0.75072	0.24915	1.26E-04	3.35871	1.23E-04	3.81E-04
(15, 5.0, 4.5, 0.2)	2.72559	0.78634	0.21347	1.90E-04	3.59247	4.14E-04	1.29E-03
(15, 4.0, 5.0, 0.2)	1.57507	0.68053	0.31939	8.28E-05	3.18040	1.30E-05	3.86E-05
(15, 4.0, 5.5, 0.2)	1.41090	0.64777	0.35214	9.13E-05	3.24541	6.18E-06	1.77E-05
(15, 4.0, 4.5, 0.3)	1.49917	0.68627	0.31369	3.98E-05	2.92344	5.57E-06	1.73E-05
(15, 4.0, 4.5, 0.4)	1.25812	0.65620	0.34378	1.85E-05	2.75961	7.81E-07	2.43E-06
(20, 4.0, 4.5, 0.2)	1.79729	0.71858	0.28135	7.30E-05	3.09563	1.56E-07	6.46E-07
(20, 4.5, 4.5, 0.2)	2.22887	0.75081	0.24907	1.26E-04	3.35916	1.21E-06	5.01E-06
(20, 5.0, 4.5, 0.2)	2.73688	0.78666	0.21315	1.90E-04	3.59396	7.09E-06	2.95E-05
(20, 4.0, 5.0, 0.2)	1.57529	0.68054	0.31938	8.28E-05	3.18045	5.26E-08	2.09E-07
(20, 4.0, 5.5, 0.2)	1.41100	0.64777	0.35213	9.13E-05	3.24544	1.94E-08	7.38E-08
(20, 4.0, 4.5, 0.3)	1.49926	0.68627	0.31369	3.98E-05	2.92346	1.50E-08	6.22E-08
(20, 4.0, 4.5, 0.4)	1.25813	0.65620	0.34378	1.85E-05	2.75961	9.57E-10	3.97E-09

Furthermore, Tables 5.1 and 5.2 provide the results of a numerical simulation by fixing the default values of parameters as $\lambda = 4$, $\mu_v = 3$, $\mu_b = 5$, $\xi = 0.3$, $\eta = 0.3$, $\gamma = 2$, and $\theta = 3$ and check the variability effects by combining varied values of several system parameters to observe the effects on various performance indices. From Table 5.1, it is cleared that the expected number of customers in the service system (L_S), probability that the server is in busy mode (P_B) and probability that server's vacation is interrupted (P_{VI}) increase as the value of λ increases. The reverse trend is observed for the probability that the server is on a working vacation attribute (P_{WV}) with the variation of λ. These observations are quite obvious in general and validate our modeling. Similarly, on increasing the value of μ_b, the value of L_S, P_B and P_{VI} decreases and the value of P_{WV} and τ_p increases as expected. In a similar manner, expected results are observed in Table 5.2 also, from which the inference for system design can be interpreted easily. The comparative study via numerical simulation is very useful for system designers in order to take a better decision for achieving the optimal service strategy.

Table 5.2: Service system characteristics wrt F, η, γ, μ_v, and θ.

