Australasian Transport Research Forum 2025 Proceedings
27-29 November, Auckland, New Zealand
Publication website: https://australasiantransportresearchforum.org.au/

Evacuation optimisation with shared autonomous vehicles

Yunlong Wang[1], Minh Kieu[1], Prakash Ranjitkar[1], Doug Wilson[1]

[1]Department of Civil and Environmental Engineering, University of Auckland, New Zealand
Email for correspondence (presenting author): yunlongwang430@gmail.com

Abstract

Traditional mass evacuation plans rely on fixed routes and manual control, making them slow to respond to the' uncertain and urgent nature of disasters. This study introduces an innovative and flexible evacuation strategy using Shared Autonomous Vehicles (SAVs) to support people without access to private transport. We propose a Transformer-Based Deep Reinforcement Learning method (T-DRL) to solve the SAV dispatching problem under changing demand, traf-fic, and vehicle availability. The model learns patterns between demand and SAV locations using attention mechanisms, while a masking process removes poor choices early to speed up decisions. Tested on real data from Auckland, New Zealand, in city-wide trials, it achieved up to 18.6% savings with route cost gaps under 1.7%. Finally, deploying several smaller TB-DRL models in parallel further cuts average pick-up time at each node by 21–27% compared with a single large model, without increasing cumulative fleet hours a lot. These findings demonstrate the strong practical value of TB-DRL for future city-scale evacuation planning and disaster response.

Keywords*: shared autonomous vehicles; evacuation strategies; disaster response; deep reinforcement learning.*

1 Introduction

Natural hazards are increasing globally, posing growing risks to lives, communities, and economies (Thakur, Ranjitkar, & Rashidi, 2022b). Auckland, New Zealand, offers a vital case study due to its dense population and proximity to multiple dormant volcanoes (Thakur, Ranjitkar, & Rashidi, 2022a). Emergency evacuation is crucial for reducing casualties in such events (Jafarian, Granberg, & Farahani, 2025). However, current evacuation strategies rely heavily on manual dispatch and the use of private cars. These methods often lead to severe congestion, delay emergency response, and exclude those without access to personal transport (Lee & Kockelman, 2024).

To address this, there is a growing need to explore public transportation options—particularly Shared Autonomous Vehicles (SAVs), assumed here to be fully autonomous Level 5 vehicles capable of operating without a human driver under all conditions—for evacuating those without access to vehicles. While large-scale deployment for disaster evacuation has not yet occurred, pilot services exist (e.g., Waymo in Phoenix, USA; Baidu Apollo in China), and this study treats SAVs as a given future technology for modelling purposes. Existing models, such as those using buses or ride-sharing services, offer partial solutions (Bolia et al., 2019; Meng, Dong, & Liu, 2023), but they often lack the flexibility and real-time adaptability required during large-scale emergencies. Additionally, vulnerable groups may face barriers in accessing information and services during government-led evacuations (Cheah, Shimul, Liang, & Phau, 2023). SAVs provide a promising alternative. They are flexible, tech-driven, and capable of addressing key gaps in current systems (Y. Wang, Kieu, & Ranjitkar, 2024; Y. Zhang, Carballo, Yang, & Takeda,

2023). For people without private transport, SAVs can offer fast and responsive evacuation support (Miller, Chng, & Cheah, 2022). However, using SAVs in disaster evacuation is a relatively new concept, with numerous challenges yet to be addressed.

We propose a coordinated evacuation model where individuals gather at assembly points—either predetermined or emerging based on demand. SAVs dynamically respond by adjusting real-time routes to transport people to safe zones or transit hubs. This is the Shared Autonomous Vehicles Evacuation Dispatching Problem (SAV-EDP). SAV-EDP is highly complex. Demand at assembly points changes constantly, and SAV availability, capacity, and location are uncertain. Disaster situations require rapid decisions, making manual or traditional algorithms insufficient (Dong, Yang, Sheng, Zhang, & Jing, 2024; W. Yang et al., 2025). A more intelligent, faster dispatch system is essential. In this study, the SAV evacuation service is assumed to be available to all residents without private vehicles, without the need for prior arrangement. Requests can be made at any time during the evacuation, and the dispatch algorithm adapts dynamically to changing demand patterns.

We introduce a novel solution: Transformer-Based Deep Reinforcement Learning (T-DRL) to address this need. This approach focuses on the critical 'first mile'—quickly picking up passengers for later evacuation stages. T-DRL learns patterns in SAV and demand locations, and adjusts real-time routes accordingly. It responds better than static models to changing traffic and demand, making it ideal for large-scale, dynamic emergencies. This study contributes three key innovations:

- It develops a distributed stochastic optimisation strategy for pre-disaster SAV deployment.
- It integrates the Transformer architecture into a deep reinforcement learning model, enabling efficient and real-time decision-making under dynamic conditions.
- It validates the model using real road networks and population data from Auckland, New Zealand, showing robust performance in multiple evacuation scenarios.

2 Literature review

This section reviews the current literature on disaster evacuation from three perspectives: public transport, Shared Autonomous Vehicles (SAVs), and evacuation strategies and algorithms.

2.1 Public transport-based evacuation

Evacuation modelling has evolved significantly in the last decade. Early work by X. Yang, Ban, and Mitchell (2018) introduced Dynamic System Optimisation to model multimodal evacuation using buses, subways, and private cars. Their method reduced evacuation time through iterative optimisation, which was validated in lower Manhattan.

Later studies shifted focus to dynamic resource allocation. For example, Bolia et al. (2019) utilised Mixed-Integer Linear Programming (MILP) to enhance both efficiency and resource utilisation. J. Alam, Habib, and Venkatadri (2019) developed a multimodal microsimulation framework integrating Dynamic Traffic Assignment. Their model showed that bus-only evacuations could be completed in 7 hours, versus 22 hours using private vehicles. Follow-up work (M. J. Alam & Habib, 2021) explored multiple scenarios and transport modes, but found that dynamic programming was often too computationally intensive for real-time use.

Recent work, such as Kabir, Mobin, Nayeem, Habib, and Rahman (2023), has adopted multi-objective and heuristic approaches to reduce evacuation times by up to 33%. These methods balance trade-offs like travel time, congestion, and capacity. However, gaps remain: current models often ignore first-mile/last-mile challenges, fail to fully integrate multimodal options, and lack flexibility in adapting to fluctuating demand and bus availability.

2.2 SAV-based evacuation

SAVs offer a promising new direction in emergency transport. Studies show they can serve as few as 10–30% of the vehicles needed in conventional evacuations (Miller et al., 2022). Early research focused on autonomous systems in normal traffic (Dresner & Stone, 2008), later expanding to emergency use cases (Paddeu, Shergold, & Parkhurst, 2020; Y. Zhang et al., 2023). M. Li, Xu, Liu, Sun, and Duan (2018) addressed shared rides in emergencies, considering intermediate pickups and vulnerable populations.

