Simen Owen Academic Proceedings Series

4th International Conference on Media, Economy, Communication and Intelligence Management (MECI 2026)

Article

Analysis of Abnormal Variations in Pedestrian Flow Intensity and Probabilistic Estimation of Retail Profit–Loss Risk in Urban Commercial Clusters Using Poisson Jump Processes and Monte Carlo Stochastic Simulation

Mo Xiao[1,*]

[1] Heilongjiang Experimental High School International Department, Harbin, China
* Correspondence: Mo Xiao, Heilongjiang Experimental High School International Department, Harbin, China

Abstract: Pedestrian-flow fluctuations in urban commercial clusters directly affect customer exposure and short-term operating risk, yet conventional count and time-series models often absorb abrupt changes into residual variation and remain disconnected from probabilistic retail-loss assessment. This study develops a seasonally adjusted compound Poisson jump framework coupled with Monte Carlo risk propagation to distinguish recurrent pedestrian intensity from discrete abnormal changes and translate flow shocks into conditional shop-level profit–loss risk. Hourly pedestrian counts and spatial business-establishment data from Melbourne are synchronized to estimate baseline intensity, jump frequency, signed magnitude, and decay, after which uncertain conversion rates, transaction values, margins, and fixed costs are propagated through 100,000 Monte Carlo paths. On the 2024 out-of-sample evaluation period, the proposed model reduced MAE by 8.3%, negative log-likelihood by 3.9%, and CRPS by 7.4% relative to the strongest benchmark, with all differences significant at the 5% level. The pooled operating-loss probability increased from approximately 0.23 under regular conditions to 0.38 under negative-jump states, while normalized 95% CVaR increased from 0.78 to 1.17 fixed-cost units under strong negative shocks. At 100,000 paths, the Monte Carlo standard error was approximately 0.0015. These results link interpretable pedestrian-flow discontinuities with uncertainty-aware retail risk measures and support probabilistic assessment of commercial exposure under abnormal urban-flow conditions.

Keywords: compound Poisson jump process; pedestrian flow; urban commercial clusters; Monte Carlo simulation; retail loss risk

Received: 08 September 2026
Revised: 22 October 2026
Accepted: 01 November 2026
Published: 05 November 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Urban commercial clusters concentrate retail, catering, services, transport access, and public activity within relatively small spatial areas, making pedestrian intensity an important indicator of short-term commercial exposure. Recent evidence shows that customer flow in commercial districts is jointly influenced by business scale, format diversity, accessibility, location, and surrounding population composition [1]. Continuous urban sensing has also increased the availability of pedestrian-count observations and strengthened the empirical basis for analysing time-varying movement patterns [2]. However, pedestrian flow is not a stationary sequence: recurrent hourly, weekday–weekend, and seasonal structures coexist with abrupt changes associated with events, weather disturbances, transport disruptions, holidays, and other external shocks. Recent analysis of hourly footfall data demonstrates that aggregate pedestrian activity

contains distinct temporal signatures and can exhibit marked structural shifts under external disturbances [3].

For retailers, such variation is economically relevant because the relationship between pedestrian exposure and profit is generally nonlinear. A 10% decline in footfall does not necessarily imply a 10% decline in profit: conversion rates, average transaction values, contribution margins, and fixed operating costs jointly determine whether reduced exposure remains above the break-even level. Consequently, average-flow forecasting alone cannot answer a more decision-relevant question: what is the probability that a shop becomes loss-making when pedestrian intensity experiences an abnormal downward jump? Existing smooth temporal models, including seasonal autoregression, Poisson regression, and negative-binomial specifications, can represent periodicity and over-dispersion, but abrupt changes are often absorbed into residual variation rather than parameterized through an explicit jump-arrival rate, jump magnitude, and persistence mechanism. Compound Poisson formulations provide a mathematically tractable basis for representing discrete stochastic shocks and their accumulated effects [4].

A second gap is the separation between pedestrian-flow modelling and retail-risk analysis. Footfall studies commonly terminate at count accuracy or likelihood-based prediction, whereas operational-risk studies focus on demand, inventory, revenue, or cost uncertainty. Recent risk-averse operations research demonstrates the importance of modelling ambiguous demand distributions together with downside-risk measures such as conditional value-at-risk [5]. A third limitation is that point predictions do not propagate uncertainty in customer conversion, transaction value, margins, and operating cost into an interpretable probability of loss; retail decision models likewise show that expected profit and conditional tail-risk criteria can lead to materially different assessments under uncertain demand [6].

