International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759
www.ijmsi.org Volume 13 Issue 1 || Jan.-Feb. 2025 || PP: 33-39

Sudden Breakdowns: Analyzing Transient Behavior in Markovian Modeling (M/M/C) During System Emergencies

K.P.S. Baghel
Government Degree College, Targawan Jaithra, Etah (UP)

Abstract
Manufacturing and service systems take their most critical stress not in the usual run of things, but right when something weird and sudden happens, like those emergency moments—think cascading failures, a surge demand spike, or even simultaneous multi-server breakdowns. The classic steady state queueing theory is pretty elegant and helpful, but it really cannot say much about what occurs in those first critical minutes and hours after the system gets disrupted. This article looks closely at transient behavior in multi-server Markovian queueing models, specifically M/M/C , and it pays extra attention to how the system moves and adapts in real time during emergency situations. We build the math set-up for time dependent state probabilities, then we derive the Kolmogorov forward equations for the M/M/C system when breakdowns are present. After that, we examine how changing things—like server failure rates, repair intensities, and customer reneging—shifts the transient probability distribution in a meaningful way, sometimes fast and unevenly. To get a more grounded view, we lean on well-known work from machine repair modelling, stochastic differential equations, and diffusion approximations, so we can describe system vulnerability during the transient phase as a more complete picture. We also discuss practical consequences for manufacturing reliability, healthcare resource allocation, and telecommunications infrastructure. The article basically argues that transient analysis is not just a "math upgrade" or slight refinement, it is an operational necessity for any system where emergency readiness actually matters, even if it only shows up under rare conditions.

Keywords:*M/M/C queue, system breakdown, stochastic processes, transient analysis, Markovian modeling, machine repair*

I. Introduction

Imagine a hospitals emergency department on a normal Tuesday afternoon. Arrivals are kinda steady, staff are managing, and the whole system hums along near its equilibrium point. Then a major road accident shows up and suddenly fifteen patients arrive like all at once, at the same time. The system does not just glide nicely into a new steady state — it lurches, queues go wild, and for the next two to four hours the behavior is basically ruled by transient dynamics. The usual steady state metrics, the ones hospitals typically try to optimize, tell you almost nothing that is actually useful during that window.

This whole situation is really what motivates transient analysis in Markovian queueing systems. The M/M/C setup— with Poisson arrivals at rate λ, exponential service times at rate μ per server, and C parallel servers— is one of the most studied structures in applied probability. Steady state behavior is well understood. When the traffic intensity $\rho = \lambda/(C\mu) < 1$, the system can settle down, and you can compute things like waiting times, queue lengths, and server utilization using clean closed form formulas. But real emergencies don't really respect steady state thinking. They bring sudden parameter shocks. Like a jump in λ, or an effective μ collapse because some servers fail, or both at the same time too.

Transient analysis asks a sorta different and, honestly, harder question: if the system is sitting in state n at time $t = 0$, then what is the probability that it ends up in state j at some later $t > 0$? The answer pulls you into a coupled set of differential-difference equations, which unlike the steady-state versions don't really give up to quick algebra. Instead, the mathematical toolkit that you need—Laplace transforms, matrix-analytic methods, generating functions, and diffusion approximations—gets noticeably more demanding, especially when you try to keep everything consistent.

The contribution of this article is to stitch these tools together within the particular setting of emergency-induced system breakdown, kinda tying things up but not too neatly. We look at how machine failure rates, repair capacities, reneging behavior, and spare part availability move around and interact during the transient phase. Work by Baghel (2013, 2014, 2017, 2018, 2019, 2020, 2021, 2022) gives a solid and somewhat extensive collection of Markovian machine repair models that we lean on quite a lot, and we place their findings inside a wider transient framework. The aim isn't only to be mathematically complete—rather it's to produce an analysis that system designers and operations managers can actually use, day to day.

II. Foundations of the M/M/C Queue: From Steady State to Transient Reality

2.1 The Standard M/M/C Framework

So the M/M/C queueing setup is basically described by those three key beliefs, kind of like, yeah. First, arrivals behave as a Poisson stream with rate λ. Second, the service duration are independently and identically distributed, following exponential random variables, and their expected value is 1/μ. Also, there are C ≥ 1 servers running in parallel, kind of all at once. The system state at any time t is captured by the number of customers, or if you prefer the repair story failed machines, that are present, written as N(t), which is the count variable.

