

Fixing the Chain Reaction: What Happens When Many Machines Stop Working

K.P.S. Baghel

Govt. Degree College Manikpur, Chitrakoot (U.P.)

Abstract

Repairs do not always start when only one machine breaks. If many units stop around the same time, the team that can fix them becomes the main bottleneck. Then the waiting line can drive a large share of the lost output. The paper looks at this case with a finite source birth death model. There are N identical machines. There are R repair workers who serve them. The study gets the steady state probabilities. It also turns those results into closed form expressions for mean downtime. It reports availability, worker use, and the mean time a job spends in the repair shop. After the base model, two changes are added. First, there is a load sharing cascade. Each failed machine makes the remaining working machines more likely to fail. The failure rate rises by a factor of $(1 + \gamma n)$, where n is the number failed. Second, there is reneging. A job that waits too long may quit the repair line. That departure happens at rate β . This idea was discussed in Baghel (2014) for similar machine setups. To keep the arithmetic easy, the text uses a three machine example with $\rho = \lambda/\mu = 0.5$. With one crew, availability is 52.6%. With two crews, availability becomes 65.5%. A third crew helps only a little, adding about 1.2 percentage points. In the cascade case, $\gamma = 0.5$ drops the one crew availability to 39.3%. It also lifts the chance that all machines are down. That probability moves from 15.8% to 32.1%. Under reneging, the mean number of machines down goes from 1.42 to 1.13. At the same time, 21.4% of failed units leave the queue before they get fixed. The paper then compares training ideas from Baghel (2013). It argues that pooling two crews over four machines cuts expected downtime by 8.0%. It also notes a tradeoff point. A pooled crew can be about 12% slower per repair before the benefit fades. A final part gives a cost based rule for choosing crew size. The final formulas are simple enough for a spreadsheet. That is why they fit maintenance planning tasks. The values shown are outputs from the model for the stated parameters. They are not measurements taken from the field.

Keywords: machine interference, birth-death process, repairable systems, reneging, cascading failures, availability

I. Introduction

1.1 Background

In a plant, breakdowns do not behave like separate events. One voltage drop can knock out several drives together. A hot afternoon can push every compressor on a feeder. If a conveyor stops, the units after it run out of work. So the repair crew often faces a pile of failed items, not one clear fault.

A common way to frame this is the machine interference model. In that view, a set of machines keeps running until they fail. Then a smaller team of repair workers brings them back. While the workers are occupied, other failed machines sit and wait. That waiting time is the cost caused by interference. Palm (1947) set this up for textile equipment. Stecké and Aronson (1985) looked at both operator and machine forms. Haque and Armstrong (2007) later reviewed work on the topic.

This setup depends on two ideas that many plants do not match. One idea is that failures are independent. Another is that a failed machine will wait through the full service time. In reliability work, it is known that failures can be linked. A usual example is load sharing, where machines that survive take extra duty (Rausand and Høyland, 2004). Work on networks also covers chain reactions. Watts (2002) treated large cascades as threshold effects in random graphs. Buldyrev and coauthors (2010) found that when two networks depend on each other, a small random hit can trigger a wide collapse. The other mismatch is impatience. If a machine waits too long, it may get taken apart, rerouted, or removed from service. Queueing research calls that reneging (Barré, 1957; Ancker and Gafarian, 1963).

1.2 Objectives and Scope

The article lists three goals. First, it lays out the finite source model that serves as the starting point. It also gives closed form formulas for the performance. Second, it modifies that base model. It adds a cascade effect and a reneging effect. Then it explains how each addition shifts the stationary distribution. Third, it looks at two choices for the design. One question is if the crews should be shared. The other question is how many crews the plant should use.

II. The Baseline Model

2.1 Assumptions

Consider N identical machines. Each working machine fails after an exponentially distributed time with rate λ , the usual lifetime assumption in reliability theory (Barlow & Proschan, 1965). There are R repair crews, and each crew repairs one machine at a time with exponential service rate μ . Failed machines form a single queue served in order of arrival, and no machine leaves the queue before repair.

Let n denote the number of machines down. The process moves between 0 and N in single steps, so it is a birth-death process (Ross, 2014). A failure moves the state up at rate

$$\lambda_n = (N - n)\lambda, n = 0, 1, \dots, N - 1 \dots (1)$$

and a repair completion moves it down at rate

$$\mu_n = \min(n, R)\mu, n = 1, \dots, N \dots (2)$$

Equation (1) expresses the defining property of a finite source. As more machines fail, fewer remain to fail, so the stream of new breakdowns thins. This separates the model from a queue with an unlimited population (Gross et al., 2008). Figure 1 shows the structure of the chain.

