Adverse Weather and Urban Transport: Travel Demand, System Performance, and Resilience

ROOZBEH EBRAHIMI GOLSHANABADI
Doctoral School of Transport
National University of Science and Technology "Politehnica" Bucharest
Splaiul Independentei 313, 060042 Bucharest
ROMANIA

Abstract: - Weather acts on urban transport at several levels at once: on individual trip decisions, on the speed and reliability of services, on the integrity of infrastructure and on the agencies that run it. Evidence for each level comes from different disciplines and data sources, and these strands seldom cite one another. This paper synthesizes studies of metro, bus, ride-sourcing, bike-sharing, road networks and transit agencies in Asia, Australia, North America and Europe. The studies are placed in a four-level framework (travel demand; operational performance and prediction; network resilience; governance and adaptation) linked to a typology of direct and indirect disruption pathways and their weather definitions, model formulations and effect estimates are compared in their native metrics. Precipitation and wind reduce aggregate demand in every study that tests them, yet the sign at a single station or zone can be positive when that unit receives travelers leaving walking, cycling or open-air waiting. Weekday/weekend moderation runs in opposite directions across studies. Bus-stop shelters offset about a quarter of the weekday loss on cold and heavy-rain days and all of it on hot days. Weather inputs change average forecast accuracy very little but reduce error sharply during rare events. On Shenzhen roads the response phase lasted two to three times as long as recovery, and the usual loss–recovery relation reversed under a super typhoon. Among US transit agencies, hurricanes were the only hazard that raised risk perception. We argue that demand robustness, network performance resilience and life-safety resilience are separate properties that need separate metrics and governance responses.

Key-Words: - Urban Transport Resilience; Extreme Weather; Public Transit Ridership; Spatiotemporal Heterogeneity; Resilience Metrics; Climate Adaptation

Received: March 19, 2026. Revised: May 21, 2026. Accepted: June 23, 2026. Published: November 11, 2026.

1 Introduction

Passengers, vehicles and infrastructure respond to weather on very different time scales. A traveller decides within minutes whether to cancel a trip, delay it or switch mode when rain starts. Road speeds fall and recover over hours during a storm. Infrastructure built for a historical climate can fail outright once that climate is exceeded [1], [2]. Markolf et al. [1] call the last problem climate non-stationarity: past records no longer predict present or future conditions, so design values such as the 100-year storm rest on a distribution that is moving. In one pilot assessment they cite, the annual flood probability at an interstate bridge is expected to rise from 1.6% to 10% during the bridge's service life.

Research on these problems has grown along four lines that rarely cite each other. Travel-behaviour studies use smart-card and GPS records to estimate how temperature, precipitation, humidity, wind and air quality shift ridership and mode choice [3], [4], [5]. Prediction studies, from tree ensembles to graph neural networks, feed weather into short-term demand or speed forecasts [6], [7], [8]. Resilience studies define and measure a system's capacity to absorb disruption and recover, mostly through topological node-removal simulations or single-event cases [2]. Governance studies ask how agencies perceive weather risk and choose adaptation measures [1], [9], [10].

Demand studies disagree on the sign of precipitation and temperature effects [4], and it is unclear whether the disagreement reflects context or simply different definitions of "extreme" weather and different spatial units. Prediction studies usually report an average accuracy gain from weather inputs, which can hide how a model performs during the rare events that operators care about most. Evidence that ridership is stable under bad weather is also sometimes read as evidence that the system is resilient. The flooding of Zhengzhou Metro Line 5 on 20 July 2021, which killed 14 people [11], happened on the mode that demand studies find least sensitive to weather [4].

The paper offers a four-level framework that places demand, performance, resilience and governance studies in one structure tied to disruption pathways. It compares weather definitions, model formulations and reported effects side by side in their original metrics, and it separates divergent findings that reflect real differences between settings from those that follow from method.

2 Methods

2.1 Corpus and review design

The study is a structured qualitative synthesis with quantitative extraction. The studies (Table 1 (Appendix)) were assembled before the analysis rather than retrieved by a systematic database search, so the synthesis makes no claim of exhaustiveness; its inferences concern patterns and contradictions that can be shown within the corpus.

2.2 Analytical framework

The framework (Figure 1) crosses four levels with the disruption pathways of Markolf et al. [1]. Direct physical pathways damage infrastructure. Direct non-physical pathways act on health, behaviour and decisions. Indirect pathways propagate through systems that are physically, geographically, digitally or logically connected. Level 1 (L1), travel demand and mode choice, is driven mainly by direct non-physical pathways. L2 covers operational performance (punctuality, waits, speed) and its prediction. L3 covers network resilience, meaning the loss and recovery of performance under extreme events and the physical and organisational failure chains behind them. L4 covers risk perception, adaptation choices and decision-support tools. Measures chosen at L4, such as shelters, dispatching rules, shutdown protocols and design standards, act back on the lower levels.

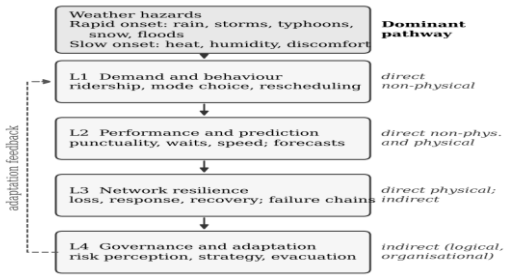


Fig. 1: Multi-level analytical framework. Labels on the right give the dominant disruption pathway in the typology of Markolf et al. [1]
Source: created by the authors.