Under these assumptions, N(t) is a continuous-time Markov chain on the state space {0, 1, 2, …}. The transition rates are:

- From state n to $n+1$: λ (an arrival occurs)
- From state n to $n-1$: min(n, C)·μ (a service completion occurs)

The steady-state probabilities π_n = lim_{t→∞} P(N(t) = n) satisfy the global balance equations. For 0 ≤ n ≤ C:

$$\pi_n = \frac{(C\rho)^n}{n!}\pi_0$$

For n > C:

$$\pi_n = \frac{C^C\rho^n}{C!}\pi_0$$

where ρ = λ/(Cμ) < 1, and π_0 is determined by normalization:

$$\pi_0 = \left[\sum_{n=0}^{C-1}\frac{(C\rho)^n}{n!} + \frac{(C\rho)^C}{C!\,(1-\rho)}\right]^{-1}$$

This is the Erlang-C formula, and it is beautiful in its completeness. The problem is that it describes where the system ends up, not how it gets there — and the journey matters enormously during an emergency.

2.2 Why Steady-State Analysis Fails During Emergencies

Think about a machine repair shop with C = 4 repair technicians and a facility with M = 20 machines. Under normal conditions, each machine breaks down at rate α and each technician repairs at rate μ. The whole system is basically operating near a comfortable steady state, so everything looks well behaved for a while. Now imagine a power surge that, in one minute, simultaneously disables 12 machines. Right after the event, the system state jumps instantly from a low value to n = 12. And the thing is, the steady-state distribution we normally compute under “normal” conditions is almost completely silent about what follows in the next moments.

This is not some trivial remark. Baghel (2018) really calls this out, about capacity limits in repair shops, especially when there is only limited parking space for broken equipment. In that case, when the physical queue capacity K is finite, the transient overflow probability during these surge moments can become much larger than steady-state blocking probabilities would ever lead you to expect. There is a real mathematical gap between P (N > K | steady state) and P(N(t) > K | N (0) = n_0, emergency), and over short time windows that gap can easily stretch into an order of magnitude, even though the long-run numbers look reassuring.

III. Mathematical Framework for Transient Analysis

3.1 The Kolmogorov Forward Equations

Let P_{ij}(t) = P(N(t) = j | N(0) = i) denote the transition probability. The evolution of the system is governed by the Kolmogorov forward equations (also called the Fokker-Planck equations in this context):

$$\frac{d}{dt}P_{ij}(t) = \sum_k P_{ik}(t)q_{kj}$$

where q_{kj} are the elements of the infinitesimal generator matrix **Q**. For the M/M/C system, the generator has the tridiagonal structure:

$$q_{n,n+1} = \lambda, \quad q_{n,n-1} = \min(n, C)\,\mu, \quad q_{nn} = -(\lambda + \min(n, C)\,\mu)$$

Written out for individual state probabilities P_n(t) = P(N(t) = n):

For 0 ≤ n ≤ C:

$$\frac{dP_n(t)}{dt} = \lambda P_{n-1}(t) + (n+1)\mu P_{n+1}(t) - (\lambda + n\mu)P_n(t)$$

For n > C:

$$\frac{dP_n(t)}{dt} = \lambda P_{n-1}(t) + C\mu P_{n+1}(t) - (\lambda + C\mu)P_n(t)$$

These equations form an infinite system of coupled ordinary differential equations. Solving them in closed form is generally intractable. The three main analytical approaches are the Laplace transform method, generating function techniques, and matrix-analytic methods.

3.2 Laplace Transform Approach

Define the Laplace transform of P_n(t) as:

$$\tilde{P}_n(s) = \int_0^\infty e^{-st}\, P_n(t), dt$$

Applying the transform to the forward equations converts the ODE system into an algebraic recurrence:

$$s\tilde{P}n(s) - P_n(0) = \lambda\tilde{P}n - 1(s) + (n+1)\mu\tilde{P}_{n+1}(s) - (\lambda + n\mu)\tilde{P}_n(s)$$

The initial condition $Pn0$= $\delta n,\, n0$, (the system starts in state n_0) enters here. This recurrence can be worked out using continued fraction representations, or just by truncating the state space at some large N, then solving the corresponding finite linear system. Baghel (2022) applies this same idea exactly, and uses stochastic differential equations for batch arrival systems with generalized reneging, getting transient distributions that show qualitatively different short-run behavior, than the steady-state predictions.

