https://doi.org/10.48047/AFJBS.8.06.2026.64-69

AFJBS African Journal of Biological Sciences
Journal homepage: http://www.afjbs.com

Research Paper Open Access

Queueing-Model based Optimal Threshold Recruitment in IT Markov Manpower Sectors

[1] R. Sivasamy, [2] Fatai Adewole Adebayo, [3] Mokgweetsi B, [4] Molefe, W. B
[1,2,3,4] *Department of Statistics, Demography and Population Science, Faculty of Natural and Applied Science, School of Computer, Mathematics and Statistics Science, University of Botswana*
[1] rssamy@gmail.com, ,[2] adebayof@ub.ac.bw , 3] mokgweetsib@ub.ac.bw, [4] molefewb@ub.ac.bw

Volume 8, Issue 6, June 2026
Received: 22 April 2026
Accepted: 05 June 2026
Published: 25 June 2026
doi:10.48047/AFJBS.8.6.2026.64-69

Abstract— Information technology (IT) or health care (HC) service organizations increasingly rely on dynamic staffing policies to balance fluctuating Information technology (IT) and health care (HC) service organizations often depend on versatile staffing methods to handle altering workloads, teamwork challenges, and the uprising costs of employee turnover. Simple manpower models mostly assume linear attrition rates and boundless hiring. Hence, such hypotheses do not consider either the congestion effects or the capacity limits witnessed in modern IT infrastructure management. To address this congestion and recruitment issues, this paper designs a finite-capacity Markov manpower model in which recruitment initiates only when the workforce level n drops below a certain level, $n < L$. Attrition increases non-linearly when staffing goes beyond this level L. The sequence, representing number of employees in the system at time t, becomes a state-dependent birth and death process, where attrition is affected by congestion parameter α, and recruitment restricted by capacity. We obtain the stationary distribution depending on α and formulate a cost function balancing recruitment and turnover costs, and inadequacies caused by α. Illustrations show there is an optimal recruitment threshold, $L*$, that minimizes long-term total expected costs. Further examples tailored to IT industry environments explore how recruitment-based thresholds can significantly reduce manpower volatility and working congestion. This article provides a transparent and relevant foundation for building effective recruitment policies in IT manpower systems.

Index Terms— Attrition rate, Congestion sensitive, Manpower level, Optimal cost, Recruitment policy,

I. INTRODUCTION

Modern essential services of IT wings and healthcare hospitals operate in environments characterized by unpredictable demands, evolving skill requirements, and rigid service level prospects. In such situations, manpower levels must be well managed by operational procedures through optimal hiring rules or recruitment policies. For instance, empirical evidence from IT support teams, software maintenance units, and cloud operations units could inform about how excessive staffing creates coordination overhead and the managerial congestion, whereas under staffing produces service degradation, burnout, and raised turnover. Such opposing pressures obviously motivate the adoption of threshold-based hiring policies, wherein recruitment is activated only when the workforce drops below a critical operational level. The work of [1] develops linear birth–death processes in random environments, obtaining conditions under which environmental volatility alters transient and asymptotic population interactions. [2] introduce a Markovian manpower-forecasting structure that splits the workforce into subgroups to yield tractable transition activities for HR (human resource) modelling. Authors of [3] study a Markovian Arrival Process (MAP)-based queueing system recruiting served customers as temporary secondary servers and obtain the steady-state characteristics relating to a threshold-based manpower control. In [4], authors investigate heterogeneous-server queues and discuss optimal threshold policies using a typical heuristic optimization integrated with neural-network-based performances approximately. [5] discusses a meticulous analysis of stochastic queueing models such as birth–death frameworks, steady-state characteristics, and core performance measures to manpower systems. [6] discusses foundational Markov-chain theory, highlighting best tools for classifying irreducible, recurrence, and stationary properties of random processes applied in manpower modelling. [7] investigate birth–death models in interactive random environments, computing in-variant measures and convergence rates via coupling for systems with state-dependent rates. [10] studies workforce planning that concerns with estimating the number of employees and their skills necessary to meet the future operational needs of an organization. [11] studies a manpower system that monitors wastages admitting recruitment when the cumulative loss of man hours crosses a breakdown threshold level. Authors of [12] investigate a mathematical model for employee replacement and promotion policies in manpower systems focusing on preventive planned promotions or retirements, and corrective replacement, which addresses un-planned exits like resignations. [13] discuss a queueing-based recruitment method for time-dependent dual-class waiting systems, deducing powerful recruiting rules under non-stationary demand. Contributions by [14] investigate a threshold epidemic model SIRS (Susceptible → Infectious → Recovered → Susceptible) with logistic birth and nonlinear incidence, deriving threshold for extinction versus persistence and stability conditions using threshold-based recruitment or attrition mechanisms.