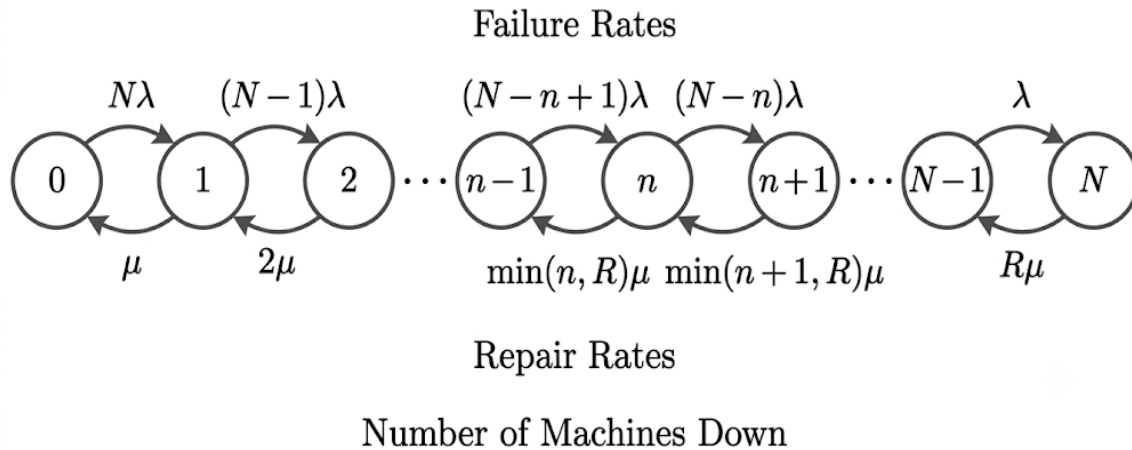


Figure 1: State-Transition Diagram of the Finite-Source Machine Repair Model with N Machines and R Crews

In the diagram, the states go from 0 up to N , left to right. Here, state n means there are n machines down at that time. The arrows that point to the right use the rate $(N-n)\lambda$. The arrows that point to the left use the rate $\min(n, R)\mu$. Notice what happens as n changes. As n gets larger, the right-side rates get smaller. For the left side, the rate rises at first, then it stops growing once n reaches R . This setup matches the usual model from Gross et al. (2008) and Cooper (1981).

2.2 Stationary Distribution

In equilibrium the probability flow across each cut between adjacent states balances, so $\lambda_{n-1}p_{n-1} = \mu_n p_n$. Solving recursively gives

$$p_n = p_0 \prod_{k=1}^n (\lambda_{k-1} / \mu_k) \dots (3)$$

Write $\rho = \lambda/\mu$ and substitute equations (1) and (2). For $n \leq R$,

$$p_n = p_0 \cdot C(N, n) \cdot \rho^n \dots (4)$$

and for $R < n \leq N$,

$$p_n = p_0 \cdot [N! / ((N - n)! R! R^{n-R})] \cdot \rho^n \dots (5)$$

Denote the multiplier of p_0 by w_n , with $w_0 = 1$. Consecutive weights satisfy

$$w_n / w_{n-1} = (N - n + 1) \rho / \min(n, R) \dots (6)$$

Normalization then gives

$$p_0 = (\sum_{n=0}^N w_n)^{-1} \dots (7)$$

The algorithm is short. Build the weights from (6), sum them, and divide. Kleinrock (1975) and Medhi (2003) give equivalent derivations, and Cooper (1981) treats the single-crew case in detail.

2.3 Performance Measures

Every measure of interest follows from the p_n . The mean number of machines down is

$$L = \sum_{n=0}^N n p_n \dots (8)$$

and the steady-state availability of the fleet is

$$A = 1 - L/N \dots (9)$$

Let $B = \sum \min(n, R) p_n$ be the mean number of busy crews. Crew utilization is then B/R . In equilibrium the rate of failures equals the rate of repair completions, so

$$\lambda_e f f = \lambda(N - L) = \mu B \dots (10)$$

Dividing by λ yields an identity that is convenient for hand calculation:

$$L = N - B/\rho \dots (11)$$

Little's (1961) law gives the mean time a failed machine spends in the shop, waiting plus repair:

$$W = L/\lambda_e f f \dots (12)$$

For a single crew, equation (5) reduces to $w_n = N! \rho^n / (N - n)!$, and $B = 1 - p_0$. Equation (11) then becomes

$$L = N - (1 - p_0)/\rho \dots (13)$$

Only p_0 is needed. Equation (10) also supplies a useful check, because the two expressions for $\lambda_e f f$ must agree. If they do not, the weights contain an arithmetic error.