$(F,\eta,\gamma,\mu_v,\theta)$	L_S	P_B	P_{WV}	P_{VI}	τ_p	P_S	P_L
(4, 0.2, 1.5, 2.5, 2.5)	1.49209	0.62607	3.71E-01	0.00314	3.00584	2.08E-04	0.00059
(4, 0.3, 1.5, 2.5, 2.5)	1.41685	0.61694	3.80E-01	0.00294	2.95646	1.03E-04	0.00027
(4, 0.4, 1.5, 2.5, 2.5)	1.35622	0.60904	3.88E-01	0.00275	2.91345	5.31E-05	0.00013
(4, 0.2, 2.0, 2.5, 2.5)	1.49219	0.62607	3.71E-01	0.00314	3.00597	1.56E-04	0.00054
(4, 0.2, 2.5, 2.5, 2.5)	1.49226	0.62607	3.71E-01	0.00314	3.00605	1.25E-04	0.00051
(4, 0.2, 1.5, 3.0, 2.5)	1.43328	0.62289	3.75E-01	0.00242	3.03340	1.88E-04	0.00054
(4, 0.2, 1.5, 3.5, 2.5)	1.38098	0.62022	3.78E-01	0.00188	3.05876	1.73E-04	0.00049
(4, 0.2, 1.5, 2.5, 3.0)	1.44788	0.66244	3.36E-01	0.00204	3.02539	1.91E-04	0.00054
(4, 0.2, 1.5, 2.5, 3.5)	1.41493	0.69272	3.06E-01	0.00136	3.04041	1.80E-04	0.00051
(7, 0.2, 1.5, 2.5, 2.5)	1.50099	0.62409	3.76E-01	0.00016	3.00368	2.18E-04	0.00046
(7, 0.3, 1.5, 2.5, 2.5)	1.42388	0.61500	3.85E-01	0.00012	2.95393	1.08E-04	0.00021
(7, 0.4, 1.5, 2.5, 2.5)	1.36195	0.60717	3.93E-01	0.00010	2.91078	5.58E-05	0.00010
(7, 0.2, 2.0, 2.5, 2.5)	1.50114	0.62410	3.76E-01	0.00016	3.00380	1.64E-04	0.00041
(7, 0.2, 2.5, 2.5, 2.5)	1.50124	0.62410	3.76E-01	0.00016	3.00387	1.31E-04	0.00037
(7, 0.2, 1.5, 3.0, 2.5)	1.43983	0.62092	3.79E-01	0.00010	3.03185	1.95E-04	0.00041
(7, 0.2, 1.5, 3.5, 2.5)	1.38546	0.61831	3.82E-01	0.00007	3.05778	1.77E-04	0.00037
(7, 0.2, 1.5, 2.5, 3.0)	1.45303	0.66131	3.39E-01	0.00008	3.02430	1.97E-04	0.00041
(7, 0.2, 1.5, 2.5, 3.5)	1.41812	0.69205	3.08E-01	0.00004	3.03986	1.83E-04	0.00039
(10, 0.2, 1.5, 2.5, 2.5)	1.50219	0.62401	3.76E-01	0.00001	3.00375	2.33E-04	0.00033
(10, 0.3, 1.5, 2.5, 2.5)	1.42452	0.61492	3.85E-01	0.00000	2.95388	1.13E-04	0.00016
(10, 0.4, 1.5, 2.5, 2.5)	1.36232	0.60710	3.93E-01	0.00000	2.91070	5.81E-05	0.00008
(10, 0.2, 2.0, 2.5, 2.5)	1.50239	0.62402	3.76E-01	0.00001	3.00386	1.77E-04	0.00028
(10, 0.2, 2.5, 2.5, 2.5)	1.50252	0.62403	3.76E-01	0.00001	3.00392	1.43E-04	0.00024
(10, 0.2, 1.5, 3.0, 2.5)	1.44074	0.62085	3.79E-01	0.00000	3.03195	2.06E-04	0.00029
(10, 0.2, 1.5, 3.5, 2.5)	1.38617	0.61826	3.82E-01	0.00000	3.05791	1.87E-04	0.00027
(10, 0.2, 1.5, 2.5, 3.0)	1.45383	0.66129	3.39E-01	0.00000	3.02443	2.08E-04	0.00030
(10, 0.2, 1.5, 2.5, 3.5)	1.41874	0.69205	3.08E-01	0.00000	3.04003	1.93E-04	0.00027

5.7.2 Sensitivity Results

In addition to theoretical illustration in section 5.6 for sensitivity of system parameter(s) on expected total, we present the graphical illustration too. For that purpose, the default value of service system design parameters are fixed as $K = 12$, $\lambda = 4$, $\mu_v = 3$, $\mu_b = 5$, $\gamma = 2$, $\theta = 3$, $\eta = 0.3$, $\xi = 0.3$, and $F = 2,5,8$ along with unit cost elements as $C_h = 5$, $C_b = 200$, $C_l = 150$, $C_s = 100$, $C_w = 60$, $C_k = 15$, $C_1 = 50$, $C_2 = 12$, $C_3 = 170$, $C_{vi} = 10$.

