Industrial Systems Engineering | IE3100R Systems Design Project

Modelling of AMR Routes for Delivery within New Emergency/NNI Building (EB)

Group 21
Subhash Nair Anandh
Koh Guo Xuan
Phylis Leung Ning Xin
Gan Bei Fang
Hou Yuxuan

NUS Supervisor Li Haobin
SGH Supervisor Cheryl Quek Mei Xian
SGH Supervisor Estee Chan Ching Lian

NUS National University of Singapore
Singapore General Hospital SingHealth

INTRODUCTION

Background

As Singapore General Hospital (SGH) expands its campus with the new Emergency/NNI Building (EB), centralising emergency care and diagnostics has become a priority. Currently, the internal logistics for medication, consumables, and admission folders rely heavily on manual portering. This labor-intensive process is increasingly difficult to scale alongside growing hospital delivery volumes. To ensure operational scalability, SGH is looking to integrate Autonomous Mobile Robots (AMRs) to provide a more reliable logistics framework.

Objective

This project develops a data-driven simulation framework to evaluate AMR feasibility in augmenting SGH's logistics within the new Emergency/NNI Building. A Python-based discrete event simulation replicates patient crowd arrivals and translates them into downstream task generation across three key delivery use cases. The objectives are as follows :

- Model and simulate the end-to-end delivery process across three key hospital use cases.
- Optimise AMR routing and dispatch to minimise turnaround times and maximise fleet utilisation.
- Quantify and statistically validate performance under stochastic operational constraints (Eg. lift congestion and human crowd delays) and stress-test configurations to produce a robust, contingency-aware deployment recommendation for SGH.

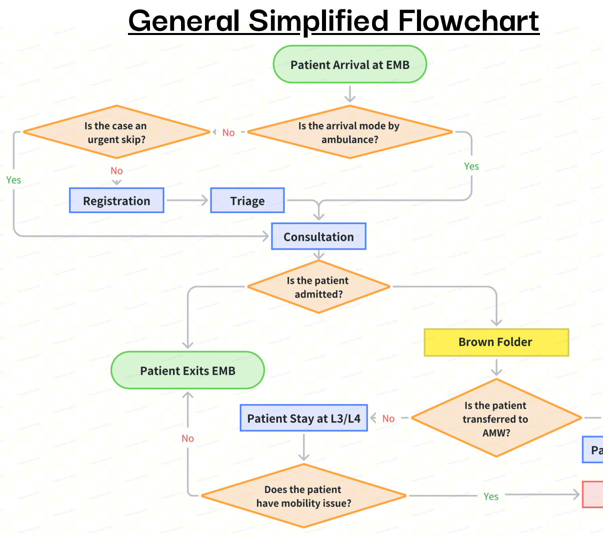
KEY CONCEPTS

- Discrete Event Simulation
- Machine Learning
- Operation Research
- Statistical Validation & Testing
- Process Control & Design
- Data Analysis
- Data Structure & Algorithms
- Mathematical Statistics
- Healthcare Systems Engineering

METHODOLOGY - DISCRETE EVENT SIMULATION

A discrete-event simulation was used to replicate the patient journey at the individual level, **enabling realistic task generation and crowd-induced AMR traversal delays** to arise endogenously from modelled patient flow.

General Simplified Flowchart

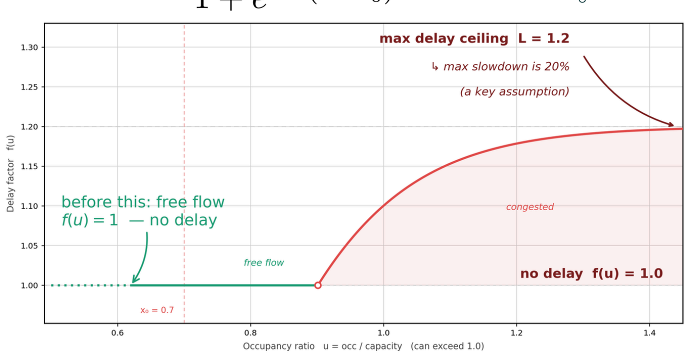


Crowd Delay - Sigmoid

A sigmoid function governs AMR traversal delay as a function of corridor occupancy, proxied from nearby clinical service nodes. Delay scales non-linearly as zone utilisation approaches capacity.

