50 years of behavioral models for transportation and logistics

Léa Ricard *		Michel Bierlaire *

February 20, 2024

Report TRANSP-OR 220408
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
`transp-or.epfl.ch`

*École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland, {lea.ricard,michel.bierlaire}@epfl.ch

Abstract

Fifty years ago, transportation and logistics problems were primarily analyzed either from a supply-side or a demand-side perspective, with the fields of operations research and demand modeling evolving separately. Since then, there has been a growing interest in behavioral optimization models, aiming to integrate both supply and demand aspects. The purpose of this survey is to offer a historical perspective on the evolution of behavioral optimization models in transportation and logistics. It does so by delving into significant works in demand modeling and choice-based optimization, highlighting their interconnections. In particular, we focus on three important optimization applications, namely the facility location problem, the airline scheduling and fleet assignment problem, and the urban transportation planning problem. Additionally, we identify potential avenues for future research to bridge existing gaps in the literature and promote behavioral models in transportation and logistics.

1 Introduction

Over the past half-century, the field of transportation and logistics has undergone an important transformation in the representation of humans and demand within optimization models. This evolution has been driven by the recognition that traditional approaches, which primarily focused on operational efficiency and profit maximization, fell short in capturing the complex interplay between supply and demand, as well as the preferences and choices of individuals within these systems. In transportation, as in other fields, demand can be characterized by aggregating the choices made by a heterogeneous population of actors (travelers, customers, etc.); the primary method for representing disaggregate demand is the use of choice models. Using this more complex representation of demand, behavioral, or choice-based, optimization models are an answer to the challenge of capturing the interactions between supply and demand. These models aim to influence the behavior and choices of individuals by incorporating their preferences into the optimization process. A prime example of behavioral change is an individual's decision to adopt a new transportation mode, such as using public transportation instead of a personal car for daily commuting or switching to a different supplier for purchasing products. In this article, we delve into the historical development of these behavioral optimization models, tracing their evolution over time and shedding light on their significance in the fields of transportation and logistics.

One key aspect of this transformation is the changing role of humans in optimization models. Initially, transportation and logistics models predominantly centered on operators, aiming to maximize revenues without considering the utility of passengers or customers. However, a paradigm shift occurred over time, leading to the emergence of passenger-centric models. This shift gained prominence at the end of the 2000s, with seminal works like that of Lu et al. (2009), and has continued to grow in popularity. As evidence of this trend, 36 publications with the term 'passenger-centric' in their title, abstract, or keywords were found within the Web of Science database, with half of them being published in the last five years.

Simultaneously, the representation of demand within transportation and logistics models has also undergone major evolution within the operations research literature. Fifty years ago, demand was often represented using simplified models, but these representations have since evolved into more complex and accurate forms. Demand can be represented through aggregated or disaggregated models. The former treats demand as aggregated flows and thus tends to overlook individual preferences. One of the pioneering works in aggregated demand modeling is Talvitie (1973), which introduced the so-called econometric model. In contrast, disaggregated models incorporate individual preferences and sometimes account for heterogeneity among individuals within a population. The state-of-the-art of disaggregated demand modeling is discrete choice models (Pacheco Paneque, 2020), with most of them being random utility models, assuming that economic agents are solving an optimization problem. The popularization of the logit model (Ben-Akiva, 1973; McFadden, 1973) marked the beginning of this trend fifty years ago, along with the development of new mathematical reformulations for the transportation problem, often referred to as the Monge–Kantorovich problem (Intrator, 1973; Appa, 1973). It is worth noting that the identification of discrete choice models is a specific instance of the Monge–Kantorovich problem (i.e., the solution of the dual of this Monge–Kantorovich problem corresponds to the individual preferences) (Chiong et al., 2016). In recent years, more complex and realistic discrete choice models, such as the mixture of logit model, have been proposed, enriching the repertoire of tools available to researchers in this domain.

These two parallel changes— the shift towards passenger-centric models and the development of more sophisticated demand representations— are closely interconnected. Models centered on passengers require a disaggregated representation of demand to accurately capture the nuanced choices and preferences of individuals within the transportation and logistics systems.

This article aims to provide transportation researchers with a historical perspective on the development of behavioral optimization models in transportation and logistics. Our objective is not to offer a comprehensive literature review but rather to offer insightful commentary on the historical evolution and prospective future of these models. We will review significant works in the fields of demand modeling and choice-based optimization from the past half-century, discussing their contributions and implications. Furthermore, we will present perspectives for future research in this dynamic and evolving field. The remainder of this paper is organized as follows. Section 2 introduces the key foundations of demand modeling in transportation and logistics and reviews major works published over the past fifty years. Section 3 discusses the methodological developments in the choice-based optimization field, makes connections between contributions for three applications in transportation and logistics, gives a general overview of the challenges inherent in this type of model. Section 4 highlights several promising research perspectives. Finally, Section 5 concludes the paper.

2 Modeling demand

Demand is defined here as "the quantity of a good that consumers are willing and able to purchase at various prices during a given time" (Wikipedia). It is intimately related to human behavior and choice, as individuals make decisions based on their preferences, needs, and constraints.

In logistics, demand can take various forms, such as the number of customers buying from a firm, the volume of products, or the count of students registering at a school. In contrast, in transportation, travel demand encompasses both short-term decisions (e.g., choosing activities, departure times, transportation modes, and itineraries) and long-term decisions (e.g., whether to purchase a car or not and housing decisions). It can be considered a 'derived' demand because individuals do not travel to get around per se, but rather to carry out activities distributed in space and time. Traditionally, it was modeled using the four-step approach, but it can also be represented using disaggregated models or discrete choice model extensions, such as activity-based modeling and dynamic choice models. In this section, we will review the latter approaches. Note that they all (except the four-step model) also apply to choices made by individuals in a logistics system.

2.1 Aggregate demand model: the four-step model

The evolution of travel demand modeling has witnessed the development of various approaches and techniques. One of the seminal models in the field is the four-step model, which emerged in the 1950s during the post-war development era in the United States (McNally, 2007). This model provides a structured framework for analyzing demand and can be applied to various transportation scenarios.

In the four-step model, demand is represented as flows, with the modeling process decomposed into four sequential subproblems:

1. trip generation— the objective is to estimate the number of trips originating from different areas. This problem can be framed as a location choice problem, where individuals decide where to start their journeys.
2. trip distribution— distribute the trips to potential destinations. This step can be conceptualized as a multi-commodity flow problem, where trips are allocated among available destinations.
3. mode split— allocate trips to different travel modes. This allocation can be based on predetermined choices or modeled using choice models.
4. trip assignment— assign trips to specific travel routes. This process can be viewed as an 'all-or-nothing' assignment, where for each trip the shortest path is selected. In order to account for congestion, it must be treated as a Nash equilibrium problem and solved using methods like the Frank-Wolfe algorithm.

While the four-step model offers a comprehensive framework for travel demand modeling, it primarily focuses on aggregated decisions. The representation of demand with flows is associated with an implicit assumption of behavioral homogeneity. Although several approaches have attempted to relax this assumption by defining multiple classes of travelers, the need to capture travel behavior at the individual level has become necessary.

2.2 Disaggregated demand models: discrete choice models

While there exist several types of disaggregated choice models, e.g., Markov chain models and rank-based models, we focus here on the state-of-the-art approaches, namely discrete choice models. These models are rooted in microeconomic principles and random utility theory. They assume that individuals are rational decision-makers who seek to maximize their utility, subject to budget constraints. Formally, each individual solves

$$\max_{q \in X} \{u(q) | p^T q \leq I\},$$

where X is the set of all possible decisions, I is the total budget or capacity, p is a vector of the prices or resource consumption, q is a decision, and $u(q)$ its utility.

When $X = \mathbb{R}^L$, i.e., when the decisions are continuous, this optimization problem corresponds to continuous optimization. However, when some or all decisions are discrete, it transforms into a mixed integer optimization problem (MIP). Since first-order (and second-order) optimality conditions do not apply to MIPs, demand functions (i.e., functions that output the quantity of each product purchased or chosen as a function of its attributes) cannot easily be derived, as is the case for the continuous case.

Consider the case with two alternatives. U_{in} is the utility associated with individual n and with alternative i and U_{jn} with alternative j. In reality, U_{in} is a random variable defined as $U_{in} = v_{in} + \varepsilon_{in}$, where v_{in} is deterministic and the error term ε_{in} is a continuous random variable. The deterministic term v_{in} is usually expressed as a linear function $v_{in} = \beta_n^T x_{in}$ of the model's parameters β_n^T and the vector of variables x_{in}, including the socioeconomic characteristics of individual n and the attributes of alternative i. Discrete choice models differ in terms of the assumptions they make about the distribution of the error term, and the fact that the whole population can be modeled by the same β parameters or not.