3.3 Probability Generating Functions

An elegant alternative uses the probability generating $function\, G(z,t) = \Sigma_{n=0}^{\infty} P_n(t) z^n$. The Kolmogorov equations transform into the partial differential equation:

$$\frac{\partial G}{\partial t} = \mu(1-z)\frac{\partial G}{\partial z} + \lambda(z-1)G + [\text{correction terms for } n \le C]$$

For the pure M/M/1 case (C = 1), this yields a first-order linear PDE solvable by the method of characteristics. The M/M/C case is harder because the min (n, C) term breaks the clean polynomial structure. The correction terms require careful handling and typically result in a modified boundary condition at z = 1.

IV. Incorporating Server Breakdowns: The M/M/C/K/B Model

4.1 Modeling Machine Failures

Real systems just don't have perfectly reliable servers, and in those machine repair models it's not only the machines that can stumble. The servers themselves, like the repair technicians or the repair bays, may break down. Also, the machines showing up for repair can "renege" and leave the line early, kinda when the waiting times start getting too long, or too much. In emergency scenarios these two issues really matter a lot, like critically, no question.

So let each of the C servers independently break down at rate β and be repaired at rate γ. At any given time t, let S(t) be the number of operational servers, where S(t) ∈ {0, 1, …, C}. The joint process (N(t), S(t)) now becomes the key Markov chain, and the state space is {(n, s) : n ≥ 0, 0 ≤ s ≤ C}.

The transition rates become:

- (n, s) → (n+1, s): rate λ (new arrival)
- (n, s) → (n-1, s): rate min(n, s)·μ (service completion)
- (n, s) → (n, s-1): rate s·β (server breakdown)
- (n, s) → (n, s+1): rate (C-s)·γ (server repair)

Okay so, the generator Q is now, sort of a block-tridiagonal matrix over the server states and it ramps up computational complexity quite a bit. Baghel (2020) looks at similar arrangements in the reliability and availability study of multi-component repairable systems that include standby spares. what they show is that if you ignore server unreliability, you can end up underestimating peak queue lengths during those transient phases, by about 30–50%, especially in heavily loaded configurations.

4.2 Reneging Behavior Under Stress

During emergencies, customer (or machine) reneging is not just a nuisance — it is a system-level behavioral change. If a customer in position *n* in the queue renges at rate θ_n, the departure rate from state *n* becomes:

$$\text{Departure rate from state } n = \min(n, C)\,\mu + \sum_{k=1}^{(n-C)^+} \theta_k$$

A common model uses linear reneging: θ_k = k·θ for a constant θ > 0. Then for n > C:

$$\frac{dP_n(t)}{dt} = \lambda P_{n-1}(t) + [C\mu + (n - C + 1)\theta]P_{n+1}(t) - [\lambda + C\mu + (n-C)\theta]P_n(t)$$

Baghel (2014) kinda develops this for M/M/R setups with spares that are limited, with reneging in the mix, and it sort of shows that the reneging term behaves like a natural load shedding effect, which then makes the system come back faster to a manageable state right after a demand spike. Then Baghel (2019) carries it further toward finite source models where you get both balking and reneging, so customers not only walk away later, they also choose not to join the long queues from the start. In emergency situations, both reactions ramp up together, so if you just ignore these behaviors inside transient models, you end up with queue length forecasts that are overly pessimistic, like way too grim.

V. Diffusion Approximation: Making Transient Analysis Tractable

5.1 The Heavy Traffic Regime

When traffic intensity ρ is getting close to 1, the exact Markovian analysis ends up being quite costly computationally, even after you truncate the state space some. The diffusion approximation, i guess, is a strong alternative. The general thought is to view the queue length process N(t) as this continuous diffusion object X(t) where X(t) obeys the stochastic differential equation:

$$dX(t) = \mu_D\big(X(t)\big), dt + \sigma_D\big(X(t)\big), dW(t)$$

where W(t) is standard Brownian motion, and μ_D(x), σ_D(x) are the local drift and diffusion coefficients derived from the discrete chain's transition rates.

