Operations Research Forum (2026) 7:58
https://doi.org/10.1007/s43069-026-00642-1

REVIEW

From Models to Mobility Systems: A Survey of Multimodal Dynamic Traffic Assignment

Elisabeth S. Fokker[1,2] · **Robert D. van der Mei**[1,2] · **Elenna R. Dugundji**[3]

Received: 13 January 2026 / Accepted: 20 April 2026

Abstract
Multimodal dynamic traffic assignment models are widely used to estimate traffic flow patterns and support decision-making in transportation supply, infrastructure planning, and traffic management. As transportation systems shift toward multimodal, service-oriented, and data-driven approaches, existing models face growing challenges in realism and scalability. This survey reviews recent advances in multimodal dynamic traffic assignment, with a focus on modeling heterogeneous traffic, incorporating multiple transportation modes, and capturing interactions between different vehicle types. The evolution of dynamic traffic assignment from single- to multimodal frameworks is reviewed, alongside recent modeling and data-driven approaches, including deep reinforcement learning for adaptive, agent-based routing. Key research gaps are identified, and a future research agenda is outlined. Priority directions include improving scalability for large-scale networks, incorporating time-varying and data-driven demand representations, and integrating real-time traffic information to bridge the simulation-to-reality gap.

Keywords Multimodal dynamic traffic assignment · Traffic flow simulation · Multi-class and mixed traffic · Dynamic network loading · Data-driven traffic modeling · Deep reinforcement learning · Markov routing games

✉ Elisabeth S. Fokker
elisabeth.fokker@cwi.nl

Robert D. van der Mei
mei@cwi.nl

Elenna R. Dugundji
elenna_d@mit.edu

1 Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

2 Vrije Universiteit, Amsterdam, The Netherlands

3 Massachusetts Institute of Technology, Cambridge, MA, USA

Published online: 08 May 2026

1 Introduction

Multimodal dynamic traffic assignment (MM-DTA) models play a central role in transportation planning and infrastructure decision-making [1–3], and the development of advanced traveler information systems [4, 5]. Unlike static traffic assignment (STA) models, DTA models estimate traffic flow patterns and travel times in a time-dependent manner, capturing the temporal variability of congestion, driver route choices, and network performance. An MM-DTA model extends this framework by considering interactions and distinctions among at least two different transportation modes. This is achieved by incorporating mode-specific attributes such as speed, size, and travel costs [4]. Notable MM-DTA applications include modeling car–truck interactions [6], integrating ride-hailing (RH) systems with parking dynamics [7], and analyzing mixed traffic with autonomous (AV) and human-driven vehicles (HDV) [8].

1.1 Framework of Multimodal Dynamic Traffic Assignment

The general framework of a traditional MM-DTA model is illustrated in Fig. 1. The model combines a multimodal traffic network representation with time-varying travel demand and consists of two core components: (1) a mode and route choice model, often based on discrete choice formulations, which governs the assignment process, and (2) a dynamic network loading (DNL) model, which captures the temporal evolution and propagation of traffic along network links given the assigned modes and routes. The transportation network is represented by links or cells corresponding to road segments and nodes representing intersections. Each link is characterized by attributes such as capacity and the transportation modes it supports. DNL models capture the spatio-temporal interaction between traffic flow and traffic density on each network link [9]. These interactions are reflected in link performance functions that describe how travel times or generalized travel costs evolve as traffic volumes and densities change over time. The resulting performance measures are used to compute time-dependent network flows, with mode and route choices iteratively updated until an equilibrium state is reached. DNL can be implemented through analytical, semi-analytical, or simulation-based approaches, including the cell transmission model (CTM), the link transmission model (LTM), and queuing-based formulations such as the point-queue model (PQM) and the spatial-queue model (SQM). Equilibrium problems are commonly formulated using variational inequalities (VI), mathematical programming (MP), fixed-point problems (FPP), or approximated through agent-based simulation. Well-established equilibrium concepts include user equilibrium (UE) and system optimum (SO), originally introduced by [10].

1.2 Research Gap and Research Directions

The above framework builds on the classical four-step procedure for traffic flow prediction introduced in the late 1950s [11], which has long provided the basis for DTA models. Although widely applied, this framework has been extended to better capture growing system complexity, yet challenges remain in representing traffic

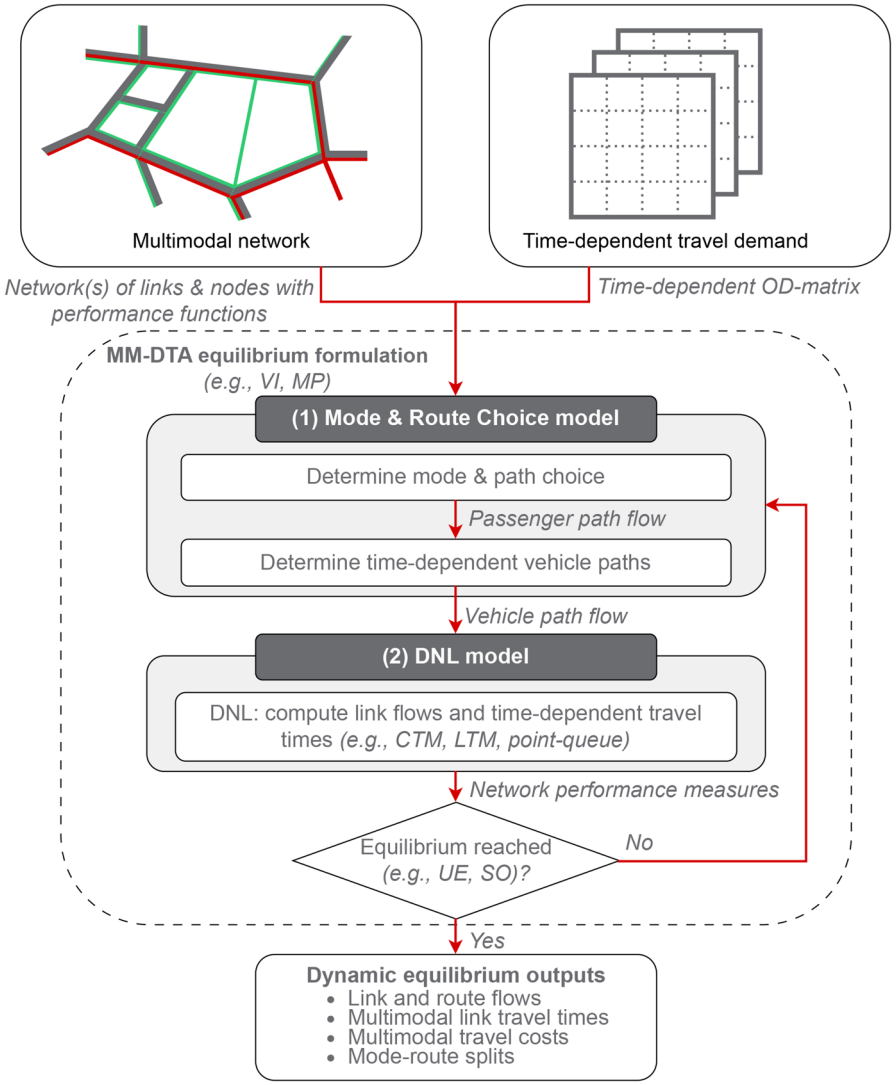

Fig. 1 General framework of the MM-DTA model

heterogeneity, network geometry, capacity constraints, and operational features such as intersection control, spillback, and queue formation [12]. In addition, many DTA studies focus on single-mode and small-scale networks, with limited use of real-world or real-time data. Recent advances in communication and sensing technologies create opportunities for MM-DTA models to evolve in two complementary directions: supporting MM and service-oriented transport systems, and strengthening the link between simulation outputs and real-world conditions. The adoption of Mobility as a Service and shared mobility increases the need for MM-DTA models that estimate MM travel times and supporting traveler decision-making [13, 14]. At the same time, data-

driven approaches offer effective means to better capture demand variability, network disturbances, and intermodal interactions, thereby improving predictive performance [15]. Against this background, this review surveys recent advances in MM-DTA, highlights emerging applications, and outlines future research directions, with particular emphasis on data-driven and deep reinforcement learning (DRL)–based approaches.

1.3 Literature Reviews on Multimodal Traffic Assignment

A well-known literature review on DTA is provided by [16], which primarily focuses on single-modal settings. However, relatively few studies explicitly address the characteristics of multimodal or heterogeneous traffic; these are summarized in Table 1. Early work by [17] synthesizes static macroscopic and dynamic microscopic models for heterogeneous traffic, but does not consider DTA within a network equilibrium framework. Subsequent reviews tend to focus on specific components of MM-DTA, most notably the DNL module, rather than the full assignment process [18, 19]. Broader methodological overviews are provided by [20] and [21]. The most recent review by [22] focuses on metaheuristic optimization for multimodal urban transport systems, with limited attention to DTA. In contrast, this review covers methods for a complete MM-DTA framework, jointly addressing route choice, DNL, and multimodal interactions. It highlights recent advances enabled by communication and sensing technologies and emphasizes real-world applicability. To the best of our knowledge, it is the first review to explicitly examine DRL within an MM-DTA context.

1.4 Structure

This paper is organized as follows. Section 2 describes the search strategy and inclusion criteria. Section 3 reviews the evolution of MM-DTA models, while Sect. 4 examines how multiple transportation modes are incorporated in MM-DTA frameworks. Section 5 explores recent DRL-based approaches and outlines a research agenda for real-world MM applications. Finally, Sect. 6 concludes the paper. A complete list of acronyms is provided in Appendix 1, while the list of notations is given in Appendix 2.

2 Methodology

The literature search was conducted in June 2023 and updated in October and December 2025, and February 2026. The search strategy combined keywords related to the main research field, traffic heterogeneity, and model characteristics. Papers were selected if these keywords appeared in the title, abstract, or keywords, while studies from unrelated domains (e.g., chemical engineering, telecommunications, and medicine) were excluded.

Keywords for the **main field** combined *dynamic* with *traffic assignment* or *traffic flow*.

Keywords for **traffic heterogeneity**, combined using *or*, included: *multi-class*, *heterogeneous*, *mixed-mode*, *mixed traffic*, *multi-modal*, *multimodal*, *multi-type*, *multi-*

Table 1 Literature reviews on multimodal dynamic traffic assignment

Paper	Period	Scope	VInt	MM	Dyn.	Contribution
Khan and Maini [17]	1955–1996	Heterogeneous traffic	×	×	S, D	Empirical, method
Jeihani [20]	1974–2006	DTA software			D	Method
Bhavathrathan and Mallikarjuna [18]	1955–2010	Macroscopic heterogeneous traffic		×	S, D	Method, agenda
Verma [19]	1955–2014	Mixed traffic (developing)	×	×	S, D	Empirical, agenda
Wang et al. [21]	1952–2018	Sustainability in road-based MM-DTA			D	Method
Chau and Gkiotsalitis [22]	2002–2025	Metaheuristic MM optimization		×	S, D	Method, agenda
This paper	1952–2025	Recent advances in mixed-traffic DTA	×	×	D	Method, real-world, agenda

Legend: VInt = vehicle interactions; MM = multimodal traffic; D = dynamic; S = static modeling

category, and *passenger-car equivalent*. In Sects. 3 and 5, selected single-modal studies were also included to reflect the historical context and the limited availability of explicitly multimodal DRL-based traffic assignment models.

Keywords for **model details**, used in Sects. 3 and 4, included: *cell transmission model*, *link transmission model*, *queue*, *user equilibrium*, *system optimum*, *social optimum*, *Nash equilibrium*, *game theory*, and *deep reinforcement learning*. Section 5 focused exclusively on DRL-related keywords.

Table 2 summarizes the search strategy per section, including time periods, keyword combinations, and the number of papers identified via Scopus, Google Scholar, and backward and forward snowballing. Backward snowballing was used to identify foundational studies, while forward snowballing captured more recent work. After removing duplicates and applying the inclusion criteria (Sect. 2.1), 44 publications were retained. In addition, methodological references and existing literature reviews (Table 1) were consulted to provide further context.