The binary logit model is the specific case where we assume the error terms ε_{in} and ε_{jn} to be independent and identically distributed across i and n and their difference $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ to follow a logistic regression. This model and its version with multiple alternatives has a closed-form probabilistic expression. It was popularized by the American economist and Economics Nobel Prize winner Daniel McFadden in McFadden (1973) and is still widely used in the literature. However, the logit model has been widely criticized because of its assumptions that can lead to unrealistic forecasts of individual choice. Particularly, it performs poorly when there are complex substitution patterns due to one of its properties called Independence of Irrelevant Alternatives (IIA), which is a direct consequence of the independence assumption for the error terms.

To address these limitations, more complex discrete choice models have been introduced over the past 50 years:

1. nested logit— alternatives are grouped into nests such that in each nest the IIA assumption holds. This model has a closed form probabilistic expression, like the logit model.
2. cross-nested logit (Small, 1987; Vovsha, 1997; Papola, 2004; Bierlaire, 2006)— alternatives are grouped into nests such that nests may overlap. This model can capture mixed interactions among alternatives and has a closed-form probabilistic expression.
3. mixture of logit— various parameters of the logit model can be assumed to be distributed, in order to capture unobserved taste heterogeneity. McFadden and Train (2000) showed that a mixture of logit model can approximate any discrete choice model. However, it does not have a closed-form probabilistic expression.

These models generally represent individual behavior better than the simple logit model, but parameter estimation is more difficult. Note that the logit, the nested logit, and the cross-nested logit models are part of the so-called multivariate extreme value (MEV) family of choice models introduced by McFadden (1978), which includes many other models (some of which have not yet been exploited). Interested readers can find complementary information on discrete choice models in the following textbooks: Ben-Akiva and Lerman (1985) and Train (2009).

2.3 Disaggregate demand models: beyond discrete choice models

Traditionally, demand modeling often focuses on isolated trip-level decisions, failing to account for correlations between multiple trips made by the same individual or within a household. Two innovative approaches have emerged to address these shortcomings: activity-based modeling and dynamic choice models.

Activity-based modeling. These models represent a shift in perspective, viewing travel not as an end in itself but as a mean to fulfill specific activities within time and space constraints. It provides a more comprehensive understanding of traveler behavior by considering the broader context of activities and their dependencies. Recent advancements in activity-based modeling, such as the framework proposed by Pougala et al. (2023), have further enhanced our ability to capture the dynamics of demand in activity-based contexts. We refer interested readers to the survey of Rasouli and Timmermans (2014) for further details on this topic.

Dynamic choice models. These models extend discrete choice modeling to capture sequential decision processes where each decision depends on previous ones. These models are particularly relevant when analyzing long-term correlations in decision-making, such as purchasing a house at time t followed by buying a car at time $t+1$, and planning a road trip for the holidays at time $t+2$. To accommodate such correlations, discrete choice models may rely on synthetic data to infer future individual or household attributes. Recent works by Aguirregabiria and Mira (2010), Bierlaire et al. (2021), and Cirillo and Xu (2011) have made significant contributions to the field of dynamic choice modeling, allowing for a more nuanced understanding of complex decision sequences.

3 Choice-based optimization

Individual choice is best represented by disaggregated choice models, for which discrete choice models are the state-of-the-art. In this section, we review optimization models accounting for endogenous demand through an explicit incorporation of discrete choice models inside the optimization problem. Note, however, that there is a large number of works introducing optimization models based on aggregated demand models, such as

1. Time-varying unimodal origin-destination matrices (see, e.g., Yin et al., 2016, 2017; Binder et al., 2017; Tong et al., 2017; Szymula and Bešinović, 2020; Polinder et al., 2022; Wang et al., 2022),
2. Spatial-temporal demand models using massive data (see, e.g., Tu et al., 2016; Ma et al., 2023),
3. Gravity models (see, e.g., Birolini, Jacquillat, Cattaneo and Antunes, 2021; Tiwari et al., 2021; Birolini et al., 2023),
4. General attraction models (see, e.g., Wei et al., 2020; Yan et al., 2022).

Yet, integrating aggregated demand models into optimization problems is much more straightforward than integrating discrete choice models, thus our focus is on the latter. Note also that many works in the literature embed another type of disaggregated choice model, namely Markov chain choice models, into optimization problems (see, e.g., Blanchet et al., 2016; Feldman and Topaloglu, 2017; Ahipaşaoğlu et al., 2019). This parallel research direction is however outside of the scope of our review.

The remainder of this section is organized as follows. First, notable methodological developments, both in terms of modeling and solution method, are presented in Section 3.1. Second, we review in Section 3.2 three applications in transportation in logistics. Based on the models for these applications, we provide general comments on challenges of choice-based optimization in Section 3.3.

3.1 Methodological development

Over the past few years, there have been methodological developments aimed at overcoming the computational complexity of choice-based optimization and incorporating realistic features of optimization problems. On the one hand, when demand modeling is too complex to be directly included in an optimization problem, two approaches can be employed: (i) simulation-based optimization (Gosavi et al., 2015) and (ii) numerical approximation (Gilbert et al., 2014b). However, while (i) is time-consuming, (ii) comes at the cost of a decrease in the quality of the demand approximation. Therefore, one must carefully weigh the advantages and disadvantages of each method to choose the most appropriate one in each context. On the other hand, several works have proposed methodological advances for choice-based optimization with competition among multiple suppliers. In the following section, we provide more details on these three crucial methodological developments.

Simulation-based optimization. Simulation-based optimization is a family of techniques that optimize a stochastic simulation to find good decisions or strategies. In the case of choice-based optimization, simulation can be used to approximate the demand for a service or a product by generating several random choices based on the utility functions and using a sample average approximation method (see, e.g., Haase and Müller, 2013; Legault and Frejinger, 2022, for works relying on simulated demand). This method avoids the use of the choice probabilities, which may be highly nonlinear and not have a closed-form probabilistic expression.

Recently, Gupta et al. (2020) and Osorio and Atasoy (2021) developed simulation-based frameworks for the transportation demand management problem, where the traffic assignment subproblems are simulated. The transportation demand management problem consists of optimizing a set of control strategies (e.g., tolls or ramp metering rates) to minimize traffic congestion. Gupta et al. (2020) proposed a simulation-based model for the real-time transportation demand management problem. A genetic algorithm is applied to solve this problem, treating the traffic prediction simulation as a black box. In Osorio and Atasoy (2021), a non-linear analytical network model is introduced. This network model can approximate the mapping between the toll vector and the network performance (given, e.g., by the congestion level). The model is integrated into a metamodel for simulation-based optimization, where the simulator is no longer treated as a black box. Analytical information from the network model is combined with information from the simulator, resulting in high computational efficiency and the ability to handle large-scale cases.

In the context of MIPs, Pacheco et al. (2021) presented a general simulation-based framework that can be applied to any problem and given any discrete choice model. In order to solve instances of relevant size, the authors extended the latter framework in Pacheco et al. (2022) to include a Lagrangian decomposition approach.

Numerical approximation. The demand (or flow) assignment can be replaced by numerical approximations that can provide demand (or flow) in closed form. On the one hand, piecewise uniform or linear approximations of the demand functions are common practice (see, e.g., Gilbert et al., 2014b, 2015; Cadarso et al., 2017; Birolini, Antunes, Cattaneo, Malighetti and Paleari, 2021). Alternatively, for a mixture of logit model, the distributions of the factors specifying the heterogeneity in the population (e.g., the price sensitivity of individuals) can also be approximated by piecewise uniform or linear distributions. In Birolini, Antunes, Cattaneo, Malighetti and Paleari (2021), the number of pieces, the breaking points, and the feasible region of the piecewise linear function are selected such that the approximation error is minimized based on a least squared fit. On the other hand, a stochastic choice model can be replaced by a deterministic choice model (also referred to as an "all-or-nothing" assignment) by dropping the error term (see, e.g., Gilbert et al., 2014b, 2015).

Competition. Bortolomiol et al. (2021), Cadarso et al. (2017), Liu et al. (2019), and Gallego and Wang (2014) have investigated supply and demand interactions involving multiple suppliers. This situation can be characterized as a noncooperative game, either a Nash-type game when all players possess the same status or a Stackelberg-type game when there is a hierarchical relationship among the players. Nash games have been extensively researched in the context of an oligopolistic marketplace (i.e., a marketplace with a small number of competing suppliers). Bortolomiol et al. (2021) introduced a simulation-based optimization heuristic that is close to a fixed-point iteration algorithm for approximating equilibria in oligopolistic-type competition. Furthermore, the existence and uniqueness conditions of equilibrium in games involving various discrete choice models are studied in Milgrom and Roberts (1990), Bernstein and Federgruen (2004), Gallego et al. (2006), Kök and Xu (2011), and Li and Huh (2011), among others.