Numerous investigations in manpower planning so far cited highlight the impacts of having both linear and non-linear employee attrition and recruitment. In many operational environments—specifically in IT and service-intensive industries, the intensity of workload can escalate speedily, causing stress, fatigue, and dissatisfaction to increase in a non-proportional way. Further, multitasking, response duties, and deadline-based operations amplify these congestion effects, making attrition highly sensitive to the prevailing workload. This motivates authors of this paper to con-sider congestion-sensitive attrition technique into the proposed Markov manpower model so that recruitment strategies remain both cost efficient and operationally realistic, especially under conditions where workforce overload can trigger accelerated departures.

In this work, we propose a ’Queueing-Model based Optimal Threshold Recruitment in IT based Manpower Sectors’ where employment process is activated only if the workforce level goes below a prefixed threshold. Further this model secures cost trade-offs and furnishes actionable tools for policy strategy.

For situations, when staffing falls below a crucial level L, proactive recruitment is triggered to avoid project holdups and cost escalations. However, when the team capacity exceeds L, coordination overhead, task interdependence, and communication delays increase steadily, forcing elevated burnout and voluntary staff replacement. Such occurrences could be well accounted by introducing a congestion sensitive parameter $\alpha > 0$, if it enhances attrition as the workforce level transits beyond the feasible threshold.

A. *Threshold Recruitment in IT Firms*

We propose a threshold-based manpower model to obtain useful insights for large Indian IT service companies like HCL Technologies and Infosys. These organizations depend on a stable group of software engineers, project managers, and domain specialists to run optimally. They operate in a fast-changing environment with random project inflows, varying client demand, and notable attrition rates, which can be between 12% and 20% each year. Hence, their workforce planning teams must always balance recruitment efforts with the risk of either being understaffed or overstaffed.

The threshold-based recruitment level, ‘L’, introduced in this paper, matches well with how these firms work. For instance, when the active project pipeline is sparse, HCL or Infosys may keep a small bench of employees and begin recruiting only when the available workforce drops below a critical point.

On the other hand, manpower operations can cause project delays, dissatisfaction among clients, and the need for costly last-minute hiring. So, IT often tries internal hiring triggers based on awaited project inflows. This framework closely resembles how we identify an optimal threshold L^* in our model.

By measuring the relationship between recruitment intensity, attrition, and activation thresholds, our proposed model establishes a sound methodology for maintaining both operational stability and cost efficiency in large manpower systems.

B. Objectives

The main objective of this proposed state-dependent birth–death model is to derive its stationary distribution, the optimal threshold L^* that minimizes a long-run total expected cost function based on recruitment costs, attrition and congestion-induced losses. The whole framework consists of (i) modelling the workforce changes subject to threshold-based recruitment and congestion-sensitive attrition, (ii) describing the steady-state behavior of the manpower model, and (iii) tracing the best threshold that yields the most cost-effective and operationally stable staffing policy for IT manpower environments.

C. Organization of the paper

This paper is organized subsequently as follows: Section II develops the methodology by formulating a controlled Markov manpower model that apprehends random attrition, recruitment activation based on thresholds, and the resulting birth and death method of workforce evolution. It also offers the cost evaluation framework. It derives long-term staffing, vacancy, and recruitment costs, and discusses the best implications of threshold-based control.

Section III introduces the Mean Time to Recruitment (MTR) and derives its closed form expression. This section also emphasizes the practical importance of MTR in IT service environments where delays in recruitment directly impact project delivery and service level compliance. Section IV discusses numeric illustrations of optimal thresholds, supported by two figures. Figure 1 shows the cost function C(L), which exhibits a U-shaped pattern with the minimum at the optimal threshold L* = 3. Figure 2 illustrates the two-parameter cost surface C (L, μ), again showing a U-shaped structure with the minimum at L* = 6. These results establish how threshold selection interacts with recruitment intensity to stabilize manpower levels.

Section V concludes the study by summarizing the methodological contributions, highlighting insights for manpower planning, and outlining future research directions, including multi-grade manpower systems, recruitment sensitive to congestion, and data-driven calibration of threshold policies.