Technological advances, such as co-simulation frameworks (Syed & Hasan, 2023), have demonstrated that even modest SAV penetration (25%) can reduce conflicts by over 40%, underscoring the importance of communication systems in conjunction with driving algorithms. Practical deployment strategies have also emerged. Lee and Kockelman (2024) proposed a "Bus + SAV" hybrid system to support car-less populations, though results showed that excessive SAV fleets could lead to inefficiency. Other works on ride-sharing evacuation (Fiedler, Čertickỳ, Alonso-Mora, & Čáp, 2018; Lu, Liu, Wang, Zhou, & Hu, 2020; Naoum-Sawaya & Yu, 2017) have introduced dynamic assignment models that improve traffic flow and reduce delays. Public attitudes are also critical. Cheah et al. (2023) found that perceived usefulness influences acceptance. Many still prefer private vehicles or staying with relatives during disasters (Wong, Broader, Walker, & Shaheen, 2023), suggesting that assembly points must be local and familiar. Meng et al. (2023) demonstrated that ride-sharing is most effective when approximately 60% of people participate, but cautioned that excessive vehicle usage can exacerbate congestion. Overall, SAVs and ride-sharing offer strong potential, but more work is needed to handle spatial imbalance, forecast dynamic demand, and align tech solutions with human behaviour.

2.3 Evacuation strategies and algorithms

Evacuation modelling combines simulation tools and optimisation algorithms. Platforms like SUMO, MATSim, and VISSIM model traffic flow and human behaviour in virtual disaster scenarios (Rohaert et al., 2023; Wahlqvist et al., 2021; Q. Wang & Nie, 2023). These tools visualise bottlenecks and patterns but require detailed data and high computing power, making generalisation difficult. Optimisation algorithms aim to find the best routes under constraints like time, capacity, and road limits. However, exact methods often fail under real-time uncertainty. Heuristic methods—such as those by Naoum-Sawaya and Yu (2017) and Fiedler et al. (2018)—have improved scalability and speed through techniques like incremental cost and maximum matching.

Advanced strategies include Approximate Dynamic Programming (Lei, Jiang, & Ouyang, 2019) and hybrid methods. For instance, Yin, Xu, Wang, Yu, and Cheng (2024) developed two-stage robust optimisation; others used Ant Colony, Genetic Algorithms, and Particle Swarm Optimisation to handle dynamic routing (Dong et al., 2024; Jafarian et al., 2025; W. Yang et al., 2025). Hybrid models combining Simulated Annealing, Tabu Search, and other metaheuristics further improved efficiency (Jiang, Shen, & Ouyang, 2023; Sadeghi, Aros-Vera, Mosadegh, & YounesSinaki, 2023). Deep learning has recently transformed the field. Rahman and Hasan (2023) used transfer learning to predict demand from sparse data. Z. Zhang, Fei, and Fu (2024) applied deep reinforcement learning to model pedestrian-vehicle interactions, improving efficiency by 23.6%. Takabatake, Asai, Kakuta, and Hasegawa (2025) merged Q-learning and agent-based models to improve decision speed by 40% under road congestion and capacity constraints.

Despite these advances, existing strategies rarely address SAV coordination for first-mile evacuation. This study fills that gap by introducing the T-DRL approach—an adaptive, intelligent solution designed to improve real-time SAV dispatch in high-pressure, uncertain environments.

3 Preliminary

3.1 Road networks and dynamic traffic modelling

To create an interactive environment for training the DRL model, we use the SUMO microsimulation software to build a realistic traffic network (Lopez et al., 2018). Figure 1(a) shows population density in Auckland, New Zealand, based on census data (Lindsay, Leonard, Smid, & Hayward, 2011). We focus on the North Shore area due to its high population density and clear one-directional evacuation flow towards the north, making it a suitable case study.
The road network for this area was extracted from real maps and imported into SUMO, as shown in Figure 1(b). The total length of motor vehicle roads is 345.3 km, with motorways accounting for 29.5% (speed limits: 80–100 km/h), and residential streets making up 45.8% (speed limit: 50 km/h). Based on road type and vehicle suitability, all motor vehicle roads were designated as potential SAV routes.

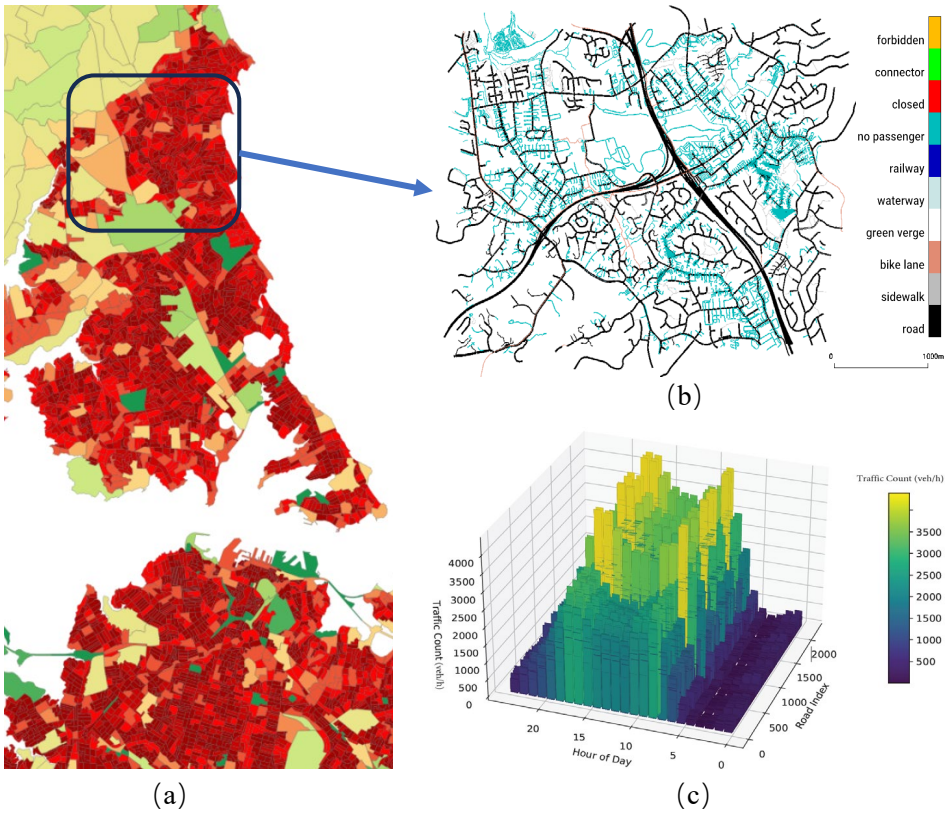

Figure 1: Modelling of the study area. (a) Population density map for Auckland (2023 census). (b) Road network imported into SUMO. (c) Hourly traffic flow for each road segment.

To enable route adaptation under congestion, we incorporate real-time traffic data into T-DRL. This allows the model to respond dynamically to varying road conditions and validate its effectiveness in realistic traffic scenarios. Daily traffic patterns for Auckland were obtained from Planning and Investment Auckland Transport (2021), including average volumes, peak periods, and congestion levels. Additional hourly data were retrieved from the Auckland Transport Asset Inventory Database (Auckland Transport, 2021), referred to as $data_{24h}$.
These datasets were matched to the study area's network (Figure 1(b)) to build a 24-hour dynamic traffic profile, shown in Figure 1(c). This traffic model forms the basis for training and evaluating the T-DRL strategy under dynamic conditions.In the event of network disruptions, such as blocked road segments caused by earthquakes or other hazards, SAVs are assumed

to re-route in real-time using updated travel time data. As travel time is a key component of the reward function (Section 4.1), such obstructions are reflected in the decision-making process, encouraging the selection of alternative feasible paths.