$$f(u) = L \cdot \frac{1}{1+e^{-k(u-x_0)}}$$

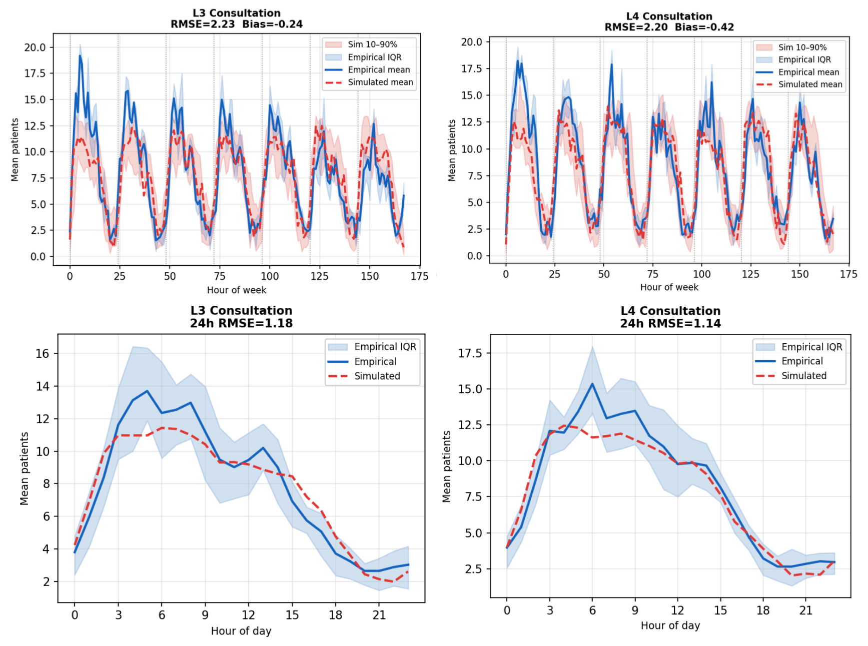
where L = 1.2, k = 8 & X_0 = 0.7



Once server occupancy (U) > ~90%, congestion will start to kick in. Below 0.901, the model treats the zone as free flow.

Calibrating Service Times – High-Dimensional Inverse Optimisation

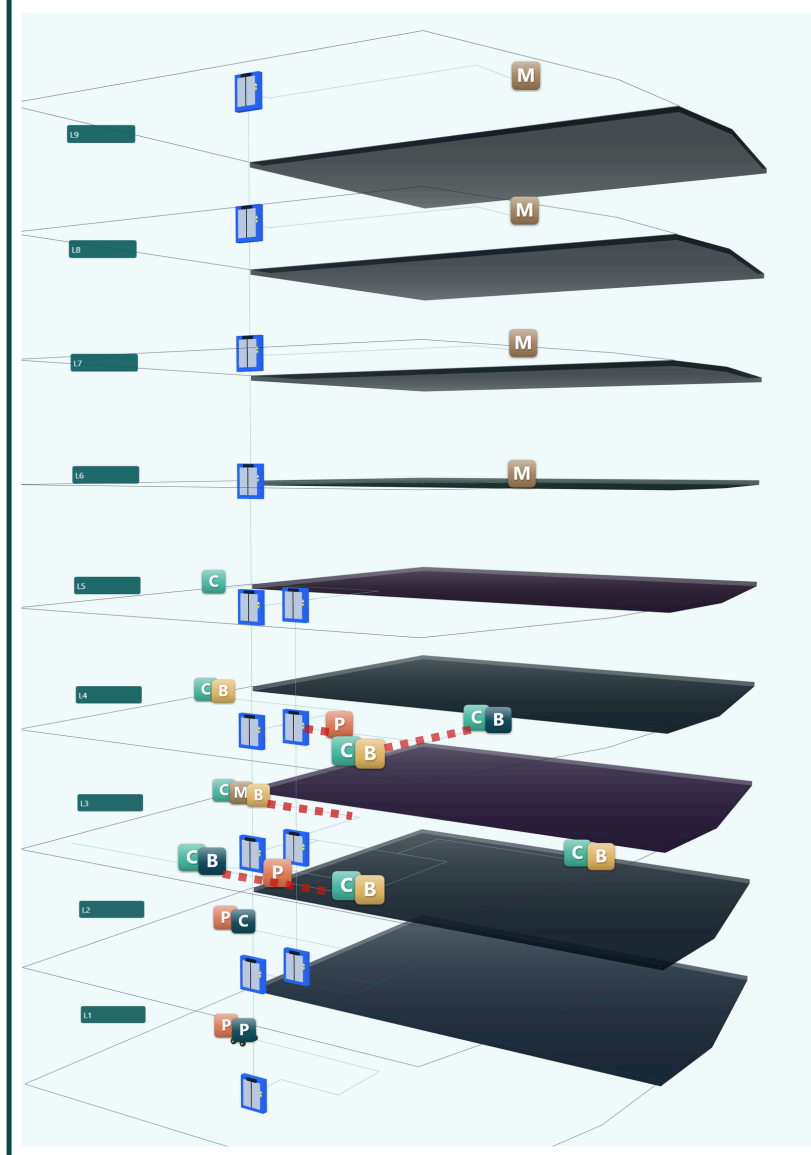
Since ground-truth service times are not recorded in the hospital's operational data, consultation service time parameters were recovered through a calibration procedure. A grid search was conducted over the lognormal parameters (μ, σ) for registration, triage, and consultation service times, and jointly over (μ, σ) and re-consultation probability p for consultation, evaluated separately for each floor. Parameter estimates were selected by minimising the RMSE between simulated and empirical hourly zone occupancy profiles aggregated across the calibration weeks.