2.1 Inclusion Criteria

This review focuses on studies of road traffic and excludes work solely addressing air or rail transport. In addition, given the emphasis on DTA, papers related to supply chain modeling, such as the planning and optimization of goods delivery processes and the first- or last-mile problem, are excluded. Section 3 covers seminal papers from 1952 to 2002 that lay the groundwork for MM-DTA. It includes both single-modal and multimodal STA and DTA models. Section 4 reviews studies that consider traffic scenarios with at least two distinct transportation modes. AVs and HDVs are treated as separate vehicle classes when explicitly modeled, as behavioral differences (e.g., reaction times) directly affect DNL. Although some studies distinguish electric and gasoline vehicles, their DNL formulations are typically identical; differences arise only in generalized costs, range constraints, or policy assumptions. As driving behavior is the same, both are treated as a single car mode. This section considers dynamic models only. In Sect. 5, tabular reinforcement learning is excluded, as it is better suited to small state and action spaces than real-world settings. We include only DRL methods where the agent makes en-route traffic decisions, excluding studies focused on traffic signal control.

3 Definitions and Evolution

This section introduces key definitions and reviews the evolution of single- and multimodal STA and DTA. Table 3 summarizes model characteristics, emphasizing dynamic and multimodal aspects such as detail levels, equilibrium concepts, and flow representations. Further discussion is provided by [16].

3.1 Evolution of Multimodal Traffic Assignment

Figure 2 illustrates the evolution of MM-DTA models across three phases: (1) offline STA, (2) offline DTA, and (3) real-time DTA. As real-time DTA models are mostly

Table 2 Search strategy per section

Sec.	Topic	Period	Keywords included			Number of papers included			
			Main field	Traffic heterogeneity	Model details	Scopus	Google Scholar	Backward snowballing	Forward snowballing
3	History of MM-DTA	until 2010	All	Both all and none	All	6	7	5	0
4	Conventional MM-DTA	2010–2025	All	All	All	5	7	3	3
5	DRL	2010–2025	All	Both all and none	Only DRL	2	4	1	1

Table 3 Overview of common classification dimensions in traffic assignment research

Category	Explanation	Subcat.	Explanation
Level of detail	Level of abstraction in space–time traffic behavior	Microscopic	Individual vehicle behavior and interactions
		Mesoscopic	Group-based modeling without individual tracking
		Macroscopic	Traffic represented as aggregated flow
Level of dynamics	Extent of time-varying properties of traffic flow	Static	Time-invariant traffic properties
		Semi-dyn	Time-dependent flows with simplified congestion dynamics
		Dynamic	Time-varying traffic demand and flows
Modes	Number of modes included	Single-mode	Focuses exclusively on one mode (e.g., car)
		Multi-mode	Includes multiple modes (e.g., car, public transit (PT))
Lane basis	Manner of ordering traffic on a network	Lane	Differentiates traffic by lanes
		Non-lane	Does not differentiate traffic by lanes
DNL model	Framework capturing time-dependent traffic flow	CTM	Kinematic wave model using discretized cells
		LTM	Kinematic wave model using cumulative link counts
		PQM	Represents congestion as point queues
		SQM	Represents queues distributed along link length
Equilibrium concept	Stable state in which travel decisions satisfy the adopted equilibrium concept	UE	No agent can reduce its own travel cost by unilaterally changing route, given the decisions of others
		SUE	Stochastic UE with probabilistic route choice
		SO	Minimizes total system travel cost through coordinated routing decisions
Equilibrium representa- tion	Representation of equilibrium linking flows, travel costs, and route choices in DTA	VI	Variational inequality; equilibrium as inequality constraints linking flows and costs
		MP	Mathematical programming; equilibrium as a constrained optimization problem, e.g., Linear Programming (LP) and Nonlinear Programming (NLP)
		FPP	Fixed point problem; equilibrium as a fixed-point condition requiring consistency between flows and costs

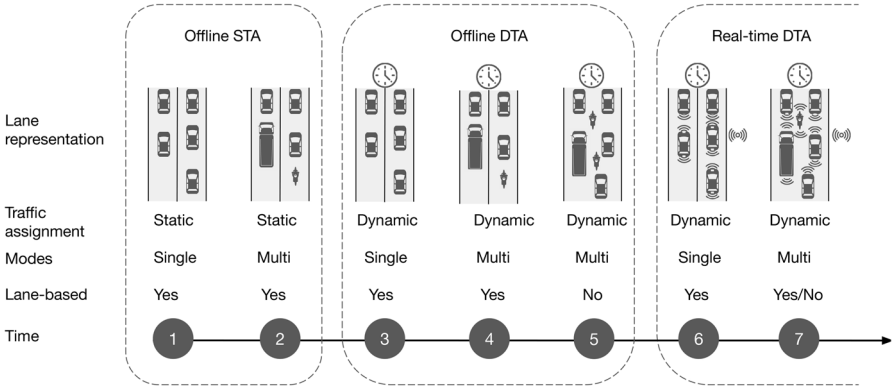

Fig. 2 Evolution of the multimodal DTA model

single-modal, this section focuses on the first two phases (Table 4), followed by a discussion of real-time MM-DTA developments in Sects. 4 and 5.

3.1.1 Static Traffic Assignment Models

STA originates from Wardrop's first principle [10], which formulates route choice as UE, where no traveler can reduce travel time by unilaterally changing routes. Wardrop's second principle, later termed SO, shifts the focus to system-wide performance by minimizing total travel time. Despite methodological advances, early traffic assignment models remained largely single-mode and car-centric, with modes treated independently due to simplifying assumptions or methodological constraints [23]. Mixed-mode formulations began to emerge in the 1960s, incorporating inter-modal vehicle interactions [24, 27], differences in speed, comfort, and pricing [25, 26], and modal split functions based on travel impedances [27]. Advances in communication and information technologies further enabled semi-dynamic formulations, including simulation-based models [29–31] and equilibrium models with elastic demand [32]. These approaches are foundational to DTA but rely largely on static assumptions. Volume–delay functions limit their ability to capture temporal dynamics, congestion spillback, and short-term demand variability. Strict FIFO assumptions and the lack of lane-level detail further constrain real-time representation.

3.1.2 Dynamic Traffic Assignment Models

The dynamic, time-varying nature of traffic makes DTA more complex than STA. Early work by [33, 34] introduced the M–N model, a nonlinear and nonconvex MP formulation for deterministic, fixed-demand, single-destination, single-commodity, and single-modal settings. They further proposed a piecewise linear formulation with a characteristic staircase structure, enabling efficient solution methods (e.g., matrix decomposition) and global optimization via a one-pass simplex algorithm without requiring branch-and-bound.

Table 4 Overview of STA and DTA models in the past

Level of dynamics	Paper	MM	Scope	Modes
Static	[10]		UE and SO	Car
	[23]		Capacity-constrained flow simulation	Car, PT
	[24]	×	5-step MM-STA	Car, PT
	[25]	×	DODOTRANS: incremental supply–demand equilibrium	Air and rail
	[26]	×	Heuristic UE	Bus and car
	[27]	×	UE with inter-modal interactions	Car, PT
	[28]	×	UE with fixed, decision-based, and route-based modal splits	Car, PT
	[29, 30]	×	SATURN: congestion simulation	Car, PT
	[31]	×	SATCHMO: SATURN extension for modal change	General MM
	[32]	×	MM equilibrium with elastic demand	General MM
Dynamic	[33, 34]		M–N model: nonlinear, nonconvex MP	Car
	[35]	×	Microscopic simulation-based DTA model	Car, minibus, bus, tram
	[36]	×	PACSIM: behavioral mesoscopic mode	General MM
	[37]		CTM model	Car
	[38]	×	Stochastic DTA for asymmetric assignment	Car, truck
	[39]	×	Cell-based DTA (Riemann-based)	Car, truck

The CTM by [37] likewise employs piecewise linear relationships and builds on hydrodynamic traffic flow theory [40, 41]. By representing traffic as a fluid across discrete cells, it captures essential phenomena such as stop-and-go waves and congestion-induced shockwaves. CTM has inspired a range of extensions, including an LP formulation for system-optimal DTA by [42], later extended by [6] to incorporate interactions between cars and trucks. Early MM-DTA developments include the simulation framework by [35], which captures heterogeneous vehicle behavior by combining physical road characteristics, vehicle properties, and stochastic driver behavior. Subsequent advances include the mesoscopic PACSIM model [36], the stochastic equilibrium analysis of [38], and the cell-based model by [39], which accommodates multiple vehicle types and captures density–flow dynamics.

3.1.3 Real-Time Dynamic Traffic Assignment

Real-time transportation modeling has traditionally focused on single-modal traffic [43–45]. For example, [44] integrated STA and DTA to evaluate network-wide conditions using real-time data, historical databases, and control center inputs. Similarly, MITSIM [45] simulates detailed traffic dynamics, incorporating car-following, lane-changing, signal control, and probabilistic route choice based on real-time information. [46] further combined DTA with adaptive signal control to generate both static and dynamic strategies. These approaches improve accuracy by continuously updating traffic conditions and control measures. However, route choices are typically determined prior to departure, with limited adaptation en route. While local adjustments (e.g., lane changes) are captured, this does not fully reflect travelers' real-time decision-making. Fully online models addressing this limitation are discussed in Sect. 5.

4 Multimodal Features in Dynamic Traffic Assignment Frameworks

This section reviews recent MM-DTA studies summarized in Table 5, focusing on how mode types are distinguished and modeled in multimodal, service-oriented traffic contexts. We examine mode-specific characteristics, choice models, test networks, mode-shift mechanisms, and cross-modal interactions, and conclude by synthesizing key design principles for practical MM-DTA applications. The reviewed models are primarily designed for strategic and policy-oriented analysis and therefore do not incorporate real-time traffic data. For service-oriented applications, closer integration of real-time information remains a promising direction to bridge the simulation-to-reality gap identified in the introduction.

4.1 Mode-Specific Characteristics

Each vehicle type $m \in \mathcal{M}$ is characterized by a set of mode-specific attributes. These include traffic flow–related parameters such as the Passenger car equivalent (PCE) factor, free-flow speed, reaction time, and capacity constraints. Additionally, attributes such as mode-specific occupancy rate and a discomfort factor may be incorporated in the objective or mode choice function. In the following, we outline how these differences can be integrated into the model.

4.1.1 Passenger Car Equivalent

To capture differences in space usage and traffic impact across vehicle types, PCE factors are commonly employed. For example, [6] introduce a truck-specific PCE factor in an extension of the CTM to distinguish between cars ($m = 1$) and trucks ($m = 2$). In their formulation, truck flows are expressed in PCE units using a factor $E_2 \geq 1$. Let $q_{m,i,t}$ denote the flow of vehicle class m in cell i at time t. Then, a flow of ten cars is represented as $q_{1,i,t} = 10$, whereas a flow of ten trucks is represented as $q_{2,i,t} = 10E_2$. In their simulations, $E_2 = 2$, implying that one truck is equivalent to two passenger cars.

Table 5 Overview of MM-DTA literature

Paper	Modes	Choice	DNL	Equil. repres	Level	Equil. concept	Case study	Scale	Mode shifts	Mode inter
[2]	PT, car-pool, car	NL	PQM	VI	Macro	UE	Freeway, transit line	S	×	×
[47]	Car, bus, metro, bike	CL/NL	SIM	–	Meso	SUE	Beijing (Chaoyang)	M	×	×
[6]	Car, truck	–	CTM	LP	Meso	SO	Nguyen-Dupuis	S		×
[48]	Car, bus	–	CTM	LP	Meso	SO	Nguyen-Dupuis	S		×
[49]	Bus, rail, walk, bike	HP	SIM	FPP	Meso	UE	Chicago	L	×	
[50]	Car, truck	–	SQM	VI	Meso	UE	Nguyen-Dupuis, Sioux Falls	S		×
[8]	HDV, AV	–	CTM	FPP	Meso	UE	Austin	M		×
[51]	Car, motor, bus	MNL	LQM	–	Meso	UE	Taoyuan	S		×
[52]	Car, truck	–	CTM	–	Macro	–	Cell-based link	S		×
[53]	Car, motor	–	CTM	–	Meso	–	Road segment	S	×	×
[7]	PT, car, RH, P&R	NL	CTM	VI	Meso	UE	Pittsburgh, Fresno	M	×	
[54]	Car, train, bus, tram, metro	–	SQM	–	Meso	–	Switzerland (road, PT)	L	×	×
[55]	HDV, AV	CNL	LTM	VI	Macro	UE	Nguyen-Dupuis, Sioux Falls, Anaheim	S		×
[56]	HDV, AV	–	CTM	NLP	Macro	UE, SO	Braess; Sioux Falls	S		
[57]	HDV, AV	–	CTM	NLP	Macro	SO	Sioux Falls	S		
[58]	RH, walk, car, PT	MNL	SQM	FPP	Meso	SUE	Jiading (Shanghai)	M	×	
[59]	HDV, AV	MNL	SIM	FPP	Meso	UE, SO	Sioux Falls	S		×
[60]	Bus, rail, bike, walk	MNL	SIM	VI	Meso	UE	Boston, Jiaxing	M	×	

Choice: NL: Nested logit, CL: C-logit, MNL: Multinomial logit, CNL: Cross-nested logit, HP: Hyperpath
Scale: network definition with S: < 100 nodes, M: 100–1k nodes, L: > 1k nodes

4.1.2 Free-Flow Speed

The free-flow speed v_m^F denotes the travel speed of vehicle class m under uncongested conditions and is determined by class-specific physical and operational characteristics, such as vehicle size, acceleration capabilities, and speed limits.