3.2 Applications

We review choice-based optimization through the prism of three key problems: facility location, airline scheduling and fleet assignment, and urban transportation planning. These problems differ in terms of their complexity, structure, and supply-demand interactions, among other things, as we discuss next.

3.2.1 Facility location

The facility location problem (or the maximum capture location problem) is to locate r new facilities in a market such that the number of customers buying from the facilities of the firm is maximized. Let $\mathcal{M}$ be the set of available locations for the facilities and $\mathcal{J} \subset \mathcal{M}$ be the facility locations selected in the problem solution to be opened. Existing competitors in the market, denoted by $\mathcal{X} \subset \mathcal{M}$, can be considered or not. The customers are located in zones $i \in \mathcal{I}$, where each zone is associated with a number q_i of customers. This problem can be formulated as

$$\max z = \sum_{i \in \mathcal{I}} q_i \sum_{j \in \mathcal{J}} P(i, j|\mathcal{J}, \mathcal{X}) \tag{1}$$

$$\tag{2}$$

where $P(i, j|\mathcal{J}, \mathcal{X})$ is the proportion of customers located in i choosing facility $j \in \mathcal{J}$. When the choice probability $P(i, j|\mathcal{J}, \mathcal{X})$ is given by a logit model, this problem can be formulated as

$$\max z = \sum_{i \in \mathcal{I}} q_i \frac{\sum_{j \in \mathcal{M}} e^{v_{ij}} y_j}{\sum_{j \in \mathcal{X}} e^{v_{ij}} + \sum_{j \in \mathcal{M}} e^{v_{ij}} y_j}, \tag{3}$$

$$\text{s.t.} \sum_{j \in \mathcal{M}} y_j = r, \tag{4}$$

$$y_j \in \{0, 1\}, \qquad \forall j \in \mathcal{J}, \tag{5}$$

where v_{ij} is the deterministic part of the utility of location j for customers in zone i (i.e., $u_{ij} = v_{ij} + \varepsilon_{ij}$) and y_j is equal to 1 if facility j is selected and 0 otherwise.

In recent years, there have been significant developments in both the formulation of the facility location problem and the solution techniques used to solve it, as detailed next. The seminal work of Benati and Hansen (2002) introduced the logit-based facility location problem which was proven to be NP-hard (Benati, 1999). Additionally, they provided the first two linear reformulations of this problem, one of which is taking advantage of the submodularity of the objective function—a key insight from an earlier work (Benati, 1997). Furthermore, an exact branch-and-bound method to solve the linearized problems and a variable neighborhood search heuristic to solve larger instances are proposed.

The logit-based facility location problem was later enhanced when Haase (2009) introduced another linear reformulation, leveraging the constant substitution pattern inherent to the logit model. Subsequently, in Haase and Müller (2014), this new reformulation was compared to the ones proposed in Benati and Hansen (2002), and it was found to be more efficient. Furthermore, Haase (2009) proposed a second formulation that considers discrete choice models with flexible substitution patterns. To address this formulation, they employed a simulation-based approach, specifically a sample average approximation framework, to derive solutions.

In Freire et al. (2016), the mixed integer linear optimization problem (MILP) reformulation of Haase (2009) is strengthened by incorporating tighter coefficients in some inequalities. Furthermore, they developed a new branch-and-bound algorithm with a greedy approach. Their method was found to be more efficient than benchmark approaches on several numerical experiments, notably for instances of the park-and-ride facility location problem.

Ljubić and Moreno (2018) made a significant contribution by introducing a new branch-and-cut algorithm that combines two types of cutting planes, namely outer-approximation cuts and submodular cuts. Their approach significantly improved the state-of-the-art exact approaches for the facility location problem. This work stayed state-of-the-art until Mai and Lodi (2020) introduced more efficient outer-approximation cuts that generate cuts for groups of zones instead of cuts for each zone. Their solution method also differs from that of Ljubić and Moreno (2018) because they use their cuts in a cutting plane approach instead of a branch-and-cut. Moreover, they tested their approach for both a logit and a mixture of logit discrete choice model.

Dam et al. (2022) studied the facility location problem employing the MEV family of discrete choice models, which includes the logit and the nested logit models, but not the mixture of logit. They demonstrated that the objective function is submodular under any MEV type of discrete choice model and developed a new algorithm combining a greedy heuristic with gradient-based local search and an exchange procedure to solve their mixed-integer nonlinear optimization problem. Legault and Frejinger (2022) extended the submodularity proof for objective functions with any random utility function maximization model. They estimated demand using a generalized version of the sample average approximation method of Haase (2009) and developed a new branch-and-cut algorithm.

A facility location problem with a sequential decision process is introduced in Méndez-Vogel et al. (2023). A nested logit models the sequential decision process of first choosing a firm, and then choosing the store to buy from (i.e., the stores are nested by their firm). The resulting mixed-integer nonlinear optimization model is solved using a branch-and-cut algorithm featuring new cuts.

Müller et al. (2009) and Haase and Müller (2013) addressed a variant of the problem for school location planning with free school choice. In Müller et al. (2009), a two-step problem for the multi-period variant is proposed. In the first step, a set of school planning scenarios is generated and, for each scenario, a mixture of logit model gives the share of students choosing the available schools given capacity constraints. Then, in the second step, the scenario that minimizes the costs is selected. Later, Haase and Müller (2013) presented a model for the school location planning problem that can accommodate any random utility function while considering capacity and budget constraints. School choice is given by simulated utility values obtained using a sample average approximation framework and the resulting problem is solved using a standard MIP solver.

Zhang et al. (2012) developed a model for the preventive care facility location and configuration problem considering constraints on mean waiting time, workload requirements, capacity, and server availability. In addition to selecting the location of the new facilities to open, the capacity of each facility (i.e., the number of servers allocated) must be decided. Each facility is modeled as a M/M/c queuing system, where c is the number of servers. The MIP formulation is solved using two heuristics, a probabilistic search-based algorithm and a genetic algorithm.

Finally, a park-and-ride facility location problem using a logit model for the mode choice is introduced in Aros-Vera et al. (2013). The linearized problem is solved using a heuristic concentration integer approach.

Table 1 summarizes the main attributes of the literature on the choice-based facility location problem reviewed. The columns display the discrete choice model used, the choices modeled by the discrete choice models, the type of problem (linear - L, non-linear - NL, and integer - INT), the solution method employed (Sol. method), and the additional constraints included (Add. constr.).

Authors	Discrete choice model			Choice	Type of problem			Sol. method	Add. constr.
	Logit	NeL	MoL		L	NL	INT		
Benati (1999)	•			firm	•		•	MILP, MH	
Haase (2009)	•			firm	•		•	MILP, SAA	
Müller et al. (2009)			•	school		•	•	QP	capacity
Zhang et al. (2012)	•			hospital	•		•	MH	capacity, service quality
Aros-Vera et al. (2013)	•			mode	•		•	MILP, MH	
Haase and Müller (2013)			•	school	•		•	MILP, SAA	capacity, budget
Freire et al. (2016)	•			firm	•		•	B&B	
Ljubić and Moreno (2018)	•			firm	•		•	B&C	
Mai and Lodi (2020)	•		•	firm	•		•	B&C	
Dam et al. (2022)	•	•		firm		•	•	H	
Legault and Frejinger (2022)	•		•	firm	•		•	B&C, SAA	
Méndez-Vogel et al. (2023)		•		firm, store		•	•	B&C	

Table 1: Main attributes of selected choice-based facility location problems. NeL - nested logit model, MoL - mixture of logit model, L - linear, NL - nonlinear, INT - (mixed-)integer. Solution method: B&B - branch-and-bound method, B&C - branch-and-cut, H - heuristic, MH - metaheuristic, MILP - mixed integer linear optimization method, QD - quadratic optimization method, SAA - sample average approximation

We can observe that our understanding of the facility location problem, especially the submodularity of its objective function under different classes of discrete choice models, has evolved over the years. Nowadays, we know that the submodularity of the objective function holds for any random utility function maximization model (including MEV's type of discrete choice model) (Legault and Frejinger, 2022). This feature allows the problem to be easily transformed into a MILP. Furthermore, efficient branch-and-cut algorithms are readily available. For specific applications, additional constraints, such as capacity or budget constraints, may make this problem more complex to solve, and efficient heuristics are often needed to tackle real-life instances.