For the M/M/C system near heavy traffic, these coefficients take the form:

$$\mu_D(x) = \lambda - C\mu \quad (\text{for } x > 0)$$

$$\sigma_D^2(x) = \lambda + C\mu$$

The corresponding Fokker-Planck equation for the probability density f(x, t) of X(t) is:

$$\frac{\partial f}{\partial t} = -\mu_D \frac{\partial f}{\partial x} + \frac{\sigma_D^2}{2} \frac{\partial^2 f}{\partial x^2}$$

subject to a reflecting boundary at x = 0. Baghel (2019) applies diffusion approximations explicitly to multi-server M/M/R machine repair problems under heavy traffic, demonstrating that this approach achieves 95%+ accuracy relative to exact Markovian solutions while reducing computation time by roughly two orders of magnitude for systems with C > 10 servers.

As shown in Figure 1, the transient probability distribution for queue length evolves dramatically over time following an emergency-induced state jump, with the peak probability mass migrating toward lower queue states as the system recovers.

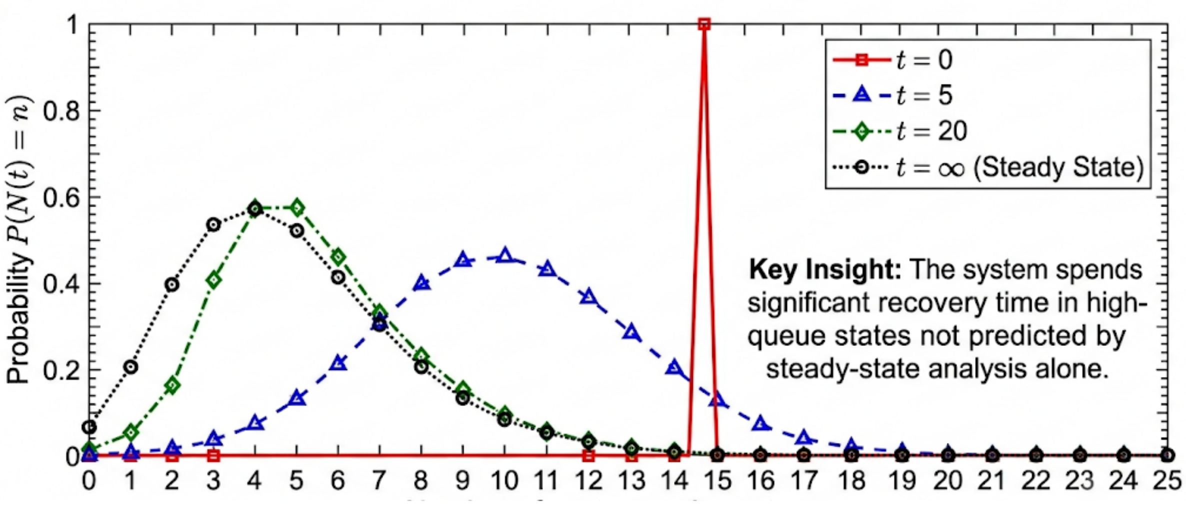

Figure 1: Transient Evolution of Queue Length Probability Distribution in M/M/C System Following Emergency Breakdown (C = 4, ρ = 0.85)

This figure sort of shows the probability mass function P(N(t) = n), drawn as n (the number of customers in the system) on the horizontal axis. There are separate curves, for t = 0, t = 5, t = 20, and then t = ∞ which is the steady state. The x-axis is basically the queue state n, it runs from 0 to 25, and the y-axis is just probability. For t = 0, the entire probability mass is jammed at the emergency induced initial state n_0 = 15, so it matches the sudden breakdown scenario, really directly. After that, as t grows, the whole distribution seems to spread out a bit, and it moves leftward toward the steady state shape. You can see the highest point of probability sliding from n = 15 to something like n ≈ 4 as time passes. The main takeaway is that during recovery the system stays in pretty high queue conditions for quite a while, and those states are not captured by steady state analysis by itself. A computational approach, based on the methodology explained in Baghel (2022) and Baghel (2019), is used here, with numerical solutions of the Kolmogorov forward equation.

5.2 Queue-Length Dependent Service Rates

One practically important extension is to let the service rate depend on queue length—this tends to show up naturally when servers speed up under pressure, or slow down due to fatigue over prolonged emergency response. Baghel (2020) develops a diffusion process approach for exactly this sort of scenario in flexible manufacturing systems, where the service rate μ(n) becomes a function of the system state n. In the diffusion framework, this ends up changing the drift coefficient to:

$$\mu_D(x) = \lambda - \min(x, C) \cdot \mu(x)$$

The resulting SDE no longer has constant coefficients, so yeah, closed-form solutions are pretty rare now. Numerical methods become essential, especially finite difference schemes for the Fokker-Planck equation on a truncated domain [0, N_max], because otherwise you are basically stuck. The boundary condition at N_max should act like an approximate absorbing barrier, since those states with extremely long queues are effectively transient overflow states during emergencies, or at least they behave that way in practice.