For instance, [6] assume that trucks have lower or equal free-flow speeds than passenger cars, and define a scaling factor $\beta_m = v_m^F / v_{\text{ref}}^F$. Taking passenger cars as the reference class ($m_{\text{ref}} = 1$), with $v_1^F = 48$ km/h, and trucks as $m = 2$ with $v_2^F = 36$ km/h, this yields $\beta_2 = v_2^F / v_1^F = 0.75$. Similarly, [48] assume free-flow speeds of 60 km/h for cars and 30 km/h for buses. These values are context-dependent and vary across networks and modes. These models include class-specific free-flow speeds but do not enforce equal speeds under congestion, capturing interactions implicitly via flow propagation and capacity constraints. In contrast, [50] explicitly model a state of homogeneous speeds under congestion. They propose a link-based formulation in which each link is partitioned into a free-flow segment, where vehicles travel at class-specific speeds v_m^F, and a congested segment, where all vehicle classes move at a common reduced speed.

4.1.3 Capacity Constraints

Several MM-DTA studies explicitly model capacity constraints to represent competition among vehicle classes for limited infrastructure. Capacity is typically imposed at the link or cell level and shared across classes, while class-specific characteristics determine how strongly each class contributes to capacity consumption. In the CTM-based formulation of [6], differences in free-flow speed and PCE imply class-dependent capacity usage, with trucks consuming more capacity than passenger cars. Due to their lower free-flow speed, trucks are also assigned a lower cell capacity, i.e., $Q_{2,i,t} \leq Q_{1,i,t}$ [6].

Similarly, capacity constraints are often represented through spatial queuing mechanisms, in which congestion and spillback arise when aggregate demand exceeds downstream receiving capacity. For example, [50] model these effects explicitly via a spatial queuing formulation. In large-scale agent-based simulators such as GEMSim [54], similar behavior emerges through queue storage limits and link flow buffers, allowing congestion and spillback to arise naturally when infrastructure capacity is exceeded. When reaction time is explicitly incorporated, capacity constraints can additionally emerge endogenously through class-specific flow limits.

Multi-class formulations such as those proposed by [52] further show that the fundamental diagram, and thus effective capacity, depends on traffic composition. Differences in vehicle characteristics, such as free-flow speed, give rise to class-specific flow–density relationships, which jointly determine congestion dynamics in mixed traffic. To formalize the dependence on traffic composition, we introduce a space-allocation framework. Consider M vehicle classes, indexed such that $v_1^F > \cdots > v_M^F$, and let $\rho = (\rho_1, \ldots, \rho_M)$ denote the class-specific densities on a road section. We define a space-allocation vector $\alpha = (\alpha_1, \ldots, \alpha_M)$, where $\alpha_m \in [0, 1]$ represents the fraction of road space occupied by class m, with $\sum_{m=1}^{M} \alpha_m \leq 1$. This formulation captures competition among vehicle classes for limited road capacity, arising from

differences in vehicle size, driving behavior, and lane usage. The perceived density for class m is then defined as

$$\rho_m^p = \begin{cases} \rho_m \dfrac{\sum_{j=1}^{M} \alpha_j}{\alpha_m}, & \text{if } \alpha_m > 0, \\ 0, & \text{if } \alpha_m = 0, \end{cases} \tag{1}$$

which represents the effective density experienced by class m per unit of allocated space. In this way, capacity constraints are captured through the relative allocation of space across vehicle classes.

4.1.4 Reaction Times

Another driver of mode-specific capacity is the difference in reaction time across vehicle classes. For instance, [8] developed a shared-road CTM for AVs and HDVs, modeling them as distinct classes with different reaction times that affect achievable flow and capacity. Human drivers are assumed to have reaction times between 1 and 1.5 s [61], leading to the following flow–density relationship (see Fig. 3b):

$$q_{m,i,t} = \min\left(v_{m,i,t}\rho_{m,i,t},\ \left(\frac{1/\rho_{m,i,t} - \ell_m}{\Delta t_m}\rho_{m,i,t}\right)\right), \tag{2}$$

where, for vehicle class m at time t at cell or link i, $v_{m,i,t}$ denotes the actual speed, $\rho_{m,i,t}$ the density, ℓ_m the vehicle length, and Δt_m the reaction time. This formulation

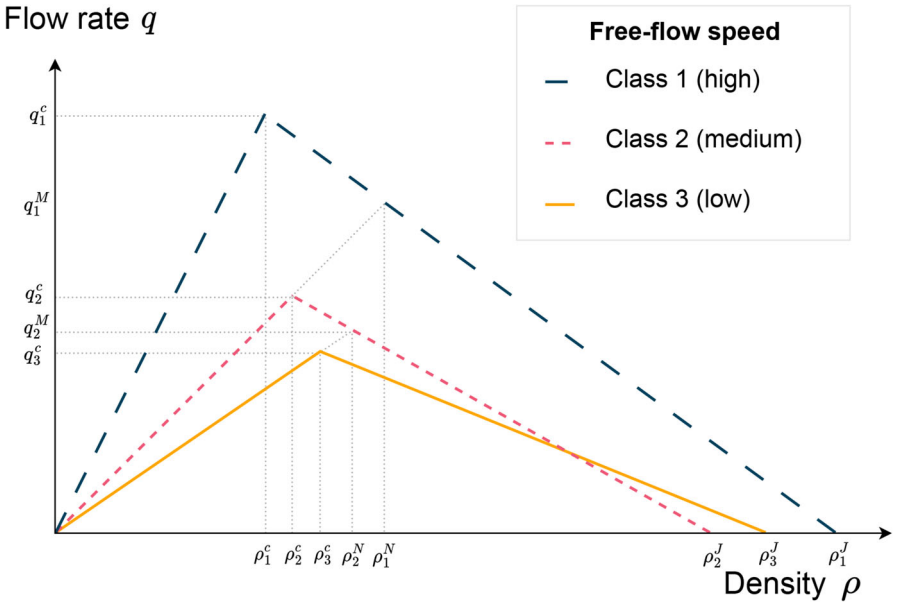

Fig. 3 Multi-class fundamental diagrams by mode characteristics

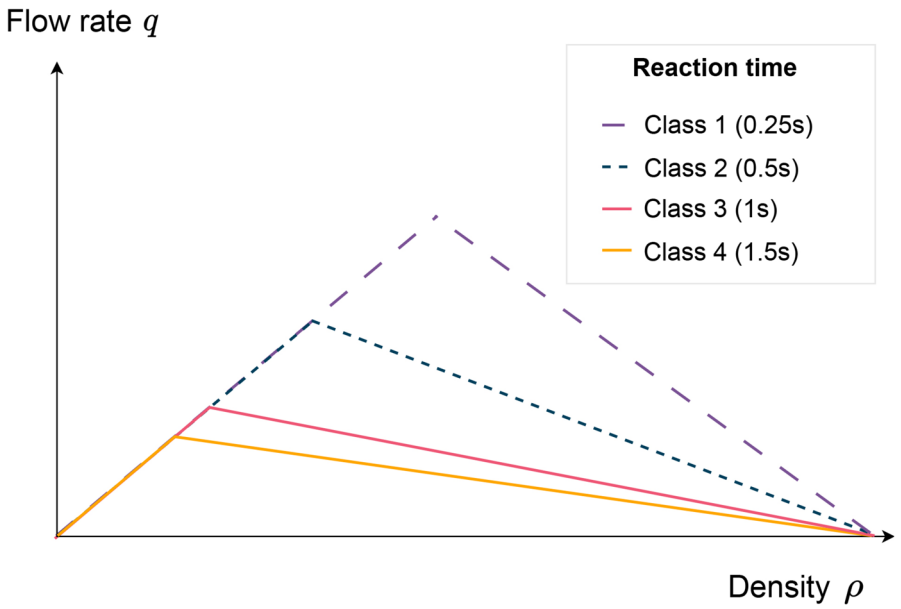

Fig. 3 continued

implies that reaction time acts as a class-specific constraint on admissible flow, directly affecting effective link capacity. Classes with shorter reaction times can sustain higher flows at the same density, delaying congestion onset. In mixed traffic, this results in heterogeneous congestion patterns, where AVs may continue to flow freely while HDVs experience reduced speeds.

4.1.5 Occupancy Rate

Vehicle occupancy can be incorporated into the objective function to prioritize higher-capacity modes. For example, [48] weight the system objective by occupancy, thereby assigning greater importance to vehicles carrying more passengers. Under this formulation, a bus carrying 40 passengers is weighted twenty times more than a car carrying two passengers.

4.1.6 Discomfort Factor

Differences across mode types can also be captured through discomfort penalties. For instance, [49] assign a higher weight to waiting time relative to in-vehicle time (two times as high) and include a fixed transfer penalty (for example, 5 min), thereby reflecting the additional inconvenience associated with waiting and transfers.

4.2 Path and Mode Choice Model

s In MM-DTA frameworks, distinguishing between transportation modes typically requires jointly modeling modal split and route choice. Discrete choice models are

widely used to represent travelers' decisions, incorporating attributes such as travel time, cost, convenience, and schedule constraints. For a detailed treatment of their theoretical foundations, see [62] and [63]. As shown in Table 5, some studies determine decisions deterministically via equilibrium principles (e.g., UE or SO), assigning travelers to minimum-cost paths with mode differences captured through class-specific assumptions rather than probabilistic choice [8, 50].

4.2.1 Nested-Logit Models

Nested logit models capture passenger mode choice by grouping alternatives into nests that reflect correlations arising from shared attributes. [2] determine the modal split of the commuter population based on differences in generalized travel time, whereas [7] extend this framework by additionally incorporating parking, park-and-ride (P&R) options, and travel costs.

Figure 4 illustrates a nested logit model. For each origin–destination (O–D) pair, travelers choose among modes such as PT, Driving, RH, and P&R. Each alternative is further decomposed into a set of lower-level options representing more detailed choices. Let $\mathcal{P}_m^{rs}$ denote the set of feasible paths for mode m between origin r and destination s, and let t denote the departure time. Then, for each $m \in \mathcal{M}$ and path $k \in \mathcal{P}_m^{rs}$, the generalized cost is given by

$$c_{m,k,t}^{rs} = \lambda_{\mathrm{TT}}\, w_{m,k,t}^{rs} + \max\left\{\lambda_{\mathrm{L}}\left(t + w_{m,k,t}^{rs} - t^*\right),\ \lambda_{\mathrm{E}}\left(t^* - t - w_{m,k,t}^{rs}\right)\right\} + \Phi_{m,k,t}^{rs}, \tag{3}$$

where $\lambda_{\mathrm{TT}} w_{m,k,t}^{rs}$ represents the travel time cost, with λ_{TT} denoting the unit cost of travel time and $w_{m,k,t}^{rs}$ the actual travel time along path k for a departure at time t from r to s. Intuitively, it captures how travelers value the time spent traveling along a given path. The second term captures schedule delay costs. It penalizes deviations from the preferred arrival time t^* (e.g., the standard work starting time), where arriving late incurs a penalty weighted by λ_{L}, and arriving early incurs a penalty weighted by λ_{E}. The last term, $\Phi_{m,k,t}^{rs}$, captures additional mode- and path-specific costs. The components contributing to the mode-specific aggregated factors to the travel time $w_{m,k,t}^{rs}$ and the additional costs $\Phi_{m,k,t}^{rs}$ for each mode m are detailed in Table 6.

