ANFIS-BASED COMPUTING FOR BULK QUEUEING FRAMEWORK WITH LOW-BATCH SERVICE, STANDBY SUPPORT AND VACATION STRATEGIES

S. Karpagam[1], N. Aarthy[2]

•

[1,2]Vel Tech Rangarajan Dr. Sagunthala R& D Institute of Science and Technology, India.
Tamil Nadu, India.
[1]karpagammaths19@gmail.com, [2]aarthyselvi2008@gmail.com

Abstract

This study investigates a non-Markovian queueing model involving a single server providing bulk and low-batch services. The model incorporates key operational features such as server breakdown and repair, a standby server and multiple vacations. The service times for the primary server, low-batch service, standby server and vacation are governed by general distributions, whereas breakdown and repair follow exponential distributions. While the primary (bulk) server is under repair, the standby server temporarily takes over service. The probability generating function for queue size at any time is determined, and various system performance metrics are evaluated. Numerical examples are used to demonstrate the analytical findings. Furthermore, the analytical results are validated using the Adaptive Neuro-Fuzzy Inference System (ANFIS), which improves the precision and dependability of the model's predictions.

Keywords: ANFIS, bulk service, repair, standby, breakdown, low-batch, and multiple vacations.

1. Introduction

Queueing theory is a mathematical framework that can be used to evaluate service systems. Queueing models can aid decision makers when making and improving performance evaluations within a very diverse list of operations like communication networks, factory manufacturing facilities, and cloud computing services. Bulk service models are useful in organizations that handle multiple customers or jobs simultaneously and help improve efficiency. In some cases, the principal machine may fail; but other servers (standbys) of the system can carry processes and jobs on. This standby server is essential to avoid disruptions and keep service steady when system failures take place unexpectedly. It is therefore responsible for both standby and maintenance aid to the queue system so that it is always reliable and operational.

In addition, vacation policies within queueing systems enable the server to take pre-scheduled breaks that may be utilized for basic jobs like cleaning up or preventive maintenance or system upgrading or handling other tasks. By strategically utilizing the planned breaks, service improves, downtime can be reduced during periods of low demand, and the overall performance and reliability of the system can be improved. This section will review the relevant literature sequentially by the major operational components of the proposed model: bulk service; standby server; server failure; recovery and multiple vacations. Much research in queueing theory has looked at each of these components, addressing different operational issues that occur in practical service systems. The foundation for the development and analysis of the current model is laid by the literature review that follows, which highlights significant contributions in these fields.

In this queueing theory field, Neuts [17] established the foundation work by presenting a general type of queueing that involved Poisson arrivals.Ayyappan and Karpagam examined the single-server queueing model in [2]. Standby servers, multiple vacations, a re-service control mechanism, batch arrivals, and server failures with repair are some of its useful features. Also, Ayyappan et al. [3] examined a MAP/PH/1 system and the service is optional including setup and phase-type repair. In order to illustrate the practical aspects of production systems, Karpagam and Lokesh [10] looked at an arrival (bulk) fixed batch service with rework, inspection, and multiple vacation policies.

To enhance reliability and prevent service disruption at breakdown or overload, it is common to use standby servers. Yeh et al. [4] studied a redundant system of standby support with controlled failure machine arrival. Madan et al. [11] and [12] formulated a queueing system with a vacation and repair of the main server where the system is covered by a standby server. Also, Murugeswari and Sundari [16] studied a bulk arrival queueing system with a mandatory vacation of the server and standby service. Jeyakumar and Senthilnathan [9] examined a bulk service queueing system with vacation and server breakdowns and formulated performance consequences under service disciplines of this type.

Doshi [8] presented a systematic study of queueing systems with vacationing servers and their role in system performance. Takagi's work [23], Queueing Analysis: A Foundation of Performance Evaluation, is a book containing a survey of vacation and priority systems. Bulk service queueing models where many vacations were considered along with setup times and service policies have been explored by Reddy et al. [20] and Ayyappan and Deepa [1]. Tamrakar et al. [24] performed the batch size-dependent bulk service queue, including multiple working vacations. Chakravarthy and Kulshrestha's study [5] focuses on a queueing model that includes multiple vacations, repair, server failures, and a backup server. Models for integrating server failure without service interruption were proposed by Jeyakumar and Senthilnathan [21].

Utilizations of idle time are among classical problems in queueing theory. Levy and Yechiali [15] discussed how vacation systems use idle time. Whether to enter into vacation or idle state based on queue level thresholds is critically significant for optimal performance of systems as addressed by Choudhury and Madan [6]. n order to improve server accessibility through two-phase imperative services and a backup server, Rathinasamy and Kirupa [19] presented a model for minimizing wait times.. Pradhan et al. [18] formulated performance in group arrival and queue-dependent vacation systems. Srivastava et al. [22] examined dual-service types and multiple vacations in a system, with increased service flexibility and decreased server idleness, respectively, while Lavanya et al. [13] talked about ceramic technology and dual-service vacation systems. To enhance the precision in numerical outcomes, current works employing ANFIS techniques were highlighted. Divya and Indhira [7], who studied a heterogeneous framework with a second optional service involving hybrid vacation, whereas Thakur et al. [25] discussed optimization of expenses through ANFIS in a Markovian queue with operational vacations.

1.1. Gap in the Existing Literature

Madan et al. [12] concentrated on single-service queueing systems with batch arrivals, which included a backup server to ensure service continuity during server breakdowns. Later, Karpagam et al. [2] applied this concept to bulk service systems, creating models that incorporated server failures and maintenance, standby service, multiple vacations, and customer reprocessing strategies. However, these studies focus exclusively on single or bulk service mechanisms, ignoring cases in which low-batch services are also required. In many real-world systems, particularly in manufacturing and production, bulk and low-batch services coexist and have a substantial impact on performance. To close this gap, the current study expands the bulk service model with standby service by incorporating low-batch service. The proposed model efficiently handles varying customer demand, reducing queue length and waiting time.

1.2. Structure of the Paper

The manuscript is structured as follows: Section 2 outlines the proposed model, encompassing the subsection Real-World Applications.Section 3 gives the distributions of queue size, incorporating subsections Computational Aspects, Stability Analysis, and Special Case. Section 4 discusses performance metrics. Section 5 illustrates the model numerically and compares the findings to those obtained using the ANFIS technique. Finally, Section 6 has the conclusion.

2. Model description

The model studied in this work involves arrivals following a compound Poisson process with rate λ. According to the general bulk service rule, both the primary and standby servers serve customers. The primary server starts service only when there are at least a units in the queue, and it can serve up to b units simultaneously. When there are fewer than a customers in the queue ($1 \leq Q \leq a - 1$), the server switches to low-batch service. If the number of customers in the queue is equal to or greater than b, they are served on the primary server. During regular service, the primary server is likely to fail, which occurs at an exponential rate of β. When a server fails, it is automatically repaired based on an exponential distribution of rate δ. During this period, the service of the current batch is interrupted. The interrupted batch is transferred to a standby unit, which resumes service for that batch. The standby unit remains operational in the system until the primary server is restored. If the standby server is still busy after the repair is complete, the current batch is transferred to the primary server, which then resumes service. After completing the bulk and low-batch service phases, if there are no customers in the queue, the server starts a random time interval. If the queue is empty at the end of this interval, the server immediately starts another interval. This cycle continues until at least a customers are in the queue. Until, the next service is triggered. Furthermore, the primary server, standby server, and low-volume service have independent service times with a general distribution.
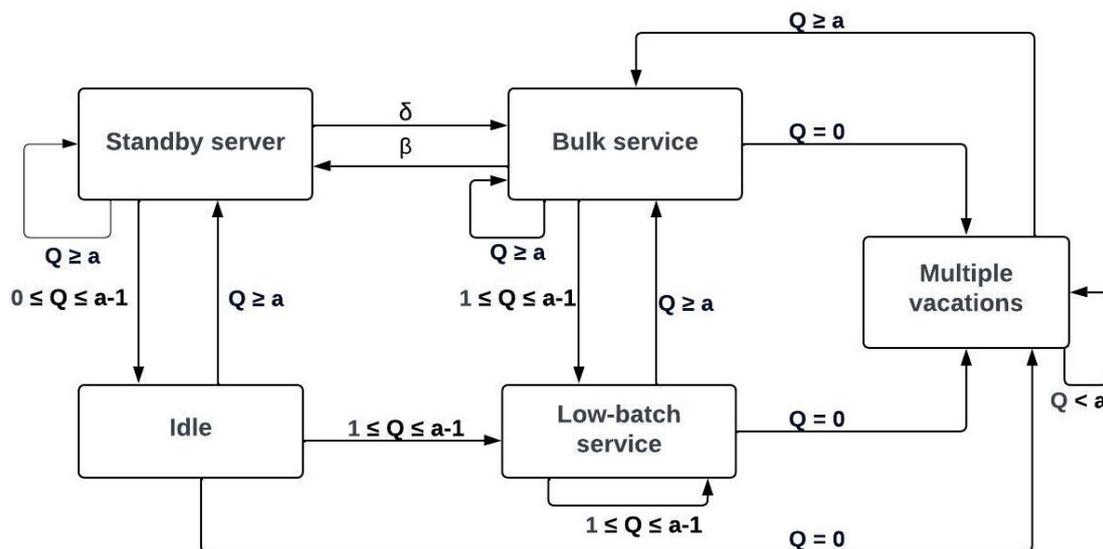
Figure 1: *Schematic Representation*

2.1. Real-Time Relevance of the Proposed Framework

In cloud computing, the system functions across different states to deliver reliability, scalability, and optimal resource management. In the bulk service state, cloud resources, such as compute instances or packaging, handle significant workloads by processing huge quantities of data or serving several concurrent user requests. Such services are typically auto-scaled to meet peak demand and managed by orchestration systems like Kubernetes or AWS Auto Scaling. When the primary compute instance fails due to system overload, software failure, or hardware breakdown, the standby server comes online. To reduce downtime and guarantee service continuity, this standby resource usually a warmed instance or hot backup node, is instantly deployed to take over the interrupted task. During low-incoming workload situations, the system switches to a low-batch service mode, in which limited functions like AWS Lambda or Azure Functions do minimum or single operations. This mode promotes economical computing since it supplies only those resources required for small jobs. During periods of no active requests or jobs in the queue, the system goes into vacation mode.