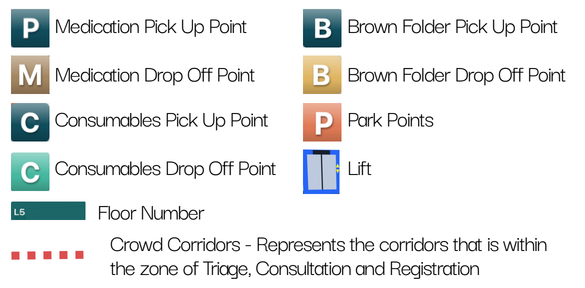


The calibrated simulation reproduces the diurnal structure of consultation zone occupancy to within a maximum of 8–9% relative error at peak, ensuring the downstream logistics model correctly replicates both AMR demand timing and crowd-induced congestion penalties. Residual deviations are attributable to stochastic seed variability rather than systematic misspecification.

METHODOLOGY - PYTHON MODEL

We developed a Python-based digital twin of the SGH EB floor plan to simulate complex hospital logistics. To ensure high-fidelity results, we **coupled the delivery workflows for Brown Folders (B), Consumables (C), and Medications (M) with real-time patient demand generated by the crowd model** (see below).

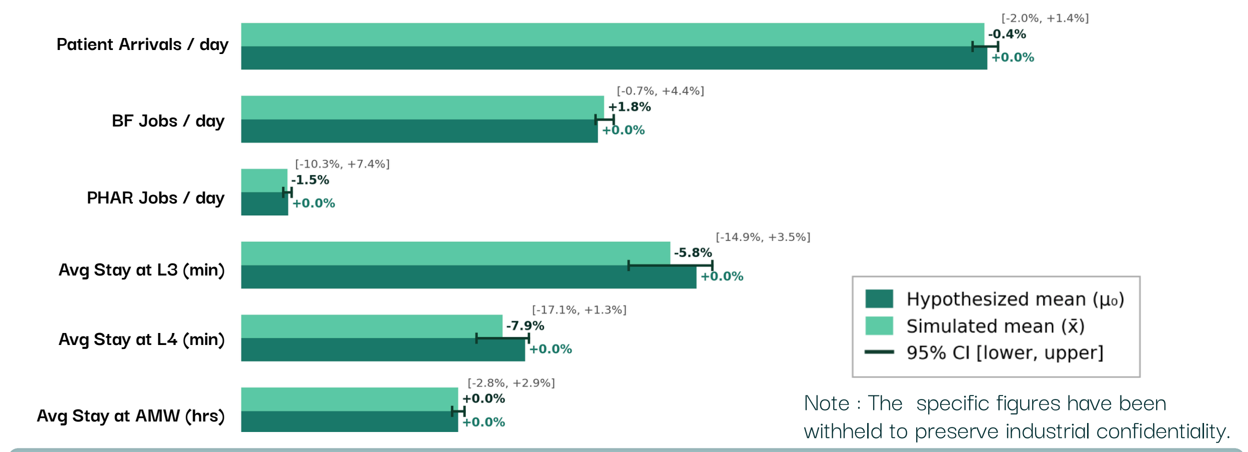




Patient arrivals at the EB are governed by a Non-Homogeneous Poisson Process (NHPP), capturing the time-varying nature of emergency demand throughout the day. Each patient progresses through registration, triage, and consultation at either L3/L4, with real-life conditional transition probabilities – conditioned on arrival mode, triage tag, and stage skipping – governing how patients propagate and eventually shaping their probability of admission. Admitted patients are assigned to L3, L4, or the Acute Medical Ward (AMW) on L6–L9. This journey endogenously drives AMR task generation: **brown folder deliveries are triggered upon admission confirmation, while medication runs are triggered by AMW transfers and selectively for mobility-limited L3/L4 patients**. The NHPP thus serves as the primary engine of the simulation, simultaneously governing crowd dynamics across the EB and AMR delivery workloads from patient flow.