3.2.2 Airline scheduling and fleet assignment

The integrated airline schedule design and fleet assignment problem consists of determining flight schedules (departure and arrival times) and assigning an airline fleet type to each scheduled flight in order to maximize profit. This problem was formulated as a multicommodity flow problem in Hane et al. (1995). We present here a general mathematical formulation for the incremental airline schedule design and fleet assignment problem, which considers an initial schedule with mandatory and optimal flight legs.

Let $\mathcal{L}$ denote the set of flight legs, with $\mathcal{L}_m$ and $\mathcal{L}_o$ the sets of mandatory and optional flights, respectively. Let also $\mathcal{F}$ be the set of all fleet types and $\mathcal{F}(\ell)$ be the set of fleet types compatible for a given flight leg $\ell \in \mathcal{L}$. For each fleet type $f \in \mathcal{F}$, we build a time-space network with node set N_f. In this time-space network, $\text{In}(v)$, $\text{Out}(v)$ are the sets of inbound and outbound flight arcs from node $v \in N_f$, respectively. Furthermore, y_v represents the number of aircraft on the ground at node v. For every fleet type $f \in \mathcal{F}$, let $T_F(f)$ and $T_N(f)$ denote the sets of flight and ground arcs in the time-space network at count time, respectively. The count time is set arbitrarily.

We denote $\mathcal{H}$ the set of cabin classes (e.g., economy and business). Let also $\mathcal{I}$ be the set of flight itineraries, and $\mathcal{I}(\ell)$ be the itineraries requiring flight $\ell \in \mathcal{L}$. For each cabin class $h \in \mathcal{H}$ and itinerary $i \in \mathcal{I}$, p_{ih} is the price of the flight ticket, and s_{ih} is the demand for the corresponding product.

The mixed-integer problem also uses the following notation. Let $x_{\ell f}$ be a binary variable equal to 1 if fleet type f is assigned to flight ℓ and 0 otherwise. This problem can be expressed as:

$$\max \sum_{i\in\mathcal{I}} \sum_{h\in\mathcal{H}} p_{ih} s_{ih} - \sum_{\ell\in L} \sum_{f\in\mathcal{F}(\ell)} c_{\ell f} x_{\ell f}, \tag{6}$$

$$\text{s.t.} \sum_{f\in\mathcal{F}(\ell)} x_{\ell f} = 1, \forall \ell \in \mathcal{L}_m, \tag{7}$$

$$\sum_{f\in\mathcal{F}(\ell)} x_{\ell f} \leq 1, \forall \ell \in \mathcal{L}_0, \tag{8}$$

$$y_{v^-} + \sum_{\ell\in\text{In}(v)} x_{\ell f} = y_v + \sum_{\ell\in\text{Out}(v)} x_{\ell f}, \quad \forall f \in \mathcal{F}, v \in N_f, \tag{9}$$

$$\sum_{v\in T_N(f)} y_v + \sum_{\ell\in T_F(f)} x_{\ell f} \leq n_f, \forall f \in \mathcal{F}, \tag{10}$$

$$\sum_{h\in\mathcal{H}} \sum_{i\in\mathcal{I}(\ell)} s_{ih} \leq \sum_{f\in\mathcal{F}(\ell)} Q_f x_{\ell f}, \forall \ell \in \mathcal{L}, \tag{11}$$

$$s_{\ell h} \geq 0, \forall \ell \in \mathcal{L}, h \in \mathcal{H}, \tag{12}$$

$$y_v \geq 0, \forall f \in \mathcal{F}, t \in \mathcal{T}, v \in N_{ft}, \tag{13}$$

$$x_{\ell f} \in \{0, 1\}, \forall f \in \mathcal{F}(\ell), \ell \in \mathcal{L}, \tag{14}$$

where v^- is the predecessor node of node v, n_f is the number of available aircraft of type f, Q_f is the capacity of aircraft type f, and $c_{\ell f}$ is the operational costs of flight ℓ when using aircraft type f. Note that p_{ih}, for $i \in \mathcal{I}, h \in \mathcal{H}$, can be fixed or considered as a decision variable. In the latter case, the problem includes a pricing component. The objective function (6) maximizes the profit, defined as the revenue minus the operational costs. Constraints (7) and (8) ensure that the mandatory flights and the optional flights are covered exactly once and at most once, respectively. The flow conservation in each time-space network is guaranteed by constraints (9). Constraints (10) limit the number of aircraft of each type used and constraints (11) are to ensure the sales on a flight do not exceed the seat capacity of the aircraft.

The choice-based integrated airline schedule design and fleet assignment problem is when the demand variables s_{ih}, for all $i \in \mathcal{I}, h \in \mathcal{H}$, are considered endogenous to scheduling decisions and represented by a discrete choice model. Atasoy et al. (2014) proposed a mixed-integer nonlinear optimization problem for the integrated airline schedule design, fleet assignment, and pricing problem approximating the passenger spill and recapture effects. The itinerary choice model (i.e., a logit model) computes the market share of each itinerary based on the price, the travel time, the number of stops, and the departure time of the itinerary. Unfortunately, the highly nonlinear parts of this model pose computational issues and limit its implementation to small-scale instances. Later, Dong et al. (2016) proposed two mixed-integer models, one of which includes the itinerary price elasticity. The nonlinear constraints associated with demand (i.e., the constraints that ensure the market share of an itinerary is proportional to its utility) are approximated by two fractional constraints. The heuristic developed can scale to larger instances and outperforms the CPLEX MIP solver.

Cadarso et al. (2017) and Xu et al. (2023) tackled the choice-based integrated airline schedule design and fleet assignment problem with competition effects. In Cadarso et al. (2017), competition between airlines and a high-speed train new to the market is examined. The demand is estimated using a nested logit model, but the prices for all itineraries are considered fixed. The nonlinear function of the demand for an airline in terms of the frequency of the service is approximated by piecewise linear functions. The resulting mixed-integer linear optimization problem is solved using CPLEX to near-optimal solutions. The work of Xu et al. (2023) is an extension that further considers endogenous pricing and a comprehensive design approach (in opposition to the commonly used incremental approach). The problem is framed as a differentiated Bertrand game and an efficient algorithm combining column generation and large neighborhood search algorithms is proposed.

Birolini, Antunes, Cattaneo, Malighetti and Paleari (2021) proposed an efficient approach for the choice-based integrated airline schedule design and fleet assignment problem able to solve mid-size instances of hub-and-spoke networks using a regular MILP solver. The demand functions are linearized by using piecewise linear functions.

Table 2 summarizes the main attributes of the literature on the choice-based integrated airline schedule design and fleet assignment problem reviewed. The columns display the discrete choice model used (DCM), consideration of passenger spill and recapture (S/R), endogeneity of pricing (pricing probl.), if the demand functions are approximated, for example using piecewise linear functions (approx. demand), implementation of complete or incremental timetabling (Incr.), whether the solution method is exact or heuristic (heur.), as well as some attributes of the largest instance solved, namely the number of itineraries ($|\mathcal{I}|$), fleet types ($|\mathcal{F}|$), fare class ($|\mathcal{H}|$), and airports (# airp.).

Overall, we observe that when considering the price as a decision variable, the problem complexity increases and heuristic solution methods must be employed. Furthermore, all the works except the one of Atasoy et al. (2014) which is limited to small-scale cases use a numerical approximation method to address the tractability issues associated with the nonlinear part of the problem.

Authors	DCM		S/R	Pricing	Approx.	Incr.	Method		Instances			
	Logit	NeL		probl.	demand		Exact	Heur.	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{F}\rvert$	$\lvert\mathcal{H}\rvert$	# airp.
Atasoy et al. (2014)	•		•	•		•		•	36	3	2	3
Dong et al. (2016)	•			•	•	•		•	876	5	3	NA
Cadarso et al. (2017)		•			•	•	•		104	3	NA	23
Birolini, Antunes, Cattaneo, Malighetti and Paleari (2021)		•			•	•	•		2,170	2	NA	100
Xu et al. (2023)		•		•	•			•	110	14	2	56

Table 2: Main attributes of selected choice-based airline scheduling and fleet assignment contributions. NeL - nested logit model.

3.2.3 Urban transportation planning

Transportation planning traditionally encompasses public transportation planning and traffic flow optimization (e.g., through efficient road pricing schemes), the latter being complemented by the planning of emerging mobility services like car-sharing, electric bike-sharing, and ride-sourcing. In this section, we will first describe the key aspects of these three urban transportation planning components before going through relevant choice-based optimization models proposed in the literature.

