NORTHWESTERN UNIVERSITY

Development and Application of a Dynamic Solution Framework for Urban Air Taxi Fleet Operation

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Civil and Environmental Engineering
Transportation Systems Analysis and Planning

By
Haleh sadat Ale-Ahmad

EVANSTON, ILLINOIS

March 2022

ABSTRACT

Development and Application of a Dynamic Solution Framework for Urban Air Taxi Fleet Operation

Haleh sadat Ale-Ahmad

Urban Air Taxi (UAT) is the use case of passenger-carrying Urban Air Mobility (UAM) at its mature state, and it offers a ubiquitous on-demand (or nearly on-demand) per-seat service that moves passengers in urban or suburban areas using groundbreaking aircraft. However, the absence of a dominant electric vertical take-off and landing (eVTOL) aircraft technology and UAT operator feeds the uncertainty around UAT. This dissertation focuses on outlining the concept of operations for UAT services, defining the UAT problem, and developing and applying a dynamic solution framework to address the stochastic and dynamic problem of UAT fleet operation. As a result, it provides the UAT operator with a decision-making tool to achieve higher network efficiency.

To accomplish this goal, the UAT concept of operations, which involves a ubiquitous service with air pooling and elimination of short repositioning flights, is first outlined. Subsequently, the entities relevant to the UAT fleet operation are specified, and their associated states and events are presented in detail. The dynamic and stochastic problem of UAT fleet operation is modeled on a rolling horizon basis. A static and deterministic problem is solved at each decision epoch to help the UAT operator make the dynamic operational decisions, including acceptance and rejection of requests, routing and scheduling the aerial fleet, and assigning the requests to flights. Based on a node-based representation of the UAT network, the snapshot problem is modeled as a *Capacitated Location-Allocation-Routing Problem with Time Windows*

and Short Repositioning Elimination (CLARPTW-SRE). For narrow time windows and relatively short minimum distance for repositioning flight legs, the corresponding MIP could be solved quickly using commercial software, enabling its real-time application.

The proposed dynamic solution framework is subsequently implemented using a discrete-event simulation. The impacts of various exogenous and design parameters on demand consolidation are examined using comprehensive sensitivity analyses in a synthetic network. Furthermore, the framework is applied to the Chicago network using a fixed fleet of UAT aircraft and Chicago Transportation Network Providers (TNPs) demand. Augmenting the devised UAT operational strategy with real-world data would validate the network efficiency assumptions (e.g., the average load factor and utilization) made by many UAM market studies and offer estimates of the said parameters for future studies.

"This is by the Grace of my Lord." (27:40)

To Mamanjoon,

Maman, Baba, Laleh,

and Mostafa,

without your sacrifices, love, and support, I could not have done this.

And to Ala,

you taught me to love the way I never knew before.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation

In January of 1951, a *personal helicopter*, which was "big enough to carry two people and small enough to land on your lawn," was featured on the cover of Popular Mechanics [1]. Since then, numerous terms have been used to describe air transportation in metropolitan areas, including *helicopter air carrier*, *air taxi*, *On-Demand Aviation (ODA)*, *On-Demand Mobility (ODM)*, *on-demand urban air transportation*, and most recently, *Urban Air Mobility (UAM)* [2].

The first attempts to provide on-demand air transportation date back to the 1960s [2], where Air General offered on-demand service in Boston with as few as 30 minutes reservation windows using over 70 helicopters. In the early 2000s, a Small Aircraft Transportation System (SATS) research project carried out jointly by the Federal Aviation Administration (FAA) and the National Aeronautics and Space Administration (NASA) advocated on-demand regional services between cities using Very Light Jets (VLJ) [2]. This idea drew a lot of interest [3], which subsequently subsided in part due to the great recession in 2008 [2].

In recent years, with the vision of eco-friendly autonomous aircraft equipped with electric propulsion (which enables a 10x reduction in energy costs [4]) and efficient batteries with short charging or swapping time, the interest in urban air transportation has resurfaced. Compared to a helicopter, electric vertical take-off and landing (eVTOL) aircraft are 4 times quieter (with Joby claiming ~100 times quieter [5]) and 10 times less expensive [6]. Benefitting from this revolutionary aircraft technology, the Advanced Air Mobility (AAM) [7] initiative is pursuing to transfer cargo and passengers between urban, local, regional, and intraregional areas, while the UAM market focuses on carrying passengers and goods within metropolitan areas [7-9]. Urban

Air Taxi (UAT), a subset of UAM, is a ubiquitous on-demand per-seat service that transfers passengers in urban or suburban areas using groundbreaking aircraft [8,9]. UAT does not have fixed routes and regular schedules, distinguishing it from *air metro* [9] or *airport shuttle* [8], which are envisioned to operate on predetermined routes.

As of February 2020, 110 UAM[1] city projects were in progress worldwide [6]. More than 250 businesses were involved in UAM as of August 2020 [10]. The ride-hailing company Uber [11], major aircraft manufacturers Airbus [12] and Boeing [13], car manufacturer Hyundai [14], and start-ups Lilium [15], Volocopter [16], Kitty Hawk [17], and Joby Aviation [18] have shown considerable interest in passenger AAM. In the first half of 2020, USD 907 million was invested in UAM start-ups, which is nearly 20 times the amount invested in the whole year of 2016 [19].

UAM would be an attractive mode of transportation if it could deliver the critical promises of being the fastest mode while safe and enjoyable and having reasonable prices while offering multi-modal service with seamless transfers [20]. Therefore, the travel time saving is a critical variable in choosing UAM. Using UAT could reduce travel time over shorter distances, where the ground network is congested or the travel time is not reliable. However, for longer trips, the difference between aerial speed and ground speed plays a significant role. In 2017, the American Community Survey (ACS) estimated the number of workers to be nearly 152.8 million, out of which 85.3% drive alone and 38.3% (or 58.5 million) have commuting time greater than 30 minutes [21]. Moreover, the 2006-2010 ACS suggests that in the top 10 mega counties, commuters have a mean travel time ranging from 102 to 116 minutes and mean travel distance between 59 and 91 miles [22], and therefore, could significantly benefit from AAM.

[1] The UAM definition in Porsche Consulting report includes city-to-city trips as well.

However, there are numerous barriers and challenges in launching UAM as a mass-scale transit system [23], including:

- The regulation and certification process from the U.S. Federal Aviation Administration (FAA) and European Aviation Safety Agency (EASA) to fly large numbers of VTOL aircraft in urban areas;
- The development of reliable aircraft (especially in inclement weather) with efficient performance for commercial use;
- The advancement of battery technology to accommodate long-distance rides;
- The conflict resolution between UAM and Unmanned Aerial Vehicle (UAV) specified and operated by Air Traffic Control (ATC);
- Providing the service at an affordable cost;
- Safety [24,25],
- Noise [26],
- Life-cycle emissions, and
- Required infrastructure.

In a UAM market study conducted by Booz Allen Hamilton [8], under an entirely unconstrained (i.e., best case) scenario, the total available market value of airport shuttles and air taxis in the U.S. is estimated to be USD 500 billion with 11 million daily trips, which corresponds to 20% of the daily work trips across the U.S. However, willingness-to-pay, availability of the infrastructure and their capacity, adverse weather, and limited operation hours could reduce this market to 55,000 daily trips (i.e., 0.1% of total daily work trips in the U.S.) with an estimated USD

2.5 billion market value in the near term. The number of UAM aircraft drops from 850,000 in the unconstrained scenario to 4,100 in the constrained scenario.

Meanwhile, the most recent study by Roland Berger [19] estimates that by 2050, the revenue generated by the passenger UAM industry worldwide will be USD 90 billion a year with 160,000 passenger UAM aircraft, a significant growth from the projected USD 1 billion in 2030. To put these numbers into perspective, the total revenues of the global commercial airline market in 2019 were USD 840 billion, and the global taxi market is estimated to be USD 300 billion in 2030.

Passenger UAM is projected to grow at a compound annual growth rate (CAGR) of 45.9% by 2040 as Frost and Sullivan estimate [27], or 35% by 2035 (starting from 2025) as Porsche Consulting forecasts [6]. Factors such as cost, travel time savings, transfers and stops, safety, and noise are influential in adopting passenger UAM [8]. Booz Allen Hamilton's UAM market study [8] argues that high network efficiency, including high aircraft utilization and load factor[2], could increase the UAM demand by more than 200% compared to their base scenario. Uber suggests that ridesharing economics is one of the three critical steps towards lowering costs [28]. Uber's argument is consistent with the Crown Consulting UAM market study [9] commissioned by NASA, which forecasts that the on-demand point-to-point air taxi market will not be profitable by 2030, stating that the assumption of one passenger per trip is one of the main barriers.

Consequently, choosing the right UAM business model is crucial to the success of industry players [19]. Aside from the UAM market studies [8,9], ConOps [7], and OpsCons [29] provided by NASA and FAA, multiple players from the industry have offered their visions on passenger

[2] Capacity utilization of an aircraft

AAM and UAM [23,30-32]. Patterson et al. [33] summarize the numerous proposed visions in five on-demand or near-on-demand passenger UAM missions: private operation, air taxi, air pooling, semi-scheduled commuter, and scheduled commuter. The lack of consensus on the vision of passenger UAM operations and the absence of a single player in the industry has provided an opportunity to coordinate the UAM research according to the industry's needs [9]. For instance, Boeing plans to address the challenges of UAM operations by "modeling and simulating multiple end-state operational scenarios" [30].

Similar calls have echoed in the UAM research community. Following the review of the recent research and developments in UAM, Straubinger et al. [34] maintain that more advanced passenger pooling and aircraft dispatching models are required. Rajendran and Srinivas [35] review the developments of passenger UAM and the future challenges. They argue that the dynamic routing of air taxis and integration of ground and air transportation scheduling is underexplored and call for more research in these areas. Garrow et al. [36] present a systematic review of the UAM literature and conclude that most studies in the literature offer a deterministic framework for dispatching and scheduling algorithms. Consequently, they call for algorithms that could be implemented online or use a rolling horizon framework to address the uncertainties encountered in the UAM operation.

To address the gap in the literature, this dissertation focuses on developing a solution framework for the stochastic and dynamic problem of ubiquitous on-demand per-seat passenger UAT with air pooling. Among the 110 ongoing UAM and AAM city projects worldwide [6], the sheer number of potential UAM infrastructures in some cities, such as Dubai, would accelerate the UAM implementation. Consequently, such cities could see significant and immediate benefits from the UAT operations solution framework devised in this study. Nonetheless, this methodology

sets a benchmark for other use cases of passenger UAM and other city projects in the planning process. Furthermore, augmenting the devised UAT operational strategy with real-world data would validate the network efficiency assumptions (e.g., the average load factor and utilization) made by many UAM market studies or offer estimates of said parameters for future studies.

Ultimately, this research aims to provide a tool for the researchers to examine various concepts of operations and evaluate various operational strategies such as sharing schemes. The outcomes of such studies could be valuable for the industry players and the regulators, and therefore, could pave the way for future collaborations.

1.2 Problem Statement

UAT evolves around travelers who utilize their smartphones to request a ride within a city or a city and its suburbs. The service is on-demand and per seat. The requests are immediate or provide short notice, and the users expect to be served within a couple of minutes from their requested time. A fleet of eVTOL aircraft and ground-based ride-hailing service provided by a centralized UAT operator is available on-demand to cover the requests at competitive costs. The UAT operator offers a multi-modal service and synchronizes its aerial and ground fleet to serve the requests.

The UAT operator manages a fleet of homogenous vehicles and UAT aircraft in a ubiquitous network, and it synchronizes these two modes to serve the customers. Each request is identified by the origin, destination, requested time to begin the service, and group size. Given that the UAT pads are ubiquitously present in a ubiquitous network, the origin and destination of the request correspond to the desired take-off and landing sites, respectively.

Each passenger group is flexible in their pick-up and drop-off UAT pads selection and could be relocated on the ground within a reasonable radius from their origin or destination. This allows the UAT operator to eliminate the short repositioning flight legs in the ubiquitous network by relocating the requests over short distances. Additionally, the customers are willing to share a UAT aircraft with other passengers, and consequently, the UAT operator could relocate the passengers to consolidate customer requests and increase the aircraft load factor. As a result of short repositioning elimination and demand consolidation, the passenger trip consists of, at most, three legs: two ground-based legs and an aerial leg. The aerial leg is non-stop, and there is no intermediate transfer point for passengers to change the UAT aircraft after boarding.

The UAT operator is unaware of all future requests at the beginning of the planning horizon, and the customer requests for service arrive in real-time. As a result, the UAT operator updates its dispatching plan through a sequential decision-making process. The UAT operator may have a relatively short period for decision-making, particularly if there is no reservation scheme and requests are expected to be served immediately. The acceptance or rejection decision of the arrived requests is made at the first decision epoch after their arrival and will not change in the subsequent decision epochs. In other words, while considering accepting a new request, the UAT operator cannot reject the requests accepted in the previous decision epochs. However, the flight legs assigned to the accepted requests (and therefore, the pick-up and drop-off UAT pads) could change as long as the customers did not leave their origin. After leaving the origin for the pick-up UAT pad, the pick-up UAT pad of the request is fixed, and its boarding time could be rescheduled.

The UAT competitive advantage is the travel time savings. As a result, if the UAT operator chooses to serve a request, the trip delay (i.e., deviation of the request's total trip time from its desired trip time) cannot exceed a prespecified value, which in turn limits the wait time for the

aerial service, the ingress and egress time, and the deviation from the shortest flight. Additionally, the UAT operator determines when the customers should leave their origin to access the pick-up UAT pad. The scheduling is such that the customer wait time is mainly incurred at the origin rather than the pick-up UAT pad. However, some customers may wait at the departure gate after a schedule change resulting from the new information (e.g., the arrival of new requests). Since the assigned flight legs and their schedules could be updated multiple times, the requests are only provided, in advance, with the time window during which they will be prompted to leave their origin or board the aircraft, not the exact time.

The revenue that the UAT operator earns from serving a request is proportional to the distance between the origin and destination of the request and its group size. The UAT operator incurs a fixed cost per flight and a variable cost proportional to the aircraft mileage. Ultimately, the UAT operator seeks a strategy that maximizes its net profit given the capacity, delay, and synchronization constraints. This strategy should address request acceptance and rejection, allocation of accepted requests to flight legs, and the sequence that UAT aircraft should serve these flight legs. It should further handle the flight scheduling, the boarding time of each passenger group, and the time by which the passenger groups need to leave their origin.

1.3 Contributions

Aiming to provide insights on the UAT fleet operations and its potential in terms of travel timing savings in the urban setting, this dissertation makes the following contributions to the passenger UAM literature:

i. This research presents a concept of operations for UAT fleet operation in its mature state. The proposed concept of operations involves demand consolidation and elimination of short repositioning flight legs while providing a prespecified level of

service. To this end, it proposes the concept of flexible pads for UAT service design, aiming to increase the aircraft load factor (i.e., capacity utilization) and decrease the operating costs.

ii. This research models and solves an optimization problem with demand consolidation, elimination of short repositioning legs, synchronized logistics, and a guaranteed level of service.

iii. This research offers a dynamic solution framework that utilizes the optimization model for dynamic and stochastic UAT fleet operations, providing the UAT operator with a decision-making tool.

iv. This research develops a discrete-event simulation framework with sequential decision-making problems for UAT operation as a proof of concept. This simulation framework provides a tool that could be enhanced with other modules and models (e.g., demand, pricing, air traffic control) to further examine UAT operation.

v. Using a synthetic network and sensitivity analyses, this research evaluates the impact of technological factors (such as aircraft cruising speed) and strategic and operational decisions (such as demand consolidation strategy and the guaranteed level of service) on UAT fleet operation.

vi. This research provides insights on the factors associated with network efficiency, such as aircraft utilization and average load factor, using real-world data in Chicago.

1.4 Organization

Chapter 1 presents the motivation for this dissertation, states the problem, and outlines the dissertation's contributions. Chapter 2 provides more details on the concept of Urban Air Mobility,

recent developments, and the use cases suggested in the literature. Chapter 3 describes the proposed concept of operations for UAT in its mature state and outlines the operational assumptions used in this research. Chapter 4 presents a literature review relevant to the proposed concept of operations for the urban air taxi problem. Chapter 5 describes the modeling components of the UAT operational problem. Chapter 6 specifies the dynamic solution framework for the dynamic and stochastic UAT fleet operation. Chapter 7 discusses the network definition for the *Capacitated Location-Allocation-Routing Problem with Time Windows And Short Repositioning Elimination (CLARPTW-SRE)* presented in Chapter 8. Chapter 9 provides the design of numerical experiments, sensitivity analyses, and the outcomes for a synthetic network, while Chapter 10 presents the numerical experiments and results using the real-world data of the Chicago network. Ultimately, Chapter 11 concludes the dissertation with a summary of the research, findings, limitations, and suggestions for future research.

Chapter 2 Passenger Urban Air Mobility

2.1 Overview

In January of 1951, a *personal helicopter*, which was "big enough to carry two people and small enough to land on your lawn," was featured on the cover of Popular Mechanics [1]. Since then, numerous terms have been used to describe air transportation in metropolitan areas, including *helicopter air carrier, air taxi*, *On-Demand Aviation (ODA)*, *On-Demand Mobility (ODM)*, *on-demand urban air transportation*, and most recently, *Urban Air Mobility (UAM)* [2]. The media has chosen the name *flying cars* regardless of the industry disapproval [37].

With the vision of autonomous aircraft equipped with electric propulsion, which enables a 10 times reduction in energy costs [4], the interest in air transportation has resurfaced. Benefitting from revolutionary aircraft technology, the Advanced Air Mobility (AAM) [7] initiative is pursuing to transfer cargo and passengers between urban, local, regional, and intraregional areas, while the UAM market focuses on carrying passengers and goods within metropolitan areas [7-9]. Even though UAM covers passengers and cargo, the terminology is mainly used to refer to passenger-carrying services in the literature [34,38-40]. As a result, passenger-carrying UAM and UAM are used interchangeably in this dissertation.

UAM Coordination and Assessment Team (UCAT) outlines 6 *UAM Maturity Level (UML)* in 3 states: initial (UML 1 and 2), intermediate (UML 3 and 4), and mature (UML 5 and 6) [41]. The capabilities of each state are defined as follows:

- UML1: Late-stage certification testing and operational demonstrations in limited environment;
- UML2: Low density and complexity commercial operations with assistive automation;

- UML3: Low density, medium complexity operations with comprehensive safety assurance automation;
- UML4: Medium density and complexity operations with collaborative and responsible automated systems;
- UML5: High density and complexity operations with highly-integrated automated networks; and
- UML6: Ubiquitous UAM operations with system-wide automated optimization.

Table 2.1 compares the details of operations across UML 3 to 6.

Table 2.1 Comparison of operations across UML 3 to 6

	Number of Simultaneous Operations	Network	Weather
UML3	-	Closely-spaced UAM ports	Weather-tolerant operations
UML4	100s	Expanded network including high capacity UAM ports	Low-visibility operations
UML5	1,000s	Highly-distributed	High-weather tolerance including icing
UML6	10,000s	Ad-hoc (ubiquitous)	-

In summary, the idea of on-demand aerial operation has been around for decades. However, UAM is a fledgling vision facilitated by recent advancements in aircraft technologies. Numerous use cases of UAM have been discussed in the news and reports. As a result, this chapter reviews the early and recent advancements in UAM, presents the envisioned use cases, and ultimately reviews the projected components and characteristics of passenger-carrying UAM.

2.2 Early Developments in UAM

Aerial operations resembling UAM could be traced back to 1963 [2] when four helicopter carriers were offering scheduled service, mainly between major airports or an airport and downtown area, in Los Angeles, San Francisco, New York, and Chicago. By 1967, over 1.2 million passengers were being transferred annually. As early "air taxi" operations, more than 100 companies provided intracity transportation with advance reservations. More interestingly, based in Boston, Air General provided on-demand service for commuters from 1962 to 1969 with a reservation scheme as little as 30 minutes ahead of departure time. Regardless of aircraft type (i.e., helicopter vs. eVTOL aircraft), the Air General business model highly resembles the on-demand passenger UAM envisioned today. However, helicopters are inefficient and noisy and have high maintenance costs [23]. As a result, financial challenges and public acceptance significantly reduced the early air operations for commuting purposes [2] to the point that today few cities such as New York and São Paulo have large-scale commercial urban aerial transportation using helicopters [23]. Currently, BLADE Urban Air Mobility, Inc. offers on-demand passenger aerial service in New York City using helicopters and fixed-wing aircraft, while SkyRyde offers a similar aerial service in Los Angeles with its fixed-wing fleet [8].

2.3 Recent Developments in UAM

Urban passenger-carrying aerial service has drawn considerable interest amongst certain companies and communities in the past couple of years. Inspired by advancements in distributed electric propulsion (DEP) and vertical take-off and landing (VTOL) technology, various companies and start-ups, including the ride-hailing company Uber [3] [11,42] and aircraft manufacturers Airbus [12,43] and Boeing [13], have shown considerable interest in the at-scale

[3] Uber Elevate was acquired by Joby Aviation in December 2020

operation of passenger UAM. The United Arab Emirates (UAE), New Zealand, and Singapore are expected to be the early adopters of passenger UAM, with Dubai's ambitious plan for launching the commercial service by 2022 [27].

Uber, one of the major players in the ride-sourcing industry, created much excitement around UAM by announcing its plan to enter shared air transportation at a price comparable to its ground ridesharing service [11,28]. UberAIR utilizes eVTOL aircraft and is envisioned as a multi-modal service, where Uber's ground-based service conducts the legs from the origin to "skyport" and from "skyport" to the destination of each trip. The flights are initially planned between suburbs and cities with the ultimate goal of intracity shared flights. Uber had initially planned to start eVTOL demonstrator flights in 2020 and launch commercial uberAIR flights in 2023 [28]. The service needs to be affordable for large-scale transportation. Currently, the cost per passenger-mile is estimated at $8.93 for helicopters. However, Uber estimates a cost per passenger-mile of $5.73, $1.84, and $0.44 for the launch period, short term, and long term UberAIR operation, respectively [28]. To provide more insights on UAM, Uber Copter launched aerial service in New York in July 2019 and offered 8-minute rides from Lower Manhattan to JFK at the average cost of nearly $200 per person. The aerial service was complemented with ground-based service on both ends of the trip [44].

Airbus has partnered with BLADE Urban Air Mobility, Inc., an air charter broker and indirect air carrier, to gain experience from BLADE's per-seat and on-demand aerial operations, where passengers use an app for booking flights in the Northeast, Los Angeles, San Francisco, and soon Mumbai [45]. BLADE also offers helicopter rides between Manhattan and commercial and private airports with a minimum flight time of 5 minutes. BLADE Urban Air Mobility recently started a continuous per-seat helicopter service between three BLADE Lounges located in

Manhattan and JFK, La Guardia Airport, and Newark Liberty International Airport [45]. Its users utilize the app to schedule their own private charter flight or crowdsourced charter, or book individual seats on an existing flight.

Volocopter [16], the German aircraft manufacturer, has announced its plan to launch commercial flights in Paris and Singapore in the next 2-3 years. Volocopter promotes "Autonomous air taxi at the press of a button" and "mobility in three dimensions: urban. autonomous. on demand". VoloCity, the two-seater aircraft developed by Volocopter, has received permits to fly in manned or unmanned configurations for conducting test flights in Germany, Singapore, Dubai, and Helsinki.

EHang [46], based in China, envisions their UAM operation as an autonomous low-altitude short-and-medium-haul service, which would resemble on-demand bus operations rather than taxi operations. In May 2019, EHang launched its passenger-carrying UAM service between a harbor and a boutique hotel on an islet in China's Zhejiang province. The service decreased the 40-minute travel time to 5 minutes [32].

2.4 UAM Use Cases

Aside from the UAM market studies [8,9], ConOps [7], and OpsCons [29] provided by NASA and FAA, multiple players from the industry have offered their visions on passenger AAM and UAM [23,30-32]. Patterson et al. [33] summarize the numerous proposed visions in five on-demand or near-on-demand passenger UAM missions: private operation, air taxi, air pooling, semi-scheduled commuter, and scheduled commuter.

Porsche Consulting suggests five missions: personal ownership, rental, on-demand air taxi (including sightseeing), air bus, and rescue operations [6]. The Roland Berger UAM study [19]

covers both intracity and intercity[4] trips and defines three use cases: *City Taxi*, *Airport Shuttles*, and intercity service. The *City Taxi* would operate within a densely populated city and is an on-demand point-to-point non-stop service carrying 1 or 2 passengers with light luggage, while *Airport Shuttle*, as the name suggests, transfer 2 to 4 passengers with luggage between airport and take-off and landing sites within an urban area. City Taxi and Airport Shuttles cover distances from 15 to 50 km (i.e., 9.3 to 31.1 mi), while intercity flights cover distances up to 250 km (i.e., 155.3 mi). They estimate that by 2050, City Taxi and Airport Shuttles would, respectively, constitute 36% and 35% of the "UAM"[4] trips worldwide. However, the Airport Shuttle and intercity flights would account for 90% of the revenue.

The Booz Allen Hamilton's UAM market study commissioned by NASA [8] identifies 36 potential markets among 16 market categories. The market categories include air commute, first response, logistics and good delivery, public service, and rentals, among others. Four markets, in turn, constitute the air commute market category: *Airport Shuttle, Air Taxi, Train,* and *Bus*. The *Airport Shuttle* has fixed routes, while the *Air Taxi* service is point-to-point without fixed routes or schedules. The Air Taxi could be viewed as the extension of Airport Shuttle with high demand and more network coverage and bigger fleet size. The *Train* would be operated along the network infrastructure (e.g., subway and train) while the *Bus* would replace public transportation lines (e.g., Greyhound).

Another UAM study commissioned by NASA [9] presents and evaluates the viability of two use cases for passenger UAM, namely, *Air Metro* and *Air Taxi*. In this study, *Air Metro* transfers 2 to 5 (with an average of 3) passengers, employing autonomous aircraft over fixed routes

[4] More accurately, the intercity trips should be classified under AAM that covers local, regional, and intraregional areas.

with regular schedules, and therefore, resembles subway or buses. In comparison, *Air Taxi* provides on-demand door-to-door per-seat service operated with a fleet of autonomous eVTOL aircraft with 2 to 5 passenger seats. However, in the Air Taxi service, the average number of passengers is assumed to be 1. The Air Taxi mission resembles the current ridesharing service.

In May 2020, the UAM Operational Concept (OpsCon) for passenger-carrying operations commissioned by NASA [29] specified three use cases over three states of UAM Maturity Level: *Human-piloted Air Medical Transport* (initial state), *Intra-Metro Air Shuttle* (intermediate state), and *Ubiquitous Air Taxi* (mature state). The *Air Ambulance* flights are unscheduled, with about 10 UAM aircraft per metro area and 2 aircraft flying simultaneously. *Intra-Metro Air Shuttle* is a scheduled or seme-scheduled service transporting 3 to 9 passengers. The trip distance is 10s of miles, with 100s of aircraft per metro area and 10s vehicles flying simultaneously. Lastly, *Ubiquitous Air Taxi* utilizes semi-autonomous or fully autonomous eVTOL aircraft with low noise, low operating costs, and passenger capacity of 1 to 4. The service is on-demand, but one could book their flight with advance notice. The flights are shared, carrying 1 or 2 passengers on a typical flight. In this state, there are 100s of take-off and landing areas and 10,000s of aircraft per metro area, with 1,000s of aircraft flying simultaneously.

In conclusion, the *Urban Air Taxi (UAT)* is a use case envisioned in all market analyses. It is a ubiquitous on-demand per-seat service operated with autonomous eVTOL aircraft with low operating costs. In this case, the network coverage is high, and numerous take-off and landing sites provide near-ubiquitous service. However, the point-to-point service may lead to a low passenger load factor per aircraft.

2.5 UAM Components and Characteristics

There are many factors involved in the UAM operations. This section discusses the passenger UAM operations characteristics in detail.

2.5.1 Aircraft

The vision of UAM is enabled by electric propulsion and vertical take-off and landing, hence the name eVTOL. There are 95 eVTOL projects in progress across the world [19]. However, no dominant design has yet emerged. These aircraft could be classified under five categories: highly distributed propulsion (multicopters), quadcopters, hybrid, tilt-wing/convertible aircraft, and fixed-wing vectored thrust concepts [19]. Table 2.2 lists some aircraft designed for passenger UAM. These aircraft could accommodate a maximum of four passengers. Their speed varies between 62 and 200 mph, with the battery range spanning from 19 to 186 miles with a single charge.

Table 2.2 Features of various aircraft for UAM operations

Company	Aircraft	Planning Year	Passenger Capacity	Crew Capacity	Maximum Speed (mph)	Range (mi)
Volocopter [16]	VoloCity	-	1	1*	68	21
Airbus [47]	CityAirbus	2023	4	A	75	19
EHang [32,46]	eHANG 216	2019	1	A	80	22
Boeing NeXt [48]	Boeing PAV	-	2	A	-	50
Kitty Hawk [17]	Heaviside	-	1	A	180	100
Hyundai [49]	S-A1	-	4	1*	180	60
Joby Aviation [18]	-	2024	4	1	200	150
Uber Elevate [23]	-	2023	2-4	1*	150-200	100+

Note: *Envisioning future autonomy, PAV: Passenger Air Vehicle, A: Autonomous

Uber has focused on mega commuters (who travel, one way, 90 minutes or more and 50 miles or more to work [50]) and planned its operation with an aircraft speed range of 150-200 mph as they argue this is the speed range that electric distributed propulsion (EDP) is most efficient. Even though Uber intends to use a homogeneous fleet in the initial stage, it contends that an eVTOL aircraft for trips less than 50 miles does not need the maximum speed required for performing long-distance travel [23]. Furthermore, Uber suggests that the aircraft's battery range should cover two 50-mile trips at maximum cruising speeding with the two corresponding take-offs and landings plus 30 minutes reserves to meet the FAA Instrument Flight Rules (IFR) [23]. Volocopter, on the other hand, argues that aircraft with a range of 30 km (21 mi) can serve the important airport routes in 93% of the world's largest cities and adds that a speed range of 80-100 km/h is a trade-off between competing with ground-based transportation and avoiding the complications resulting from higher speed [31].

The batteries could either be charged or swapped. Uber requires a charging period of fewer than 7 minutes for continuous operation [51], while Volocopter specifies the battery swapping time of 5 minutes for its passenger UAM aircraft, VoloCity [16]. Similarly, the Crown Consulting market study assumes a 2-4 minutes period for swapping batteries [9].

Volocopter claims that they have conducted the world's first autonomous eVTOL flight in Dubai in 2017 and public test flights at Singapore's Marina Bay in October 2019 [16]. Wisk asserts that in 2017, they became the first company in the U.S. to successfully fly a passenger autonomous eVTOL aircraft [52].

There is no consensus in the community on the term used to refer to UAM passenger aircraft. Table 2.3 lists some of the names used in the literature. *UAM aircraft/vehicle* is a term used by some studies commissioned by NASA. Even though autonomous [9] and electric [33]

VTOL aircraft are viewed as the enabler of UAT operation, in this dissertation, we use the term UAT aircraft (or aircraft in short) for the UAT operation since the proposed methodology is independent of the aircraft type.

Table 2.3 Terms used for referring to UAM passenger aircraft

Term	Study
Personal Air Vehicle (PAV)	Moore [53], Hyundai [14]
Passenger Air Vehicle (PAV)	Boeing [48]
Electric Vertical Aircraft (EVA)	BLADE Urban Air Mobility Inc. [45]
Autonomous Aerial Vehicle (AAV)	EHang [46]
Passenger Drone	Porsche Consulting [6], Roland Berger [19]
VTOL aircraft	Uber [23] Crown Consulting, Inc. [9]
UAM aircraft/vehicle	Booz Allen Hamilton, Inc. [8] Patterson et al. [33] FAA [7] Deloitte Consulting LLP [54] Price et al. [29]

2.5.2 Infrastructure

eVTOLs operations require sites for take-off and landing, picking up or dropping off passengers, parking, and charging. For instance, Heaviside by Kitty Hawk needs a 30-foot by 30-foot area for take-off and landing as a pad, which does not have to be paved [17]. In the urban setting, the required infrastructure could be active helipads, the roof of the public parking [23], the space available in cloverleaf interchanges [55], or a new infrastructure built for UAM operations.

Table 2.4 lists the various names used in the UAM community for referring to UAM infrastructure. The use of *vertiport* could be traced back to 1967 [56], and it is the term used most frequently in the literature. Regardless of the name, there is a consensus that passenger UAM operation requires two types of facilities: a smaller one just big enough for an aircraft to land and take off safely and a bigger one with all the supporting facilities for aircraft and passengers,

including charging stations and parking spots. The former is often referred to as a *pad*, while the latter may have a *port* or *hub* as part of its name.

Table 2.4 Terms used for referring to UAT take-off and landing sites

Term	Study
Take-off and Landing Area	Vascik, 2017 [2]
Vertiport/Vertistop	Cheyno, 1967 [56] Uber, 2016 [23] Holmes et al., 2017 [57] Crown Consulting, Inc., 2018 [9] Booz Allen Hamilton, Inc., 2018 [8] Porsche Consulting, 2018 [6] Patterson et al., 2018 [33] Vascik and Hansman, 2019 [58] Lilium [15]
VertiPad/VertiHub	Airbus [43], McKinsey & Company, 2020 [10]
Skyport/Skystop	Uber, 2018 [51]
VoloPort	Volocopter [16]
Base Point	EHang [32]
BLADE Lounge	BLADE Urban Air Mobility Inc. [45]
UAM Port/Pad	NASA UCAT, 2020 [41] Price et al., 2020 [29]
Aerodrome	FAA, 2020 [7] Deloitte Consulting LLP, 2020 [54]

UAM pads and UAM ports are primarily used in the literature for passenger-carrying aerial operations resembling the UAT use case. Consequently, we adopt the terms UAT pads and UAT ports to specify the use case of the infrastructure clearly. *UAT pad* is a general term for an area designated for take-off and landing of a single UAT aircraft and could operate as a stand-alone facility, while *UAT port* is used for a facility with multiple UAT pads and all the required supporting systems such as charging stations [29]. That being said, for simplicity, we use the term UAT pad to refer to the UAT infrastructure, regardless of its size and available resources.