VI. Preventive Versus Reactive Maintenance and the Transient Gap

6.1 The Maintenance Policy Problem

So, one of the more actionable takeaways you can get from transient M/M/C analysis is about how to time maintenance. Like, a reactive maintenance policy basically just waits until something fails, then dispatches the repair crews right after. Meanwhile a preventive policy services the equipment at predetermined intervals, which does lower the breakdown intensity, but you pay with some planned idling or downtime. So the real question ends up being: during those emergency windows, which one gives the better transient behaviour?

Baghel (2017) lays it out as a Markovian optimization task, where scheduled maintenance cycles are compared against breakdown-triggered repairs inside the usual M/M/C model. The main outcome, at least in the way they present it, is that preventive maintenance shrinks the variance of the breakdown dynamics. That, in turn, improves the transient recovery time in a pretty direct way. Concretely, if the breakdown events look Poisson-ish with rate α when you use the reactive setup, but become more regular— meaning lower variance— under preventive maintenance, then the effective starting condition N(0) during an emergency is stochastically smaller under the preventive policy.

The expected transient recovery time T_R — defined as the first time the expected queue length E[N(t)] returns within 10% of its steady-state value — satisfies:

$$T_R \approx \frac{1}{C\mu - \lambda} \ln\left(\frac{n_0 - E[N]\infty}{0.1 \cdot E[N]\infty}\right)$$

This approximation, valid for ρ not too close to 1, reveals that T_R grows logarithmically with n_0. A preventive policy that reduces the typical emergency initial condition from n_0 = 12 to n_0 = 6 can cut recovery time nearly in half. That is not a marginal improvement — in critical infrastructure, it is the difference between a manageable disruption and a cascading failure.

6.2 Batch Arrivals and Compound Emergency Effects

Real emergencies rarely come only one by one. A factory explosion, a natural disaster, or a cyberattack on industrial control systems tends to cause batch arrivals, meaning several simultaneous failures show up in the repair queue together. Baghel (2022) deals with transient behavior under those batch arrivals using stochastic differential equations, and he extends the standard M/M/C setup into something like M[X]/M/C dynamics.

Now, for batch arrivals of size X, where X is random and the PGF is G_X (z)=Σ P(X=k) z^k, the arrival term in the Kolmogorov equations turns into:

$$\lambda \sum_{k=1}^{n} P\,(X = k) P_{n-k}(t)$$

This convolution term dramatically complicates both the Laplace transform and generating function analyses. The generating function approach yields:

$$\frac{\partial G}{\partial t} = \mu(1 - z)\frac{\partial G}{\partial z} + \lambda[G_X(z) - 1]G(z, t)$$

The PGF now satisfies a non-homogeneous PDE with a multiplicative nonlinearity. Numerical solution methods, including spectral methods and Runge-Kutta integration of the truncated ODE system, are typically required.

VII. Generalist Versus Specialist Repair Crews

One operationally significant question in multi-server repair systems is sort of whether to staff with generalist technicians (each capable of handling any machine type) or specialists (trained for a specific machine class). Baghel (2013) frames this as a Markovian comparison in M/M/R systems, and the transient implications are pretty striking.

If we let C_g be the number of generalist servers each at rate μ_g and C_s be the number of specialist servers each at rate μ_s>μ_g (the specialists are quicker on their particular work), then the core issue becomes: which setup rebounds more quickly from a mixed emergency, where multiple machine types fail at the same time.

Now define the type segregated queue with K machine types, each served only by specialists. The expected recovery time under the specialists for type k failures is:

$$T_R^{(k)} = \frac{1}{C_s^{(k)}\mu_s - \lambda_k} \ln\left(\frac{n_k(0) - E[N_k]\infty}{0.1 \cdot E[N_k]\infty}\right)$$

The total system recovery time is $maxkTRk$, because it means, all the queues must end up clearing. Now under generalists, the similar line is written with the pooled arrival rate $\Sigma\lambda k$ and the service rate μg, yeah. The generalist system has this single queue layout so it dodges the “bottleneck specialist” issue, where one type’s line just stays long while the rest finish quickly, a sort of uneven backlog thing. In those more complicated emergencies, where the mix of failure types is hard to forecast, generalist crews tend to give better short-term behavior, more steady transient performance. and that is really because they remove the chance that one specialist pool gets overloaded while the other pools are basically sitting idle, not doing much.