3.2 SAV evacuation dispatch problem modelling

This paper models the network using a weighted directed graph $G = (N, E)$. The vertex set N includes critical geographical and functional nodes within the evacuation system, primarily divided into the following two categories:

- V: Starting locations and attributes SAVs
  - Each node $v_i \in V$ represents the initial position of an available AV, with its passenger capacity c_i being a random variable, typically drawn from the discrete set {4,6,8}.
  - SAVs that have picked up passengers will travel to a safe node and be marked as unavailable. After the delivery task, they will be marked as available again.
- X: Assembly nodes
  - Each node $x_j \in X$ represents a predetermined assembly location for receiving evacuees.
  - The corresponding demand d_j denotes the number of people needing transport, fluctuating with changes in time and population distribution.

In this study, assembly points are assumed to be fixed, pre-identified locations such as bus stops, transit hubs, metro or train stations where evacuees naturally gather. SAVs provide direct, point-to-point transport from these locations to nearby safe nodes, and the optimisation focuses solely on this first-mile stage with dynamic SAV allocation. After completing a trip to a safe node, the SAV is marked as available again and immediately reassigned to the next assembly point according to dynamic demand and operational priorities (e.g., hazard level, number of evacuees, or presence of vulnerable groups). Safe nodes are fixed locations such as bus stations, transit hubs, or metro/train stations located outside the immediate hazard area, serving as transfer points for onward evacuation.
Edge set E describes feasible paths from SAVs' locations to assembly nodes.

- Each edge $e_{ij} \in E$ corresponds to a segment of the travel route, with a non-negative weight w_{ij}, representing the distance and travel time from i to j.
- Some edges may experience congestion due to unforeseen events, and such paths will be assigned additional travel time.

In this study, 'safe nodes' are assumed to be pre-designated nearby safe zones or transit hubs outside the immediate hazard area, serving as transfer points to other modes (e.g., buses) for subsequent evacuation. The optimisation focuses on this first-mile stage, dynamically assigning SAVs to assembly points and safe nodes to maximise pick-up efficiency.

4 Methodology

In the high-time-demand evacuation scenario, demand dynamics and the stochastic SAV state often lead to difficulties balancing the solution efficiency and algorithmic stability of classical optimisation methods. To address the limitations of MILP, this paper proposes a DRL strategy based on the Transformer(T-DRL). T-DRL aims to facilitate fast and optimal decision-making while ensuring evacuation objectives are met, minimising travel costs and balancing the utilisation of excess capacity.

4.1 Markov decision process-based DRL framework

Deep Reinforcement Learning (DRL) provides strong adaptability in uncertain and dynamic environments by learning from interactions and continuously updating its policy. To solve the evacuation dispatching problem using DRL, we model it as a Markov Decision Process (MDP), defined as $M = \{S, A, \tau, R\}$.

State S_t: Each state $S_t = (V, X)$ includes the positions and statuses of available SAVs (V) and the evacuation demand at assembly nodes (X). The network structure follows the graph model $G = (N, E)$ from Section 3.

- V includes: *loc_nor* (normalised node coordinates), *capacity* (SAV_i's seat count), and *eva_a* (availability status: 1 = available, 0 = unavailable).
- X includes: *demand_left* (remaining evacuees at node j), *dis* (distance between SAV_i and node j), *ttm* (travel time), *SAV_sets* (SAVs assigned to node j), and *loc_node* (location of node j).

Action A_t: At each step, the agent selects a pair (v_{i_t}, x_{j_t}), assigning SAV_i to assembly node x_j. This constructive approach allows step-by-step route building under uncertainty.

Transition τ_t: The next state S_{t+1} updates based on the current action. The demand at node j and the set of assigned SAVs are updated using:

$$\text{demand}_{j_{t+1}} = \begin{cases} \text{demand_ini}, & \text{if } i_{t+1} = \text{SAV}_{\text{virtual}} \\ \text{demand}_{j_t} - \text{capacity}_{ij_t}, & \text{if } i_{t+1} = v_i \\ \text{demand}_{j_t}, & \text{otherwise} \end{cases} \tag{1}$$

$$\text{SAV}_{\text{set}_{t+1}} = \begin{cases} [\text{SAV}_{\text{set}_t}, v_{ij_{t+1}}], & \text{if } i_{t+1} = v_i \\ [\text{SAV}_{\text{set}_t}, v_{i_t}], & \text{otherwise} \end{cases} \tag{2}$$

Reward R_t: The reward function encourages efficient and balanced evacuation. It is defined as:

$$R_t = \sum_{i=1}^{m} \sum_{j=1}^{n} [\alpha \cdot (\text{dis}_i + \text{ttm}_i) + \beta \cdot |\text{capacity}_i - \text{demand}_j|] \tag{3}$$

Policy p_t: We use a stochastic policy $p(s_{T_{\text{end}}} \mid s_{T_{\text{ini}}})$ based on an attention-based encoder-decoder. It models the joint probability of decisions from the initial state to the terminal state. The probabilities of choosing $SAVs$ and nodes are computed as:

$$\begin{aligned} p(s_{V_{T_{\text{end}}}} | s_{V_0}) &= \prod_{t=0}^{T_{\text{end}}} \pi_\theta(a_{V_t} | s_{V_t}) \cdot p(s_{V_{t+1}} | s_{V_t}, a_{V_t}) \\ p(s_{X_{T_{\text{end}}}} | s_{X_0}) &= \prod_{t=0}^{T_{\text{end}}} \pi_\theta(a_{X_t} | s_t) \cdot p(s_{X_{t+1}} | s_{X_t}, a_{X_t}) \end{aligned} \tag{4}$$

Here, $\pi_\theta(a_{V_t} | s_{V_t})$ is the policy for selecting an SAV, and $\pi_\theta(a_{X_t} | s_t)$ is for selecting an assembly node, both based on the current state.

Encoder

This study builds on the attention model proposed by Kool, Van Hoof, and Welling (2018), introducing a tailored encoder to process the heterogeneous information from SAVs and assembly nodes (Fig. 2). The goal is to embed global context into each node by stacking multi-head attention (MHA) and feed-forward (FF) layers.