4.2.2 C-Logit Model

While the nested logit model captures correlation through a hierarchical structure, it does not explicitly account for similarity based on shared network components.

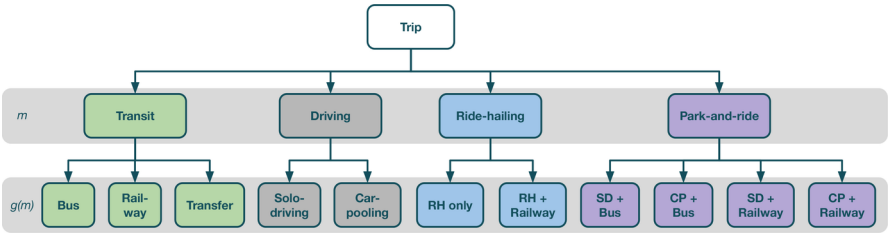

Fig. 4 Nested logit choice model, adapted from [7]

Table 6 Components of travel time $w^{rs}_{m,k,t}$ and additional costs $\Phi^{rs}_{m,k,t}$ (Eq. 3)

Mode	Components of total travel time $w^{rs}_{m,k,t}$	Additional costs $\Phi^{rs}_{m,k,t}$
Transfer	Possible PT waiting time, PT travel time, possible walking time	$\delta^{rs}_{k} + \sigma^{rs}_{k,t}$
Driving	Driving time, possible parking cruising time, possible walking time	$\frac{\delta_{\mathrm{P},i}}{n} + \kappa^{rs}_{k,t}(n) + \xi$
RH	Driving time, possible waiting time to pick up	$\delta^{rs}_{\mathrm{RH},k,t}$
P&R	Driving time, possible parking cruising time, PT travel time and possible walking times	$\frac{\delta_{\mathrm{P},i}}{n} + \kappa^{rs}_{k,t}(n) + \delta^{rs}_{k} + \sigma^{rs}_{k,t} + \xi$

From r to s on path k at time t: δ^{rs}_{k} denotes the PT fare; $\sigma^{rs}_{k,t}$ the crowding cost; $\kappa^{rs}_{k,t}(n)$ the carpooling impedance ($\kappa^{rs}_{k,t}(1) = 0$); $\delta_{\mathrm{P},i}$ the parking fee at location i; n the number of pooled travelers; $\delta^{rs}_{\mathrm{RH},k,t}$ the RH fare; and ξ car accessibility (0 if unavailable, else a large constant)

In route choice settings, this may lead to an overestimation of highly overlapping alternatives. The C-logit model addresses this limitation by introducing a commonality factor that penalizes route overlap. An example is provided by [47], who, similar to the aforementioned studies, use travel costs and travel times as attributes to determine mode and route choice, and additionally incorporate transit schedules.

4.2.3 Multinomial Logit Model

Moreover, [51] employ a bilevel choice–equilibrium model with a multinomial logit formulation to simulate the decision-making dynamics of travelers attending planned special events (e.g., sporting events, concerts, and conventions). The upper level models departure time window and mode choices, while the lower level consists of an MM-DTA model that captures traffic flow patterns and travel-related parameters.

For the mode choice between feeder buses and private modes, the authors define disutility functions that capture travel time, schedule delay, and travel costs. Specifically, for a trip from origin r to destination s, departing at time t and arriving in time window n', the disutility for the feeder bus mode is given by:

$$c^{rs}_{n',t} = \lambda_{0,t} + \lambda_{\mathrm{TT}} w^{rs}_{n',t} + \lambda_{\mathrm{C}} \left[\Theta_{n'}\left(\tau^{rs}_{n',t}\right) + \delta^{rs}_{\mathrm{FB}} \right], \quad \forall r, s, n', t, \tag{4}$$

where $w^{rs}_{n',t}$ denotes the average path travel time between r and s for travelers departing at time t and arriving in time window n', and $\tau^{rs}_{n',t} = t + w^{rs}_{n',t}$ is the corresponding arrival time. The function $\Theta_{n'}(\tau^{rs}_{n',t})$ represents the schedule delay penalty associated with arriving earlier or later than the desired arrival time corresponding to time window n'. Furthermore, $\delta^{rs}_{\mathrm{FB}}$ denotes the feeder bus fare. The parameter $\lambda_{0,t}$ is a departure-time-specific constant, while λ_{TT} and λ_{C} represent the marginal disutility of travel time and travel cost, respectively.

For private modes $m \in \mathcal{M}_{\text{private}}$, disutility is specified as a linear function of travel time, schedule delay, costs, and distance, and is given by:

$$c^{rs}_{m,n',t} = \lambda_{0,m,t} + \lambda_{\text{TT}} w^{rs}_{m,n',t} + \lambda_{\text{C}} \left[\Theta_{n'}\left(\tau^{rs}_{m,n',t}\right) + \delta^{rs}_{m} \right] + \lambda_{\text{D}} d^{rs}_{m}, \quad \forall r, s, m, n', t, \tag{5}$$

where $w^{rs}_{m,n',t}$ denotes the average path travel time for mode m, and $\tau^{rs}_{m,n',t} = t + w^{rs}_{m,n',t}$ is the corresponding arrival time. The function $\Theta_{n'}(\tau^{rs}_{m,n',t})$ represents the schedule delay penalty relative to the desired arrival time of window n'. Furthermore, δ^{rs}_{m} denotes the variable operating cost (e.g., fuel and parking costs), and d^{rs}_{m} denotes the travel distance. The parameter $\lambda_{0,m,t}$ is a mode- and departure-time-specific constant, while λ_{TT}, λ_{C}, and λ_{D} represent the marginal disutility of travel time, travel cost, and travel distance, respectively. A detailed specification on the schedule delay function and lower-level model is provided in [51].

4.2.4 Hyperpath-Based Transit Assignment

In a hyperpath formulation, travelers do not select a single fixed route, but instead adopt a travel strategy that specifies feasible boarding and transfer options based on expected costs and service frequencies. Such formulations are commonly used in MM-DTA models where PT is represented by service frequencies rather than explicit timetables. Following [49], let $\mathcal{A}_i$ denote the set of access and transfer links departing from node i, and $\mathcal{L}_i$ the set of transit lines available for boarding at node i. In contrast to the deterministic path cost $c^{rs}_{m,k,t}$ defined earlier, we introduce the expected generalized cost-to-go V_i from node i to the destination. This value represents the minimum expected generalized cost under optimal travel behavior and uncertainty in service arrivals, and satisfies the following Bellman-type recursion:

$$V_i = \min \left\{ \min_{a \in \mathcal{A}_i} \left(w_a + V_{j(a)} \right), \ \frac{\sum_{l \in \mathcal{L}_i} f_l \left(w_l + V_{j(l)} \right)}{\sum_{l' \in \mathcal{L}_i} f_{l'}} \right\}, \tag{6}$$

where w_a denotes the travel time on link a, and w_l the expected waiting time for transit line $l \in \mathcal{L}i$, typically determined by its service frequency f_l. The mapping $j(\cdot)$ gives the downstream node after an action: $j(a)$ is the node reached via link a, with expected remaining cost $V\,j(a)$, and $j(l)$ the node reached after boarding line l, with cost $V_{j(l)}$. The first term represents a deterministic access or transfer link, where the total cost equals travel time plus downstream cost. The second term captures a boarding strategy over attractive transit lines, where travelers board the first arriving service, yielding a frequency-weighted expected cost. Passenger flows are therefore distributed across lines proportional to their frequencies. The probability of boarding line $l \in \mathcal{L}_i$ at node i is given by

$$p_l = \frac{f_l}{\sum_{l' \in \mathcal{L}_i} f_{l'}}, \quad \forall l \in \mathcal{L}_i, \tag{7}$$

where f_l denotes the service frequency of line l, and the denominator represents the total frequency of all lines available at node i.

4.2.5 Combination-Based Mode–Route Choice

MM-DTA models commonly represent mode and route choice as a hierarchical decision process, separating intermodal choice across modes from intramodal route choice within each mode. [60] follow this approach by modeling the intermodal split with a logit function based on the minimum generalized cost of each mode, while within-mode route choice follows a deterministic shortest-path condition. For an OD pair (r, s) at departure time t, the intermodal split is given by

$$\frac{q_{m,t}^{rs}}{q_t^{rs}} = \frac{\exp\left(-\theta \mu_{m,t}^{rs}\right)}{\sum_{m' \in \mathcal{M}} \exp\left(-\theta \mu_{m',t}^{rs}\right)}, \tag{8}$$

where q_t^{rs} denotes the total departure flow between r and s at time t, $q_{m,t}^{rs}$ the portion assigned to mode $m \in \mathcal{M}$, $\mu_{m,t}^{rs}$ the minimum generalized cost over all feasible paths of mode m, and $\theta > 0$ a dispersion parameter reflecting sensitivity to cost differences. The exponential transformation converts generalized costs into positive choice weights, such that lower-cost alternatives receive a higher probability. Within each mode m, flows are assigned to minimum-cost paths according to the complementarity condition

$$q_{m,k,t}^{rs}\left(c_{m,k,t}^{rs} - \mu_{m,t}^{rs}\right) = 0, \quad \forall k \in \mathcal{P}_m^{rs}, \tag{9}$$

where $q_{m,k,t}^{rs}$ denotes the path flow and $c_{m,k,t}^{rs}$ the corresponding generalized cost, ensuring that only minimum-cost paths carry flow, while allowing multiple paths with equal minimum cost to be used simultaneously. This combination of probabilistic intermodal choice and deterministic intramodal routing is common in MM-DTA models. Similar principles also arise in CTM- and SQM-based formulations [8, 50], where travel times emerge from traffic propagation. Compared to large-scale agent-based simulations [54], the formulation of [60] provides a compact equilibrium representation in which stochasticity is confined to the intermodal decision layer.

4.3 Differences by Equilibrium Concept

Beyond parameter-based distinctions, recent studies differentiate vehicle classes through distinct routing principles. For instance, [59] model HDVs as UE seekers, while AVs follow SO routing based on marginal travel times. This approach introduces behavioral heterogeneity directly into the equilibrium concept, rather than confining class differences to link performance functions. Similarly, [56] formulate a dynamic SO framework in which automated vehicles are routed to support system-level objectives, whereas HDVs remain part of the mixed traffic flow without individual optimization behavior.

4.4 Networks

The table shows most models are evaluated on small- to medium-sized networks, with common test cases including the Nguyen–Dupuis network (13 nodes, 20 links) and the Sioux Falls network (24 nodes, 76 links). While these benchmarks are intuitive and convenient for small-scale studies, scalability considerations are often overlooked. An exception is [54], who address scalability in a real-world, large-scale setting by modeling the entire Swiss road network and PT schedules, comprising hundreds of thousands of links and nodes and approximately 5.2 million agents. Scalability is achieved through a combination of parallel processing, optimized memory access, and efficient algorithms (here, based on SQM) designed for massively parallel execution on graphics processing units. Co-evolutionary learning and spatio-temporal partitioning further enhance the model's ability to handle complex, large-scale scenarios. In other multimodal network formulations, PT often operates on dedicated links or lanes, although buses may also share infrastructure with other vehicle types [2, 7, 60]. For instance, [7] model a virtual bus network that uses the same road links as private vehicles. Similarly, [60] represent PT, walking, and bike-sharing within a unified multimodal network, where interactions occur through shared nodes and transfer links rather than fully separated infrastructures. Intermodal travel (e.g., bus–rail–bike-sharing) is enabled via transfer nodes, allowing transitions across modes while preserving mode-specific costs and constraints.