First, public transport service planning involves the identification of routes to serve, the planning of trips along these routes, and the allocation of resources (both material and human) to cover the planned trips. In this context, the primary objective of public authorities is to provide a given population with the highest level of service, including service availability and quality, while adhering to budgetary constraints. This can be interpreted as aiming to maximize the user's utility, where the utility can be a function of spatial, temporal, and capacity availability, as well as safety, reliability, average travel time, average waiting time, and comfort, among other factors. The overall public transportation planning problem is too complex to be addressed as a whole and is therefore usually divided into sub-problems that are solved either sequentially or in an integrated manner. These sub-problems are:

1. *Network design*, which involves identifying the links to be integrated into the transportation network and aggregating these links to form coherent routes,
2. *Frequency setting and timetabling*, which entails determining the frequency of service on each route and establishing the exact schedule,
3. *Vehicle scheduling*, aimed at assigning vehicles to timetabled trips,
4. *Crew scheduling*, involving the establishment of anonymous driver workdays and personalized driver schedules that are valid for several weeks,
5. *Maintenance scheduling*, relating to the planning and management of maintenance tasks and downtime for vehicles and infrastructure,
6. *Real-time control*, which allows for adjustments to the operational plan during execution.

The network design as well as the frequency setting and timetabling sub-problems both rely on demand models, whether they are aggregated or disaggregated, to determine the demand on each route within a given network and route frequencies. Consequently, these problems can be formulated as choice-based optimization problems. On-demand (or demand-responsive) transportation (see, e.g., Nelson et al., 2010; Liu and Ceder, 2015; Vansteenwegen et al., 2022) can complement fixed-line and scheduled public transportation systems by serving low-density areas and feeding the traditional public transportation system. The planning of an on-demand transportation system involves determining the number of vehicles in the system, the location of stations (if using any), and the route (if using any). Subsequently, vehicles are dispatched to meet the real-time demand or requests placed a few hours in advance.

The second facet of transportation planning is optimizing traffic congestion. The primary objective of traffic congestion optimization is typically either reducing the average waiting time in traffic or lowering greenhouse gas emissions resulting from traffic. For many decades, the predominant approach to mitigating traffic congestion was expanding road capacity. However, it has been demonstrated that this approach, due to induced demand (Goodwin and Noland, 2003; Duranton and Turner, 2011), does not necessarily improve the situation and may even make it worse. Today, transportation demand management strategies aims at influencing the choice of travel mode, departure time, route, and even the decision to engage or not in an activity (thus necessitating travel). One of the most commonly applied measures involves the implementation of road pricing, often referred to as congestion pricing, network pricing, or toll pricing.

Finally, emerging mobility services, like bike-sharing (Fishman, 2016), car-sharing (Shaheen and Cohen, 2013), scooter-sharing, and ride-haling (Tirachini, 2020) have complemented traditional transportation systems in recent years. The planning decisions for such a service are fixing the number of vehicles in the system and determining the stations and their capacity (if using any).

The network pricing problem based on the choice of individuals has been studied in Gilbert et al. (2014a), Gilbert et al. (2015), and Gilbert et al. (2014b). A nonlinear logit network pricing problem is proposed in Gilbert et al. (2014a). Leveraging the model's analytical properties, new network simplification rules based on the notion of "cells" are proposed and a class of network typologies that is unimodal is identified. A heuristic yielding quasi-optimal solutions is proposed in the companion paper Gilbert et al. (2015). This heuristic combines MILP approximations (either by considering the deterministic problem or a piecewise linear function approximation of the demand) and local ascent methods (a line search and a trust region method are implemented). Gilbert et al. (2014b) extended these works by incorporating a mixture of logit choice model to accommodate for the nonuniform perception of travel time (or equivalently price) among the population. This new formulation is challenging to solve as the demand (i.e., the flow) on each arc of the network has no closed-form expression. The MILP approximation step of Gilbert et al. (2015) is therefore adapted to approximate path flows.

The works of Gilbert et al. (2014a), Gilbert et al. (2015), and Gilbert et al. (2014b) highlight the challenging nature of the choice-based network pricing problem, even when congestion is not taken into account (i.e., without incorporating equilibrium constraints). In Wu et al. (2012), an optimization problem with equilibrium constraints for the congestion pricing problem with credit schemes is proposed. Their model incorporates heterogeneous mode and itinerary choices, represented by a nested logit model. Income is factored into this choice model to capture the distributional impact of congestion on different income groups. Due to the numerical integration involved in their model, a derivative-free solution algorithm is employed. In the context of ride-sourcing, Li et al. (2021) addressed the spatial pricing problem with congestion charges. In this problem, the ride-sourcing platform sets ride prices, while a regulatory agency determines toll prices. The authors proposed a nonconvex network economic equilibrium model in which individuals' mode choices are determined by a logit model. The equilibrium constraints considered in this problem imply that travel time in a congested area depends on the number of ride-sourcing trips. They introduced a heuristic based on grid search and a dual decomposition algorithm, along with a tight upper bound to evaluate the performance of their solution approach.

In Robenek et al. (2016), a model addressing cyclic and non-cyclic train timetabling problems, taking into account passenger satisfaction through a deterministic passenger satisfaction function is proposed. The bi-objective problem is reformulated as a single-objective problem with an ϵ-constraint, where the main objective is the operators' profit, calculated as the revenues minus the operating costs. The ϵ-constraint ensures a minimum level of satisfaction. The challenge in this problem lies in routing passengers within the network and adjusting train capacities to meet capacity constraints. This work was further extended in Robenek et al. (2018), by considering stochastic and endogenous demand and passenger satisfaction. This extension allows for a more accurate representation of train occupation levels, but at the cost of nonlinear constraints to account for the logit choice model. In addition, the work of Robenek et al. (2018) incorporates a competing operator and the ticket pricing decisions to capture the elasticity in passenger demand. A simulated annealing heuristic is proposed for solving a real-world case study.

Recently, Liu et al. (2019) and Sharif-Azadeh et al. (2022) proposed models for the design of on-demand transportation systems. On the one hand, Liu et al. (2019) introduced a bilevel problem aimed at optimizing supply-side parameters, such as fleet size and the discount factor for shared trips, within an on-demand transportation system comprised of both private and shared vehicles with varying capacities. The outer loop of the bilevel problem optimizes these supply-side parameters while considering the transportation system's simulation as a black-box function. The inner loop updates service-specific attributes at traffic equilibrium by simulating the transportation system. They employed a Bayesian optimization approach to maximize the operator's profit and passenger surplus. On the other hand, Sharif-Azadeh et al. (2022) presented a MILP model for the integrated on-demand and fixed transportation scheduling problem. In this context, on-demand transportation has fixed pick-up and drop-off locations but flexible routes. The problem revolves around assigning stations to either the on-demand transportation network, the fixed transportation network, or both. This model does not account for congestion equilibrium and vehicle capacity constraints, and it assumes fixed fares. Consequently, the choice probabilities can be provided as input to the model and the expected ridership is computed using linear equations in the optimization model.

Table 3 summarizes the main attributes of the literature on choice-based urban planning problems reviewed. The columns display the discrete choice model used, the choices modeled by discrete choice models, consideration of traffic equilibrium constraints (Traffic equili.), consideration of vehicle capacity constraints, endogeneity of pricing (Pricing prob.), the type of problem (linear - L, non-linear - NL, integer - INT, and BL - bilevel), and the solution method employed (Sol. method).

Authors	Discrete choice model			Choice	Traffic equili.	Vehicle capacity	Pricing prob.	Type of problem				Sol. Method
	Logit	NeL	MoL					L	NL	INT	BL	
Wu et al. (2012)		•		mode, itinerary	•		•		•			DF
Gilbert et al. (2014a)	•			itinerary			•		•			H
Gilbert et al. (2014b)			•	itinerary			•		•			H
Robenek et al. (2018)	•			itinerary		•	•		•	•		H
Liu et al. (2019)	•			mode	•	•	•				•	BO
Li et al. (2021)	•			mode	•		•		•			H
Sharif-Azadeh et al. (2022)	•			mode				•		•		LP

Table 3: Main attributes of selected choice-based urban planning problems. NeL - nested logit model, MoL - mixture of logit model, L - linear, NL - nonlinear, INT - (mixed-)integer, BL - bilevel. Solution method: BO - Bayesian optimization, DF - derivative-free algorithm, H - heuristic, LP - linear optimization method.