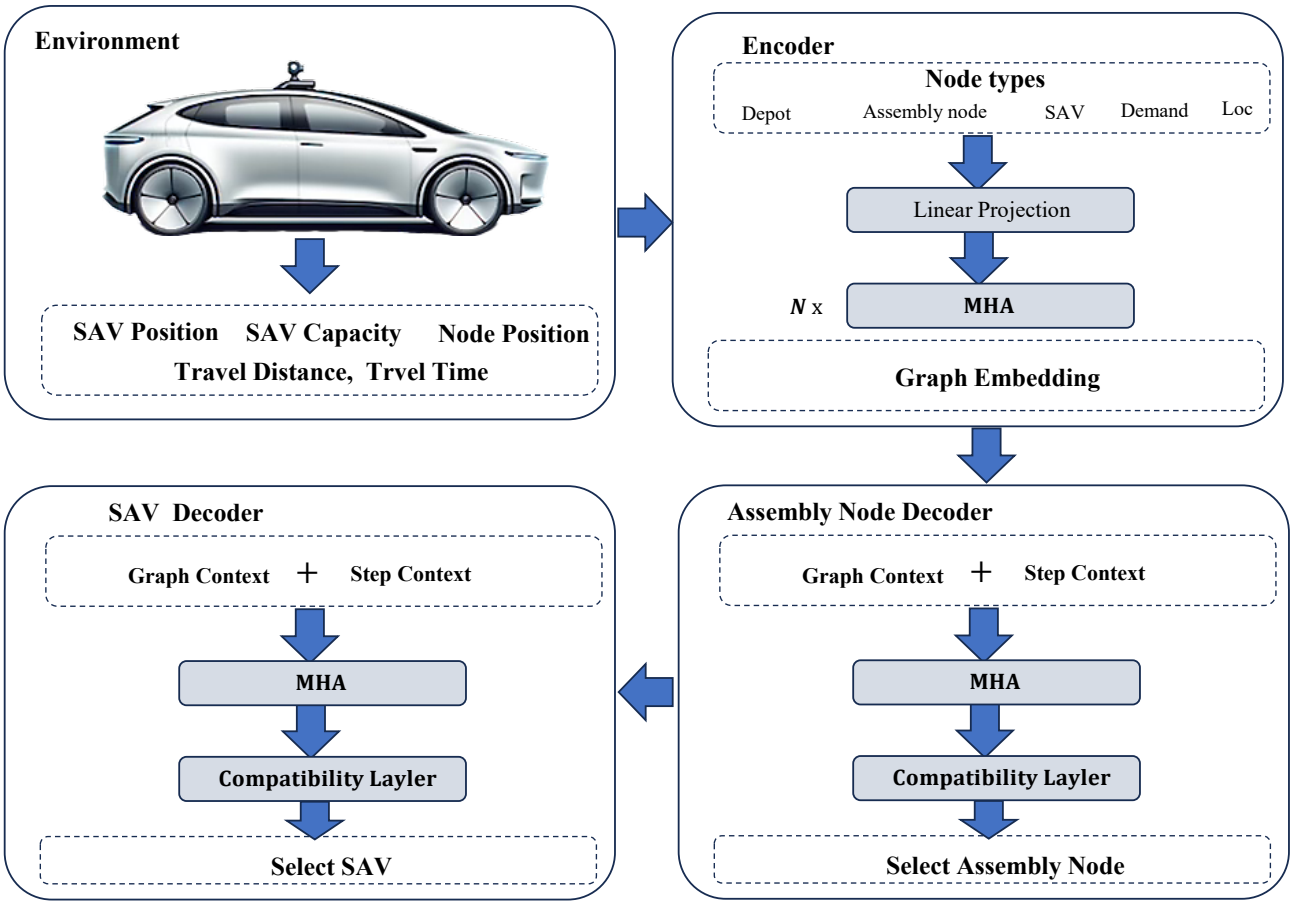

Figure 2: T-DRL strategy framework.

We first encode SAV states and assembly node features separately. Let $x_{SAV} \in \mathbb{R}^m$ denote the SAV state vector. It is projected into a higher-dimensional space using a linear layer:

$$e_{SAV} = f_{SAV}(x_{SAV}) = W_{SAV} \cdot x_{SAV} + b_{SAV} \tag{5}$$

Similarly, assembly nodes are encoded as:

$$e_{node} = f_{node}(x_{node}) = W_{node} \cdot x_{node} + b_{node} \tag{6}$$

These are combined to form the graph embedding:

$$e_{graph} = e_{SAV} + e_{node} \tag{7}$$

The combined embedding is passed through N stacked attention layers. Each layer includes MHA and FF modules. Node states are batch-normalised and updated as:

$$\hat{h}^i(l) = \mathrm{BN}(l)\big(h^i(l-1) + \mathrm{MHA}_i(l)(h^1(l-1), ..., h^n(l-1))\big) \tag{8}$$

After N layers, the final embedding $h_i^{(n)}$ reflects both SAV and node information. The average of all node embeddings, $\bar{h}^{(n)}$, is used to represent the global graph context.

Decoder

The decoder uses encoder outputs to compute selection probabilities for assembly nodes and SAVs. As SAV locations, capacities, and node demands vary, T-DRL must effectively match SAVs to high-need locations.

Assembly node decoder: To select an assembly node, we consider the current demand distribution, SAV assignments, distances, and travel times. We define a context vector:

$$N_t = [d_{r,t}, Rs_t, dis_t, ttm_t, loc_t] \tag{9}$$

where $d_{r,t}$ is the remaining demand, Rs_t is the current SAV assignment status, and others represent spatial-temporal features.
The node embedding is:

$$FN_t = \mathsf{FF}(W_1 N_t + b_1) \tag{10}$$

For the SAV distribution, the spatial features of each SAV are:

$$G_{v_j} = [hv_0, ..., hv_j, ..., hv_m] \tag{11}$$

and the pooled distribution embedding is:

$$FSV_t = \mathsf{FF}(W_2 \cdot \max(G_{v_j}) + b_2) \tag{12}$$

The final node selection probability is computed via:

$$p_{node} = \mathsf{softmax}(W_3[FN_t, FSV_t] + b_3) \tag{13}$$

SAV decoder: Given a selected node, the model selects an SAV based on its context:

$$F_{vt} = [\bar{h}_t^{(n)}, h_{last,t-1}, D_t] \tag{14}$$

where D_t is the demand at the selected node and $h_{last,t-1}$ is the previous step's embedding.
An MHA layer evaluates compatibility between this context and the global embedding:

$$\bar{p}_t = \mathsf{softmax}\left(C \cdot \tanh(\hat{F}_{SVt} W_{Q_v} \cdot (h_g^{(n)} W_{K_v})^U) - \infty \cdot \mathsf{mask}_{general}\right) \tag{15}$$

Here, W_{Q_v} and W_{K_v} are learnable parameters. The constant C regulates entropy, and the masking term enforces constraints.

4.2 Training

We train the T-DRL model using a REINFORCE-based policy gradient method with a rollout baseline, following the approach in Kool et al. (2018); Y. Wang et al. (2024). At each step, random scenario parameters ξ (e.g., SAV availability, demand) generate new instances. The current policy p_θ samples a solution trajectory π_i, and the baseline policy $p_{\theta_{\mathsf{BL}}}$ produces a greedy trajectory π_{BL_i}. Rewards for both are computed, and the policy is updated based on their difference using the Adam optimiser. After each epoch, a one-sided paired t-test is used to compare the new policy against the baseline. If a significant improvement is detected $(p < \gamma)$, the baseline is updated.

5 Numerical experiments

This section describes the results of our numerical experiments for the case study in Auckland, New Zealand.
We evaluate the performance of the Transformer-based Deep Reinforcement Learning (T-DRL) strategy and a classical MILP model on static instances of the SAV-EDP. T-DRL is tested using two action selection methods: *Greedy* and *Sample*, as described by J. Li et al. (2021). The Greedy strategy selects the action with the highest predicted reward, while the Sample strategy selects actions probabilistically based on the reward distribution.