MODEL VALIDATION

The model's outputs successfully pass hypothesis testing, with all p-values exceeding 0.05, suggesting that there is a lack of statistical evidence that the simulation does not accurately replicate real-world operational patterns.



Note : The specific figures have been withheld to preserve industrial confidentiality.

Patient Arrivals	BF Jobs	PHAR Jobs	Avg Stay L3	Avg Stay L4	Avg Stay AMW
P = 0.81	P = 0.15	P = 0.71	P = 0.11	P = 0.089	P = 0.99

PROCESS CONTROL EXPERIMENTATION - MEAN/VARIANCE MODELLING

Mean and variance regression models were fitted for all three delivery turnaround times ($R^2 \geqslant 0.89$), enabling systematic identification of the optimal AMR configuration and contingency plan under stressed operating conditions.

Model	Fitted equation	R²
Brown folder turnaround		
Mean	$\hat{\mu}_{BF} = 15.55 + 6.18A - 6.58B_{lin} + 2.92B_{quad} + 0.76D + 2.38AB_{lin} - 1.34B_{lin}D$	0.971
Variance	$\ln(\hat{s}^2)_{BF} = 0.80 + 0.54A - 2.57B_{lin} - 0.52C + 0.87D - 0.42CD + 1.21AB_{lin}$	0.889
Consumables turnaround		
Mean	$\hat{\mu}_{CONS} = 37.37 + 4.77A - 17.85B_{lin} + 8.85B_{quad} + 1.76D$	0.965
Variance	$\ln(\hat{s}^2)_{CONS} = 4.21 + 0.29A - 0.96B_{lin} + 0.35B_{quad} + 0.97D - 0.20AD + 0.32AB_{lin}$	0.931
Pharmacy turnaround		
Mean	$\hat{\mu}_{PHAR} = 27.75 + 7.43A - 11.86B_{lin} + 4.21B_{quad} + 1.82D + 2.63AB_{lin}$	0.967
Variance	$\ln(\hat{s}^2)_{PHAR} = 3.26 + 0.16A - 0.79B_{lin} + 0.43B_{quad} - 0.46C + 1.29D - 0.29CD + 0.37AB_{lin}$	0.972

FACTOR ENCODING : A = dispatch (fifo = -1, weight_time = +1) · B_{lin} = fleet size linear contrast (2 AMRs = -0.707, 3 AMRs = 0, 4 AMRs = +0.707) · B_{quad} = fleet size quadratic contrast (2 AMRs = +0.408, 3 AMRs = -0.816, 4 AMRs = +0.408) · C = crowd (low = -1, high = +1) · D = lift (low = -1, high = +1)

Key Source of Variability

Lift congestion remains the primary robustness boundary no AMR configuration can fully neutralise. Simulation reveals an average of ~9 medication jobs arising during a full lift breakdown – **a single porter deployment serves as a cost effective workaround to ultimately deliver a robust and reliable logistics system.**

Legend

- 2 AMR : 1 L3_BF, 1 GEN
- 3 AMR : 1 L3_BF, 2 GEN
- 4 AMR : 1 L3_BF, 1 L4_BF, 2 GEN

Review on AMR Deployment

The positive B_quad term confirms diminishing returns: most performance gain occurs in the 2→3 AMR step, with 3→4 providing only incremental improvement. This translates to a clear performance dislocation between 2-AMR and 3/4-AMR configurations. 2 AMR delivers roughly 2× higher mean turnaround times across all delivery types and catastrophic variance under lift congestion – **2 AMRs is not a viable deployment option**. Given that 3 AMR and 4 AMR deliver broadly comparable performance across all delivery types, our team **recommends 4 AMRs as the primary deployment plan with 3 AMRs as the contingency plan.**

OUTPUT ANALYSIS

From workload to dispatching to allocation, each level builds toward system performance evaluation.