Accordingly, this study makes three experimentally testable contributions: (1) It develops a seasonally adjusted compound Poisson jump model that separates recurrent pedestrian intensity from discrete abnormal increases and decreases, with jump frequency, magnitude, and decay estimated explicitly. (2) It constructs a commercial-exposure and stochastic cash-flow coupling mechanism that translates pedestrian states into conditional shop-level profit and loss rather than treating footfall as revenue directly. (3) It propagates jump, conversion, transaction-value, margin, and cost uncertainty through Monte Carlo sampling to estimate loss probability, VaR, CVaR, confidence intervals, and Monte Carlo error.

The technical route proceeds from hourly pedestrian observations through temporal–spatial synchronization, seasonal baseline estimation, jump inference, business-exposure matching, stochastic cash-flow mapping, and Monte Carlo risk propagation. The framework is evaluated through jump-characteristic analysis, out-of-sample model comparison, mechanism ablation, cross-zone robustness tests, and simulation-convergence analysis. Its main academic value lies in linking discontinuous urban-mobility dynamics to probabilistic retail risk within one interpretable stochastic framework, while its practical value lies in providing uncertainty-aware indicators without requiring individual-level pedestrian identification.

2. Related Works

2.1. Statistical Modelling of Urban Pedestrian Flow

Statistical models remain widely used for pedestrian-flow analysis because count-based formulations are computationally efficient, interpretable, and naturally aligned with non-negative observations. Poisson and negative-binomial regression can represent mean intensity and over-dispersion, while seasonal time-series specifications capture recurrent hourly, weekday–weekend, and monthly structures. Recent urban pedestrian studies show that temporal regularity coexists with substantial spatial heterogeneity across streets and commercial environments [7]. More recent urban-flow models improve short-horizon accuracy by learning periodic and dynamic spatio-temporal dependence [8].

Sensor-based footfall research also demonstrates that weather, built-environment characteristics, and local context affect observed pedestrian intensity [9]. These approaches are useful for prediction, but their main objective is usually conditional-mean or distributional forecasting under gradually evolving dependence. Abrupt deviations are therefore commonly represented through enlarged residuals, adaptive normalization, or latent temporal features rather than an explicit stochastic mechanism describing jump timing, magnitude, and persistence.

2.2. Point Processes and Discontinuous Demand Modelling

Point-process methods provide a different representation by treating irregular events as realizations of an intensity-driven stochastic process. Poisson processes offer tractable arrival models, compound Poisson formulations attach random magnitudes to arrivals, and jump-state models permit discontinuous changes in an otherwise smoother process. Hawkes processes extend this logic by allowing previous events to increase subsequent conditional intensity; recent spatio-temporal applications demonstrate their ability to separate background intensity from event-triggered excitation while retaining periodic components [10]. Extended Hawkes formulations have also been used to represent disruption, recovery, and cascading mobility responses under externally disturbed conditions [11]. In operations settings, Poisson arrival processes remain useful for connecting stochastic demand occurrences with inventory states and shortage costs [12]. Nevertheless, self-exciting formulations are best suited to clustered or mutually triggering events. For commercial footfall, where the objective is to distinguish recurring seasonal intensity from discrete exogenous shocks, a compound Poisson jump representation offers a more direct decomposition of arrival frequency, signed magnitude, and post-shock decay.

2.3. Probabilistic Retail and Commercial Risk Assessment

Retail and operations research increasingly recognizes that expected demand or expected profit alone can obscure downside exposure. Distributionally robust retail models incorporate incomplete demand information and conditional tail-risk criteria to balance expected performance against adverse outcomes [13]. Related inventory-risk research shows that mean-CVaR formulations can materially change operating decisions when service delays or demand uncertainty generate asymmetric costs [14]. Monte Carlo experiments are also used to evaluate uncertain demand realizations and compare operating-policy robustness when analytical distributions are insufficient [15]. Their common limitation for the present problem is that demand uncertainty is typically specified exogenously through assumed distributions, scenarios, or historical residuals. They therefore quantify the consequences of uncertain demand without identifying an upstream urban-mobility process that generates the uncertainty.

2.4. Critical Comparison and Research Gap

The three research streams solve adjacent but incomplete parts of the same problem. Pedestrian models estimate urban flow but usually terminate at forecasting accuracy; point-process models characterize stochastic arrivals but seldom translate pedestrian discontinuities into commercial exposure; and retail-risk models evaluate uncertain cash flow while generally taking demand distributions as given. A missing link is therefore a coherent probabilistic chain in which pedestrian discontinuity is estimated first and then propagated into economic exposure. This study addresses that gap through jump intensity, commercial exposure, stochastic cash flow, and tail risk. Seasonal intensity and compound Poisson jumps generate demand-side uncertainty endogenously; business composition maps pedestrian states to exposure; and Monte Carlo propagation converts uncertainty in jumps, conversion, transaction value, margin, and cost into loss probability, VaR, and CVaR. This design preserves interpretability at each stage while allowing the final risk estimate to reflect both mobility shocks and commercial uncertainty.