3.3 General comments

The overview of these three applications, which involve increasing levels of complexity - the facility location problem, airline scheduling and fleet assignment, and urban transportation planning - provides insights into the challenges of incorporating individual choice into optimization problems. The facility location problem possesses a unique structure that facilitates the integration of discrete choice models. Specifically, its objective function is submodular under any random utility maximization model (including the MEV family of discrete choice models and the mixture of logit model) (Legault and Frejinger, 2022). However, this favorable structure is not present in the latter two problems. The airline scheduling and fleet assignment problem is more complex due to its inclusion of fleet capacity constraints and pricing components that interact with the demand functions. Finally, urban transportation planning problems can additionally involve traffic equilibrium constraints. In this case, the optimization problems transform into noncooperative games, making them even more challenging to solve.

We have seen that the implementation of the logit model in optimization problems is relatively straightforward due to its closed-form probabilistic expression, which offers an analytical form for modeling choices. Consequently, it is by far the most common option for modeling individuals' choices. Many optimization problems, including the facility location problem, can effectively integrate a homogeneous logit model by exploiting the problem structure in order to linearize the model. However, when dealing with more complex discrete choice models that lack an analytical form, their implementation in optimization problems becomes inherently challenging. This holds true even for problems with a desirable structure like the facility location problem. To address this complexity, methods such as simulation and numerical approximation must be employed, as detailed in Section 3.1.

4 Research perspectives

Choice-based optimization is a growing field that intersects operations research, transportation, logistics, and economics. As researchers continue to delve into this area, several promising research directions emerge, each with the potential to advance our understanding and practical applications of choice-based optimization. In this section, we outline some of these promising research avenues.

Assess the impact of complex discrete choice models. One research avenue involves assessing the potential gains of considering complex discrete choice models in choice-based optimization problems. While the logit model is commonly used due to its simplicity, more advanced models may offer improved accuracy in capturing decision behaviors. In this train of thought, some studies evaluate the impact of not considering certain parameters affected by demand as endogenous. For example, Lurkin et al. (2017) assessed the impact of considering exogenous versus endogenous price in the itinerary choice model of the airline network planning problem, highlighting the shortcomings of not accounting for price endogeneity. Similarly, Talluri and van Ryzin (2004) conducted such an analysis for airline revenue management problems. Extending these studies to compare objective value losses when employing the logit model or aggregate demand models instead of complex discrete choice models would provide valuable insights.

Moreover, conducting a comprehensive study across various applications to assess the added value of considering complex discrete choice models over simpler alternatives is crucial. This would provide evidence encouraging the development of methodological contributions to choice-based optimization using complex discrete choice models.

Exploring decomposition methods and simulation-based optimization. Continuing to explore the interplay between decomposition methods and simulation-based optimization is another promising direction. Various decomposition methods have been proposed in the context of choice-based optimization in the past, such as column generation (see, e.g., Bront et al., 2009; Szymula and Bešinović, 2020; Van Den Eeckhout et al., 2021), dynamic programming decomposition (see, e.g., Koch et al., 2017), Lagrangian decomposition (see, e.g., Pacheco et al., 2022), and Benders' decomposition (see, e.g., Yan et al., 2022). However, to the best of our knowledge, leveraging the power of decomposition methods to handle large-scale choice-based optimization problems solved by simulation-based optimization methods has only been proposed for Lagrangian and Benders decomposition in Pacheco et al. (2022) and Haering et al. (2023), respectively. This avenue is promising because simulation-based optimization is one of the two alternative methods to incorporate complex discrete choice models into choice-based optimization (as explained in Section 3.1). Still, it has the disadvantage of increasing solving time. Decomposition methods could help reduce this computing time, making it possible to solve large-scale optimization problems based on complex discrete choice models.

Modeling combined choices. In reality, individuals often face multiple combined choices that should not be treated independently. For example, the choice of transportation mode and the choice of trip itinerary are interdependent and often selected simultaneously by individuals. However, for reasons of tractability, these decisions are almost always considered sequentially and independently in the choice-based optimization literature. Addressing the modeling of such correlated choices is a promising research direction. Developing methods to capture and optimize interdependencies between choices could lead to more realistic and effective decision-support tools in various domains.

Integration of activity-based modeling and dynamic choice models. Taking choice-based optimization a step further, integrating activity-based modeling or dynamic choice models is an intriguing avenue. These approaches consider the dynamic nature of decision-making processes and account for individual preferences and constraints over time. Incorporating such models into choice-based optimization can enhance its applicability in dynamic and evolving environments.

Machine learning for choice-based optimization. Machine learning techniques have been increasingly adopted in choice modeling (Hillel et al., 2021). While much of the focus has been on demand estimation using machine learning (e.g., works by Ben-Elia and Shiftan, 2010; Chen et al., 2017; Ke et al., 2017; Dabiri and Heaslip, 2018; Lhéritier et al., 2019; Cheng et al., 2019; Yan et al., 2020; Zhao et al., 2020; Ke et al., 2021; Zhang and Zhao, 2022), another promising avenue is the integration of machine learning within choice-based optimization itself. This area remains relatively unexplored, with notable exceptions like the work of Clarke et al. (2023) in the context of simultaneous and non-cooperative games. Leveraging machine learning within the optimization process itself has the potential to reduce computing time and thus address the current scalability issues present in most choice-based optimization problems.

5 Conclusions

Fifty years ago, the analysis of the *supply-side* and the *demand-side* of transportation problems was conducted by two disconnected communities. In their PhD thesis on the so-called "transportation problem", Jakob Iterator from the Weizmann Institute described the *demand-side* in the first sentence of their manuscript as follows:

> In this research, we consider the L.P. transportation problem which can be formulated as follows. Let a_i be the quantity of a certain product available at an origin O_i and b_j be the quantity of the same product required at a destination D_j (Intrator, 1973).

The same year, Moshe Ben-Akiva, who first introduced the nested-logit model, referred in their PhD thesis to the *supply-side* using the general and particularly vague term "transport services":

> Decision making in transportation planning, as in any other planning activity, requires the prediction of impacts from proposed policies. One of the inputs to the prediction process is the demand function which describes consumers' expected usage of transport services (Ben-Akiva, 1973).

Jakob Iteraror and Moshe Ben-Akiva considered demand and supply as a simple inputs to their respective models. This practice was common at the time. Today, both communities acknowledge the need to leverage the remarkable developments made by the other community over the past 50 years. This is an exciting era where integrating operations research and discrete choice is recognized as essential, presenting intriguing research challenges. However, while choice-based optimization has received some attention in the fields of transportation and logistic over the past fifty years, we have observed that in three important optimization applications - namely, the facility location problem, the airline scheduling and fleet assignment problem, and the urban transportation planning problem - the literature on discrete choice models and choice-based optimization is still not harmoniously connected.

First, while several advanced discrete choice models have been proposed in the demand modeling literature, these models are almost non-existent in the choice-based optimization literature. The logit model remains by far the most common choice model in choice-based optimization, probably due to its closed-form probabilistic expression and the relatively straightforward estimation of its parameters. Second, extensions of discrete choice models (e.g., activity-based modeling and dynamic choice models) have not been implemented in the operations research literature to the best of our knowledge. Yet, such models are becoming increasingly prevalent in the choice modeling literature as they offer an alternative to modeling multiple decisions in isolation and independently.

To bridge this gap in the literature, we argued that more general solution approaches tailored to choice-based optimization (i.e., methodological contributions) are needed. Several potential methodological contributions, including exploring decomposition methods for simulation-based optimization, modeling correlated choices, and developing machine learning tools for choice-based optimization, have been identified. A new emphasis on this topic appears essential for the flourishing of behavioral models in transportation and logistics.

References

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey, *Journal of Econometrics* **156**(1): 38–67. Structural Models of Optimization Behavior in Labor, Aging, and Health.

Ahipaşaoğlu, S. D., Arıkan, U. and Natarajan, K. (2019). Distributionally robust markovian traffic equilibrium, *Transportation Science* **53**(6): 1546–1562.

Appa, G. (1973). The transportation problem and its variants, *Journal of the Operational Research Society* **24**: 79–99.

Aros-Vera, F., Marianov, V. and Mitchell, J. E. (2013). p-hub approach for the optimal park-and-ride facility location problem, *European Journal of Operational Research* **226**(2): 277–285.

Atasoy, B., Salani, M. and Bierlaire, M. (2014). An integrated airline scheduling, fleeting, and pricing model for a monopolized market, *Computer-Aided Civil and Infrastructure Engineering* **29**(2): 76–90.

Ben-Akiva, M. E. (1973). *Structure of passenger travel demand models*, PhD thesis, Massachusetts Institute of Technology.

Ben-Akiva, M. E. and Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*, Vol. 9, MIT press.

Ben-Elia, E. and Shiftan, Y. (2010). Which road do I take? A learning-based model of route-choice behavior with real-time information, *Transportation Research Part A: Policy and Practice* **44**(4): 249–264.