II. METHODOLOGY: MARKOV MANPOWER MODEL

Let X(t) represent the number of employees in a manpower organization at time t, where the workforce is restricted on the finite state space {0, 1, ..., K}, K being maximum capacity. Assume that employee attrition (departure from a queue of employed preseason) occurs according to a Poisson process with rate λ > 0, and recruitment (arrival to the queue) follows an exponential process at rate μ > 0. Further, recruitment is only active if and only if the manpower level satisfies X(t) ≤ L, L < K is called the recruitment activation threshold.

Under this recruitment constraints, it can be shown that the sequence {X(t): t ≥ 0} forms a finite birth–death process with state–dependent arrival and death rates. The downward transition rate from n to n − 1 (attrition) is

$$\lambda_n = \begin{cases} \lambda n, & 1 \le n \le L, \\ \lambda n(1 + \alpha(n - L)), & n > L, \end{cases} \tag{1}$$

where α > 0, a parameter representing congestion sensitivity, increases attrition during the manpower exceeds the threshold L. On the other hand, upward transition rate from n to n + 1 (recruitment) is

$$\mu_n = \begin{cases} \mu, & n \le L, \\ \mu/(K - n), & n > L, \end{cases} \qquad n = 0, 1, \dots, K - 1. \tag{2}$$

Let $\pi_n = \lim_{t\to\infty} \mathbb{P}\{X(t) = n\}$ and the sequence of probabilities $\{\pi_n\}_{n=0}^{K}$ with $\sum_{n=0}^{K} \pi_n = 1$ be the stationary distribution of the $X(t)$ process. Under steady state conditions, the global balance equations for the birth–death process {X(t); t≥ 0} are

$$\text{State } 0: \quad \pi_1\lambda_1 = \pi_0\mu_0, \implies \frac{\pi_1}{\pi_0} = \frac{\mu_0}{\lambda_1},$$

$$\text{State } n, \quad 1 \le n \le K - 1: \ \pi_{n-1}\mu_{n-1} + \pi_{n+1}\lambda_{n+1} = \pi_n(\lambda_n + \mu_n),$$

$$\implies \frac{\pi_{n+1}}{\pi_n} = \frac{\mu_n}{\lambda_{n+1}}, \qquad n = 0, 1, \dots, K - 1.$$

$$\text{State } K: \quad \pi_{K-1}\mu_{K-1} = \pi_K\lambda_K, \implies \frac{\pi_K}{\pi_{K-1}} = \frac{\mu_{K-1}}{\lambda_K}$$

together with the normalization condition process, the stationary probabilities satisfy

$$\pi_n = \pi_0 \prod_{k=0}^{n-1} \frac{\mu_k}{\lambda_{k+1}}, \qquad n = 1, 2, \dots, K, \tag{3}$$

where π_0 is determined by normalization. Using the specific death rates $\{\lambda_n\}$ and birth rates $\{\mu_n\}$, we obtain

$$\pi_n = \begin{cases} \pi_0 \prod_{k=0}^{n-1} \frac{\mu}{\lambda(k+1)}, & 1 \le n \le L, \\ \pi_0 \left[\prod_{k=0}^{L-1} \frac{\mu}{\lambda(k+1)}\right] \prod_{k=L}^{n-1} \frac{\mu/(K-k)}{\lambda(k+1)(1+\alpha(k+1-L))}, & L < n \le K. \end{cases} \tag{4}$$

Finally, π_0 is given by

$$\pi_0 = \left(1 + \sum_{n=1}^{K} \prod_{k=0}^{n-1} \frac{\mu_k}{\lambda_{k+1}}\right)^{-1}. \tag{5}$$

Proposition 2.1 (Impact of Congestion on Stationary Structure). *Consider the threshold-based manpower model with recruitment active only for $n \le L$ and congestion-sensitive attrition for $n > L$, governed by the rates $\{\mu_n\}$ and $\{\lambda_n\}$ defined in (1) and (2). Then, the stationary probabilities $\{\pi_n\}$ satisfy*

$$\frac{\pi_{n+1}}{\pi_n} = \frac{\mu_n}{\lambda_{n+1}}, \qquad n = 0, 1, \dots, K - 1. \tag{6}$$

Since λ_{n+1} increases in n for $n \ge L$ due to the congestion term $1 + \alpha(n - L)$, the ratio π_{n+1}/π_n decreases for $n \ge L$. Hence, π_n becomes strictly decreasing beyond the threshold, reflecting the instability of congested workforce levels.