Uber envisions that ports have a maximum capacity of 12 VTOLs [23]. McKinsey & Company [10] envision three types of UAM infrastructure: *Vertipad* (small structure), *Vertibase* (medium structure), and *Vertihub* (large structure). Vertipad and Vertibase could be either new or retrofit, while Vertihub is a new structure. In each infrastructure, the ratio of landing and take-off pads to the parking and charging spots is devised as 1 to 2. Vertipad, Vertibase, and Vertihub have 1, 3, and 10 landing and take-off pads, respectively. Furthermore, they assume that the infrastructure charge is USD 150 per trip, excluding the fuel charging costs. In a small and premium UAM market, they conclude that there should be a 24-hour average of 1 trip per hour per UAT pad or 1 trip every 20 minutes per UAT pad during peak periods in a large and densely populated city to break even on the fixed costs. In a medium-size and less dense city, the average over 24 hours per UAT pad is 1 trip every 100 minutes or 1 trip every 30 minutes during peak periods. However, to achieve very low costs and make UAM available to the public, UAT pads across the network should accommodate one trip every 5 minutes during the peak, which could be challenging.

2.5.3 Network Coverage

A higher number of UAT pads spread across the network leads to more significant travel time savings. In the mature state, the *ubiquitous* network of UAT pads provides an opportunity for *point-to-point* (also referred to as *door-to-door*, *end-to-end*) service. However, in the initial state, the operation would be *hub-to-hub* [6]. A UAM market study [8] shows that the number of UAT pads and ports significantly impacts the UAM demand.

In the 1960s, Air General utilized over 70 heliports in Boston [2]. São Paulo has 193 active helipads [23]. McKinsey & Company [10] estimate that 85 to 100 stand-alone UAT pads or UAT ports are required for large and densely populated cities, and the number drops to 38 to 65 for

medium-size cities. Porsche Consulting [6] projects that cities with a population of five to ten million or more will have up to 100 stand-alone UAT pads or ports in the fully developed phase. However, in the first phases, 5 existing heliports are sufficient for providing attractive routes, and in the next phase, 40 UAT pads will be available in some areas. Decreasing the number of UAT pads from 100-200 per city to 15-40 would decrease the market size by 40% [6]. The UAM OpsCon for passenger-carrying operations commissioned by NASA [29] projects 10s of UAT pads or ports for the intermediate state and 100s for the mature state.

2.5.4 Travel Time Saving and Trip Distance

The competitive advantage of passenger UAM is in the travel savings. Uber assumes that a UAM trip should be at least 40% faster than the corresponding ground-based trip, while Porsche Consulting [6] suggests UAM needs to offer at least 20% travel time savings to be competitive with other modes. Booz Allen Hamilton's market study finds no significant demand for mandatory (i.e., work-related) trips that take less than 30 minutes on the ground. Furthermore, most of the UAM demand is captured for trips that are at least 45 minutes on the ground.

Air taxi service would be more beneficial over long distances [8]. Porsche Consulting suggests that UAM could outperform other modes of transportation for trips that are at least 20 km (i.e., 12.4 mi) long, which is almost twice as long as the average trip distance of 11 km in the urban settings [6]. The Roland Berger UAM study [19] specifies that trips should be at least 15 to 25 km (i.e., 9.5 to 15.5 mi), while the intercity service will be provided for distances between 15 to 50 km (i.e., 9.3 mi to 31 mi). Additionally, a UAM market study commissioned by NASA [9] envisions UAT trips between 10 to 70 miles.

2.5.5 Fleet Size

The fleet size directly impacts the level of service and utilization of the fleet [59]. As a result, the sensitivity to fleet size is often investigated in related studies [59-61]. Uber Elevate aims for a low volume of UAM aircraft in 2023, a number between 10 to 50 aircraft [23]. Porsche Consulting estimates that by 2035 a megacity with a population of five to ten million will not have more than 1,000 UAM aircraft in operation [6]. For instance, they estimate that São Paulo requires 5 UAT pads and 120 aircraft in the initial phase, 40 pads and 390 aircraft in the expansion phase, 100 pads and 1050 aircraft in the full-service phase. In comparison, São Paulo has the largest registered fleet of 420 helicopters [23].

2.5.6 First-mile and Last-mile Service

Offering the first-mile and last-mile service would alleviate the inconvenience of the multi-modal trips. BLADE Urban Air Mobility currently offers point-to-point service using helicopters, seaplanes, jets, and SUVs [45]. Uber Elevate announced that uberAIR would conduct the first and last mile of trips with its ground ride-hailing service or walk [23]. In Hyundai's Smart Solution Mobility vision, the UAM aircraft are connected to *Purpose Built Vehicle (PBV)* at *hubs* to provide a multi-modal service [14].

2.5.7 Flight Sharing

The air taxi operation will be most efficient and cost-effective when aircraft are highly utilized, and thus, passengers share the flights. Booz Allen Hamilton's UAM market study shows that passengers are willing to share the flight with passengers they do not know as long as they receive a discount [8].

However, it remains unclear whether ridesharing should be limited to passengers with the same pick-up and drop-off UAT pads or the operation should resemble ground-based ride-sourcing

programs where multiple passengers share an aircraft despite having different pick-up or drop-off locations. In the former case, the passengers are *pooled* and use one flight to get from starting UAT pad to the ending UAT pad, and there is no intermediate stop to pick up or drop off other passengers. We refer to this mode of sharing as *air pooling*. On the other hand, the latter case has the attributes of ground-based ridesharing services and is what we refer to as *air sharing* hereafter.

BLADE Urban Air Mobility currently offers crowdsourcing, in which case the passenger books the flight for their desired departure time and allows BLADE to sell any available seat. Thus, the passenger could receive credits back for the purchased seats by other passengers [45]. Furthermore, Uber Elevate has incorporated shared flights in its envisioned operational strategy [23].

2.5.8 Advance Reservation Time Window

Reservation schemes and the time window that customers are given to request a flight ahead of desired departure time are significant indicators of how dynamic the evolution of information in the system is. The more time the operator has, the more operationally optimal strategies they could employ. *Immediate requests* are the ones that need to be served immediately as opposed to *advance requests*, which give the operator some time before the desired service time [62].

The emergence of very light jets (VLJs) in the early 2000s prompted *Dial-A-Flight* business models on a regional level where passengers could book their flights one or a few days in advance [3]. As an on-demand aerial mobility service, fractional ownership programs offer a reservation window of 4 or 48 hours [63] or 8 hours [64] ahead of departure time.

In urban air taxi operations, Air General's users could request a flight with as little as 30-minute notice in the 1960s. Today, BLADE Urban Air Mobility offers both immediate and advance reservation options for its air taxi service in New York City. Immediate requests for helicopters with just "minutes notice" to or from any airport in the area would cost between $1,575 - $1,775, while a 24-hour notice for charter flights to JFK or another local area airport would reduce the cost to $795 - $995 [45]. Uber Elevate discusses no advance reservation [23] or one-hour lead time [28], consistent with its on-demand service philosophy. Similarly, the UAM OpsCon for passenger-carrying UAM envisions operations to be generally on-demand with the option to schedule trips in advance [29].

2.5.9 Operating Costs and Passenger Price

The operating cost per passenger mile is one of the significant factors in the viability of UAT. Efficient UAT operations, which involve high aircraft utilization, high passenger load factor, and low empty mileage, as well as high cruise speed, decrease operating cost per passenger mile [8].

McKinsey & Company [10] assert that the costs of UAM trips should decrease by around 80% of current helicopter rides for the service to be competitive with ground-based transportation. They estimate the operating cost of $0.5-2.5 per seat-mile for UAM, compared to $6-8 for current helicopter service. Moreover, the energy cost for the electric motor is estimated at $0.13/kWh vs. the $5.5/gallon for the combustion engine.

In comparison, the cost per passenger mile is estimated at $8.93 for helicopters by Uber Elevate, while they estimate the cost per passenger mile of $5.73, $1.84, and $0.44 for the launch period, short term, and long term uberAIR operation [28]. Booz Allen Hamilton's UAM market study estimates the median operating cost per passenger mile of $9.5, $7.0, $5.5, and $4.75 for air

pooling business model with 1, 2, 3, and 4 passenger seats, respectively. In the near future, it expects a 5-Seat eVTOL to cost $6.25 per passenger mile. Porsche Consulting [6] projects that a trip from the airport to the city with a 10-min flight time will be priced at $123. They estimate the operating cost of air taxi service will be $1.8 per km (i.e., $2.9 per mile). To put these numbers into perspective, a 22-minute flight by Skyway Air Taxi costs $950 for up to three passengers and baggage [65].

In February 2021, Joby Aviation [5] estimated an operating cost of $95 over a 25-mi trip using their four-seater aircraft (corresponding to 0.95¢ per available seat-mile), a 4 times cost-per-mile improvement over a twin-engine helicopter with an operating cost of $393 for a similar trip. Subsequently, in October 2021, Joby Aviation [66] presented cost drivers of service unit economics at scale in 2026, shown in Table 2.5. Cost per available seat-mile (CASM) is estimated to be 0.86¢, out of which 0.22¢ (around 25%) is the pilot cost. McKinsey and Company [67] estimate that the cost per passenger-seat-kilometer of a piloted UAM flight could be two times higher than an autonomous flight.

Table 2.5 Cost drivers of service unit economics at scale in 2026 by Joby Aviation

Service Component	**Estimated Cost per Available Seat Mile (cents)**
Pilot	22
Maintenance (including labor)	19
Skyport Support and Landing Fee	11
Battery and Charging	13
Aircraft and Insurance	9
Other Expenses	12
Cost per Available Seat Mile (CASM)	**$0.86**

For pricing, McKinsey & Company [10] project that, in a small and premium UAM market, the charges should be $50 to $75 per passenger, depending on the number of passengers per trip, for an intracity and metropolitan UAM travel with a distance of under 50 miles. They further assert that for at-scale operations, the price per passenger trip should be around $25, which requires 10,000 trips per day in a large, dense, high-income city and approximately 3,500 trips per day in a medium-size, less dense city. Porsche Consulting [6] assumes a price between $8 to $18 per minute for the on-demand air taxi service. Additionally, in the Booz Allen Hamilton's market study, the maximum revenue was achieved at passenger price of ~$2.50-$2.85 per mile for 10 study areas, including Dallas, Los Angeles, New York, and Washington D.C. Lastly, Joby Aviation estimates the price of $3 per passenger mile given the revenue drivers of the service unit economics at scale in 2026, shown in Table 2.6. Subsequently, the average load factor of 57.7% would result in Passenger revenue per available seat-mile (PRASM) of $1.733.

Table 2.6 Revenue drivers of service unit economics at scale in 2026 by Joby Aviation

Average Flight Length	24 miles
Cruising Speed	~165 mph
Average Load Factor	2.3 Passengers for a 4-seater Aircraft (i.e., 57.5%)
Turnaround Time	~6 minutes
Price per Passenger-Mile	$3.00
Passenger Revenue per Available Seat Mile (PRASM)	**$1.73**

2.6 Concluding Remarks

UAM is a nascent idea at the core of numerous discussions by NASA and FAA, news coverage, reports from consulting companies, and white papers and reports by original equipment manufacturers (OEMs). Nonetheless, the vision has not been fully developed, and it is not currently employed. As a result, this chapter presents the developments in UAM and reviews the envisioned

use cases. Furthermore, it reviews the relevant components and the projected characteristics of UAM from the industry perspective. Chapter 3 provides a review of the literature related to UAM and UAT and aims to offer an academic perspective.

Chapter 3 Literature Review

3.1 Overview

Transportation on Demand (TOD) [68] is the concept of moving goods or passengers from their origin to their destination when the service is provided based on customer requests. TOD includes ridesharing, bike-sharing, carsharing, taxi service, Transportation Network Companies/Providers (TNCs and TNPs), and on-demand air mobility. In formulating TOD problems, three conflicting objectives manifest themselves: minimizing operating costs, maximizing the number of requests served (and thus maximizing the revenue), and maximizing the level of service. The level of service could be defined in terms of deviation from desired pick-up or delivery times. In passenger transport, wait time or excess ride time are alternative measures of the level of service.

Decisions regarding managing TOD systems typically have three intertwined components: request clustering, request routing, and request scheduling [68]. Request clustering aims to reduce operational costs by creating groups of requests that are close in time and space. Request routing finds the routes of vehicles for serving the customers. Lastly, request scheduling determines the exact timing of each visit.

This chapter aims to review the literature related to UAM and UAT, and is organized as follows. First, the *Vehicle Routing Problem with Pick-up and Delivery (VRPPD)*, the class of problems that UAT fleet operation belongs to, is discussed. Next, the relevant studies on on-demand air mobility are reviewed. Lastly, an overview of the literature on synchronized logistics is presented.

3.2 Vehicle Routing Problem with Pick-up and Delivery (VRPPD)

Vehicle Routing Problem (VRP) refers to a class of problems where a set of locations (nodes) should be visited only once by identical vehicles located at depots such that it minimizes transportation costs. *Vehicle Routing Problem with Pick-up and Delivery (VRPPD)* is a generalization of VRP, where goods should be picked up or dropped off at specific locations. When the units moved in the transportation network are passengers, one should take the user's inconvenience into account, which shapes other variants of VRPPD, namely, *Dial-A-Ride-Problem (DARP)* for ground transport and *Dial-A-Flight-Problem (DAFP)* for air transport.

VRP and its variants are generally classified based on the *Quality* and *Evolution of information* [69]. *Quality of Information* reflects the uncertainty in the input data available to the decision-maker, while the *Evolution of Information* reflects how the available information changes during the execution of the plan. Based on these two dimensions, VRPs are classified as follows:

- Quality of Information: Deterministic vs. Stochastic

  In deterministic problems, the input data available to the decision-maker is deterministic, while in stochastic problems, the input data is uncertain and a random variable. The stochasticity could be associated with demand (such as its location, timing, or intensity), travel times, or service breakdown [70].

- Evolution of Information: Static vs. Dynamic

  In static problems, all the information (regardless of its quality) is available to the decision-maker before the planning phase. On the other hand, a problem is classified as dynamic when one or some of its input data varies with time [70], and consequently, part of the information is revealed during the design or execution of

routes [69]. As the information becomes known in dynamic systems, the routes should be adjusted in real-time in response to the new information.

Dynamic problems either have *dynamic data* or *time-dependent data* [70]. Dynamic data changes with time and may include customer demand or travel times. Time-dependent data, however, are known in advance and may include *Vehicle Routing Problem with Time Windows (VRPTW)*. Additionally, customer demand or travel times might be considered time-dependent data if they are defined as functions of time, and therefore, they are all known in advance.

It is crucial to distinguish between dynamic *problems*, *models,* and *applications* [70]. A model is considered dynamic if it explicitly models the changes of input data over time. Nonetheless, it is possible to have static or dynamic applications of static or dynamic models. If a dynamic model is solved only once and the analyst selects one strategy regardless of changes of the input data over time, that constitutes a static application of a dynamic model. In contrast, if a static model is solved repeatedly as new information is revealed in the system, it is considered a dynamic application of a static model. In practice, dynamic and deterministic models are often solved as a sequence of static and deterministic models [68], which is regarded as a dynamic application of a static model.

Degrees of dynamism [71] is a measure that seeks to explain the *frequency of changes* and *the urgency of requests* for a problem with dynamic requests. The frequency of changes refers to how often changes in requests happen. For instance, how often a new request arrives or the attributes of a request change. The urgency of the requests reflects the available response time. In other words, it shows the available time window between receiving a request and serving it. Accordingly, VRP could be *weakly, moderately,* or *strongly dynamic*. This information helps to identify the appropriate solution methods given the trade-off between solution time and accuracy.

For instance, in a weakly dynamic system, one has the time to achieve an optimal or near-optimal solution. However, in a strongly dynamic system, time constraint limits the accuracy of the solution. The *degree of structural diversity* [71] reflects the spatial and temporal dynamism of the requests and highlights the high value of using stochastic information about future requests in a network with a high degree of structural diversity.

As the new information, such as a new request for service, becomes known in a dynamic model, three methods could be used to adjust the solution [71]:

1. *Policy:* In this method, a policy or rule is used to obtain new solutions. Examples of these rules include First-Come-First-Served (FCFS) or the nearest idle vehicle for assigning new requests to vehicles.
2. *Local Heuristic Search:* In this approach, the static problem is solved at the beginning of the planning horizon using the information available to the analyst at the time. As new requests arrive, the current solution is adjusted by employing heuristic methods such as insertion heuristics, deletion heuristics, or interchange [72].
3. *Re-optimization:* In this case, the problem could be re-optimized every time new information becomes available. Depending on the size of the problem, degree of dynamism, and the time available for solving the problem, exact, approximate, or heuristic methods could be employed to update the current solution with the new information.

Dynamic VRPPD (D-VRPPD) has numerous applications in *Truckload Pick-up and Delivery Problems (TLPDP)*, *Dial-A-Ride Problem (DARP)*, and *Dial-A-Flight Problem (DAFP)*, which are discussed in the following sections.

3.2.1 Freight Truckload Pick-up and Delivery Problem (TLPDP)

In *Truckload Pick-up and Delivery Problem (TLPDP)* [73-78], a trucking company with a fleet size of K aims to serve the requests given pick-up and drop-off locations, the earliest pick-up time, and the latest delivery time of the *job*. A truck cannot serve a new request until it completes its previous job, and therefore, TLPDP resembles on-demand services with no ride-sharing. However, TLPDPs are associated with transferring goods, and consequently, the quality of service is limited to deviations from desired pick-up or delivery time. Desired pick-up and delivery time could be formulated as a *hard constraint* in the model, or alternatively, the delay could be included in the objective function with a penalty term as a *soft constraint* [75,78]. The operator has the option of either accepting or rejecting the requests. The acceptance and rejection decision-making process could be rule-based [73] or be incorporated in the optimization model [75]. The objective functions of TLPDP could cover empty distance traveled [73,75,78], penalty cost for the delay from desired pick-up or delivery time [75,78], and revenue loss resulting from rejecting requests [78].

Yang et al. [78] model the TLPDP as an assignment problem with timing constraints. The offline Mixed Integer Programming (MIP) formulation seeks to find the least cost assignments between all the nodes defined as $\{1, \dots, K, K+1, \dots, K+N\}$, where K is the number of vehicles and N is the number of jobs. The objective function covers costs associated with empty mileage, delays in delivery times, and request rejections. Furthermore, they utilize rolling horizon strategies for real-time implementations and compare the optimization-based methodology with three heuristic methods. In their earlier paper, Yang et al. [75] conclude that even though optimization-based strategy outperforms other heuristic approaches, some of the heuristic approaches are competitive given their low computational requirements for solving the problem.

A special class of TLPDPs is the routing and scheduling of the drayage operations, which refers to the regional movement of trailers and containers, either empty or loaded, by tractors between rail yards, shippers, consignees, and equipment yards [79]. Smilowitz [79] models the drayage operations as a *Multi-Resource Routing Problem (MRRP)*, where multiple *resources* are used to perform a series of *tasks*. The tasks are either *well-defined* or *flexible*. The origin, destination, and time window of well-defined tasks are known, while either the origin or destination of flexible tasks is unspecified. She presents a node-based model and a set partitioning formulation, where requested tasks are partitioned into resource (i.e., tractor) routes. However, a conservative time window is placed on all tasks to remove the time dependency between tasks. In other words, the resources are assumed to be unavailable throughout the entire time window, even if the duration of a task is shorter than its time window.

Smilowitz [79] uses a constant radius around the origin or destination to define the flexible tasks. However, when a fixed radius is used for every node, a node in a dense area may have a higher set of possible executions of a flexible task than in a sparse area [80]. To address this issue, Francis et al. [80] propose the Variable Radius (VR) method, which limits choices for nodes in dense regions and increases choices for nodes in more sparse areas.

The frameworks developed for TLPDPs are applicable in ubiquitous operations of UAT without air pooling since these frameworks take the operating costs and a notion of user inconvenience into account. Some variations of TLPDP also formulate the cost of rejecting a request, which is relevant to UAT problems. However, the most significant difference is the sharing economy, a critical factor to the viability of the UAT business model [9].

3.2.2 Dial-A-Ride Problem (DARP)

Dial-A-Ride Problem (DARP) is similar to VRPPD but puts more emphasis on customer inconvenience since it deals with passengers, not goods. DARP was initially designed for non-profit services to senior citizens and people with disabilities, where most people share either the same origin or destination. DARP has recently gained more tractions in health care and *demand-responsive transportation (DRT)* to complement scheduled public transit [81].

In DARP, requests from customers are characterized by pick-up location, drop-off location, desired pick-up time, and desired drop-off time. The operator must design the routes and schedules of K vehicles such that it minimizes the operator's objective function and meet the service constraints. These service constraints may take wait time and ride time into account in addition to the delay in the desired pick-up and delivery time in VRPPDWT [68]. In these problems, the capacity of the vehicle is an operational constraint. The operator also has the option of rejecting the requests [81].

The objective function of DARP often seeks to minimize the operating costs (such as total travel distance, travel time, or fleet size) and user inconvenience. However, other objectives, including maximizing operator's profit or passenger occupancy rate and minimizing vehicle emissions, have been studied in the literature [81]. There are three methods of formulating a multi-objective DARP [81]:

1. In the first approach, the objective function is a weighted sum of different objectives. This approach is most appealing when the weight of one objective relative to another is well-defined, or all the objectives could be converted to the same evaluation unit. For instance, when the operating cost per unit mile and the

cost of wait time per minute are known, all the objectives could be defined in monetary values.

2. The second approach takes advantage of a hierarchical method. It first optimizes the most important criteria, then the second one, and so on. For instance, the DARP is first optimized by minimizing the operating costs, and then if possible, wait times of the passengers are minimized. This approach does not require the relative importance of objectives and is well-suited when one objective is dominant.

3. The third approach aims to obtain the Pareto frontier of the problem. As a result, it provides the analyst with multiple solutions and the trade-off between the conflicting objectives. For this reason, it is not suitable for instances where one solution is needed in a short time, for example, in the case of a dynamic DARP.

First studies on DARP date back to 1978 with Stein's static and dynamic DARP. Later in 1980, Psaraftis used dynamic programming and developed an exact algorithm for static and dynamic DARP with a single vehicle and immediate requests [81]. For the dynamic case, the static case is re-optimized when a new request arrives.

The solution algorithms of DARP are classified into three groups [81]. *Construction insertion heuristics* are basic, policy-based heuristics based on greedy insertion. They seek to insert a new request to the vehicle's route where the insertion cost is the cheapest. Even though these heuristics are basic, they are fast and, therefore, appropriate for dynamic DARP (D-DARP) applications with a high degree of dynamism. The second group of solution methods are *exact algorithms* and include branch-and-cut, branch-and-price, and branch-and-price-and-cut algorithms and reduction approach. The exact methods guarantee optimality and are most

appropriate for static DARP when the problem is solved once during the planning phase, and thus, a solution time in the range of hours is justifiable. In contrast, *heuristics and metaheuristics* are utilized to obtain a solution in a shorter amount of time. These methods include Tabu Search (TA), Simulated Annealing (SA), Variable Neighborhood Search (VNS), Large Neighborhood Search (LNS), Genetic Algorithm (GA), and hybrid methods. The readers are referred to Cordeau and Laporte [82] for a survey of DARP models and algorithms prior to 2007 and Ho et al. [81] for a survey of recent developments in the field.

DARPs address the pick-up and delivery problems in the context of passenger transportation. Therefore, they cover the constraints regarding passengers' wait time, maximum ride time, and the excess time incurred due to the detours. However, DARPs include ridesharing in their formulation. As a result, besides the fact that DARPs deal with vehicle routing problems with pick-up and delivery of passengers, they are not directly relevant to the UAT concept of operations defined in this dissertation since DARPs consider ridesharing where passengers have either the same origin or destination, and they may experience multiple stops along their route.

3.2.3 Dial-A-Flight Problem (DAFP)

Dial-A-Flight Problem (DAFP) was introduced by the emergence of on-demand air mobility. Analogies could be drawn between DAFP and DARP in many aspects. However, some of the characteristics are different [3,68]:

- In DAFP, the service is often offered at a specified set of airports, and therefore, the operator could take advantage of the fixed network structure.
- In DARP, the requests typically share a common origin or destination, which is barely the case with DAFP.

- Requests in DAFP are typically placed a couple of hours in advance, giving the operator more time to obtain an optimal solution to the routing and scheduling problem.
- Given that cost per mile of aerial operations is significantly higher than ground-based transportation, achieving optimal or near-optimal solutions is more critical in DAFP.
- DARP is traditionally for non-profit and social services, whereas DAFP is more common in commercial settings. Therefore, the level of service in the two problems is expected to be different.
- Weight constraints are considered in DAFP.
- DAFP should consider strict rules regarding pilot's and crew flying and duty hours and aircraft maintenance imposed by FAA.

DAFP rises in the context of per-seat on-demand air mobility, and it has been studied in [3,59,83,84,85]. Similar to DARPs, DAFPs address vehicle routing problems with pick-up and delivery of passengers with *air sharing*. Consequently, analogies could be drawn between DAFPs and UAT fleet operations if passengers in DAFPs experience no stops on their route. However, in DAFPs, pooling the passengers is not as challenging given that the airports are spaced far enough that passengers do not have multiple choices for the pick-up and drop-off. Additionally, the distances between one airport to another are long enough to warrant an empty repositioning of the aircraft.

3.3 On-Demand Air Mobility

On-demand air mobility is a service that is offered in response to the customer's request, not the operator's schedule. The first applications of on-demand air mobility using helicopters could be traced back to the 1960s. Since then, it has appeared in the literature under various names

[2]: Helicopter Air Carrier, Air Taxi, Metro taxi, Metrobus Intracity Air Transportation, Interurban Short-Haul Air Transportation, Personal Air Transportation, On-Demand Aviation, On-Demand Air Mobility, Zip Aviation, Sky Transit, On-Demand Mobility, Air Mobility on Demand, On-Demand Urban Air Transportation, and Urban Air Mobility[5]. This dissertation distinguishes between *On-Demand Air Mobility* and *Urban Air Mobility (UAM)*. On-Demand air mobility refers to on-demand aerial operations, either in urban or regional settings, while UAM refers to aerial operations in an urban setting, which are not necessarily on demand.

On the regional scale where scheduled flights through commercial airports is an alternative mode of travel, the demand for on-demand service mainly arises from limited schedules, congested airports and parking lots, wait times at security checks, flight delays, missed connections, and the distance between commercial airports and desired origin and destination [3,59]. However, on the urban scale, the travel time saving compared to ground-based transportation is the primary drive. On the supply side, advances in aircraft technology have reduced operating costs and environmental impacts, and increased efficiency. The combination of these factors in demand and supply has led to an increase in on-demand air services. Taxonomy of on-demand air mobility is presented as follows:

- Network extent: urban vs. regional
- Aircraft type: helicopter, small piston aircraft, very light jet (VLJ), vertical take-off and landing (VTOL)
- Per-seat vs. per-aircraft service
- Non-stop flight vs. multi-stop flight

[5] The concept of Urban Air Mobility (UAM) defined by NASA is not limited to on-demand transportation.

The two primary business models of passenger-carrying on-demand air mobility are fractional ownership programs and air taxi operations. In fractional ownership programs, the service is offered on a regional level, and customers own a share of aircraft. On the other hand, air taxi operates on both urban and regional scale, the operator has the liberty of rejecting the requests, and the customer might share the aircraft with other passengers. The following sections discuss the studies on these two types of operations.

3.3.1 Fractional Aircraft Ownership Program

High acquisition, operating, and maintenance costs of aircraft have led to the emergence of fractional ownership programs. These programs offer flights among 5500 airports compared to 500 airports for commercial airlines [64]. Fractional ownership programs are most appealing to small- to medium-size private companies that need to fly frequently but cannot justify purchasing and operating an entire aircraft [86].

The fractional owner orders a service with as little as 4-hour notice [63] and is entitled to fly for certain hours annually. For instance, one-sixteenth shareowners are entitled to 50 hours flying annually [63]. In addition to a one-time share purchase fee, the fractional owner pays a monthly management fee and is charged an hourly usage fee for flying the aircraft [86]. On the other hand, the operator guarantees to provide service at the customer's request while being responsible for crew scheduling and aircraft maintenance. There are five costs associated with the program's operation [64]: repositioning the aircraft while empty to the desired departure location, upgrading to a bigger aircraft, transporting the crew using commercial airlines, the crew working overtime, and chartering additional aircraft to serve a request. It is worth noting that repositioning time may comprise 35% or even more of the total flight time [64], and therefore, repositioning cost is a significant part of the operating costs. Finally, the period between the arrival time of a

request and the requested time for service allows the operator to optimize the operation while considering the new requests.

Multiple studies have focused on fractional ownership programs [63,64,86-90]. These studies differ in aircraft homogeneity, planning horizon, maintenance and crew scheduling constraints, and the objective function. Since fractional ownership programs are operated per aircraft, the flights are not shared, there is no intermediate stop to pick up or drop off other passengers, and there is no need for capacity and weight constraints, as is the case in DAFP.

In these problems, the network could be presented in two ways: the nodes are the airports (arc-based representation), or the nodes are the requests (node-based representation) [90]. Hicks et al. [63] develop an integer multi-commodity network flow problem and then employ a branch-and-bound approach to solve it. Yang et al. [88] introduce NETIP, a network flow model for the aircraft scheduling problem. They show that for randomly generated data with 200 aircraft and 400 requests over a 24-hour planning horizon, NETIP would take about 7 CPU minutes to obtain the solution. Yao et al. [64] formulate the crew pairing problem as set partitioning and obtain the solution using column generation. They also investigate the effects of modifying demands on improving aircraft utilization, and consequently, increasing profitability. To this end, they show that the charter costs would be reduced significantly if slight flexibility (in the order of minutes) on departure times were allowed.

Munari [89] develops a MIP model for per-aircraft services, which is solved using CPLEX. The objective function only considers operating costs, and maintenance schedules are implicitly considered as a request. The planning horizon consists of 7 days, and in total, 12 instances are solved. Solving instances with nearly 100 requests and 50 aircraft would take about 10 minutes. In some instances, the proposed model reduces the empty leg (i.e., ferry leg, deadhead, non-

revenue flight) up to 16 hours compared to the solution used by the operator at the time. This improvement is significant since flying costs are estimated between €3000 to €8000 per hour.

Munari and Alvarez [90] build on the previous aircraft routing and scheduling work by incorporating maintenance events and service upgrade costs in the model. The objective is to minimize the total operating cost, including repositioning and service upgrades. The network representation is similar to Keskinocak and Tayur [87] and Martin et al. [86], where nodes of the network are the requests that have to be served. The compact MIP is solved using GLPK, an open-source general optimization software. The planning horizon spans over three days, and the average number of requests in four different periods is 35.1, 40.8, 37, and 109.5 over a 3-day planning horizon, with the average number of aircraft being 18.4, 22, 21.3, and 49.7. Yao et al. [64] had previously examined the value of minutes of flexibility in desired pick-up and drop-off times. Munari and Alvarez [90] further build upon this idea and consider anticipation or postponement of the starting time of flights for 15 minutes and maintenance events for one day in their model. Lastly, they compare the computation time of open-source GLPK with commercial CPLEX. The results show that CPLEX could solve all the instances to optimality in less than 30 seconds. Interestingly, for three cases with upgrades where GLPK could not find the optimal solution within one hour, CPLEX obtained optimal solutions within seconds. These cases have combination of (118, 93, 52), (125, 93, 51), and (121, 88, 52) as (#requests, #airports, #aircraft). These findings highlight the significant role of the solver in obtaining an optimal solution for bigger instances within a reasonable time.

Fractional ownership program problems (FOPPs) share many features of ubiquitous UAT operations. However, the models in *FOPPs* tend to be more complex given the heterogeneous fleet and additional constraints. The differences between FOPP and UAT are explained below:

- UAT operation is envisioned as an autonomous service, and therefore, there is no need to consider crew flight hours and scheduling in the model.
- In contrast to FOPP, the UAT fleet is assumed to be homogenous. So, there is only one type of aircraft, and thus, no upgrade cost in the objective function.
- The UAT operator is not obligated to serve all the incoming requests, and therefore, it can reject requests when serving them is not feasible or profitable. However, in FOPPs, the operator has to serve all the incoming requests at the cost of upgrading the aircraft or chartering the flight.
- UAT operations could be per seat, while FOPPs are per aircraft.
- FOPPs do not have a strong level of dynamism, while UAT problems could be strongly dynamic.

3.3.2 Air Taxi

With many similarities to fractional ownership programs, on-demand air taxi offers regional [60,91] or urban [92] services. The major difference between the two is the ability of the air taxi operator to reject a request, and therefore, not incur a charter or upgrade cost. Additionally, fractional ownership programs offer per-aircraft service and non-stop flights. However, in air taxi operations, the operator could offer per-seat services. Consequently, in air taxi operations, the flights might be shared, there could be a transit stop, or the route may include an intermediate stop to pick up or drop off other passengers. If air taxi operation is per seat where requests with the same origin and destination are pooled together (i.e., air pooling), it resembles fractional ownership programs. On the other hand, if the service is per seat with intermediate stops, the framework is similar to DAFP.