Fig. 5.8 represents the sensitivity of the expected total cost function with respect to λ, μ_v, μ_b, θ, γ, η, and ξ for the range $\lambda \sim [1.0,1.1]$, $\mu_b \sim [5.0,6.0]$, $\mu_v \sim [3.0,5.0]$, $\theta \sim [1.0,3.0]$, $\gamma \sim [1.0,3.0]$, $\eta \sim [0,1.0]$, and $\xi \sim [0,1.0]$ respectively. With sign and magnitude of the first-order partial derivative of the expected total cost function with respect to parameter depict the sensitivity corresponding to that parameter. It gives inference about incremental changes with the rate in expected total cost. In Fig. 5.8(i), it is observed that $\frac{\partial TC}{\partial \lambda}$ is monotonically increasing from negative magnitude to positive magnitude. We infer that TC is increasing function of λ and major

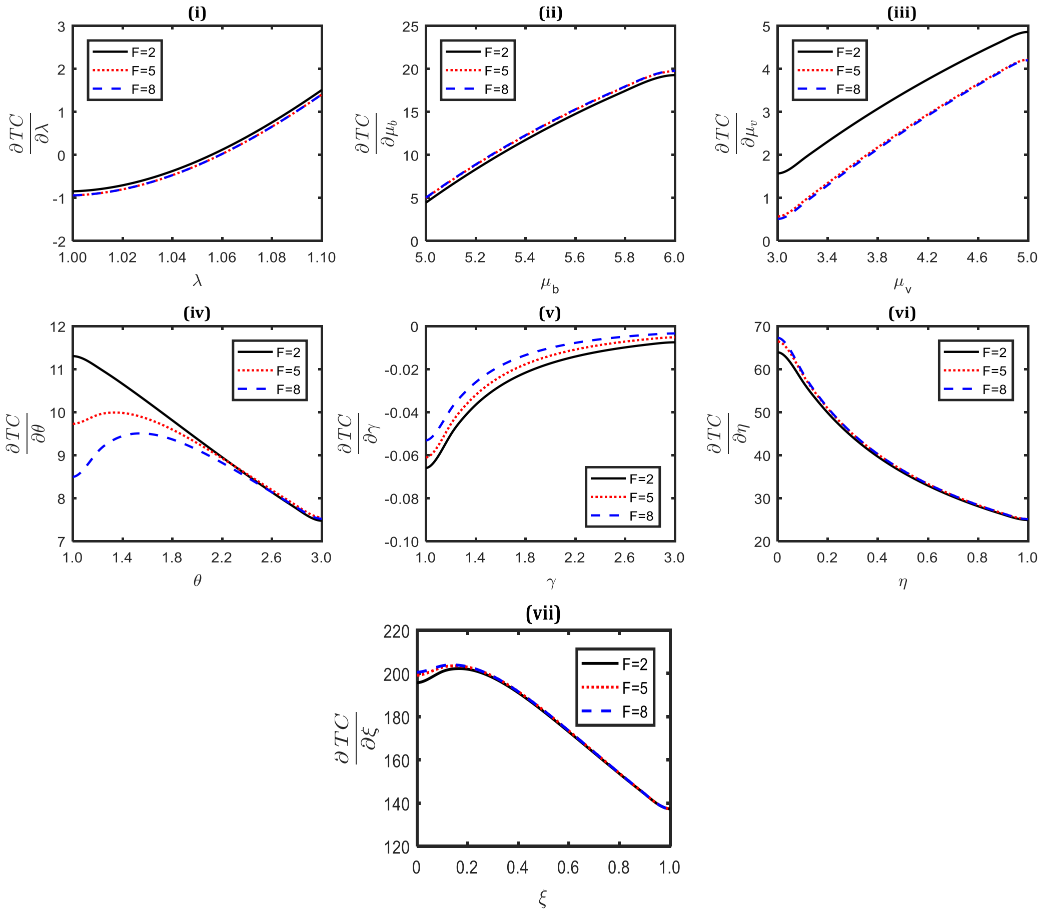

Figure 5.8: Sensitivity analysis of expected cost function (TC) wrt to (i) λ, (ii) μ_b, (iii) μ_v, (iv) θ, (v) γ, (vi) η, and (vii) ξ for different values of F.