2.3 Model formulations

To compare studies, we restate their core formulations in common notation while keeping the original definitions. Jiang and Cai [4] detrend daily metro ridership Y_t against a nine-term moving average taken over the same weekday (±4 weeks):

$$\Delta Y_t = \frac{Y_t - \bar{Y}_t^{\mathrm{MA}}}{\bar{Y}_t^{\mathrm{MA}}}, \qquad \bar{Y}_t^{\mathrm{MA}} = \frac{1}{9}\sum_{\tau=-4}^{4} Y_{t+7\tau}. \quad (1)$$

Temperature, relative humidity (RH), wind and the air quality index (AQI) are transformed in the same way. ΔY_t is regressed on these residuals, on rain-class dummies (light 0–10 mm, moderate 10–25 mm, heavy ≥25 mm per day), on season dummies and on a comfort index ranked 0–3:

$$\mathrm{CCI} = 0.6\,|T-24| + 0.07\,|RH-70| + 0.5\,|V-2|. \quad (2)$$

Weekday and weekend models are estimated separately, and 30 holidays are excluded.

Wei [3] fit a geographically weighted negative binomial regression for each suburb i and passenger type p:

$$\ln R_i^{(p)} = \beta_{i0} + \beta_{i1}\,\mathrm{Rain}_i + \beta_{i2}\,\mathrm{Wind}_i + \beta_{i3}\,\mathrm{HI}_i + \varepsilon_i, \quad (3)$$

with weather interpolated from six stations by inverse-distance weighting (power 2) and HI the heat index. An effect is called heterogeneous when the coefficient of variation of the local estimates, $\mathrm{CV} = \sigma/|\mu|$, exceeds 1, or when the sign changes across locations.

Miao et al. [12] estimate a stop-level panel:

$$\ln R_{ict} = \beta_1 W_{ct} + \beta_2\,(W_{ct} \times S_i) + \delta_m + \eta_y + \gamma_i + \mu_{ict}, \quad (4)$$

where W_{ct} marks county-days of extreme heat, cold or heavy rain, S_i marks a sheltered stop, δ_m and η_y are month and year dummies, and γ_i are stop fixed effects; standard errors are clustered by stop. β_1 is the effect at unsheltered stops and $\beta_1 + \beta_2$ the effect at sheltered stops. Because S_i does not vary over time, its main effect is absorbed by γ_i. Otim et al. [5] estimate multinomial logit and probit mode-choice models with walking as the reference alternative and report marginal effects $\partial P_{in}/\,\partial z_n$.

Rowe et al. [6] forecast one-hour-ahead stop ridership from 43 predictors:

$$R_{i,t+1} = f\left(R_{it}, R_{i,t-24}, R_{i,t-168}, W_{it}, C_t, L_i, G_i, F_i\right) + \varepsilon_{it}, \quad (5)$$

covering weather W, calendar C, trip-duration shares L, location G and urban amenities F. Yang et al. [7] combine a spatiotemporal branch D_g (multi-graph convolution followed by an LSTM) with a weather branch D_e (fully connected network):

$$D_p^{(t)} = w_g\,D_g^{(t)} + w_e\,D_e^{(t)}, \qquad w_g = 0.4,\ w_e = 0.6, \quad (6)$$

with the weights chosen empirically. Wang et al. [8] use a diffusion convolutional recurrent neural network (DCRNN) on a directed road graph that maps 24 past hours of speed and meteorology to 24 future hours.

Wang et al. [8] also measure resilience on observed speeds (Figure 2a). For a precipitation event, m_0 is the speed at onset t'_0, m_1 the minimum at t'_1, and m_2 the recovered speed at t'_2. The baseline $U_{wk}(t)$ is the mean hourly speed on dry days of the same weekday:

$$\mathrm{LoR} = \int_{t'_0}^{t'_2} \left[U_{wk}(t) - U_{day}(t)\right] dt, \qquad (7)$$

$$\mathrm{RST} = t'_1 - t'_0, \qquad \mathrm{RCT} = t'_2 - t'_1, \qquad (8)$$

$$\mathrm{RSR} = \frac{m_0 - m_1}{t'_1 - t'_0}, \qquad \mathrm{RCR} = \frac{m_2 - m_1}{t'_2 - t'_1}. \qquad (9)$$

LoR is the loss of resilience, RST and RCT are response and recovery times, and RSR and RCR are response and recovery rates. Replacing U_{day} with forecast speeds gives predicted indices.

For decision support, Deveci et al. [10] combine objective MEREC weights w_j^M and subjective rank-sum weights w_j^R for 12 criteria,

$$w_j = \tau\, w_j^M + (1 - \tau)\, w_j^R, \qquad (10)$$

and rank alternatives with an interval-valued Fermatean fuzzy MULTIMOORA procedure. Yazdani et al. [16] compare algorithms by the relative deviation index:

$$\mathrm{RDI} = 100\, \frac{\mathrm{Alg} - \mathrm{Best}}{\mathrm{Worst} - \mathrm{Best}}. \qquad (11)$$

2.4 Extraction and comparison

For each study we recorded setting, data and period, temporal and spatial unit, weather definition, model, outcome, estimates and stated limitations. Comparison used four criteria: direction and significance of each weather variable; moderators (day type, time of day, passenger type, location, infrastructure); the added predictive value of weather, taken from ablation or feature-importance results; and resilience behaviour (degradation, response, recovery). Estimates stay in their native metrics (percentage residuals, log-count coefficients, marginal probabilities, seconds, dollars and error statistics) because they cannot be put on a common scale, and no pooled effect is computed. Each statement is tagged as observed, model-derived, assumption or interpretation.