4.5 Mode Shifts

Understanding mode changes is essential in transportation studies, especially in metropolitan areas with multimodal trips. Ignoring transfers limits real-world applicability and overlooks interactions between mode choice and trip assignment. For example, [47] propose a dynamic combined trip assignment model that jointly determines mode and route choices. The model allows up to two transfers per trip and considers five travel modes: private car, P&R, a single bus line, bus-to-transit (bus or subway), and bicycle-to-transit (bus or subway). The multimodal network distinguishes between route-choice nodes, where travelers may change both mode and route, and non-route-choice nodes, which restrict movement to a single downstream link of the same mode. Mode–route decisions are therefore made only at route-choice nodes. Route choice is modeled using a nested logit formulation combined with a time-dependent shortest path algorithm that dynamically assigns trips as network conditions evolve. In a related study, [7] formulate a UE model encompassing multiple modes, including solo driving, carpooling, ride-hailing, bus transit, railway transit, and P&R. Mode choice is governed by a multi-layer nested logit structure (Fig. 4), enabling a detailed representation of multimodal travel decisions.

At a larger scale, [54] develop an agent-based framework comprising a multimodal supply network and an optional PT schedule, coupled with a demand model representing individual travelers. Each agent is characterized by a set of daily plans consisting of spatially distributed activities and connecting travel legs. Mode changes are permitted exclusively at activity locations. To enable transfers, the authors introduce dedicated

"merge" activities that impose a short delay, representing actions such as PT transfers or transitions from walking to PT.

4.6 Cross-Vehicle Interaction

The interaction column indicates whether different mode types influence each other's performance. This is important because slower or larger vehicles, such as trucks and buses, can disproportionately reduce car speeds, leading to asymmetric interactions.

4.6.1 Car-Truck Interaction

Cars and trucks can interact asymmetrically, as slower and larger trucks may disproportionately reduce car speeds and effective capacity. [6] explicitly model this car–truck interaction within a CTM-based SO-DTA framework by introducing a unitary car-capacity reduction per unit of truck flow. In their formulation, the presence of trucks reduces car capacity by a factor ν, set to $\nu = 0.75$ in the experiments. This interaction is implemented as a constraint in the LP formulation, allowing the asymmetric impact of trucks on car traffic to be captured directly.

4.6.2 Car-Bus Interaction

Similarly, [48] incorporate the impact of slow vehicles (e.g., buses) on overall traffic capacity. In their formulation, slow vehicles act as moving bottlenecks that reduce the effective capacity of the traffic stream. Assuming a single slow vehicle class, the effective capacity $Q_{i,t}$ of road segment (or cell) i at time t is given by

$$Q_{i,t} = \frac{q_{U,i,t}}{1 - \exp\left(-q_{U,i,t}\left(1 - \frac{q_{U,i,t}}{Q_i}\right)\eta_{i,t}\,\tilde{\tau}_i\right)}, \tag{10}$$

where $q_{U,i,t}$ denotes the upstream flow entering segment i at time t, Q_i is the nominal capacity in the absence of slow vehicles, and $\eta_{i,t}$ is the proportion of slow vehicles in the traffic stream on segment i. The disturbance time $\tilde{\tau}_i$ represents the duration over which a slow vehicle affects following traffic, and is given by

$$\tilde{\tau}_i = \frac{\ell_{\text{slow}}\left(v_i^S + v_{\text{slow}}^F\right)}{v_i^S\, v_{\text{slow}}^F}, \tag{11}$$

where ℓ_{slow} is the vehicle length of the slow vehicle class, v_i^S denotes the backward shockwave speed on segment i, and v_{slow}^F is the free-flow speed of the slow vehicle.

4.6.3 Interaction in Traffic Queues

Another example of cross-class interaction is provided by [50]. While travel times are initially computed separately for each vehicle class, interactions are captured through

the outflow calculation. Specifically, the queue length is defined as the aggregate flow across all classes and is then used to determine class-specific outflow rates, ensuring that congestion effects reflect interactions among all traffic classes.

4.7 Applications for Multimodal and Service-Oriented Systems

Overall, incorporating multimodal characteristics, capacity mechanisms, and behavioral heterogeneity enhances both the realism and policy relevance of MM-DTA models. The reviewed approaches support the analysis and design of multimodal, service-oriented transport systems in several key ways:

1. **PCE**: PCE-based formulations capture differences between vehicle types, enabling realistic use of shared infrastructure.
2. **Free-flow speed and capacity heterogeneity**: Accounting for differences in free-flow speed and capacity across modes captures fundamental disparities in traffic dynamics and supports the evaluation of multimodal interactions.
3. **Capacity constraints and congestion formation**: Link- or cell-level capacity constraints allow MM-DTA models to represent congestion, spillback, and queue propagation under mixed traffic conditions, which is essential for stress-testing system performance and service-oriented strategies.
4. **Reaction time effects**: Class-specific reaction times (e.g., AVs versus HDVs) capture the impact of automation on admissible flow, congestion onset, and effective capacity, enabling scenario analysis of emerging technologies.
5. **Occupancy-based prioritization**: Occupancy-weighted objectives shift the focus from vehicles to persons, aligning traffic assignment with service-oriented goals such as maximizing person throughput and supporting high-capacity modes.
6. **Discomfort and transfer penalties**: Incorporating discomfort, waiting, and transfer penalties reflects perceived service quality and traveler experience, thereby improving behavioral realism in multimodal settings.
7. **Dynamic mode shifts**: Modeling mode shifts and combined mode–route choices captures multimodal trip chains and transfers, supporting the evaluation of integrated mobility services.
8. **Equilibrium concepts and behavioral heterogeneity**: Allowing different equilibrium principles across classes (e.g., UE versus SO routing) enables the representation of heterogeneous decision-making and coordinated routing strategies relevant to centrally guided or automated systems.
9. **Cross-vehicle interactions**: Explicitly modeling interactions between vehicle types (e.g., slower or larger vehicles reducing capacity) improves realism and supports multimodal analysis.

5 Deep Reinforcement Learning for Traffic Modeling

This section outlines the applications of DRL in traffic flow modeling and DTA. We highlight its potential to reduce the simulation-to-reality gap in conventional models, and investigate extensions to mixed-mode settings. As this research area is still emerg-

ing, most existing studies focus on single-mode traffic, and multimodal applications remain limited. As traditional DTA models rely on simplifying assumptions (e.g., perfect network knowledge and fixed route choices), their ability to represent realistic traffic behavior under dynamic or uncertain conditions remains limited. Therefore, recent studies embed DRL within Markov Routing Game (MRG) frameworks [64], enabling agents to learn adaptive, state-dependent routing policies through interaction with a simulated traffic environment. These approaches can learn adaptive policies directly from interaction with the environment, reducing the need for explicit travel-time functions or detailed prior knowledge of system dynamics. Although training can be computationally demanding, model deployment is typically efficient and scalable.

5.1 Deep Reinforcement Learning Applications

Table 7 provides an overview of DRL-based traffic flow models, focusing on how key components such as agents, states, actions, and rewards are defined across studies. In these models, DRL is used to capture adaptive, state-dependent decision-making in dynamic traffic environments. To facilitate interpretation of the table, we briefly introduce the core components of a DRL framework:

- **Agent** $i \in \mathcal{N}$: The decision-making entity. In traffic modeling, this may represent an individual driver, a group of drivers, or a centralized controller.
- **Environment**: The system with which the agent interacts, typically the road network and other vehicles, modeled via a traffic simulator or DNL model.
- **State** $s_t \in \mathcal{S}$: The information available at time t, describing the current situation (e.g., location, time, congestion).
- **Action** $a_t \in \mathcal{A}$: A decision taken in state s_t, such as selecting an outbound link, route, or departure-time adjustment.
- **Reward** $r_t \in \mathbb{R}$: A feedback signal indicating performance, often based on (negative) travel cost or a weighted combination of objectives.

As shown in Table 7, two main types of agents can be distinguished. First, a centralized agent, i.e., a traffic controller, which observes aggregate traffic conditions and makes system-level decisions. Second, decentralized agents in single- or multi-agent settings represent (groups of) drivers who make en-route decisions by selecting actions that maximize expected cumulative rewards via the Q-function. States are commonly defined by location, time, and congestion levels, while rewards often reflect travel time, though multi-objective formulations also exist (e.g., [65]). After each action, the environment—modeled through a DNL model or traffic simulator (e.g., SUMO)—updates and returns a reward. The resulting system behavior is typically associated with Nash equilibria (analogous to UE) or Stackelberg equilibria (analogous to SO). However, most studies are limited to relatively small-scale networks, as computational complexity constrains the number of agents. A notable exception is [66], who scale to a large urban network by aggregating travelers into representative agents, limiting actions to mode and departure-time choices, and outsourcing DNL to an external microscopic simulator rather than learning detailed routing policies.

Table 7 Overview of DRL-based traffic flow models

Paper	Alg.	Agent	State	Action	Reward	Traffic model	Equil.	Case study	Mode inter.
[67]	DQL	Central agent (single)	Demand (cur., rescheduled, remaining)	OD flow assignment	System travel time	Meso sim	SO	Sioux Falls (S)	
[68]	DQL	Car, truck (multi)	Count, speed, position, dest	Link choice	Travel time	Micro sim	Nash	Synthetic, Liverpool (S)	
[64]	Mean-field DQL	Driver (multi)	Time, node	Link choice	Travel cost	Micro sim	Nash	Columbia area (S)	×
[69]	DQL	AV (multi)	Local density, traffic state	Link choice	Network efficiency	Micro sim	Nash	Synthetic grid (S)	
[65]	DQN	Central agent (multi)	Region, dest., trip length	Link choice	Speed gain	Macro MFD & queue	SO	Synthetic multi-region (S)	×
[70]	DDQN	AV, HDV (multi)	Time, speed, gap, position	Accel., decel., lane change	Safety–efficiency	Micro sim	Nash	San Fran on-ramp (S)	×
[66]	DQL	Driver (multi)	Experience, network attributes	Mode, dep. time	Travel utility	Micro sim	Nash	Suzhou (L)	×
[71]	AC	Central agent (single)	Flow, density, speed, local network	OD path assignment	Total travel time	Meso CTM	SO	Maryland, stylized (M)	

Case study: network scale with small (S) < 100 nodes, medium (M) 100–1000, large (L) > 1000

5.2 Deep Reinforcement Learning Methodology

In the following, we examine the methodological foundations of the DRL-based traffic flow models summarized in Table 7.

5.3 Deep Q-Learning

Traditional RL models rely on tabular Q-learning, which is suitable for problems with small, discrete state and action spaces where value functions can be explicitly stored in lookup tables. An example in the transportation context is the study by [72], who applied standard tabular Q-learning in a single-agent setting on the small Sioux Falls network. In their formulation, the agent represents an individual driver, the state is defined by the current time, node, and congestion level, the action space corresponds to successor nodes, and the reward is given by the negative travel time. Despite its simplicity, transparency, and low computational cost, tabular Q-learning does not scale well to continuous or high-dimensional state and action spaces. As a result, its applicability to large-scale, real-world transportation networks is limited. Deep Q-learning (DQL) overcomes these limitations by replacing the explicit Q-table with a neural network that approximates the action–value function. This allows DQL to handle large and continuous state spaces by learning a mapping from states to Q-values for each feasible action, thereby enabling decision-making in more complex environments. The neural network takes the current agent state as input and outputs Q-values associated with the available actions. The objective is to learn a policy that maximizes the expected cumulative reward over time, commonly defined as a discounted sum:

$$R_t = \sum_{\tau=t}^{\infty} \gamma^{\tau-t} r_\tau, \tag{12}$$

where the discount factor $\gamma \in [0, 1)$ controls the relative importance of future rewards. Values of γ close to one place greater emphasis on long-term outcomes. Because route choice and driving decisions are inherently long-term, aiming to minimize total travel time rather than immediate travel time at each step, the discount factor γ is typically set close to one (e.g., 0.99 [68]). In some applications with a finite planning horizon, $\gamma = 1$ may also be adopted [64]. The Q-function represents the expected cumulative reward associated with taking action $a_t \in \mathcal{A}$ in state $s_t \in \mathcal{S}$ at time step t, and subsequently following an optimal policy. In tabular Q-learning, the Q-values are iteratively updated according to

$$Q^{new}(s_t, a_t) = Q^{old}(s_t, a_t) + \alpha \left(r_t + \gamma \max_{a \in \mathcal{A}} Q(s_{t+1}, a) - Q^{old}(s_t, a_t) \right), \tag{13}$$