incremental change is observed for the high value of λ. From this, one can interpret that for the high demand for service, service provider or analyst must make an arrangement of a better or alternative service facility. This is a decreasing function of F that makes F as decision design parameter. In Fig. 5.8(ii) $\frac{\partial TC}{\partial \mu_b}$ is positive and increasing with respect to μ_b and F, which means that the incremental change in μ_b increases the expected total cost. This implies that μ_b and F are decision parameters to limit the expected total cost of the queue-based service system. Similar implication and inference can be drawn for μ_v from Fig. 5.8(iii). The shape of $\frac{\partial TC}{\partial \theta}$ with respect to θ in Fig. 5.8(iv) makes θ a prime parameter for a service system. The permissible duration of vacation is really crucial for decision and making the service system better. In Fig. 5.8(v), $\frac{\partial TC}{\partial \gamma}$ is negative, which means the incremental change in γ improves the expected cost. In a similar fashion from Fig. 5.8(vi) and Fig. 5.8(vii), it is observed that $\frac{\partial TC}{\partial \eta}$ and $\frac{\partial TC}{\partial \xi}$ are positive respectively but having opposite nature of change. These results resemble the observation in Figs. 5.6 and 5.7. Our observations validate and tractable our modeling and methodology used for the solution.

5.7.3 Optimal Analysis

It is prompt in previous subsections that the design parameters K, F, and service rates μ_b, μ_v are prime parameters for the optimal strategy of the studied queue-based service system and its expected total cost. The particle swarm optimization (PSO) technique is employed for determining the optimal value of decision parameters K^*, F^*, μ_b^*, and μ_v^* at the minimum expected total cost (TC^*). For this purpose, we ensure that the expected total cost function of the service system is convex graphically in Fig. 5.9 since the analytical proof is not possible due to the complexity of the governing function. Fig. 5.9 describes the structure of the expected total cost function for decision variables F, μ_v, and μ_b by fixing the cost elements and system parameters as mentioned above.

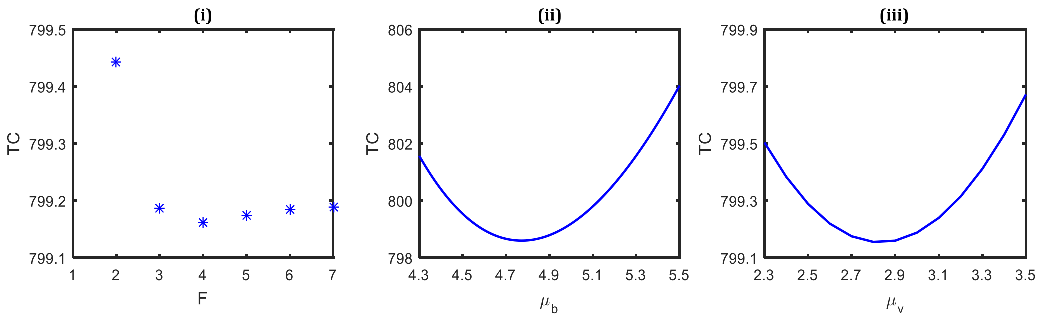

Figure 5.9: Convexity of expected total cost function (TC) wrt parameters (i) F, (ii) μ_b, and (iii) μ_v.

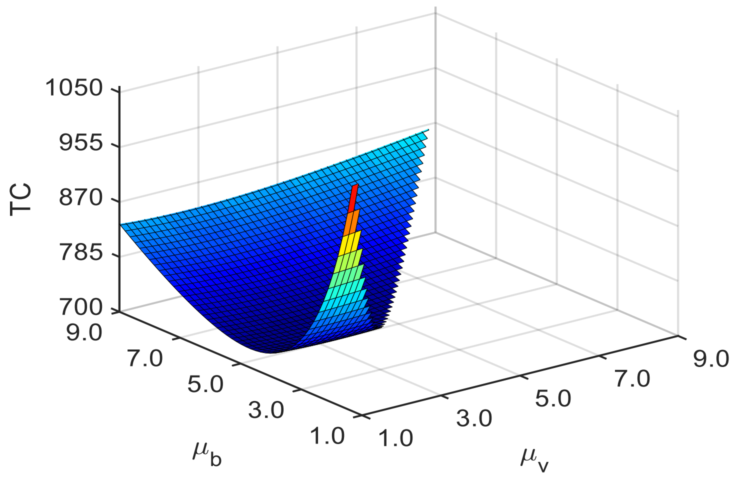

Figure 5.10: Surface plot of expected total cost function for joint values of μ_v, and μ_b.