Definitions of adverse weather vary widely. Some thresholds are absolute, such as heavy rain at ≥25 mm per day [4] or more than 3 mm per hour [6]. Some are relative to the local distribution: Miao et al. [12] define hot days at or above the 90th percentile (≈23 °C daily mean), cold days at or below the 10th (≈−3 °C) and heavy rain at or above the 75th (≈19 mm), while Rowe et al. [6] use ±1.5 standard deviations from the hourly mean. Others are categorical: reported phenomena such as fog, rain or snow [13], [15], official alert levels [8], or counts of recorded storm events per county [9].

3 Results

3.1 L1: weather and travel demand

Table 2. (Appendix) lists the main demand-side estimates. Moderate and heavy rain and wind lower detrended metro ridership in Beijing, Shanghai and Shenzhen, with larger coefficients for heavy rain [4]. Rain also lowers adult ridership in Brisbane [3] and boardings at Salt Lake City bus stops [12]. No other result in the corpus is as consistent, although the effects are small: an extremely cold weekday removes about 1.6 of every 100 boardings at an unsheltered stop [12]

The sign does change in particular places and times. Light rain on Shanghai weekends goes with higher metro ridership [4]. In Brisbane, rain raises tertiary-student ridership near campuses in the morning peak and adult ridership in the inner city in the evening peak [3]. During extreme cold and high humidity, bus ridership rises on commuter corridors and busway stations and falls at retail and education stops [6]. In Philadelphia, extreme weather on weekdays brings longer pick-up waits and higher ride-sourcing fares, which the authors read as higher demand [13].

Warmer-than-usual days raise weekday metro ridership in Beijing and Shanghai but not in subtropical Shenzhen [4]. The heat index raises weekday adult ridership in Brisbane and lowers weekend senior ridership [3]. In Beijing GPS trajectories, walking shares are higher in every temperature class than in cold conditions, while cycling drops above 25 °C [5]. The comfort index is significant in all weekday metro models, but the source reads it one way in its results and the opposite way in its discussion [4], so we report only its sign.

Five moderators recur. For day type, weather effects on metro are larger at weekends [4] and the ride-sourcing indicators flip sign between weekdays and weekends [13], yet the Brisbane adult rain coefficient is larger on weekdays [3] and shelters reduce losses only on weekdays [12]. For time of day, adult demand in the morning peak shows no weather response [3]. Among passenger types, seniors are the most weather-sensitive, and their response is similar across space [3]. By location, city centres, subcentres and campuses respond differently and sometimes in opposite directions [3], [6]. As for infrastructure, shelters reduce losses, more so at low-frequency stops and stops with few transfers, where riders wait longer [12]. Shelters are not placed at random: they are more likely at high-frequency stops and in higher-income areas, areas with more

elderly residents and areas with fewer white residents.

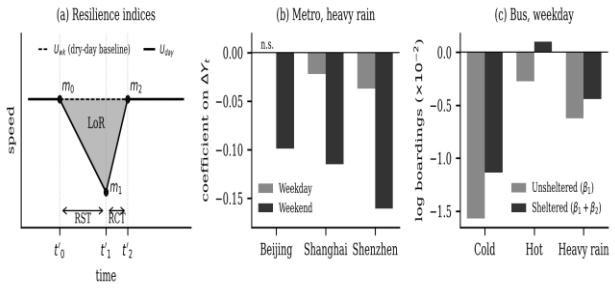


Fig. 2: b and c show two of these moderators. The net effect at sheltered stops is. On extreme hot weekdays it turns slightly positive (), in line with the authors' statement that shelters cancel the heat-related loss. For cold and heavy rain, shelters offset about 27% and 29% of the loss$\beta_1 + \beta_2 + 0.097 \times 10^{-2}$

Source: created by the authors.

Fig. 2: (a) Resilience indices of Wang et al. [8], redrawn from Eqs. (7)–(9); the shaded area is the loss of resilience. (b) Heavy-rain coefficients in the metro models of Jiang and Cai [4]; Beijing weekday not significant. (c) Weekday extreme-weather effects at unsheltered and sheltered bus stops from the panel estimates of Miao et al. [12]; sheltered values are model-derived sums

3.2 L2: operational performance and prediction

Adverse weather is associated with earlier, not later, bus arrivals in Winnipeg [15]. Support for "very early" arrival is 0.35 under rain and snow and 0.32 under fog, against 0.26 under clear skies. On-time support drops from 0.48 to 0.41–0.44, and late arrivals stay at or below about 7% in all conditions. The analysis uses association rules without controls and proposes no mechanism, so the result is descriptive.

Three regularities appear in the predictive studies. First, weather adds little on average. In Brisbane, lagged ridership, amenities and calendar events mattered more than weather; random forest, XGBoost and Tweedie XGBoost reached test RMSE of 2.497, 2.488 and 2.447, with a correlation of 0.985 and 81–82% of predictions exactly correct [6]. In Shenzhen, dropping meteorological and forecast inputs raised the 24-hour DCRNN RMSE only from 1.08 to 1.14 [8].