III. A Worked Example

3.1 One Crew

Take $N = 3$, $\rho = 0.5$, and $R = 1$. Equation (6) gives $w_1 = 3(0.5) = 1.5$, $w_2 = 1.5 \times 2(0.5) = 1.5$, and $w_3 = 1.5 \times 1(0.5) = 0.75$. With $w_0 = 1$ the weights sum to 4.75, so

$$p_0 = 0.2105, p_1 = 0.3158, p_2 = 0.3158, p_3 = 0.1579.$$

From equation (8), $L = (1.5 + 3.0 + 2.25)/4.75 = 1.421$, and equation (9) gives $A = 1 - 1.421/3 = 52.6\%$. As a check, $B = 1 - p_0 = 0.7895$, and equation (13) gives $L = 3 - 0.7895/0.5 = 1.421$. The two routes agree. Equation (12) gives $W = 1.421/0.7895 = 1.80$ mean repair times. A failed machine therefore spends 1.8 repair times in the shop, of which only one is actual repair.

3.2 Two and Three Crews

With $R = 2$, equation (6) divides by $\min(n, 2)$. The weights are $w_1 = 1.5$, $w_2 = 1.5 \times 2(0.5)/2 = 0.75$, and $w_3 = 0.75 \times 1(0.5)/2 = 0.1875$, with a sum of 3.4375. Then $L = (1.5 + 1.5 + 0.5625)/3.4375 = 1.036$, and availability rises to 65.5%. Mean busy crews are $B = 0.982$, so utilization falls from 78.9% with one crew to 49.1% with two.

With $R = 3$, every failed machine has its own repairer. Each machine then behaves as an independent up-down unit that is down with probability $\rho/(1 + \rho)$, so

$$L = N\rho/(1 + \rho) = 3(0.5)/1.5 = 1.000 \dots (14)$$

and $A = 66.7\%$. The second crew adds 12.8 percentage points of availability, and the third adds 1.2. Table 1, in Section 6.4, collects these results with the ones that follow. Section 7 returns to the gap between 12.8 and 1.2.

3.3 Sensitivity to Load

Availability depends strongly on ρ . Keep $N = 3$ and $R = 1$ and vary only ρ . At $\rho = 0.25$ the weights are 0.75, 0.375, and 0.094, which sum with w_0 to 2.219. Then $L = 0.803$ and $A = 73.2\%$. At $\rho = 1.0$ the weights are 3, 6, and 6, which sum with w_0 to 16. Then $L = 33/16 = 2.063$ and $A = 31.3\%$.

Raising ρ from 0.25 to 0.5 costs 20.6 points of availability. Raising it from 0.5 to 1.0 costs a further 21.4. The relationship is nonlinear, so a shop that coped at one load can struggle at twice that load. Ageing equipment raises λ and therefore ρ , which can justify a second crew sooner than average figures suggest.

IV. Cascading Failures

4.1 Load Sharing as a Chain Reaction

Independent failure is a convenient simplification. When one machine stops, its work shifts to the others, which then run harder and wear faster. Reliability engineers describe this as load sharing, a standard source of dependent failures (Rausand & Høyland, 2004).

A simple encoding multiplies each survivor's failure rate by a stress factor that grows with the number of failures:

$$\lambda_n = (N - n)\lambda(1 + \gamma n), \gamma \geq 0 \dots (15)$$

At $\gamma = 0$ the model returns to the independent baseline. At $\gamma > 0$ each breakdown makes the next more likely. This single term represents the chain reaction of the title.

The stress term competes with the shrinking source. The factor $(N - n)$ falls as n rises, while $(1 + \gamma n)$ grows. Comparing λ_1 with λ_0 shows that the failure stream increases after the first breakdown exactly when $(N - 1)(1 + \gamma) > N$, that is, when

$$\gamma > 1/(N - 1) \dots (16)$$

For $N = 3$ the threshold is $\gamma = 0.5$. Our example therefore sits at the boundary, where the stream is flat after the first failure instead of falling.