In this phase, background tasks such as software updates, patching of security vulnerabilities, archiving of logs and system checks are executed while the resources that are idle are reduced or halted. It ensures the infrastructure is tuned and is secure to handle future workloads.Idle state refers to when resources are committed but are not being run, ready state to deploy instantly when needed. These resources may be periodically checked for health or readiness. All cloud computing paradigm states, as a general perspective, are essential to ensure operational efficiency, cost reduction, and continuity of service delivery under varying workloads and system configurations.

Table 1: *Description of Notations*

Symbol	Description
λ	Rate at which customers arrive.
X	Random variable representing the group size.
$Pr(X = k) = g_k$	Probability that a group of size k arrives.
$X(z)$	Probability generating function (PGF) of X.
$S(.), U(.), C(.), V(.)$	Bulk service, standby, low-batch service and vacation periods cumulative distribution functions (CDFs).
$s(\varkappa), u(\varkappa), c(\varkappa), v(\varkappa)$	Probability density functions (PDFs) of S, U, C, and V.
$S^0(t), U^0(t), C^0(t), V^0(t)$	Remaining times of bulk service, standby service, low-batch service and vacation at time t.
$\tilde{S}(\zeta), \tilde{U}(\zeta), \tilde{C}(\zeta), \tilde{V}(\zeta)$	Laplace-Stieltjes transforms (LSTs) of S, U, C, and V.
$N_{\mathfrak{s}}(t)$	Customers currently receiving service at time t.
$N_{\mathfrak{q}}(t)$	Number of customers in the queue at time t.
$Z(t) = \mathfrak{F}$	Server is taking its $\mathfrak{F}^{th}$ vacation at time t.

For the further development of the queueing system, let us define the random variable as:

$$\Re(t) = \begin{cases} 0, & \text{when the server is on vacation} \\ 1, & \text{while the server is providing bulk service} \\ 2, & \text{when the server is engaged in low-batch processing} \\ 3, & \text{while the standby server is serving customers (bulk service)} \\ 4, & \text{when the standby server is idle} \end{cases}$$

The probabilities of states are defined as follows:

$$I_\mathfrak{h}(t)dt = Pr\{N_\mathfrak{q}(t) = \mathfrak{h}, \Re(t) = 4\}, 0 \leq \mathfrak{h} \leq a-1,$$

$$\mathcal{B}_{\mathfrak{e},\mathfrak{F}}(\varkappa, t)dt = Pr\{N_\mathfrak{s}(t) = \mathfrak{e}, N_\mathfrak{q}(t) = \mathfrak{F}, \varkappa \leq S^0(t) \leq \varkappa + dt, \Re(t) = 1\}, a \leq \mathfrak{e} \leq b, \quad \mathfrak{F} \geq 0,$$

$$T_{\mathfrak{e},\mathfrak{F}}(\varkappa, t)dt = Pr\{N_\mathfrak{s}(t) = \mathfrak{e}, N_\mathfrak{q}(t) = \mathfrak{F}, \varkappa \leq U^0(t) \leq \varkappa + dt, \Re(t) = 3\}, a \leq \mathfrak{e} \leq b, \quad \mathfrak{F} \geq 0,$$

$$L_{\mathfrak{e},\mathfrak{F}}(\varkappa, t)dt = Pr\{N_\mathfrak{s}(t) = \mathfrak{e}, N_\mathfrak{q}(t) = \mathfrak{F}, \varkappa \leq C^0(t) \leq \varkappa + dt, \Re(t) = 2\}, 1 \leq \mathfrak{e} \leq a-1, \quad \mathfrak{F} \geq 0,$$

$$Q_{l,\mathfrak{F}}(\varkappa, t)dt = Pr\{N_\mathfrak{q}(t) = \mathfrak{F}, \varkappa \leq V^0(t) \leq \varkappa + dt, \Re(t) = 0\}, \quad l \geq 1, \quad \mathfrak{F} \geq 0.$$

3. Queue size distribution

Given the stated probabilities, the steady-state equations are as follows:
The standby server idle

$$(\lambda + \delta)I_0 = \sum_{m=a}^{b} T_{m,0}(0), \tag{1}$$

$$(\lambda + \delta)I_\mathfrak{h} = \sum_{m=a}^{b} T_{m,\mathfrak{h}}(0) + \sum_{k=1}^{\mathfrak{h}} I_{\mathfrak{h}-k}\lambda g_k, \quad 1 \leq \mathfrak{h} \leq a-1. \tag{2}$$

The primary server busy

$$-\mathcal{B}'_{\mathfrak{e},0}(\varkappa) = -(\lambda + \beta)\mathcal{B}_{\mathfrak{e},0}(\varkappa) + s(\varkappa)\sum_{m=a}^{b} \mathcal{B}_{m.\mathfrak{e}}(0) + s(\varkappa)\sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0)$$

$$+ \sum_{l=1}^{\infty} Q_{l,\mathfrak{e}}(0)s(\varkappa) + \delta\int_0^\infty T_{\mathfrak{e},0}(y)dys(\varkappa), \quad a \leq \mathfrak{e} \leq b, \tag{3}$$

$$-\mathcal{B}'_{\mathfrak{e},\mathfrak{F}}(\varkappa) = -(\lambda + \beta)\mathcal{B}_{\mathfrak{e},\mathfrak{F}}(\varkappa) + \sum_{k=1}^{\mathfrak{F}} \mathcal{B}_{\mathfrak{e},\mathfrak{F}-k}(\varkappa)\lambda g_k + \delta\int_0^\infty T_{\mathfrak{e},\mathfrak{F}}(y)dys(\varkappa),$$

$$\mathfrak{F} \geq 1, \quad a \leq \mathfrak{e} \leq b-1, \tag{4}$$

$$-\mathcal{B}'_{b,\mathfrak{F}}(\varkappa) = -(\lambda + \beta)\mathcal{B}_{b,\mathfrak{F}}(\varkappa) + \sum_{k=1}^{\mathfrak{F}} \mathcal{B}_{b,\mathfrak{F}-k}(\varkappa)\lambda g_k + s(\varkappa)\sum_{m=a}^{b} \mathcal{B}_{m,b+\mathfrak{F}}(0)$$

$$+ s(\varkappa)\sum_{m=1}^{a-1} L_{m.b+\mathfrak{F}}(0) + \sum_{l=1}^{\infty} Q_{l,b+\mathfrak{F}}(0)s(\varkappa)$$

$$+ \delta\int_0^\infty T_{b,\mathfrak{F}}(y)dys(\varkappa), \quad \mathfrak{F} \geq 1. \tag{5}$$

The standby server busy

$$-T'_{\mathfrak{e},0}(\varkappa) = -(\lambda + \delta)T_{\mathfrak{e},0}(\varkappa) + \sum_{m=a}^{b} T_{m,\mathfrak{e}}(0)u(\varkappa) + \sum_{k=0}^{a-1} I_k\lambda g_{\mathfrak{e}-k}u(\varkappa)$$

$$+ \beta\int_0^\infty \mathcal{B}_{\mathfrak{e},0}(y)dyu(\varkappa), \quad a \leq \mathfrak{e} \leq b, \tag{6}$$

$$-T'_{\mathfrak{e},\mathfrak{F}}(\varkappa) = -(\lambda + \delta)T_{\mathfrak{e},\mathfrak{F}}(\varkappa) + \sum_{k=1}^{\mathfrak{F}} T_{\mathfrak{e},\mathfrak{F}-k}(\varkappa)\lambda g_k + \beta\int_0^\infty \mathcal{B}_{\mathfrak{e},\mathfrak{F}}(y)dyu(\varkappa),$$

$$a \leq \mathfrak{e} \leq b-1, \quad \mathfrak{F} \geq 1, \tag{7}$$

$$-T'_{b,\mathfrak{F}}(\varkappa) = -(\lambda + \delta)T_{b,\mathfrak{F}}(\varkappa) + \sum_{k=1}^{\mathfrak{F}} T_{b,\mathfrak{F}-k}(\varkappa)\lambda g_k + \sum_{m=a}^{b} T_{m,b+\mathfrak{F}}(0)u(\varkappa)$$

$$+ \sum_{k=0}^{a-1} I_k\lambda g_{b+\mathfrak{F}-k}u(\varkappa) + \beta\int_0^\infty \mathcal{B}_{b,\mathfrak{F}}(y)dyu(\varkappa), \quad \mathfrak{F} \geq 1. \tag{8}$$