Benati, S. (1997). Submodularity in competitive location problems, *Ricerca Operativa* **26**: 3–34.

Benati, S. (1999). The maximum capture problem with heterogeneous customers, *Computers & Operations Research* **26**(14): 1351–1367.

Benati, S. and Hansen, P. (2002). The maximum capture problem with random utilities: Problem formulation and algorithms, *European Journal of Operational Research* **143**(3): 518–530.

Bernstein, F. and Federgruen, A. (2004). Dynamic inventory and pricing models for competing retailers, *Naval Research Logistics (NRL)* **51**(2): 258–274.

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model, *Annals of Operations Research* **144**(1): 287–300.

Bierlaire, M., Frejinger, E. and Hillel, T. (2021). Dynamic choice models, *Technical Report TRANSP-OR 210305*, Transport and Mobility Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Binder, S., Maknoon, Y. and Bierlaire, M. (2017). The multi-objective railway timetable rescheduling problem, *Transportation Research Part C: Emerging Technologies* **78**: 78–94.

Birolini, S., Antunes, A. P., Cattaneo, M., Malighetti, P. and Paleari, S. (2021). Integrated flight scheduling and fleet assignment with improved supply-demand interactions, *Transportation Research Part B: Methodological* **149**: 162–180.

Birolini, S., Jacquillat, A., Cattaneo, M. and Antunes, A. P. (2021). Airline network planning: Mixed-integer non-convex optimization with demand–supply interactions, *Transportation Research Part B: Methodological* **154**: 100–124.

Birolini, S., Jacquillat, A., Schmedeman, P. and Ribeiro, N. (2023). Passenger-centric slot allocation at schedule-coordinated airports, *Transportation Science* **57**(1): 4–26.

Blanchet, J., Gallego, G. and Goyal, V. (2016). A markov chain approximation to choice modeling, *Operations Research* **64**(4): 886–905.

Bortolomiol, S., Lurkin, V. and Bierlaire, M. (2021). A simulation-based heuristic to find approximate equilibria with disaggregate demand models, *Transportation Science* **55**(5): 1025–1045.

Bront, J. J. M., Méndez-Díaz, I. and Vulcano, G. (2009). A column generation algorithm for choice-based network revenue management, *Operations Research* **57**(3): 769–784.

Cadarso, L., Vaze, V., Barnhart, C. and Marín, A. (2017). Integrated airline scheduling: Considering competition effects and the entry of the high speed rail, *Transportation Science* **51**(1): 132–154.

Chen, X. M., Zahiri, M. and Zhang, S. (2017). Understanding ridesplitting behavior of on-demand ride services: An ensemble learning approach, *Transportation Research Part C: Emerging Technologies* **76**: 51–70.

Cheng, L., Chen, X., De Vos, J., Lai, X. and Witlox, F. (2019). Applying a random forest method approach to model travel mode choice behavior, *Travel Behaviour and Society* **14**: 1–10.

Chiong, K. X., Galichon, A. and Shum, M. (2016). Duality in dynamic discrete-choice models, *Quantitative Economics* **7**(1): 83–115.

Cirillo, C. and Xu, R. (2011). Dynamic discrete choice models for transportation, *Transport Reviews* **31**(4): 473–494.

Clarke, S., Dragotto, G., Fisac, J. F. and Stellato, B. (2023). Learning rationality in potential games, *arXiv preprint arXiv:2303.11188* .

Dabiri, S. and Heaslip, K. (2018). Inferring transportation modes from GPS trajectories using a convolutional neural network, *Transportation Research Part C: Emerging Technologies* **86**: 360–371.

Dam, T. T., Ta, T. A. and Mai, T. (2022). Submodularity and local search approaches for maximum capture problems under generalized extreme value models, *European Journal of Operational Research* **300**(3): 953–965.

Dong, Z., Chuhang, Y. and Lau, H. H. (2016). An integrated flight scheduling and fleet assignment method based on a discrete choice model, *Computers & Industrial Engineering* **98**: 195–210.

Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities, *American Economic Review* **101**(6): 2616–52.

Feldman, J. B. and Topaloglu, H. (2017). Revenue management under the markov chain choice model, *Operations Research* **65**(5): 1322–1342.

Fishman, E. (2016). Bikeshare: A review of recent literature, *Transport Reviews* **36**(1): 92–113. Cycling as Transport.

Freire, A. S., Moreno, E. and Yushimito, W. F. (2016). A branch-and-bound algorithm for the maximum capture problem with random utilities, *European Journal of Operational Research* **252**(1): 204–212.

Gallego, G., Huh, W. T., Kang, W. and Phillips, R. (2006). Price competition with the attraction demand model: Existence of unique equilibrium and its stability, *Manufacturing & Service Operations Management* **8**(4): 359–375.

Gallego, G. and Wang, R. (2014). Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities, *Operations Research* **62**(2): 450–461.

Gilbert, F., Marcotte, P. and Savard, G. (2014a). Logit network pricing, *Computers & Operations Research* **41**: 291–298.

Gilbert, F., Marcotte, P. and Savard, G. (2014b). Mixed-logit network pricing, *Computational Optimization and Applications* **57**: 105–127.

Gilbert, F., Marcotte, P. and Savard, G. (2015). A numerical study of the logit network pricing problem, *Transportation Science* **49**(3): 706–719.

Goodwin, P. and Noland, R. B. (2003). Building new roads really does create extra traffic: A response to Prakash et al., *Applied Economics* **35**(13): 1451–1457.

Gosavi, A. et al. (2015). *Simulation-based optimization*, Springer.

Gupta, S., Seshadri, R., Atasoy, B., Prakash, A. A., Pereira, F., Tan, G. and Ben-Akiva, M. (2020). Real-time predictive control strategy optimization, *Transportation Research Record* **2674**(3): 1–11.

Haase, K. (2009). Discrete location planning.

Haase, K. and Müller, S. (2013). Management of school locations allowing for free school choice, *Omega* **41**(5): 847–855.

Haase, K. and Müller, S. (2014). A comparison of linear reformulations for multinomial logit choice probabilities in facility location models, *European Journal of Operational Research* **232**(3): 689–691.

Haering, T., Legault, R., Torres, F., Ljubic, I. and Bierlaire, M. (2023). Exact algorithms for continuous pricing with advanced discrete choice demand models, *Technical Report TRANSP-OR 231211*, Transport and Mobility Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Hane, C. A., Barnhart, C., Johnson, E. L., Marsten, R. E., Nemhauser, G. L. and Sigismondi, G. (1995). The fleet assignment problem: Solving a large-scale integer program, *Mathematical Programming* **70**: 211–232.

Hillel, T., Bierlaire, M., Elshafie, M. Z. and Jin, Y. (2021). A systematic review of machine learning classification methodologies for modelling passenger mode choice, *Journal of Choice Modelling* **38**: 100221.

Intrator, J. (1973). *Transportation problem*, PhD thesis, The Weizmann Institute of Science.

Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z. and Ye, J. (2021). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network, *Transportation Research Part C: Emerging Technologies* **122**: 102858.

Ke, J., Zheng, H., Yang, H. and Chen, X. M. (2017). Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach, *Transportation Research Part C: Emerging Technologies* **85**: 591–608.

Koch, S., Gönsch, J. and Steinhardt, C. (2017). Dynamic programming decomposition for choice-based revenue management with flexible products, *Transportation Science* **51**(4): 1046–1062.

Kök, A. G. and Xu, Y. (2011). Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models, *Management Science* **57**(9): 1546–1563.

Legault, R. and Frejinger, E. (2022). A model-free approach for solving choice-based competitive facility location problems using simulation and submodularity, *arXiv preprint arXiv:2203.11329* .

Lhéritier, A., Bocamazo, M., Delahaye, T. and Acuna-Agost, R. (2019). Airline itinerary choice modeling using machine learning, *Journal of Choice Modelling* **31**: 198–209.

Li, H. and Huh, W. T. (2011). Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications, *Manufacturing & Service Operations Management* **13**(4): 549–563.

Li, S., Yang, H., Poolla, K. and Varaiya, P. (2021). Spatial pricing in ride-sourcing markets under a congestion charge, *Transportation Research Part B: Methodological* **152**: 18–45.

Liu, T. and Ceder, A. A. (2015). Analysis of a new public-transport-service concept: Customized bus in China, *Transport Policy* **39**: 63–76.

Liu, Y., Bansal, P., Daziano, R. and Samaranayake, S. (2019). A framework to integrate mode choice in the design of mobility-on-demand systems, *Transportation Research Part C: Emerging Technologies* **105**: 648–665.