where $\alpha \in (0, 1]$ denotes the learning rate, $r_t \in \mathbb{R}$ the immediate reward observed after taking action a_t in state s_t, and $\gamma \in [0, 1)$ the discount factor. Furthermore, $s_{t+1} \in \mathcal{S}$ denotes the subsequent state, and $\max_{a \in \mathcal{A}} Q(s_{t+1}, a)$ represents the estimated value of the best action available in the next state. The term in parentheses is referred to as the

temporal-difference error, capturing the discrepancy between the current estimate and a target value based on newly observed information. The Q-function is parameterized by a neural network with parameters $\theta \in \mathbb{R}^d$:

$$Q(s, a) \approx Q(s, a; \theta). \tag{14}$$

The parameters are learned through trial-and-error interactions with the environment by minimizing the expected squared temporal-difference error. Specifically, the loss function is defined as

$$\mathcal{L}(\theta) = \mathbb{E}\left[\left(r_t + \gamma \max_{a'} Q(s_{t+1}, a'; \theta^-) - Q(s_t, a_t; \theta)\right)^2\right], \tag{15}$$

where $\theta \in \mathbb{R}^d$ denotes the vector of neural network parameters, and $r_t \in \mathbb{R}$ represents the immediate reward observed after taking action $a_t \in \mathcal{A}$ in state $s_t \in \mathcal{S}$ at time step t, and $s_{t+1} \in \mathcal{S}$ denotes the subsequent state. The discount factor $\gamma \in [0, 1)$ determines the relative importance of future rewards. The expression $\max_{a' \in \mathcal{A}} Q(s_{t+1}, a'; \theta^-)$ represents the estimated value of the best action in the next state under the target network parameterization. The expectation operator $\mathbb{E}[\cdot]$ is taken over the stochastic transitions of the environment and, in practice, is approximated using sampled experience tuples. The resulting approximation converges toward the optimal action–value function Q^*, corresponding to the policy that maximizes the expected cumulative reward.

5.3.1 Dueling Deep Q-Network

An extension of DQL is the dueling deep Q-network (DDQN) architecture, which decomposes the Q-function into separate value and advantage components,

$$Q(s, a; \theta) = V(s; \theta_1) + A(s, a; \theta_2), \tag{16}$$

where $V(s; \theta_1)$ captures the value of being in state s, and $A(s, a; \theta_2)$ represents the relative advantage of selecting action a in that state. This architecture uses two jointly trained network streams to separate the value of a state from the effect of individual actions. The dueling architecture is particularly effective in settings where actions do not consistently exert a strong influence on the environment. In traffic modeling, this is often the case when (i) rewards are sparse, such that many state–action transitions provide limited learning signal, and (ii) the environment is highly stochastic, with randomness arising from vehicle arrivals, turning movements, and other exogenous factors that may hinder stable training [69]. A key drawback, however, is the increased computational burden associated with training multiple network components rather than a single Q-network.

5.3.2 Actor-Critic Network

Another class of DRL methods is the Actor–critic (AC) framework, which combines value-based and policy-based approaches. The architecture consists of two compo-

nents: an *actor*, representing the policy and producing a probability distribution over actions, and a *critic*, which estimates the expected return of a state or state–action pair. The actor is parameterized by a policy $\pi(a_t \mid s_t; \phi)$, while the critic approximates a value function, typically the state–value function $V(s_t; \theta_2)$ or the action–value function $Q(s_t, a_t; \theta_2)$. Both components are trained jointly, with the critic providing a learning signal that guides the actor's updates. The actor parameters are updated using a policy-gradient method of the form

$$\theta_{t+1} = \theta_t + \alpha \, \nabla_\theta \log \pi(a_t \mid s_t; \theta) \, \hat{A}(s_t, a_t), \tag{17}$$

where $\theta_t \in \mathbb{R}^d$ denotes the vector of actor parameters at time t, with d the number of parameters, $\alpha > 0$ the learning rate, and ∇_θ the gradient with respect to θ. The term $\pi(a_t \mid s_t; \theta)$ denotes the probability of selecting action $a_t \in \mathcal{A}$ in state $s_t \in \mathcal{S}$. The logarithm ensures that the update increases the probability of advantageous actions and decreases that of disadvantageous ones. The quantity $\hat{A}(s_t, a_t)$ is an estimate of the advantage function, measuring the relative benefit of taking action a_t in state s_t. A common choice is

$$\hat{A}(s_t, a_t) = Q(s_t, a_t) - V(s_t), \tag{18}$$

where $Q(s_t, a_t)$ denotes the expected cumulative reward of taking action a_t in state s_t and subsequently following the policy, and $V(s_t)$ the expected return of state s_t.

Alternative estimators based on temporal-difference errors are commonly used in practice. AC methods reduce variance compared to pure policy-gradient approaches by leveraging value estimates from the critic, often resulting in improved sample efficiency and faster convergence. Moreover, temporal-difference learning enables online updates without requiring episode termination. Although applications of AC methods to large-scale traffic assignment problems remain limited, the framework is well suited to traffic systems with high-dimensional state spaces and complex network interactions.

In traffic modeling, the state s_t typically represents current network conditions, such as link densities, speeds, or queue lengths. The action a_t corresponds to traffic control decisions, such as routing choices or the distribution of flows across alternative paths. The reward r_t reflects system-level performance, for example, as negative total travel time or congestion, thereby encouraging efficient network utilization. In the AC framework of [71], traffic routing is formulated as a sequential decision-making problem, where the actor learns a policy that determines routing actions based on observed network states, while the critic evaluates these routing actions by estimating their effect on future network performance and cumulative system cost. The reward signal is defined in terms of system efficiency, such that minimizing total travel time corresponds to maximizing cumulative reward. By iteratively updating both components, the method learns routing strategies that adapt to evolving traffic conditions and approximate system-optimal behavior.

Despite these advantages, AC methods present several challenges. Inaccuracies in the critic's value estimates can bias the policy gradient and degrade performance. In addition, training both actor and critic increases computational cost and algorithmic

complexity. Finally, their simultaneous learning may introduce instability, particularly when combined with function approximation in large-scale settings.

5.3.3 Mean-Field Deep Q-Learning

Mean-field games, originally introduced by [73] and [74], study the behavior of large populations of interacting agents and enable scalable analysis as the number of agents grows. In traffic modeling, this framework is particularly attractive for representing and coordinating interactions among large numbers of vehicles in a computationally tractable manner. In large-scale traffic systems, explicitly modeling interactions between all individual vehicles quickly becomes intractable. Mean-field approaches address this challenge by approximating the influence of other agents through an aggregate quantity, typically the average behavior of the surrounding population. As a result, each agent interacts not with every other agent individually, but with a representative "mean field," capturing the collective effect of the system. In contrast to single-agent settings, an agent $i \in \mathcal{N}$ belonging to class $c \in \mathcal{C}$ interacts both with the environment and with this aggregate population effect. Each agent collects experience tuples of the form $(o_i, a_i, o_i', r_i, \bar{a}_i)$, where o_i denotes the agent's private observation, a_i the chosen action, o_i' the subsequent observation after all agents act, r_i the received reward, and $\bar{a}_i$ the mean action of neighboring agents, representing the population influence. While the global state s is not directly observable, each agent receives a correlated private observation $o_i \in \mathcal{O}_i$, and the joint observation space is given by $O = O_1 \times \cdots \times O_N$, where N denotes the number of agents. In traffic applications, such observations typically encode information such as location and time [64]. Under the mean-field formulation, the Q-function is learned by minimizing the following loss:

$$\mathcal{L}(\theta_c) = \mathbb{E}_{(o_i, a_i, o_i', \bar{a}_i)}\left[\left(r_i + \gamma \max_{a_i'} \mathbb{E}_{\bar{a}_i' \sim \mu_{-i}^*}\left[Q^c(o_i', a_i', \bar{a}_i' \mid \theta_c^-)\right] - Q^c(o_i, a_i, \bar{a}_i \mid \theta_c)\right)^2\right], \quad (19)$$

where r_i denotes the immediate reward received by agent i after taking action a_i, $\gamma \in [0, 1)$ denotes the discount factor, μ_{-i}^* the equilibrium policy of all agents except agent i, $Q^c(o_i, a_i, \bar{a}_i \mid \theta_c)$ the class-specific action–value function, and $Q^c(\cdot \mid \theta_c^-)$ a target network used to stabilize training. The term $\bar{a}_i'$ denotes the mean action in the next state. The expectation $\mathbb{E}_{\bar{a}_i' \sim \mu_{-i}^*}[\cdot]$ captures the expected influence of the population under the equilibrium policy. The resulting optimal policy is given by

$$\mu_i^*(o_i) = \arg\max_{a_i} \mathbb{E}_{\bar{a}_i \sim \mu_{-i}^*}\left[Q^c(o_i, a_i, \bar{a}_i \mid \theta_c)\right], \quad \forall o_i \in \mathcal{O}_i, \quad (20)$$

where $\mu_i^*(o_i)$ denotes the optimal policy of agent i, mapping observations to actions. The operator $\arg\max_{a_i \in \mathcal{A}}$ selects the action that maximizes the expected action–value, where the expectation $\mathbb{E}_{\bar{a}_i \sim \mu_{-i}^*}[\cdot]$ is taken over the distribution of mean actions induced by the equilibrium policy μ_{-i}^* of all other agents, capturing the aggregate population effect. The function $Q^c(o_i, a_i, \bar{a}_i \mid \theta_c)$ represents the expected cumulative discounted reward given observation o_i, action a_i, and mean neighboring action $\bar{a}_i$. By optimizing

this expectation, each agent anticipates the average behavior of others and selects the best response accordingly. The condition $\forall o_i \in \mathcal{O}_i$ indicates that this holds for all observations.

Mean-field DQL reduces the complexity of multi-agent interactions by approximating individual interactions through the population mean. This formulation mitigates noise arising from exploratory behavior and leads to more stable learning dynamics. As a result, it improves scalability and convergence speed in systems with many agents [75]. However, solving mean-field games in large, continuous state spaces remains challenging. Many existing approaches rely on discretization, which suffers from the curse of dimensionality and limits scalability in high-resolution traffic networks [76].

5.4 Integrating Deep Reinforcement Learning with Multimodal Traffic Assignment

Figure 5 illustrates an example framework for integrating MM-DTA with DRL, informed by insights from the existing literature. The process begins with the collection of real-world data, including population characteristics, activity schedules, network topology, and traffic observations from heterogeneous sensors. These data are used to estimate mode-specific OD matrices and to extract vehicle and mode characteristics, such as free-flow speed, reaction time, and spatio-temporal modal splits. These inputs define a simulation-based traffic environment, implemented through a DNL model such as SUMO or CTM, which aims to capture the complexity of real-world traffic dynamics. This environment interacts with a multi-agent DRL model, typically based on DQL, in which vehicles of different modes are represented as agents. Agents make en-route decisions by selecting actions (e.g., outgoing links) that maximize expected

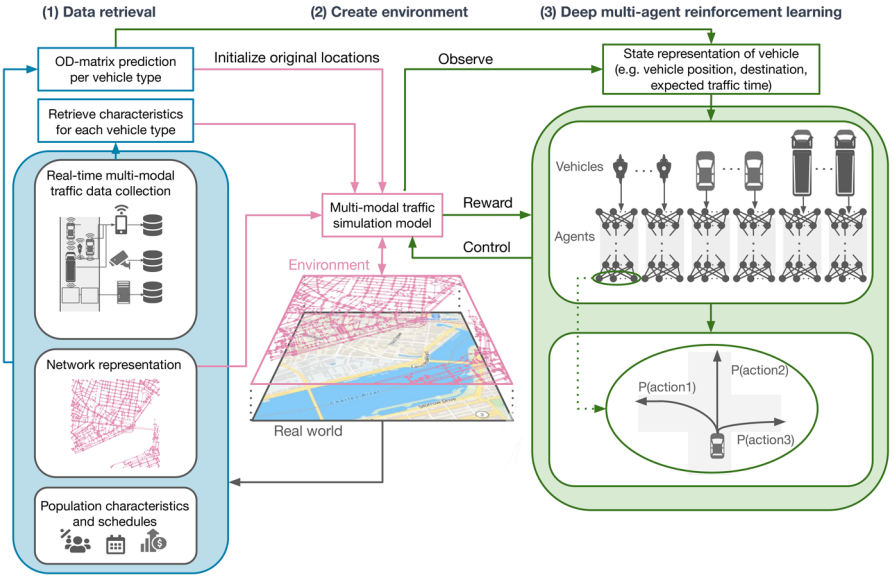

Fig. 5 Schematic illustration of a multi-agent DRL model

cumulative rewards, such as minimizing generalized travel costs, based on feedback from the simulator. Model parameters can be continuously updated as new real-world data become available. Building on concepts from the MM-DTA literature reviewed in the previous section, the proposed framework distinguishes agents by mode type and allows for mode-specific reward functions (see Eqs. 3 and 5). Modal split mechanisms can be introduced at designated activity nodes where mode changes are permitted, following approaches such as [54]. Interactions between different modes can be further captured using mean-field DQL, enabling the representation of multimodal trips in metropolitan networks.