A. Cost evaluation and optimal implications

The steady state distribution $\{\pi_n\}$ depends on both threshold-based recruitment and congestion-sensitive attrition. Any cost landscape created with linear or nonlinear interactions between workforce shortages and overload effects can guarantee for obtaining an economically optimal policy. A typical perception, formalized in the corollary (2.2) below, is that C(L) remains discretely convex despite this non-linearity.

The long-run cost function in (7) provides a natural economic interpretation of the recruitment threshold L as a control parameter balancing the opposing risks of understating and congestion.

Corollary 2.2 (Convexity of the cost in the recruitment threshold). *Let X denote the steady-state workforce level with stationary distribution $\{\pi_n\}_{n=0}^{K}$ induced by the threshold-based recruitment policy with congestion-sensitive attrition. Consider a long-run expected cost of the form*

$$C(L) = c_u \mathbb{E}\big[(L - X)^+\big] + c_o \mathbb{E}\big[(X - L)^+\big], \qquad c_u > 0,\ c_o > 0, \tag{7}$$

where c_u penalizes understaffing and c_o penalizes congestion and over-staffing. Then $C(L)$ is convex in L on $\{0, 1, \ldots, K-1\}$, i.e.,

$$\Delta^2 C(L) := C(L+1) - 2C(L) + C(L-1) \ \geq\ 0, \qquad L = 1, 2, \ldots, K-2.$$

In particular, the optimal recruitment threshold

$$L^* = \arg\min_{0 \leq L < K} C(L) \tag{8}$$

exists and is (discrete) unique whenever $\Delta^2 C(L) > 0$ for all L in the interior of the feasible range.

Contents of (7) and (8) ensure that the cost admits no extraneous local minima, that the optimal threshold L^* in (8) is well-defined and operationally stable, and that the search for L^* can be performed efficiently. From a managerial view, convexity reveals the basic trade-off that very low thresholds cause continued understaffing penalties, while excessively high thresholds magnify congestion and overstaffing costs, certifying that the minimal cost optimal level lies at a single balance landmark

III. MEAN TIME TO RECRUITMENT (MTR)

For the threshold–based manpower system with $\mu_n = \mu$, recruitment occurs only in states {0, 1, . . . , L}, the long-run average recruitment rate, is defined by

$$R(L, \mu) = \sum_{n=0}^{L} \pi_n \mu = \mu \sum_{n=0}^{L} \pi_n. \tag{9}$$

Also $R(L, \mu)$ represents the expected number of recruitment events per unit time, the reciprocal quantity gives the *mean time between recruitment events*. Hence, the *mean time to recruitment (MTR)* is

$$\mathrm{MTR}(L, \mu) = \frac{1}{R(L, \mu)} = \frac{1}{\mu \sum_{n=0}^{L} \pi_n}, \quad \sum_{n=0}^{L} \pi_n = \pi_0 \left(1 + \sum_{n=1}^{L} \prod_{k=0}^{n-1} \frac{\mu_k}{\lambda_{k+1}}\right), \tag{10}$$

Using the closed-form stationary probabilities derived in (4) and (5), we obtain explicit formulas for MTR:

$$\mathrm{MTR}(L, \mu) = \frac{1 + \sum_{n=1}^{K} \prod_{k=0}^{n-1} \frac{\mu_k}{\lambda_{k+1}}}{\mu \left(1 + \sum_{n=1}^{L} \prod_{k=0}^{n-1} \frac{\mu_k}{\lambda_{k+1}}\right)}. \tag{11}$$

Obviously, the size MTR (L, μ) of (11) declines if the manpower system spends more time in the recruitment-active region {0, . . . , L}, and increases while congestion effects captured by α push the stationary distribution toward higher states n > L, where recruitment is inactive.

A. *Sensitivity of MTR*

Treating L as fixed and using the chain rule, we obtain $\frac{\partial}{\partial \mu} \mathrm{MTR}(L, \mu) = \frac{\partial}{\partial \mu}\left[\frac{1}{\mu P_L(L,\mu)}\right]$ and thus prove that $\frac{\partial}{\partial \mu} \mathrm{MTR}(L, \mu) < 0$, which shows that higher recruitment intensity μ strictly decreases the MTR.