Furthermore, from the surface plot of TC with respect to service rates μ_b and μ_v in Fig. 5.10, it is noticed that the expected cost function remains convex for joint parameters μ_v and μ_b also. Therefore, Figs. 5.9 and 5.10 deduce that the expected total cost function is convex in nature with respect to decision variables F, μ_v and μ_b.

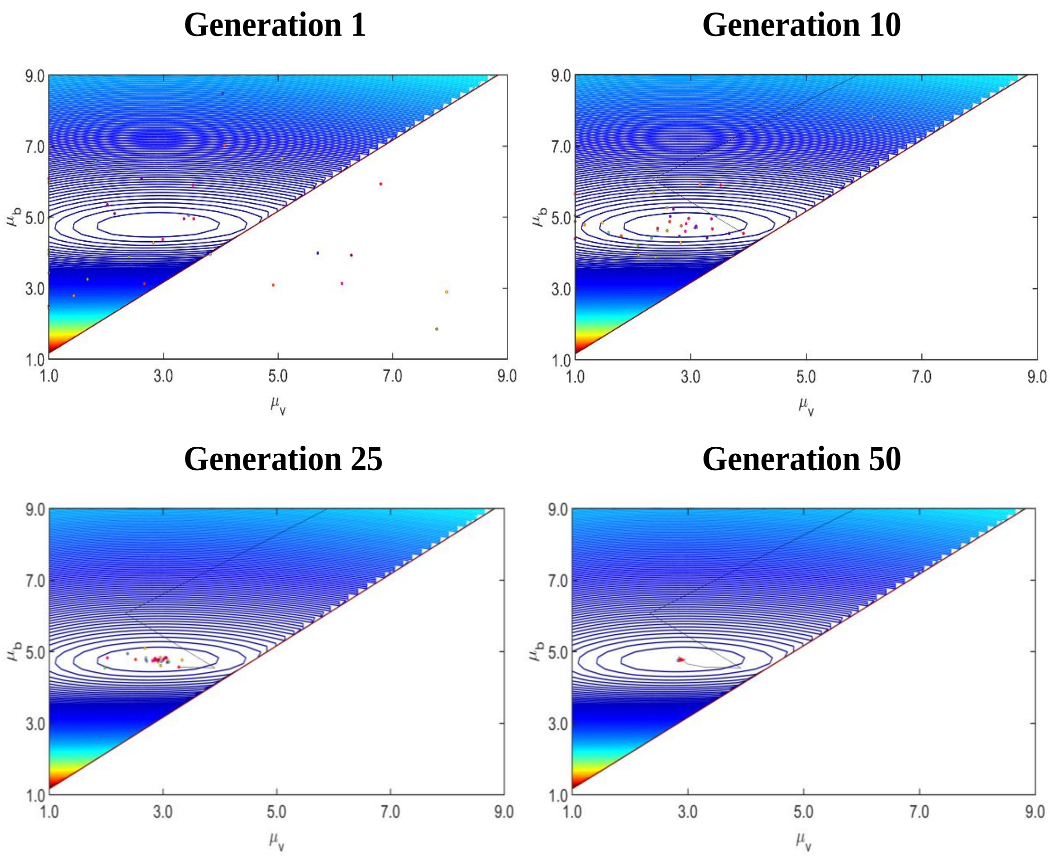

Figure 5.11: Various generations of PSO algorithm over contour plot of the objective function.

PSO is a generation based technique for computing the optimal value of the decision variables. The search particles move along with the directed path (exploitation) and a considerable amount randomness factor (exploration). For the illustration purpose, the default value of the parameters involved in setting the service system is same as for the example of Fig. 5.10 and optimize TC for μ_v and μ_b. The default parameters of particle swarm optimization κ_1, κ_2 are set as 2 and the standard value of the inertia factor between 0.5 to 0.9. The lower and upper bounds for discrete decision variables F and K are taken as $\{3, 10\}$ and $\{4, 15\}$, respectively and [1, 8] for both continuous decision variables μ_v and μ_b. 50 independent runs for each numerical experiment by the PSO algorithm are repeatedly performed. Here, We depict the selective generation of PSO in search of optimal values for demonstration purposes in Fig. 5.11. Initially, all these particles are randomly interspersed across the whole space, as in generation 1 in Fig. 5.11 and come closer and closer through generations (generation 5, 15, 25, 40) in Fig. 5.11 to the optimum (generation 50) along with exploring extensively at the untested regions in the search space. As the generations pass, all the search agents converge to the optimum and the coordinates of the position of the best particle is defined as the optimal value $[\mu_v^*, \mu_b^*] = [2.876073, 4.761424]$ for a minimum expected total cost $TC^* = 724.593796$ as shown in Fig. 5.11.