Currently, passenger-carrying air taxi operations defined by the US Department of Transportation (DOT) include on-demand flights conducted with *small aircraft* (i.e., aircraft with 60 or fewer passenger seats and a maximum payload capacity of 18,000 lbs or less) [93]. Depending on the aircraft seating and payload capacity, air taxi operations are conducted under 14 CFR Part 135 on-demand or Part 121 supplemental operations of the FAA regulations [94,95]. Part 135 on-demand operations cover airplanes with 30 seats or fewer and 7,500 lbs payload or less or rotorcraft.

A single-entity charter (also known as a private jet charter or air charter) flight is a per-aircraft service in which an individual charters the entire aircraft on demand. A single entity charter would also apply to a case where a group of individuals self-aggregate and charter an aircraft as a single entity [96]. Therefore, single entity charters fall under air taxi operations for the fleet of aircraft that meet the seating and payload capacity specifications. Despite the DOT's definition of air taxi operations, some distinguish between per-aircraft air taxi and air charter, and view the air taxi business model as the less expensive option that utilizes a new generation of small aircraft compared to air charter [97,98].

On-demand UAM would fall under the air taxi operations currently defined by FAA and DOT. However, per-seat on-demand UAM service faces some legal challenges given the current regulation since the UAM operator cannot play any role in consolidating the demand [99]. Current per-seat air taxi operators circumvent these legal challenges by acting as web hosts where the *lead passenger* could form a group by notifying other members [99].

The studies on air taxis cover the market and demand, facility location problems, routing and scheduling, and system analysis using a simulation framework, each discussed in the following sections. Since no dominant business model or aircraft type for on-demand UAM has emerged,

studies have used various assumptions to describe or model on-demand UAM operations. Table 3.1 summarizes business models and operational assumptions of the service in the relevant studies.

Table 3.1 Business models and operational assumptions of on-demand UAM in the literature

Study	Cruise Speed (mph)	Detour Factor (%)	Boarding/Deboarding Time (min)	Vertical Ascend/ Descend Time (sec)	Ground Speed (mph)	Max Ingress/ Egress (mi or min)	No. Passenger Seats	Shared Flights	No. Intermediate Stops	Autonomy
Antcliff et al. [100]	120	-	-	114	21	0.66, 0.99 mi	1-2	Y	0	A, P
Holden and Goel (Uber Elevate) [23]	170	-	3 (B), 2 (D)	60 (T) 75 (L)	-	-	-	-	0	-
Porsche Consulting [6]	124	-	3	-	24.8	5 min	-	-	0	-
Goyal et al. (Booz Allen Hamilton) [8]	125	5-15	3-5 (B) 2-3 (D)	-	-	-	2-4	Y	0	P
Rothfeld et al. [101]	93	-	≤ 2.5	50	-	3.1 mi	2	-	-	-
Rajendran and Zack [92]	170	-	-	-	6	10 min	-	-	0	-
Rajendran and Shulman [102]	160	-	3 (B), 2 (D)	60 (T) 75 (L)	-	-	4	Y	0	-
Ale-Ahmad and Mahmassani [103]	150	10	3 (B), 2 (D)	45 (T) 45 (L)	20	~10 min	1-4	Y	0	A

Note: Dashes indicate that the item is either not applicable, not available, or both. B = boarding; D = deboarding; T = take-off; L = landing; Y = yes; A = autonomous; P = piloted

3.3.2.1 Market

Baik et al. [104] propose a Transportation Systems Analysis Model (TSAM), a four-step modeling process that could predict the demand for air taxis. Kreimeier et al. [105] study the feasibility of on-demand air mobility from an economic perspective. Depending on the aircraft's speed and the trip distance, they conclude that willingness-to-pay is in the range of 0.5-0.8 €$_{2015}$/km

for operations in 2030. Sun et al. [91] compare the competitiveness of the air taxi with car, rail, and traditional air transportation regionally in the area covered by 29 countries in Europe. Their goal is first to find the dominating mode of travel between air taxi and rail in 500 European cities, and next, to identify origin-destination (OD) pairs with high demands that could benefit the most from travel time savings.

Multiple market studies and stated preference surveys have focused on passenger UAM adoption and mode choice [8,9,38,106-109]. Booz Allen Hamilton's UAM market study [8] finds that respondents were more interested in using UAM for recreational trips or trips to airports than commuting. Younger male survey respondents who were already familiar with the concept of UAM and individuals with higher income were more inclined to use UAM. Additionally, respondents from Los Angeles were willing to pay nearly $0.85 more for one additional mile than Houston respondents.

Garrow et al. [108] administered a stated preference survey with 1,405 full-time workers with the minimum annual household incomes of $75K in Atlanta, Boston, Dallas-Fort Worth, San Francisco Bay Area, and Los Angeles in 2019. They report that early adopters are more likely to be frequent air travelers and frequent users of ride-hailing services. Boddupalli et al. [109] report that younger and male individuals who are frequent users of ride-hailing services and are characterized as tech-savvy are more likely to choose air taxis for commuting. In comparison, the market study with 248 respondents conducted by Fu et al. [39] in Munich suggests that the air taxi adoption rate will be higher among younger individuals (18–35 years old) and older respondents (56–65 years old) with high income. However, they did not identify any difference in the adoption rate based on gender.

Al Haddad et al. [38] develop a stated preference survey with a sample size of 221 to identify the factors that affect UAM adoption. They report that the adoption of UAM is highly influenced by travel time savings and perceived costs, which is in line with the findings of the UAM market studies by Booz Allen Hamilton [8], Fu et al. [39], and Boddupalli et al. [109] on the significance of time and cost in the passenger UAM mode choice.

Boddupalli et al. [109] study air taxi mode choice for commuting purposes in the early stages after launch using a stated preference survey. The air taxi mode in the study is battery-powered, piloted, includes no transfer, and has two to four passenger seats. The minimum annual individual income for the respondents is set to $100K. Among the 2,499 sampled respondents, 8 percent reported a one-way commute time of 90 minutes or more, while 27 percent and 19 percent reported a commute time of 40 to 59 minutes and 60 to 89 minutes, respectively. They report that the average values of in-vehicle travel time for transit, air taxi, and auto are $23.94, $26.38, and $28.21, respectively. Finally, when individuals are offered a guaranteed ride home with Uber or Lyft in the event the trip using transit or air taxi gets canceled, they are 1.8 times more likely to choose the said modes.

3.3.2.2 Facility Location

Keysan [110] studies the tactical level base location and fleet allocation problem for per-seat on-demand air transportation enabled by advances in Very Light Jets (VLJs). In the basic solution approach, the number of required jets at each location represents the demand. However, in the integrated solution approach, a more detailed model is developed, integrating the operational flight scheduling with the location problem.

Using New York City (NYC) Taxi and Limousine Commission data, Rajendran and Zack [92] estimate the potential demand for urban air taxi services based on travel time savings.

Furthermore, they use k-means clustering to identify 21 potential UAT pads or ports based on estimated demand distribution. They recommend that the UAT ports at the South Central Park and JFK International Airport should accommodate nearly 150 take-offs and landings per hour. Rath and Chow [111] formulate the UAT pad location problem as a hub location problem. The results suggest that at least 9 UAT pads in NYC are required to achieve at least 10% market penetration. Lastly, Fadhil [112] offers a GIS-based analysis to select the UAT pads.

3.3.2.3 Routing and Scheduling

Some studies use a simple policy for routing and scheduling of aircraft within the air taxi simulation framework and assign the nearest idle aircraft to the request [61,113]. Fagerholt et al. [59] employ an insertion heuristic coupled with a local search heuristic to solve the per-seat on-demand air taxi operations with no stops. The optimization-based studies focus on the DAFP [3,68,83-85,110,114,115]

Espinoza et al. [3] formulate the per-seat dial-a-flight problem as an integer multicommodity network flow model with capacity, weight, and time window constraints and solve instances with 8 aircraft and 81 requests using commercial software. They subsequently propose [114] a parallel local search scheme to solve instances with over 300 aircraft and over 2800 requests close to optimality. Engineer et al. [83] develop a relaxation-based dynamic programming algorithm for DAFP that, in combination with other techniques, could solve the column generation relaxation for cases with up to 200 aircraft and 1613 requests. La Foy [115], Campbell [84], and Reddy [85] present variations of DAFP to formulate regional air taxi operations in Southern Africa.

3.3.2.4 Simulation Framework

Simulation studies on the regional operations of on-demand air taxis are motivated by the emergence of VLJs in the 2000s. Bonnefoy [60] developed a simulation framework, called Air Taxi Network Simulator (ATNS), to duplicate the on-demand air taxi operations over one year on the regional level in the US using a fleet size of 25 to 100 aircraft and up to 780 airports. In this framework, passenger's request includes the earliest pick-up and the latest drop-off time, and the willingness of passengers to share a flight with others. Three heuristics for maintenance routing, pilot routing, and aircraft routing are devised. The objective function for aircraft routing and pilot assignment includes the ratio of non-revenue to revenue-generating mileage, idle time, and location of vehicles at the end of the day. Additionally, the demand for air taxis is generated using the gravity model.

Bonnefoy [60] further studies the impacts of demand intensity and network size. First, he shows that for a constant number of aircraft, an increase in demand intensity results in an increase in revenue-generating mileage. However, since the passengers have a desired time window for the flight, this increase in demand leads to more rejected requests. More importantly, he investigates the impact of demand intensity on system performance. For uniform distribution, the ratio of non-revenue to revenue-generating mileage is slightly below 0.3. As the demand intensity increases, this ratio decreases while the average number of flights per day increases. Second, he studies network size effects using four networks ranging from 400 to 700 miles around a metropolitan area using a fleet size of 75 aircraft. The results indicate that a bigger network results in higher average revenue mileage and fewer flights per day per aircraft. However, the ratio of non-revenue-to revenue-generating distances does not change. He concludes that the system's performance is more sensitive to demand intensity than the size of the network.

Boyd et al. [61] similarly investigate the on-demand regional air taxi service for one week on a small scale. They devise scenarios with 2-3 airports and 2 and 4 aircraft, while the city distances are 70 and 140 miles. Their simulation framework is developed using the Arena software package of Rockwell Software, which noticeably limits the modeler's flexibility. The requests arrive in real-time at each airport. When the number of passengers for a given destination reaches aircraft capacity or wait time for one passenger exceeds one hour, a request for a flight is placed. The assumptions regarding the demand limit the analysis. They disregard the spatiotemporal distribution of the demand. Demand is set as a constant for a given origin and does not change if more destinations are added for each origin. Furthermore, the demand is distributed uniformly over all destinations. No maintenance is assumed for the aircraft. Additionally, the aircraft are assigned to passengers based on the shortest distance.

The results suggest that customers' wait time increases by decreasing the fleet size or increasing aircraft capacity since more requests are needed to reach aircraft capacity. Additionally, increasing the number of airports increases the number of non-revenue-generating flights. However, the simulation logic makes an aircraft leave the gate after dropping off passengers without picking up the passengers currently waiting at the gate.

Lee et al. [116] compare the results of a discrete event simulation model with a flow model for air taxi operations. The flow model provides an aggregate model of air taxi operations in a medium-range planning horizon without explicit modeling of events such as passenger arrivals or non-revenue flights. The discrete event simulation framework is similar to Boyd et al. [61] with some modifications. Passengers arrive at each airport according to a non-homogeneous Poisson Process at a fare-dependent rate to control flow for each route. No maintenance or downtime is assumed, so the aircraft are in service 24/7. The framework assumes that the duration of the delay

and ground time of each flight are random variables. However, the numerical experiments assume zero delay and ground time. The aircraft are homogenous with a capacity of 4 and a speed of 345 mph. The passengers wait for a maximum of W^{max} before an aircraft is assigned to them based on the shortest-distance policy.

Fagerholt et al. [59] develop a simulation framework for the air taxi service in Norway. The operator does not have to accept all the incoming requests. However, if a request is accepted, the exact pick-up time should be announced at the time of booking. As a result, the operator seeks to obtain a solution to this dynamic problem in a very short time. They formulate the per-seat on-demand air taxi operations as a special case of DAFP with no intermediate stops, which is subsequently solved using an insertion heuristic coupled with a local search heuristic. For the insertion cost, a multi-criteria objective based on the flight cost and a measure, which estimates the probability of accepting future requests, is used. As a result, the acceptance and rejection scheme developed in this study only considers the feasibility of serving a new request and not its profitability.

An instance with 10 aircraft, 3 airports, and 200 requests over the planning horizon of one day is solved in 1 CPU minute. The authors also examine the impacts of several strategic decisions such as fleet size and booking policy on the number of accepted requests, time utilization, and distance utilization. Time utilization and distance utilization increase as the number of arrived requests increases. For arrived requests more than 50, the time utilization varies between 65 to 75%, while the distance utilization changes between 65 to 70%.

Rothfeld et al. [113] present an agent-based simulation framework as a passenger UAM extension of the multi-agent transport simulation, MATSim. In this framework, four events are defined for an eVTOL: staying put (i.e., being idle), pick-up passengers, drop-off passengers, and

flying. The customers could use a car, bike, walk, or public transportation to ingress or egress the aerial service. Each customer looks for a UAT pad within a prespecified radius of its origin and destination as the potential pick-up and drop-off UAT pads. Subsequently, the customer considers the cost and travel time associated with each combination of these UAT pads and their corresponding available ground-based modes. Therefore, the customer is the agent who chooses the pick-up and drop-off UAT pad.

Moreover, the operator receives the request for aerial service when the customer finishes its previous activity and sets out toward the pick-up UAT pad. If there are available UAT aircraft, the nearest one is assigned to the request. Otherwise, the request is placed in a queue to be assigned in future decision epochs. Lastly, the framework allows for requests to be pooled together as long as they have the same pick-up and drop-off UAT pads and reach the pick-up UAT pad before the scheduled boarding time of the first request.

In the following study, Rothfeld et al. [101] implement this framework in the Sioux Falls network with 10 UAT pads and a homogenous fleet of 100 UAT aircraft. The flight level for the cruise is set to 500 meters with a vertical speed of 10 m/s, translating to the ascend or descend duration of 50 seconds. They further define *ground-based UAM process time* as a process that includes elevator usage, security screening, or boarding the aircraft, and its duration is set to 2.5 minutes.

The results show that more than 75% of the flights are shorter than 3.1mi (or 5 km). The mean travel time for the aerial leg is 20 minutes. The mean total trip distance is 4.5 mi (or 7.2 km), which is on average 2.4 times longer than the straight-line distance between origin and destination of customers. While the maximum ingress or egress distance is set to 5 km (i.e., 3.1 miles), there is no constraint on the ingress/egress time. The lack of temporal constraints on the customer's

delay is the main drawback of the framework since it has caused unacceptable ingress and egress time. For instance, while the mean travel time for the aerial leg is 20 minutes, the mean ingress time is 71 minutes. Lastly, they conduct sensitivity analyses to the cruising speed, vertical speed, process time, aircraft capacity, fleet size, and the number of UAT pads.

Rajendran and Shulman [102] offer a per-seat multimodal concept of operations for UAT. In the proposed framework, if the UAT operator accepts the customer request, it provides the customer with the cost and duration of the service. Subsequently, the customer can reject the operator's offer for the aerial service. As the customers, who have chosen to use the service, arrive at the UAT pad, they enter a queue based on their destination. The customers will wait a maximum time of W^{max} before leaving the system without being served. A UAT aircraft would arrive at a UAT pad only when dropping off the passengers at that pad. Therefore, no empty flight is conducted to reposition the aircraft. However, in cases where the incoming demand at a given UAT pad is disproportionately lower than the outgoing demand, this assumption causes excessive wait time. Additionally, a customer in the queue should wait until the number of waiting customers reach the capacity of the aircraft, which in turn could lead to unacceptable wait times for OD pairs with low demands.

They develop a discrete-event simulation in SIMIO. There are 500 replications, and each simulates the air taxi operations for 30 days with 24 operating hours per day. The warm-up period is three hours. The numerical experiment includes 5 UAT pads with 60 homogeneous UAT aircraft. The take-off, landing, passenger loading, and passenger unloading durations are assumed 60, 75, 120, and 180 seconds, and the maximum wait time of a customer at the UAT pad before they leave is set to 20 minutes. The average daily demand for the air taxi in the base scenario is about 193,000, translating to an average interarrival time of 0.44 seconds. The aircraft utilization,

mean customer's total trip time, mean customer's wait time, and the average number of customers created per week are reported as the performance measures. However, it is not clear how many passengers are served. The results show that an average trip would take 40 minutes while the average passenger's waiting time is 15 minutes. Aerial fleet utilization is about 34%. Lastly, the sensitivity analyses examine the impacts of demand density, aircraft capacity, number of aircraft, and maximum wait time on the performance measures.

Table 3.2 summarizes the simulation studies on on-demand air mobility on the regional level.

Table 3.2 Comparison of simulation studies on on-demand air mobility

Study	Dynamic	Requests Rejection	Air Pooling (P) vs. Air Sharing (S)	Intermediate Stops	Demand Model	Planning Horizon	Number of Instances (Replications)	Number of Requests	Number of Aircraft	Number of Airports	Airports Distance (mi)	Model	Solution Method	Solver/ Programming Language	Solution Time (min)
Bonnefoy (2005)	N	Y	S	0	Gravity Model	1 Year	-	-	25-100	780	-	Aircraft Routing	Heuristic	-	-
Boyd et al. (2007)	Y	N	P	0	Stationary Poisson	1 week	10k	0.5, 4 /hour	2,4	2,3	<160	Policy	Shortest Distance	-	-
Lee et al. (2008)	Y	N	P	0	Non-stationary Poisson	1 week	1000	(i) 1 – 5 /hour, (ii) 50 & 200/day	-	5	<600	Policy	Shortest Distance	-	-
Fagerholt et al. (2009)	Y	Y	P	0	Random	1 day	100	Up to 200	5, 10, 15, 20	3	-	DAFP	Insertion Heuristic	JAVA	~ 1

3.4 Synchronized Logistics

Covering the first and last leg of the trip on the ground by the UAT operator in addition to the flight would require coordinated transportation between the ground-based and aerial modes. In the literature, this has been referred to *vehicle routing problem with multiple synchronization constraints (VRPMS)* [117]. These synchronizations could be about tasks, operations, movement, and load. Synchronization in operation concerning time could be broken into pure spatial operation synchronization, operation synchronization with precedence, and exact operation synchronization [117]. *Location-Routing Problems (LRPs)* could be classified under VRPMS, where locations and routes are determined simultaneously. In the following sections, LRPs and some applications of synchronized logistics are discussed.

3.4.1 Location-Routing Problems

Location-Routing Problem (LRP) [118] is a class of location problem where locations are planned while considering the aspects of tour planning, and therefore, the relation between these two decisions is taken into account. In the classic LRP, the locations of facilities are determined in conjunction with vehicle tours from these locations to cover the customers. There is a cost associated with opening each facility, and there are no vehicle tours between the facilities. In these problems, the problem of finding the locations is strategic, while the problem of finding the routes is tactical [118].

If the location and routing problems are solved *sequentially* without considering the interrelation of these two, the problem is no longer classified as LRP. In sequential approaches, first, the facility location problem is solved by minimizing the sum of distances between customers and the facility, and then the routing problem is solved based on the location of the facilities in the first step.

Three heuristic solution approaches for solving LRPs exist [118]: clustering-based, iterative, and hierarchical heuristics. In the *cluster-based approach*, the customers are grouped into clusters. Afterward, two possible methods exist: (i) for each cluster, the location of the facility is determined, and then VRP (or TSP in case of one vehicle) is solved for customers in each cluster, or (ii) TSP is solved for each cluster, and then the location of the facility (i.e., depot) is determined. Cluster-based heuristics are similar to sequential methods since there is no feedback between routing and location problems. However, the clustering is performed while considering some aspects of routing. As a result, cluster-based heuristics are classified under solution approaches of LRPs. In the iterative approach, the problem is divided into two consecutive subproblems, solved iteratively with feedback from one subproblem to another. In hierarchical heuristics, the main problem involves solving facility location while referring to a routing subproblem in each step.

There are problems related to LRPs that are not classified as classic LRP problems. In the following sections, the two variations that are relevant to UAT are discussed:

1. *Location-Allocation-Routing Problem (LARP)* or *Vehicle Routing-Allocation Problem (VRAP)*: This variation includes problems with vehicle routing just between facilities (i.e., hubs or depots) but not between facilities and customers. So, the customers are allocated to the facility simply with direct transport, and consequently, radial distances are most relevant.

2. *N-echelon Location-Routing Problem (LRP-NE)*: This variation includes problems with multi-level (or *echelon*) intermediary facilities where each echelon has its own vehicles that form tours to visit the facilities from the next echelon. These vehicles could be homogeneous or heterogeneous and have their own attributes, such as operating costs per mile, speed, and capacity. Additionally, tours at each echelon

could have different fixed costs, and the costs associated with opening/using/stopping at facilities could differ for each echelon.

Two-echelon location-routing problems (LRP-2Es) seek to determine the location of the facilities while considering tour planning between facilities and between customers. Some variations allow customer visits in the first-level routes while others do not [119]. The first study on LRP-2E was done by Jacobsen and Madsen [120] on a newspaper delivery system. In this system, the newspapers are first delivered to transfer points and then from transfer points (TP) to sales points (SP). The primary route is constructed between TPs, and the secondary route is constructed between a TP and SP. There is a capacity constraint for the vehicles performing primary tours. Additionally, secondary tours have the latest delivery time and tour duration constraints. The cost of each tour consists of the fixed cost, cost of stops, and the cost associated with tour length. In other words, Jacobsen and Madsen [120] seek to minimize the number of routes (i.e., required vehicles), the number of stops, and the length of the route.

It is worth mentioning that there is a subtle difference between *N-echelon Location Routing Problem (LRP-NE)* and *N-echelon Vehicle Routing Problem (VRP-NE)*. In *VRP-NE*, there is no fixed cost associated with opening or using a facility, and all locations are assumed to be open [119], while in *LRP-NE*, the cost of opening a facility is included in the objective function.

An analogy could be drawn between synchronized logistics of UAT (i.e., multimodal operation with ground-based first- and last-mile service) and the two variations of LRPs discussed above. In the UAT framework, the aircraft needs to visit the nodes corresponding to flights while the requests are allocated to these nodes, which is similar to the tour planning between intermediary

facilities in VARP or LRP-NE. In addition to the operating cost of the tour, there is a fixed cost associated with each tour (i.e., acquisition and maintenance cost of the aircraft) and a cost associated with each stop along the way (i.e., take-off and landing costs).

Since there are multiple UAT aircraft, each capable of forming one tour, multi-depot variations of LRPs are more applicable to the UAT operations. Additionally, the limited capacity of the aircraft calls for capacitated LRP. Finally, some variations of LRP-NE allow vehicles in the primary tour to visit the customer's node directly. These variations are more relevant to UAT operations.

For the ground-based operation of UAT, two conceptual frameworks are inspired by VARP and LRP-NE. In the first one, the requests are simply assigned to the flights, while in the second one, there is tour planning to cover the transportation of customers to the flights.

Despite all the similarities between synchronized logistics of UAT and LRPs, there are significant differences that make the two problems distinct. LRPs are developed in the context of freight transportation where mostly only deliveries (e.g., newspapers delivery [120]) or only pick-ups (e.g., milk collection [121]) are involved. However, UAT addresses pick-ups and drop-offs of the passengers. Aside from the difference in network representation, the boarding at the UAT pads should be synchronized such that the vehicles transferring passengers on the ground arrive at the pad before the aircraft does, or the aircraft should be held until the vehicle arrives. Either way, there should be a limit on the wait time of the passengers or aircraft holdings. For these reasons, LRPs should be adjusted to be able to model UAT fleet operations.

3.4.2 Truck and Trailer Routing Problem (TTRP)

Truck-and-Trailer Routing Problems (TTRPs) or *road-train problems* [122] address a class of problems where a truck and trailer attached together leave the depot to serve the customer demand. Due to accessibility restrictions, the trailer cannot visit all the customers. The stores (i.e., customers) that can be served by trailer-truck are called *trailer stores*, while the stores that can be served by only a truck are called *truck stores*. Therefore, trailer points [123] are defined where the trailer is detached from the truck, and the truck performs a tour to cover the truck stores. In the classic TTRP developed by Semet and Taillard [122], the trailer is detached from the truck at the trailer store. Therefore, the trailer points are selected among trailer stores. In a later variation, with milk collection in Norway, the trailer is detached at a parking place so that the truck could serve the farms. Therefore, no customer is served directly on the primary tour. TTRP models and the solution algorithms are surveyed in [119] and [124].

TTRP is a special case of a more general class of problems, i.e., *Vehicle Routing Problem with Trailers and Transshipments (VRPTT)* [125], with a fixed truck-trailer assignment, meaning that each truck is attached to one trailer, and the transfer of the load could happen only between this truck and the trailer at the *transshipment locations* [125].

TTRPs resemble synchronized logistics of the UAT because the vehicles have limited capacity, and there are time window constraints on the deliveries. Additionally, in the classic definition, the transfer points are located at the customers, which is similar to the operational frameworks of UAT.

3.4.3 School Bus Routing Problem (SBRP)

School Bus Routing Problem (SBRP) involves planning the routes for a fleet of school busses to pick up students from bus stops and drop them off at their destination (i.e., their school)

under various constraints such as the capacity of the bus and maximum ride time of students [126]. From the 27 studies reviewed by Park and Kim [126], only 6 of them consider the subproblem of defining bus stop location, from which only two studies take the maximum walking distance into account.

There are two heuristic approaches for bus stop location in SBRP [126]: the *location-allocation-routing (LAR)* strategy or the *allocation-routing-location (ARL)* strategy. In LAR heuristics, first, the locations of bus stops are determined, and the students are allocated to these locations. Next, the tour is planned for these selected bus stops. This approach is similar to *sequential methods* discussed under LRP. Since the bus stop locations are determined without taking the routing (for instance, the capacity of each vehicle) into account, LAR tends to generate excessive routes. In ARL heuristics, first, the students are partitioned into clusters while considering the vehicle's capacity, and then the location of bus stops within each cluster is determined. Next, the vehicle route is calculated for these selected bus stops. Finally, the students in each cluster are allocated to the stops. ARL approach is similar to *cluster-based methods* for solving LRP. Chapleau et al. [127] use clustering within the ARL framework for the bus stop location problem. Interestingly, their objective function in clustering the students involves minimizing the number of clusters (i.e., routes) instead of minimizing the total distance. This is most useful when vehicles' fixed cost dominates the costs associated with the operating mileage of the vehicle.

Bus stop location subproblem in SBRPs covers a class of problems where the demand is consolidated at intermediate facilities. However, it is different from LRPs since these intermediate facilities are neither located at customer locations nor are selected from a potential set of facilities. Additionally, in the SBRP framework, there is a limit on the maximum walking distance from the

selected bus stop, which resembles the constraint on the maximum ingress and egress time or distance of customers in the UAT operation. However, in SBRPs, the students share the same destination (i.e., school) and are either picked up or dropped off, but not both. The solution algorithms of SBRPs should be adjusted for UAT operations with demand consolidation to identify consolidation locations for pick-up and drop-off of passengers who have a sufficiently close origin and destination.

3.4.4 Other Applications of Synchronized Logistics

Synchronized logistics have other various applications, many of which are reviewed in [117]. The most prominent application is in hybrid transit, which aims to synchronize flexible demand-responsive service and fixed-route service. Another application involves synchronizing a moving truck and a drone in a continuous network [128].

3.5 Concluding Remarks

This chapter reviews the literature relevant to UAT operations. UAT is a nearly on-demand aerial service in a ubiquitous network UAT pads, and therefore, in many aspects, it resembles the ground-based on-demand service. As a result, this chapter first presents the literature on vehicle routing problems with pick-up and delivery. Subsequently, it reviews the literature specific to on-demand air mobility use cases, namely, the fractional ownership programs and the air taxi. Lastly, UAT is a multi-modal service, which requires synchronization between the ground and aerial modes. Consequently, this chapter further reviews the literature on synchronized logistics.

Given the current information and visions on passenger-carrying UAM and UAT use case presented in Chapter 2 and the relevant literature reviewed in Chapter 3, Chapter 4 defines the underlying UAT concept of operations in this research and outlines the problem we aim to address.

Chapter 4 Urban Air Taxi: Concept of Operations and Problem Definition

4.1 Overview

Urban Air Taxi (UAT), a subset of UAM, is a *ubiquitous* on-demand per-seat service that moves passengers in urban or suburban areas using groundbreaking aircraft [8,9]. UAT does not have fixed routes or regular schedules, distinguishing it from *air metro* [9] or *airport shuttle* [8], which are envisioned to operate on predetermined routes.

Despite all the excitement around UAT, the specifics of many aspects of operations, including the aircraft, infrastructure, network coverage, and sharing strategies, remain unclear. The sheer number of terms used when referring to UAT aircraft and infrastructure vouch for the nascent UAT concept. As a result, this chapter presents and introduces the concept of operations for the UAT operations studied in this dissertation. Accordingly, the problem is defined, and the corresponding assumptions are presented.

4.2 Urban Air Taxi Concept of Operations

This section discusses the concepts and assumptions related to the aircraft, UAT infrastructure, network coverage, flexible pads, repositioning flight legs, demand consolidation, and first- and last-mile trips.

4.2.1 Aircraft

The aircraft in this research is assumed to be autonomous and electric, and have vertical take-off and landing. Hence, the UAT aircraft are autonomous eVTOLs. The cruise speed of these aircraft is assumed to be 150 mph with at most four passenger seats. These specifications resemble the tiltrotor aircraft developed by Joby Aviation, which has a maximum cruising speed of 200 mph [18]. We also assume that aircraft batteries could be swapped quickly, i.e., in less than 5 minutes.

Therefore, the batteries could be changed while the passengers are boarding or deboarding the aircraft, eliminating the need for scheduling events specific to battery swapping.

4.2.2 Infrastructure

Even though UAT pads and UAT ports are envisioned to differ in landing, parking, charging capacity, and available resources, we assume all the infrastructure in the network has a sufficient number of UAT pads with full batteries and resources required for swapping the batteries. Consequently, the UAT aircraft do not need to be routed to specific UAT pads to change their batteries.

4.2.3 Network Coverage

UAM Coordination and Assessment Team (UCAT) outlines 6 UAM Maturity Levels (UMLs) in 3 states: initial, intermediate, and mature [41]. As the system evolves from the initial state to the mature state, the density of UAT pads increases to the point that in UML-6, *ubiquitous* UAT service with 10,000s of simultaneous operations and ad hoc landing sites is envisioned.

Having a network with a selected number of UAT pads would help the UAT operator to aggregate the demand, increase aircraft load factor, and consequently, improve the efficiency of the operation. However, it would limit the service to the users that are relatively close to the UAT pads. As the number of UAT pads increases, the coverage of UAT service increases. In the long term, when the density of UAT pads is high, the UAT operator could provide a point-to-point service. We refer to the first setting with a limited number of UAT pads in the initial and intermediate states, a *limited network*, and the second setting with a highly distributed network of UAT pads, a *ubiquitous network*. By definition, the UAT use case is associated with a ubiquitous network [9].

4.2.4 Flexible Pick-up and Drop-off Pads

From the UAT operator's standpoint, one of the significant drawbacks of the ubiquitous network is the potential for extremely short repositioning flight legs from the drop-off UAT pad of one passenger to the pick-up UAT pad of another passenger. Furthermore, the operator would have fewer opportunities to consolidate the demand and reduce operational costs. To address this issue, the operator designs the UAT service with *flexible pads* for the passengers, where passengers have a desired pick-up and drop-off UAT pad, but they are flexible and could be *relocated* (or *transferred*) to a location within an acceptable radius of their origin or destination for the aerial service. Let Δ^{ACCESS} denote the radius surrounding the origin or destination of requests within which passengers are willing to be relocated.

4.2.5 Repositioning Flight Legs

While the requests are being served in a ubiquitous network, some empty flight legs might be too short to justify the repositioning of the aircraft. Therefore, the UAT operator could benefit from the concept of flexible UAT pads and relocate the passengers on the ground to eliminate the empty flight legs shorter than a threshold. Let Δ^{EMPTY} denote the minimum required (straight-line) distance between two UAT pads to justify the repositioning leg.

4.2.6 Demand Consolidation

The Crown Consulting UAM market study commissioned by NASA envisions a limited potential market for UAT by 2030 [9]. However, the study asserts that the assumption of one passenger per trip in their model is one of the main barriers. Hence, the UAT profitability hinges on the aircraft load factor, and consequently, a notion of demand consolidation should be incorporated in the UAT operation.

There are two models of demand consolidation: *air sharing* and *air pooling*. Air sharing service resembles ground-based ridesharing operations, where a passenger might have one or multiple stops per trip while other passengers on the aircraft are being served. DAFP presents a framework for modeling the air sharing service. In comparison, air pooling provides a service where all passengers assigned to one flight leg board and deboard the aircraft at the same locations. These locations may differ from their desired pick-up and drop-off UAT pads.

Consequently, with air pooling, passengers with sufficiently close origins and destinations could be pooled by being picked up and subsequently dropped off at the same UAT pads. Figure 4.1 demonstrates the concept of *air pooling in a ubiquitous network with flexible pads*. In Figure 4.1(a), three flights must be conducted to serve the three corresponding requests in a ubiquitous network without flexible pads. However, in Figure 4.1(b) with flexible pads, the upper and lower requests are relocated on the ground to the desired pick-up and drop-off UAT pad of the request in the middle, and therefore, only one flight is conducted to serve all three requests.