Algorithm 1 T-DRL policy training with rollout baseline

Require: Epochs E, steps per epoch K, transition step T, batch size B, significance level γ

Initialise policy θ, set baseline $\theta_{\text{BL}} \leftarrow \theta$
for epoch = 1 to E **do**
  for step = 1 to K **do**
    for each batch item $i = 1$ to B **do**
      Generate instance s_i with random scenario ξ_i
      if step $\leq T$ **then**
        Sample π_i using p_θ
      else
        Sample π_i using standard mask and MHA
      end if
      Generate baseline π_{BL_i} using greedy rollout
      Compute rewards R_i, R_{BL_i}
    end for
    Compute gradient: $\nabla L = \sum (R_i - R_{\text{BL}_i}) \nabla \log p_\theta(\pi_i)$
    Update θ via Adam optimiser
  end for
  if t-test result $< \gamma$ **then**
    Update baseline: $\theta_{\text{BL}} \leftarrow \theta$
  end if
end for

Figure 3 presents the total cost and route cost for MILP and both T-DRL variants. Several observations can be made. First, the Sample strategy consistently outperforms Greedy in both cost metrics, indicating improved decision diversity and robustness. Second, while MILP achieves slightly lower route costs, T-DRL—particularly with Sample—offers better total cost efficiency, making it a practical alternative. Finally, runtime analysis on 128 static V100-N10 instances shows that Sample requires an average of 7.68 seconds per instance, while Greedy completes in 1.38 seconds, offering flexibility depending on time constraints. These results demonstrate that T-DRL is effective in solving static evacuation dispatch problems, striking a balance between solution quality and computational efficiency. The remainder of this paper reports results based on the T-DRL-Sample strategy.

5.1 Practical population-based evacuation simulation

In addition to evaluating evacuation strategies based on a real road network, this study also verifies the practical applicability of the proposed strategies using real population distribution data from Auckland, New Zealand (Humanitarian Data Exchange, 2023). Figure 3a shows the population distribution within the study area, with a total population of 38,993 people.

Following M. J. Alam and Habib (2021); Meng et al. (2023), the proportion of the population requiring evacuation by SAVs is set to 20%. This setting considers both the generally high preference for private vehicles and the demonstrated public acceptance of officially arranged transport among non-vehicle residents in Auckland (Thakur et al., 2022a, 2022b). SAVs, as a new form of government-coordinated shared mobility, are expected to be accepted by this group, especially when safety and reliability are ensured.

Subsequently, the K-means clustering method is employed to determine the assembly points, as shown in Fig. 3b. Since this paper focuses on SAV dispatch strategies, the following assumptions are made for the evacuation simulation:

1. SAVs are only responsible for the “first-mile" transport (Lee & Kockelman, 2024). Des-

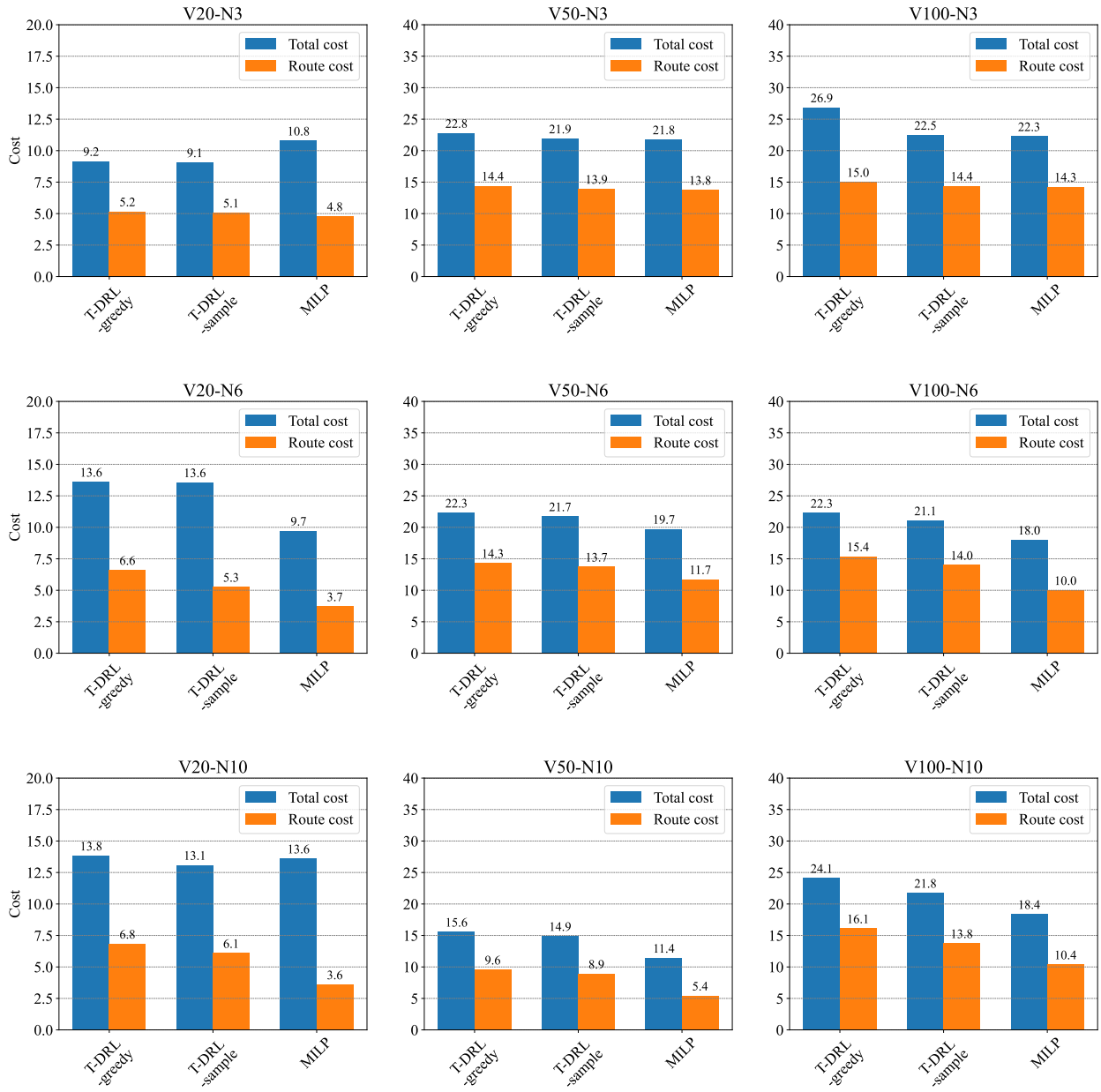

Figure 3: Different SAV-EDP strategies results comparison.

tinations may include temporary transfer stations or public transport stations, as buses or trains have greater advantages for evacuating over longer distances to safe locations (Kabir et al., 2023).

2. The total number of SAVs within the study area remains constant. After completing their assigned trips, SAVs become available again, and the average travel time for each SAV from the assembly node to the public transport station is 30 minutes.

3. Evacuation will follow a predefined schedule, with residents arriving at the assembly points according to the timetable, ensuring that SAVs do not need to wait.

4. Demand at each assembly node will remain relatively stable within specific time intervals, avoiding scenarios where demand significantly exceeds SAV capacity.