The low-batch server busy

$$-Ł'_{\mathfrak{e},0}(\varkappa) = -\lambda L_{\mathfrak{e},0}(\varkappa) + \sum_{m=1}^{a-1} L_{m.\mathfrak{e}}(0)c(\varkappa) + \sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0)c(\varkappa) + c(\varkappa)\delta I_{\mathfrak{e}}, \quad 1 \le \mathfrak{e} \le a-1, \quad (9)$$

$$-L'_{\mathfrak{e},\mathfrak{F}}(\varkappa) = -\lambda L_{\mathfrak{e},\mathfrak{F}}(\varkappa) + \sum_{k=1}^{\mathfrak{F}} L_{\mathfrak{e},\mathfrak{F}-k}(\varkappa)\lambda g_k, \quad 1 \le \mathfrak{e} \le a-1, \quad \mathfrak{F} \ge 1. \quad (10)$$

The Vacation

$$-Q'_{1,0}(\varkappa) = -\lambda Q_{1,0}(\varkappa) + \sum_{m=a}^{b} \mathcal{B}_{m,0}(0)v(\varkappa) + v(\varkappa) \sum_{m=1}^{a-1} L_{m,0}(0) + \delta I_0 v(\varkappa), \quad (11)$$

$$-Q'_{1,\mathfrak{F}}(\varkappa) = -\lambda Q_{1,\mathfrak{F}}(\varkappa) + \sum_{k=1}^{n} Q_{1,\mathfrak{F}-k}(\varkappa)\lambda g_k, \quad \mathfrak{F} \ge 1, \quad (12)$$

$$-Q'_{l,0}(\varkappa) = -\lambda Q_{l,0}(\varkappa) + v(\varkappa)Q_{l-1,0}(0), \quad l \ge 2, \quad (13)$$

$$-Q'_{l,\mathfrak{F}}(\varkappa) = -\lambda Q_{l,\mathfrak{F}}(\varkappa) + Q_{l-1,\mathfrak{F}}(0)v(\varkappa) + \sum_{k=1}^{n} Q_{l,\mathfrak{F}-k}(\varkappa)\lambda g_k, \quad 1 \le \mathfrak{F} \le a-1, \quad l \ge 2, \quad (14)$$

$$-Q'_{l,\mathfrak{F}}(\varkappa) = -\lambda Q_{l,\mathfrak{F}}(\varkappa) + \sum_{k=1}^{n} Q_{l,\mathfrak{F}-k}(\varkappa)\lambda g_k, \quad l \ge 2, \quad \mathfrak{F} \ge a. \quad (15)$$

Equations (1)- (15), when transformed using the Laplace-Stieltjes Transform, yield:

$$\zeta\tilde{\mathcal{B}}_{\mathfrak{e},0}(\zeta) - \mathcal{B}_{\mathfrak{e},0}(0) = (\beta+\lambda)\tilde{\mathcal{B}}_{\mathfrak{e},0}(\zeta) - \sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0)\tilde{S}(\zeta) - \sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0)\tilde{S}(\zeta)$$
$$- \sum_{l=1}^{\infty} \tilde{Q}_{l,\mathfrak{e}}(0)\tilde{S}(\zeta) - \delta \int_0^{\infty} T_{\mathfrak{e},0}(y)dy\tilde{S}(\zeta), \quad a \le \mathfrak{e} \le b, \quad (16)$$

$$\zeta\tilde{\mathcal{B}}_{\mathfrak{e},\mathfrak{F}}(\zeta) - \mathcal{B}_{\mathfrak{e},\mathfrak{F}}(0) = (\beta+\lambda)\tilde{\mathcal{B}}_{\mathfrak{e},\mathfrak{F}}(\zeta) - \sum_{k=1}^{\mathfrak{F}} \tilde{\mathcal{B}}_{\mathfrak{e},\mathfrak{F}-k}(\zeta)\lambda g_k - \delta \int_0^{\infty} T_{\mathfrak{e},0}(y)dy\tilde{S}(\zeta),$$
$$\mathfrak{F} \ge 1, \quad a \le \mathfrak{e} \le b-1, \quad (17)$$

$$\zeta\tilde{\mathcal{B}}_{b,\mathfrak{F}}(\zeta) - \mathcal{B}_{b,\mathfrak{F}}(0) = (\beta+\lambda)\tilde{\mathcal{B}}_{b,\mathfrak{F}}(\zeta) - \sum_{m=a}^{b} \mathcal{B}_{m,b+\mathfrak{F}}(0)\tilde{S}(\zeta) - \sum_{m=1}^{a-1} L_{m,b+\mathfrak{F}}(0)\tilde{S}(\zeta)$$
$$- \sum_{k=1}^{\mathfrak{F}} \tilde{\mathcal{B}}_{b,\mathfrak{F}-k}(\zeta)\lambda g_k - \delta \int_0^{\infty} T_{\mathfrak{e},0}(y)dy\tilde{S}(\zeta), \quad \mathfrak{F} \ge 1, \quad (18)$$

$$\zeta\tilde{L}_{\mathfrak{e},0}(\zeta) - L_{\mathfrak{e},0}(0) = \lambda\tilde{L}_{\mathfrak{e},0}(\zeta) - \sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0)\tilde{C}(\zeta) - \sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0)\tilde{C}(\zeta) - \delta I_{\mathfrak{e}}\tilde{C}(\zeta),$$
$$1 \le \mathfrak{e} \le a-1, \quad (19)$$

$$\zeta\tilde{L}_{\mathfrak{e},\mathfrak{F}}(\zeta) - L_{\mathfrak{e},\mathfrak{F}}(0) = \lambda\tilde{L}_{\mathfrak{e},\mathfrak{F}}(\zeta) - \sum_{k=1}^{\mathfrak{F}} \tilde{L}_{\mathfrak{e},\mathfrak{F}-k}(\zeta)\lambda g_k, \quad 1 \le \mathfrak{e} \le a-1, \quad \mathfrak{F} \ge 1, \quad (20)$$

$$\zeta\tilde{T}_{\mathfrak{e},0}(\zeta) - T_{\mathfrak{e},0}(0) = (\lambda+\delta)\tilde{T}_{\mathfrak{e},0}(\zeta) - \sum_{m=a}^{b} T_{m,\mathfrak{e}}(0)\tilde{U}(\zeta) - \beta \int_0^{\infty} \mathcal{B}_{\mathfrak{e},0}(y)dy\tilde{U}(\zeta)$$
$$- \sum_{k=0}^{a-1} I_k\lambda g_{\mathfrak{e}-k}\tilde{U}(\zeta), \quad a \le \mathfrak{e} \le b, \quad (21)$$

$$\zeta\tilde{T}_{\mathfrak{e},\mathfrak{F}}(\zeta) - T_{\mathfrak{e},\mathfrak{F}}(0) = (\lambda+\delta)\tilde{T}_{\mathfrak{e},\mathfrak{F}}(\zeta) - \sum_{k=1}^{\mathfrak{F}} \tilde{T}_{\mathfrak{e},\mathfrak{F}-k}(\zeta)\lambda g_k - \beta \int_0^{\infty} \mathcal{B}_{\mathfrak{e},\mathfrak{F}}(y)dy\tilde{U}(\zeta),$$
$$a \le \mathfrak{e} \le b-1, \quad \mathfrak{F} \ge 1, \quad (22)$$

$$\zeta \tilde{T}_{b,\mathfrak{F}}(\zeta) - T_{b,\mathfrak{F}}(0) = (\lambda + \delta)\tilde{T}_{b,\mathfrak{F}}(\zeta) - \sum_{m=a}^{b} T_{m,b+\mathfrak{F}}(0)\tilde{U}(\zeta) - \sum_{k=1}^{\mathfrak{F}} \tilde{T}_{b,\mathfrak{F}-k}(\zeta)\lambda g_k$$

$$- \beta \int_0^\infty \mathcal{B}_{b,\mathfrak{F}}(y)dy\tilde{U}(\zeta) - \sum_{k=0}^{a-1} I_k\lambda g_{b+\mathfrak{F}-k}\tilde{U}(\zeta), \quad \mathfrak{F} \geq 1, \quad (23)$$

$$\zeta\tilde{Q}_{1,0}(\zeta) - Q_{1,0}(0) = \lambda\tilde{Q}_{1,0}(\zeta) - \sum_{m=a}^{b} \mathcal{B}_{m,0}(0)\tilde{V}(\zeta) - \sum_{m=1}^{a-1} L_{m,0}(0)\tilde{V}(\zeta) - \delta I_0\tilde{V}(\zeta), \quad (24)$$

$$\zeta\tilde{Q}_{1,\mathfrak{F}}(\zeta) - Q_{1,\mathfrak{F}}(0) = \lambda\tilde{Q}_{1,\mathfrak{F}}(\zeta) - \sum_{k=1}^{n} \tilde{Q}_{1,\mathfrak{F}-k}(\zeta)\lambda g_k, \quad \mathfrak{F} \geq 1, \quad (25)$$

$$\zeta\tilde{Q}_{l,0}(\zeta) - Q_{l,0}(0) = \lambda\tilde{Q}_{l,0}(\zeta) - Q_{l-1,0}(0)\tilde{V}(\zeta), \quad l \geq 2, \quad (26)$$

$$\zeta\tilde{Q}_{l,\mathfrak{F}}(\zeta) - Q_{l,\mathfrak{F}}(0) = \lambda\tilde{Q}_{l,\mathfrak{F}}(\zeta) - Q_{l-1,\mathfrak{F}}(0)\tilde{V}(\zeta) - \sum_{k=1}^{n} \tilde{Q}_{l,\mathfrak{F}-k}(\zeta)\lambda g_k,$$

$$\mathfrak{F} = 1,2,...,a-1, \quad l \geq 2, \quad (27)$$

$$\zeta\tilde{Q}_{l,\mathfrak{F}}(\zeta) - Q_{l,\mathfrak{F}}(0) = \lambda\tilde{Q}_{l,\mathfrak{F}}(\zeta) - \sum_{k=1}^{n} \tilde{Q}_{l,\mathfrak{F}-k}(\zeta)\lambda g_k, \quad \mathfrak{F} \geq 2, \quad l \geq 2. \quad (28)$$