4.2 The Example with Cascading

Keep $N = 3$, $\rho = 0.5$, and $R = 1$, and set $\gamma = 0.5$. In units of μ , the upward rates are $\lambda_0 = 3(0.5) = 1.5$, $\lambda_1 = 2(0.5)(1.5) = 1.5$, and $\lambda_2 = 1(0.5)(2) = 1.0$. Equation (3) gives $w_1 = 1.5$, $w_2 = 1.5 \times 1.5 = 2.25$, and $w_3 = 2.25 \times 1.0 = 2.25$, which sum with w_0 to 7.0. Then

$$L = (1.5 + 2(2.25) + 3(2.25))/7.0 = 12.75/7.0 = 1.821,$$

and availability falls to $1 - 1.821/3 = 39.3\%$. The probability that all three machines are down is $2.25/7.0 = 32.1\%$, compared with 15.8% in the independent case.

The repair shop is identical in both cases, and only the coupling differs. Availability falls by 13.3 points, about one quarter of the baseline value. Added crews recover part of the loss. With two crews, the downward rates become $\mu_2 = \mu_3 = 2\mu$, and the weights are 1.5, 1.125, and 0.5625. This gives $L = 1.298$ and $A = 56.7\%$. With three crews, the weights are 1.5, 1.125, and 0.375, giving $L = 1.219$ and $A = 59.4\%$. A plant whose history shows clustered breakdowns should therefore distrust the independent model, which overstates its availability.

4.3 Estimating γ and the Link to Network Cascades

The parameters can be estimated from a maintenance log. Let T_n be the total time spent with n machines down, and let F_n be the number of failures observed in those states. The per-machine failure rate in state n is $r_n = F_n/[(N - n)T_n]$. Under equation (15), $r_n = r_0(1 + \gamma n)$, so

$$\hat{\gamma} = (r_n/r_0 - 1)/n \dots (17)$$

Here r_0 is used to estimate λ . The symbol $\hat{\mu}$ is found by taking the count of completed repairs and dividing by the total time the crew is busy. Before we trust the linear form, the estimate of γ should look steady as n changes. If γ moves up or down, then we should switch to a different functional form.

The link to network models is mainly a way to think, not a strict match. Watts (2002) treated cascades as a threshold process. In that view, a unit fails when enough of its neighbors have already failed. Buldyrev et al. (2010) then studied how two networks interact. They found that a small random hit can lead to a fast breakdown when the networks are coupled. The parameter γ here works as a mean field stand in. It does not tell us which machine depends on which other machine. It only says that each extra failure makes the hazard higher for the machines that are still running. If we know the dependency graph, a network model would give clearer results. If we do not know it, equation (15) serves as a reasonable starting point.

V. Reneging from the Repair Queue

5.1 Model Definition

The baseline assumes that a failed machine waits as long as necessary. In practice, a machine may be cannibalized for parts, sent to an outside workshop, or written off after a long wait. Barrer (1957) and Ancker and Gafarian (1963) developed the classical treatments of impatient customers, and Baghel (2014) applied reneging to machine systems with limited spares.

Suppose each machine waiting in the queue, not in service, leaves at rate β . The number waiting is $\max(n - R, 0)$, so the downward rate becomes

$$\mu_n = \min(n, R)\mu + \max(n - R, 0)\beta \dots (18)$$

A failed machine now has two exits, repair or abandonment. In the chain below, an abandoned machine is replaced immediately by a working one, so N stays constant. This is our modeling choice. The chain does not price the cost of replacement, and it is not the limited-spares formulation of Baghel (2014).

5.2 The Example with Reneging

Set $N = 3$, $R = 1$, $\rho = 0.5$, and $\beta = 0.5\mu$, with $\mu = 1$. The downward rates are $\mu_1 = 1$, $\mu_2 = 1 + 0.5 = 1.5$, and $\mu_3 = 1 + 2(0.5) = 2$. The upward rates equal those of the baseline: 1.5, 1.0, and 0.5.

Equation (3) gives $w_1 = 1.5$, $w_2 = 1.5 \times 1.0/1.5 = 1.0$, and $w_3 = 1.0 \times 0.5/2 = 0.25$. Together with w_0 they sum to 3.75. The probabilities are $p_0 = 0.2667$, $p_1 = 0.4$, $p_2 = 0.2667$, and $p_3 = 0.0667$, so $L = (1.5 + 2.0 + 0.75)/3.75 = 1.133$.

This is lower than the baseline value of 1.421. A manager watching only the queue would conclude that the shop had improved.