Considering the area size and population, we tested two scenarios with 100 and 200 SAVs,

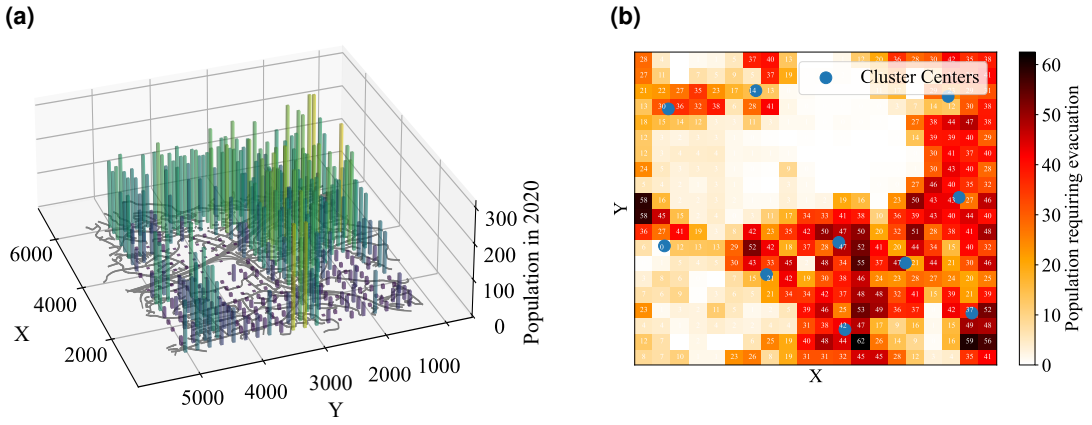

Figure 4: Visualisation of population distribution .(a) Total population distribution in the study area. (b) Clustering based on population requiring evacuation.

respectively, with distributed model strategies. Here, distributed modelling refers to a unique strategy where we segment the TB-DRL decision space into smaller, manageable parts and then utilise models trained on smaller instances to address larger problems. This strategy enables parallel processing, allowing the system to scale efficiently without incurring excessive computational load. For the scenario with 100 SAVs, the selected TB-DRL strategies are V100-N10, 2 × V50-N10, and 5 × V20-N10. For the scenario with 200 SAVs, the selected strategies were 2 × V100-N10, 4 × V50-N10, and 10 × V20-N10. Figure 5 shows the results of multi-scenario evacuation with 100 and 200 SAVs when the traffic is low, medium and high.

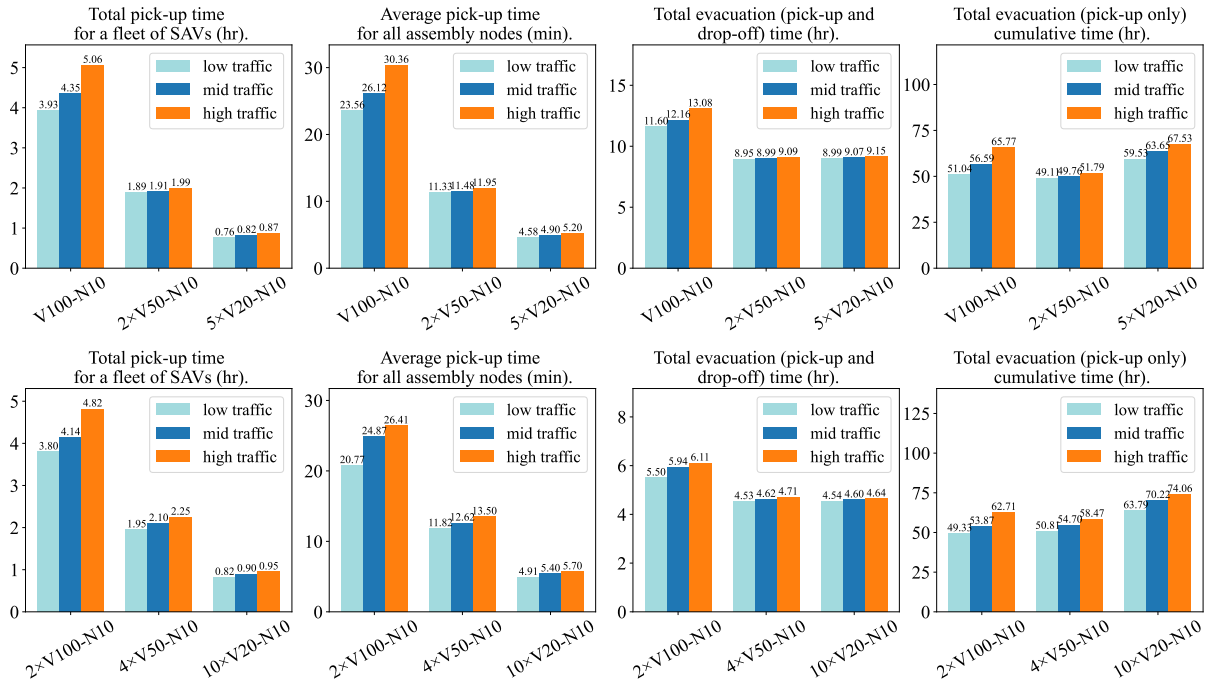

Figure 5: Test results of multi-scenario evacuation for 100 and 200 SAVs.

Figure 5 shows that traffic flow has a clear impact on all outcomes, as increased traffic congestion leads to longer evacuation times to varying degrees. Compared to the V100-N10 scenario, the 2 × V50-N10 and 5 × V20-N10 scenarios experience less impact from traffic flow, for example, in terms of total pick-up time, average pick-up time, and total evacuation time (the first

three columns). This is because using more, smaller models increases the number of assembly nodes, which spreads out the evacuees and allows more parallel pick-ups. As a result, each SAV travels a shorter route to serve fewer people, reducing delays and improving overall traffic flow. A similar effect is seen in the 200 SAV case, where distributed models help reduce total evacuation time by enabling more simultaneous evacuations.
However, in the scenario with 200 SAVs, the total cumulative evacuation time-which is the combined operating time (pickup and drop-off) of all SAVs—increases, as shown in the last column of Fig. 5. This is because more assembly nodes lead to more dispersed SAV routes, including those to less populated or remote areas. These longer and less efficient routes increase the total operating time, even if each individual evacuation is faster. Detailed results from these tests are provided in Table 1 and 2.

Table 1: Test results of multi-scenario evacuation for 100 SAVs.

Test scenarios	Scale of instances	Total pick-up time for a fleet of SAVs (hr).	Average pick-up time for all assembly nodes (min).	Total evacuation (pick-up and drop-off) time (hr).	Total evacuation (pick-up only) cumulative time (hr).
100 SAVs low traffic	V100-N10	3.93	23.56	11.60	51.04
	V50-N10-1	1.82	10.93	8.87	23.68
	V50-N10-2	1.96	11.74	9.04	25.43
	V20-N10-1	0.55	3.32	8.66	8.64
	V20-N10-2	0.76	4.55	8.98	11.84
	V20-N10-3	0.72	4.34	8.93	11.28
	V20-N10-4	0.65	3.87	8.80	10.06
	V20-N10-5	1.14	6.82	9.57	17.72
100 SAVs mid traffic	V100-N10	4.35	26.12	12.16	56.59
	V50-N10-1	1.82	10.93	8.87	23.67
	V50-N10-2	2.01	12.04	9.11	26.08
	V20-N10-1	0.79	4.76	9.04	12.37
	V20-N10-2	0.90	5.40	9.20	14.03
	V20-N10-3	0.81	4.84	9.06	12.57
	V20-N10-4	0.80	4.78	9.04	12.44
	V20-N10-5	0.78	4.71	9.02	12.24
100 SAVs high traffic	V100-N10	5.06	30.36	13.08	65.77
	V50-N10-1	1.93	11.57	9.00	25.06
	V50-N10-2	2.06	12.34	9.17	26.73
	V20-N10-1	0.82	4.93	9.08	12.82
	V20-N10-2	0.97	5.85	9.32	15.21
	V20-N10-3	0.86	5.14	9.14	13.37
	V20-N10-4	0.84	5.06	9.11	13.16
	V20-N10-5	0.83	4.99	9.10	12.97

In summary, using multiple distributed smaller models can reduce the average evacuation time at each assembly node but increase the total cumulative evacuation time. Medium-sized distributed models such as V50-N10 would balance this trade-off. Multiple V50-N10 models reduce evacuation time by 26.63% compared to one V100-N10 model with 100 SAVs, and 21.04% with 200 SAVs. On the other hand, V50-N10 models reduce the total cumulative time by 21.01% compared to V20-N10 models with 100 SAVs and by 21.19% with 200 SAVs.
Although the distributed TB-DRL model described above demonstrates significant advantages in improving evacuation efficiency and resource utilisation, its real-world deployment still requires careful consideration of several challenges in emergency scenarios. These include the limitations of communication network capacity and stability during disasters, constraints on computational resources affecting model inference, and the difficulties of resource reallocation when some nodes fail. These challenges can be gradually addressed by leveraging the computing platforms of Level-3 SAVs and designing scheduling mechanisms with dynamic resource allocation capabilities.