To compute the PGF of the queue length, the subsequent PGFs are constructed:

$$\left.\begin{array}{l} \tilde{\mathcal{B}}_\mathfrak{e}(\mathfrak{z},\zeta) = \sum_{\mathfrak{F}=0}^\infty \tilde{\mathcal{B}}_{\mathfrak{e},\mathfrak{F}}(\zeta)\mathfrak{z}^\mathfrak{F}, \quad \mathcal{B}_\mathfrak{e}(\mathfrak{z},0) = \sum_{\mathfrak{F}=0}^\infty \mathcal{B}_{\mathfrak{e},\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F}; \quad a \leq \mathfrak{e} \leq b, \\[3mm] \tilde{L}_\mathfrak{e}(\mathfrak{z},\zeta) = \sum_{\mathfrak{F}=0}^\infty \tilde{L}_{\mathfrak{e},\mathfrak{F}}(\zeta)\mathfrak{z}^\mathfrak{F}, \quad L_\mathfrak{e}(\mathfrak{z},0) = \sum_{\mathfrak{F}=0}^\infty L_{\mathfrak{e},\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F}; \quad 1 \leq \mathfrak{e} \leq a-1, \\[3mm] \tilde{T}_\mathfrak{e}(\mathfrak{z},\zeta) = \sum_{\mathfrak{F}=0}^\infty \tilde{T}_{\mathfrak{e},\mathfrak{F}}(\zeta)\mathfrak{z}^\mathfrak{F}, \quad T_\mathfrak{e}(\mathfrak{z},0) = \sum_{\mathfrak{F}=0}^\infty T_{\mathfrak{e},\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F}; \quad a \leq \mathfrak{e} \leq b, \\[3mm] \tilde{Q}_l(\mathfrak{z},\zeta) = \sum_{\mathfrak{F}=0}^\infty \tilde{Q}_{l,\mathfrak{F}}(\zeta)\mathfrak{z}^\mathfrak{F}, \quad Q_l(\mathfrak{z},0) = \sum_{\mathfrak{F}=0}^\infty Q_{l,\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F}; \quad l \geq 1. \end{array}\right\} \quad (29)$$

Applying the appropriate powers of $\mathfrak{z}^n$ to Equations (16)-(28), summing over $n = 0$ *to* ∞ and using Equation (29), yields:

$$(\zeta - l(\mathfrak{z}))\tilde{\mathcal{B}}_\mathfrak{e}(\mathfrak{z},\zeta) = \mathcal{B}_\mathfrak{e}(\mathfrak{z},0) - \tilde{S}(\zeta)\Big[\sum_{m=a}^b \mathcal{B}_{m,\mathfrak{e}}(0) + \sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0) + \sum_{\mathfrak{F}=0}^\infty Q_{l,\mathfrak{F}}(0) + \delta\tilde{T}_\mathfrak{e}(\mathfrak{z},0)\Big],$$

$$a \leq \mathfrak{e} \leq b-1, \quad (30)$$

$$\mathfrak{z}^b(\zeta - l(\mathfrak{z}))\tilde{\mathcal{B}}_b(\mathfrak{z},\zeta) = (\mathfrak{z}^b - \tilde{S}(\zeta))\mathcal{B}_b(\mathfrak{z},0)$$

$$- \tilde{S}(\zeta)\Big[\sum_{m=a}^{b-1} \mathcal{B}_m(\mathfrak{z},0) + \sum_{m=1}^{a-1} L_m(\mathfrak{z},0) + \sum_{l=1}^\infty Q_l(\mathfrak{z},0) + \mathfrak{z}^b\delta\tilde{T}_b(\mathfrak{z},0)$$

$$- \sum_{\mathfrak{F}=0}^{b-1}\Big(\sum_{m=a}^b \mathcal{B}_{m,\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F} + \sum_{m=1}^{a-1} L_{m,\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F} + \sum_{\mathfrak{F}=0}^\infty Q_{l,\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F}\Big)\Big], \quad (31)$$

$$(\zeta - m(\mathfrak{z}))\tilde{T}_\mathfrak{e}(\mathfrak{z},\zeta) = T_\mathfrak{e}(\mathfrak{z},0) - \tilde{U}(\zeta)\Big[\sum_{m=a}^b T_{m,\mathfrak{e}}(0) + \sum_{k=0}^{a-1} I_k\lambda g_{i-k} + \beta\tilde{\mathcal{B}}_\mathfrak{e}(\mathfrak{z},0)\Big], \quad a \leq \mathfrak{e} \leq b-1,$$

$$(32)$$

$$\mathfrak{z}^b(\zeta - m(\mathfrak{z}))\tilde{T}_b(\mathfrak{z},\zeta) = (\mathfrak{z}^b - \tilde{U}(\zeta))T_b(\mathfrak{z},0) - \tilde{U}(\zeta)\Big[\sum_{m=a}^{b-1} T_m(\mathfrak{z},0) + \mathfrak{z}^b\beta\tilde{\mathcal{B}}_b(\mathfrak{z},0)$$

$$+ \sum_{k=0}^{a-1}\sum_{\mathfrak{F}=b}^\infty I_k\mathfrak{z}^k g_{\mathfrak{F}-k}\mathfrak{z}^{\mathfrak{F}-k} - \sum_{\mathfrak{F}=0}^{b-1}\sum_{m=a}^b T_{m,\mathfrak{F}}(0)\mathfrak{z}^\mathfrak{F}\Big], \quad (33)$$

$$(\zeta - n(\mathfrak{z}))\tilde{L}_{\mathfrak{e}}(\mathfrak{z},\zeta) = L_{\mathfrak{e}}(\mathfrak{z},0) - \tilde{C}(\zeta)\Big[\sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0) + \sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0) + \delta I_{\mathfrak{e}}\Big], \quad 1 \le \mathfrak{e} \le a-1, \quad (34)$$

$$(\zeta - n(\mathfrak{z}))\tilde{Q}_1(\mathfrak{z},\zeta) = Q_1(\mathfrak{z},0) - \tilde{V}(\zeta)\Big[\sum_{m=a}^{b} \mathcal{B}_{m,0}(0) + \sum_{m=1}^{a-1} L_{m,0}(0) + \delta I_0\Big], \quad (35)$$

$$(\zeta - n(\mathfrak{z}))\tilde{Q}_l(\mathfrak{z},\zeta) = Q_l(\mathfrak{z},0) - \tilde{V}(\zeta)\sum_{\mathfrak{F}=0}^{a-1} Q_{l-1,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}}, \quad l \ge 2. \quad (36)$$

Where $l(\mathfrak{z}) = n(\mathfrak{z}) + \beta$, $m(\mathfrak{z}) = n(\mathfrak{z}) + \delta$, $n(\mathfrak{z}) = -\lambda X(\mathfrak{z}) + \lambda$
Substituting $\zeta = l(\mathfrak{z})$ in (30) and (31), $\zeta = m(\mathfrak{z})$ in (32) and (33), and $\zeta = n(\mathfrak{z})$ in (34)-(36), the corresponding results are obtained.