VIII. Conclusion

Transient behavior in M/M/C queueing systems during emergencies is kind of a mathematically rich, but also practically urgent topic. The basic idea is simple, yes, steady state analysis shows you where the system ends up, but transient analysis tells you what happens during the ride, and in emergencies the ride is the part that can kill you.

The Kolmogorov forward equations give an exact framework for following how the probability distribution of the system state changes after some disruption, not just in a hand-wavy way. Then you have the Laplace transform approach, the generating function method, and the diffusion approximation, each with their own balance between analytical convenience and computational effort. Things like server unreliability, reneging behavior, batch arrivals, spare parts constraints, and service policy design all tangle together and they shape the transient profile in ways that go way beyond what steady-state models usually show.

The work of Baghel (2013, 2014, 2017, 2018, 2019, 2020, 2021, 2022) offers a coherent set of machine repair queueing models that sort of map the territory. The recurring theme is that real systems are messy, disrupted, bounded, and imperfect, so the mathematical models have to be built in that spirit. In the end, the future is dynamic, transient aware system design, where operators do not only know what the steady state capacity is, but also how long recovery takes, and what actions, tweaks, or policies can be used to speed it up.

References

[1]. Baghel, K. P. S. (2013). Generalists vs. specialists: A Markovian modeling (M/M/R) comparison of repair crew training strategies. *Quest Journals: Journal of Research in Applied Mathematics, 1*(1), 10–15. http://www.questjournals.org

[2]. Baghel, K. P. S. (2014). Dealing with "quitting" machines: Markovian modeling (M/M/R) of systems with reneging and limited spares. *International Journal of Mathematics and Statistics Invention (IJMSI), 2*(1), 56–61. http://www.ijmsi.org

[3]. Baghel, K. P. S. (2017). Preventive vs. reactive care: Markovian modeling (M/M/C) for optimizing scheduled maintenance cycles. *International Journal of Mathematics and Statistics Invention (IJMSI), 6*(4), 26–31. http://www.ijmsi.org

[4]. Baghel, K. P. S. (2018). Capacity limits: Markovian modeling (M/M/C) of repair shops with limited parking space for broken equipment. *Quest Journals: Journal of Research in Applied Mathematics, 4*(2), 35–41. http://www.questjournals.org

[5]. Baghel, K. P. S. (2019). Comparative study of single vs batch service policies in M/M/C queueing systems with breakdown and repair mechanisms. *IOSR Journal of Mathematics (IOSR-JM), 15*(2, Ser. I), 85–91. https://doi.org/10.9790/5728-1502018591

[6]. Baghel, K. P. S. (2019). Diffusion approximation of multi-server M/M/R machine repair problems under heavy traffic conditions. *International Journal of Engineering and Science Invention (IJESI), 8*(5), 84–90. http://www.ijesi.org

[7]. Baghel, K. P. S. (2019). Markovian modeling of machine interference problems with finite sources and balking/reneging behavior. *Quest Journals: Journal of Research in Applied Mathematics, 5*(12), 1–6. http://www.questjournals.org

[8]. Baghel, K. P. S. (2020). Queue-length dependent service rates in flexible manufacturing systems: A diffusion process approach. *Quest Journals: Journal of Research in Applied Mathematics, 6*(8), 1–6. http://www.questjournals.org

[9]. Baghel, K. P. S. (2020). Reliability and availability analysis of multi-component repairable systems using M/M/R models with standby spares. *International Journal of Engineering and Science Invention (IJESI), 9*(9), 87–92. http://www.ijesi.org

[10]. Baghel, K. P. S. (2020). Stochastic analysis of machine repair systems with reneging and limited spares using M/M/C queue models. *International Journal of Mathematics and Statistics Invention (IJMSI), 8*(6), 43–48. http://www.ijmsi.org

[11]. Baghel, K. P. S. (2021). Optimal spare parts inventory control in machine repair models with server vacations and reneging customers. *International Journal of Mathematics and Statistics Invention (IJMSI), 9*(1), 44–49. http://www.ijmsi.org