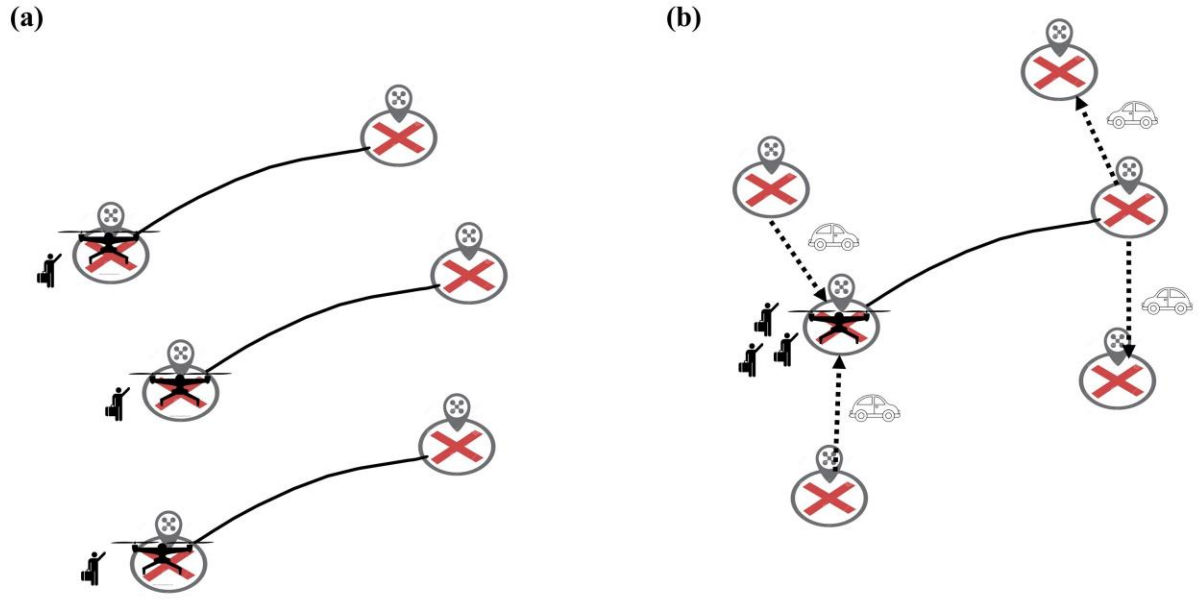

Figure 4.1 Concept of air pooling in a ubiquitous network with flexible pads, (a) without, and (b) with air pooling.

Air pooling takes advantage of the existing flight legs. As demonstrated in Figure 4.1, with air pooling, one request is served with its desired flight leg, and therefore, it undergoes no relocation, while the other experiences two relocations, one to the pick-up UAT pad and one from the drop-off UAT pad. To address this issue, the UAT operator could offer the service at a discounted rate to the requests that are not served with their desired flight leg. Nonetheless, in a more equitable setting, new flight legs should be defined between pick-up (or drop-off) of one request and drop-off (or pick-up) of another request. In this case, each request experiences one relocation.

The Booz Allen Hamilton's UAM market study [8] commissioned by NASA has reported the customers' concerns for the high cost of the service, multiple stops per passenger trip, and the relocations resulting from the ground-based transportation in the multimodal UAT operations [8]. Implementing air pooling in a ubiquitous network would reduce the operating costs by decreasing the aerial mileage. Additionally, the passengers would not have multiple stops. However, it adds a maximum of two relocations per trip compared to the point-to-point service without flexible pads. Lastly, Booz Allen Hamilton's UAM market study employs the air pooling business model, where all the passengers sharing one flight are picked up and subsequently dropped off at the same UAT pad.

4.2.7 First- and Last-Mile Service

The key to a successful UAT operation is a seamless multimodal operation where customers can smoothly access the UAT pad on the ground and use a ride-hailing service with negligible wait time [6]. Currently, BLADE offers ground-based transportation between the helicopter and the aircraft [45]. Furthermore, Uber Elevate announced that uberAIR would perform the first and last mile of trips with its ride-hailing service or walk [23]. Consequently,

UAT can be defined as a multimodal service where the UAT operator covers the ground-based legs of the trip in addition to the aerial leg.

Figure 4.2 depicts the concept of multimodal UAT service, where the UAT operator is informed of the request while they are at their origin. This enables the UAT operator to assign the pick-up UAT pad (i.e., starting UAT pad of the flight) and the drop-off UAT pad (i.e., ending UAT pad of the flight), and schedule the aircraft while considering the availability of its fleet and the congestion of the UAT pads. In contrast, when passengers are not flexible in the pick-up and the drop-off UAT pads and place a request only when they reach their desired pick-up UAT pad, the operator loses some of the flexibility it would otherwise have to operate the system more efficiently.

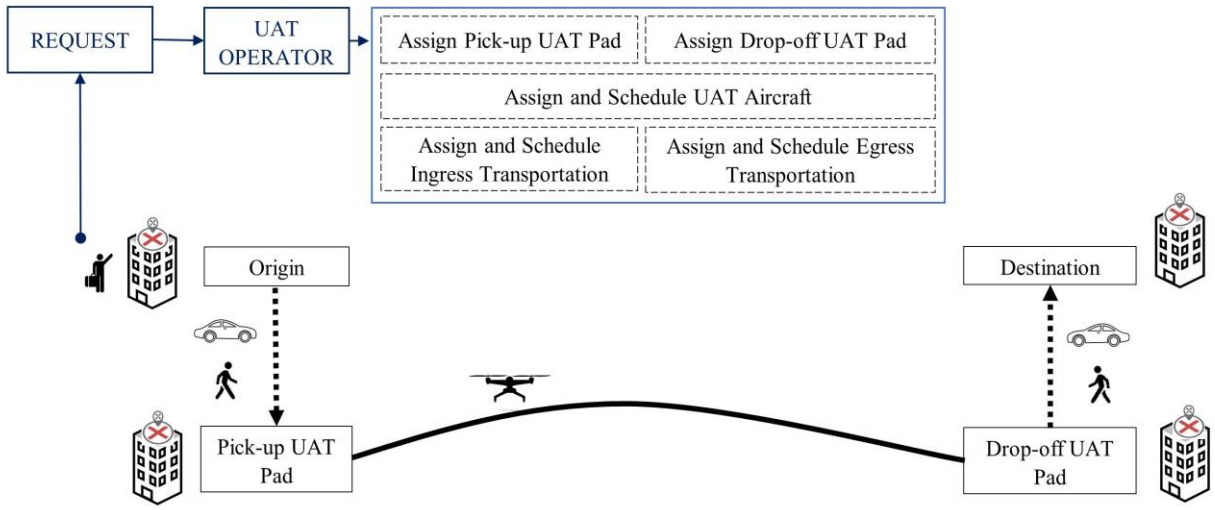

Figure 4.2 Concept of multimodal UAT operation

4.3 Problem Definition

The UAT operator manages a fleet of homogenous vehicles on the ground and UAT aircraft in a ubiquitous network (i.e., UML-6), and it synchronizes these two modes to serve the customer requests. Each request is identified by the origin, destination, desired pick-up and drop-off UAT

pads, requested time to begin the service, and group size. Given that the UAT pads are ubiquitously present in a ubiquitous network, the origin and destination of the request coincide with the desired pick-up and drop-off UAT pad, respectively.

Each passenger group is flexible in their pick-up and drop-off UAT pads and could be relocated on the ground within a reasonable radius from their origin or destination, which enables the UAT operator to eliminate the short repositioning flight legs in the ubiquitous network by relocating the passengers over short distances. Additionally, the customers are willing to share a UAT aircraft with other passengers, and consequently, the UAT operator could relocate the passengers to consolidate the customer requests and increase the aircraft load factor. As a result of short repositioning elimination and demand consolidation, each request trip consists of at most three legs: two ground-based legs and one aerial leg. The aerial leg is non-stop, and there is no intermediate transfer point for passengers to change the UAT aircraft after boarding.

The UAT operator is unaware of all future requests at the beginning of the planning horizon, and the customer requests for service arrive in real-time. As a result, the UAT operator updates its dispatching plan through a sequential decision-making process. The UAT operator may have a relatively short period for decision-making, particularly if there is no reservation scheme and requests are expected to be served immediately. The acceptance or rejection decision of the arrived requests is made at the first decision epoch after their arrival and will not change in the subsequent decision epochs. In other words, while considering accepting a new request, the UAT operator cannot reject the requests accepted in the previous decision epochs. However, the flight legs assigned to the accepted requests (and therefore, the pick-up and drop-off UAT pads) could change as long as the customers have not left their origin. After leaving the origin for the pick-up

UAT pad, the pick-up UAT pad of the request is fixed, and only its boarding time could be rescheduled.

The UAT competitive advantage is the travel time savings. As a result, if the UAT operator chooses to serve a request, the trip delay (i.e., deviation of the request's total trip time from its desired trip time) cannot exceed a prespecified value, which in turn limits the wait time for the aerial service, the ingress and egress time, and the deviation from the desired flight leg. Additionally, the UAT operator determines when the customers should leave their origin to access the pick-up UAT pad. The scheduling is designed so that the customer wait time is mainly incurred at the origin rather than the pick-up UAT pad. However, some customers may have to wait at the departure gate after a schedule change resulting from the new information (e.g., the arrival of new requests). Since the assigned flight legs to the requests and their schedules could be updated multiple times, the requests are only provided, in advance, with the time window during which they will be prompted to leave their origin or board the aircraft, not the exact time.

The revenue that the UAT operator earns from serving a request is proportional to the distance between the origin and destination of that request and its group size. The UAT operator incurs a fixed cost per flight and a variable cost proportional to the aircraft mileage. Ultimately, the UAT operator seeks a strategy that maximizes its net profit given the capacity, delay, and synchronization constraints. This strategy should address request acceptance and rejection, allocation of accepted requests to flight legs, and the sequence that UAT aircraft should serve these flight legs. It should further handle the flight scheduling, the boarding time of each passenger group, and the time by which the passenger groups need to leave their origin.

4.4 Assumptions

The analysis of UAT fleet operation in this research is based on the following assumptions across four categories of UAT service, operational policies, fleet, and customer requests:

4.4.1 UAT Service

- The UAT service is the envisioned use case of passenger UAM in the mature state (i.e., UML-6), and therefore, the UAT pad network is ubiquitous.
- The UAT service is nearly on-demand, and short advance reservation windows are allowed.
- The UAT service is per seat. Air pooling, where passengers share an aircraft as long as they are picked up and subsequently dropped off at the same UAT pads, is envisioned.
- The empty repositioning flight legs need to satisfy a minimum-mileage constraint, and therefore, short repositioning flight legs are eliminated.
- The UAT operator conducts the first and last mile of the trip on the ground.

4.4.2 Operational Policies

- The UAT operator guarantees a predefined level of service for the accepted requests. To this end, the trip delay (i.e., deviation of the passenger's total trip time from the desired trip time) cannot exceed a prespecified time. Consequently, the accepted requests are provided with the time window during which they will be prompted to leave their origin or board the aircraft.
- Each flight leg in a ubiquitous network is defined to serve a specific request. As a result, if an *intended request* of a flight leg is not assigned to that flight, the flight leg will not be served.

- The operator could either accept or reject the requests and does so in a prespecified period. As a request gets accepted, it is guaranteed to be served regardless of the arrival of new requests in the future.
- The requests could be reassigned to a new flight leg or another aircraft as long as the passengers have not left their origin. When a passenger group leaves its origin to the pick-up UAT pad, its pick-up UAT pad can no longer change.
- The passenger group of requests could be relocated on the ground to eliminate a short repositioning leg or to consolidate the demand. However, to implicitly minimize the number of relocations and avoid searching the entire space with ubiquitous UAT pads, the location they are being relocated to should be the desired pick-up or drop-off UAT pad of another request. In other words, the passengers cannot be relocated to an intermediary UAT pad in the space, which is not the desired pick-up or drop-off UAT pad of any other request.
- There is no intermediate transit stop. In other words, passengers do not change their aircraft in an intermediate UAT pad.
- Each passenger trip includes only one flight leg, and therefore, the passengers do not have multiple stops while being on board.
- The UAT aircraft do not have a hub, and they are spread throughout the network at the beginning of the planning horizons and do not need to go back to a specific UAT pad at the end of the planning horizon.

4.4.3 Fleet

- The fleet is homogeneous, and thus, the speed and capacity in each class of the aerial and ground-based vehicles are identical.
- UAT aircraft is fully autonomous, and therefore, no crew scheduling is necessary.

- The UAT aircraft takes advantage of eVTOL technology.
- The aerial fleet does not have an assigned UAT port. They are assumed to be randomly located in the network at the beginning of the planning horizon, and they do not need to go back to a UAT port at the end of the planning horizon.
- The time required for swapping the aircraft battery is short, and the battery could be swapped while the passengers are deboarding the aircraft. As a result, no charging slots are scheduled.
- No maintenance operations are scheduled, and therefore, the aircraft are available over the planning horizon.

4.4.4 Customer Requests

- Customer requests are known by their origin, destination, desired pick-up and drop-off UAT pads, and group size. In the ubiquitous network, the desired pick-up and drop-off UAT pads are, respectively, the same as the origin and destination.
- Customers are flexible in pick-up and drop-off UAT pads for aerial service as long as the ingress and egress time is reasonable.
- The number of passengers in each request is smaller than the aircraft's capacity. Additionally, the passengers in one group do not split.

4.5 Limitations

The modeling assumptions in the previous section have some drawbacks:

- The pricing scheme is per seat and does not provide a discount for shared flights, which does not encourage the customers to have a bigger group size. However, Booz Allen

Hamilton's UAM market study shows that passengers are willing to share the flight with passengers they do not know as long as they receive a discount [8].

- It is assumed that every UAT port has enough capacity for UAT operations, and therefore, the congestion at the UAT ports is not considered.
- No maintenance event is scheduled.
- Selecting an intermediary UAT pad in the ubiquitous network of UAT pads, as opposed to the one that is the desired UAT pad of a request, may reduce the aerial mileage.
- The requests are not provided with the exact boarding time, and there is no upper bound on the number of times a request could be rescheduled.

4.6 Concluding Remarks

UAT embodies the passenger-carrying UAM in its mature state, and therefore, the concept of operations has not been clearly specified. At the same time, no dominant player in the industry has yet emerged. Consequently, this chapter first discusses the aircraft, infrastructure, and network coverage associated with the UAT service. It further proposes the concept of flexible UAT pads for UAT service design in a ubiquitous network, suggesting that the passengers are flexible towards their pick-up and drop-off UAT pads within a reasonable distance of their origin and destination. This idea allows the UAT operator to move the passengers for two purposes: eliminating the short repositioning flight legs and consolidating the demand.

Based on the proposed concept of operations for the UAT service, the problem is defined, and the corresponding assumptions are outlined. Chapter 5 presents the Urban Air Taxi modeling framework.

Chapter 5 Urban Air Taxi Model

5.1 Overview

The Urban Air Taxi (UAT) operation has many elements that work together to provide aerial service. However, including and tracking all these elements are not essential to modeling the UAT fleet operation. Consequently, this chapter presents the relevant components of the UAT fleet operation, namely, entities, state variables, events, activities and delays, and transition functions. To this end, we borrow some of the concepts and terms used in the simulation literature. However, the model is defined independently of a simulation framework and, therefore, could be employed to model real-time UAT operations.

5.2 UAT Entities

Entities are discrete components that require explicit representation in the model [129]. They flow through the system and have attributes. Entities are classified into *objects* and *agents* [130]. Objects are passive entities that do not have intelligence. In contrast, agents are active entities that take actions, interact, learn, and adapt. In other words, agents are objects with attitudes [130]. The UAT model involves four entity classes: dispatcher, customer request, UAT aircraft, and flight leg; among which the last three are considered objects. The entities are discussed in further detail as follows.

5.2.1 Dispatcher

The dispatcher is a centralized unit that receives the customer request, accepts or rejects them, assigns the accepted requests to UAT aircraft, and schedules the ground and aerial fleet to serve the request. At each decision epoch, the dispatcher has the information on the status and location of customers and the status and location of its fleet, and it uses the available information

at the time to manage the system. As a result, the dispatcher wishes to solve a static and deterministic model of the dynamic and stochastic UAT problem at each decision epoch.

The principal component associated with the dispatcher is the policy it uses to assign the requests to aircraft and routing and scheduling of the aerial fleet. These policies could be rule-based such as first-come-first-served (FCFS) [74], optimization-based [78], or a hybrid approach that employs a combination of rules and optimization [76,131]. The *Capacitated Location-Allocation-Routing Problem with Time Windows And Short Repositioning Elimination (CLARPTW-SRE)* is the proposed optimization-based policy in this dissertation and is defined in Chapter 8.

5.2.2 UAT Aircraft

The UAT operator employs K UAT aircraft for the aerial service. Let $\mathcal{K} = \{a_1, a_2, \dots, a_k, \dots, a_K\}$ denote the set of functioning UAT aircraft that the operator could dispatch over the planning horizon, where $K = |\mathcal{K}|$. At the beginning of the planning horizon, the availability location and time of these aircraft are known. The static attributes of $a_k \in \mathcal{K}$ are represented by $\mathbb{A}_k^{eVTOL} = (Q_k, v_k^{AIR})$, where:

Q_k = capacity of aircraft k. With a homogenous fleet of aircraft, Q denotes the capacity of aircraft;

v_k^{AIR} = cruising speed of aircraft k. v^{AIR} denotes the cruising speed of a homogeneous fleet of aircraft.

5.2.3 Flight Leg

Flight legs are the constituent of a UAT aircraft itinerary. A flight leg could be either an *empty* (also known as *deadhead*, *ferry leg*, or *non-revenue* [90]) flight or a *revenue-generating* leg.

The latter is associated with flights that move passengers, while the former refers to repositioning flights that relocate a UAT aircraft to the pick-up UAT pad of a request. In aircraft routing, the sequence of flight legs is determined, and subsequently, the routes are assigned to the aircraft. The scheduling, on the other hand, determines the time that each flight should start.

Let $\not{f}_i$ denote flight leg i. The static attributes $\mathbb{A}_i^{LEG} = (\boldsymbol{S}_i, \boldsymbol{E}_i, H_i)$ of a candidate flight leg i must be available to the UAT operator. $\boldsymbol{S}_i$ is the starting point (i.e., UAT pad) of $\not{f}_i$, $\boldsymbol{E}_i$ is the ending point of $\not{f}_i$, and H_i indicates whether the flight type is empty or revenue-generating:

$$H_i = \begin{cases} 0 & \text{empty or deadhead} \\ 1 & \text{revenue-generating} \end{cases}$$

A revenue-generating flight leg has three additional attributes. $\mathbb{A}_i^{REVLEG} = \left(\mathscr{r}_i^{INTND}, \tau_i^{MIN}, \tau_i^{MAX}\right)$ denotes the static attributes of $\not{f}_i$ where $H_i = 1$. In a ubiquitous network, each revenue-generating flight leg is created with the intention of serving a specific request. Consequently, $\mathscr{r}_i^{INTND}$ denotes the intended request of revenue-generating $\not{f}_i$. Moreover, τ_i^{MIN} and τ_i^{MAX} are, respectively, the earliest and latest time that flight leg i could be served. τ_i^{MIN} and τ_i^{MAX} are defined in Equations (7.26) and (7.27), respectively.

Lastly, let $\mathbb{F}_i(\boldsymbol{S}_i, \boldsymbol{E}_i, \mathscr{r}_i^{INTND})$ denote a function that defines $\not{f}_i$ such that it starts at $\boldsymbol{S}_i$ and ends at $\boldsymbol{E}_i$, with intended request $\mathscr{r}_i^{INTND}$.

5.2.4 Customer Request

The *customer request* or *request*, in short, represents a group of passengers who wish to travel together from their origin to destination and request the multi-modal UAT service. A request is said to *arrive* in the system when it places its request for UAT service and becomes known to

the UAT operator. The accepted requests are assigned to flight legs, and their boarding time and the time they need to leave their origin are outputs of the dispatching strategy.

Let $\mathcal{r}_r$ denote request *r*. When $\mathcal{r}_r$ arrives at time τ_r^{ARV}, its attributes are defined by the vector $\mathbb{A}_r^{REQ} = (\boldsymbol{O}_r, \boldsymbol{D}_r, \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_r^{DSRD}, q_r, \tau_r^{REQ})$, where:

τ_r^{ARV}: the time $\mathcal{r}_r$ arrives;

$\boldsymbol{O}_r$: origin of $\mathcal{r}_r$;

$\boldsymbol{D}_r$: destination of $\mathcal{r}_r$;

$\boldsymbol{S}_r^{DSRD}$: the desired pick-up UAT pad of $\mathcal{r}_r$;

$\boldsymbol{E}_r^{DSRD}$: the desired drop-off UAT pad of $\mathcal{r}_r$;

q_r: the group size of $\mathcal{r}_r$. In other words, each *request* r consists of q_r *passengers*.

τ_r^{REQ}: the requested time for service by $\mathcal{r}_r$;

Consequently, given the desired pick-up and drop-off UAT pads of request r, the UAT operator defines $\mathcal{f}_r^{DSRD}$, the desired flight leg of request r, as $\mathcal{f}_r^{DSRD} = \mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, \mathcal{r}_i^{INTND} = \mathcal{r}_r)$. In a ubiquitous network, $\boldsymbol{S}_r^{DSRD} = \boldsymbol{O}_r$ and $\boldsymbol{E}_r^{DSRD} = \boldsymbol{D}_r$ since the UAT pads are ubiquitously present in the space; however, the UAT model and operational policy

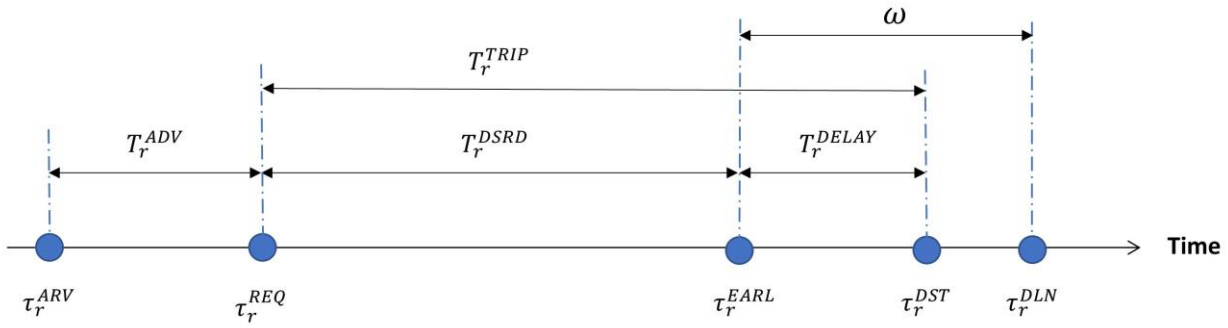

Figure 5.1 Illustration of temporal elements associated with request r

presented in this research are not dependent on this assumption and, therefore, they could be adjusted for a network with limited UAT pads.

The tuple $(\tau_r^{ARV}, \mathbb{A}_r^{REQ})$ represents the static information associated with r_r. Furthermore, Figure 5.1 illustrates the temporal elements related to r_r, where:

T_r^{ADV}: advance reservation time for request r, which is specified by the difference between the arrival time of r_r and its requested time of service (i.e., $T_r^{ADV} = \tau_r^{REQ} - \tau_r^{ARV}$);

T_r^{DSRD}: the trip time of the desired flight leg of r_r. It is equal to the total trip time of request *r* when the trip starts immediately at τ_r^{REQ} and the passenger group of r_r board the aircraft at their desired pick-up UAT pad (i.e., $\boldsymbol{S}_r^{DSRD}$) and deboard at their desired drop-off UAT pad (i.e., $\boldsymbol{E}_r^{DSRD}$) without any ground-based transportation; T_r^{DSRD} is defined in Equation (7.28).

T_r^{TRIP}: total trip time of each passenger in r_r, including ingress and egress time, aerial wait time, and aerial service time. $T_r^{TRIP} = \tau_r^{DST} - \tau_r^{REQ}$, where τ_r^{DST} is the time the passenger group of r_r reach their destination.

T_r^{DELAY}: the total delay experienced by a passenger of r_r, defined as the deviation of the trip time of r_r from the desired trip time (i.e., $T_r^{DELAY} = T_r^{TRIP} - T_r^{DSRD}$);

Consequently, τ_r^{EARL} denotes the earliest time the group passenger of r_r could reach their destination. Let ω represent the maximum allowed delay. The value of ω is prespecified by the operator as a proxy for the level of service. Lastly, τ_r^{DLN} is the latest time by which the UAT operator guarantees the passenger group of r_r would reach their destination. τ_r^{EARL} and τ_r^{DLN} are defined in Equations (5.1) and (5.2), respectively.

$$\tau_r^{EARL} = \tau_r^{REQ} + T_r^{DSRD} \quad (5.1)$$

$$\tau_r^{DLN} = \tau_r^{EARL} + \omega = \tau_r^{REQ} + T_r^{DSRD} + \omega \quad (5.2)$$

5.3 State Variables

While many attributes of the entities change over time, not all are required to study the UAT fleet operation. State variables track the values of attributes of interest and provide the information required for describing the system at any time. From the decision-making perspective, "A state variable is the minimally dimensioned function of history that is necessary and sufficient to compute the decision function, the transition function, and the utility function" [132].

Let S_t represent the state of the system at time t. S_t is defined by using the state of three entities, namely, requests, UAT aircraft, and flight legs. As a result, S_t = ($S_t^{REQ}, S_t^{eVTOL}, S_t^{LEG}$), where S_t is the state of the system at time t, and S_t^{REQ}, S_t^{eVTOL}, and S_t^{LEG}represent, respectively, the state of requests, UAT aircraft, and flight legs at time t. These state variables remain constant unless a relevant event in the future prompts their values to change. The state variables associated with requests, UAT aircraft, and flight legs are discussed in the following sections.

5.3.1 Customer Request

S_t^{REQ} specifies the state of all the requests that have been placed by time t, i.e., $S_t^{REQ} = \left(S_{rt}^{REQ}\right)_{r_r \in \mathcal{R}_t^{ARV}}$, where $\mathcal{R}_t^{ARV}$ denotes all the requests arrived by time t. $S_{rt}^{REQ} = \left(\varsigma_{rt}^{REQ}, \varphi_{rt}, \tau_{rt}^{ORG}\right)$, where ς_{rt}^{REQ} is the status of request r at time t, φ_{rt} is the flight leg assigned to request r as of time t and τ_{rt}^{ORG} is the time passenger group of request r are scheduled to leave their origin as of time t. ς_{rt}^{REQ}, φ_{rt}, and $\tau_{rt}^{ORG}$are defined in Equations (5.3)-(5.5), respectively.

$$\varsigma_{rt}^{REQ} = \begin{cases} -1 & \text{rejected} \\ 0 & \text{waiting for acceptance} \\ 1 & \text{accepted} \\ 2 & \text{waiting for service} \\ 3 & \text{en-route to pick-up UAT pad} \\ 4 & \text{ingression to the departure gate} \\ 5 & \text{waiting for boarding} \\ 6 & \text{boarding} \\ 7 & \text{on-board} \\ 8 & \text{deboarding} \\ 9 & \text{egression from the arrival gate} \\ 10 & \text{waiting for ground transportation to destination} \\ 11 & \text{en-route to destination} \\ 12 & \text{reached destination} \end{cases} \tag{5.3}$$

The status of request r at time t, ς_{rt}^{REQ}, could take 14 values, as defined in Equation (5.3). When $\mathscr{r}_r$ is accepted, a flight leg is assigned to it. This initial assignment could change from acceptance until the passengers start the boarding process. Let $\tilde{\varphi}_{rt}$ denote the flight leg assigned to request r at time t when $\varsigma_{rt}^{REQ} \in \{1, \dots, 5\}$. The assigned flight leg will remain unchanged afterward. Thus, let $\bar{\varphi}_r$ represent the flight leg that the passenger group of request r takes to reach their destination. To summarize, the value of φ_{rt} is defined in Equation (5.4).

$$\varphi_{rt} = \begin{cases} \text{N\A} & \varsigma_{rt}^{REQ} \in \{-1, 0\} \\ \tilde{\varphi}_{rt} & \varsigma_{rt}^{REQ} \in \{1, \dots, 5\} \\ \bar{\varphi}_r & \varsigma_{rt}^{REQ} \in \{6, \dots, 12\} \end{cases} \tag{5.4}$$

Additionally, when $\mathscr{r}_r$ is accepted and a flight leg is assigned to it, the time that its passengers should leave their origin to reach the pick-up UAT pad is also scheduled. This initial schedule could change from the acceptance until the time passenger group of $\mathscr{r}_r$ leaves their origin. Let $\tilde{\tau}_{rt}^{ORG}$ denote the *scheduled* time that the passenger group of $\mathscr{r}_r$ should leave the origin as of time t for $\varsigma_{rt}^{REQ} \in \{1, 2\}$. After the passengers leave the origin, this value will remain constant.

Thus, let $\bar{\tau}_r^{ORG}$ denote the *realized* time that passengers of r_r left their origin. To summarize, the value of τ_r^{ORG} is defined in Equation (5.5).

$$\tau_{rt}^{ORG} = \begin{cases} \text{N\A} & \varsigma_{rt}^{REQ} \in \{-1, 0\} \\ \tilde{\tau}_{rt}^{ORG} & \varsigma_{rt}^{REQ} \in \{1, 2\} \\ \bar{\tau}_r^{ORG} & \varsigma_{rt}^{REQ} \in \{3,4, \dots, 12\} \end{cases} \tag{5.5}$$

Furthermore, let τ_{rt}^{SRVC} denote the start of service time of r_r as of time t, and therefore, $\tau_{rt}^{SRVC} \geq \tau_r^{REQ}$. For the requests that their passenger groups have not left their origin, τ_{rt}^{SRVC} represents the earliest time passenger group of r_r could leave their origin. On the other hand, for requests that have already left their origin, τ_{rt}^{SRVC} is the actual time that providing the service to r_r starts. Therefore, the value of τ_{rt}^{SRVC} is equal to the time the passenger group of request r left the origin (i.e., $\bar{\tau}_r^{ORG}$). As a result, τ_{rt}^{SRVC} is defined in Equation (5.6), where W_{rit}^{INGRS} is the wait time for ground-based transportation to the starting point of flight i from the origin of request r if the were to served at time t. However, the wait time for walking is zero, and we could reasonably assume that if the reservation window is sufficiently wide, the wait time for ride-hailing will be zero as well.

$$\tau_{rt}^{SRVC} = \begin{cases} N\backslash A & \varsigma_{rt}^{REQ} = -1 \\ \max\left(\tau_r^{REQ}, t\right) + W_{ri,\max\left(\tau_r^{REQ}, t\right)}^{INGRS} & \varsigma_{rt}^{REQ} \in \{0,1,2\} \\ \bar{\tau}_r^{ORG} & \varsigma_{rt}^{REQ} \in \{3,4,..,12\} \end{cases} \tag{5.6}$$

5.3.2 Flight Leg

S_t^{LEG} represents the state of all the flight legs available flight legs by time t. Therefore, $S_t^{LEG} = \left(S_{it}^{LEG}\right)_{\forall i \in \mathcal{F}_t^{CAND}}$, where $\mathcal{F}_t^{CAND}$ is the set of candidate flight legs for assignment as of time t and $S_{it}^{LEG} = \left(\varsigma_{it}^{LEG}, \tau_{it}^{STRT}\right)$, where ς_{it}^{LEG} is the status of flight leg i at time t and τ_{it}^{STRT} is the

scheduled start time of flight leg i as of time t. The status of flight leg i at time t is specified by ς_{it}^{LEG}, which could take four values as specified in Equation (5.7).

$$\varsigma_{it}^{LEG} = \begin{cases} 0 & \text{unassigned} \\ 1 & \text{waiting for service} \\ 2 & \text{in service} \\ 3 & \text{served} \end{cases} \tag{5.7}$$

At the end of decision epochs, the start time of flight legs that should be conducted are determined. As long as a flight leg has not started, its scheduled start time could change. Let $\tilde{\tau}_{it}^{STRT}$ be the *scheduled* start time of flight leg i for $\varsigma_{it}^{LEG} = 1$ (i.e., waiting for service). Clearly, the start time of flight leg i will be no longer changed after the flight leg starts the service. Therefore, $\bar{\tau}_{i}^{STRT}$ presents the *realized* start time of flight leg i for $\varsigma_{it}^{LEG} \in \{2,3\}$. In summary, the value of τ_{it}^{STRT} is defined in Equation (5.8).

$$\tau_{it}^{STRT} = \begin{cases} \text{N\A} & \varsigma_{it}^{LEG} = 0 \\ \tilde{\tau}_{it}^{STRT} & \varsigma_{it}^{LEG} = 1 \\ \bar{\tau}_{i}^{STRT} & \varsigma_{it}^{LEG} \in \{2,3\} \end{cases} \tag{5.8}$$

When the start time of a flight leg passes, the corresponding status of the flight is either in service or served. In other words, $\varsigma_{it}^{LEG} \in \{2,3\}$ implies $t \geq \tau_{it}^{STRT}$. For $\varsigma_{it}^{LEG} = 0$, τ_{it}^{STRT} is undefined, implying that no start time is scheduled for an unassigned flight. On the contrary, when $\varsigma_{it}^{LEG} = 1$, τ_{it}^{STRT} has been set, but it could be updated in future decision epochs. Therefore, for $t \neq t'$, it is possible that $\varsigma_{it}^{LEG} = \varsigma_{it'}^{LEG} = 1$ while $\tau_{it}^{STRT} \neq \tau_{it'}^{STRT}$.