$$\mathcal{B}_{\mathfrak{e}}(\mathfrak{z},0) = \tilde{S}(l(\mathfrak{z}))\Big[\sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0) + \sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0) + \sum_{\mathfrak{F}=0}^{\infty} Q_{l,\mathfrak{F}}(0) + \delta\tilde{T}_{\mathfrak{e}}(\mathfrak{z},0)\Big], \quad a \le \mathfrak{e} \le b-1, \quad (37)$$

$$\Big(\mathfrak{z}^b - \tilde{S}(l(\mathfrak{z}))\Big)\mathcal{B}_b(\mathfrak{z},0) = \tilde{S}(l(\mathfrak{z}))\Big[\sum_{m=a}^{b} \mathcal{B}_m(\mathfrak{z},0) + \sum_{m=1}^{a-1} L_m(\mathfrak{z},0) + \sum_{l=1}^{\infty} Q_l(\mathfrak{z},0) + \mathfrak{z}^b\delta\tilde{T}_b(\mathfrak{z},0)$$
$$- \sum_{\mathfrak{F}=0}^{b-1}\Big(\sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{F}}(0) + \sum_{m=1}^{a-1} L_{m,\mathfrak{F}}(0) + \sum_{\mathfrak{F}=0}^{\infty} Q_{l,\mathfrak{F}}(0)\Big)\Big], \quad (38)$$

$$T_{\mathfrak{e}}(\mathfrak{z},0) = \tilde{U}(m(\mathfrak{z}))\Big[\sum_{m=a}^{b} T_{m,\mathfrak{e}}(0) + \sum_{k=0}^{a-1} I_k\lambda g_{\mathfrak{e}-k} + \beta\tilde{\mathcal{B}}_{\mathfrak{e}}(\mathfrak{z},0)\Big], \quad a \le \mathfrak{e} \le b-1, \quad (39)$$

$$\Big(\mathfrak{z}^b - \tilde{U}(m(\mathfrak{z}))\Big)T_b(\mathfrak{z},0) = \tilde{U}(m(\mathfrak{z}))\Big[\sum_{m=a}^{b-1} T_m(\mathfrak{z},0) + \mathfrak{z}^b\beta\tilde{\mathcal{B}}_b(\mathfrak{z},0)$$
$$+ \sum_{k=0}^{a-1}\sum_{\mathfrak{F}=b}^{\infty} I_k\mathfrak{z}^k g_{\mathfrak{F}-k}\mathfrak{z}^{\mathfrak{F}-k} - \sum_{\mathfrak{F}=0}^{b-1}\sum_{m=a}^{b} T_{m,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}}\Big] \quad (40)$$

$$L_{\mathfrak{e}}(\mathfrak{z},0) = \tilde{C}(n(\mathfrak{z}))\Big[\sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0) + \sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0) + \delta I_{\mathfrak{e}}\Big], \quad 1 \le \mathfrak{e} \le a-1, \quad (41)$$

$$Q_1(\mathfrak{z},0) = \tilde{V}(n(\mathfrak{z}))\Big[\sum_{m=a}^{b} \mathcal{B}_{m,0}(0) + \sum_{m=1}^{a-1} L_{m,0}(0) + \delta I_0\Big], \quad (42)$$

$$Q_l(\mathfrak{z},0) = \tilde{V}(n(\mathfrak{z}))\sum_{\mathfrak{F}=0}^{a-1} Q_{l-1,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}}, \quad l \ge 2. \quad (43)$$

Substitute Equations (37)-(43) in Equations (30)-(36) after simplication, and we get,

$$(\zeta - l(\mathfrak{z}))\tilde{B}_i(\mathfrak{z},\zeta) = (\tilde{S}(l(\mathfrak{z})) - \tilde{S}(\zeta))\Big[\sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{e}}(0) + \sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0) + \sum_{\mathfrak{F}=0}^{\infty} Q_{l,\mathfrak{F}}(0) + \delta\tilde{T}(\mathfrak{z},0)\Big],$$
$$a \le \mathfrak{e} \le b-1, \quad (44)$$

$$(\zeta - l(\mathfrak{z}))(\mathfrak{z}^b - \tilde{S}(l(\mathfrak{z})))\tilde{\mathcal{B}}_b(\mathfrak{z},\zeta) = (\tilde{S}(l(\mathfrak{z})) - \tilde{S}(\zeta))\Big[\sum_{m=a}^{b} \mathcal{B}_m(\mathfrak{z},0) + \sum_{m=1}^{a-1} L_m(\mathfrak{z},0) + \sum_{\mathfrak{F}=0}^{\infty} Q_l(\mathfrak{z},0)$$
$$+ \mathfrak{z}^b\delta\tilde{T}(\mathfrak{z},0) - \sum_{\mathfrak{F}=0}^{b-1}\Big(\sum_{m=a}^{b} \mathcal{B}_{m,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}} + \sum_{m=1}^{a-1} L_{m,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}} + \sum_{\mathfrak{F}=0}^{\infty} Q_{l,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}}\Big)\Big], \quad (45)$$

$$(\zeta - m(\mathfrak{z}))\tilde{T}_{\mathfrak{e}}(\mathfrak{z},\zeta) = (\tilde{U}(m(\mathfrak{z})) - \tilde{U}(\zeta))\Big[\sum_{m=a}^{b} T_{m,\mathfrak{e}}(0) + \sum_{k=0}^{a-1} I_k\lambda g_{\mathfrak{e}-k} + \beta\tilde{\mathcal{B}}_{\mathfrak{e}}(\mathfrak{z},0)\Big], \quad a \le \mathfrak{e} \le b-1, \quad (46)$$

$$(\zeta - m(\mathfrak{z}))(\mathfrak{z}^b - \tilde{U}(m(\mathfrak{z})))\tilde{T}_b(\mathfrak{z},\zeta) = (\tilde{U}(m(\mathfrak{z})) - \tilde{U}(\zeta))\Big[\sum_{m=a}^{b} T_m(\mathfrak{z},0) + \mathfrak{z}^b \beta \tilde{\mathcal{B}}_b(\mathfrak{z},0)$$

$$+ \sum_{k=0}^{a-1}\sum_{\mathfrak{F}=b}^{\infty} I_k \mathfrak{z}^k g_{\mathfrak{F}-k}\mathfrak{z}^{\mathfrak{F}-k} - \sum_{\mathfrak{F}=0}^{b-1}\sum_{m=a}^{b} T_{m,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}}\Big], \quad (47)$$

$$(\zeta - n(\mathfrak{z}))\tilde{L}_{\mathfrak{e}}(\mathfrak{z},\zeta) = (\tilde{C}(n(\mathfrak{z})) - \tilde{C}(\zeta))\Big[\sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0) + \sum_{m=a}^{b}\mathcal{B}_{m,\mathfrak{e}}(0) + \delta I_{\mathfrak{e}}\Big], \quad 1 \le \mathfrak{e} \le a-1, \quad (48)$$

$$(\zeta - n(\mathfrak{z}))\tilde{Q}_1(\mathfrak{z},\zeta) = (\tilde{V}(n(\mathfrak{z})) - \tilde{V}(\zeta))\Big[\sum_{m=a}^{b} B_{m,0}(0) + \sum_{m=1}^{a-1} L_{m,0}(0) + \delta I_0\Big], \quad (49)$$

$$(\zeta - n(\mathfrak{z}))\tilde{Q}_l(\mathfrak{z},\zeta) = (\tilde{V}(n(\mathfrak{z})) - \tilde{V}(\zeta))\sum_{\mathfrak{F}=0}^{a-1} Q_{l-1,\mathfrak{F}}(0)\mathfrak{z}^{\mathfrak{F}}, \quad l \ge 2. \quad (50)$$

In summary,the probability generating function associated with the queue size is

$$P(\mathfrak{z}) = \sum_{\mathfrak{e}=a}^{b}\tilde{\mathcal{B}}_{\mathfrak{e}}(\mathfrak{z},0) + \sum_{\mathfrak{e}=a}^{b}\tilde{T}_{\mathfrak{e}}(\mathfrak{z},0) + \sum_{\mathfrak{e}=1}^{a-1}\tilde{L}_{\mathfrak{e}}(\mathfrak{z},0) + \sum_{\mathfrak{F}=1}^{\infty}\tilde{Q}_{\mathfrak{F}}(\mathfrak{z},0) + I(\mathfrak{z}). \quad (51)$$

Substituting $\zeta = 0$ in Equations (44)-(50), Equation (51) simplifies to:

$$P(\mathfrak{z}) = \frac{\left\{\begin{array}{l} A_1(\mathfrak{z})\sum_{\mathfrak{e}=a}^{b-1}[\mathfrak{z}^b - \mathfrak{z}^{\mathfrak{e}}]r_{\mathfrak{e}} + A_2(\mathfrak{z})\sum_{\mathfrak{e}=a}^{b-1}[\mathfrak{z}^b - \mathfrak{z}^{\mathfrak{e}}]c_{\mathfrak{e}} \\[2mm] + (1-\tilde{V}(n(\mathfrak{z})))A_3(\mathfrak{z})\sum_{\mathfrak{e}=0}^{a-1} q_{\mathfrak{e}}\mathfrak{z}^{\mathfrak{e}} - A_1(\mathfrak{z})\sum_{\mathfrak{e}=0}^{a-1} s_{\mathfrak{e}}\mathfrak{z}^{\mathfrak{e}} \\[2mm] + [B_1(\mathfrak{z}) - \tilde{C}(n(\mathfrak{z}))A_3(\mathfrak{z})]\sum_{\mathfrak{e}=1}^{a-1}[s_{\mathfrak{e}} + \delta I_{\mathfrak{e}}] \\[2mm] + [B_1(\mathfrak{z}) - \tilde{V}(n(\mathfrak{z}))A_3(\mathfrak{z})][s_0 + \delta I_0] \\[2mm] + [n(\mathfrak{z})B_1(\mathfrak{z}) - m(\mathfrak{z})A_2(\mathfrak{z})]\sum_{\mathfrak{e}=0}^{a-1} I_{\mathfrak{e}}\mathfrak{z}^{\mathfrak{e}} \end{array}\right\}}{n(\mathfrak{z})B_1(\mathfrak{z})} \quad (52)$$

where $p_{\mathfrak{e}} = \sum_{m=a}^{b} B_{m,\mathfrak{e}}(0), l_{\mathfrak{e}} = \sum_{m=1}^{a-1} L_{m,\mathfrak{e}}(0), q_{\mathfrak{e}} = \sum_{l=1}^{\infty} Q_{l,\mathfrak{e}}(0), v_{\mathfrak{e}} = \sum_{m=a}^{b} T_{m,\mathfrak{e}}(0), r_{\mathfrak{e}} = p_{\mathfrak{e}} + l_{\mathfrak{e}} + q_{\mathfrak{e}},$

$s_{\mathfrak{e}} = p_{\mathfrak{e}} + l_{\mathfrak{e}}$ and $c_{\mathfrak{e}} = v_{\mathfrak{e}} + \sum_{k=0}^{a-1} l_k \lambda g_{\mathfrak{e}-k}.$

Appendix A provides the expressions for $B_1(\mathfrak{z}), A_1(\mathfrak{z}), A_2(\mathfrak{z})$ and $A_3(\mathfrak{z})$.