Second, the gain is concentrated in rare conditions. Brisbane errors were largest below 10 °C, below 40% relative humidity and in heavy rain, all unusual there. Under rain extremes, Tweedie XGBoost kept stop-level errors between about −50 and +25 riders, against below −100 to above +100 for the other two models, at 2.7 times the XGBoost training time (176.5 vs 65.6 min). During Typhoon Mangkhut, DCRNN followed an abnormal speed drop that sequence models without graph structure missed (R^2 0.95 against 0.84–0.89). Meteorological inputs let the model anticipate the typhoon but exaggerated its effect when used alone; historical speeds restored the daily and weekly pattern [8].

Third, aggregate metrics hide where and when errors occur. Random forest looked best on aggregate metrics in Brisbane, but event- and location-specific errors reversed the ranking [6]. For dockless bike-sharing in Tianjin, the fused model beat historical-average, ARIMA, LSTM and GRU baselines (MAE 8.209, RMSE 11.527, R^2 0.737), and removing the weather branch raised MAE to 10.903 [7]. The ablated models, however, report higher R^2 (0.762–0.851) than the full model, which makes the ablation hard to interpret.

3.3 L3: network resilience under extreme events

Wang et al. [8] is the only study in the corpus that measures resilience from observed data across many events: 300 precipitation-related events in 649 days of Shenzhen speeds. Loss of resilience grows with precipitation intensity and alert level and depends more on network density than on road class, with sparse areas losing more. Response time is two to three times recovery time and is the index most sensitive to severity, while recovery time varies less. Under ordinary precipitation, links that lose more recover faster, which the authors call elastic behaviour. Under Typhoon Mangkhut the relation reversed, and losses jumped where accumulated rainfall exceeded 50 mm; the authors interpret this as an elastic limit. Predicted loss of resilience had errors near 7%, against about 11% for response rate and 13% for recovery rate, so the scale of a disruption is easier to forecast than its timing.

Zhengzhou [11] shows a different kind of failure. Peak hourly rainfall reached 201.9 mm. Physical defects included an unauthorised design change that left the Wulongkou depot area 1.973 m lower, substandard or temporary flood walls, and a drainage ditch blocked by a 58 m cover plate and construction waste. Organisational failures compounded them. Line 5 kept running after other lines had closed despite consecutive red warnings. The control centre ordered the train to reverse 30 m, after which it lost power at a low point. The branch company did not report to the group company until 19:48, and evacuation ended at 22:00. More than 900 passengers were trapped, over 400 of them for

more than three hours. Three accident-causation models yield 12 governance factors across natural, technological, institutional and organisational resilience, ten of them the responsibility of government.
Ji et al. [2] note the lack of a unified resilience index, the focus on single rather than compound events, design based on historical return periods, and topological methods that cannot represent post-event behaviour.

3.4 L4: perception, adaptation and decision support

Zhang [9] relates 2017–2018 storm exposure to risk perception and adaptation in 144 US transit agencies. Hurricanes were the only hazard that raised risk perception (coefficient 0.260, $p < 0.01$) and the only significant predictor of incremental adaptation; adding the hazard variables raised R^2 from 0.209 to 0.253 (joint Wald $F = 2.29$, $p = 0.04$). Floods and winter storms recur often, affect wide areas and arrive quickly, attributes that theory links to higher perceived risk, yet they had no effect. Agencies experienced about six such events on average, against 0.29 and 0.12 federal disaster declarations. The author attributes this to normalising bias. Extreme heat, the slow-onset hazard, did not raise perception and was negatively associated with coping adaptation. Markolf et al. [1] report a similar pattern: of about two dozen agencies in a US federal climate-resilience pilot, only three considered indirect or non-physical pathways, and most adaptation aimed at robustness. They argue that non-stationarity, interconnection and unpredictable human response limit robustness-only strategies, and they propose flexibility (capacity for anticipated change) and agility (capacity for unanticipated change).
In an expert assessment for a hypothetical congested city, operational preparedness ranked above climate-resistant design, which ranked above alternative routing. The order held for every $\tau \in [0,1]$ in Eq. (10) and under fuzzy TOPSIS, WSM and WASPAS variants; routing was penalised because it may add traffic to congested networks [10]. For bus evacuation, a simheuristic that simulates only elite solutions reached an average RDI of 34.9% and 38.9% at route-disruption probabilities of 0.1 and 0.3, compared with 84.9% and 84.4% when every iteration was simulated; all 20 instances are synthetic [16]. Brum-Bastos et al. [17] propose adding a weather channel based on ERA5 reanalysis to an agent-based mobility model, a hypothesis they have not yet tested.

4 Discussion

4.1 Why demand findings diverge

Most L1 contradictions become consistent once two mechanisms are separated, as Jiang and Cai [4] suggest. Bad weather suppresses discretionary trips and makes access and egress harder. It also moves travellers from exposed modes (walking, cycling, waiting in the open) to protected ones. The sign observed at any unit is the net of these two flows. Units near the origins of substituted short trips gain riders: Shanghai metro on light-rain weekends, Brisbane campuses and the inner city at certain peaks, and Brisbane commuter corridors. System-wide and suburban units lose them. This is an interpretation, but it matches the spatial evidence, and Wei [3] suggest that earlier coarse-scale studies found weaker weather effects because aggregation cancelled opposing local responses.