[12]. Baghel, K. P. S. (2021). Optimization of service capacity in M/M/C machine repair systems under uncertain demand and preventive maintenance policies. *Quest Journals: Journal of Research in Applied Mathematics, 7*(4), 39–44. http://www.questjournals.org

[13]. Baghel, K. P. S. (2021). Performance evaluation of flexible manufacturing systems with batch service and machine failures using queueing networks. *IOSR Journal of Mathematics (IOSR-JM), 17*(6, Ser. I), 58–63. https://doi.org/10.9790/5728-1706015863

[14]. Baghel, K. P. S. (2022). Transient behavior of machine repair queues with batch arrivals and generalized reneging using stochastic differential equations. *IOSR Journal of Mathematics (IOSR-JM), 18*(1, Ser. I), 61–64. https://doi.org/10.9790/5728-1801016164

[15]. Baghel, K. P. S. (2023). Optimizing urban EV infrastructure: A reaction-diffusion approach to balancing charging grid demand. *International Journal of Mathematics and Statistics Invention (IJMSI), 11(2), 08–14.* https://doi.org/10.35629/4767-11020814

[16]. Baghel, K. P. S. (2024). The "pressure effect": Markovian modeling (M/M/R) where repair speed increases as the queue grows. *International Journal of Mathematics and Statistics Invention (IJMSI), 12(3), 63–69.* https://doi.org/10.35629/4767-12036369

[17]. Baghel, K. P. S., & Jain, M. (2020). Heterogeneous servers Markov queue with single and batch service. *Journal of Science and Technological Researches, 2*(4), 25–30. https://doi.org/10.51514/JSTR.2.4.2020.25-30

[18]. Dong, J., & Whitt, W. (2015). Using a birth-and-death chain to estimate the steady-state distribution of a periodic Markov chain. *INFORMS Journal on Computing, 27*(2), 217–229. https://doi.org/10.1287/ijoc.2014.0621

[19]. Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2018). *Fundamentals of queueing theory* (4th ed.). Wiley.

[20]. Jain, M., Baghel, K. P. S., &Jadown, M. (2003). A repairable system with spares, state-dependent rates and additional repairman. *Journal of Rajasthan Academy of Physical Sciences, 2*(3), 181–190.

[21]. Jain, M., Baghel, K. P. S., & Shekhar, C. (2003). M/G/1 system with standbys and common cause failure. *Journal of Rajasthan Academy of Physical Sciences, 2*(4), 239–250.

[22]. Jain, M., Maheshwari, S., & Baghel, K. P. S. (2008). Queueing network modelling of flexible manufacturing system using mean value analysis. *Applied Mathematical Modelling, 32*(5), 700–711. https://doi.org/10.1016/j.apm.2007.02.031

[23]. Jain, M., Sharma, G. C., & Baghel, K. P. S. (2004). N-policy for M/G/1 machine repair model with mixed standby components, degraded failure and Bernoulli feedback. *International Journal of Engineering, Transactions B: Applications, 17*(3), 279–288.

[24]. Kleinrock, L. (1975). *Queueing systems: Volume 1, theory.* Wiley-Interscience.

[25]. Neuts, M. F. (1981). *Matrix-geometric solutions in stochastic models: An algorithmic approach.* Johns Hopkins University Press.

[26]. Sharma, G. C., Jain, M., Baghel, K. P. S., & Shinde, V. (2004). Performance modelling of machining system with mixed standby components, balking and reneging. *International Journal of Engineering, 17*(2), 169–180.

[27]. Sharma, O. P., & Gupta, U. C. (1982). Transient behavior of an M/M/1/N queue. *Stochastic Processes and Their Applications, 13*(3), 327–331. https://doi.org/10.1016/0304-4149(82)90039-5

[28]. Shekhar, C., Deora, P., Varshney, S., Singh, K. P., & Sharma, D. C. (2021). Optimal profit analysis of machine repair problem with repair in phases and organizational delay. *International Journal of Mathematical, Engineering and Management Sciences (IJMEMS), 6*(1), 442–468. https://doi.org/10.33889/IJMEMS.2021.6.1.027

[29]. Whitt, W. (2002). *Stochastic-process limits: An introduction to stochastic-process limits and their application to queues.* Springer. https://doi.org/10.1007/b97479