Example of Config (BF_L3 + 2 GEN): 1 AMR assigned to Level 3 to handle Brown Folders Tasks + 2 AMR assigned to handle all pending tasks

Level 1: Scenario

Mixed Workload (All job types generated)

Individual Cases
- Medicine Jobs
- Consumables Jobs
- Brown Folder Jobs

Level 2: Dispatching

FIFO: first come, first serve

Weight - Time Prioritise < 15 min

Priority Based
- **Consumables** first
- **Medicine** first
- **Brown Folders** first

Level 3: Allocation Strategies

Fleet Size Only All AMRs serve any job type 4,3,2,1

Tagged use case mix AMR is either assigned to a use case / used across different use cases

Fixed Scenario Dedicated AMRs / job category

Highlighted cases : Allocations that give the best performance and shortlisted for statistical validation.

#	CONFIGURATION	POLICY	AMRS	AVG TAT	MEDS TAT	BF TAT	CONS TAT	MAX TAT	DELAY	FLEET UTIL	UTIL RANGE	LIFT-FREE
1	4R: 4 roaming	Weight-time	4	26.7	32.2	20.9	32.0	298	593k	17.5%	2.8pp	0/4
2	4R: BF_L3+3GEN	Weight-time	4	24.9	31.8	20.8	32.3	298	608k	17.4%	4.8pp	1/4
3	3R: BF_L3+2GEN	FIFO	3	16.4	17.4	6.6	25.1	261	1054k	36.6%	9.3pp	1/3
4	3R: BF_L3+2GEN	Weight-time	3	26.3	36.5	21.3	37.1	300	604k	23.5%	4.2pp	1/3
5	3R: BF_L3+2GEN	BF-first	3	16.8	18.6	6.1	26.0	256	1041k	36.6%	9.3pp	1/3
6	4R: BF_L3+BF_L4+2GEN	FIFO	4	14.4	14.6	4.6	24.0	263	916k	27.1%	22.3pp	2/4
7	4R: BF_L3+BF_L4+2GEN	Weight-time	4	26.3	32.9	20.6	35.5	321	523k	16.6%	13.6pp	2/4
8	4R: BF_L3+BF_L4+2GEN	BF-first	4	14.4	14.8	4.4	24.2	263	916k	26.9%	22.3pp	2/4
9	3R: BF_L3+BF_L4+GEN	FIFO	3	18.7	24.6	7.7	43.9	328	712k	31.3%	41.1pp	2/3
10	3R: BF_L3+BF_L4+GEN	Weight-time	3	28.0	41.6	21.5	56.0	357	487k	21.3%	27.4pp	2/3

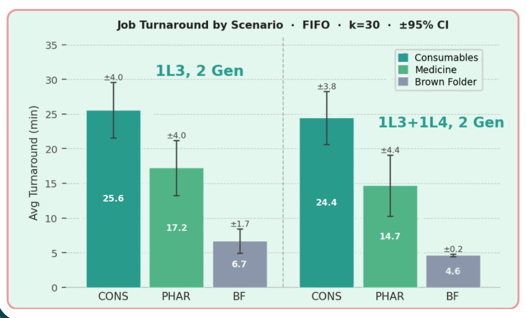
1 L3, 2 General (Utilisation Rate)

	08-11	11-14	14-17	17-20	20-23	23-02	02-05	05-08
1	20.3%	33.1%	37.8%	40.0%	38.3%	38.3%	26.7%	10.9%
2	30.9%	54.0%	50.6%	56.7%	54.7%	40.8%	20.9%	8.0%
3	26.4%	56.2%	54.1%	57.9%	53.7%	36.9%	20.1%	6.4%
4	25.7%	46.8%	49.7%	51.5%	49.1%	38.4%	21.1%	9.0%

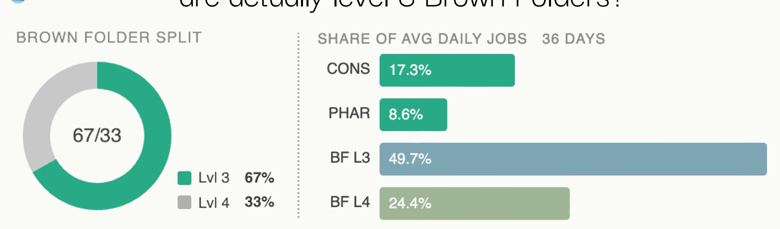
1 BF L3 #1 Brown Folder Level 3 · 2 GEN #2 General AMR · 3 GEN #3 General AMR · 4 Fleet Avg All AMRs combined