The day-type contradiction separates in a similar way. Demand sensitivity, the proportional loss of trips, is higher at weekends when trips are discretionary [4], [13]. Sensitivity to mitigation, the share of riders an amenity retains, appears only on weekdays, when commuters must travel and shelter is worth more to them [12]. The larger weekday rain coefficient for Brisbane adults does not fit this pattern. One likely reason is that Wei [3] classify days as school and non-school days, so their weekday sample includes off-peak discretionary trips, although the corpus cannot test this. Weekday versus weekend is a weak proxy for trip purpose, and studies that observe purpose directly would tell us more.
Temperature results differ because the comfort optimum depends on local climate (Section 3.1), and forecast errors peak in conditions that are locally rare. Markolf et al. [1] make the same point with the 2014 Atlanta snowstorm, which paralysed a city unused to snow. Linear terms and single absolute thresholds fit this behaviour poorly. Comfort and heat indices move toward climate-relative modelling, but their mixed results show the problem is not solved.

4.2 Methodological choices

Moving-average residuals (Eq. 1) express all outcomes as percentage deviations and so allow pooling across cities of very different size. They also remove weather variation that is itself seasonal, which means they identify the effect of anomalous weather rather than of climate. Fixed effects with month and year dummies (Eq. 4) use within-stop variation over six years and give a stronger basis for a causal reading of the shelter interaction, provided

the interaction is treated as exogenous; Miao et al. [12] argue that it is, since the stop fixed effects already condition on shelter status. We treat the fixed-effects estimates as the most credible magnitudes in the demand evidence and the residual estimates as the most comparable across cities.

A global coefficient averages opposing local responses. The geographically weighted model (Eq. 3) exposes this heterogeneity, though its CV > 1 rule is a heuristic that applies only where all local coefficients share a sign. Given the substitution mechanism, local estimation is the better default whenever the question concerns specific stops or zones.

Wang et al. [8] pooled 300 events of all intensities for training. Their own results support this for loss and response but not for recovery under the typhoon, so pooling seems acceptable for predicting whether and how much a network degrades, while recovery under catastrophic events needs separate validation. The same caution applies to any demand model whose extreme category rests on a few days.

No two demand studies share a definition of extreme weather: "heavy rain" means ≥25 mm per day in one study, about 19 mm per day in another and more than 3 mm per hour in a third. Distribution-relative thresholds express rarity, which drives forecast error and possibly behaviour, and so compare better across climates. Absolute thresholds relate more directly to flooding and to operating rules. Reporting both would help. Extreme categories are also small by construction (8% of ride-sourcing observations; a three-month window in the Brisbane prediction study), so a null result for extremes in a short series indicates low power, not robustness.

For prediction, the DCRNN ablation shows why weather matters mainly in rare events: meteorology anticipates the typhoon but overstates it, and periodic speeds restrain the overreaction. Model comparisons should therefore report errors stratified by weather extremity, time and location alongside global RMSE or MAPE, because the model preferred on global metrics can be the wrong one for operations [6]. Ablations also need to be internally consistent before they count as evidence, and the fusion weight $w_e = 0.6$ in Eq. (6) is a tuned hyperparameter, not a measure of how much weather matters.

4.3 Demand robustness is not system resilience

The corpus uses "resilience" for three different properties. Demand robustness is the stability of ridership under adverse weather; metro scores high, with detrended ridership moving by fractions of a percent for most weather residuals [4]. Network performance resilience is the shape of the degradation–recovery curve in speed or travel time [8]. Life-safety and asset resilience is the avoidance of catastrophic failure [11]. Zhengzhou Metro failed on the third property, which demand studies do not observe. The failure came from latent technical and organisational weaknesses that never appear in normal operation, and neither a ridership regression nor a speed index would have detected them.

The resilience indices carry two further lessons. A response phase two to three times longer than recovery means that the larger operational gains lie in response, through earlier warnings, restrictions and rerouting. The elastic limit means that relationships estimated from frequent moderate events may not hold for rare severe ones. In the terms of Ji et al. [2], response and recovery correspond to static and dynamic resilience; in those of Markolf et al. [1], the elastic limit is where a robustness regime fails and graceful extensibility becomes necessary.

4.4 From evidence to institutions

Agency behaviour departs from the evidence in three ways. Floods and winter storms, the hazards agencies record most often, do not raise their risk perception [9], yet the L1–L3 studies suggest that frequent moderate events account for most cumulative ridership loss and network degradation. Heat did not raise agency risk perception and was negatively associated with coping adaptation, even though it reduces boardings at exposed stops [12], changes seniors' travel [3] and endangers people walking to and waiting at stops [1]. And agencies favour strengthening assets against direct physical damage [1], whereas Zhengzhou failed through decisions, communication and coordination. Xie and Peng [11] name complacency, overconfidence and fatigue as the institutional causes, which resemble the normalising bias in Zhang [9].

The decision tools point the same way, but both rest on hypothetical or synthetic cases [10], [16], so they show which decisions matter rather than giving calibrated answers. Table 3 (Appendix) summarises the findings by evidence type and strength.

Within these limits, some practical steps follow. Shelters should go first to low-frequency stops with few transfer options, where waits are longest, and their siting should be checked for equity [12]. Weather-responsive service changes can be targeted in space, moving capacity toward centres, campuses and commuter corridors that gain riders in rain [3], [6]. Forecasting systems should be judged on event-conditional errors and use count-appropriate objectives. Underground systems need trigger-based shutdown and evacuation protocols that do not de-

pend on ridership targets [11]. Internal risk communication in agencies should address the normalisation of moderate floods and the low salience of heat [9].