The flows tell a different story. Completed repairs occur at rate $\mu(1 - p_0) = 0.733$. Abandonments occur at rate $\beta(p_2 + 2p_3) = 0.5(1.0 + 0.5)/3.75 = 0.200$. Their sum, 0.933, equals $\sum \lambda_n p_n = 3.5/3.75$, so the flows balance. The abandonment share is

$$\phi = 0.200/0.933 = 21.4\% \dots (19)$$

Roughly one failed machine in five leaves without being repaired.

5.3 Interpretation

A shorter line can mean workers are moving faster. It can also mean people are walking away. On a simple queue-length chart, the two cases can look the same. But they cost very different amounts. That is why the abandonment fraction, ϕ , should be tracked with L .

There are also a couple of other notes to keep in mind. Using a fixed β is just for convenience. In real life, how impatient someone feels can depend on how long that person has been waiting at that specific machine. Once you add that kind of timing, the Markov setup no longer fits well. You then need more general patience models (Barrer, 1957).

The availability figure used here is $1 - 1.133/3 = 62.2\%$. But this number also treats replaced units as if they were working. So you cannot line it up with the baseline 52.6% value unless you also assume that replacement does not add any cost.

VI. Pooled versus Dedicated Crews

6.1 The Pooling Question

Think about a plant that has four machines and two repair crews. In the dedicated setup, each crew is assigned two machines. A crew fixes only its own machines. In the pooled setup, the crews are not stuck to any fixed machines. Either repairer can work on any of the four machines. Baghel (2013) used this sort of comparison to study crew training plans in an M/M/R model. For the dedicated setup, it is treated as two separate systems. Each one has $N = 2$ and $R = 1$. For the pooled setup, it is treated as one system. That case uses $N = 4$ and $R = 2$. The machines, the repairers, and the load ρ are the same in both cases. The only real change is whether a repairer is allowed to switch and help with other machines.

6.2 The Calculation

Use $\rho = 0.5$. For each dedicated group, $w_1 = 2(0.5) = 1$ and $w_2 = 1 \times 1(0.5) = 0.5$. The sum with w_0 is 2.5, and $L = (1 + 2(0.5))/2.5 = 0.8$. Two groups give 1.6 machines down out of four, an availability of 60.0%.

For the pooled system, $w_1 = 2$, $w_2 = 2 \times 3(0.5)/2 = 1.5$, $w_3 = 1.5 \times 2(0.5)/2 = 0.75$, and $w_4 = 0.75 \times 1(0.5)/2 = 0.1875$. The sum with w_0 is 5.4375, and $L = (2 + 3 + 2.25 + 0.75)/5.4375 = 8.0/5.4375 = 1.471$.

Availability is 63.2%. Pooling reduces expected downtime from 1.600 to 1.471 machines, a reduction of 8.0%, and raises availability by 3.2 points. No staff were added. The gain comes from idle repairers being free to serve a neighboring queue.

6.3 Mechanism and Break-Even Penalty

The mechanism is easy to state. In the dedicated layout, one repairer can sit idle while the other group has a queue. Let n_1 and n_2 be the numbers down in the two groups. The dedicated service rate is $[\min(n_1, 1) + \min(n_2, 1)]\mu$, and the pooled rate is $\min(n_1 + n_2, 2)\mu$. For all n_1 and n_2 ,

$$\min(n_1 + n_2, 2) \geq \min(n_1, 1) + \min(n_2, 1) \dots (20)$$

For any configuration of failures, the pooled system therefore serves at least as fast as the dedicated one. Pooling carries a cost. A generalist trained on every machine type is more expensive and may be slower on each task. Represent this by lowering the pooled service rate to $\mu_g = \mu(1 + \delta)$, which raises the load to $\rho_g = \rho(1 + \delta)$. For $N = 4$ and $R = 2$, the pooled weights are $w_1 = 4\rho_g$, $w_2 = 6\rho_g^2$, $w_3 = 6\rho_g^3$, and $w_4 = 3\rho_g^4$. The pooled mean number down is then

$$L(\rho_g) = (4\rho_g + 12\rho_g^2 + 18\rho_g^3 + 12\rho_g^4)/(1 + 4\rho_g + 6\rho_g^2 + 6\rho_g^3 + 3\rho_g^4) \dots (21)$$

Setting $L(\rho_g) = 1.6$, the dedicated value, and solving numerically gives $\rho_g \approx 0.561$, so $\delta \approx 0.12$. The pooled crew can be about 12% slower per repair and still match the dedicated arrangement on mean machines down. Beyond that penalty, specialist speed outweighs the benefit of flexibility.