1 L3, 1 L4, 2 General (Utilisation Rate)

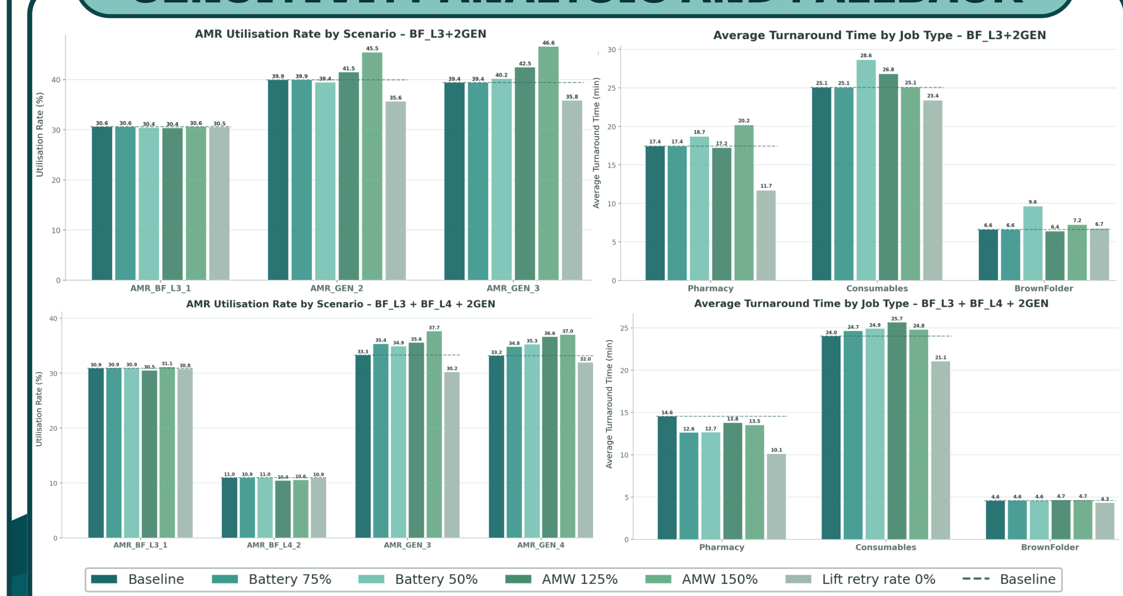
	08-11	11-14	14-17	17-20	20-23	23-02	02-05	05-08
BF L3 #1 Brown Folder Level 3	20.2%	34.9%	37.7%	41.2%	38.3%	37.9%	26.8%	10.9%
BF L4 #2 Brown Folder Level 4	8.0%	15.8%	15.0%	15.7%	14.3%	10.6%	3.8%	3.4%
GEN #3 General AMR	24.1%	47.8%	42.2%	48.2%	43.4%	32.3%	20.7%	5.1%
GEN #4 General AMR	21.8%	49.0%	44.7%	45.8%	42.9%	34.4%	18.3%	5.3%
Fleet Avg All AMRs combined	18.9%	35.5%	36.6%	37.4%	34.8%	28.4%	16.1%	6.7%



Fun fact! : Do you know that ~50% of all jobs generated are actually level 3 Brown Folders?



SENSITIVITY ANALYSIS AND FALLBACK



- Battery degradation & AMW admission rate → minimal impact → Robust system
- Eliminating Lift Retry rate → Significant reduction in Turnaround time
- BF_L3+BF_L4+2GEN demonstrates lower sensitivity to parameter variations vs BF_L3+2GEN, further reinforcing our overall reccomendation.

CONCLUSION AND RECOMMENDATION

Optimal Fleet Configuration: BF_L3 + BF_L4 + 2 GEN under FIFO with 1 Porter Fallback Simulation results across nominal and disruption scenarios identify this as the optimal fleet configuration. The floor-dedicated BF agents sustain brown folder throughput on L3 and L4 independently of vertical infrastructure, while the two GEN AMRs handle pharmacy and consumables demand in parallel – eliminating cross-task contention under surge conditions. The porter complement closes the resilience gap: mid-mission battery failures and lift disruptions are recovered with minimum task loss. Even under fleet reduction, the system degrades gracefully to BF_L3 + 2 GEN, which continues to sustain acceptable performance. This configuration is recommended for its throughput, disruption resilience, and operational simplicity.