3. Methodology

3.1. Overall Architecture

The proposed Poisson Jump–Monte Carlo framework integrates pedestrian-flow anomaly modelling with conditional retail profit–loss risk estimation. As illustrated in Figure 1, the framework contains six sequential modules: Pedestrian Count and Business Data Input, Temporal–Spatial Synchronization, Seasonal Baseline Intensity Estimation, Compound Poisson Jump Layer, Commercial Exposure and Cash-Flow Mapping, and Monte Carlo Risk Estimation. A separate Parameter Updating and Convergence Diagnostics module feeds estimation results back to the baseline-intensity and jump-process components until the likelihood and parameter values stabilize.

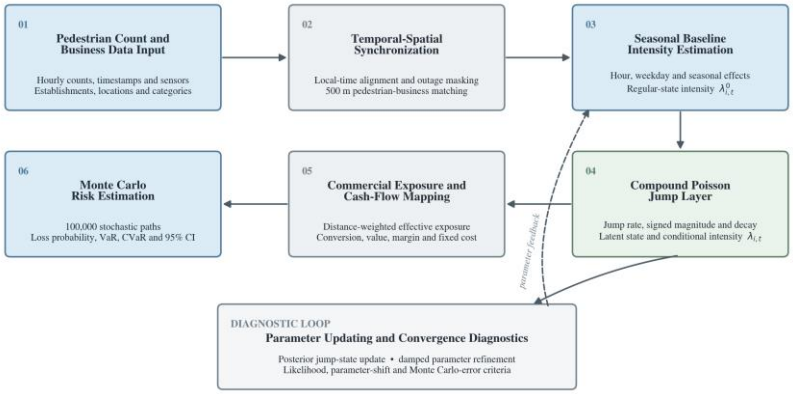


Figure 1. Overall Architecture of the Poisson Jump–Monte Carlo Framework for Pedestrian-Flow Anomaly Modelling and Retail Loss-Risk Estimation.

Hourly pedestrian records are first aligned by sensor identifier and local timestamp. Commercial establishments are geocoded to the same spatial reference system and associated with nearby pedestrian sensors using a fixed catchment radius. The modelling structure is statistical rather than neural: the observed hourly count $Y_{i,t}$ is the observable target, the jump state $Z_{i,t}$ is a latent state inferred from the count process, and business density B_i and category composition are spatial covariates. Missing sensor periods are masked rather than treated as zero counts, because artificial zeros would otherwise be incorrectly interpreted as negative jumps.

3.2. Seasonal Baseline Pedestrian Intensity

Pedestrian counts exhibit strong recurrent structure before abnormal variations are considered. The baseline component therefore estimates the expected intensity under regular temporal conditions as

$$\lambda_{i,t}^{(0)} = \exp\left(\beta_0 + \beta_h H_t + \beta_d D_t + \beta_m M_t + \boldsymbol{\gamma}^{\top}\mathbf{x}_{i,t}\right), \qquad (1)$$

where i denotes a pedestrian sensor or associated commercial zone, t denotes an hourly observation, and $\lambda_{i,t}^{(0)}$ is the expected pedestrian intensity in the absence of an abnormal jump. H_t , D_t , and M_t represent hour-of-day, day-of-week, and month/season effects, respectively. $\mathbf{x}_{i,t}$ contains additional observable covariates available for the corresponding zone and period, while β_0 , β_h , β_d , β_m , and $\boldsymbol{\gamma}$ are estimated coefficients.

The log-link guarantees a positive intensity while allowing multiplicative temporal effects. Importantly, seasonal variation is removed before jump inference. A high Friday-evening count is therefore not automatically classified as abnormal when it is consistent with the historical Friday-evening baseline.

3.3. Compound Poisson Jump Mechanism

Abrupt pedestrian-flow changes are represented as discrete stochastic arrivals. For zone i , the number of jumps occurring over interval Δ follows

$$N_i(t+\Delta) - N_i(t) \sim \text{Poisson}\left(\kappa_{i,t}\Delta\right), \qquad (2)$$

where $N_i(t)$ is the cumulative number of jump arrivals and $\kappa_{i,t}$ is the jump-arrival intensity. Larger $\kappa_{i,t}$ therefore indicates a zone with more frequent abnormal changes.