Formally treating L as a continuous design parameter, we differentiate $\mathrm{MTR}(L, \mu)$ with respect to L and using the fact that the finite birth–death process under study, enlarging the recruitment-active set $\{0, \ldots, L\}$ increases the probability of being in

that set, we can establish that

$$\frac{\partial P_L(L, \mu)}{\partial L} > 0 \quad \Longrightarrow \quad \frac{\partial}{\partial L} \mathrm{MTR}(L, \mu) < 0.$$

Thus, increments in the threshold L (allowing recruitment to remain active in more states) reduces the size of MTR.

Since L is an integer design parameter, useing the forward difference $\Delta_L \mathrm{MTR}(L, \mu) = \mathrm{MTR}(L+1, \mu) - \mathrm{MTR}(L, \mu)$, one can show that $\Delta_L \mathrm{MTR}(L, \mu) < 0$,which confirms that raising the recruitment threshold L strictly decreases the amount of MTR.

IV. DISCUSSION ON NUMERICAL ILLUSTRATIVE OPTIMAL THRESHOLD

A. *Numerical Example: Stationary distribution*

To compute the stationary distribution and threshold-based optimal policy, consider a system with capacity K = 4, threshold L = 2, and parameters λ = 1, μ = 1, and α = 1. We then obtain: $\pi_1 = \pi_0$, $\pi_2 = 12\, \pi_0$, $\pi_3 = 112\, \pi_0$, $\pi_4 = 1144\, \pi_0$, with $\pi_0 \approx 0.386$.

Table 1: Cost and Mean Time to Recruitment (MTR) for each threshold L, $c_u = 1$, $c_o = 5$

No.	L	Cost	MTR
1	0	1.828848	1.528993
2	1	1.502108	1.132378
3	2	2.504021	1.036111

A simple numerical search over the L values of Table 1 identifies the optimal manpower level $L^* = 1$ as the minimizer for a wide range of ($c_u = 1$, $c_o = 5$) values. This illustrates how the congestion sensitivity (through large λ_n for n > L) pushes the optimal threshold towards values that avoid persistent operation in highly congested states.

B. *MTR in IT operations.*

The mean time to recruitment (MTR) can be viewed as an interpretable operational indicator of how frequently an IT service organization is expected to trigger hiring interventions subject to a fixed threshold L. Under scenarios dominated by multitasking, emergency protocols, and deadline-based workloads, shorter MTR sizes correspond to more repeated recruitment cycles, informing prolonged pressure on the working employee. Table 1 vales show that the MTR decreases monotonically from 1.528993 at L = 0 to 1.007812 at L = 3, as the recruitment threshold increases, which ensures that the system spends a larger proportion of time in states where the recruitment is active. This phenomenon is compatible with congestion-sensitive attrition in IT operations. Our conclusion is that MTR serves as an additional performance measure next to the cost function C(L) to offer useful insight concerning staffing pattern in IT manpower companies.

The joint behaviour of the MTR and C(L) therefore furnishes a deeper understanding of threshold performance, informing that thresholds jointly with very low MTR may still be sub-optimal from a cost aspect in congestion-sensitive IT workplaces.

C. *IT Workforce Dynamics and Operational Congestion*

We now demonstrate another illustration to fit into an information technology (IT) services industry is characterized by highly dynamic workload patterns, rapid task switching, and persistent deadline-driven operations, where workforce attrition is strongly influenced by the operational congestion experienced by employees.

The cost profile in Figure 1illustrates this trade-off for the parameter setting K = 10, λ = 3, μ = 12, α = 1.5, cu = 1, and co = 5. The convexity of the cost function C(L) reflects the balance between understaffing penalties at low thresholds and congestion penalties at high thresholds. The minimum cost occurs at L = 3, indicating that the optimal recruitment activation level is L∗ = 3.

This value represents the point at which the IT workforce can maintain operational stability while mitigating both service degradation and excessive congestion-induced attrition.

Figure 1 together illustrate the convex behaviour of the long-run cost C(L) as the recruitment activation threshold L varies from 0 to 9. For small values of L, the system experiences substantial understaffing, reflected in the relatively high costs at L = 0 and L = 1. As L increases, the expected understaffing penalty decreases, and the cost reaches its minimum at L = 3, indicating that this threshold balances the trade-off between insufficient staffing and congestion.