Ljubić, I. and Moreno, E. (2018). Outer approximation and submodular cuts for maximum capture facility location problems with random utilities, *European Journal of Operational Research* **266**(1): 46–56.

Lu, A., Aievoli, S., Ackroyd, J., Carlin, C. and Reddy, A. (2009). Passenger environment survey: Representing the customer perspective in quality control, *Transportation Research Record* **2112**(1): 93–103.

Lurkin, V., Garrow, L. A., Higgins, M. J., Newman, J. P. and Schyns, M. (2017). Accounting for price endogeneity in airline itinerary choice models: An application to continental U.S. markets, *Transportation Research Part A: Policy and Practice* **100**: 228–246.

Ma, W., Zeng, L. and An, K. (2023). Dynamic vehicle routing problem for flexible buses considering stochastic requests, *Transportation Research Part C: Emerging Technologies* **148**: 104030.

Mai, T. and Lodi, A. (2020). A multicut outer-approximation approach for competitive facility location under random utilities, *European Journal of Operational Research* **284**(3): 874–881.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior, *Frontier in Econometrics* .

McFadden, D. (1978). Modelling the choice of residential location, *in* A. Karlquist *et al.* (ed.), *Spatial interaction theory and residential location*, North-Holland, Amsterdam, pp. 75–96.

McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response, *Journal of Applied Econometrics* **15**(5): 447–470.

McNally, M. G. (2007). The four step model, *in* D. A. Hensher and K. J. Button (eds), *Handbook of Transport Modelling*, Vol. 1, Emerald Group Publishing Limited, pp. 33–53.

Milgrom, P. and Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities, *Econometrica* **58**(6): 1255–1277.

Méndez-Vogel, G., Marianov, V. and Lüer-Villagra, A. (2023). The follower competitive facility location problem under the nested logit choice rule, *European Journal of Operational Research* **310**(2): 834–846.

Müller, S., Haase, K. and Kless, S. (2009). A multiperiod school location planning approach with free school choice, *Environment and Planning A: Economy and Space* **41**(12): 2929–2945.

Nelson, J. D., Wright, S., Masson, B., Ambrosino, G. and Naniopoulos, A. (2010). Recent developments in flexible transport services, *Research in Transportation Economics* **29**(1): 243–248.

Osorio, C. and Atasoy, B. (2021). Efficient simulation-based toll optimization for large-scale networks, *Transportation Science* **55**(5): 1010–1024.

Pacheco, M., Bierlaire, M., Gendron, B. and Sharif-Azadeh, S. (2021). Integrating advanced discrete choice models in mixed integer linear optimization, *Transportation Research Part B: Methodological* **146**: 26–49.

Pacheco, M., Gendron, B., Sharif-Azadeh, S. and Bierlaire, M. (2022). A lagrangian decomposition scheme for choice-based optimization, *Computers & Operations Research* **148**: 105985.

Pacheco Paneque, M. (2020). *A general framework for the integration of complex choice models into mixed integer optimization*, PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland.

Papola, A. (2004). Some developments on the cross-nested logit model, *Transportation Research Part B: Methodological* **38**(9): 833–851.

Polinder, G.-J., Cacchiani, V., Schmidt, M. and Huisman, D. (2022). An iterative heuristic for passenger-centric train timetabling with integrated adaption times, *Computers & Operations Research* **142**: 105740.

Pougala, J., Hillel, T. and Bierlaire, M. (2023). OASIS: Optimisation-based activity scheduling with integrated simultaneous choice dimensions, *Transportation Research Part C: Emerging Technologies* **155**: 104291.

Rasouli, S. and Timmermans, H. (2014). Activity-based models of travel demand: promises, progress and prospects, *International Journal of Urban Sciences* **18**(1): 31–60.

Robenek, T., Maknoon, Y., Sharif-Azadeh, S., Chen, J. and Bierlaire, M. (2016). Passenger centric train timetabling problem, *Transportation Research Part B: Methodological* **89**: 107–126.

Robenek, T., Sharif-Azadeh, S., Maknoon, Y., de Lapparent, M. and Bierlaire, M. (2018). Train timetable design under elastic passenger demand, *Transportation Research Part B: Methodological* **111**: 19–38.

Shaheen, S. A. and Cohen, A. P. (2013). Carsharing and personal vehicle services: Worldwide market developments and emerging trends, *International Journal of Sustainable Transportation* **7**(1): 5–34.

Sharif-Azadeh, S., Atasoy, B., Ben-Akiva, M. E., Bierlaire, M. and Maknoon, Y. (2022). Choice-driven dial-a-ride problem for demand responsive mobility service, *Transportation Research Part B: Methodological* **161**: 128–149.

Small, K. A. (1987). A discrete choice model for ordered alternatives, *Econometrica* **55**(2): 409–424.

Szymula, C. and Bešinović, N. (2020). Passenger-centered vulnerability assessment of railway networks, *Transportation Research Part B: Methodological* **136**: 30–61.

Talluri, K. and van Ryzin, G. (2004). Revenue management under a general discrete choice model of consumer behavior, *Management Science* **50**(1): 15–33.

Talvitie, A. (1973). A direct demand model for downtown work trips, *Transportation* **2**(2): 121–152.

Tirachini, A. (2020). Ride-hailing, travel behaviour and sustainable mobility: an international review, *Transportation* **47**(4): 2011–2047.

Tiwari, R., Jayaswal, S. and Sinha, A. (2021). Alternate solution approaches for competitive hub location problems, *European Journal of Operational Research* **290**(1): 68–80.

Tong, L. C., Zhou, L., Liu, J. and Zhou, X. (2017). Customized bus service design for jointly optimizing passenger-to-vehicle assignment and vehicle routing, *Transportation Research Part C: Emerging Technologies* **85**: 451–475.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*, 2 edn, Cambridge University Press.

Tu, W., Li, Q., Fang, Z., lung Shaw, S., Zhou, B. and Chang, X. (2016). Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach, *Transportation Research Part C: Emerging Technologies* **65**: 172–189.

Van Den Eeckhout, M., Vanhoucke, M. and Maenhout, B. (2021). A column generation-based diving heuristic to solve the multi-project personnel staffing problem with calendar constraints and resource sharing, *Computers & Operations Research* **128**: 105163.

Vansteenwegen, P., Melis, L., Aktaş, D., Montenegro, B. D. G., Sartori Vieira, F. and Sörensen, K. (2022). A survey on demand-responsive public bus systems, *Transportation Research Part C: Emerging Technologies* **137**: 103573.

Vovsha, P. (1997). Application of cross-nested logit model to mode choice in Tel Aviv, Israel, metropolitan area, *Transportation Research Record* **1607**(1): 6–15.

Wang, P., Zhu, Y. and Corman, F. (2022). Passenger-centric periodic timetable adjustment problem for the utilization of regenerative energy, *Computers & Industrial Engineering* **172**: 108578.

Wei, K., Vaze, V. and Jacquillat, A. (2020). Airline timetable development and fleet assignment incorporating passenger choice, *Transportation Science* **54**(1): 139–163.

Wu, D., Yin, Y., Lawphongpanich, S. and Yang, H. (2012). Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks, *Transportation Research Part B: Methodological* **46**(9): 1273–1287.

Xu, Y., Adler, N., Wandelt, S. and Sun, X. (2023). Competitive integrated airline schedule design and fleet assignment, *European Journal of Operational Research* .

Yan, C., Barnhart, C. and Vaze, V. (2022). Choice-based airline schedule design and fleet assignment: A decomposition approach, *Transportation Science* **56**(6): 1410–1431.

Yan, X., Liu, X. and Zhao, X. (2020). Using machine learning for direct demand modeling of ridesourcing services in Chicago, *Journal of Transport Geography* **83**: 102661.

Yin, J., Tang, T., Yang, L., Gao, Z. and Ran, B. (2016). Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: An approximate dynamic programming approach, *Transportation Research Part B: Methodological* **91**: 178–210.

Yin, J., Yang, L., Tang, T., Gao, Z. and Ran, B. (2017). Dynamic passenger demand oriented metro train scheduling with energy-efficiency and waiting time minimization: Mixed-integer linear programming approaches, *Transportation Research Part B: Methodological* **97**: 182–213.

Zhang, X. and Zhao, X. (2022). Machine learning approach for spatial modeling of ridesourcing demand, *Journal of Transport Geography* **100**: 103310.

Zhang, Y., Berman, O. and Verter, V. (2012). The impact of client choice on preventive healthcare facility network design, *OR spectrum* **34**: 349–370.

Zhao, X., Yan, X., Yu, A. and Van Hentenryck, P. (2020). Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models, *Travel Behaviour and Society* **20**: 22–35.