A jump does not have to disappear immediately after its arrival. The conditional pedestrian intensity is consequently defined as

$$\log \lambda_{i,t} = \log \lambda_{i,t}^{(0)} + \sum_{\tau_{ik} \le t} J_{ik}\, e^{-\rho_i (t-\tau_{ik})}, \quad (3)$$

where τ_{ik} is the occurrence time of jump k, J_{ik} is its signed magnitude, and $\rho_i > 0$ controls post-jump decay. Positive J_{ik} represents an abnormal increase, whereas negative J_{ik} represents an abnormal decline. The exponential term separates short-lived disturbances from persistent changes. The latent state $Z_{i,t}$ indicates whether the observation is associated with an active jump and is inferred probabilistically rather than supplied as a manually labelled anomaly.

3.4. Likelihood-Based Parameter Estimation

Parameters are estimated jointly from observed pedestrian counts and the latent jump process. The objective log-likelihood is

$$\mathcal{L}(\Theta) = \sum_{i,t} \left[Y_{i,t} \log \lambda_{i,t} - \lambda_{i,t} - \log\left(Y_{i,t}!\right) \right] + \log p\,(\mathbf{J}, \mathbf{N} \mid \Theta), \quad (4)$$

where $Y_{i,t}$ is the observed count and $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\kappa}, \boldsymbol{\rho}, \mu_J, \sigma_J\}$ contains the baseline coefficients, jump-arrival rates, decay parameters, and parameters governing the jump-magnitude distribution.

The first term evaluates how well the conditional intensity explains observed counts. The second prevents arbitrary residual deviations from being classified as jumps by imposing a probabilistic structure on jump occurrence and magnitude. Estimation alternates between evaluating posterior probabilities for latent jump states and updating continuous parameters. This separation is important because the framework must distinguish a genuine discontinuity from ordinary Poisson variability.

3.5. Commercial Exposure and Profit Mapping

Each pedestrian sensor is spatially linked to commercial establishments within a predefined 500 m catchment. Establishments are grouped into retail, food and beverage, personal services, and accommodation categories. Effective exposure $F_{s,t}$ is obtained from the conditional pedestrian intensity surrounding shop s, with closer sensors receiving greater weight and local business density controlling the division of aggregate footfall among establishments.

The conditional operating profit is then defined by

$$\Pi_{s,t} = F_{s,t} q_s v_s m_s - C_{s,t}^F, \quad (5)$$

where $F_{s,t}$ is effective pedestrian exposure, q_s is the pedestrian-to-customer conversion probability, v_s is average transaction value, m_s is the contribution margin, and $C_{s,t}^F$ denotes fixed operating cost.

This formulation deliberately avoids assuming that pedestrian count equals revenue. A negative jump first reduces effective exposure and then affects profit through conversion and margin conditions. It also creates a nonlinear risk response near the break-even level: the same percentage reduction in footfall can have a limited effect on a high-margin establishment but move a low-margin establishment from positive to negative profit.

Because the public spatial data do not contain verified shop-level financial statements, the resulting quantity is interpreted as conditional profit–loss risk under specified operating-parameter distributions, rather than observed accounting profit.

3.6. Monte Carlo Loss-Probability Estimation

Uncertainty is propagated through $M = 100{,}000$ independent Monte Carlo paths. Each path samples jump occurrence, jump magnitude and duration together with uncertain commercial parameters. Conversion probability and contribution margin are represented by bounded distributions, while transaction value and operating cost are sampled from positive-valued distributions whose parameters are varied in the sensitivity analysis.

For shop s, the estimated probability of operating loss is

$$\hat{P}_{\text{loss},s} = \frac{1}{M}\sum_{m=1}^{M} \mathbb{I}\left(\Pi_s^{(m)} < 0\right), \quad (6)$$

where $\Pi_s^{(m)}$ is the simulated profit in Monte Carlo path m, and $\mathbb{I}(\cdot)$ is an indicator equal to one when the simulated profit is negative.

Loss probabilities are estimated separately under normal, positive-jump, and negative-jump states. Monte Carlo standard errors and 95% confidence intervals are reported so that differences in estimated risk are not interpreted from point estimates alone.

3.7. Tail-Loss Severity

Loss probability measures the frequency of negative-profit outcomes but does not measure their severity. Tail exposure is therefore evaluated using Conditional Value-at-Risk:

$$\text{CVaR}_\alpha = E[-\Pi \mid -\Pi \geq \text{VaR}_\alpha],\ \alpha = 0.95. \quad (7)$$

Here, VaR_α is the α-quantile of the simulated loss distribution, and CVaR_α is the expected loss conditional on outcomes lying beyond that threshold. P_{loss} therefore describes whether a loss occurs, whereas CVaR describes the expected severity of adverse tail outcomes. Reporting both avoids treating a frequently small loss and an infrequent but severe loss as equivalent risks.