5.3.3 UAT Aircraft

The state of the UAT aircraft at time t is specified by the vector $S_t^{eVTOL} = (\varsigma_{kt}^{eVTOL}, \tau_{kt}^{AVL}, \boldsymbol{L}_{kt}^{AVL}, \mathcal{Q}_{kt}, \mathfrak{E}_{kt}^{NDSRD})_{k \in \mathcal{K}}$, where:

ς_{kt}^{eVTOL}: the status of the UAT aircraft k at time t;

τ_{kt}^{AVL}: the earliest time the subsequent itinerary of UAT aircraft k could be modified, and therefore, the time aircraft k becomes available for the future service as of time t;

$\boldsymbol{L}_{kt}^{AVL}$: the location of UAT aircraft k at τ_{kt}^{AVL}; and

$\mathcal{Q}_{kt}$: the ordered list of non-completed flight legs assigned to UAT aircraft k as of time t;

$\mathfrak{E}_{kt}^{NDSRD}$: a binary variable, which is 1 if $\boldsymbol{L}_{kt}^{AVL}$ is the drop-off UAT pad of the passengers, but not their desired one, 0 in any other case.

The status of eVTOL k at time t, ς_{kt}^{eVTOL}, could take seven values, as defined in Equation (5.9).

$$\varsigma_{kt}^{eVTOL} = \begin{cases} 0 & \text{idle} \\ 1 & \text{boarding} \\ 2 & \text{take-off clearance} \\ 3 & \text{in flight} \\ 4 & \text{landing clearance} \\ 5 & \text{deboarding} \\ 6 & \text{holding} \end{cases} \tag{5.9}$$

$\mathcal{Q}_{kt} = \{q_{kt1}, \dots, q_{ktn}, \dots, q_{kt|\mathcal{Q}_{kt}|}\}$, where q_{ktn} is the nth leg on $\mathcal{Q}_{kt}$, is the ordered list of non-completed flights (i.e., scheduled itinerary) of UAT aircraft k as of time t, and is subject to change at each decision epoch. If $\mathcal{Q}_{kt} = \emptyset$, then $\varsigma_{kt}^{eVTOL} = 0$, meaning that an idle UAT aircraft has no non-completed flight leg assigned to it. For $\mathcal{Q}_{kt} \neq \emptyset$, let $\not{f}_i = q_{kt1}$ represent the first flight leg on $\mathcal{Q}_{kt}$. If $\varsigma_{it}^{LEG} = 2$, flight leg i is the flight leg currently performed by aircraft k as of time t. As a result, $\not{f}_{kt}^{CRNT}$ denotes the flight leg in service by aircraft k as of time t. $\not{f}_{kt}^{CRNT}$ is specified in Equation (5.10).

$$\mathbb{f}_{kt}^{CRNT} = \begin{cases} q_{kt1} & q_{kt1} = \mathbb{f}_i \text{ and } \varsigma_{it}^{LEG} = 2 \\ N\backslash A & \text{otherwise} \end{cases} \tag{5.10}$$

Additionally, let $\mathcal{Q}_{kt}^{WAIT}$ denote the ordered list of flight legs on $\mathcal{Q}_{kt}$ that have not started their service as of time t, i.e., $\mathcal{Q}_{kt}^{WAIT} = \{q_{ktn} = \mathbb{f}_i : \varsigma_{it}^{LEG} = 1\}$. Therefore, $\mathcal{Q}_{kt} = \{\mathbb{f}_{kt}^{CRNT}\} \cup \mathcal{Q}_{kt}^{WAIT}$ if $\mathbb{f}_{kt}^{CRNT}$ is well defined (i.e., not N\A), and $\mathcal{Q}_{kt} = \mathcal{Q}_{kt}^{WAIT}$, otherwise.

Let τ_{kt}^{AVL} denote the earliest time after t that UAT aircraft k would become available for the service. If aircraft k is idle or on holding at time t, it will be immediately available for future service, and therefore, $\tau_{kt}^{AVL} = t$. Otherwise, if UAT aircraft k is currently serving a flight, τ_{kt}^{AVL} is the time aircraft k completes its current flight and either becomes idle or put on hold. In other words, $\tau_{kt}^{AVL} = \min \tau$ such that $\tau \geq t$ and $\varsigma_{k\tau}^{eVTOL} \in \{0, 6\}$. $\boldsymbol{L}_{kt}^{AVL}$ is the location of aircraft k at τ_{kt}^{AVL}, and is a two-dimensional vector in a two-dimensional space.

At the beginning of the planning horizon, i.e., $t = t_0$, $\boldsymbol{L}_{kt_0}^{AVL}$ is the initial location of the aircraft k and $\tau_{kt_0}^{AVL}$ is the earliest time that it could start serving the requests. Subsequently, the value of $\boldsymbol{L}_{kt}^{AVL}$ and τ_{kt}^{AVL} will be updated every time aircraft k starts a flight leg. Additionally, $\mathfrak{E}_{kt_0}^{NDSRD} = 0$.

5.4 Events

While UAT fleet operation involves many events, not all are required to model the UAT fleet operation. We only consider the events that prompt a change in the state of the entities. Consequently, the states of the entities change only when an *event* occurs. Otherwise, they remain constant between two consecutive events. Therefore, when event v at time τ_v occurs, the relevant states get updated, and future events will be scheduled. In this section, we discuss the events associated with the UAT entities.

5.4.1 Dispatcher

The dispatcher is involved in two events: the start of the decision epoch e (v_e^S) and the end of the decision epoch e (v_e^E). With the start of the decision epoch e at time $t = \tau_{v_e^S}$, the process of routing and scheduling the UAT aircraft and assigning the requests to the flights given the policy π starts. When the decision epoch e ends at e at time $t = \tau_{v_e^E}$, the dispatcher (re)assigns the requests to flight legs, updates the itinerary of UAT aircraft, and (re)schedules the boarding time of flights and the time by which the passengers should leave their origin. Figure 5.2 depicts the events associated with the dispatcher.

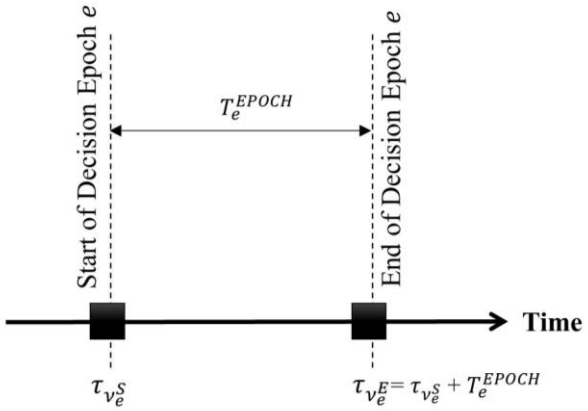

Figure 5.2 The events associated with the dispatcher

5.4.2 Flight Leg

All flight legs are involved in four events: the start of air traffic control (ATC) clearance for take-off, the departure of the UAT Aircraft, the landing of the UAT Aircraft, and the flight completion. If the flight leg is revenue-generating, it is further involved in the boarding and deboarding processes. Therefore, for an empty leg, the flight starts with ATC take-off clearance and finishes with flight completion, whereas for a revenue-generating leg, the flight starts with the boarding processes. These events and their sequence are shown in Figure 5.3.

The numbered square boxes in Figure 5.3 depict the corresponding status of a flight leg between events, i.e., ς_{it}^{LEG} in Equation (5.7). Before a flight leg is assigned, its status is *unassigned*. From the assignment until the beginning of the service, the status is *waiting for service*. The flight is *in service* between the start and completion of the flight. Ultimately, when the flight leg is complete, its status changes to *served*.

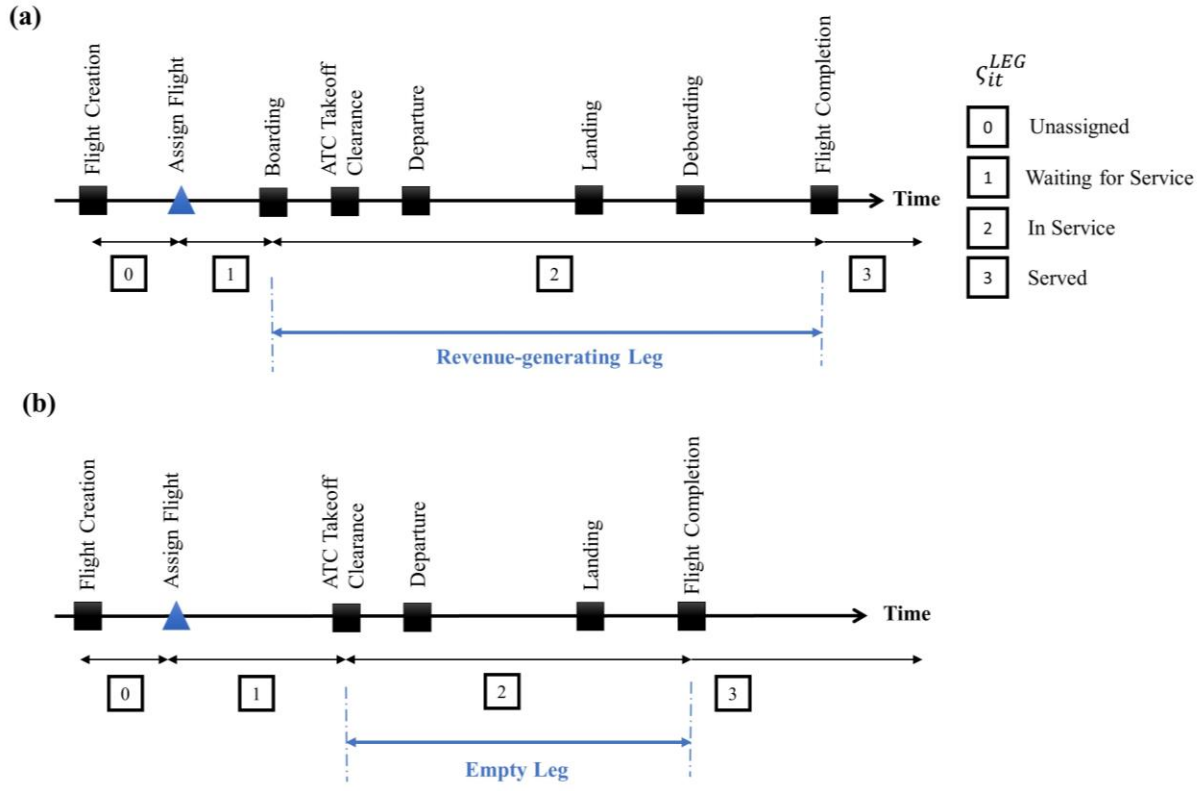

Figure 5.3 Depiction of flight leg's events for (a) revenue-generating flight, (b) empty flight

5.4.3 UAT Aircraft

Sequencing one or multiple flight legs forms the itinerary of a UAT aircraft. As a result, the events of a UAT aircraft are similar to those of flight legs. The additional event is related to an aircraft *becoming idle*, which is triggered when no event is scheduled for UAT aircraft *k* after completing a flight at time t (i.e., $\mathcal{Q}_{kt} = \emptyset$). The events associated with a UAT aircraft are illustrated in Figure 5.4(b). The numbered square boxes at the bottom of Figure 5.4 depict the

corresponding status of a UAT aircraft between events, i.e., ς_{kt}^{eVTOL} in Equation (5.9). Other events such as maintenance or charging could be scheduled for a UAT aircraft; however, we have not included those events in this dissertation.

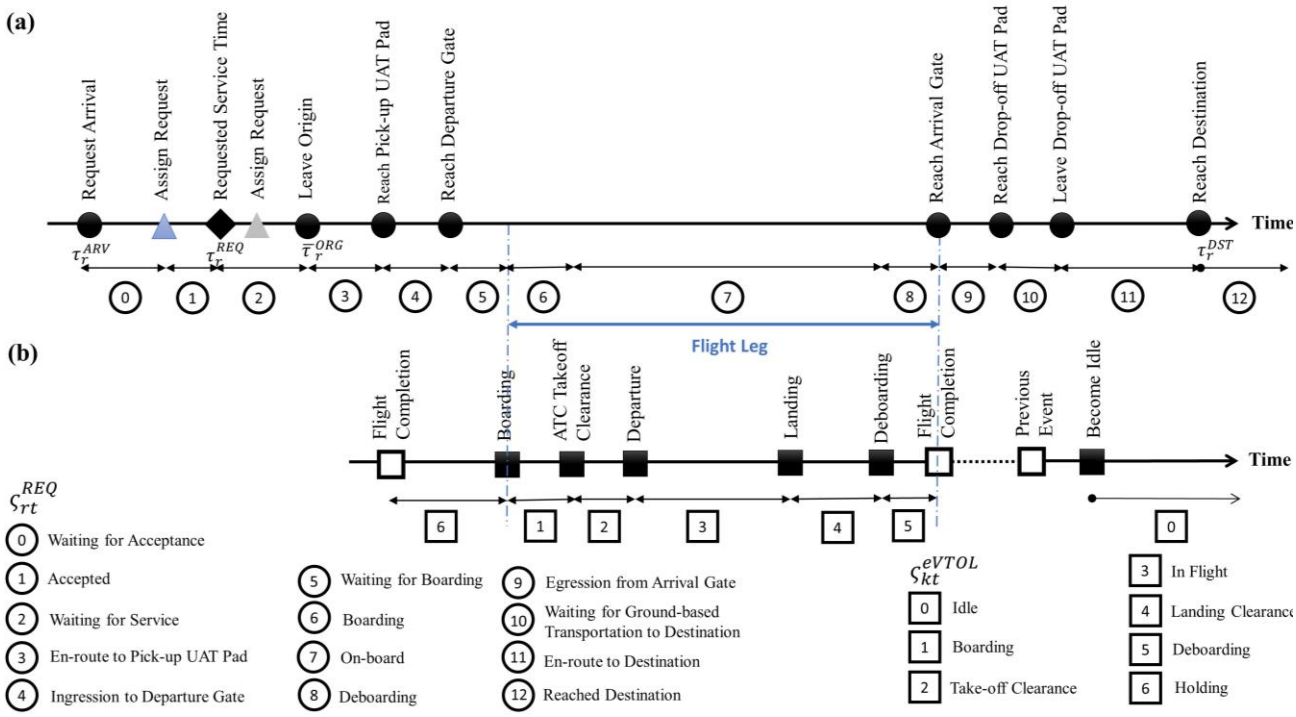

Figure 5.4 Depiction of events for (a) a request, and (b) its assigned UAT aircraft

5.4.4 Customer Request

The events directly related to the accepted requests, in chronological order, are: request arrival, leaving origin, reaching the pick-up UAT pad, reaching the departure gate, arriving at the arrival gate, reaching the designated area for ground transportation at the drop-off UAT pad, leaving the drop-off UAT pad, and finally, reaching the destination. Figure 5.4 illustrates the events associated with a request and its assigned UAT aircraft from the moment a request arrives until it reaches the destination. The numbered circles show the status of request r, i.e., ς_{rt}^{REQ} in Equation (5.3), between the events.

5.5 Activities and Delays

Let $\mathcal{T}$ denote the set of all discrete points in time when events occur. Thus, $\mathcal{T} = \{\tau_1, \tau_2, \dots, \tau_v, \dots, \tau_{|\mathcal{T}|}\}$, where τ_v represents the time event v starts. The *interevent interval* $\tau_{v+1} - \tau_v$ for $v \in \{1, \dots, |\mathcal{T}| - 1\}$ could be deterministic or stochastic. *Activity*, also known as *unconditional wait*, characterizes a period specified by the modeler and is known when it begins (e.g., the boarding duration of passengers). Activity could be deterministic (e.g., 3 minutes), have a statistical distribution (e.g., uniformly distributed with the range of (2,4), or be a function of the system's state and entity's attributes (e.g., 2 minutes per boarding passengers). In contrast, *delay*, also called *conditional wait*, describes a period that is not specified by the modeler and is determined by the system's conditions. A delay may end when a specific event occurs or a logical condition becomes true [129]. For instance, the time a passenger might wait at the departure gate to board a UAT aircraft is a type of delay.

Table 5.1 shows the duration of the activity between two consecutive events associated with request r taking flight leg i. If the times are deterministic, the varibales in the table specify the activity duration, whereas in stochastic cases, they are the mean of the activity duration. The parameters associated with activity duration are defined in detail in Chapter 7.

The N\A in Table 5.1 implies that the duration between the two events is unknown and, therefore, is considered a delay. As a result, a request could encounter delays at two points: First, the wait time between the requested time and leaving the origin (i.e., when $\varsigma_{rt}^{REQ} = 2$), and second, the wait time between arriving at the departure gate and the beginning of the boarding process (i.e., when $\varsigma_{rt}^{REQ} = 5$). The operation is designed so that if a passenger had to wait, they would do so at their origin and not at the departure gate. Nonetheless, since the boarding time could be updated

at each decision epoch, the passengers may end up waiting at the departure gate. Similarly, Table 5.2 and Table 5.3, respectively, specify the activity duration of an aircraft for empty and revenue-generating flight legs.

Table 5.1 Duration of the activity between two consecutive events associated with request r taking flight leg i

Current Event	Next Event	Activity Duration
Requested Time	Leave Origin	$N\backslash A$
Leave Origin	Reach Pick-up UAT Pad	T_{ri}^{INBND}
Reach Pick-up UAT Pad	Reach Departure Gate	T_{ri}^{DGATE}
Reach Departure Gate	Boarding	$N\backslash A$
Boarding	ATC Take-off Clearance	T^{BOARD}
ATC Take-off Clearance	Departure	$T^{TAKEOFF}$
Departure	Landing	T_i^{FLIGHT}
Landing	Deboarding	$T^{LANDING}$
Deboarding	Reach Arrival Gate	$T^{DEBOARD}$
Reach Arrival Gate	Reach Drop-off UAT Pad	T_{ri}^{AGATE}
Reach Drop-off UAT Pad	Leave Drop-off UAT Pad	W_{ri}^{EGR}
Leave Drop-off UAT Pad	Reach Destination	T_{ri}^{OUTBND}

Table 5.2 Duration of the activity between two consecutive events associated with empty flight leg i

Current Event	Next Event	Activity Duration
ATC Take-off Clearance	Departure	$T^{TAKEOFF}$
Departure	Landing	T_i^{FLIGHT}
Landing	Flight Completion	$T^{LANDING}$

Table 5.3 Duration of the activity between two consecutive events associated with revenue-generating flight leg i

Current Event	Next Event	Activity Duration
Boarding	ATC Take-off Clearance	T^{BOARD}
ATC Take-off Clearance	Departure	$T^{TAKEOFF}$
Departure	Landing	T_i^{FLIGHT}
Landing	Deboarding	$T^{LANDING}$
Deboarding	Flight Completion	$T^{DEBOARD}$

5.6 Transition Function

The transition function specifies how the state of the system evolves from event v to event $v+1$. We denote the transition function by $S^M(\cdot)$. Therefore, the dynamics of the system could be represented with $S_{\tau_{v+1}} = S^M\left(S_{\tau_v}, W_{\tau_{v+1}}\right)$, where S_{τ_v} is the state of the system during event v and $W_{\tau_{v+1}}$ is the exogenous information that arrives between event v and event $v+1$.

The state of the system S_t remains unchanged unless an event prompts the change of one or more components of S_t. Table 5.4 summarizes how the occurrence of event v at the time τ_v changes the status of accepted request r to $\varsigma_{r\tau_v}^{REQ}$ in Equation (5.3). Similarly, Table 5.5 depicts how an event changes the status of flight leg i for empty ($H_i = 0$) and revenue leg ($H_i = 1$) and status of aircraft k after occurrence of event v at time τ_v , i.e., $\varsigma_{i\tau_v}^{LEG}$ in Equation (5.7) and $\varsigma_{k\tau_v}^{eVTOL}$ in Equation (5.9), respectively.

Table 5.4 Status of request r after occurrence of event v at time τ_v ($\varsigma_{r\tau_v}^{REQ}$)

Event v	**$\varsigma_{r\tau_v}^{REQ}$**	**Definition**
Request Arrival	0	Waiting for acceptance
Passing of Requested Service Time	2	Waiting for service
Leave Origin	3	En-route to pick-up UAT pad
Reach Pick-up UAT Pad	4	Accessing departure gate
Reach Departure Gate	5	Waiting for boarding
Boarding	6	Boarding
ATC Take-off Clearance	7	On-board
Deboarding	8	Deboarding
Reach Arrival Gate	9	Accessing designated area for ground transportation
Reach Drop-off UAT Pad	10	Waiting for ground transportation to destination
Leave Drop-off UAT Pad	11	En-route to destination
Reach Destination	12	Reached destination

Table 5.5 Status of flight leg i for empty and revenue leg ($\varsigma_{i\tau_v}^{LEG}$) and status of aircraft k after occurrence of event v at time τ_v ($\varsigma_{k\tau_v}^{eVTOL}$)

Event v	$\varsigma_{k\tau_v}^{eVTOL}$	$\varsigma_{i\tau_v}^{LEG}$	
		$H_i = 0$ (Empty)	$H_i = 1$ (Revenue-generating)
Boarding	1	-	2
ATC Take-off Clearance	2	2	-
Departure	3	-	-
Landing	4	-	-
Deboarding	5	-	-
End of Deboarding	0 or 6	3	3

Moreover, the occurrence of the following events prompts additional changes in the system state at time t:

i. *Leave origin*: When the passengers of request r leave their origin at time t, set $\bar{\tau}_r^{ORG} = t$.

ii. *Start of flight leg*: Let $\not{f}_{kt^+}^{CRNT}$ be the first flight leg on $\mathcal{Q}_{kt^+}$. In other words, $\not{f}_{kt^+}^{CRNT}$ is the flight leg currently performed by aircraft k as of time t. When $\not{f}_i = \not{f}_{kt^+}^{CRNT}$ starts at time t, the earliest availability of aircraft k would correspond to the completion time of flight leg i. Therefore:

$$\tau_{kt}^{AVL} = \begin{cases} t + T_i^{SRVEMP} & H_i = 0 \\ t + T_i^{SRVREV} & H_i = 1 \end{cases} \tag{5.11}$$

Where T_i^{SRVEMP} and T_i^{SRVREV} are, respectively, the time it takes to serve an empty and revenue-generating flight leg i. Additionally, $\boldsymbol{L}_{kt}^{AVL} = \boldsymbol{E}_i$. Subsequently, the value of $\mathfrak{E}_{kt}^{NDSRD}$, as defined in Equation (5.12), will be updated to 1 for revenue-generating flight leg i if it is heading to a UAT pad other than the desired UAT pad of its intended request.

$$\mathfrak{E}_{kt}^{NDSRD} = \begin{cases} 1 & \boldsymbol{S}_i \neq \boldsymbol{S}_r^{DSRD} \text{ if } H_i = 1 \text{ and } \mathcal{r}_r = \mathcal{r}_i^{INTND} \\ 0 & \text{otherwise} \end{cases} \quad (5.12)$$

iii. *Completion of flight leg*: With the completion of $\mathcal{f}_i = \mathcal{f}_{kt^-}^{CRNT}$ at time t, the completed flight leg will be removed from $\mathcal{Q}_{kt^-}$. Therefore, $\mathcal{Q}_{kt} = \mathcal{Q}_{kt^-} \backslash \{\mathcal{f}_i\}$. Subsequently, if $\mathcal{Q}_{kt} = \emptyset$, then $\varsigma_{kt}^{eVTOL} = 0$ and $\mathfrak{E}_{kt}^{NDSRD} = 0$. However, if $\mathcal{Q}_{kt} \neq \emptyset$, then $\varsigma_{kt}^{eVTOL} = 6$.

The decisions made during the decision epochs also change the state of the system. The transition function associated with the decision epochs is specified in Section 6.6.

5.7 Concluding Remarks

The UAT operation involves numerous components and events, many of which are irrelevant to the problem of UAT fleet operation. For this reason, this chapter presents the entities necessary to model the UAT fleet operation, namely, the dispatcher, UAT aircraft, flight legs, and customer requests. Each entity is associated with a set of variables required to define its state in the system. Furthermore, these entities are involved in some events with prespecified or undetermined duration, which in turn changes their states. As a result, this chapter discusses the state variables, the events, activities and delays, and transition functions. Based on the UAT model defined in this chapter, Chapter 6 presents the dynamic solution framework for the UAT fleet operation problem.

Chapter 6 Dynamic Solution Framework

6.1 Overview

The UAT fleet operation is a dynamic and stochastic problem. Dynamic problems are often solved as a sequence of static models on a rolling horizon approach [68], which is regarded as a dynamic application of a static model. As a result, the UAT fleet operation problem involves a *sequential decision-making process*. In these problems, the operator observes the state of the system at a point in time and makes a decision in response. Ensuing this decision, the operator receives a reward or incurs a cost, and the state of the system evolves as a result of the decisions or exogenous information [133].

Consequently, in the UAT dynamic model with a sequential decision-making process, a *policy* is repeatedly called to solve the deterministic and static *snapshot* (also called *off-line* or *static*) problem. This chapter presents the solution framework with a sequential decision-making process to address the dynamic and stochastic UAT fleet operation.

6.2 Decision Epoch

The periods of time when decisions are made are called *decision epochs*. Simple decisions could be made at points in time. However, making more complicated decisions would take more time. A change in the state of the system (e.g., the arrival of a new request) could trigger the beginning of a decision epoch, or decision epochs could be scheduled at prespecified times (e.g., every 15 minutes). Let $\mathcal{E} = \{1, 2, \dots, e, \dots, E\}$ represent the set of all decision epochs, where $E = |\mathcal{E}|$.

Let v_e^S and v_e^E, respectively, denote the events corresponding to the start and end of decision epoch e. Consequently, $\tau_{v_e^S}$ and $\tau_{v_e^E}$ represent the start time and end time of decision

epoch e, respectively. Therefore, $\tau_{v_e^E} = \tau_{v_e^S} + T_e^{EPOCH}$ denotes the end of decision epoch *e*, where $T_e^{EPOCH}$is the duration of decision epoch e. We further assume that decision epochs are scheduled at prespecified times. Let $\Delta^{UPDATE} = \tau_{v_e^S} - \tau_{v_{e-1}^S}$ denote the fixed interval between the start of two decision epochs. Consequently, in a sequential decision-making process, $T_e^{EPOCH} \leq \Delta^{UPDATE}$ for $e \in \mathcal{E}$.

With rule-based decisions, such as first-come-first-served, or more straightforward decisions that could be achieved within seconds, the state of the system over the decision epoch remains unchanged. However, with more complicated decisions, the state of the system during the decision-making process may change. This impacts the quality of the solution since the solution was calculated given the system's state at the beginning of the decision epoch (i.e., $S_{\tau_{v_e^S}}$) but should be implemented based on the state at the end of the decision epoch (i.e., $S_{\tau_{v_e^E}}$). As a result, in these cases, the decision should be made based on an estimate of the system's state at the end of the decision epoch [77]. In this research, we assume that the state of the system does not change during the decision epoch.

6.3 Policy

Policy (or *decision rule*) π during the decision epoch e is a rule or function that specifies how to select a set of actions given the state $S_{\tau_{v_e^S}}$ and the exogenous information $W_{\tau_{v_e^S}}$, where $\tau_{v_e^S}$ represents the start time of decision epoch e. Therefore, $x_{\tau_{v_e^E}} = X^{\pi}\left(S_{\tau_{v_e^S}}, W_{\tau_{v_e^S}}\right)$ defines a policy function that returns the decision $x^*_{\tau_{v_e^E}}$ at time $\tau_{v_e^E}$ given the state $S_{\tau_{v_e^S}}$ and the exogenous information $W_{\tau_{v_e^S}}$. In general, π or $X^{\pi}\left(S_{\tau_{v_e^S}}, W_{\tau_{v_e^S}}\right)$ are interchangeably referred to as the policy [132].

Decision rules are classified based on how they use past information (memoryless vs. history-dependent) and how they make a decision (deterministic vs. randomized) [133]. In the *memoryless* (i.e., Markovian) process, the previous states and decisions are reflected only through the current state of the system. On the other hand, a *history-dependent* decision rule is dependent on the previous states and decisions in the system. With deterministic decision rules, decisions are made with certainty, while in randomized decision rules, a probability distribution is defined for the set of possible actions.

In dynamic programming, the most elementary class of policies is the myopic policy [132], where during each decision epoch, the decision is made using the information available at the time and without considering any information about the future state of the system. As new information becomes available in the subsequent epochs, the decision will be updated. Let $x_{\tau_{v_e^E}}$ denote the decision yielded at the end of decision epoch e at time $\tau_{v_e^E}$. Furthermore, let $C_{\tau_{v_e^E}}\left(S_{\tau_{v_e^S}}, x_{\tau_{v_e^E}}, W_{\tau_{v_e^S}}\right)$ denote the utility function of decision epoch e at time $\tau_{v_e^S}$ given the state $S_{\tau_{v_e^S}}$, exogenous information $W_{\tau_{v_e^S}}$, and the decision $x_{\tau_{v_e^E}}$. Therefore, the myopic policy π^{MYOPT} is defined as $X^{\pi^{MYOPT}}\left(S_{\tau_{v_e^S}}, W_{\tau_{v_e^S}}\right) = \min_{x_{\tau_{v_e^S}}} C_{\tau_{v_e^E}}\left(S_{\tau_{v_e^S}}, x_{\tau_{v_e^S}}, W_{\tau_{v_e^S}}\right)$ subject to the various constraints on the decisions at decision epoch *e*.

We use the CLARPTW-SRE formulation (specified in Chapter 8) as the myopic policy for the *online* (or *dynamic*) implementation of the UAT problem. CLARPTW-SRE policy, denoted as $\pi^{CLARPTW-SRE}$, is a myopic policy since it captures the state of the system at each decision epoch and does not include any information on the future state of the system.

6.4 Dynamic Input Parameters

In a dynamic setting, the demand and resources are continuously changing. At the beginning of each decision epoch, the demand for service and the available resources for serving this demand must be known. Therefore, the dynamic input parameters should be calculated for decision epoch e starting at time $t = \tau_{v_e^s}$. The input parameters at each decision epoch are either retrieved from exogenous sources or are calculated given the state of the system.

6.4.1 Exogenous Information

Exogenous information is the data that becomes known to the system over time from exogenous resources. Let $W_{\tau_{v_e^s}}$ denote the exogenous information that arrives between the start of decision epoch e at time $\tau_{v_e^s}$ and decision epoch $e-1$ at time $\tau_{v_{e-1}^s}$. The exogenous information could be sampled from a known probability distribution or retrieved from a file of sample paths [132]. In the UAT problem, there are two sources of exogenous information, namely, customer request (demand) and travel times (supply). If the travel times are not known in advance, they should be retrieved as time passes. The UAT problem has a scheduling component (e.g., boarding times). With stochastic travel times, the actual duration of an event would deviate from the scheduled time. Therefore, time variables have a *scheduled* time and a *realized* time, where the latter accounts for the stochasticity and delays.

This research assumes the customer requests are stochastic while the travel times are deterministic and known in advance. As a result, $\mathcal{R}_t^{UNASGN}$, which denotes the set of newly arrived customer requests that have not been assigned as of time t, is the only exogenous information to the UAT system.

6.4.2 System State

Let $t = S_{\tau_{v_e^S}}$ denote the state of the system at the beginning of the decision epoch e. Therefore, $S_t = \left(S_t^{REQ}, S_t^{eVTOL}, S_t^{LEG}\right)$. The following sections address the state of each UAT entity.

6.4.2.1 Customer Requests

As a passenger group starts the boarding process, it no longer needs to be included in the decision-making process. As a result, the state of the requests that have been accepted in previous decision epochs but have not started the boarding process is an integral part of the decision-making process. Let $\mathcal{R}_t^{FLXSTRT}$ denote the requests that have been accepted as of time t but their associated passenger groups have not left the origin to the pick-up UAT pad. Since these requests have not left their origin, they could be (re-)assigned to any UAT pad and, therefore, they have flexible pick-up (or start) UAT pad.

Once the passenger group leaves the origin, their pick-up UAT pad cannot be modified. Consequently, let $\mathcal{R}_t^{FXDSTRT}$ denote the accepted requests as of time t that their associated passenger groups have left the origin to access the pick-up UAT pad but have not started the boarding process. $\bar{\tau}_r^{ORG}$ represents the actual time passenger group of request r left their origin.

For all the requests in the decision-making process, τ_{rt}^{SRVC}, which is the earliest time serving request r could start as of time t, is calculated as follows:

$$\tau_{rt}^{SRVC} = \begin{cases} \max\left(\tau_r^{REQ}, t\right) + W_{ri,\max\left(\tau_r^{REQ},t\right)}^{INGRS} & r_r \in \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT} \\ \bar{\tau}_r^{ORG} & r_r \in \mathcal{R}_t^{FXDSTRT} \end{cases} \quad (5.6)$$

6.4.2.2 Flight Legs

At each decision epoch e, the requests should be assigned to *candidate flight legs* while the aircraft should be routed and scheduled to serve these flight legs. Since the UAT service is nearly on-demand, the candidate flight legs in the decision-making process could vary from one decision epoch to another, depending on the current state of the requests in the system. Let $\mathcal{F}_t^{CAND}$ denote the candidate flight legs as of time t. The definition of $\mathcal{F}_t^{CAND}$ is explained in detail in Section 7.2.2.