3.1. Computational Aspects

Equation (52) involves a total of $2b + a$ unknowns, specifically the sets $\{r_a, r_{a+1}, \ldots, r_{b-1}\}$, $\{c_a, c_{a+1}, \ldots, c_{b-1}\}$, $\{s_0, s_1, \ldots, s_{a-1}\}$, $\{q_0, q_1, \ldots, q_{a-1}\}$, and $\{I_0, I_1, \ldots, I_{a-1}\}$. Theorem(1) establishes that each $q_{\mathfrak{e}}$ can be expressed as a function of $S_0 + \delta I_0$, with the numerator containing only $2b$ constants. Using Rouche's theorem, it is shown that the function $B_1(\mathfrak{z})$ possesses $2b - 1$ zeros strictly inside the unit circle and one zero exactly on the unit circle, i.e., $|\mathfrak{z}| = 1$. Given that $P(\mathfrak{z})$ is analytic within and on the unit circle, this analyticity condition leads to a system of $2b$ equations in $2b$ unknowns. To satisfy these conditions, the numerator of the function must vanish at the corresponding points. These resulting equations can be solved using the Gauss elimination method.

Theorem 1. Let q_{κ}, for $(0 \le \kappa \le a - 1)$, be expressed in terms of $(s_0 + \delta I_0)$ as follows:

$$q_n = \frac{\left(\beta_n(s_0 + \delta I_0) + \sum_{\kappa=0}^{n-1} q_{\kappa}\beta_{n-\kappa}\right)}{1 - \beta_0}, \quad n = 1, 2, \ldots, a-1; \qquad q_0 = \frac{\beta_0}{1 - \beta_0}(s_0 + \delta I_0)$$

where β_κ is the probability that κ customers arriving while on vacation.

3.2. Stability Analysis

The queue length's probability generating function (PGF) must meet $P(1) = 1$ in order to meet the normalization criteria. To ensure this, L' Hopital's rule is applied to evaluate the limit $\lim\limits_{\mathfrak{z}\to 1} P(z)$, which yields $\mathbb{G} = \pounds_{12}$. It follows that $\mathbb{G}$ must be positive, and hence, the condition $P(1) = 1$ is satisfied if and only if $\pounds_{12} > 0$. Furthermore, the traffic intensity ρ is given by

$$\rho = \frac{\lambda X_1(\beta + \delta)(1 - \tilde{S}(\beta))(1 - \tilde{U}(\delta))}{b\beta\delta[\tilde{S}(\beta)(1 - \tilde{U}(\delta)) + \tilde{U}(\delta)(1 - \tilde{S}(\beta))]}. \tag{53}$$

The condition $\rho < 1$ must hold for stability.

3.3. Special Case

When $a = 1$, $b = 1$ and no breakdown, then equation (52) reduces to,

$$P(\mathfrak{z}) = \frac{[\tilde{V}(n(\mathfrak{z})) - 1](1 - \mathfrak{z})[q_0 + s_0]}{n(\mathfrak{z})(\mathfrak{z} - \tilde{S}(n(\mathfrak{z})))} \tag{54}$$

The equation provides the queue size distribution for a queueing system with multiple vacations of $M^{[X]}/G/1$. With N=1, the outcome coincides with the queue size distribution of Lee et al. [14].

4. Performance metrics

4.1. Expected queue length

The mean queue size $E(\mathfrak{Q})$ at any given moment can be determined through evaluating the differential of $P(\mathfrak{z})$ with respect to $\mathfrak{z}$.

$$E(Q) = \frac{\left\{\begin{array}{l} \Im_1 \sum\limits_{\mathfrak{e}=a}^{b-1} r_\mathfrak{e}[-\mathfrak{e}(\mathfrak{e}-1) + b(b-1)] + \Im_2 \sum\limits_{\mathfrak{e}=a}^{b-1} r_\mathfrak{e}(-\mathfrak{e}+b) \\ + \Im_1 \sum\limits_{\mathfrak{e}=a}^{b-1} c_\mathfrak{e}[-\mathfrak{e}(\mathfrak{e}-1) + b(b-1)] + \Im_3 \sum\limits_{\mathfrak{e}=a}^{b-1} c_\mathfrak{e}(-\mathfrak{e}+b) \\ + \Im_4 \sum\limits_{\mathfrak{e}=0}^{a-1} \mathfrak{e}q_\mathfrak{e} + \Im_5 \sum\limits_{\mathfrak{e}=0}^{a-1} q_\mathfrak{e} + \Im_6 \sum\limits_{\mathfrak{e}=1}^{a-1} (s_\mathfrak{e} + \delta I_\mathfrak{e}) \\ + \Im_7(s_0 + \delta I_0) + \Im_8 \sum\limits_{\mathfrak{e}=0}^{a-1} I_\mathfrak{e} + \Im_9 \sum\limits_{k=0}^{a-1} \mathfrak{e}I_\mathfrak{e} - \Im_{10} \sum\limits_{\mathfrak{e}=0}^{a-1} s_\mathfrak{e} \end{array}\right\}}{3[\pounds_{12}]^2} \tag{55}$$

Appendix B provides the expressions for $\Im_1$ to $\Im_{10}$.

4.2. Expected waiting time

According to Little's formula, $E(Q)$ can be determined as follows:

$$E(Q) = \frac{E(Q)}{\lambda E(X)} \tag{56}$$

where $E(\mathfrak{Q})$ is given in Eq.(55)

5. NUMERICAL EXAMPLE

The proposed model is numerically demonstrated using the following assumptions. Arrivals occur in batches, with the batch size following a geometric distribution with mean 2. Service times (main server, standby server, and low-batch service) follow an Erlang-2 distribution with parameters m_1, m_2, and m_3 respectively. With a rate of $v = 5$, vacation time distributes exponentially. Tables 2 - 5 show the results of calculating the expected queue length, and the expected waiting time for different arrival and service rates.

Based on these tables, the following can be observed:

- Table 2 indicates that an increase in the arrival rate results in increases expected queue length
- Tables 3 and 4 indicate that an increase in the service rates of the main server and standby server results in a decrease in both the expected queue length and waiting time
- Table 5 indicates that when the service rate is zero, the expected queue length and waiting time are high. As the service rate of the low-batch service gradually increases, there is a noticeable decrease in both the expected queue length and waiting time compared to the case with zero service rate

Furthermore, the model compares the results produced by the Adaptive Neuro-Fuzzy Inference System to those acquired using the SVT/PGF analytical approach. To assess performance with the ANFIS model, five linguistic values were assigned to input parameters, interpreted as fuzzy variables, and trained across five learning epochs.

Each input variable is connected with a Gaussian membership function that represents the following linguistic categories: low, moderate, medium, high, and extremely high. These functions determine the degree to which each input value corresponds to a specific linguistic phrase.

Tables 2 - 5 and Figures 2 - 5 show that the difference in results between the ANFIS and the analytical is negligible, with both graphs almost identical.

Table 2: *Performance Metrics vs. (λ) for $a = 5$, $b = 8$, $m_1 = 15$, $m_2 = 14$, $m_3 = 10$, $\beta = 1$, $v = 5$, and $\delta = 2$.*

λ	ρ	$E(Q)$		$E(W)$	
		Analytical	ANFIS	Analytical	ANFIS
5.0	0.0872	3.9394	3.9963	0.3939	0.3971
5.5	0.0959	4.6953	4.6052	0.4269	0.4225
6.0	0.1046	5.6039	5.6231	0.4670	0.4648
6.5	0.1133	6.4456	6.4444	0.4958	0.4995
7.0	0.1220	8.0544	8.1043	0.5753	0.5762
7.5	0.1307	9.7471	9.7077	0.6498	0.6488
8.0	0.1395	11.9231	11.7800	0.7452	0.7384
8.5	0.1482	14.7864	15.2189	0.8698	0.8886
9.0	0.1569	18.7363	18.2796	1.0419	1.0220
9.5	0.1656	24.4725	24.7282	1.2880	1.2995
10.0	0.1743	33.5637	33.4798	1.6782	1.6744

Table 3: *Performance Metrics vs. (m_1) for $a = 5$, $b = 8$, $\lambda = 5$, $m_2 = 7$, $m_3 = 5$, $\beta = 1$, $v = 5$, and $\delta = 2$.*

m_1	ρ	$E(Q)$		$E(W)$	
		Analytical	ANFIS	Analytical	ANFIS
7.5	0.3563	80.6337	79.6274	8.0634	8.0544
8.0	0.3402	53.9373	53.2782	5.3937	5.4239
8.5	0.3254	41.2900	40.5498	4.1290	4.0568
9.0	0.3119	34.2405	34.2952	3.4241	3.5145
9.5	0.2993	29.5711	29.6083	2.9571	2.9096
10.0	0.2880	26.3219	26.6564	2.6322	2.6465
10.5	0.2773	23.9263	24.0915	2.3926	2.3766
11.0	0.2674	22.0967	21.8730	2.2097	2.2227
11.5	0.2582	20.6582	20.6882	2.0658	2.0606
12.0	0.2496	19.5012	19.5090	1.9501	1.9520