3.8. Parameter Updating and Convergence Control

After latent jump probabilities are updated, continuous parameters are refined using a damped quasi-Newton step:

$$\Theta^{(r+1)} = \Theta^{(r)} - \eta_r (H_r + \varepsilon I)^{-1} \nabla \mathcal{J}\left(\Theta^{(r)}\right), \quad (8)$$

where r denotes the iteration, η_r is the step size, H_r is the Hessian or quasi-Newton approximation, I is the identity matrix, ε is a small stabilization constant, and $\mathcal{J} = -\mathcal{L}$ is the minimized negative log-likelihood.

The update is accepted only when it improves the objective or satisfies the damping rule. Iteration terminates when the relative objective change falls below 10^{-5} and the maximum parameter shift falls below 10^{-4}. Monte Carlo estimation is additionally regarded as stable when the standard error of the estimated loss probability is below 0.005. These criteria separate statistical parameter convergence from simulation convergence, which are evaluated independently in Chapter 4.

Algorithm 1. Poisson Jump–Monte Carlo Risk Estimation

Input: hourly pedestrian counts Y; timestamps T; sensor coordinates S; business records B; Monte Carlo size M

Output: jump states Z; conditional intensity λ; P_{loss}; VaR; CVaR

1: Synchronize pedestrian observations by sensor and local timestamp.

2: Remove duplicated observations and mask confirmed sensor outages.

3: Match commercial establishments to pedestrian catchment zones.

4: Construct hourly, weekday, monthly, and spatial covariates.

5: Estimate seasonal baseline intensity $\lambda^{(0)}$.

6: Initialize jump rate κ, jump magnitude parameters, and decay ρ.

7: Evaluate posterior probabilities of latent jump states Z.

8: Update jump-arrival and jump-magnitude parameters.

9: Re-estimate conditional pedestrian intensity λ.

10: Repeat Steps 7–9 until likelihood and parameter convergence.

11: Map conditional intensity to category-specific commercial exposure.

12: Sample conversion, transaction value, margin, cost, and jump uncertainty.

13: Generate M pedestrian-flow and cash-flow trajectories.

14: Estimate P_{loss}, VaR, CVaR, confidence intervals, and Monte Carlo error.

15: Return jump characteristics, conditional intensities, and probabilistic risk outputs.

4. Results and Analysis

4.1. Data and Experimental Setting

The empirical analysis used two public datasets from the City of Melbourne. Hourly pedestrian observations were obtained from the Melbourne Pedestrian Counting System, while establishment locations and ANZSIC4 industry classifications were obtained from the CLUE Business Establishments dataset. The study period covered January 2019 to December 2024. The pedestrian dataset provided the observed count $Y_{i,t}$, timestamp, sensor identifier, and sensor location, whereas CLUE records were used to construct local business density and industry composition within the 500 m catchment defined in Chapter 3.

All timestamps were converted to Australia/Melbourne local time. Duplicate sensor-hour records were removed, confirmed sensor-outage periods were masked, and missing observations were not replaced with zeros. Genuine extreme counts were retained because abrupt variation constituted the primary object of analysis. Hour-of-day, day-of-week, and monthly indicators were constructed before fitting the seasonal component. Establishments were spatially matched to pedestrian catchments and classified into retail, food and beverage, personal services, and accommodation activities. Financial inputs were treated as uncertain operating parameters rather than observed CLUE variables. Conversion probabilities and margins were sampled from bounded distributions, while transaction values and fixed costs were represented by positive-valued distributions; monetary outcomes were normalized by daily fixed-cost units to avoid interpreting the scenario parameters as observed shop accounts.

A chronological split was applied: 2019–2022 for parameter estimation, 2023 for model selection and jump-threshold determination, and 2024 for out-of-sample evaluation. Four benchmarks were used: Seasonal Poisson Regression, Negative Binomial Regression, Poisson Autoregression (PAR), and a Hawkes-type intensity model. Performance was assessed using mean absolute error (MAE), negative log-likelihood (NLL), continuous ranked probability score (CRPS), and an additional jump-period MAE calculated only for abnormal-flow intervals. Ten blocked bootstrap repetitions ($n = 10$) were conducted, and results are reported as mean ± standard deviation. Pairwise significance was evaluated using paired block-bootstrap tests at ($p<0.05$).

4.2. Pedestrian-Flow Jump Characterization

Table 1 summarizes the inferred jump characteristics for four representative commercial-zone types. Baseline hourly intensity differed substantially across zones, but abnormality was not simply proportional to average volume. The transport-commercial zone exhibited the highest estimated jump-arrival rate, ($\kappa = 0.53 \pm 0.06$) events per day, whereas the office-commercial zone had the lowest rate at (0.29 ± 0.04). This indicates that high temporal variability and high average pedestrian intensity are related but distinct properties.