6.4.2.3 UAT Aircraft

τ_{kt}^{AVL}, defined as the earliest time as of time t after which the routing and scheduling of a UAT aircraft could be modified, is another crucial input parameter at each decision epoch. If aircraft k is idle or on holding at time t, it is available for future service immediately, and therefore, $\tau_{kt}^{AVL} = t$. Otherwise, if UAT aircraft k is currently serving a flight, τ_{kt}^{AVL} is the time aircraft k completes its current flight and becomes either idle or held. $\boldsymbol{L}_{kt}^{AVL}$ represents the location of aircraft k at time τ_{kt}^{AVL}.

At the beginning of the planning horizon, i.e., $t = t_0$, $\boldsymbol{L}_{kt_0}^{AVL}$ is the initial location of the aircraft k and $\tau_{kt_0}^{AVL}$ is the earliest time that it could start flying. Subsequently, the value of $\boldsymbol{L}_{kt}^{AVL}$ and τ_{kt}^{AVL} will be updated every time aircraft k starts a flight leg. The details of updating τ_{kt}^{AVL} and $\boldsymbol{L}_{kt}^{AVL}$ are discussed in Section 5.6.

6.5 Dynamic Decisions

At the end of decision epoch e (i.e., $t = \tau_{v_e^E}$), the following variables should be either determined for the first time or get updated based on the outputs of the decision epoch.

I. $\bar{\mathcal{R}}_e^{ACCPT}$: the subset of unassigned requests that get accepted during decision epoch e. Subsequently, $\bar{\mathcal{R}}_e^{REJCT} = \mathcal{R}_{\tau_{v_e^S}}^{UNASGN} \setminus \bar{\mathcal{R}}_e^{ACCPT}$, where $\bar{\mathcal{R}}_e^{REJCT}$ denotes the set of rejected requests during decision epoch e.

II. $\tilde{\tau}_{rt}^{ORG}$: the time when $r_r \in \mathcal{R}_{\tau_{v_e^S}}^{UNASGN} \cup \mathcal{R}_{\tau_{v_e^S}}^{FLXSTRT}$ must leave its origin to reach the pick-up UAT pad as of time $t = \tau_{v_e^E}$.

III. $\tilde{\varphi}_{rt}$: the flight leg assigned to $r_r \in \bar{\mathcal{R}}_e^{ACCPT}$ as of time $t = \tau_{v_e^E}$.

IV. $\mathcal{Q}_{kt}^{WAIT}$: an ordered list (i.e., queue) of all the flight legs (empty or revenue-generating) to be served by aircraft k starting from $\boldsymbol{L}_{kt}^{AVL}$ as scheduled at time $t = \tau_{v_e^E}$. For $\mathcal{Q}_{kt}^{WAIT} \neq \emptyset$, let $\mathcal{Q}_{kt}^{WAIT} = \left\{q_{kt1}^W, \dots, q_{ktn}^W, q_{kt,n+1}^W, \dots, q_{kt|\mathcal{Q}_{kt}^{WAIT}|}^W\right\}$, where q_{ktn}^W is the nth leg on $\mathcal{Q}_{kt}^{WAIT}$. Let q_{ktn}^W and $q_{kt,n+1}^W$ represent two consecutive flight legs on $\mathcal{Q}_{kt}^{WAIT}$. If q_{ktn}^W is a revenue-generating flight leg, $q_{kt,n+1}^W$ is either a revenue-generating or an empty leg. If q_{ktn}^W is an empty leg, $q_{kt,n+1}^W$ could only be a revenue-generating flight leg. Moreover, for the two consecutive flight leg $\not{f}_i = q_{ktn}^W \in \mathcal{Q}_{kt}^{WAIT}$ and $\not{f}_j = q_{kt,n+1}^W \in \mathcal{Q}_{kt}^{WAIT}$, $\boldsymbol{E}_i = \boldsymbol{S}_j$, while $\tau_{jt}^{STRT} \geq \tau_{it}^{COMP}$, where τ_{it}^{COMP} is the scheduled completion time of flight leg i as of time $t = \tau_{v_e^E}$. The time between the scheduled completion of flight leg i and the scheduled start time of the flight leg j is the *holding time* before flight leg j. Therefore, $T_{jt}^{HOLD} = \tau_{jt}^{STRT} - \tau_{it}^{COMP}$ for $\not{f}_i = q_{ktn}^W \in \mathcal{Q}_{kt}^{WAIT}$, $\not{f}_j = q_{kt,n+1}^W \in \mathcal{Q}_{kt}^{WAIT}$, and $n \neq \left|\mathcal{Q}_{kt}^{WAIT}\right|$, where T_{jt}^{HOLD} is the holding time before flight leg j as of time $t = \tau_{v_e^E}$.

V. τ_{it}^{STRT}: the starting time of flight leg i (either empty or revenue-generating) as of time $t = \tau_{v_e^E}$.

6.6 Transition Function of Decision Epochs

If an event yields a decision, the system's state is also a function of that decision. Therefore, $S_{\tau_{v+1}} = S^M\left(S_{\tau_v}, W_{\tau_{v+1}}, x_{\tau_{v+1}}\right)$, where $x_{\tau_{v+1}}$ is the decision made at τ_{v+1} that results in the system's state at τ_{v+1}, i.e., $S_{\tau_{v+1}}$. When decision epoch e ends at $t = \tau_{\nu_e^E}$, the decision variables are either determined for the first time or updated. Accordingly, the state of UAT aircraft, requests, and flight legs will be modified as follows.

6.6.1 Customer Requests

The state of $\mathscr{r}_r \in \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT} \cup \mathcal{R}_t^{FXDSTRT}$ at $t = \tau_{\nu_e^E}$ is specified by $S_{rt}^{REQ} = \left(\varsigma_{rt}^{REQ}, \varphi_{rt}, \tau_{rt}^{ORG}\right)$. At the end of the decision epoch $e \in \mathcal{E}$, ς_{rt}^{REQ} should be updated only for the requests that have arrived between the beginning of decision epoch $e-1$ at $\tau_{\nu_{e-1}^S}$ and the beginning of decision epoch e at $\tau_{\nu_e^S}$. These requests had not been assigned to any flight legs prior to $\tau_{\nu_e^S}$, and therefore, $\mathscr{r}_r \in \mathcal{R}_{\tau_{\nu_e^S}}^{UNASGN}$. At the end of the decision epoch e, all the unassigned requests should be either accepted or rejected. For the subset of $\mathcal{R}_{\tau_{\nu_e^S}}^{UNASGN}$ that got rejected during decision epoch e (i.e., $\mathscr{r}_r \in \bar{\mathcal{R}}_e^{REJCT}$), set $\varsigma_{rt}^{REQ} = -1$, while for the requests that got accepted (i.e., $\mathscr{r}_r \in \bar{\mathcal{R}}_e^{ACCPT}$), set $\varsigma_{rt}^{REQ} = 1$.

Furthermore, φ_{rt} and τ_{rt}^{ORG} for $\mathscr{r}_r \in \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT} \cup \mathcal{R}_t^{FXDSTRT}$ will be updated based on the decision made by using $\pi^{CLARPTW-SRE}$. Details are outlined in Section 8.7.

6.6.2 Flight Legs

The state of $\mathscr{f}_i \in \mathcal{F}_t^{CAND}$ at $t = \tau_{\nu_e^E}$ is specified by $S_{it}^{LEG} = \left(\varsigma_{it}^{LEG}, \tau_{it}^{STRT}\right)$. If a request is assigned to $\mathscr{f}_i$ at the end of the decision epoch (i.e., $\mathscr{f}_i = \varphi_{rt}$ for $\mathscr{r}_r$), set $\varsigma_{it}^{LEG} = 1$. Otherwise, set

$\varsigma_{it}^{LEG} = 0$. Furthermore, the start time of flight leg i, τ_{it}^{STRT}, will be updated based on the decision made by using $\pi^{CLARPTW-SRE}$. Details are outlined in Section 8.7.

6.6.3 UAT Aircraft

The state of $a_k \in \mathcal{K}$ at $t = \tau_{v_e^E}$ is specified by $S_{rt}^{eVTOL} = (\varsigma_{kt}^{eVTOL}, \tau_{kt}^{AVL}, \boldsymbol{L}_{kt}^{AVL}, \mathcal{Q}_{kt})$, among which, ς_{kt}^{eVTOL} and $\mathcal{Q}_{kt}^{WAIT} \subseteq \mathcal{Q}_{kt}$ could be impacted by the decision made during the decision epoch. The list of flight legs that are waiting to be served by aircraft k as of time t, $\mathcal{Q}_{kt}^{WAIT}$, will be updated based on the decision made by using $\pi^{CLARPTW-SRE}$. Details are outlined in Section 8.7.

When $(\mathcal{Q}_{kt})_{a_k \in \mathcal{K}}$ is determined, ς_{kt}^{eVTOL} could change based on the two following cases:

i. All the flight legs of a holding UAT aircraft are removed. In other words, $\varsigma_{k\tau_{v_e^S}}^{eVTOL} = 6$ (i.e., holding), and therefore, $\mathcal{Q}_{k\tau_{v_e^S}} \neq \emptyset$, while $\mathcal{Q}_{k\tau_{v_e^E}} = \emptyset$. In this case, set $\varsigma_{k\tau_{v_e^S}}^{eVTOL} = 0$, cancel $\psi_{k\tau_{v_e^S}}^{eVTOL}$, where ψ_{kt}^{eVTOL} denotes the next event scheduled for UAT aircraft k as of time t;

ii. A new flight leg is scheduled for an idle UAT aircraft. In other words, $\varsigma_{k\tau_{v_e^S}}^{eVTOL} = 0$ (i.e., idle), and therefore, $\mathcal{Q}_{k\tau_{v_e^S}} = \emptyset$, while $\mathcal{Q}_{k\tau_{v_e^E}} \neq \emptyset$. In this case, set $\varsigma_{k\tau_{v_e^E}}^{eVTOL} = 6$, and set and schedule $\psi_{k\tau_{v_e^E}}^{eVTOL}$ as the "start of flight leg."

6.7 Limitations

The dynamic solution framework presented in this chapter has some limitations, as listed below:

I. This framework covers deterministic travel times and does not tackle stochasticity in travel times.

II. The boarding time and the time the passengers need to leave their origin is subject to change. Therefore, the passengers would be provided with a time window to leave their origin or start boarding, rather than the exact time, which could inconvenience the passengers.

III. The state of the system is assumed to remain constant during the decision epoch. This assumption could cause an issue if the decision epoch is long enough for the state of the system to change.

6.8 Concluding Remarks

In the UAT fleet operation, as the new requests arrive, the UAT operator should decide to accept or reject the requests and update the itinerary and schedule of the aerial fleet accordingly. Therefore, the solution framework for the UAT problem involves sequential decision making. Consequently, a dynamic solution framework with sequential decision-making is defined in this chapter to address the dynamic problem of UAT fleet operation.

The decision epochs are defined in advance and are spaced equally throughout the planning horizon. The dynamic and stochastic problem of UAT fleet operation is approached on a rolling horizon basis. The policy at each decision epoch is an optimization-based myopic policy, and it uses the information available to the UAT operator at the time, without any attempts to include any forecast about the future. Consequently, a Capacitated Location-Allocation-Routing Problem with Time Windows and Short Repositioning Elimination (CLARPTW-SRE) formulation is employed at each decision epoch to re-optimize the UAT fleet operation problem at each decision epoch.

Given the dynamic nature of the problem, the input information at the beginning of each decision epoch changes. Furthermore, the specified policy provides the UAT operator with a tool to make dynamic decisions at the end of the decision epoch. These decisions, in turn, impact the state of the system. Therefore, this chapter presents the dynamic input parameters, the decisions to be made based on the outputs of the decision epochs, and the transition functions at the end of the decision epochs.

To elaborately define the optimization-based policy used at each decision epoch, Chapter 7 presents the network representation associated with CLARPTW-SRE. Subsequently, CLARPTW-SRE formulation is presented in Chapter 8.

Chapter 7 Capacitated Location-Allocation-Routing Problem with Time Windows and Short Repositioning Elimination: Network Representation

7.1 Overview

This section defines the network representing the *Capacitated Location-Allocation-Routing Problem with Time Windows and Short Repositioning Elimination (CLARPTW-SRE).* With this model, the UAT operator could decide which flight legs to perform (i.e., location in LRPs), how to assign the requests to the flight legs (i.e., allocation in LRPs), and how to route the capacitated aircraft to conduct these flight legs (i.e., routing in RRPs) while respecting the time windows of the flight legs and avoiding short empty repositioning flight legs.

Prior to presenting CLARPTW-SRE formulation, we define the network entities, including candidate requests and candidate flight legs. We next present how distances and times are defined in our model. Subsequently, we transform the UAT physical network to a node-based network by defining the corresponding nodes and arcs. Lastly, we reduce the size of the network for faster computational performance.

7.2 Network's Entities

This section specifies the three sets of entities required to define the CLARPTW-SRE at the start of decision epoch $e \in \mathcal{E}$, i.e., $t = \tau_{v_e^s}$: candidate requests, candidate flight legs, and available UAT aircraft.

7.2.1 Candidate Requests

Let $\mathcal{R}_t^{ARV}$, defined in Equation (7.1), denote the set of requests that have arrived as of time t. *Candidate requests*, i.e., $\mathcal{R}_t^{CAND} \subseteq \mathcal{R}_t^{ARV}$ defined in Equation (7.2), are the requests that have

arrived by time t but their passenger groups have not started the boarding process. Either the itinerary or schedule of these requests could undergo some changes, and therefore, they have to be involved in the decision-making process at time t.

$$\mathcal{R}_t^{ARV} = \{r_r : \tau_r^{ARV} \leq t\} \tag{7.1}$$

$$\mathcal{R}_t^{CAND} = \left\{r_r : r_r \in \mathcal{R}_t^{ARV}, \varsigma_{rt}^{REQ} \in \{0,1,2,3,4,5\}\right\} \tag{7.2}$$

Candidate requests are the union of unassigned requests ($\mathcal{R}_t^{UNASGN}$), requests with flexible pick-up UAT pads ($\mathcal{R}_t^{FLXSTRT}$), and requests with fixed pick-up UAT pads ($\mathcal{R}_t^{FXDSTRT}$), as defined in Equations (7.3)-(7.5), respectively. Therefore, $\mathcal{R}_t^{CAND} = \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT} \cup \mathcal{R}_t^{FXDSTRT}$.

$$\mathcal{R}_t^{UNASGN} = \left\{r_r : r_r \in \mathcal{R}_t^{CAND}, \varsigma_{rt}^{REQ} = 0\right\} \tag{7.3}$$

$$\mathcal{R}_t^{FLXSTRT} = \left\{r_r : r_r \in \mathcal{R}_t^{CAND}, \varsigma_{rt}^{REQ} \in \{1,2\}\right\} \tag{7.4}$$

$$\mathcal{R}_t^{FXDSTRT} = \left\{r_r : r_r \in \mathcal{R}_t^{CAND}, \varsigma_{rt}^{REQ} \in \{3,4,5\}\right\} \tag{7.5}$$

Unassigned requests, i.e., $\mathcal{R}_t^{UNASGN}$, are the requests that have arrived since the last decision epoch and are waiting for acceptance or rejection. They could either get rejected, or get accepted and have a flight assigned to them. Let $\bar{\mathcal{R}}_e^{REJCT}$ and $\bar{\mathcal{R}}_e^{ACCPT}$ denote the rejected and accepted requests during decision epoch e, respectively. Consequently, $\mathcal{R}_{\tau_{v_e^S}}^{UNASGN} = \bar{\mathcal{R}}_e^{REJCT} \cup \bar{\mathcal{R}}_e^{ACCPT}$, where $\tau_{v_e^S}$ denotes the starting time of decision epoch e.

Requests with flexible pick-up (or *start*) *UAT pads*, i.e., $\mathcal{R}_t^{FLXSTRT}$, are the requests that have been accepted in the previous decision epochs, but the passenger groups have not left their

origin as of time t. Since the passengers are at their origin, they can be reassigned to any pick-up UAT pad without rerouting and additional inconvenience. Hence, they have flexible pick-up UAT pads.

On the other hand, *Requests with fixed pick-up* (or *start*) *UAT pads*, i.e., $\mathcal{R}_t^{FXDSTRT}$, are the requests that have been accepted in the previous decision epochs, and their passenger groups have left the origin to the starting point of the assigned flight leg, but they have not started the boarding process as of time t. Since the passengers have left their origin and are on their way to the pick-up UAT pad, they will not be rerouted to another UAT pad to avoid further inconvenience. Thus, these requests have fixed pick-up UAT pads. Nonetheless, the scheduled boarding time of the flight legs assigned to these requests could be updated.

7.2.2 Candidate Flight Legs

In a network with a limited number of UAT pads, the flight legs available to serve the requests are limited and could be defined in advance. In contrast, in a ubiquitous network, a flight leg could be conducted between any two points in space, and therefore, the list of available flights is not known at the beginning of the planning horizon. Consequently, in such networks, the flight legs are defined based on the candidate requests that should be served. *Candidate flight legs* at time t, i.e., $\mathcal{F}_t^{CAND}$, are the flight legs that the UAT operator could offer to serve the candidate requests $\mathcal{R}_t^{CAND}$. $\mathcal{F}_t^{CAND}$ is the union of *desired flight legs* $\mathcal{F}_t^{DSRD}$ and *connecting flight legs* $\mathcal{F}_t^{CNCT}$, and is defined in Equation (7.6). $\mathcal{F}_t^{DSRD}$ are well-defined flight legs (or *jobs*, *tasks*) specified for each candidate request, while $\mathcal{F}_t^{CNCT}$ are composed of flexible flight legs with a starting point or ending point (or both) different from the desired ones of the request they intend to serve.

$$\mathcal{F}_t^{CAND} = \mathcal{F}_t^{DSRD} \cup \mathcal{F}_t^{CNCT} \tag{7.6}$$

Additionally, let $\mathcal{F}_t^S$ denote all the flight legs $\not{f}_i \in \mathcal{F}_t^{CAND}$ that start at the desired pick-up UAT pad of their intended request, i.e., r_i^{INTND}. Similarly, let $\mathcal{F}_t^E$ denotes all the flight legs $\not{f}_i \in \mathcal{F}_t^{CAND}$ that end at the desired drop-off UAT pad of r_i^{INTND}. Equations (7.7) and (7.8) define $\mathcal{F}_t^S$ and $\mathcal{F}_t^E$, respectively. Therefore, $\overline{\mathcal{F}_t^S} = \mathcal{F}_t^{CAND} \backslash \mathcal{F}_t^S$ is the set of candidate flight legs with the starting point different from the desired pick-up UAT pad of their intended request, and $\overline{\mathcal{F}_t^E} = \mathcal{F}_t^{CAND} \backslash \mathcal{F}_t^E$ is the set of candidate flight legs with the ending point different from the desired UAT pad of their intended request.

$$\mathcal{F}_t^S = \{\not{f}_i \in \mathcal{F}_t^{CAND} : \boldsymbol{S}_i = \boldsymbol{S}_r^{DSRD}, r_r = r_i^{INTND}\} \tag{7.7}$$

$$\mathcal{F}_t^E = \{\not{f}_i \in \mathcal{F}_t^{CAND} : \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, r_r = r_i^{INTND}\} \tag{7.8}$$

Desired and connecting flight legs are discussed in the following sections.

7.2.2.1 Desired Flight Legs

Request r (i.e., $r_r \in \mathcal{R}_t^{CAND}$) has a desired pick-up UAT pad and a desired drop-off pad, which are denoted by $\boldsymbol{S}_r^{DSRD}$ and $\boldsymbol{E}_r^{DSRD}$, respectively. Consequently, $\not{f}_r^{DSRD}$ denotes the desired flight leg, which aims to move r_r from $\boldsymbol{S}_r^{DSRD}$ to $\boldsymbol{E}_r^{DSRD}$ without any ground-based transportation. Thus, $\not{f}_r^{DSRD}$ is defined as $\mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, r_i^{INTND} = r_r)$.

In cases where r_r is an unassigned request or has no restriction on its pick-up UAT pad (i.e., $r_r \in \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT}$), $\not{f}_r^{DSRD}$ is a feasible flight leg that could be assigned to r_r. In contrast, for $r_r \in \mathcal{R}_t^{FXDSTRT}$, the pick-up UAT pad is fixed, and therefore, only flight legs that

have the same starting pad as the one already assigned to r_r could be included as feasible candidate flight legs. Let $\mathcal{F}_t^{DSRD}$ denote the set of all the *feasible* desired flight legs for $r_r \in \mathcal{R}_t^{CAND}$. $\mathcal{F}_t^{DSRD}$ is specified in Equation (7.9), where $\mathbb{S}(f_i)$ is a function that returns the starting point of f_i, and φ_{rt} denotes the flight leg that has been assigned to $r_r \in \mathcal{R}_t^{FLXSTRT} \cup \mathcal{R}_t^{FXDSTRT}$ as of time t.

$$\begin{aligned}\mathcal{F}_t^{DSRD} &= \{f_r^{DSRD}: r_r \in \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT}\} \\ &\cup \{f_r^{DSRD}: \mathbb{S}(f_r^{DSRD}) = \mathbb{S}(\varphi_{rt}), r_r \in \mathcal{R}_t^{FXDSTRT}\}\end{aligned} \tag{7.9}$$

7.2.2.2 Connecting Flight Legs

While the requests are being served in a ubiquitous network, some empty flight legs might be too short to justify the repositioning. Let Δ^{EMPTY} denote the minimum distance on the ground between two UAT pads to sanction a repositioning flight leg. Figure 7.1 illustrates how introducing *connecting flight legs* could eliminate short repositioning empty flight legs. Figure 7.1(a) depicts the first availability UAT pad of aircraft *k* as of time *t* (i.e., $\boldsymbol{L}_{kt}^{AVL}$) and two requests (i.e., r_r and r_s) as well as their corrsponding desired pick-up UAT pads, desired drop-off UAT pads, and their desired flight legs. Additionally, let the function $dist(a, b)$ measure the distance (as the crow flies) between point *a* and point *b* in the space.

If aircraft k were to serve only request r, a possible itinerary would be $\boldsymbol{L}_{kt}^{AVL} \rightarrow \boldsymbol{S}_r^{DSRD} \rightarrow \boldsymbol{E}_r^{DSRD}$. If the distance between $\boldsymbol{L}_{kt}^{AVL}$ and $\boldsymbol{S}_r^{DSRD}$ is below the minimum UAT flight distance, i.e., $dist(\boldsymbol{L}_{kt}^{AVL}, \boldsymbol{S}_r^{DSRD}) < \Delta^{EMPTY}$, r_r could be *relocated* or *transferred* to $\boldsymbol{L}_{kt}^{AVL}$ on the ground, and consequently, a revenue-generating flight leg from $\boldsymbol{L}_{kt}^{AVL}$ to $\boldsymbol{E}_r^{DSRD}$, i.e., $\mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{L}_{kt}^{AVL}, \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, r_i^{INTND} = r_r)$, could serve the passenger group of r_r while eliminating the short repositioning flight leg from $\boldsymbol{L}_{kt}^{AVL}$ to $\boldsymbol{S}_r^{DSRD}$.

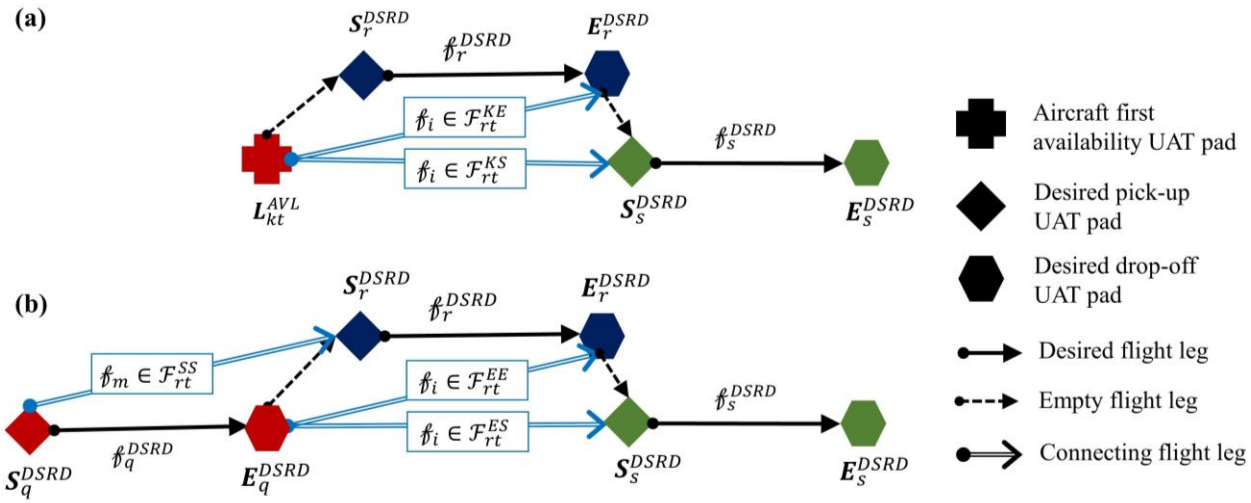

Figure 7.1 Concept of connecting legs

On the other hand, if aircraft k were to serve r_r and r_s consecutively, the itinerary as depicted in Figure 7.1(a), would be $\boldsymbol{L}_{kt}^{AVL} \rightarrow \boldsymbol{S}_r^{DSRD} \rightarrow \boldsymbol{E}_r^{DSRD} \rightarrow \boldsymbol{S}_s^{DSRD} \rightarrow \boldsymbol{E}_s^{DSRD}$. If the distance between $\boldsymbol{L}_{kt}^{AVL}$ and $\boldsymbol{S}_r^{DSRD}$ and the distance between $\boldsymbol{E}_r^{DSRD}$ and $\boldsymbol{S}_s^{DSRD}$ are both shorter than the minimum UAT flight distance, i.e., $dist(\boldsymbol{L}_{kt}^{AVL}, \boldsymbol{S}_r^{DSRD}) < \Delta^{EMPTY}$ and $dist(\boldsymbol{E}_r^{DSRD}, \boldsymbol{S}_s^{DSRD}) < \Delta^{EMPTY}$, the passengers of r_r could be relocated from $\boldsymbol{S}_r^{DSRD}$ to $\boldsymbol{L}_{kt}^{AVL}$ and from $\boldsymbol{S}_s^{DSRD}$ to $\boldsymbol{E}_r^{DSRD}$ on the ground if they were served by a connecting flight leg that goes directly from $\boldsymbol{L}_{kt}^{AVL}$ to $\boldsymbol{S}_s^{DSRD}$, i.e., $\mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{L}_{kt}^{AVL}, \boldsymbol{E}_i = \boldsymbol{S}_s^{DSRD}, r_i^{INTND} = r_r)$, eliminating the two short repositioning flight legs.

Analogously, Figure 7.1(b) illustrates requests r_q, r_r, and r_s. If r_r were to be served immediately after r_q, the aircraft itinerary would include $\boldsymbol{S}_q^{DSRD} \rightarrow \boldsymbol{E}_q^{DSRD} \rightarrow \boldsymbol{S}_r^{DSRD} \rightarrow \boldsymbol{E}_r^{DSRD}$, which involves the empty repositioning flight between $\boldsymbol{E}_q^{DSRD}$ and $\boldsymbol{S}_r^{DSRD}$. In this case, two connecting flight legs could eliminate the need for the empty flight leg: (1) $\mathbb{F}_i\left(\boldsymbol{S}_i = \boldsymbol{S}_q^{DSRD}, \boldsymbol{E}_i = \boldsymbol{S}_r^{DSRD}, r_i^{INTND} = r_q\right)$ which connects the desired pick-up UAT pad of r_q to the desired pick-up UAT pad of r_r while serving passengers of r_q, and (2) $\mathbb{F}_i\left(\boldsymbol{S}_i = \boldsymbol{E}_q^{DSRD}, \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, r_i^{INTND} = r_r\right)$ which connects the desired drop-off UAT pad of r_q to the desired drop-off UAT pad of r_r

while serving passengers of r_r. In the former case, the passengers of r_q and in the latter case the passengers of r_r would be relocated on the ground from $\boldsymbol{S}_r^{DSRD}$ to $\boldsymbol{E}_q^{DSRD}$.

On the other hand, if an aircraft were to serve r_q, r_r, and r_s consecutively, the itinerary, as depicted in Figure 7.1(b), would be $\boldsymbol{S}_q^{DSRD} \rightarrow \boldsymbol{E}_q^{DSRD} \rightarrow \boldsymbol{S}_r^{DSRD} \rightarrow \boldsymbol{E}_r^{DSRD} \rightarrow \boldsymbol{S}_s^{DSRD} \rightarrow \boldsymbol{E}_s^{DSRD}$. If the distance between $\boldsymbol{E}_q^{DSRD}$ and $\boldsymbol{S}_r^{DSRD}$ and the distance between $\boldsymbol{E}_r^{DSRD}$ and $\boldsymbol{S}_s^{DSRD}$ are both shorter than the minimum UAT flight distance, i.e., $dist\left(\boldsymbol{E}_q^{DSRD}, \boldsymbol{S}_r^{DSRD}\right) < \Delta^{EMPTY}$ and $dist(\boldsymbol{E}_r^{DSRD}, \boldsymbol{S}_s^{DSRD}) < \Delta^{EMPTY}$, the passengers of r_r would be relocated from $\boldsymbol{S}_r^{DSRD}$ to $\boldsymbol{E}_q^{DSRD}$ and from $\boldsymbol{S}_s^{DSRD}$ to $\boldsymbol{E}_r^{DSRD}$ on the ground if they were served by a connecting flight leg that goes directly from $\boldsymbol{E}_q^{DSRD}$ to $\boldsymbol{S}_s^{DSRD}$, i.e., $\mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{E}_q^{DSRD}, \boldsymbol{E}_i = \boldsymbol{S}_s^{DSRD}, r_i^{INTND} = r_r)$, eliminating the two short repositioning flight legs.

In conclusion, to eliminate a short repositioning flight leg, the passenger group of a request could be relocated on the ground within the radius of Δ^{EMPTY} from their desired UAT pads. It is worth noting that the relocation should be within Δ^{ACCESS} of either the origin or destination of the request. However, in a ubiquitous network, the origin and destination of a request coincide with the desired pick-up and drop-off UAT pads. Introducing the connecting legs and eliminating the short repositioning flight legs would transform the CLARPTW to CLARPTW-SRE, where SRE stands for *short repositioning elimination*. Subsequently, we define five classes of connecting legs, namely, $\mathcal{F}_{rt}^{KE}, \mathcal{F}_{rt}^{KS}, \mathcal{F}_{rt}^{SS}, \mathcal{F}_{rt}^{EE}$, and $\mathcal{F}_{rt}^{ES}$ for $r_r \in \mathcal{R}_t^{CAND}$ as follows.

Figure 7.2 illustrates $\boldsymbol{S}_r^{DSRD}$ and $\boldsymbol{E}_r^{DSRD}$, the desired pick-up and drop-off UAT pads of $r_r \in \mathcal{R}_t^{CAND}$, respectively, as well as the desired pick-up UAT pads of all the candidate requests within a radius of Δ^{EMPTY} of $\boldsymbol{E}_r^{DSRD}$. The connecting legs are defined so that they could serve r_r from $\boldsymbol{S}_r^{DSRD}$ to these pads. Let $\mathcal{F}_{rt}^{SS}$ denote the set of connecting legs that intend to serve r_r by flying

between $\boldsymbol{S}_r^{DSRD}$ and the desired pick-up pad of $r_s \in \mathcal{R}_t^{CAND}$ within Δ^{UAT} of $\boldsymbol{E}_r^{DSRD}$. As a result, $\mathcal{F}_{rt}^{SS}$ for $r_r \in \mathcal{R}_t^{CAND}$ is defined in Equation (7.10).