5.5 Future Research

While recent DRL-based traffic models show promise, several directions remain open for improving their realism and practical relevance. Future work may focus on better use of real-world data, richer state and action representations, more realistic reward formulations, and improved integration with traffic simulation and DNL models.

5.5.1 Use of Real-World Data

One important direction is the increased use of real-world data to better reflect actual traffic conditions and traveler behavior. Most studies summarized in Table 7 rely on synthetic demand generated by DNL models or traffic simulators, and only a small number incorporate observed traffic data. Real-time traffic measurements, as well as demographic and socio-economic information, could help capture heterogeneity in travel behavior and improve external validity. Incorporating population-based datasets may allow agents to learn behavior that is more representative of real travelers rather than stylized drivers.

5.5.2 State Representation and Real-Time Information

Another avenue concerns the state representation of DRL agents. Many models assume full compliance with recommended routes. This is reasonable for connected or autonomous vehicles, but less so for human drivers. Incorporating real-time information (e.g., congestion, incidents, weather, events) could better capture deviations between assigned and realized behavior. Despite its practical relevance, such integration in MM-DTA and DRL-based models remains limited, indicating scope for further research.

5.5.3 Mode Choice and Mode Shifts

Despite the multimodal nature of urban transport systems, mode shifts are rarely modeled explicitly in current DRL studies. Extending action spaces to include mode-specific decisions, such as switching between private vehicles and public transport or park-and-ride options, could improve behavioral realism. Mode changes could be modeled at designated transfer locations, such as transit stops or parking facilities, following approaches used in conventional MM-DTA studies.

5.5.4 Reward Design

Most DRL-based traffic models use travel time as the primary reward signal. Future work could explore richer, mode-specific reward functions that also reflect factors such as operating costs, comfort, reliability, safety, or environmental impacts. Introducing rewards or penalties for transfers or mode changes may further support realistic multimodal decision-making.

5.5.5 Traffic simulation and Dynamic Network Loading realism

Model realism also depends on the quality of the underlying traffic simulator or DNL model. Future research may benefit from closer integration between DRL and advanced DNL formulations that capture spillback, queue interactions, and heterogeneous traffic more accurately. Relatively little work has addressed DNL models that jointly represent motorized and non-motorized traffic, which remains an open research area.

5.5.6 Interactions Across Mode Types

Finally, interactions between different transportation modes deserve further attention. While some DRL studies account for interactions among agents of the same mode, cross-mode interactions are rarely modeled explicitly. Extending multi-agent or mean-field DRL frameworks to distinguish between mode-specific agent groups could help capture the complex interactions observed in real-world multimodal traffic systems.

6 Conclusion

This survey reviews recent advances in MM-DTA with the objective of identifying how existing models can evolve toward (i) multimodal, service-oriented mobility systems and (ii) stronger integration with real-world conditions. In particular, the paper makes the following contributions:

- It provides an overview of MM-DTA models, covering conventional formulations, mixed-traffic representations, and emerging DRL approaches.
- It compares modeling choices across modes, including route and mode choice, traffic flow dynamics, interaction mechanisms, and equilibrium concepts.
- It reviews DRL-based traffic flow and routing models formulated as Markov routing games, focusing on agent design, state–action spaces, rewards, and scalability.
- It identifies challenges in multimodality, scalability, and the simulation-to-reality gap, and outlines directions for integrating data-driven learning into MM-DTA.

The literature shows that MM-DTA frameworks have significantly advanced in representing heterogeneous traffic through mode-specific characteristics, such as passenger car equivalents, free-flow speeds, reaction times, occupancy rates, discomfort penalties, and the explicit modeling of cross-mode interactions. Combined with improved route and mode choice models, these models enable detailed analysis of multimodal networks involving private vehicles, public transport, and shared mobility

services. However, most applications remain strategic, rely on offline demand, and are tested mainly on small to medium-sized networks.

Recent DRL-based approaches model traffic assignment as a stochastic routing game, where agents learn en-route decisions by interacting with a simulated environment. These methods relax assumptions of perfect information and fixed routing behavior, allowing travelers to respond dynamically to congestion and network conditions. While most DRL applications remain limited in scale and modal scope, extensions such as dueling networks, actor–critic methods, and mean-field formulations show potential for improving scalability, stability, and behavioral realism in large populations of interacting agents. Across both conventional and DRL-based MM-DTA models, scalability and realism remain key challenges. Large-scale applications often rely on aggregation, simplified action spaces, or external simulation engines to remain computationally tractable, while most models continue to depend on stylized demand and limited integration of real-time data. Addressing these limitations will require closer coupling between MM-DTA frameworks and empirical data sources, richer state representations incorporating real-time traffic information, and reward structures aligned with multimodal service quality and system-level objectives.

Overall, MM-DTA has evolved into a flexible framework for multimodal transport analysis, while DRL offers promising tools for adaptive traffic management. Future research may benefit from further exploring how the structural rigor of MM-DTA can be complemented by the adaptive capabilities of DRL to develop scalable, data-driven models that better support traffic management, infrastructure planning, and policy evaluation in increasingly complex mobility systems.

Appendix 1. List of Acronyms

Below we list the acronyms used in this paper, in alphabetical order.

AC	Actor-critic
AV	Autonomous vehicles
CTM	Cell transmission model
DDQN	Dueling deep Q-network
DNL	Dynamic network loading
DQL	Deep Q-learning
DRL	Deep reinforcement learning
FIFO	First-in-first-out
FPP	Fixed-point problems
HDV	Human-driven vehicles
LP	Linear programming
LTM	Link transmission model
MITSIM	Microscopic traffic simulator
MM	Multimodal
MM-DTA	Multimodal dynamic traffic assignment
MP	Mathematical programming
MRG	Markov routing game
NLP	Non-linear programming
OD	Origin-destination

PCE	Passenger car equivalent
PQM	Point-queue model
P&R	Park-and-ride
PT	Public transport
RH	Ride-hailing
SO	System optimum
SQM	Spatial-queue model
STA	Static traffic assignment
SUE	Stochastic user equilibrium
UE	User equilibrium
VI	Variational inequalities

Appendix 2. Notation

This appendix summarizes the notation used in the paper.

Symbol	Description
Multimodal dynamic traffic assignment	
Indices, sets, and mappings	
r, s	Origin and destination
i, j	Nodes, links, or cells
$t,\ n'$	Departure time, arrival window
m	Mode or vehicle class
k	Path index
$l,\ a$	Transit line, access/transfer link
$\mathcal{M},\ \mathcal{P}_m^{rs},\ \mathcal{L}_i,\ \mathcal{A}_i$	Sets of modes, paths, transit lines, and access links
$j(\cdot)$	Downstream node mapping
Demand, flow, and capacity	
$q_t^{rs},\ q_{m,t}^{rs},\ q_{m,k,t}^{rs},\ q_{m,i,t},\ q_{U,i,t}$	Flows (total, mode, path, link/cell, upstream)
$p_l,\ f_l$	Boarding probability, service frequency
$v_m^F,\ v_{\text{slow}}^F,\ v_{m,i,t},\ v_i^S$	Speed (class and slow free-flow, actual, shockwave)
β_m	Free-flow speed ratio
$\rho_{m,i,t},\ \rho_m^p,\ \rho$	Density (class, perceived, vector)
$\alpha_m,\ \alpha$	Space allocation (scalar, vector)
E_m	Passenger car equivalent
$Q_{m,i,t},\ Q_{i,t},\ Q_i$	Capacity (class, effective, nominal)
$\ell_m,\ \ell_{\text{slow}}$	Vehicle length (class-specific, slow vehicle)
$\Delta t_m,\ \tilde{\tau}_i$	Reaction and disturbance time
$\eta_{i,t},\ \nu$	Slow-vehicle share and capacity reduction
Travel cost components	
$w_{m,k,t}^{rs},\ w_{m,n',t}^{rs},\ w_a,\ w_l$	Travel and waiting time
$c_{m,k,t}^{rs},\ c_{m,n',t}^{rs}$	Generalized cost/disutility
$\mu_{m,t}^{rs},\ V_i$	Minimum cost and cost-to-go
$t^*,\ \tau_{m,n',t}^{rs}$	Preferred and realized arrival time
$\Phi_{m,k,t}^{rs}$	Additional cost
$\delta_k^{rs},\ \delta_{\text{RH},k,t}^{rs},\ \delta_{\text{P},i},\ \delta_m^{rs},\ \delta_{\text{FB}}^{rs}$	Monetary costs (PT, RH, parking, operating, feeder bus)
$\kappa_{k,t}^{rs}(n),\ n,\ \sigma_{k,t}^{rs}$	Carpooling impedance, group size, and crowding costs
d_m^{rs}	Distance
ξ	Car accessibility

Symbol	Description
$\lambda_{\mathrm{TT}},\ \lambda_{\mathrm{L}},\ \lambda_{\mathrm{E}},\ \lambda_{\mathrm{C}},\ \lambda_{\mathrm{D}},\ \lambda_{0,t},\ \lambda_{0,m,t}$	Disutility parameters and constants
$\Theta_{n'}(\cdot)$	Schedule-delay function
θ	Logit scale parameter
Deep reinforcement learning	
$i \in \mathcal{N}$	Agent index
$s_t,\ s_{t+1} \in \mathcal{S}$	Current and next-state
$a_t,\ a_i,\ a_i' \in \mathcal{A}$	Actions (current, agent-specific, next-state agent-specific)
$\bar{a}_i,\ \bar{a}_i'$	Mean neighboring action (current and next-state)
r_t	Reward
γ	Discount factor
α	Learning rate
$o_i,\ o_i' \in \mathcal{O}_i$	Observations of agent i (current and next-state)
$c \in \mathcal{C}$	Agent class index
$\mathcal{S},\ \mathcal{A},\ \mathcal{O},\ \mathcal{C},\ \mathcal{N}$	Sets of states, actions, observations, classes, and agents
$Q(s,a),\ V(s),\ \hat{A}(s,a),\ \mathcal{L}(\theta)$	Functions (action–value, state–value, advantage, loss)
$\pi(a_t \mid s_t)$	Policy (action distribution over $\mathcal{A}$)
$\mu_i,\ \mu_{-i}^*$	Policies (agent i, and population excluding i)
$\theta,\ \theta^-,\ \theta_1,\ \theta_2$	Network parameters (Q, target, value, advantage networks)
$\mathbb{E}[\cdot]$	Expectation (over stochastic transitions/policies)

Acknowledgements The authors acknowledge Prof. Lori Tavasszy (TU Delft) for leading and funding the project, and Prof. Bert van Wee (TU Delft) for valuable insights provided during the course Writing a Literature Review in the TIL Domain.

Author Contributions E.S.F. conducted the literature review and wrote the manuscript; R.v.d.M. and E.R.D. provided supervision and critical feedback.