Beyond L = 3, the cost begins to rise due to the increasing likelihood of overstaffing, which is penalized through the congestion cost parameter co = 5. The steep growth in C(L) for $L \geq 6$ reflects the rapidly increasing probability of operating in high-congestion states, where attrition is amplified by the congestion sensitivity parameter $\alpha = 1.5$. The resulting U-shaped pattern confirms the discrete convexity of the cost function and demonstrates the existence of a unique optimal threshold. For the chosen parameter values, the minimum cost occurs at L∗ = 3. which there-fore represents the optimal recruitment activation level for this healthcare staffing environment.

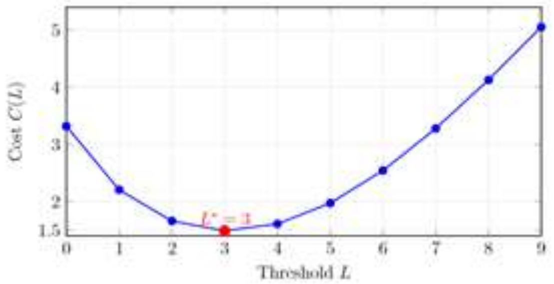

Figure 1: Cost function C(L)exhibiting a U-shaped pattern, with the minimum attained at the optimal threshold L^*=3

D. Optimal threshold– rate pair (L, μ)

The interplay between L, and μ implicitly shapes important metrics such as MTR, and E(X)=expected staffing levels, and α, congestion risk. To achieve an optimal balance between MTR and recruitment size L in an IT manpower levels X, we propose another systematic augmented cost function that annex a few technical trade-offs to jointly discover the long-run execution of the manpower planning. Let

$$C(L,\mu) = w_R\,(MTR\,(L, \mu) - \tau_R\,)^2 + w_X\,(E[X] - \tau_X)^{\,2} + w_C\,P_{cong}(L, \mu) + w_L\,(L^2) + w_\mu(\mu^2),$$

The optimization problems are then:

$$(L*, \mu*) = \text{argmin}_{0\leq L\leq K,\, \mu>0}\, C\,(L, \mu), \text{ and}$$
$$\mu*(L) = \text{argmin}_{\mu>0}\, C(L, \mu).$$

Since the set L ∈ {0,1, ..., K} is a finite, it can be shown that the global optimum exits for evaluating C(L, μ∗(L)) varying L thresholds. Further, the manpower penalty $w_L(L^2)$ and MTR(L, μ) have opposing effects, we can expect that the mapping L → C (L, μ∗(L)) is U-shaped, and therefore the global minimizer L* and the pair (L∗, μ∗) are all uniquely defined.

To obtain the solutions of these optimization exercises, we consider a hypothetical set of input values of C(L,μ) as K = 20, λ=0.6, α=0.05, w_R = 1, w_X = 1, w_C = 0.5, w_L = 0.2, w_μ = 0.1, τ_R = 5, and τ_X = 8, for tracing the optimum pair (L, μ). For this case, first we compute the minimizing recruitment rate μ∗(L) = $\text{argmin}_{\mu>0}$C(L, μ), and record the corresponding minimum cost C (L, μ∗(L)) using R-programming code.

The resulting sequence C∗ versus L is plotted in Figure 1 which exhibits an U-shaped curve: the cost is high for small L due to large MTR(L,μ), decreases to a unique minimum at L = L∗=6, with a minimum cost as 45.299 and then increases for larger thresholds due to the quadratic penalty $w_L(L^2)$.

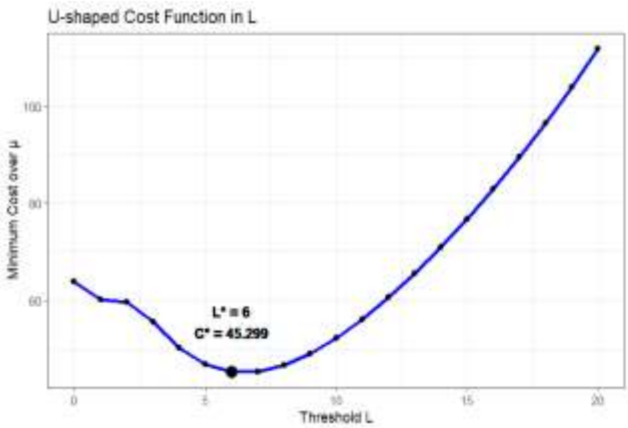

Figure 2: U-shaped cost C(L, μ) showing how the minimum cost occurs at the optimal threshold L*=6

Figure 2 shows the convexity characteristic of the cost C (L, μ): cost declines as L increases from low values, reaches a unique minimum at L^*=6, and then rises again. This property confirms the presence of a well-defined optimal threshold that balances recruitment and attrition costs efficiently.