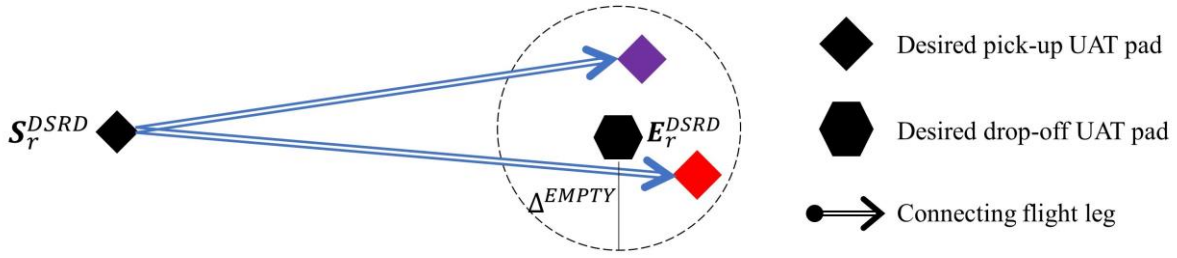

Figure 7.2 Connecting legs intending to serve request r that start at its desired pick-up UAT pad

$$\mathcal{F}_{rt}^{SS} = \{\not{p}_i : \not{p}_i = \mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_i = \boldsymbol{S}_s^{DSRD}, r_i^{INTND} = r_r), \; 0 < dist(\boldsymbol{E}_r^{DSRD}, \boldsymbol{S}_s^{DSRD}) < \Delta^{EMPTY}, r_s \in \mathcal{R}_t^{CAND}\} \qquad r_r \in \mathcal{R}_t^{CAND} \quad (7.10)$$

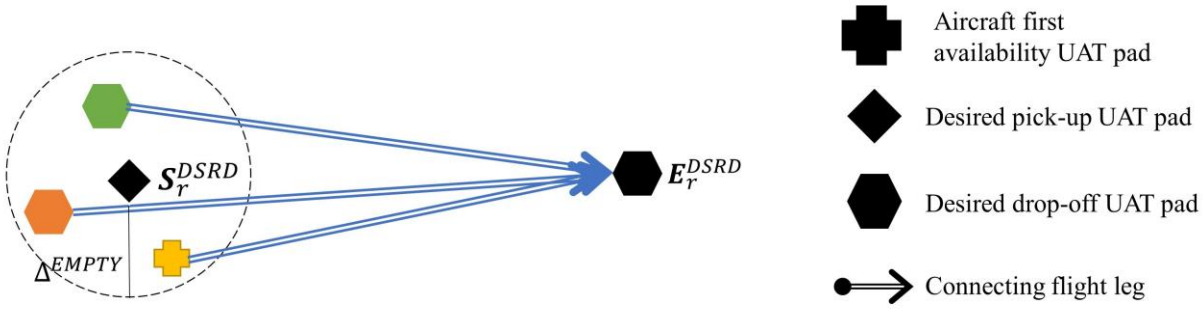

Figure 7.3 Connecting legs intending to serve request r that end at its desired drop-off UAT pad

Similarly, Figure 7.3 illustrates $\boldsymbol{S}_r^{DSRD}$ and $\boldsymbol{E}_r^{DSRD}$, the desired pick-up and drop-off UAT pads of $r_r \in \mathcal{R}_t^{CAND}$, respectively. Moreover, the first availability UAT pads of aircraft as well as the desired drop-off UAT pads of all candidate requests within the Δ^{EMPTY} of $\boldsymbol{S}_r^{DSRD}$ are depiceted. The connecting legs are defined so that they could serve the passengers of r_r from these pads to $\boldsymbol{E}_r^{DSRD}$. Let $\mathcal{F}_{rt}^{KE}$ denote the set of connecting legs that serve r_r by flying between $\boldsymbol{L}_{kt}^{AVL}$ (i.e., the fisrt availability UAT pad of aircraft k) within Δ^{EMPTY} of $\boldsymbol{S}_r^{DSRD}$ and $\boldsymbol{E}_r^{DSRD}$, and let $\mathcal{F}_{rt}^{EE}$ denote

the set of connecting legs introduced between $\boldsymbol{E}_q^{DSRD}$ (i.e., the desired UAT pad) of $r_q \in \mathcal{R}_t^{CAND}$ within Δ^{EMPTY} of $\boldsymbol{S}_r^{DSRD}$ and $\boldsymbol{E}_r^{DSRD}$, intending to serve r_r. $\mathcal{F}_{rt}^{KE}$ and $\mathcal{F}_{rt}^{EE}$ are defined in Equations (7.11) and (7.12), respectively.

$$\mathcal{F}_{rt}^{KE} = \{f_i: f_i = \mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{L}_{kt}^{AVL}, \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, r_i^{INTND} = r_r), \; 0 < dist(\boldsymbol{L}_{kt}^{AVL}, \boldsymbol{S}_r^{DSRD}) < \Delta^{EMPTY}, a_k \in \mathcal{K}\} \quad r_r \in \mathcal{R}_t^{CAND} \quad (7.11)$$

$$\mathcal{F}_{rt}^{EE} = \{f_i: f_i = \mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{E}_q^{DSRD}, \boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}, r_i^{INTND} = r_r), \; 0 < dist(\boldsymbol{E}_q^{DSRD}, \boldsymbol{S}_r^{DSRD}) < \Delta^{EMPTY}, r_q \in \mathcal{R}_t^{CAND}\} \quad r_r \in \mathcal{R}_t^{CAND} \quad (7.12)$$

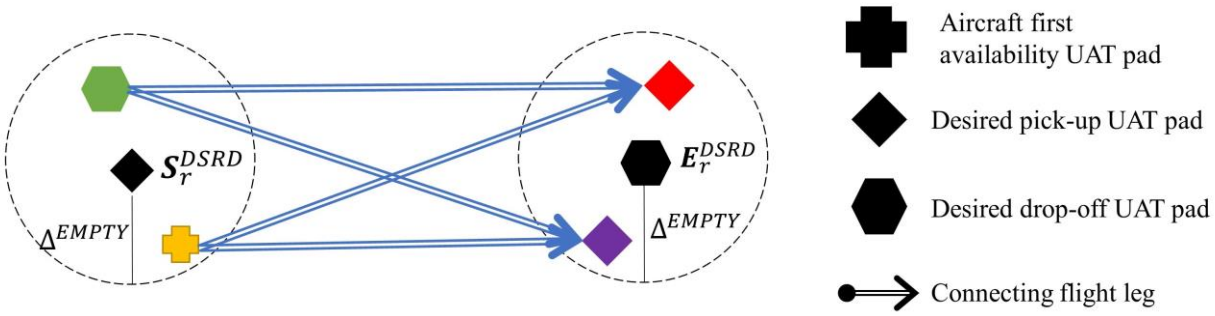

Figure 7.4 Connecting legs intending to serve request r that neither start nor end at at its desired UAT pads

Lastly, Figure 7.4 depicts $\boldsymbol{S}_r^{DSRD}$ and $\boldsymbol{E}_r^{DSRD}$, the desired pick-up and drop-off UAT pads of $r_r \in \mathcal{R}_t^{CAND}$, respectively. Additionally, the availability UAT pads and desired drop-off UAT pads of candidate requests within Δ^{EMPTY} radius of $\boldsymbol{S}_r^{DSRD}$ and the desired pick-up UAT pads of candidate requests within Δ^{EMPTY} radius of $\boldsymbol{E}_r^{DSRD}$ are illustrated. Let $\mathcal{F}_{rt}^{KS}$ denote the set of connecting legs that serve r_r by flying between the availability UAT pad $a_k \in \mathcal{K}$ within Δ^{EMPTY} of $\boldsymbol{S}_r^{DSRD}$ and the desired pick-up UAT pad of $r_s \in \mathcal{R}_t^{CAND}$ within Δ^{EMPTY} of $\boldsymbol{E}_r^{DSRD}$. Furthermore, let $\mathcal{F}_{rt}^{ES}$ denote the set of connecting legs that aim to serve r_r while connecting the desired drop-off UAT pad of $r_q \in \mathcal{R}_t^{CAND}$ within Δ^{EMPTY} of $\boldsymbol{S}_r^{DSRD}$ to the desired pick-up UAT pad of $r_s \in$

$\mathcal{R}_t^{CAND}$ within Δ^{EMPTY} of $\boldsymbol{E}_r^{DSRD}$. $\mathcal{F}_{it}^{KS}$ and $\mathcal{F}_{it}^{ES}$ are defined in Equation (7.13) and (7.14), respectively.

$$\begin{aligned}\mathcal{F}_{rt}^{KS} = \{\not{f}_i : \not{f}_i = \ & \mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{L}_{kt}^{AVL}, \boldsymbol{E}_i = \boldsymbol{S}_s^{DSRD}, r_i^{INTND} = r_r), \\ & 0 < dist(\boldsymbol{L}_{kt}^{AVL}, \boldsymbol{S}_r^{DSRD}) < \Delta^{EMPTY}, \\ & 0 < dist(\boldsymbol{E}_r^{DSRD}, \boldsymbol{S}_s^{DSRD}) < \Delta^{EMPTY}, \\ & a_k \in \mathcal{K}, r_s \in \mathcal{R}_t^{CAND}\}\end{aligned} \qquad r_r \in \mathcal{R}_t^{CAND} \qquad (7.13)$$

$$\begin{aligned}\mathcal{F}_{rt}^{ES} = \{\not{f}_i : \not{f}_i = \ & \mathbb{F}_i(\boldsymbol{S}_i = \boldsymbol{E}_q^{DSRD}, \boldsymbol{E}_i = \boldsymbol{S}_s^{DSRD}, r_i^{INTND} = r_r), \\ & 0 < dist(\boldsymbol{E}_q^{DSRD}, \boldsymbol{S}_r^{DSRD}) < \Delta^{EMPTY}, \\ & 0 < dist(\boldsymbol{E}_r^{DSRD}, \boldsymbol{S}_s^{DSRD}) < \Delta^{EMPTY}, r_q, r_s \in \mathcal{R}_t^{CAND}\}\end{aligned} \qquad r_r \in \mathcal{R}_t^{CAND} \qquad (7.14)$$

In conclusion, five classes of connecting legs, namely, $\mathcal{F}_{rt}^{KE}, \mathcal{F}_{rt}^{KS}, \mathcal{F}_{rt}^{SS}, \mathcal{F}_{rt}^{EE}$, and $\mathcal{F}_{rt}^{ES}$, could be defined for $r_r \in \mathcal{R}_t^{CAND}$. However, for $r_r \in \mathcal{R}_t^{FXDSTRT}$, the pick-up UAT pad is set as fixed, and therefore, only the connecting flight legs that have the same starting pad as the one of the flight already assigned to r_r are feasible connecting flight legs. Let $\mathcal{F}_t^{CNCT}$ denote all the *feasible* connecting legs at time t. Therefore, Equation (7.15) defines $\mathcal{F}_t^{CNCT}$, where $\mathbb{S}(\not{f}_i)$ is a function that returns the starting point of $\not{f}_i$.

$$\begin{aligned}\mathcal{F}_t^{CNCT} = & \{\not{f}_i : \not{f}_i \in (\mathcal{F}_{rt}^{KE} \cup \mathcal{F}_{rt}^{KS} \cup \mathcal{F}_{rt}^{SS} \cup \mathcal{F}_{rt}^{EE} \cup \mathcal{F}_{rt}^{ES}), \\ & \qquad r_r \in \mathcal{R}_t^{UNASGN} \cup \mathcal{R}_t^{FLXSTRT}\} \\ & \cup \{\not{f}_i : \mathbb{S}(\not{f}_i) = \mathbb{S}(\varphi_{rt}), \\ & \qquad \not{f}_i \in (\mathcal{F}_{rt}^{KE} \cup \mathcal{F}_{rt}^{KS} \cup \mathcal{F}_{rt}^{SS} \cup \mathcal{F}_{rt}^{EE} \cup \mathcal{F}_{rt}^{ES}), \\ & \qquad r_r \in \mathcal{R}_t^{FXDSTRT}\}\end{aligned} \qquad (7.15)$$

Connecting legs are introduced to eliminate the short repositioning flight legs, and therefore, a revenue-generating flight leg must come immediately either before or after them. *Preceding flight leg* $\not{f}_m$ refers to the revenue-generating flight leg that must be served immediately before a connecting leg $\not{f}_i$, suggesting $\not{f}_i \in \overline{\mathcal{F}_t^S}$ and $\not{f}_m \in \mathcal{F}_t^E$. On the other hand, *succeeding*

flight leg $\not{f}_j$ refers to the revenue-generating flight leg that must be served immediately after a connecting leg, implying $\not{f}_i \in \overline{\mathcal{F}_t^E}$ and $\not{f}_j \in \mathcal{F}_t^S$. In contrast, *free flight legs* can freely follow each other without any constraint. As can be seen in Figure 7.1, $\not{f}_i \in \mathcal{F}_{rt}^{EE}$ has preceding flight legs, $\not{f}_i \in \mathcal{F}_{rt}^{KS}$ and $\not{f}_i \in \mathcal{F}_{rt}^{SS}$ have succeeding flight legs, and $\not{f}_i \in \mathcal{F}_{rt}^{ES}$ has both preceding and succeeding flight legs.

7.2.3 Available UAT Aircraft

The UAT operator employs a fixed set of UAT aircraft, i.e., $\mathcal{K}$, over the planning horizon. Let $\overline{\mathcal{K}_t^E}$ denote all the UAT aircraft that are conducting or have recently conducted flight leg i to a UAT pad that is different from the desired drop-off UAT pad of r_i^{INTND}. $\overline{\mathcal{K}_t^E}$ is defined in Equation (7.16). As a result, $\mathcal{K}_t^E = \mathcal{K} \backslash \overline{\mathcal{K}_t^E}$ represents the set of UAT aircraft that are conducting or have recently conducted flights leg i to a UAT pad that is the same as the desired drop-off UAT pad of r_i^{INTND}.

$$\overline{\mathcal{K}_t^E} = \{a_k \in \mathcal{K}: \mathfrak{E}_{kt}^{NDSRD} = 1\} \tag{7.16}$$

7.3 Network's Metrics

This section specifies how the distances and times are defined and calculated in the network.

7.3.1 Distances

Let the function $dist(a, b)$ measure the distance (as the crow flies) between point a and point b in the space. For smaller areas, it measures the straight-line distance in Euclidean space. However, the geodesic distance is used in bigger regions to capture the earth's curvature. Additionally, we define *aerial distances* to take into account the deviation of the aerial routes from

the shortest path due to prohibited areas, noise reduction, and legal restrictions [91]. Let $\epsilon \geq 0$ denote the detour factor due to prohibited areas, noise reduction, and legal restrictions. Consequently, we define the following distances:

$D_r^{OD} = dist(\boldsymbol{O}_r, \boldsymbol{D}_r)$: the distance (as the crow flies) between the origin and destination of request r;

$D_i^{LEG} = (1 + \epsilon) \times dist(\boldsymbol{S}_i, \boldsymbol{E}_i)$: the aerial distance between the starting and ending point of flight leg i;

$D_{kit}^0 = (1 + \epsilon) \times dist(\boldsymbol{L}_{kt}^{AVL}, \boldsymbol{S}_i)$: the aerial distance between the first availability UAT pad of aircraft k as of time t and the starting point of flight leg i; and

$D_{ij} = (1 + \epsilon) \times dist(\boldsymbol{E}_i, \boldsymbol{S}_j)$: the aerial distance between flight leg i and flight leg j, defined as the aerial distance between the ending point of flight leg i and the starting point of flight leg j.

7.3.2 Times

This section presents the formal notations and definitions of temporal components of the UAT problem, including the ground-based and aerial times.

7.3.2.1 Ground-based Times

Ingress duration is the time it takes to get to the departure gate from the request's origin as the service starts. If the passenger group of a request needs to be relocated to a pad other than its origin, ground-based transportation will be conducted using two ground-based modes: walking and ride-hailing. This time is shown by the circles numbered as 3 in Figure 5.4(a). Let Δ^{WALK} denote the maximum walking distance of the passengers. If $dist(\boldsymbol{O}_r, \boldsymbol{S}_i) \leq \Delta^{WALK}$, request r walks to the starting point of flight leg i; otherwise, the UAT operator assigns a car to the request

to perform the relocation. Moreover, some time is spent from when the passenger group arrives at the UAT pad or port on the ground until they reach the departure gate. This event's duration is denoted by the circle numbered as 4 in Figure 5.4(a) and includes the security screening before the flight, taking the elevator, etc. Equation (7.17) defines the ingress duration of request r to flight leg i.

$$T_{ri}^{INGRS} = T_{ri}^{INBND} + T_{ri}^{DGATE} \tag{7.17}$$

$$T_{ri}^{INBND} = \frac{dist(\boldsymbol{O}_r, \boldsymbol{S}_i)}{v^m} \tag{7.18}$$

Where:

T_{ri}^{INGRS}: ingress duration, defined as the total time spent from when passengers of request r leave their origin until they reach the departure gate of flight leg i;

T_{ri}^{DGATE}: the time between arriving at the UAT pad or port on the ground and reaching the departure gate of flight i for request r;

T_{ri}^{INBND}: the elapsed time associated with the ground-based transportation for accessing the starting UAT pad of flight i from the origin of request r, as defined in Equation (7.18); and

v^m: speed of ground-based mode m, where $m \in$ {WALK, DRIVE}. It is worth noting that since the distances are calculated over Euclidean space, the speed is also estimated for traversing the Euclidean distance. Therefore, the actual speed in the network is greater than or equal to v^m.

If the requests are boarded at their desired UAT pad located at their origin, no ground-based transportation is required, and therefore, the minimum ingress duration is equal to the time it takes to access the departure gate (i.e., $T_{ri}^{INGRS} = T_{ri}^{DGATE}$).

Similarly, the *egress duration* is the elapsed time between the end of the deboarding process and reaching the destination. The egress duration is comprised of three components: (1) the elapsed time from reaching the arrival gate until arriving at the location designated for ground transportation, (2) the wait time for ground-based transportation, and (3) the time it takes on the ground to reach the destination. Accordingly, the egress time is defined in Equation (7.19):

$$T_{ri}^{EGRS} = T_{ri}^{AGATE} + W_{ri}^{EGRS} + T_{ri}^{OUTBND} \tag{7.19}$$

$$T_{ri}^{OUTBND} = \frac{dist(\boldsymbol{E}_i, \boldsymbol{D}_r)}{\nu^m} \tag{7.20}$$

Where:

T_{ri}^{EGRS}: egress duration, the total time spent from the end of the deboarding of flight leg i until reaching the destination of request r;

T_{ri}^{AGATE}: the elapsed time for reaching the area of ground transportation from the arrival gate flight leg i for request r;

T_{ri}^{OUTBND}: the elapsed time associated with the ground-based transportation for reaching the destination of request r from the ending UAT pad of flight leg i, as defined in Equation (7.20); and

W_{ri}^{EGRS}: wait time for ground transportation to the destination of request r from the ending point of flight leg i. This value is assumed to be zero since the UAT operator knows the arrival time of request r well in advance, i.e., $W_{ri}^{EGRS} = 0$. Therefore, $T_{ri}^{EGRS} = T_{ri}^{AGATE} + T_{ri}^{OUTBND}$.

7.3.2.2 Aerial Times

The aerial times could be defined in terms of either the time a UAT aircraft spends in the air or the time it takes to serve a flight leg. *Flight duration* is defined as the time elapsed between the take-off and landing of the UAT aircraft and comprises three components: ascending hover, cruise, and descending hover. In comparison, the *flight service time* refers to the time it takes to serve a flight leg from its start until its completion, and therefore, its duration depends on whether the flight leg is revenue-generating (with passengers) or empty.

Flight duration is defined in Equation (7.21), while flight service times for empty and revenue-generating legs are defined in Equations (7.22) and (7.23), respectively. Flight service time includes ATC departure clearance and landing clearance. However, for the revenue-generating legs, the duration of boarding and deboarding procedures should also be included in the flight service time. Clearly, T_i^{FLIGHT}, T_i^{SRVEMP}, and T_i^{SRVREV} are zero if the distance of flight leg i is zero.

$$T_i^{FLIGHT} = T^{ASCEND} + T_i^{CRUISE} + T^{DESCEND} \qquad \forall i: D_i^{LEG} \neq 0 \qquad (7.21)$$

$$T_i^{SRVEMP} = T^{TAKEOFF} + T_i^{FLIGHT} + T^{LANDING} \qquad (7.22)$$

$$T_i^{SRVREV} = T^{BOARD} + T^{TAKEOFF} + T_i^{FLIGHT} + T^{LANDING} + T^{DEBOARD} \qquad (7.23)$$

Where:

T_i^{FLIGHT}: the flight duration of leg i, which is defined as the period between take-off and touch down of a UAT aircraft;

$T^{ASCEND}, T^{DESCEND}$: the time required by eVTOL aircraft to vertically or close to vertically ascend and descend, respectively;

$T^{TAKEOFF}$: the time it takes before the departure to receive take-off clearance from ATC;

$T^{LANDING}$: the time it takes after the landing to declare the UAT pad area safe;

T_i^{SRVEMP}: the flight service time of an empty flight leg, i.e., the time it takes to serve the empty flight leg i from its start to its completion;

T_i^{SRVREV}: the flight service time of a revenue-generating flight leg, i.e., the time it takes to serve the revenue-generating flight leg i from its start to its completion; and

T_i^{CRUISE}: the time spent in cruising mode from the starting point of flight leg i to its ending point. It is estimated in Equation (7.24).

$$T_i^{CRUISE} = \frac{D_i^{LEG}}{v^{AIR}} = \frac{(1+\epsilon) \times dist(\boldsymbol{S}_i, \boldsymbol{E}_i)}{v^{AIR}} \tag{7.24}$$

Additionally, T_{kit}^0 denotes the T_m^{SRVEMP} of flight leg m where aircraft k repositions from its first availability UAT pad as of time t to the starting point of flight i. Similarly, T_{ij} denotes T_m^{SRVEMP} of flight leg m, which is the leg performed for repositioning a UAT aircraft from the ending point of flight leg i to the starting point of flight leg j.

Let T^{TAT} denote the *turnaround time,* which in aviation terminology refers to the time elapsed between the landing of an aircraft and its take-off to serve a new flight leg. Therefore, $T^{TAT} \geq T^{LANDING} + T^{DEBOARD} + T^{BOARD} + T^{TAKEOFF}$.

In summary, Figure 7.5 depicts the temporal components associated with a location change in UAT operation. W_{rit}^{INGRS} and W_{ri}^{EGRS} would occur at point 1 and point 7, respectively. $T^{TAKEOFF}$ and T^{BOARD} are incurred at point 3, while $T^{LANDING}$ and T^{BOARD} happens at point 6.

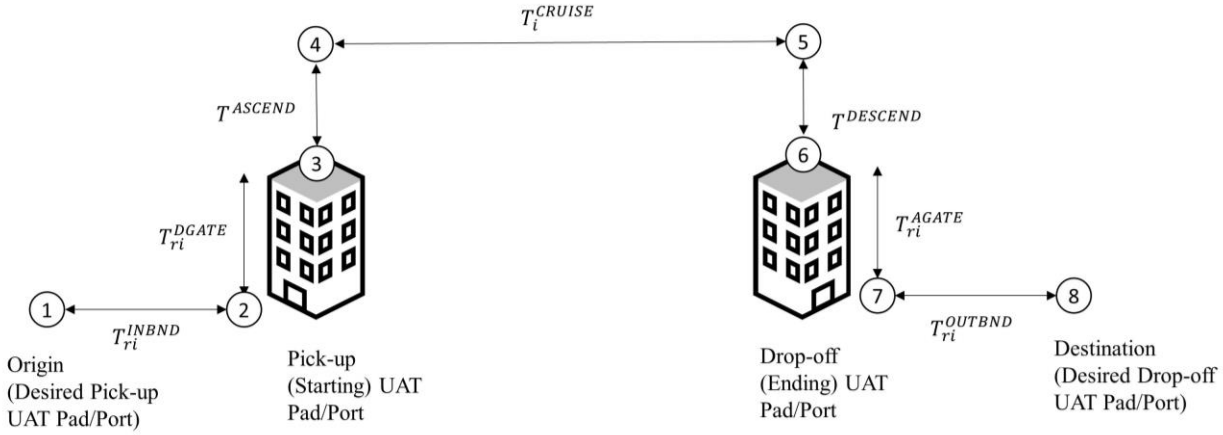

Figure 7.5 Illustration of temporal components associated with a location change

7.3.2.3 Time Windows

When passengers are assigned to a flight that differs from their desired flight leg, they incur an additional delay. Let Ω_{ri}^{MIN} denote the minimum delay incurred when passengers of $\mathcal{r}_r$ take $\not{f}_i$ instead of $\not{f}_{i^*} = \not{f}_r^{DSRD}$. Ω_{ri}^{MIN} is defined in Equation (7.25). For $\mathcal{r}_r$ to be assigned to flight leg $\not{f}_i$, it is necessary that $\Omega_{ri}^{MIN} \leq \omega$. Clearly, $\Omega_{ri}^{MIN} = 0$ for $\not{f}_i = \not{f}_r^{DSRD}$. It is worth noting that $T_{ri^*}^{INBND}$ and $T_{ri^*}^{OUTBND}$ are zero in a ubiquitous network, where the origin and destination coincide with the desired pick-up and drop-off UAT pads, respectively.

$$\begin{aligned}\Omega_{ri}^{MIN} &= \left(T_{ri}^{INBND} + T_i^{CRUISE} + T_{ri}^{OUTBND}\right) \\ &- \left(T_{ri^*}^{INBND} + T_{i^*}^{CRUISE} + T_{ri^*}^{OUTBND}\right)\end{aligned} \tag{7.25}$$

Moreover, τ_i^{MIN} and τ_i^{MAX} are, respectively, the earliest and latest time that flight leg i could be served. τ_i^{MIN} and τ_i^{MAX} are calculated in Equations (7.26) and (7.27), respectively.

$$\tau_i^{MIN} = \tau_r^{REQ} + T_{ri}^{INGRS} \qquad \forall(r,i): \mathcal{r}_r = \mathcal{r}_i^{INTND} \quad (7.26)$$

$$\begin{aligned} \tau_i^{MAX} &= \tau_r^{DLN} - T_i^{SRVREV} - T_{ri}^{EGRS} \\ &= \tau_i^{MIN} + \omega - \Omega_{ri}^{MIN} \end{aligned} \qquad \forall(r,i): \mathcal{r}_r = \mathcal{r}_i^{INTND} \quad (7.27)$$

7.3.2.4 Desired Trip Time

T_r^{DSRD}, the minimum trip time corresponding to the trip time of desired flight leg of $\mathcal{r}_r$ without any wait time, is defined in Equation (7.28). Since in a ubiquitous network, T_{ri}^{INBND} and T_{ri}^{OUTBND} for $\mathcal{f}_i = \mathcal{f}_r^{DSRD}$ are zero, $T_r^{DSRD} = T_{ri}^{DGATE} + T_i^{SRVREV} + T_{ri}^{AGATE}$.

$$T_r^{DSRD} = T_{ri}^{INGRS} + T_i^{SRVREV} + T_{ri}^{EGRS} \qquad \forall(r,i): \mathcal{f}_i = \mathcal{f}_r^{DSRD} \quad (7.28)$$

7.4 Network Representation

7.4.1 Node-based Network Representation

In the transportation problems where a *task* or *job* must be performed between pairs of origin and destination points (e.g., transporting goods and people or traversing a street), the network could be modeled in two ways: (1) arc-based and (2) node-based. In the arc-based representation, each physical arc in the transportation network corresponds to an arc in the modeled network. However, in node-based models, each task is collapsed into one node. For instance, in an arc-based representation, a flight from point a to point b is shown with node a, node b, and an arc from a to b, while in a node-based representation, the flight is modeled by node ab.

As a result, in the arc-based network representation, the arcs should be traversed (i.e., arc routing) to complete the corresponding tasks, while in node-based representations, the nodes must be visited (i.e., node routing). The arc-based networks are typically formulated as a capacitated multi-commodity network flow problem or *Arc Routing Problem (ARP)*. Specifically, ARPs could

take advantage of the structure of a graph representing a road network since, in such cases, each vertex represents a road junction where the degree of each vertex is usually small [134]. However, in a ubiquitous network of UAT pads, each vertex is a UAT pad and could be linked to numerous UAT pads.

Furthermore, the presence of time windows complicates the modeling choice. To incorporate the time windows in the arc-based network representation, the problem could be modeled in a *time-expanded network* [135] (also known as a *time-space network*). In time-expanded networks, the time is discretized over the planning horizon, and subsequently, the network is duplicated at each interval. Therefore, each node corresponds to a location at a specific time, while each arc shows the movements in space and time. Time-expanded networks are acyclic. However, the time discretization makes the network flow models challenging due to the problem size. *Fine* discretization provides a good approximation to the continuous-time problem; however, the problem could quickly get intractably large [135]. On the other hand, *coarse* discretizations are computationally more manageable while yielding poor approximations. Boland et al. [135] demonstrate that the loss of solution quality, i.e., the relative gap between the discretized and continuous-time optimal solutions, in *service network design problems (SNDP)* could be greater than 20%. Moreover, in the context of ARPs, a recent survey [134] suggests that many studies on ARP with time windows transform the problem into a node routing problem, and only one of the reviewed papers attempts to solve ARP with time windows exactly using an arc routing formulation.

Another complicating factor in the modeling choice is the flexibility of the passengers towards the pick-up and drop-off UAT pads. To incorporate this flexibility in a node-based

representation, each combination of possible pick-up and drop-off pairs should be modeled as one node, leading to a significant increase in the problem size.

Additionally, some attributes of the proposed UAT concept of operations contribute to the model selection. For instance, whether two arcs (i.e., flights) could be served immediately after each other depends on the requests they intend to serve (see Section 7.4.3 for more detail). This feature is easier to incorporate in a node-based representation. The reason is that in the node representation, an arc that intends to serve two requests is modeled using two separate nodes (i.e., flights), while it is represented as a single arc in an arc-based representation that should be traversed two times to serve the two requests.

In a closely related problem, Espinoza et al. [3] formulate the per-seat on-demand dial-a-flight problem (DAFP) as a multi-commodity network flow problem. However, they use a *time-activity* network, where the nodes include activities such as performing a flight (*gate* nodes), boarding a passenger (*direct* and *indirect loading* nodes), and being idle (*standby* nodes). Therefore, the network representation is node-based rather than arc-based, even though it is formulated as a network flow problem.

Consequently, we choose a node-based representation to model UAT fleet operation and formulate it as a capacitated location-allocation-routing problem with time windows and short repositioning elimination (CLARPTW-SRE). The UAT operator needs to decide which flight legs to perform (or which nodes to visit) (i.e., location), how to assign the requests to the flight legs (i.e., allocation), and how to route and schedule the capacitated aircraft to conduct these flight legs (i.e., routing). The CLARPTW-SRE at time t is represented by a directed graph $\mathcal{G}_t = (\mathcal{N}_t, \mathcal{A}_t)$ with the set of nodes $\mathcal{N}_t$ and the set of arcs $\mathcal{A}_t$. The arcs definitions for routing and scheduling resemble those presented in Yang et al. [78] and Bertsimas et al. [136].

Figure 7.6 depicts the transformation of the UAT network into the corresponding LARP network. Figure 7.6(a) illustrates two requests, namely, r_r and r_s, two flight legs, namely, f_i and f_j, and one aircraft, a_k, while Figure 7.6 (b) demonstrates the node-based model of the UAT network. The distance from vehicle k (i.e., UAT aircraft k) to location i (i.e., flight leg i) is the distance from $\boldsymbol{L}_{kt}^{AVL}$ to $\boldsymbol{S}_i$, i.e., D_{kit}^0. Lastly, the distance from location i to location j in Figure 7.6(b) is the distance between $\boldsymbol{E}_i$ and $\boldsymbol{S}_j$, i.e., D_{ij}.

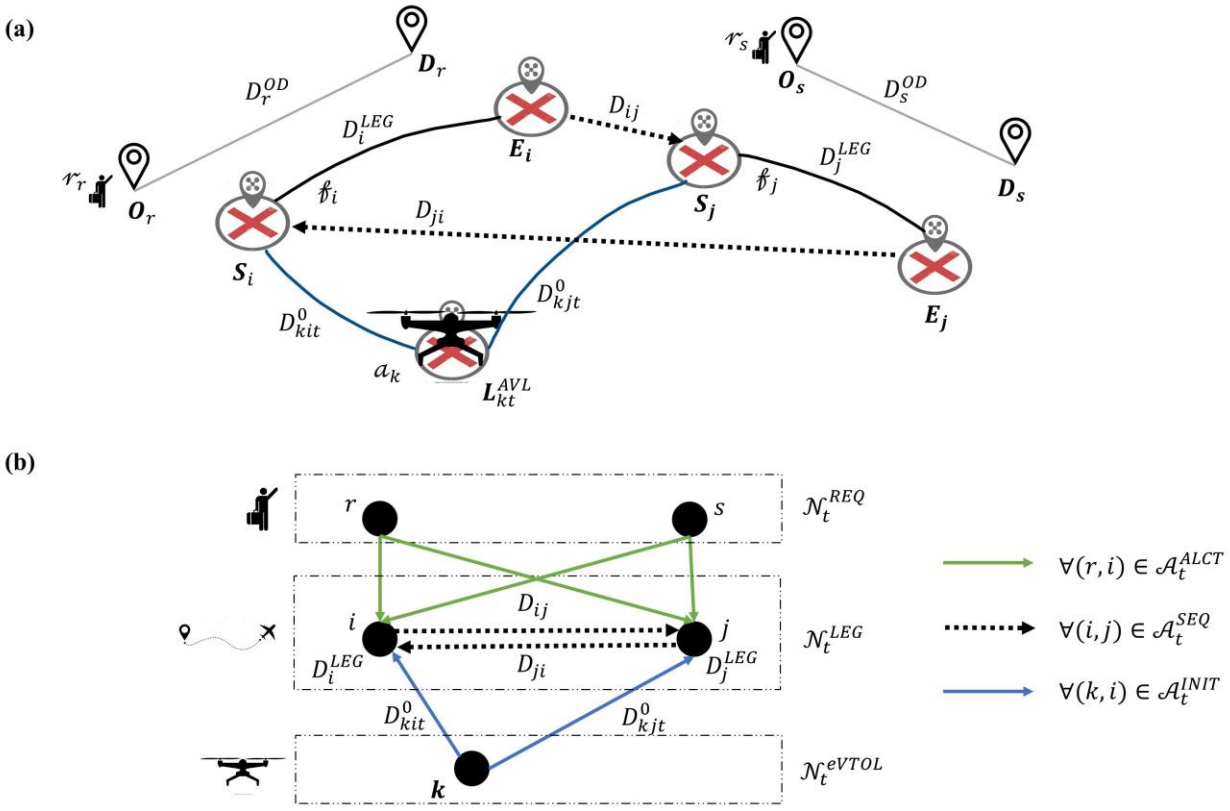

Figure 7.6 UAT network transformation, (a) UAT network, (b) node-based model

The following sections specify how the nodes and arcs are defined in the CLARPTW-SRE network.