Funding This research was conducted within the Freight Traffic Management as a Service research theme and was funded by the corresponding project.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Jin H, Hao H, Yang X, Wang Z (2023) Joint optimization of stop design and lane assignment for buses with different directions at a signalized intersection. J Trans Eng Part A Syst 149(1):04022113

2. Qian ZS, Zhang HM (2011) Modeling multi-modal morning commute in a one-to-one corridor network. Trans Res Part C Emerging Technol 19(2):254–269
3. Zhang Y, Atasoy B, Akkinepally A, Ben-Akiva M (2019) Dynamic toll pricing using dynamic traffic assignment system with online calibration. Transp Res Rec 2673(10):532–546
4. Kumari KVP, Kumar VNR (2018) Traveler information system to optimized journey using multi agent in a co-modal frame work
5. Gao S, Huang H (2012) Real-time traveler information for optimal adaptive routing in stochastic time-dependent networks. Trans Res Part C Emerging Technol 21(1):196–213
6. Mesa-Arango R, Ukkusuri SV (2014) Modeling the car-truck interaction in a system-optimal dynamic traffic assignment model. J Intell Trans Syst 18(4):327–338
7. Pi X, Ma W, Qian ZS (2019) A general formulation for multi-modal dynamic traffic assignment considering multi-class vehicles, public transit and parking. Trans Res Part C Emerging Technol 104:369–389
8. Levin MW, Boyles SD (2016) A multiclass cell transmission model for shared human and autonomous vehicle roads. Trans Res Part C Emerging Technol 62:103–116
9. Sheffi Y (1985) Urban transportation networks. Prentice-Hall, Englewood Cliffs, NJ
10. Wardrop JG (1952) Road paper. Some theoretical aspects of road traffic research. Proc Inst Civ Eng 1(3):325–362
11. Carroll Jr JD (1959) A method of traffic assignment to an urban network. Highway Res Board Bull (224)
12. Mladenovic M, Trifunovic A (2014) The shortcomings of the conventional four step travel demand forecasting process. J Road Traffic Eng 60(1):5–12
13. Docherty I, Marsden G, Anable J (2018) The governance of smart mobility. Trans Res Part A Pol Pract 115:114–125
14. Paiva S, Ahad MA, Tripathi G, Feroz N, Casalino G (2021) Enabling technologies for urban smart mobility: recent trends, opportunities and challenges. Sensors 21(6):2143
15. Asakura Y, Kashiwadani M (1991) Road network reliability caused by daily fluctuation of traffic flow. In: PTRC summer annual meeting, 19th, 1991, University of Sussex, United Kingdom
16. Peeta S, Ziliaskopoulos AK (2001) Foundations of dynamic traffic assignment: the past, the present and the future. Netw Spat Econ 1:233–265
17. Khan SI, Maini P (1999) Modeling heterogeneous traffic flow. Transp Res Rec 1678(1):234–241
18. Bhavathrathan B, Mallikarjuna C (2012) Evolution of macroscopic models for modeling the heterogeneous traffic: An Indian perspective. Trans Lett 4(1):29–39
19. Verma A (2016) Review of studies on mixed traffic flow: perspective of developing economies. Trans Dev Econ 2:1–16
20. Jeihani M (2007) A review of dynamic traffic assignment computer packages. J Trans Res Forum 46:34–46
21. Wang Y, Szeto WY, Han K, Friesz TL (2018) Dynamic traffic assignment: a review of the methodological advances for environmentally sustainable road transportation applications. Trans Res Part B Methodol 111:370–394
22. Chau ML, Gkiotsalitis K (2025) A systematic literature review on the use of metaheuristics for the optimisation of multimodal transportation. Evol Intel 18(2):1–37
23. Irwin N, Dodd N, Von Cube H (1961) Capacity restraint in assignment programs. Highway Res Board Bull (297)
24. Irwin N, Van Cube H (1962) Capacity restraint in multi-travel mode assignment programs. Highway Res Board Bull (347)
25. Manheim ML, Ruiter ER (1970) Dodotrans i: a decision-oriented computer language for analysis of multimode transportation systems. Technical report
26. Wigan MR, Bamford T (1973) An equilibrium model of bus and car travel over a road network. Technical report
27. Florian M (1977) A traffic equilibrium model of travel by car and public transit modes. Transp Sci 11(2):166–179
28. Abdulaal M, LeBlanc LJ (1979) Methods for combining modal split and equilibrium assignment models. Transp Sci 13(4):292–314
29. Hall M, Willumsen L (1980) Saturn-a simulation-assignment model for the evaluation of traffic management schemes. Traffic Eng Control 21(4)
30. Van Vliet D (1982) Saturn-a modern assignment model. Traffic Eng Control 23(HS-034 256)

31. Willumsen LG, Bolland J, Hall M, Arezki Y (1993) Multi-modal modelling in congested networks: SATURN and SATCHMO. Traffic Eng Control 34(6):294–301
32. Dafermos S (1982) The general multimodal network equilibrium problem with elastic demand. Networks 12(1):57–72
33. Merchant DK, Nemhauser GL (1978) A model and an algorithm for the dynamic traffic assignment problems. Transp Sci 12(3):183–199
34. Merchant DK, Nemhauser GL (1978) Optimality conditions for a dynamic traffic assignment model. Transp Sci 12(3):200–207
35. Marwah B, Bandyopadhyay S (1983) Development of a traffic simulation model for an Indian city street. J Indian Roads Congress 46:655–695
36. Dehoux M, Toint L (1990) PACSIM: a dynamical traffic assignment model. II: functional analysis. GRT Report 90/na, Department of Mathematics, FUNDP Namur, Belgium
37. Daganzo CF (1994) The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Trans Res Part B Methodol 28(4):269–287
38. Watling D (1996) Asymmetric problems and stochastic process models of traffic assignment. Trans Res Part B Methodol 30(5):339–357
39. Zhang H, Jin W (2002) Kinematic wave traffic flow model for mixed traffic. Transp Res Rec 1802(1):197–204
40. Lighthill MJ, Whitham GB (1955) On kinematic waves II. A theory of traffic flow on long crowded roads. Proc R Soc London Ser A Math Phys Sci 229(1178):317–345
41. Richards PI (1956) Shock waves on the highway. Oper Res 4(1):42–51
42. Ziliaskopoulos AK (2000) A linear programming model for the single destination system optimum dynamic traffic assignment problem. Transp Sci 34(1):37–49
43. Mahmassani H (1992) Dynamic traffic assignment and simulation for advanced network informatics (DYNASMART). In: The 2nd international seminar on urban traffic networks 1992
44. Ben-Akiva M, Cascetta E, Gunn HF (1995) An on-line dynamic traffic prediction model for an inter-urban motorway network. In: Gartner NH, Improta G (eds) Urban traffic networks: dynamic flow modeling and control. Springer, Berlin, Heidelberg, pp 83–122. https://doi.org/10.1007/978-3-642-79641-8_4
45. Yang QI, Koutsopoulos HN (1996) A microscopic traffic simulator for evaluation of dynamic traffic management systems. Trans Res Part C Emerging Technol 4(3):113–129
46. Gartner NH, Stamatiadis C (1998) Integration of dynamic traffic assignment with real-time traffic adaptive control system. Transp Res Rec 1644(1):150–156
47. Meng M, Shao C, Zeng J, Dong C (2014) A simulation-based dynamic traffic assignment model with combined modes. Promet-Traffic Trans 26(1):65–73
48. Liu H, Wang J, Wijayaratna K, Dixit VV, Waller ST (2015) Integrating the bus vehicle class into the cell transmission model. IEEE Trans Intell Transp Syst 16(5):2620–2630
49. Verbas İÖ, Mahmassani HS, Hyland MF (2015) Dynamic assignment-simulation methodology for multimodal urban transit networks. Transp Res Rec 2498(1):64–74
50. Jiang Y, Szeto W, Long J, Han K (2016) Multi-class dynamic traffic assignment with physical queues: intersection-movement-based formulation and paradox. Transportmetrica A Trans Sci 12(10):878–908
51. Lin Y-Z, Chen W-H (2017) A simulation-based multiclass, multimodal traffic assignment model with departure time for evaluating traffic control plans of planned special events. Trans Res Proc 25:1352–1379
52. Qian ZS, Li J, Li X, Zhang M, Wang H (2017) Modeling heterogeneous traffic flow: a pragmatic approach. Trans Res Part B Methodol 99:183–204
53. Mayakuntla SK, Verma A (2019) Cell transmission modeling of heterogeneous disordered traffic. J Trans Eng Part A Syst 145(7):04019027
54. Saprykin A, Chokani N, Abhari RS (2019) GEMSim: a GPU-accelerated multi-modal mobility simulator for large-scale scenarios. Simul Model Pract Theory 94:199–214
55. Wang J, Peeta S, He X (2019) Multiclass traffic assignment model for mixed traffic flow of human-driven vehicles and connected and autonomous vehicles. Trans Res Part B Methodol 126:139–168
56. Ngoduy D, Hoang N, Vu H, Watling D (2021) Multiclass dynamic system optimum solution for mixed traffic of human-driven and automated vehicles considering physical queues. Trans Res Part B Methodol 145:56–79

57. Barzegari V, Edrisi A, Nourinejad M (2023) Fleet cost and capacity effects of automated vehicles in mixed traffic networks: a system optimal assignment problem. Trans Res Part C Emerging Technol 148
58. Liu B, Ji Y, Cats O (2025) Integrating ride-hailing services with public transport: a stochastic user equilibrium model for multimodal transport systems. Transportmetrica A Trans Sci 21(1):2236240
59. Bamdad B, Erdmann J, Sgambi L, Seyedabrishami S, Snelder M (2025) A multiclass simulation-based dynamic traffic assignment model for mixed traffic flow of connected and autonomous vehicles and human-driven vehicles. Transportmetrica A Trans Sci 21(2):2257805
60. Shen Y, Qian Z (2025) Multi-modal dynamic transit assignment for transit networks incorporating bike-sharing. Future Trans 5(4):148
61. Johansson G, Rumar K (1971) Drivers' brake reaction times. Hum Factors 13(1):23–27
62. McFadden D (1974) Conditional logit analysis of qualitative choice behavior. Academic Press, pp 105–142
63. Ben-Akiva ME, Lerman SR, Lerman SR et al (1985) Discrete choice analysis: theory and application to travel demand, vol 9. MIT press, Cambridge, MA
64. Shou Z, Chen X, Fu Y, Di X (2022) Multi-agent reinforcement learning for Markov routing games: a new modeling paradigm for dynamic traffic assignment. Trans Res Part C Emerging Technol 137
65. Jiang S, Tran CQ, Keyvan-Ekbatani M (2024) Regional route guidance with realistic compliance patterns: application of deep reinforcement learning and MPC. Trans Res Part C Emerging Technol 158
66. Gu Z, Wang Y, Ma W, Liu Z (2024) A joint travel mode and departure time choice model in dynamic multimodal transportation networks based on deep reinforcement learning. Multimodal Trans 3(3):100137
67. Schultz L, Sokolov V (2018) Deep reinforcement learning for dynamic urban transportation problems. arXiv preprint arXiv:1806.05310
68. Koh S, Zhou B, Fang H, Yang P, Yang Z, Yang Q, Guan L, Ji Z (2020) Real-time deep reinforcement learning based vehicle navigation. Appl Soft Comput 96:106694
69. Pei H, Zhang J, Zhang Y, Xu H, Li L (2023) Self-organized routing for autonomous vehicles via deep reinforcement learning. IEEE Trans Veh Technol
70. Li L, Zhao W, Wang C, Fotouhi A, Liu X (2024) Nash double q-based multi-agent deep reinforcement learning for interactive merging strategy in mixed traffic. Expert Syst Appl 237:121458
71. Ke Z, Zou Q, Liu J, Qian S (2025) Real-time system optimal traffic routing under uncertainties—can physics models boost reinforcement learning? Trans Res Part C Emerging Technol 173:105040
72. Mao C, Shen Z (2018) A reinforcement learning framework for the adaptive routing problem in stochastic time-dependent network. Trans Res Part C Emerging Technol 93:179–197
73. Lasry J-M, Lions P-L (2007) Mean field games. Japan J Math 2(1):229–260
74. Huang M, Malhamé RP, Caines PE (2006) Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle
75. Yang Y, Luo R, Li M, Zhou M, Zhang W, Wang J (2018) Mean field multi-agent reinforcement learning. In: International conference on machine learning. PMLR, pp 5571–5580
76. Ruthotto L, Osher SJ, Li W, Nurbekyan L, Fung SW (2020) A machine learning framework for solving high-dimensional mean field game and mean field control problems. Proc Natl Acad Sci 117(17):9183–9193