4.5 Limitations

Metrics, units and thresholds differ too much for meta-analysis, so comparisons rely on direction, moderation and within-study contrasts. Most L1 estimates are associational. The fixed-effects design of Miao et al. [12] controls for time-invariant stop characteristics but not for time-varying confounders, and its sampled ridership measure may be biased downward. Several series are short, from weeks [13] to one year [3], [4]. We found internal inconsistencies in four sources: the interpretation of the comfort and air-quality coefficients in Jiang and Cai [4], the ablation R^2 values in Yang et al. [7], implausible monthly figures confounded with school holidays in Yuldoshev et al. [14], and policy simulations that insert continuous shares into a dummy-coded model in Pujiati et al. [18]. These results are reported with caveats or excluded. The L4 tools rest on hypothetical or synthetic cases, and the L3 evidence consists of one quantitative study and one qualitative case.

5 Conclusions

Across studies, precipitation and wind reduce aggregate urban transit demand more consistently than any other weather factor, but the sign at a particular location depends on whether the unit of analysis receives people moving out of exposed modes. Day type, time of day, passenger type, location and shelter all moderate the response, and temperature effects depend on local climate. Weather inputs add little to average forecast accuracy and a great deal during rare events, provided models use count-appropriate objectives, graph-based spatial dependence and both meteorological and periodic inputs. On observed roads, response lasts two to three times as long as recovery, and the loss–recovery relation can reverse under a super typhoon. Agency risk perception follows salient, high-consequence hazards rather than the frequent or slow-onset ones that drive cumulative impacts, and adaptation still favours robustness.

Demand robustness, network performance resilience and life-safety resilience should therefore be measured and governed separately. Future work could use simultaneous multimodal data to separate suppression from substitution [4], [6]; report both climate-relative and absolute thresholds over series long enough to give adequate power for extremes; evaluate forecasts on event-conditional errors; extend multi-event resilience measurement to traffic volumes, public transport and compound events [2], [8]; and test the decision tools and the weather-aware mobility framework on real networks [10], [16], [17]. Research on how such evidence can counter normalising bias inside transport agencies [9] belongs on the same agenda.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work, the author(s) utilized artificial intelligence (AI) technologies solely for the purpose of linguistic refinement and grammatical correction. Following the application of these tools, the author(s) comprehensively reviewed, verified, and edited the resulting text. The author(s) maintain full accountability and responsibility for the accuracy, integrity, and final content of the publication.

References:

[1] Markolf, S.A., Hoehne, C., Fraser, A., Chester, M.V., Underwood, B.S., Transportation resilience to climate change and extreme weather events – beyond risk and robustness, Transport Policy, Vol. 74, 2019, pp. 174-186. doi: 10.1016/j.tranpol.2018.11.003

[2] Ji, T., Yao, Y., Dou, Y., Deng, S., Yu, S., Zhu, Y., Liao, H., The impact of climate change on urban transportation resilience to compound extreme events, Sustainability, Vol. 14, 2022, Art. no. 3880. doi: 10.3390/su14073880

[3] Wei, M., How does the weather affect public transit ridership? A model with weather-passenger variations, Journal of Transport Geography, Vol. 98, 2022, Art. no. 103242. doi: 10.1016/j.jtrangeo.2021.103242

[4] Jiang, S., Cai, C., The impacts of weather conditions on metro ridership: an empirical study from three mega cities in China, Travel Behaviour and Society, Vol. 31, 2023, pp. 166-177. doi: 10.1016/j.tbs.2022.12.003

[5] Otim, T., Dörfer, L., Bousdar Ahmed, D., Munoz Diaz, E., Modeling the impact of weather and context data on transport mode choices: a case study of GPS trajectories from Beijing, Sustainability, Vol. 14, 2022, Art. no. 6042. doi: 10.3390/su14106042

[6] Rowe, F., Mahony, M., Tao, S., Assessing machine learning algorithms for near-real time bus ridership prediction during ex-

treme weather, manuscript, University of Liverpool and Beijing Normal University, n.d.

[7] Yang, Y., Shao, X., Zhu, Y., Yao, E., Liu, D., Zhao, F., Short-term forecasting of dockless bike-sharing demand with the built environment and weather, Journal of Advanced Transportation, Vol. 2023, Art. no. 7407748. doi: 10.1155/2023/7407748

[8] Wang, H.-W., Peng, Z.-R., Wang, D., Meng, Y., Wu, T., Sun, W., Lu, Q.-C., Evaluation and prediction of transportation resilience under extreme weather events: a diffusion graph convolutional approach, Transportation Research Part C, Vol. 115, 2020, Art. no. 102619. doi: 10.1016/j.trc.2020.102619

[9] Zhang, F., Not all extreme weather events are equal: impacts on risk perception and adaptation in public transit agencies, Climatic Change, Vol. 171, 2022, Art. no. 3. doi: 10.1007/s10584-022-03323-0

[10] Deveci, M., Gokasar, I., Mishra, A.R., Rani, P., Ye, Z., Evaluation of climate change-resilient transportation alternatives using fuzzy Hamacher aggregation operators based group decision-making model, Engineering Applications of Artificial Intelligence, Vol. 119, 2023, Art. no. 105824. doi: 10.1016/j.engappai.2023.105824