Table 1. Statistical Characteristics of Pedestrian-Flow Jumps Across Commercial Zones (Mean ± SD, N = 10)

Commercial zone	Baseline intensity (ped./h)	Jump rate κ\kappa (day−1)	Mean positive jump	Mean negative jump	Mean half-life (h)	Jump-hour proportion
CBD retail core	1186 ± 84	0.41 ± 0.05	+31.6 ± 6.8%	−36.9 ± 7.3%	3.4 ± 0.7	8.7 ± 1.1%
Transport-commercial mixed	1038 ± 92	0.53 ± 0.06	+38.2 ± 8.1%	−34.7 ± 6.9%	2.8 ± 0.6	10.4 ± 1.3%

Dining-oriented	824 ± 77	0.36 ± 0.05	+27.4 ± 5.9%	−42.1 ± 8.5%	4.1 ± 0.8	7.9 ± 1.0%
Office-commercial	941 ± 69	0.29 ± 0.04	+24.8 ± 5.4%	−39.6 ± 7.7%	3.7 ± 0.7	6.6 ± 0.9%

Jump magnitudes in Table 1 are expressed as percentage deviations implied by the log-intensity jump parameter. Negative changes were larger than positive changes in three of the four zones, particularly in the dining-oriented zone. The transport-commercial zone was an exception, where positive shocks were both relatively large and short-lived. Estimated half-lives ranged from 2.8 to 4.1 h, demonstrating that treating every abnormal observation as an isolated one-hour residual would discard measurable persistence. The longer persistence of negative dining-area shocks also suggests that the economic consequences of abnormal flow depend on both magnitude and duration.

4.3. Model Comparison and Jump-Mechanism Verification

Table 2 reports out-of-sample performance. The proposed model achieved the lowest aggregate MAE, NLL, CRPS, and jump-period MAE. Relative to the strongest full benchmark, the Hawkes model, MAE decreased by 8.3%, NLL by 3.9%, and CRPS by 7.4%. The corresponding paired tests gave $p = 0.008$, $p = 0.031$, and $p = 0.012$, respectively. The improvement was more pronounced during abnormal periods, where jump-period MAE decreased from 329.8 to 301.7 ($p = 0.006$).

Table 2. Out-of-sample Pedestrian-Flow Modelling and Mechanism Verification (Mean ± SD, N = 10)

Model	MAE ↓	NLL ↓	CRPS ↓	Jump-period MAE ↓
Seasonal Poisson	238.7 ± 10.9	6.02 ± 0.14	157.6 ± 8.7	403.5 ± 17.8
Negative Binomial	216.9 ± 9.6	5.71 ± 0.12	146.1 ± 7.5	381.2 ± 16.5
PAR	205.7 ± 8.8	5.58 ± 0.10	139.4 ± 6.8	352.6 ± 14.9
Hawkes-type model	196.4 ± 8.1	5.44 ± 0.09	132.8 ± 6.3	329.8 ± 13.6
Proposed w/o jump	203.6 ± 8.5	5.52 ± 0.10	137.2 ± 6.5	356.9 ± 15.1
Proposed w/o decay	188.9 ± 7.9	5.31 ± 0.09	127.4 ± 6.0	315.8 ± 13.1
Proposed full model	180.1 ± 7.4	5.228 ± 0.08	122.97 ± 5.8	301.7 ± 12.4

Removing the jump component increased jump-period MAE by 18.3%, showing that seasonal variation alone could not account for discontinuities. Removing exponential decay produced a smaller but significant deterioration ($p = 0.021$), indicating that abnormal changes frequently extended beyond their arrival hour. The Hawkes model remained competitive in transport-dominated locations, where event clustering was pronounced, but the compound Poisson representation was more stable across zones because it did not require each shock to generate subsequent shocks.

4.4. Retail Profit–Loss Risk under Different Flow States

Figure 2 presents the Monte Carlo loss probabilities for normal, mild positive, strong positive, mild negative, and strong negative pedestrian states. Under regular conditions, estimated one-day loss probabilities were 0.21 for retail, 0.25 for food and beverage, and 0.23 for personal services. Mild negative jumps increased these estimates to 0.28, 0.34, and 0.31, respectively, while strong negative jumps produced probabilities of 0.35, 0.42, and 0.38. When identified negative-jump states were pooled according to their empirical occurrence weights, the mean probability increased from approximately 0.23 under regular conditions to 0.38.