The result challenges a common assumption that specialization is always better because specialists work faster. When machines are few and failures arrive in bursts, idle time between repairs can matter as much as speed within a repair. These are separate levers, and the second is often overlooked.

6.4 Summary of Computed Cases

Table 1 collects every case computed so far. Each row uses only the parameters stated in the table.

Table 1: Computed Performance of the Example Systems ($\rho = 0.5$; $\gamma = 0.5$ for cascade; $\beta = 0.5\mu$ for reneging)

System	Crews	Mean machines down, L	Availability, A	P(all machines down)
N = 3, independent	1	1.421	52.6%	15.8%
N = 3, independent	2	1.036	65.5%	5.5%
N = 3, independent	3	1.000	66.7%	3.7%
N = 3, cascade	1	1.821	39.3%	32.1%
N = 3, cascade	2	1.298	56.7%	13.4%
N = 3, cascade	3	1.219	59.4%	9.4%
N = 3, reneging (abandoned units replaced)	1	1.133	62.2%	6.7%
N = 4, dedicated (two groups of N = 2)	2	1.600	60.0%	4.0%
N = 4, pooled	2	1.471	63.2%	3.4%

Source note: Values were computed from equations (3) to (10), (15), and (18), using the finite-source model of Palm (1947) as presented in Gross et al. (2008). They are model outputs, not measurements. The reneging row counts replaced units as working and is not directly comparable with the independent rows.

VII. Choosing the Number of Crews

7.1 A Cost Function

More crews reduce downtime but add cost. Let c_r be the cost per crew per unit time, and let c_d be the cost per machine down per unit time. The expected cost rate with R crews is

$$C(R) = c_r R + c_d L(R) \dots (22)$$

Crews should be added while each one saves more than it costs. The smallest R satisfying

$$c_d [L(R) - L(R + 1)] < c_r \dots (23)$$

is the cost-minimizing choice, provided the marginal reductions in L decrease as R grows. They do in all our examples. The left side of (23) is the marginal drop in mean downtime multiplied by the cost of downtime. The time units of λ , μ , c_r , and c_d must match, or the rule gives meaningless thresholds.

7.2 Diminishing Returns

From Section 3, $L(1) = 1.421$, $L(2) = 1.036$, and $L(3) = 1.000$. The marginal reductions are

$$L(1) - L(2) = 0.385 \text{ and } L(2) - L(3) = 0.036.$$

The second crew removes more than a third of a machine from the mean downtime. The third removes under four hundredths. The ratio is about ten to one. Figure 2 plots the independent and cascading cases.