[11] Xie, Z., Peng, B., A framework for resilient city governance in response to sudden weather disasters: a perspective based on accident causation theories, Sustainability, Vol. 15, 2023, Art. no. 2387. doi: 10.3390/su15032387

[12] Miao, Q., Welch, E.W., Sriraj, P.S., Extreme weather, public transport ridership and moderating effect of bus stop shelters, Journal of Transport Geography, Vol. 74, 2019, pp. 125-133. doi: 10.1016/j.jtrangeo.2018.11.007

[13] Shokoohyar, S., Sobhani, A., Sobhani, A., Impacts of trip characteristics and weather condition on ride-sourcing network: evidence from Uber and Lyft, Research in Transportation Economics, 2020, Art. no. 100820. doi: 10.1016/j.retrec.2020.100820

[14] Yuldoshev, D., Abdullaev, B., Yusufkhonov, Z., Muminov, T., Determining the impact of weather indicators on passenger traffic in public transport, Universum: Technical Sciences, Vol. 1, No. 94, 2022, pp. 65-70.

[15] Nguyen, B., Shinnie, M.J.D., Leung, C.K., Kuzie, M.R., Kokilev, N.N., Kaur, S., A data science solution for mining weather data and transportation data for smart cities, in 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), 2022, pp. 1672-1676. doi: 10.1109/COMPSAC54236.2022.00266

[16] Yazdani, M., Mojtahedi, M., Loosemore, M., Enhancing evacuation response to extreme weather disasters using public transportation systems: a novel simheuristic approach, Journal of Computational Design and Engineering, Vol. 7, No. 2, 2020, pp. 195-210. doi: 10.1093/jcde/qwaa017

[17] Brum-Bastos, V., Smolak, K., Rohm, W., Sila-Nowicka, K., A weather-aware framework for population mobility modelling, in 15th International Conference on Spatial Information Theory (COSIT 2022), LIPIcs, 2022, pp. 17:1-17:9. doi: 10.4230/LIPIcs.COSIT.2022.17

[18] Pujiati, A., Murniawaty, I., Nihayah, D.M., Muarrifah, I., Damayanti, N., A simulated policy towards green public transportation in a metropolitan in Indonesia, International Journal of Energy Economics and Policy, Vol. 12, No. 5, 2022, pp. 162-168. doi: 10.32479/ijeep.13121

[19] Tsavachidis, M., Le Petit, Y., Re-shaping urban mobility – key to Europe's green transition, Journal of Urban Mobility, Vol. 2, 2022, Art. no. 100014. doi: 10.1016/j.urbmob.2022.100014

[20] During the preparation of this work, the author(s) utilized artificial intelligence (AI) technologies solely for the purpose of linguistic refinement and grammatical correction. Following the application of these tools, the author(s) comprehensively reviewed, verified, and edited the resulting text. The author(s) maintain full accountability and responsibility for the accuracy, integrity, and final content of the publication.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)
Roozbeh Ebrahimi Golshanabadi, as the sole author, carried out all stages of the research, from the formulation of the problem and the evidence synthesis to the final findings, and wrote the manuscript.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself
No funding was received for conducting this study.

Conflict of Interest
The author has no conflicts of interest to declare that are relevant to the content of this article.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)
This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US

APPENDIX

Table 1. Corpus of studies. L1 = travel demand and behaviour; L2 = operational performance and prediction; L3 = network resilience; L4 = governance, adaptation and decision support; C = context only

Study	Setting and data	Unit	Method	Level
Jiang and Cai [4]	Beijing, Shanghai, Shenzhen metro, 2019	Daily; system	Moving-average residual MLR; pooled GLME	L1
Wei [3]	Brisbane smart card, 2015–16; 151.7 M trips	Daily, half-hourly; 196 suburbs	GWNBR by passenger type	L1
Miao et al. [12]	Salt Lake City bus, 2010–15; ≈5,800 stops	Daily; stop	Panel fixed effects with shelter interactions	L1/L4
Otim et al. [5]	Beijing Geolife GPS, 2007–12; 2,671 trips	Trip	Multinomial logit and probit	L1
Shokoohyar et al. [13]	Philadelphia Uber/Lyft API, June 2018	Hourly; 25-node grid	Time-fixed-effect OLS	L1/L2
Yuldoshev et al. [14]	Tashkent bus, 2019	Monthly; one depot	Descriptive comparison	L1
Nguyen et al. [15]	Winnipeg Transit on-time GPS	Stop arrival	Association rules	L2
Rowe et al. [6]	Brisbane bus smart card, Mar–May 2016	Hourly; 5,226 stops	RF, XGBoost, Tweedie XGBoost	L2
Yang et al. [7]	Tianjin dockless bike-sharing, May–Jun 2019	Hourly; 53 grids	Multi-graph GCN + LSTM + FCN	L2
Wang et al. [8]	Shenzhen road speeds (DiDi), 2017–19	Hourly; road link	DCRNN; resilience indices	L2/L3
Xie and Peng [11]	Zhengzhou Metro Line 5 flood, 2021	Event	Accident-causation models	L3/L4
Ji et al. [2]	Review	–	Conceptual review	L3
Markolf et al. [1]	Review; US pilot programmes	–	Pathways-of-disruption typology	L3/L4
Zhang [9]	144 US transit agencies, 2016 and 2019	Agency	CFA; OLS, robust SE	L4
Deveci et al. [10]	Hypothetical city; expert panel	Alternative	Fermatean fuzzy MULTI-MOORA	L4
Yazdani et al. [16]	20 synthetic evacuation instances	Bus trip	Simheuristic (DE + simulation)	L4
Brum-Bastos et al. [17]	Rio de Janeiro (planned)	Hourly; grid	Agent-based framework (no results)	L1
Pujiati et al. [18]	Semarang BRT; 100 workers	Individual	Binary logit	C
Tsavachidis and Le Petit [19]	EU urban mobility vision	–	Position paper	C

Source: created by the authors.