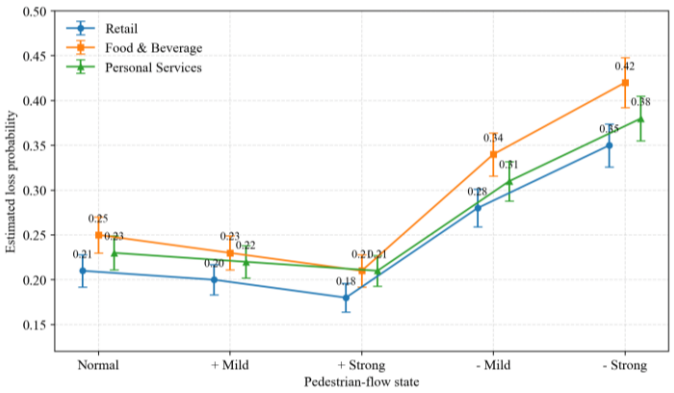


Figure 2. Estimated Retail Loss Probability under Different Pedestrian-Flow Jump States (Mean and 95% CI).

The response was not linear. When effective pedestrian exposure was reduced by 5%, 10%, 20%, 30%, and 40%, pooled loss probability was approximately 0.24, 0.25, 0.29, 0.35, and 0.44, respectively. The steeper increase beyond a 20% reduction occurred because simulated revenue approached the fixed-cost break-even boundary. Food and beverage establishments showed the greatest sensitivity to negative states, whereas small positive jumps produced only limited risk reductions for personal services. Tail severity changed in the same direction: normalized 95% CVaR increased from 0.78 fixed-cost units under regular conditions to 1.17 under strong negative jumps. Thus, jump magnitude affected both the occurrence and severity of operating losses.

4.5. Monte Carlo Stability, Sensitivity and Robustness

Figure 3(a) evaluates Monte Carlo convergence. The estimated pooled negative-jump loss probability was 0.391, 0.374, 0.382, 0.379, 0.381, and 0.380 for 1,000, 5,000, 10,000, 25,000, 50,000, and 100,000 paths, respectively. The sequence was not artificially monotonic, but dispersion narrowed rapidly as path count increased. At 100,000 paths, the Monte Carlo standard error was approximately 0.0015, below the predefined 0.005 threshold.

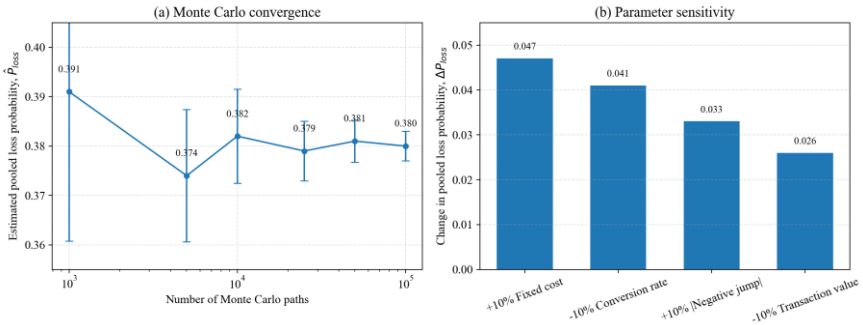


Figure 3. Monte Carlo Convergence and Sensitivity of Retail Loss-Risk Estimates: (A) Monte Carlo Convergence of Pooled Loss Probability; (B) Sensitivity of Pooled Loss Probability to Key Operating and Jump Parameters.

Figure 3(b) shows parameter sensitivity. A 10% increase in fixed operating cost increased P_{loss} by 0.047, while a 10% reduction in conversion rate increased it by 0.041. Increasing the absolute magnitude of negative jumps by 10% changed loss probability by 0.033, compared with 0.026 for a 10% reduction in transaction value. The responses were asymmetric because the profit function crosses a break-even boundary rather than varying proportionally around zero.

Cross-zone evaluation provided a less uniform result. The proposed model achieved lower CRPS in three of four commercial-zone types, whereas the Hawkes specification produced slightly lower point MAE in the transport-commercial zone. In an additional

robustness check, masking 10% of otherwise available sensor-hours increased aggregate CRPS by 4.8%, while the pooled loss-probability estimate shifted by less than 0.018. These results suggest that the principal findings were not dependent on a single commercial zone or an exact pedestrian-data coverage pattern, although local transport-driven event clustering remained a source of model-specific variation.

5. Conclusion

Explicit separation of recurrent seasonal intensity from discontinuous pedestrian-flow jumps produced consistently better out-of-sample probabilistic estimates than smooth count-process baselines. Relative to the strongest benchmark, the full specification reduced MAE from 196.4 to 180.1, NLL from 5.44 to 5.228, and CRPS from 132.8 to 122.97, corresponding to improvements of 8.3%, 3.9%, and 7.4%, respectively. These differences were statistically significant at the 5% level. The largest gain occurred during abnormal periods, where jump-period MAE declined from 329.8 to 301.7. Removing the jump component produced the strongest deterioration, while removing exponential decay caused a smaller but still significant loss in accuracy, confirming that both discontinuity and persistence contributed to the observed flow dynamics.