Table 4: *Performance Metrics vs. (m_2) for $a = 5$, $b = 8$, $\lambda = 4$, $m_1 = 12$, $m_3 = 5$, $\beta = 1$, $v = 5$, and $\delta = 2$.*

m_2	ρ	$E(Q)$		$E(W)$	
		Analytical	ANFIS	Analytical	ANFIS
6.0	0.2582	62.0072	61.9598	7.7509	7.7450
6.5	0.2538	56.4153	56.5322	7.0519	7.0665
7.0	0.2496	50.8083	50.6777	6.3510	6.3347
7.5	0.2456	45.3206	45.4477	5.6651	5.6810
8.0	0.2416	40.0282	39.8810	5.0035	4.9851
8.5	0.2378	34.9704	35.1150	4.3713	4.3894
9.0	0.2341	30.1625	30.0399	3.7703	3.7550
9.5	0.2305	25.6054	25.7381	3.2007	3.2173
10.0	0.2270	21.2924	21.1629	2.6616	2.6454
10.5	0.2237	17.2121	17.2679	2.1515	2.1585

Table 5: *Performance Metrics vs. (m_3) for $a = 5$, $b = 8$, $\lambda = 3$, $m_1 = 12$, $m_2 = 10$, $\beta = 1$, $v = 5$, and $\delta = 2$.*

m_3	$E(Q)$		$E(W)$	
	Analytical	ANFIS	Analytical	ANFIS
0	13.4068	13.4065	2.2345	2.2345
1	11.7402	11.7410	1.9567	1.9568
2	10.9469	10.9502	1.8245	1.8251
3	10.5581	10.5422	1.7597	1.7570
4	10.3446	10.3658	1.7241	1.7276
5	10.1874	10.1391	1.6979	1.6898
6	9.4319	9.5542	1.5719	1.5923
7	9.3811	9.2979	1.5635	1.5496
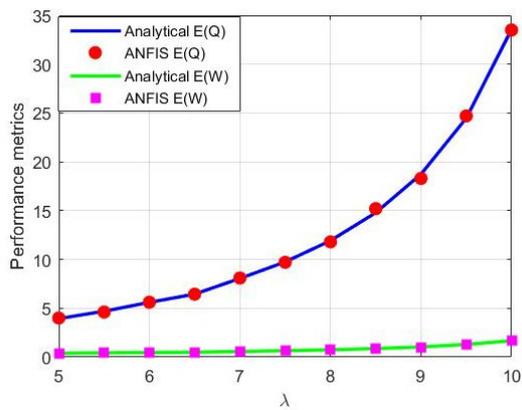
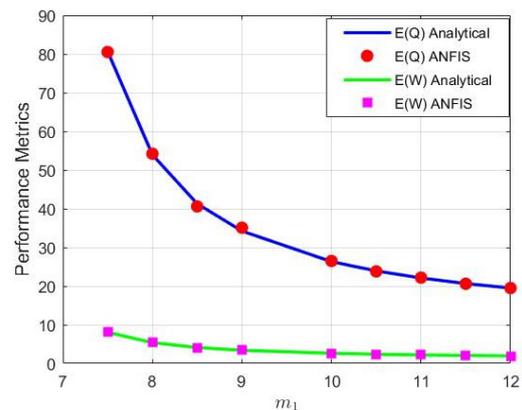**Figure 2:** *Arrival rate (λ) vs E(Q) and E(W)*

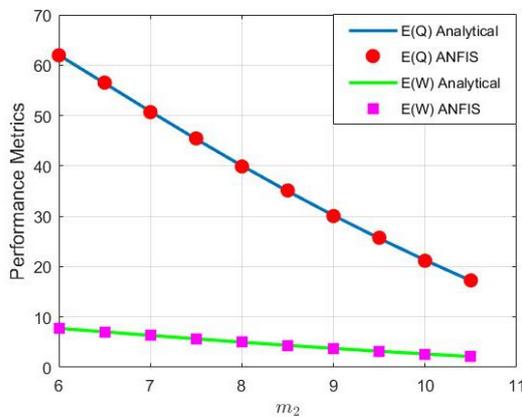**Figure 3:** *Service rate (m_1) vs E(Q) and E(W)*

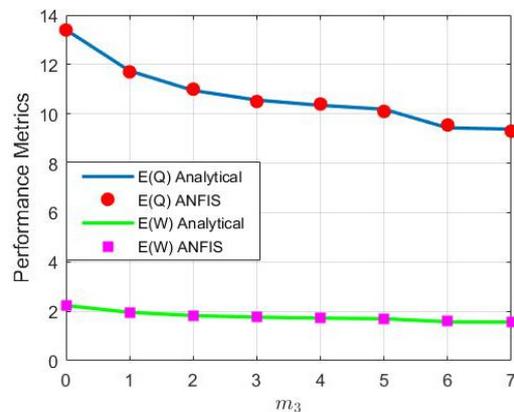**Figure 4:** *Service rate (m_2) vs E(Q) and E(W)*

Figure 5: *Service rate (m_3) vs E(Q) and E(W)*

6. Conclusion and future work

In summary, this study describes a novel way to merge low-batch and bulk services that includes a standby server and multiple vacation mechanisms. This combination is extremely useful for simulating real-world environments, especially in cloud systems and manufacturing and production systems. Table 5 shows that low-batch service significantly reduces queue length and customers waiting time. Moreover, when the analytical results are compared to the ANFIS-based predictions, the suggested model shows excellent agreement, indicating that the ANFIS technique is accurate and robust. Currently, the standby server supports only the bulk service; in future work, it could be extended to support low-batch service as well, and its impact on performance metrics could be further evaluated.

References

[1] Ayyappan, G. and Deepa, T. (2019). Analysis of batch arrival bulk service queue with additional optional service multiple vacation and setup time. *International Journal of Mathematics in Operational Research*, 15(1):1–25.

[2] Ayyappan, G. and Karpagam, S. (2018). An $M^{[X]}/G(a,b)/1$ queueing system with breakdown and repair, stand-by server, multiple vacation and control policy on request for re-service. *Mathematics*, 6(6),101.

[3] Ayyappan, G. Gurulakshmi, G. A. and Somasundaram, B.(2022). Analysis of MAP/PH/1 queueing model with multiple vacations, optional service, close-down, setup, breakdown, phase type repair and impatient customers *Reliability: Theory & Applications*, 17(68):447–468

[4] Chang, F. M. Lee, Y. T. Chang, C. J. and Yeh, C. (2018). Analysis of a standby redundant system with controlled arrival of failed machines. *International Journal of Industrial and Systems Engineering*, 28(1):117–134.

[5] Chakravarthy, S. R. and Kulshrestha, R. (2020). A queueing model with server breakdowns, repairs, vacations, and backup server. *Operations Research Perspectives*, 7, 100131.

[6] Choudhury, G. and Madan, K. C. (2005). A two-stage batch arrival queueing system with a modified Bernoulli schedule vacation under N-policy. *Mathematical and Computer Modelling*, 42(1-2):71–85.

[7] Divya, K. and Indhira, K. (2023).Cost-revenue analysis and ANFIS computing of heterogeneous queuing model with a second optional service with feedback under hybrid vacation *Reliability: Theory & Applications*, 18(74):176-195.

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

[9] Jeyakumar, S. and Senthilnathan, B. (2017). Modelling and analysis of a bulk service queueing model with multiple working vacations and server breakdown. *RAIRO-Operations Research*, 51:485–508.

[10] Karpagam, S.and Lokesh, R. (2025). Steady state analysis of bulk arrival, fixed Batch service queue with inspection, rework, and multiple vacations. *Reliability: Theory & Applications*, 20(84):577–589.

[11] Khalaf, R. F. Madan, K. C. and Lukas, C. A. (2011). On a Batch Arrival Queuing System Equipped with a Stand-by Server during Vacation Periods or the Repairs Times of the Main Server. *Journal of Probability and Statistics*, 2011(1),812726.

[12] Khalaf, R. F. Madan, K. C. and Lukas, C. A. (2012). On an $M^{[X]}/G/1$ queuing system with random breakdowns, server vacations, delay times and a standby server. *International Journal of Operational Research*, 15:30–47.

[13] Lavanya, B. Vennila, R. and Sankoh, M. (2022). Mathematical Modelling of $M^{[X]}/G(a,b)/1$ Bulk Service Queue Model with Two Vacations and Setup Time in Ceramic Technology. *Mathematical Problems in Engineering*, 2022(1),2771494.

[14] Lee, H. W. Lee, S. S. Park, J. O. and Chae, K. C. (1994). Analysis of the $M^{[X]}/G/1$ Queue by N-Policy and Multiple Vacations. *Journal of Applied Probability*, 31(2):476–496.

[15] Levy, Y. and Yechiali, U. (1975). Utilization of idle time in an $M/G/1$ queue with server vacations. *Management Science*, 22:202–211.

[16] Murugeswari, N.and Sundari, S. M. (2017). A standby server bulk arrival queuing model of compulsory server vacation. *International Journal of Engineering Development and Research*, 5(4):337–341.