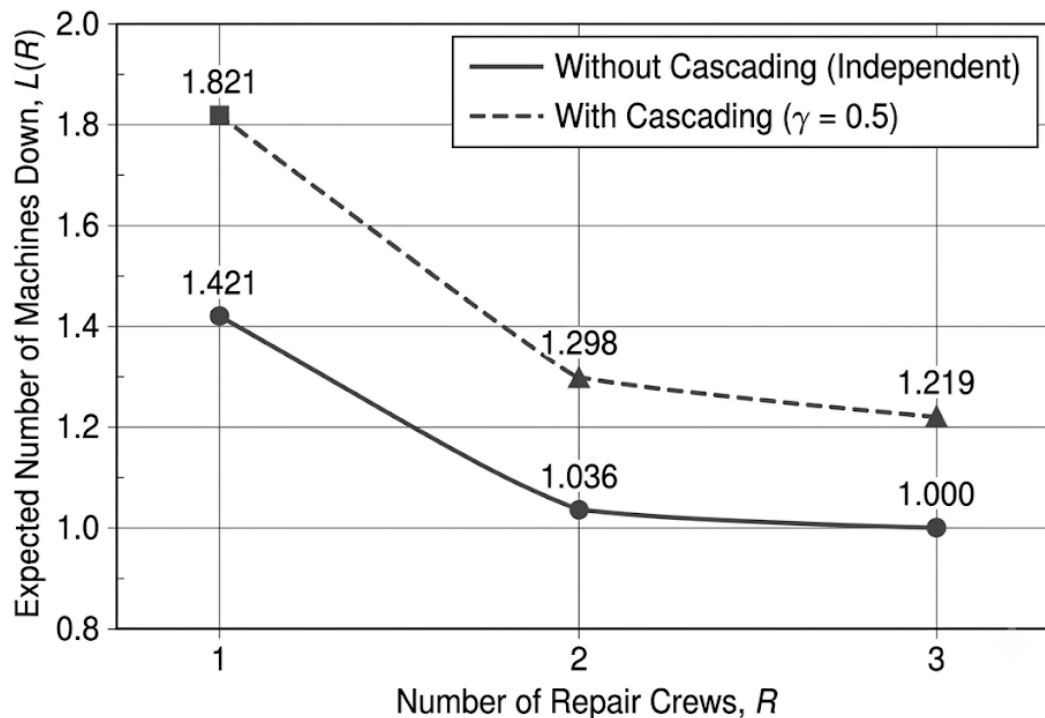


Figure 2: Mean Number of Machines Down, $L(R)$, versus Number of Repair Crews for $N = 3$ and $\rho = 0.5$, With and Without Cascading ($\gamma = 0.5$)

The graph uses the horizontal axis for the number of crews, with R taking values 1, 2, and 3. The vertical axis is the expected count of machines that are down. The first curve drops from 1.421 to 1.036 to 1.000. The cascading curve drops from 1.821 to 1.298 to 1.219. After R gets to N , both curves level off. In that region, the cascading curve sits around 0.22 machines higher than the independent curve.

The numbers come from the finite source model of Palm (1947), shown in Gross et al. (2008). The calculations use equations (3) to (8) and (15) along with the parameters that were given. These are model outputs, not data from tests.

The leveling off has a clear reason. When R equals N , no failed machine has to wait for service. So adding one more crew cannot further reduce L . In the independent case, equation (14) gives the value at that point, which is 1.000 for these settings. To push L lower, the load must be smaller, since $\rho = \lambda/\mu$. Practically, that means reducing λ , like through preventive maintenance or improved design, or increasing μ , meaning each repair takes less time. Just adding crews will not solve it.

7.3 Applying the Rule

By (23), the second crew is justified when $c_r < 0.385c_d$, and the third only when $c_r < 0.036c_d$. Suppose a machine down costs 1,000 per hour and a crew costs 300 per hour. The second crew pays for itself, since 385 exceeds 300, and the third does not, since 36 is below 300. These costs are placeholders. A plant's own data set the thresholds.

Cascading changes the thresholds. In the cascading case the marginal reductions are $1.821 - 1.298 = 0.523$ and $1.298 - 1.219 = 0.080$. The second crew is justified at any cost ratio below 0.523 instead of 0.385, and the third below 0.080 instead of 0.036. At the placeholder costs, the second crew now wins by a wider margin, 523 against 300, and the third still fails, 80 against 300. A plant that ignores cascades will understaff its maintenance team in the conditions where staffing matters most.

VIII. Limitations

The model has limits. It assumes time to run and time to fix follow exponential laws. Also, Bunday and Scraton (1980) changed that first part. They used the G/M/r setup, so run times do not have to be only one kind. Another strong simplification is that every machine is the same. Steckle and Aronson (1985) looked at cases that drop that idea. They covered nonidentical units, different priority rules, and walking time between stations. These things show up on real shop floors. There is more realism in vacation models too. In those models, the repair person can be away at times. Ke et al. (2010) discuss this. We do not take that case here.

The model also keeps key parameters fixed. In practice, wear can change, shift schedules vary, and seasonal heat can matter. Those effects move λ and μ through time. The cascade piece grows in a linear way with n . The reneging rate stays the same. Because of that, neither part shows the full set of timing effects or how impatience might act.

These choices are made so the results stay simple enough to check by hand. That is why the conclusions only apply within the stated scope.

IX. Conclusion

When several machines fail at once, it is not just a bigger copy of one breakdown. It is a full system, with rules of its own. Those rules can give people clear answers if they use the model with care.

We started with a finite source repair setup. In that model, the steady chances come from multiplying rate ratios. The availability comes from the average count of machines that are down. In one test with three machines, one crew mode led to 52.6% availability. When two crew modes were used, the result was 65.5%. When a third crew was added, the gain was small, about 1.2 points. The main reason was that the count of machines down had nearly hit its low bound for the given failure and repair rates.

After that, two changes made the results look different. With load sharing, the parameter γ linked the failures. What was once separate became a knock-on chain. In that case, single-crew availability fell to 39.3%. Then we added reneging, set by β . It pushed the mean number of machines down from 1.42 to 1.13. It also meant that 21.4% of failed machines left the waiting line before repair. These effects can be missed if you only look at the average queue length. So it helps to compute the full steady state distribution.

Pooling two repairers over four machines lowered the expected outage by 8.0%. No extra workers were added. With pooling, a crew could also be around 12% less fast per job before the gain disappeared.