Commercial risk results showed that negative pedestrian shocks materially increased conditional operating-loss probability. The pooled probability rose from approximately 0.23 under regular conditions to 0.38 under identified negative-jump states, while strong negative jumps increased normalized 95% CVaR from 0.78 to 1.17 fixed-cost units. The response was nonlinear: reductions in effective pedestrian exposure beyond approximately 20% produced increasingly larger changes in loss probability as simulated revenue approached the break-even region. Probability estimates stabilized substantially after 50,000 Monte Carlo paths, and the standard error at 100,000 paths was approximately 0.0015.

Several limitations remain. Pedestrian counts measure exposure rather than realized purchases, verified store-level financial records were unavailable, and conversion, margin, and operating-cost parameters therefore remained distributional assumptions rather than directly observed quantities. The results are also geographically dependent on the selected Melbourne commercial zones. Future work should incorporate POS transactions, anonymized store-level revenue, weather and event variables, multivariate spatial jump processes, and external-city validation to test the transferability of the identified risk relationships.

References

1. X. Peng, Y. Y. Niu, B. Meng, Y. Tao, and Z. Huang, "Big geo-data unveils influencing factors on customer flow dynamics within urban commercial districts," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 134, p. 104231, 2024.
2. C. Han, P. Seshadri, Y. Ding, N. Posner, B. W. Koo, A. Agrawal, ..., and S. Guhathakurta, "Understanding pedestrian movement using urban sensing technologies: the promise of audio-based sensors," *Urban Informatics*, vol. 3, no. 1, p. 22, 2024.
3. J. Liddle, W. Jiang, and N. Malleson, "Leveraging principal component analysis to uncover urban pedestrian dynamics," *J. Geogr. Syst.*, vol. 27, no. 3, pp. 425–453, 2025.
4. D. Goyal and M. Xie, "On the compound Poisson phase-type process and its application in shock models," *J. Comput. Appl. Math.*, vol. 446, p. 115852, 2024.
5. Y. Feng, A. Che, and N. Tian, "Robust inventory routing problem under uncertain demand and risk-averse criterion," *Omega*, vol. 127, p. 103082, 2024.
6. D. Wood and S. Çetinkaya, "A comparative analysis of how to handle stockout scenarios: the impact of newsvendor's risk attitude," *Ann. Oper. Res.*, vol. 338, no. 2, pp. 1187–1218, 2024.
7. A. Angel and P. Plaut, "Tempo-spatial analysis of pedestrian movement in the built environment based on crowdsourced big data," *Cities*, vol. 149, p. 104917, 2024.
8. L. Du, H. Liu, and W. Li, "Pedestrian flow prediction using a spatiotemporal multi-head attention graph convolutional network integrated with knowledge graph," *Appl. Intell.*, vol. 55, no. 13, p. 896, 2025.
9. M. Asher, Y. Oswald, and N. Malleson, "Understanding pedestrian dynamics using machine learning with real-time urban sensors," *Environ. Plan. B: Urban Anal. City Sci.*, vol. 52, no. 8, pp. 1994–2017, 2025.
10. P. Alaimo Di Loro, M. Mingione, and P. Fantozzi, "Semi-parametric spatio-temporal Hawkes process for modelling road accidents in Rome," *J. Agric. Biol. Environ. Stat.*, vol. 30, no. 1, pp. 8–38, 2025.

11. M. Kieu, A. Comber, T. B. Quang, and N. Malleson, "Modelling social mobility disruptions and recovery during disasters: A mobile phone data approach," *Int. J. Disaster Risk Reduct.*, p. 105812, 2025.
12. N. Drezner and Y. Barron, "An emergency supply policy for an inventory replenishment model with returns and partial backorders," *Ann. Oper. Res.*, vol. 343, no. 1, pp. 175–221, 2024.
13. W. Jammernegg, P. Kischka, and L. Silbermayr, "Risk preferences, newsvendor orders and supply chain coordination using the Mean-CVaR model," *Int. J. Prod. Econ.*, vol. 270, p. 109171, 2024.
14. T. Ahmadi, A. F. Hesaraki, A. Mahmoodi, and A. Marandi, "Managing customer waiting times in an inventory system using Conditional Value-at-Risk measure," *Ann. Oper. Res.*, vol. 344, no. 1, pp. 1–45, 2025.
15. A. Garrido, F. Pongutá, and H. García-Reyes, "Zero-inventory plans, constant workforce, or hybrid approach? Analysing pure production strategies for enhancing factory resilience with demand variability," *Int. J. Prod. Res.*, vol. 63, no. 10, pp. 3589–3607, 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.