Table 5.3: Optimal expected total cost for F^*, K^*, μ_v^*, and μ_b^*.

$(\lambda,\xi,\eta,\gamma,\theta)$	F^*	K^*	μ_v^*	μ_b^*	TC^*	Mean $\{\frac{TC}{TC^*}\}$	Max $\{\frac{TC}{TC^*}\}$
(4.0, 0.3, 0.3, 2, 3)	8	9	4.710115	4.720133	739.294775	1.0000001626	1.0000004643
(4.5, 0.3, 0.3, 2, 3)	8	9	5.158815	5.168821	786.479111	1.0000001524	1.0000004019
(5.0, 0.3, 0.3, 2, 3)	8	10	5.604838	5.614838	848.948901	1.0000000866	1.0000001503
(4.0, 0.1, 0.3, 2, 3)	7	9	4.360302	4.370335	769.214655	1.0000038405	1.0000112339
(4.0, 0.5, 0.3, 2, 3)	5	6	4.575546	5.108877	666.745578	1.0000000001	1.0000000002
(4.0, 0.3, 0.1, 2, 3)	5	6	4.647428	4.733439	689.904124	1.0000000002	1.0000000007
(4.0, 0.3, 0.5, 2, 3)	8	9	4.678164	4.688187	745.257597	1.0000000751	1.0000001634
(4.0, 0.3, 0.3, 3, 3)	6	7	2.876073	4.761424	724.593796	1.0000000000	1.0000000001
(4.0, 0.3, 0.3, 4, 3)	6	7	1.000021	4.761634	726.797971	1.0000000000	1.0000000000
(4.0, 0.3, 0.3, 2, 4)	7	8	4.706578	4.716635	724.345017	1.0000000730	1.0000001695
(4.0, 0.3, 0.3, 2, 5)	7	8	4.705824	4.715838	724.292885	1.0000007442	1.0000020526

Table 5.3 shows the optimal values F^*, K^*, μ_v^*, and μ_b^* along with minimal expected total cost (TC^*) for different values combination of various system parameters λ, ξ, η, γ, and θ for the unit cost elements assumed for PSO illustrative example. Moreover, for justification of the metaheuristic method PSO's results, the statistical measures mean ratio and max ratio are calculated. The mean and max ratios are determined out by average and maximum of ratio $\left(\frac{TC}{TC^*}\right)$, where TC is the solution generated by the PSO algorithm in each independent run and TC^* is the best solution among 50 independent runs for the minimized cost optimization problem. It is clear from the depicted results in Table 5.3 that mean and max ratios varying between 1.0000000000 and 1.0000112339 which implies that PSO search for the optimality is very strong and converges to the optimal solution within a reasonable time with population size 50 as compared to other metaheuristic algorithms in practice.

5.8 Conclusion and Future Scope

In this chapter, the mathematical modeling of the finite capacity Markovian service system with a working vacation for a random period, randomized vacation interruption, and randomized controlled arrival policy is presented. To compute the stationary probabilities, the matrix analytic approach is employed, and the computer code is developed in MATLAB. For the optimal and sensitive analysis, the expected total cost function is developed in terms of decision parameters K, F, μ_v, and μ_b.

The metaheuristic algorithm, particle swarm optimization, has been implemented to determine optimal operating conditions and global minimum of expected total cost function. Sensitivity analysis of the various service system performance measures has also been done for varied values of the system parameters. The results rendered in this chapter are beneficial in the contexts of modeling of queue-based service systems, production systems, computer communication systems, automated machine quality control systems, and many other related applications.

We can extend our present study in the future, including some more realistic service control policies namely N-policy, T-policy, for finite population queueing models with various queueing terminologies such as geometric abandonment and feedback policy, or, non-Markovian models with variant working vacation, etc.