Table 2: Test results of multi-scenario evacuation for 200 SAVs.

Test scenarios	Scale of instances	Total pick-up time for a fleet of SAVs (hr).	Average pick-up time for all assembly nodes (min).	Total evacuation (pick-up and drop-off) time (hr).	Total evacuation (pick-up only) cumulative time (hr).
200 SAVs low traffic	V100-N10-1	3.91	21.41	5.57	25.43
	V100-N10-2	3.68	20.13	5.43	23.90
	V50-N10-1	1.96	11.75	4.52	12.72
	V50-N10-2	2.34	14.06	4.77	15.23
	V50-N10-3	1.52	9.09	4.23	9.85
	V50-N10-4	2.00	12.38	4.59	13.00
	V20-N10-1	0.73	4.37	4.47	5.68
	V20-N10-2	0.75	4.51	4.48	5.86
	V20-N10-3	0.86	5.19	4.57	6.74
	V20-N10-4	0.75	4.48	4.48	5.82
	V20-N10-5	0.93	5.55	4.62	7.21
	V20-N10-6	0.75	4.53	4.49	5.88
	V20-N10-7	0.83	4.95	4.54	6.44
	V20-N10-8	0.87	5.20	4.58	6.76
	V20-N10-9	0.85	5.12	4.57	6.66
	V20-N10-10	0.86	5.18	4.57	6.73
200 SAVs mid traffic	V100-N10-1	4.35	26.12	6.08	28.30
	V100-N10-2	3.94	23.61	5.81	25.57
	V50-N10-1	1.76	10.59	4.40	11.47
	V50-N10-2	2.80	16.81	5.07	18.21
	V50-N10-3	1.83	11.00	4.44	11.91
	V50-N10-4	2.02	12.10	4.56	13.11
	V20-N10-1	0.70	4.20	4.44	5.45
	V20-N10-2	0.90	5.39	4.60	7.00
	V20-N10-3	1.04	6.23	4.71	8.10
	V20-N10-4	0.89	5.33	4.59	6.93
	V20-N10-5	0.96	5.78	4.65	7.52
	V20-N10-6	0.72	4.34	4.46	5.64
	V20-N10-7	0.88	5.25	4.58	6.83
	V20-N10-8	0.95	5.72	4.64	7.44
	V20-N10-9	1.13	6.79	4.78	8.82
	V20-N10-10	0.83	4.99	4.55	6.48
200 SAVs high traffic	V100-N10-1	5.06	27.69	6.25	32.89
	V100-N10-2	4.59	25.12	5.97	29.83
	V50-N10-1	1.87	11.22	4.46	12.15
	V50-N10-2	3.04	18.22	5.22	19.73
	V50-N10-3	1.97	11.80	4.53	12.78
	V50-N10-4	2.13	12.75	4.63	13.81
	V20-N10-1	0.73	4.41	4.47	5.73
	V20-N10-2	0.95	5.67	4.64	7.38
	V20-N10-3	1.11	6.63	4.76	8.62
	V20-N10-4	0.95	5.69	4.64	7.40
	V20-N10-5	1.02	6.13	4.70	7.97
	V20-N10-6	0.77	4.62	4.50	6.00
	V20-N10-7	0.92	5.49	4.61	7.14
	V20-N10-8	1.00	5.98	4.68	7.77
	V20-N10-9	1.18	7.10	4.82	9.23
	V20-N10-10	0.87	5.24	4.58	6.82

6 Conclusion

This study introduces a distributed stochastic optimisation strategy for pre-disaster evacuation using Shared Autonomous Vehicles (SAVs), addressing the SAV Evacuation Dispatching Problem (SAV-EDP) for populations without private transport. We propose a Transformer-Based Deep Reinforcement Learning (T-DRL) framework that leverages an attention-based architec-

ture to dynamically assign SAVs to high-demand areas under uncertainty. Using a real-world case study based on Auckland's North Shore—comprising a 345 km road network, population data for 38,993 residents, and 24-hour traffic flow profiles—we evaluated the performance of T-DRL in static scenarios. Compared with a classical MILP baseline, T-DRL (especially under the Sample strategy) demonstrated superior total cost efficiency while achieving competitive routing performance. Additionally, T-DRL offers faster computation and greater adaptability, making it a promising approach for time-sensitive evacuation planning.

This paper represents the first known application of DRL to pre-disaster evacuation using SAVs. However, several avenues remain for future research. First, individual differences in evacuees—such as age, health status, or mobility—should be integrated to prioritise vulnerable groups. Second, this study focuses solely on the initial dispatch stage; future work could explore hybrid scheduling models combining SAVs with buses or other modes. Lastly, our scenario assumed that 20% of the population lacked vehicle access; future studies could assess the feasibility of SAVs as the primary mode of transport for wider-scale evacuations.

References

Alam, J., Habib, M. A., & Venkatadri, U. (2019). Development of a multimodal microsimulation-based evacuation model. *Transportation research record*, *2673*(10), 477–488.

Alam, M. J., & Habib, M. A. (2021). A dynamic programming optimization for traffic microsimulation modelling of a mass evacuation. *Transportation research part D: transport and environment*, *97*, 102946.

Auckland Transport. (2021). *Auckland transport open gis data: Average daily traffic counts.* Open GIS Data Portal. Retrieved from `https://data-atgis.opendata.arcgis.com/datasets/ATgis::trafficservice/explore?layer=0&showTable=true` (This data is obtained from Auckland Transport Asset Inventory Database)

Bolia, N. B., et al. (2019). Operating strategies of buses for mass evacuation. *Safety science*, *111*, 167–178.

Cheah, I., Shimul, A. S., Liang, J., & Phau, I. (2023). Consumer attitude and intention toward ridesharing. In *Relationship marketing in franchising and retailing* (pp. 3–24). Routledge.

Dong, K., Yang, D., Sheng, J., Zhang, W., & Jing, P. (2024). Dynamic planning method of evacuation route in dam-break flood scenario based on the aco-ga hybrid algorithm. *International Journal of Disaster Risk Reduction*, *100*, 104219.

Dresner, K., & Stone, P. (2008). A multiagent approach to autonomous intersection management. *Journal of artificial intelligence research*, *31*, 591–656.

Fiedler, D., Čertickỳ, M., Alonso-Mora, J., & Čáp, M. (2018). The impact of ridesharing in mobility-on-demand systems: Simulation case study in prague. In *2018 21st international conference on intelligent transportation systems (itsc)* (pp. 1173–1178).