Table 2. Main weather effects on urban travel demand (observed estimates in native metrics; n.s. = not significant)

Study / mode	Variable	Reported effect	Moderation
Jiang and Cai [4], metro	Heavy rain	Weekday coefficient on ΔY: Shanghai −0.022, Shenzhen −0.037, Beijing n.s.; weekend: Beijing −0.099, Shanghai −0.115, Shenzhen −0.161	Weekend effects larger; weekend models explain more variance in Shanghai (adj. R^2 0.625 vs 0.297) and Shenzhen (0.429 vs 0.327)
	Light rain	n.s. on weekdays; weekend Beijing −0.033, Shanghai +0.037	Positive sign read as a shift from exposed modes
	RH, wind, temperature	RH +10% → −0.30% (Shanghai weekday); wind negative in all weekday models; temperature +10% → +0.08% (Beijing), +0.42% (Shanghai), n.s. in Shenzhen	Socio-economic covariates significant on weekdays only (pooled GLME)
Wei [3], bus/rail/ferry	Rain	Adult daily: weekday −0.017, weekend −0.003; stronger from CBD to suburbs (CV = 1.17); tertiary students AM peak +0.043 near campuses	No adult effect in AM peak; PM peak positive in inner city, negative in suburbs
	Heat index	Adult weekday +0.010; seniors +0.006 weekday, −0.007 weekend	Seniors most sensitive, spatially homogeneous (CV < 1)
Miao et al. [12], bus	Extreme cold / hot / rain	Weekday, unsheltered: −0.0157 / −0.00276 / −0.00623; shelter interaction +0.00430 / +0.00373 / +0.00181	Interactions n.s. on weekends; larger at low-frequency stops and stops with fewer transfers
Otim et al. [5], GPS, all modes	Temperature, air quality	Walking vs. cold: mild +23.4, warm +32.0, >25 °C +27.9 pp; bike −21.6 pp above 25 °C; air quality >300 µg/m³: bike +31.2, car −12.5, bus −9.6 pp	Distance dominant; air-quality effect attributed to traffic restrictions
Shokoohyar et al. [13], ride-sourcing	Extreme weather	Weekday: wait +21.6 s, in-vehicle +6.66 s, fare +$0.65; weekend: −20.16 s, −3.33 s, −0.86	Read as higher weekday and lower weekend demand

Source: created by the authors.

Table 3. Principal findings by evidence type and strength. Strength reflects design, sample, replication within the corpus and internal consistency

Finding	Evidence type	Supporting studies	Main qualification	Strength
Heavy rain and wind reduce aggregate transit demand	Observed	Jiang and Cai [4]; Wei [3]; Miao et al. [12]	Small magnitudes; light rain mixed	High
Net effect can be positive near origins of substituted short trips	Observed + interpretation	Jiang and Cai [4]; Wei [3]; Rowe et al. [6]	Substitution not observed directly	Moderate
Weekend (discretionary) demand more weather-sensitive	Observed	Jiang and Cai [4]; Shokoohyar et al. [13]	Reversed for Brisbane adults; Beijing exception	Moderate
Shelters offset part of weekday extreme-weather losses	Observed (panel FE)	Miao et al. [12]	Single region; sampled ridership	Moderate
Seniors most weather-sensitive; AM-peak adults insensitive	Observed	Wei [3]	Single city, one year	Moderate
Temperature effects depend on local climate	Observed + interpretation	Jiang and Cai [4]; Otim et al. [5]; Rowe et al. [6]	Comfort-index interpretation inconsistent	Moderate
Weather adds little to average accuracy, much during extremes	Model-derived	Rowe et al. [6]; Wang et al. [8]	Short windows; Yang et al. ablation inconsistent	Moderate
Adverse weather associated with earlier bus arrivals	Observed (descriptive)	Nguyen et al. [15]	No controls; no mechanism	Low
Response phase 2–3× recovery; elastic limit under typhoon	Observed + model-derived	Wang et al. [8]	One city, one super typhoon, speed only	Moderate
Catastrophic metro failure from aligned technical and organisational defects	Observed (case)	Xie and Peng [11]	Single event; qualitative	Moderate
Only salient rapid-onset hazards raise agency risk perception	Observed (survey)	Zhang [9]	Two waves; county-level exposure	Moderate
Robustness dominates agency adaptation practice	Document review + interpretation	Markolf et al. [1]; Zhang [9]	US-centric	Moderate
Operational preparedness preferred to design or rerouting	Expert judgement	Deveci et al. [10]	Hypothetical city	Low
Transit evacuation optimisable within 60 s under disruption	Model-derived (synthetic)	Yazdani et al. [16]	No real network or behaviour	Low

Source: created by the authors.