[17] Neuts, M. F. (1967). A general class of bulk queues with Poisson input. *The Annals of Mathematical Statistics*, 38:759–770.

[18] Pradhan, S. Nandy, N. and Gupta, U. C. (2024). Performance analysis of a versatile bulk-service queue with group-arrival, batch-size-dependent service time and queue-length-dependent vacation. *Quality Technology & Quantitative Management*, 22(4): 683–726.

[19] Rathinasamy, P. and Kirupa, K. (2025). Optimizing Wait Times: Enhancing Server Accessibility with Two-Phase Imperative Services and Backup Server. *Calcutta Statistical Association Bulletin*, 77(1):91–116.

[20] Reddy, G. V. K. Nadarajan, R. and Arumuganathan, R. (1998). Analysis of a bulk queue with N policy multiple vacations and setup times. *Computers & Operations Research*, 25:957–967.

[21] Senthilnathan, B. and Jeyakumar, S. (2012). A study on the behaviour of the server breakdown without interruption in a $M^{[X]}/G(a,b)/1$ queueing system with multiple vacations and closedown time. *Applied Mathematics and Computation*, 219:2618–2633.

[22] Srivastava, R. K. Singh, S. and Singh, A. (2020). Bulk arrival Markovian queueing system with two types of services and multiple vacations. *International Journal of Mathematics and Computer Research*, 8(8):2130–2136.

[23] Takagi, H. (1991). Vacation and Priority Systems. In: *Queueing Analysis: A Foundation of Performance Evaluation*, Volume I. North-Holland, Amsterdam.

[24] Tamrakar, G. K. Banerjee, A., and Gupta, U. C. (2022). Analysis of batch size-dependent bulk service queue with multiple working vacation. *International Journal of Computer Mathematics: Computer Systems Theory*, 7(3):149–171.

[25] Thakur, S.Jain, A. and Jain, M. (2021). ANFIS and cost optimization for Markovian queue with operational vacation. *Int. J. Math. Eng. Management Sci*, 6(3):894–910.

Appendix A

The following expressions are employed in Eq.(52):

$$B_1(z) = l(z)m(z)(z^b - \tilde{S}(l(z)))(z^b - U(m(z))) - z^{2b}\beta\delta(1 - \tilde{S}(l(z)))(1 - \tilde{U}(m(z))),$$

$$A_1(z) = n(z)(1 - \tilde{S}(l(z)))\Big(m(z)(z^b - \tilde{U}(m(z))) + z^b\beta(1 - \tilde{U}(m(z)))\Big),$$

$$A_2(z) = (1 - \tilde{U}(m(z)))n(z)\Big(l(z)(z^b - \tilde{S}(l(z))) + z^b\delta(1 - \tilde{S}(l(z)))\Big),$$

$$A_3(z) = B_1(z) - A_1(z).$$

Appendix B

The following expressions are employed in Eq.(55):

$$\Im_1 = 3\aleph_1\pounds_{12},$$
$$\Im_2 = 3\pounds_7\pounds_{12} - 2\aleph_1\pounds_{13},$$
$$\Im_3 = 3\pounds_9\pounds_{12} - 2\aleph_1\pounds_{13},$$
$$\Im_4 = 6b\beta\delta V_1\aleph_2\pounds_{12},$$
$$\Im_5 = 3\pounds_{12}[V_1(\pounds_7 - \pounds_2) + b\beta\delta\aleph_2 V_2] - 2b\beta\delta\aleph_2 V_1\pounds_{13},$$
$$\Im_6 = 3\pounds_{12}[V_1(\pounds_7 - \pounds_2) + b\beta\delta\aleph_2 C_2] - 2b\beta\delta\aleph_2 C_1\pounds_{13},$$
$$\Im_7 = 3\pounds_{12}[V_1(\pounds_7 - \pounds_2) + b\beta\delta\aleph_2 V_2] - 2b\beta\delta\aleph_2 V_1\pounds_{13},$$
$$\Im_8 = 3\aleph_8\pounds_{12} - [2\lambda X_1 b\beta\delta\aleph_2 - \delta\pounds_9]\pounds_{13},$$
$$\Im_9 = 3\aleph_9\pounds_{12} - 2\delta\aleph_1\pounds_{13},$$
$$\Im_{10} = 3[\pounds_7 + \mathfrak{e}(\mathfrak{e} - 1)\aleph_1]\pounds_{12} - 2\aleph_1\pounds_{13}.$$

where:

$$\aleph_1 = -\lambda X_1(\beta + \delta)\left(1 - \tilde{S}(\beta)\right)\left(1 - \tilde{U}(\delta)\right),$$
$$\aleph_2 = \tilde{S}(\beta)\left(\tilde{U}(\delta) - 1\right) + \tilde{U}(\delta)\left(\tilde{S}(\beta) - 1\right),$$
$$\aleph_3 = \tilde{S}_1\left(1 - \tilde{U}(\delta)\right) - U_1\left(1 - \tilde{S}(\beta)\right),$$
$$\aleph_4 = -\lambda X_1\left(1 - \tilde{U}(\delta)\right),$$
$$\aleph_5 = -\left(1 - \tilde{U}(\delta)\right)\lambda X_2 + 2\lambda X_1 U_1,$$
$$\aleph_6 = -\left(1 - \tilde{S}(\beta)\right)\lambda X_1,$$
$$\aleph_7 = -\left(1 - \tilde{S}(\beta)\right)\lambda X_2 + 2\lambda X_1 S_1,$$
$$\aleph_8 = b\beta\delta\aleph_2\lambda X_2 - \lambda X_1\pounds_2 - (\delta/3)\pounds_{10} + \lambda X_1\pounds_9,$$
$$\aleph_9 = 2\lambda X_2 b\beta\delta\aleph_2 - \delta\pounds_9 - \delta\aleph_1,$$

$$\pounds_1 = \aleph_1 - b\beta\delta\aleph_2,$$

$$\pounds_2 = \left(1 - \tilde{S}(\beta)\right)\left(1 - \tilde{U}(\delta)\right)\left[2(\lambda X_1)^2 - \lambda X_2(\beta + \delta) - 2\beta\delta b(b-1)X_1\right]$$
$$+ \left[\beta\delta b(b-1) - 2b\lambda X_1(\beta + \delta)\right]$$
$$+ \left(1 - \tilde{U}(\delta)\right) + \left(1 - \tilde{S}(\beta)\right)$$
$$+ 2\left[2b\beta\delta + \lambda X_1(\beta + \delta)\right]\aleph_3 + 2\beta\delta b\left(b - S_1 - U_1\right),$$

$$\pounds_3 = (\beta + \delta)(b - U_1) + \aleph_4 - b\beta\tilde{U}(\delta),$$

$$\pounds_4 = (\beta + \delta)(-U_2 + b(b-1)) - b\beta\left[(b-1)\tilde{U}(\delta) + 2U_1\right] - 2b\lambda X_1 + \aleph_5,$$

$$\pounds_5 = (\beta + \delta)(b - S_1) + \aleph_6 - b\delta\tilde{S}(\beta),$$

$$\pounds_6 = (\beta + \delta)(b(b-1) - S_2) - b\delta\left[\tilde{S}(\beta)(b-1) + 2S_1\right] + \lambda(-2b)X_1 + \aleph_7,$$

$$\pounds_7 = \aleph_7(\beta + \delta)(1 - U_1) + 2\aleph_6\pounds_3,$$

$$\pounds_8 = (\beta + \delta)\aleph_6(1 - \tilde{U}(\delta)) + 3\pounds_3\aleph_7 + 3\aleph_6\pounds_4,$$

$$\pounds_9 = \aleph_5(\beta + \delta)(1 - S_1) + 2\aleph_4\pounds_5,$$

$$\pounds_{10} = (\beta + \delta)\aleph_6(1 - \tilde{S}(\beta)) + 3\pounds_5\aleph_5 + 3\aleph_4\pounds_6,$$

$$\pounds_{11} = \delta\pounds_7 - \delta\pounds_9 + 2\lambda X_1 b\beta\delta\aleph_2,$$

$$\pounds_{12} = -2\lambda X_1\pounds_1,$$

$$\pounds_{13} = -3\left[\lambda X_1\pounds_2 + \lambda X_2\pounds_1\right],$$

$$\mathbb{E} = 2\aleph_1\sum_{\mathfrak{e}=a}^{b-1}r_{\mathfrak{e}}(b - \mathfrak{e}) + 2\aleph_1\sum_{\mathfrak{e}=a}^{b-1}C_{\mathfrak{e}}(b - \mathfrak{e}) + 2b\beta\delta\aleph_2 V_1\sum_{n=0}^{a-1}q_n + 2b\beta\delta\aleph_2 C_1\sum_{\mathfrak{e}=1}^{a-1}(s_{\mathfrak{e}} + \delta I_{\mathfrak{e}})$$
$$+ 2b\beta\delta\aleph_2 V_1(s_0 + \delta I_0) + \aleph_{11}\sum_{\mathfrak{e}=0}^{a-1}l_{\mathfrak{e}} - 2\aleph_1\left(\delta\sum_{\mathfrak{e}=0}^{a-1}\mathfrak{e}l_{\mathfrak{e}} + \sum_{\mathfrak{e}=0}^{a-1}\mathfrak{e}s_{\mathfrak{e}}\right).$$