Humanitarian Data Exchange. (2023). *New zealand: High resolution population density maps + demographic estimates.* `https://data.humdata.org/dataset/new-zealand-high-resolution-population-density-maps-demographic-estimates`. (Accessed: date-of-access)

Jafarian, A., Granberg, T. A., & Farahani, R. Z. (2025). The effect of geographic risk factors on disaster mass evacuation strategies: A smart hybrid optimization. *Transportation Research Part E: Logistics and Transportation Review*, *193*, 103825.

Jiang, Z., Shen, S., & Ouyang, Y. (2023). Planning of reliable targeted evacuation under the threat of disasters. *Transportation research part C: emerging technologies*, *153*, 104197.

Kabir, M., Mobin, J., Nayeem, M. A., Habib, M. A., & Rahman, M. S. (2023). Multi-objective optimization and heuristic based solutions for evacuation modeling. *Transportation research interdisciplinary perspectives*, *18*, 100798.

Kool, W., Van Hoof, H., & Welling, M. (2018). Attention, learn to solve routing problems! *Proc. Int. Conf. Learn. Represent.*, 1–25.

Lee, J., & Kockelman, K. M. (2024). Leveraging shared autonomous vehicles for vulnerable populations during pre-disaster evacuation. *Transportation Planning and Technology*, *47*(8), 1331–1363.

Lei, C., Jiang, Z., & Ouyang, Y. (2019). Path-based dynamic pricing for vehicle allocation in ridesharing systems with fully compliant drivers. *Transportation research procedia*, *38*, 77–97.

Li, J., Xin, L., Cao, Z., Lim, A., Song, W., & Zhang, J. (2021). Heterogeneous attentions for solving pickup and delivery problem via deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, *23*(3), 2306–2315.

Li, M., Xu, J., Liu, X., Sun, C., & Duan, Z. (2018). Use of shared-mobility services to accomplish emergency evacuation in urban areas via reduction in intermediate trips—case study in xi'an, china. *Sustainability*, *10*(12), 4862.

Lindsay, J., Leonard, G., Smid, E., & Hayward, B. (2011). Age of the auckland volcanic field: a review of existing data. *New Zealand Journal of Geology and Geophysics*, *54*(4), 379–401.

Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., ... Wießner, E. (2018). Microscopic traffic simulation using sumo. In *2018 21st international conference on intelligent transportation systems (itsc)* (pp. 2575–2582).

Lu, W., Liu, L., Wang, F., Zhou, X., & Hu, G. (2020). Two-phase optimization model for ridesharing with transfers in short-notice evacuations. *Transportation research part C: emerging technologies*, *114*, 272–296.

Meng, L., Dong, Z. S., & Liu, D. (2023). Ridesharing evacuation models of disaster response. *Computers & Industrial Engineering*, *184*, 109565.

Miller, K., Chng, S., & Cheah, L. (2022). Understanding acceptance of shared autonomous vehicles among people with different mobility and communication needs. *Travel Behaviour and Society*, *29*, 200–210.

Naoum-Sawaya, J., & Yu, J. Y. (2017). Ridesharing for emergency evacuation. *INFOR: Information Systems and Operational Research*, *55*(4), 339–358.

Paddeu, D., Shergold, I., & Parkhurst, G. (2020). The social perspective on policy towards local shared autonomous vehicle services (lsavs). *Transport Policy*, *98*, 116–126.

Planning and Investment Auckland Transport. (2021). *Auckland region transport strategic case 2021-2031.* Auckland Transport. Retrieved from `https://at.govt.nz/media/1988303/auckland-region-transport-strategic-case.pdf` (Final - endorsed by Waka Kotahi (NZ Transport Agency) on 11 March 2022)

Rahman, R., & Hasan, S. (2023). A deep learning approach for network-wide dynamic traffic prediction during hurricane evacuation. *Transportation research part C: emerging technologies*, *152*, 104126.

Rohaert, A., Kuligowski, E. D., Ardinge, A., Wahlqvist, J., Gwynne, S. M., Kimball, A., ... Ronchi, E. (2023). Traffic dynamics during the 2019 kincade wildfire evacuation. *Transportation research part D: transport and environment*, *116*, 103610.

Sadeghi, A., Aros-Vera, F., Mosadegh, H., & YounesSinaki, R. (2023). Social cost-vehicle routing problem and its application to the delivery of water in post-disaster humanitarian logistics. *Transportation research part E: logistics and transportation review*, *176*, 103189.

Syed, Z. E. M., & Hasan, S. (2023). A co-simulation study to assess the impacts of connected and autonomous vehicles on traffic flow stability during hurricane evacuation. *arXiv preprint arXiv:2311.06267*.

Takabatake, T., Asai, K., Kakuta, H., & Hasegawa, N. (2025). Optimizing evacuation paths using agent-based evacuation simulations and reinforcement learning. *International Journal of*

Disaster Risk Reduction, *117*, 105173.

Thakur, S., Ranjitkar, P., & Rashidi, S. (2022a). Investigating evacuation behaviour under an imminent threat of volcanic eruption using a logistic regression-based approach. *Safety science*, *149*, 105688.

Thakur, S., Ranjitkar, P., & Rashidi, S. (2022b). Modelling evacuation decisions under a threat of volcanic eruption in auckland. *Transportation research part D: transport and environment*, *109*, 103374.

Wahlqvist, J., Ronchi, E., Gwynne, S. M., Kinateder, M., Rein, G., Mitchell, H., ... Kuligowski, E. (2021). The simulation of wildland-urban interface fire evacuation: The wui-nity platform. *Safety science*, *136*, 105145.

Wang, Q., & Nie, X. (2023). A location-inventory-routing model for distributing emergency supplies. *Transportation Research Part E: Logistics and Transportation Review*, *175*, 103156.

Wang, Y., Kieu, M., & Ranjitkar, P. (2024). Enabling continuous operation of shared autonomous vehicles with dynamic wireless charging. *IEEE Transactions on Intelligent Transportation Systems*, 1-10. doi: 10.1109/TITS.2024.3416412

Wong, S. D., Broader, J. C., Walker, J. L., & Shaheen, S. A. (2023). Understanding california wildfire evacuee behavior and joint choice making. *Transportation*, *50*(4), 1165–1211.

Yang, W., Yan, X., Hu, D., Jiang, X., Song, J., Lang, R., ... Lian, J. (2025). A novel emergency evacuation route optimization model in flood disasters using hydrodynamic model and intelligent algorithm. *Safety Science*, *182*, 106709.

Yang, X., Ban, X. J., & Mitchell, J. (2018). Modeling multimodal transportation network emergency evacuation considering evacuees' cooperative behavior. *Transportation research part A: policy and practice*, *114*, 380–397.

Yin, Y., Xu, X., Wang, D., Yu, Y., & Cheng, T. (2024). Two-stage recoverable robust optimization for an integrated location–allocation and evacuation planning problem. *Transportation research part B: methodological*, *182*, 102906.

Zhang, Y., Carballo, A., Yang, H., & Takeda, K. (2023). Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, *196*, 146–177.

Zhang, Z., Fei, Y., & Fu, D. (2024). A deep reinforcement learning traffic control model for pedestrian and vehicle evacuation in the parking lot. *Physica A: Statistical Mechanics and its Applications*, *646*, 129876.