Equation (23) set a quick yardstick for how many people to assign. After that, cascading effects made each extra crew count more than you would expect at first.

Use the method in a simple order. First, estimate λ and μ from your maintenance logs. Next, check if failures clump in time by fitting γ . Also watch abandonment, not just how long you wait. After you have those pieces, calculate what staffing changes do before you act.

Keep in mind the limits of the examples. They are small, and they are not real plant records. If you move to bigger sites, use non exponential job times, or build dependency graphs by hand, you will need stronger models. Still, the core steps remain: set the move rates, solve the balance equations, and take the performance numbers from the result.

References

- [1]. Ancker, C. J., Jr., & Gafarian, A. V. (1963). Some queuing problems with balking and reneging. I. *Operations Research*, 11(1), 88–100. <https://doi.org/10.1287/opre.11.1.88>
- [2]. Baghel, K. P. S. (2013). Generalists vs. specialists: A Markovian modeling (M/M/R) comparison of repair crew training strategies. *Journal of Research in Applied Mathematics*, 1(1), 10–15. <https://doi.org/10.35629/0743-01011015>
- [3]. 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*, 2(1), 56–61. <https://doi.org/10.35629/4767-02015661>
- [4]. Barlow, R. E., & Proschan, F. (1965). *Mathematical theory of reliability*. Wiley.
- [5]. Barrer, D. Y. (1957). Queuing with impatient customers and ordered service. *Operations Research*, 5(5), 650–656. <https://doi.org/10.1287/opre.5.5.650>
- [6]. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291), 1025–1028. <https://doi.org/10.1038/nature08932>
- [7]. Bunday, B. D., & Scraton, R. E. (1980). The G/M/r machine interference model. *European Journal of Operational Research*, 4(6), 399–402. [https://doi.org/10.1016/0377-2217\(80\)90121-3](https://doi.org/10.1016/0377-2217(80)90121-3)
- [8]. Cooper, R. B. (1981). *Introduction to queueing theory* (2nd ed.). North Holland.
- [9]. Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2008). *Fundamentals of queueing theory* (4th ed.). Wiley.
- [10]. Haque, L., & Armstrong, M. J. (2007). A survey of the machine interference problem. *European Journal of Operational Research*, 179(2), 469–482. <https://doi.org/10.1016/j.ejor.2006.02.036>
- [11]. 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.
- [12]. 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(1), 45–56.
- [13]. Jain, M., & Maheshwari, 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.2007.02.007>
- [14]. Jain, M., & Preeti. (2009). N-policy for M/G/1 machine repair model with mixed standby components, degraded failure and Bernoulli feedback. *Journal of King Abdulaziz University-Engineering Sciences*, 21(2), 23–45.
- [15]. Jain, M., Sharma, G. C., & Baghel, K. P. S. (2002). Performance prediction of a flexible manufacturing system. *International Journal of Engineering (IJE), Transactions A: Basics*, 15(4), 355–364.
- [16]. Ke, J.-C., Wu, C.-H., & Zhang, Z. G. (2010). Recent developments in vacation queueing models: A short survey. *International Journal of Operations Research*, 7(4), 3–8.
- [17]. Kleinrock, L. (1975). *Queueing systems: Vol. 1. Theory*. Wiley.
- [18]. Little, J. D. C. (1961). A proof for the queueing formula: $L = \lambda W$. *Operations Research*, 9(3), 383–387. <https://doi.org/10.1287/opre.9.3.383>

- [19]. Medhi, J. (2003). *Stochastic models in queueing theory* (2nd ed.). Academic Press.
- [20]. Palm, C. (1947). The distribution of repairmen in servicing automatic machines (in Swedish). *Industritidningen Norden*, 75, 75–80, 90–94, 119–123.
- [21]. Rausand, M., & Høyland, A. (2004). *System reliability theory: Models, statistical methods, and applications* (2nd ed.). Wiley.
- [22]. Ross, S. M. (2014). *Introduction to probability models* (11th ed.). Academic Press.
- [23]. Sharma, G. C., Shinde, V., & Jain, M. (2004). Performance modelling of machining system with mixed standby components, balking and reneging. *Journal of King Abdulaziz University-Engineering Sciences*, 16(1), 3–20.
- [24]. Stecke, K. E., & Aronson, J. E. (1985). Review of operator/machine interference models. *International Journal of Production Research*, 23(1), 129–151. <https://doi.org/10.1080/00207548508904696>
- [25]. Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9), 5766–5771. <https://doi.org/10.1073/pnas.082090499>