V. Conclusion and Future Research

This article investigated a queueing-model based threshold-recruitment policies for IT manpower forces subject to congestion-driven attrition. This workforce-model turned out a finite birth–death process with nonlinear departure intensities, We de-rived closed-form mathematical expressions for the stationary distribution, under different recruitment thresholds. The mean time to recruitment (MTR) was computed as an operational performance measure, linking the random fluctuations of the model to the expected timing of hiring events. Establishing the convexity of the long-run cost function C(L), our results support a rigorous foundation for identifying the optimal recruitment activation level that balances understaffing risks against congestion-induced attrition pressures.

Numerical illustrations were demonstrated to know how the interplay between workload intensity, congestion effects, and threshold selection shapes both economic performance and recruitment frequency in IT service organizations.

Our future research may be extended to multi-tier IT teams using analysis of cross-trained workforce pools and differentiated service levels. Further, empirical calibration using real IT operations data would be considered to strengthen the practical relevance of the model and to support the development of decision-support tools for workforce planning in large-scale digital service organizations.

REFERENCES

[1] N. Bacaër and A. Ed-Darraz, "On linear birth-and-death processes in a random environment," J. Math. Biol., vol. 69, pp. 73–90, 2014. doi: 10.1007/s00285-013-0702-3.

[2] R. Belhaj and M. Tkiouat, "A Markov model for human resources supply forecast dividing the HR system into subgroups," J. Service Sci. Manage., vol. 6, pp. 211–217, 2013. doi: 10.4236/jssm.2013.63023.

[3] S. R. Chakravarthy, A. N. Dudin, S. A. Dudin, and O. S. Dudina, "Queueing system with potential for recruiting secondary servers," Mathematics, vol. 11, p. 624, 2023. doi: 10.3390/math11030624.

[4] D. Efrosinin and N. Stepanova, "Estimation of the optimal threshold policy in a queue with heterogeneous servers using a heuristic solution and artificial neural networks," Mathematics, vol. 9, p. 1267, 2021. doi: 10.3390/math9111267.

[5] J. Medhi, Stochastic Models in Queueing Theory, 2nd ed. Amsterdam, The Netherlands: Elsevier, 2003.

[6] J. R. Norris, Markov Chains. Cambridge, UK: Cambridge Univ. Press, 1997.

[7] G. Pang, A. Sarantsev, and Y. Suhov, "Birth and death processes in interactive random environments," arXiv, 2022, arXiv:2203.10411. [Online]. Available: https://arxiv.org/abs/2203.10411

[8] R. J. Boucherie and N. M. van Dijk, Queueing Networks: A Fundamental Approach. New York, NY, USA: Springer, 2017. doi: 10.1007/978-1-4419-6472-4.

[9] D. Gross and C. M. Harris, Fundamentals of Queueing Theory, 3rd ed. New York, NY, USA: Wiley, 1998.

[10] W. J. Peck, "Workforce planning: A review of methodologies," Ann. Oper. Res., vol. 352, no. 1–2, pp. 251–294, 2025. doi: 10.1007/s10479-025-06734-1.

[11] R. Sivasamy, P. Tirupathi Rao, and K. Thaga, "Manpower systems operating under heavy and light tailed inter-exit time distributions," Appl. Math., vol. 5, no. 2, pp. 291–300, 2014. doi: 10.4236/am.2014.52029.

[12] R. Sivasamy, G. Manyeagae, and F. A. Adebayo, "A mathematical model for employee replacement and promotion policies in manpower systems," Int. Conf. Math. Modeling for Cryptographic Protocols (IC2MCP-26), vol. 13, no. 2, 2026. ISSN 2456-1304.

[13] J. L. Vile, J. W. Gillard, P. Harper, and V. Knight, "A queueing theoretic approach to set staffing levels in time-dependent dual-class service systems," Decision Sciences, vol. 48, pp. 766–794, 2016. doi: 10.1111/deci.12236.

[14] H. Wang, G. Zhang, T. Chen, and Z. Li, "Threshold analysis of a stochastic SIRS epidemic model with logistic birth and nonlinear incidence," Mathematics, vol. 11, p. 1737, 2023. doi: 10.3390/math11071737.