7.4.2 Nodes

Each available UAT aircraft, candidate flight leg, and candidate request is represented as a node in $\mathcal{N}_t$. Thus, $\mathcal{N}_t$ includes three subsets: UAT aircraft ($\mathcal{N}_t^{eVTOL}$), flight legs ($\mathcal{N}_t^{LEG}$), and the requests $\left(\mathcal{N}_t^{REQ}\right)$, which are defined in Equations (7.29)-(7.31), respectively. $\mathcal{K}_t$ denotes the set of available UAT aircraft at time t, and is the same as $\mathcal{K}$if all the aircraft are functional at time t.

$$\mathcal{N}_t^{eVTOL} = \{k: a_k \in \mathcal{K}_t\} \tag{7.29}$$

$$\mathcal{N}_t^{LEG} = \{i: \not{p}_i \in \mathcal{F}_t^{CAND}\} \tag{7.30}$$

$$\mathcal{N}_t^{REQ} = \{r: \mathcal{r}_r \in \mathcal{R}_t^{CAND}\} \tag{7.31}$$

Additionally, $\mathcal{N}_t^{FLXSTRT} \subseteq \mathcal{N}_t^{REQ}$ and $\mathcal{N}_t^{FXDSTRT} \subseteq \mathcal{N}_t^{REQ}$, respectively defined in Equations (7.32) and (7.33), represent the nodes associated with requests with flexible pick-up UAT pad (i.e., $\mathcal{r}_r \in \mathcal{R}_t^{FLXSTRT}$) and requests with fixed pick-up UAT pad (i.e., $\mathcal{r}_r \in \mathcal{R}_t^{FXDSTRT}$).

$$\mathcal{N}_t^{FLXSTRT} = \{r: \mathcal{r}_r \in \mathcal{R}_t^{FLXSTRT}\} \tag{7.32}$$

$$\mathcal{N}_t^{FXDSTRT} = \{r: \mathcal{r}_r \in \mathcal{R}_t^{FXDSTRT}\} \tag{7.33}$$

Lastly, $\overline{\mathcal{N}_t^E} \subseteq \mathcal{N}_t^{LEG}$ denote the set of nodes associated with flight legs that do not end at the desired drop-off UAT pad of their intended request. $\overline{\mathcal{N}_t^E}$ is defined in Equation (7.34).

$$\overline{\mathcal{N}_t^E} = \{i \in \mathcal{N}_t^{LEG} \left| \boldsymbol{E}_i \neq \boldsymbol{E}_r^{DSRD}, \mathcal{r}_r = \mathcal{r}_i^{INTND} \right.\} \tag{7.34}$$

7.4.3 Arcs

$\mathcal{A}_t$ is composed of three subsets: the initial arcs from aircraft to flight legs ($\mathcal{A}_t^{INIT}$), the sequencing arcs between flight legs $\left(\mathcal{A}_t^{SEQ}\right)$, and the allocation arcs between requests and flight

legs ($\mathcal{A}_t^{ALCT}$). The existence of arc $(k, i) \in \mathcal{A}_t^{INIT}$ between aircraft k and flight leg i suggests that flight leg i could potentially be served as the first flight on aircraft k's route starting from $\boldsymbol{L}_{kt}^{AVL}$, while the arc $(i, j) \in \mathcal{A}_t^{SEQ}$ between the nodes of flight leg i and flight leg j specifies that flight leg j could potentially be served immediately after flight leg i. Lastly, arc $(r, i) \in \mathcal{A}_t^{ALCT}$ between request r and flight leg i implies that request r could be served by flight leg i. The following sections further outline the definitions of the arc.

7.4.3.1 Initial Arcs ($\mathcal{A}_t^{INIT}$)

Candidate flight legs have either the same starting UAT pad as or a different one from their intended request (i.e., $\mathcal{F}_t^{CAND} = \mathcal{F}_t^S \cup \overline{\mathcal{F}_t^S}$). Additionally, the first availability UAT pad of the UAT aircraft either coincides with the drop-off UAT pad of its passengers or not (i.e., $\mathcal{K}_t = \mathcal{K}_t^E \cup \overline{\mathcal{K}_t^E}$). Consequently, the tuple $(k, i) \in \mathcal{A}_t^{INIT}$ is defined depending on whether $\mathscr{f}_i \in \overline{\mathcal{F}_t^S}$ or not and $a_k \in \overline{\mathcal{K}_t^E}$ or not. As a result, $\mathcal{A}_t^{INIT}$, specified in Equation (7.35), is defined as the union of three subsets, namely, ${\mathcal{A}^0}_t^{PREC}$, ${\mathcal{A}^0}_t^{SUCC}$, and ${\mathcal{A}^0}_t^{FREE}$, defined in Equations (7.36)-(7.38), respectively.

$$\mathcal{A}_t^{INIT} = {\mathcal{A}^0}_t^{PREC} \cup {\mathcal{A}^0}_t^{SUCC} \cup {\mathcal{A}^0}_t^{FREE} \tag{7.35}$$

Figure 7.7 features connecting flight legs $\mathscr{f}_i \in \overline{\mathcal{F}_t^S}$, which aim to serve $\mathscr{r}_r \in \mathcal{R}_t^{CAND}$ but start at $\boldsymbol{L}_{kt}^{AVL}$ for $a_k \in \mathcal{K}$. Serving $\mathscr{r}_r = \mathscr{r}_i^{INTND}$ with $\mathscr{f}_i \in \overline{\mathcal{F}_t^S}$, which has a starting point different from the desired pick-up UAT pad of its intended request r, imposes a relocation to an undesired pick-up pad on $\mathscr{r}_r$. The relocation is justified if the starting point of $\mathscr{f}_i$ is an availability UAT pad that coincides with the desired drop-off UAT pad of the most recent request the aircraft is serving or has served. Therefore, $\mathscr{f}_i \in \overline{\mathcal{F}_t^S}$ could be conducted by aircraft k if $\boldsymbol{S}_i = \boldsymbol{L}_{kt}^{AVL}$ for $a_k \in \mathcal{K}_t^E$. The ending point of $\mathscr{f}_i$ could be either the desired drop-off UAT pad of $\mathscr{r}_r$ (i.e., $\boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}$) or the

desired pick-up point of $r_s \in \mathcal{R}_t^{CAND}$ (i.e., $\boldsymbol{E}_i = \boldsymbol{E}_s^{DSRD}$). Correspondingly, $f_i \in \mathcal{F}_{rt}^{KE}$ or $f_i \in \mathcal{F}_{rt}^{KS}$.

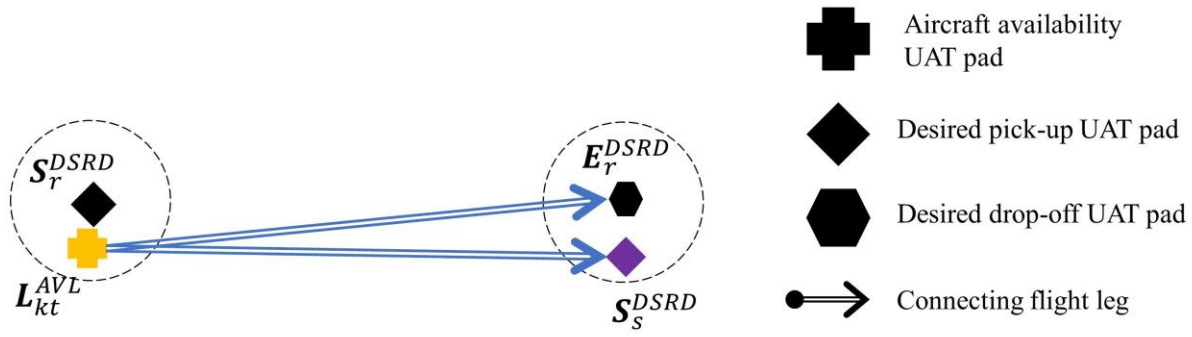

Figure 7.7 Connecting legs intending to serve request r which start at a first availability UAT pad of aircraft k

Let ${\mathcal{A}0}_t^{PREC}$ denote the set of (k, i) tuples where $f_i \in \overline{\mathcal{F}_t^S}$ starts at the availability UAT pad $a_k \in \mathcal{K}_t^E$. ${\mathcal{A}0}_t^{PREC}$ is defined in Equation (7.36).

$$ {\mathcal{A}0}_t^{PREC} = \{(k, i): a_k \in \mathcal{K}_t^E, f_i \in \overline{\mathcal{F}_t^S}, \boldsymbol{L}_{kt}^{AVL} = \boldsymbol{S}_i\} \tag{7.36} $$

On the other hand, the first availability UAT pad of $a_k \in \overline{\mathcal{K}_t^E}$ is different from the desired drop-off UAT pad of r_r (i.e., the most recent request it is serving or has served), imposing a relocation from an undesired drop-off pad on r_r. The ground-based relocation is justified if the first availability UAT pad of $a_k \in \overline{\mathcal{K}_t^E}$ is the desired pick-up UAT pad of another candidate request. Therefore, $f_i \in \mathcal{F}_t^S$ could succeed the first availability of $a_k \in \overline{\mathcal{K}_t^E}$ as long as $\boldsymbol{L}_{kt}^{AVL} = \boldsymbol{S}_j$.

Let ${\mathcal{A}0}_t^{SUCC}$ denote the set of (k, i) tuples where $f_i \in \mathcal{F}_t^S$ starts at the availability UAT pad $a_k \in \overline{\mathcal{K}_t^E}$. ${\mathcal{A}0}_t^{SUCC}$ is defined in Equation (7.37).

$$ {\mathcal{A}0}_t^{SUCC} = \{(k, i): a_k \in \overline{\mathcal{K}_t^E}, f_i \in \mathcal{F}_t^S, \boldsymbol{L}_{kt}^{AVL} = \boldsymbol{S}_i\} \tag{7.37} $$

The flight legs that start at the desired pick-up UAT pad of their intended request (i.e., $\not{f}_i \in \mathcal{F}_t^S$) are not constrained to be conducted by any specific UAT aircraft. Therefore, they could be served by any $a_k \in \mathcal{K}_t^E$, whose availability UAT pad coincides with the desired drop-off UAT pad of the most recent request it is serving or has served. ${\mathcal{A}^0}_t^{FREE}$ denotes the set of (k, i) tuples where $\not{f}_i \in \mathcal{F}_t^S$ is the flight leg served from the availability UAT pad of $a_k \in \mathcal{K}_t^E$. Since $\boldsymbol{L}_{kt}^{AVL}$ is not necessarily the same as $\boldsymbol{S}_i$ for $(k, i) \in {\mathcal{A}^0}_t^{FREE}$, an empty flight leg, which distance is not within the *short* range, could be conducted prior to serving flight leg i by aircraft k. ${\mathcal{A}^0}_t^{FREE}$ is defined in Equation (7.38).

$$
\begin{aligned}
{\mathcal{A}^0}_t^{FREE} = \{(k,i):\, & a_k \in \mathcal{K}_t^E, \not{f}_i \in \mathcal{F}_t^S, \\
& D_{kit}^0 \notin (0, (1+\epsilon)\Delta^{EMPTY})\}
\end{aligned}
\tag{7.38}
$$

7.4.3.2 Sequencing Arcs ($\mathcal{A}_t^{SEQ}$)

Candidate flight legs could be classified as three (not mutually exclusive) groups: (1) flight legs that start at a different UAT pad from the desired pick-up UAT pad of their intended request (i.e., $\not{f}_i \in \overline{\mathcal{F}_t^S}$), (2) flight legs that end at a different UAT pad from the desired dop-off UAT pad of their intended request (i.e., $\not{f}_i \in \overline{\mathcal{F}_t^E}$), (3) flight legs that start at the desired pick-up UAT pad (i.e., $\not{f}_i \in \mathcal{F}_t^S$) or end at the desired drop-off UAT pad (i.e., $\not{f}_i \in \mathcal{F}_t^E$) of their intended request.

Sequencing arcs $(i, j) \in \mathcal{A}_t^{SEQ}$, which cover how two flight legs can follow each other, are divided into three classes, namely, *preceding* ($\mathcal{A}_t^{PREC}$), *succeeding* ($\mathcal{A}_t^{SUCC}$), and *free* ($\mathcal{A}_t^{FREE}$), depending on the groups $\not{f}_i$ and $\not{f}_j$ are classified as. $\mathcal{A}_t^{SEQ}$ is specified in Equation (7.39).

$$
\mathcal{A}_t^{SEQ} = \mathcal{A}_t^{PREC} \cup \mathcal{A}_t^{SUCC} \cup \mathcal{A}_t^{FREE} \tag{7.39}
$$

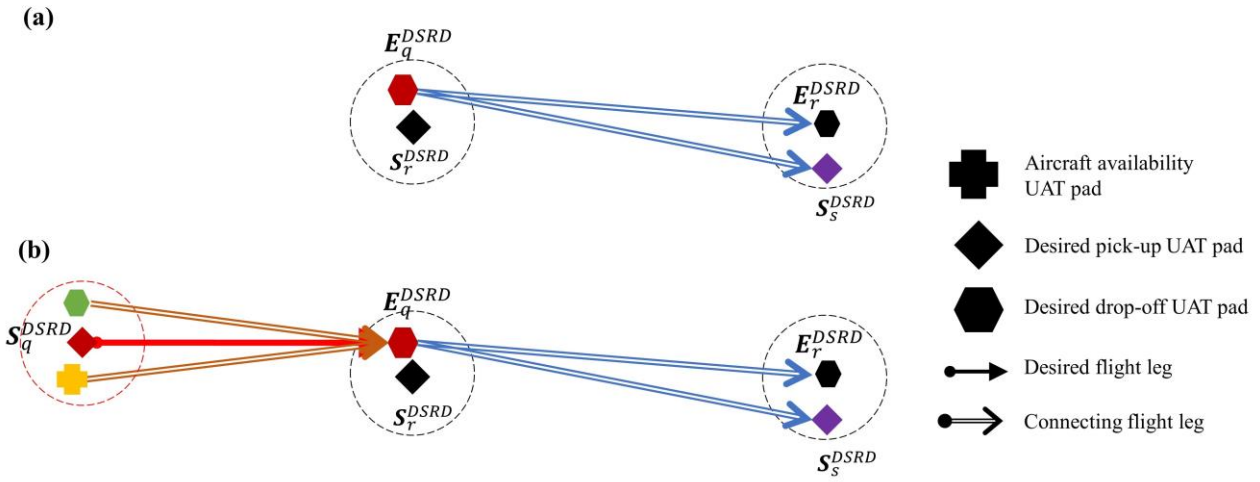

Figure 7.8 Connecting legs intending to serve request r that do not start at the desired pick-up UAT pad of request r

Figure 7.8(a) features connecting flight legs $\not{f}_i \in \overline{\mathcal{F}_t^S}$, which aim to serve $r_r \in \mathcal{R}_t^{CAND}$ but starts at a UAT pad different from the desired pick-up pad of r_r (i.e., $\boldsymbol{S}_r^{DSRD}$). Serving $r_r = r_i^{INTND}$ with $\not{f}_i \in \overline{\mathcal{F}_t^S}$, which has a starting point different from the desired pick-up UAT pad of its intended request r, imposes a relocation to an undesired pick-up UAT pad. This ground-based relocation is justified if the starting point of $\not{f}_i$ is the desired drop-off UAT pad of another candidate request. Therefore, $\not{f}_m \in \mathcal{F}_t^E$ could precede $\not{f}_i$ as long as $\boldsymbol{S}_i = \boldsymbol{E}_m$. The ending point of $\not{f}_i$, as depicted in Figure 7.8(b), is either the desired drop-off UAT pad of r_r (i.e., $\boldsymbol{E}_i = \boldsymbol{E}_r^{DSRD}$) or the desired pick-up UAT pad of $r_s \in \mathcal{R}_t^{CAND}$ (i.e., $\boldsymbol{E}_i = \boldsymbol{S}_s^{DSRD}$). Correspondingly, $\not{f}_i \in \mathcal{F}_{rt}^{EE}$ or $\not{f}_i \in \mathcal{F}_{rt}^{ES}$.

Let $\mathcal{A}_t^{PREC}$ denote the set of (m, i) tuples where $\not{f}_m \in \mathcal{F}_t^E$ could precede $\not{f}_i \in \overline{\mathcal{F}_t^S}$. $\mathcal{A}_t^{PREC}$ is defined in Equation (7.40).

$$\mathcal{A}_t^{PREC} = \{(m, i): \not{f}_i \in \overline{\mathcal{F}_t^S}, \not{f}_m \in \mathcal{F}_t^E, \boldsymbol{E}_m = \boldsymbol{S}_i\} \tag{7.40}$$

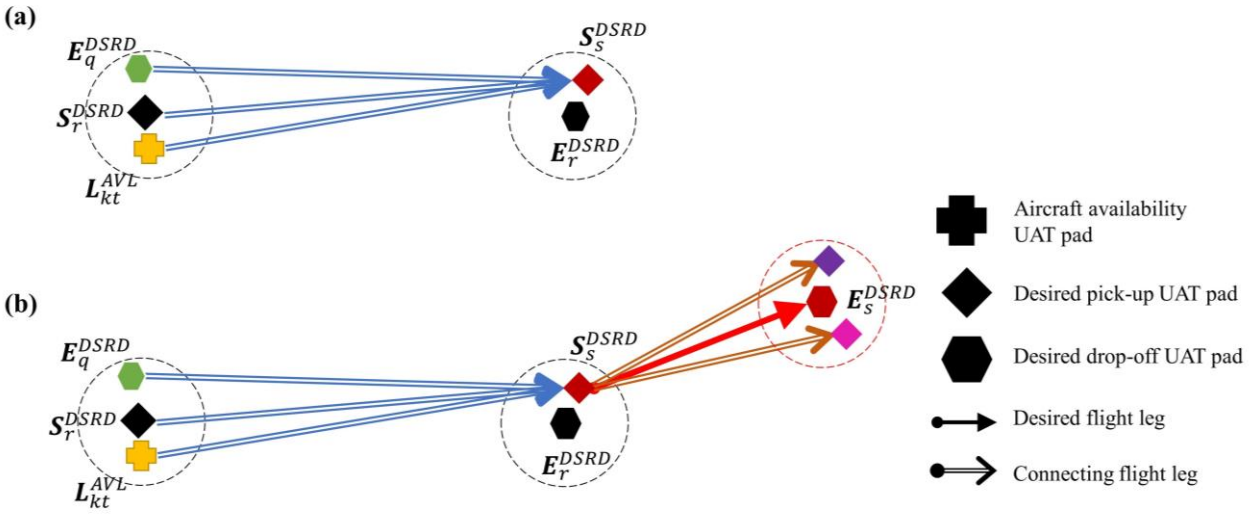

Figure 7.9 Connecting legs intending to serve request r that do not end at the desired drop-off UAT pad of request r

Figure 7.9(a) depicts flight legs $f_i \in \overline{\mathcal{F}_t^E}$, which aim to serve $r_r \in \mathcal{R}_t^{CAND}$ but ends at a UAT pad different from the desired drop-off UAT pad of r_r (i.e., $\boldsymbol{E}_r^{DSRD}$). Serving $r_r = r_i^{INTND}$ with $f_i \in \overline{\mathcal{F}_t^E}$, which has an ending point different from the desired drop-off UAT pad of its intended request r, imposes a relocation to an undesired drop-off UAT pad on r_r. This ground-based relocation is justified if the ending point of f_i is the desired pick-up UAT pad of another candidate request. Therefore, $f_j \in \mathcal{F}_t^S$ could succeed f_i as long as $\boldsymbol{E}_i = \boldsymbol{S}_j$. The starting point of f_i, as depicted in Figure 7.9(b), is either the desired pick-up UAT pad of r_r (i.e., $\boldsymbol{S}_i = \boldsymbol{S}_r^{DSRD}$), the desired drop-off pad of $r_q \in \mathcal{R}_t^{CAND}$ (i.e., $\boldsymbol{S}_i = \boldsymbol{E}_q^{DSRD}$), or the first availability UAT pad of $a_k \in \mathcal{K}_t^E$ (i.e., $\boldsymbol{S}_i = \boldsymbol{L}_{kt}^{AVL}$). Correspondingly, $f_i \in \mathcal{F}_{rt}^{SS}$, $f_i \in \mathcal{F}_{rt}^{ES}$, or $f_i \in \mathcal{F}_{rt}^{KS}$.

Let $\mathcal{A}_t^{SUCC}$ denote the set of (i, j) tuples where $f_j \in \mathcal{F}_t^S$ could succeed $f_i \in \overline{\mathcal{F}_t^E}$. $\mathcal{A}_t^{SUCC}$ is defined in Equation (7.41).

$$\mathcal{A}_t^{SUCC} = \{(i,j): f_i \in \overline{\mathcal{F}_t^E}, f_j \in \mathcal{F}_t^S, \boldsymbol{E}_i = \boldsymbol{S}_j\} \tag{7.41}$$

Lastly, there is no constraint on flight legs that could succeed $\not{f}_i \in \mathcal{F}_t^E$ since these flight legs end at the desired drop-off UAT pad of their intended request. Similarly, there is no constraint on the flight legs that could precede $\not{f}_j \in \mathcal{F}_t^S$ since these legs start at the desired pick-up UAT pad of their intended request. Let $\mathcal{A}_t^{FREE}$ denote (i,j) tuples where $\not{f}_j \in \mathcal{F}_t^S$ could follow $\not{f}_i \in \mathcal{F}_t^E$. These two are not necessarily connected directly, and an empty flight leg might be conducted between the two legs. $\mathcal{A}_t^{FREE}$ is defined in Equation (7.42).

$$\mathcal{A}_t^{FREE} = \{(i,j): \not{f}_i \in \mathcal{F}_t^E, \not{f}_j \in \mathcal{F}_t^S, \; D_{ij} \notin (0, (1+\epsilon)\Delta^{EMPTY})\} \tag{7.42}$$

It is worth noting that $\mathcal{A}_t^{PREC}$ and $\mathcal{A}_t^{SUCC}$ are defined so that *intermediary relocation pad*, the UAT pad that is not the desired one for any of the involved requests, is avoided. To further elaborate, when defining $\mathcal{A}_t^{PREC}$, not all the flight legs that end at $\boldsymbol{S}_i$ could precede $\not{f}_i \in \overline{\mathcal{F}_t^S}$. In other words, $\{(m,i) \mid \not{f}_i \in \overline{\mathcal{F}_t^S}, \not{f}_m \in \mathcal{F}_t^{CAND}, \boldsymbol{E}_m = \boldsymbol{S}_i\} \not\subset \mathcal{A}_t^{PREC}$ since the ending point of $\not{f}_m \in \mathcal{F}_t^{CAND}$ is not necessarily the desired drop-off pad of its intended request (i.e., $\boldsymbol{E}_m \neq \boldsymbol{E}_q^{DSRD}$ for $r_q = r_m^{INTND}$). If $\not{f}_m \in \overline{\mathcal{F}_t^E}$ were to be followed by $\not{f}_i \in \overline{\mathcal{F}_t^S}$, $r_q (= r_m^{INTND})$ and r_r $(= r_i^{INTND})$ are both relocated to $\boldsymbol{S}_i = \boldsymbol{E}_m$, and therefore, an unnecessary relocation is incurred. As a result, $\boldsymbol{S}_i = \boldsymbol{E}_m$ becomes an intermediary relocation pad in the network for r_q and r_r since it is not the one desired by either r_q or r_r. In a network with a limited number of UAT pads, looking for such intermediary pads, where multiple requests could be relocated to or from, could be an alternative operational policy. However, in a ubiquitous network, with a theoretically infinite number of UAT pads over the space, relocating requests to an intermediary pad is more challenging since it requires a search over the entire accessible space. Furthermore, one of the objectives is to

minimize the number of relocations and, thus, passengers inconvenience. Utilizing intermediary UAT pads would result in two requests being relocated rather than one, which is not desirable.

To further elaborate, Figure 7.10 demonstrates how a flight leg ending at $\boldsymbol{S}_i$ could not precede $\not{f}_i \in \mathcal{F}_t^{CAND}$. In this figure, four requests, namely, r_r, r_s, r_p, and r_q, and their desired pick-up and drop-off UAT pads are depicted. Additionally, $\boldsymbol{S}_s^{DSRD}$ and $\boldsymbol{E}_p^{DSRD}$ overlap. In Figure 7.10(a), a connecting leg from $\boldsymbol{S}_r^{DSRD}$ to $\boldsymbol{S}_s^{DSRD}$ is introduced to serve r_r, while in Figure 7.10(b), another connecting leg from $\boldsymbol{E}_p^{DSRD}$ to $\boldsymbol{E}_q^{DSRD}$ is created to serve r_q. As a result, $\boldsymbol{S}_r^{DSRD} \rightarrow \boldsymbol{S}_s^{DSRD} \rightarrow \boldsymbol{E}_s^{DSRD}$ and $\boldsymbol{S}_p^{DSRD} \rightarrow \boldsymbol{E}_p^{DSRD} \rightarrow \boldsymbol{E}_q^{DSRD}$ are valid itineraries in Figure 7.10(a) and Figure 7.10(b), respectively. However, as shown in Figure 7.10(c), $\boldsymbol{S}_r^{DSRD} \rightarrow \boldsymbol{S}_s^{DSRD} \rightarrow \boldsymbol{E}_q^{DSRD}$ is not a valid sequence of the connecting legs to serve r_r and r_q since, in this case, $\boldsymbol{S}_s^{DSRD}$ would act as an intermediary relocation pad where both r_r and r_q have to relocate to.

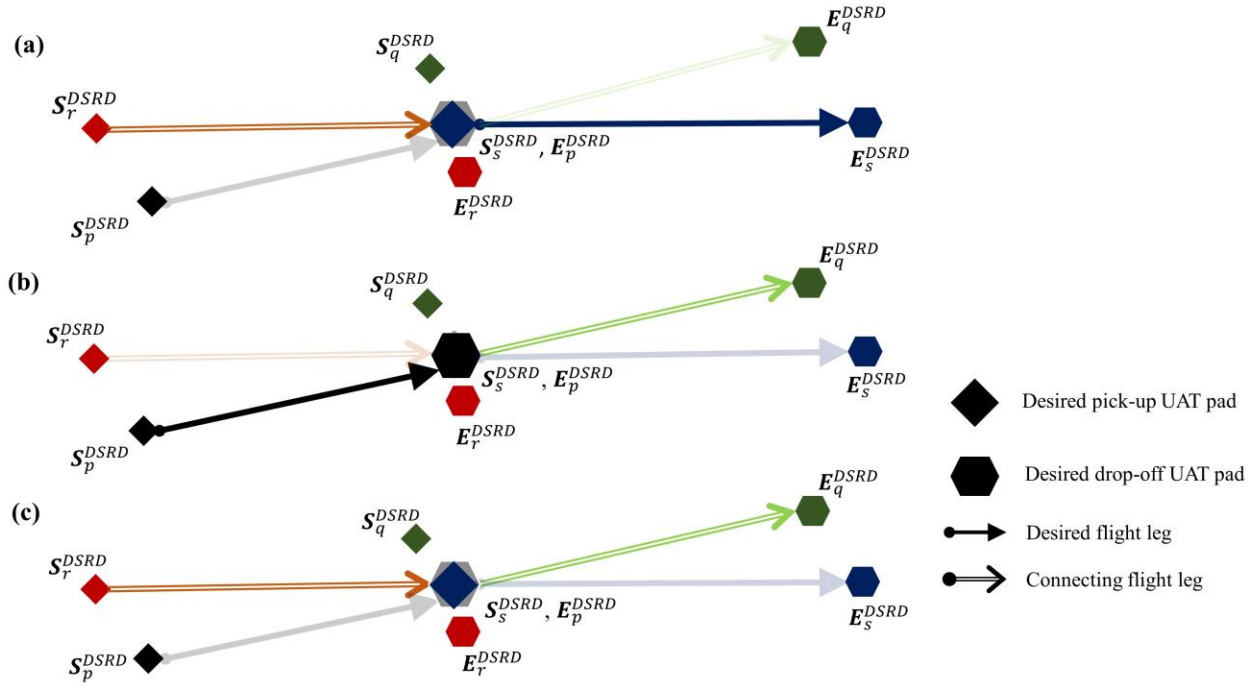

Figure 7.10 Depiction of preceding flight legs (a) valid, (b) valid, and (c) invalid sequence of flight legs

Similarly, when defining $\mathcal{A}_t^{SUCC}$, not all the flight legs starting at $\boldsymbol{E}_i$ could succeed $\not{f}_i \in \overline{\mathcal{F}_t^E}$. In other words, $\{(i,j)|\not{f}_i \in \overline{\mathcal{F}_t^E}, \not{f}_j \in \mathcal{F}_t^{CAND}, \boldsymbol{E}_i = \boldsymbol{S}_j\} \not\subset \mathcal{A}_t^{SUCC}$. Furthermore, for $\not{f}_i \in \overline{\mathcal{F}_t^E}$ ending at $\boldsymbol{E}_i$, it is possible all $\not{f}_j \in \mathcal{F}_t^S$ with $\boldsymbol{E}_i = \boldsymbol{S}_j$ get rejected. In other words, the intended request of $\not{f}_i$ might be unnecessarily relocated to $\boldsymbol{E}_i$, where no succeeding legs starting at $\boldsymbol{E}_i$ will be served. Let $\mathcal{N}_{it}^{SUCC}$ denote the set of indices of flight legs that could succeed $\not{f}_i \in \overline{\mathcal{F}_t^E}$ as of time t. $\mathcal{N}_{it}^{SUCC}$ is defined in Equation (7.43).

$$\mathcal{N}_{it}^{SUCC} = \{j: (I,j) \in \mathcal{A}_t^{SUCC}, I = i\} \tag{7.43}$$

It is worth noting that connecting leg $\not{f}_i \in \mathcal{F}_{rt}^{KE} \cup \mathcal{F}_{rt}^{EE} \cup \mathcal{F}_{rt}^{ES}$ should be served only if an aircraft is available at its starting pad or one of its preceding flight legs is conducted. This constraint is implicitly addressed in network construction by allowing $\not{f}_i$ to be reached in the network only from an available aircraft at $\boldsymbol{S}_i$ or its preceding legs. As a result, when connecting leg $\not{f}_i \in \mathcal{F}_{rt}^{KE} \cup \mathcal{F}_{rt}^{EE} \cup \mathcal{F}_{rt}^{ES}$ is served, *preceding constraints* are implicitly satisfied. In contrast, serving connecting leg $\not{f}_i \in \mathcal{F}_{rt}^{KS} \cup \mathcal{F}_{rt}^{ES}$ does not guarantee that any of its succeeding flight legs will be served. Therefore, *succeeding constraints* must be explicitly specified in the optimization problem. Nonetheless, in rare cases, while $\not{f}_i \in \mathcal{F}_{rt}^{ES}$ is in service, none of its succeeding flight legs may get served in the re-optimization.

7.4.3.3 *Allocation Arcs* ($\mathcal{A}_t^{ALCT}$)

Allocation arcs specify if $r_r \in \mathcal{R}_t^{CAND}$ could be served by $\not{f}_i \in \mathcal{F}_t^{CAND}$. The tuple $(r,i) \in \mathcal{A}_t^{ALCT}$ implies that starting UAT pad of $\not{f}_i$ is within accessible distance of the origin of r_r and the ending UAT pad of $\not{f}_i$ is within accessible distance of the destination of r_r. $\mathcal{A}_t^{ALCT}$ is defined in Equation (7.44).

$$\mathcal{A}_t^{ALCT} = \{(r,i): dist(\boldsymbol{O}_r, \boldsymbol{S}_i) \leq \Delta^{ACCESS}, \\ dist(\boldsymbol{D}_r, \boldsymbol{E}_i) \leq \Delta^{ACCESS}, \\ \mathcal{r}_r \in \mathcal{R}_t^{CAND}, \mathcal{f}_i \in \mathcal{F}_t^{CAND}\} \tag{7.44}$$

Furthermore, let $\mathcal{A}_t^{INTND} \subseteq \mathcal{A}_t^{ALCT}$ denote the set of (r,i) tuples where request r is the intended request of flight leg i. $\mathcal{A}_t^{INTND}$ is defined in Equation (7.45).

$$\mathcal{A}_t^{INTND} = \{(r,i): \mathcal{f}_i \in \mathcal{F}_t^{CAND}, \mathcal{r}_r = \mathcal{r}_i^{INTND}\} \tag{7.45}$$

7.5 Network Reduction

Given the time windows associated with the flight legs and requests, the network presented in the previous section could be reduced to increase computational efficiency. $\mathcal{A}_t^{ALCTTW}$, defined in Equation (7.46), specifies the allocation arcs after taking the time windows into account. Request r could be assigned to flight leg i if:

i. The minimum delay incurred by assigning $\mathcal{r}_r$ to $\mathcal{f}_i$ is smaller than the maximum acceptable delay (i.e., $\Omega_{ri}^{MIN} \leq \omega$);

ii. The passenger group of $\mathcal{r}_r$ should be able to get to $\mathcal{f}_i$ before the lastest start time of $\mathcal{f}_i$ (i.e., $\tau_{rt}^{SRVC} + T_{ri}^{INGRS} \leq \tau_i^{MAX}$); and

iii. $\mathcal{f}_i$ must start before the request r maxes out the modified maximum acceptable delay (i.e., $\tau_{rt}^{SRVC} + T_{ri}^{DGATE} + \omega' \geq \tau_i^{MIN}$);

$$\mathcal{A}_t^{ALCTTW} = \big\{(r,i) \in \mathcal{A}_t^{ALCT}: \Omega_{ri}^{MIN} \leq \omega, \\ \tau_{rt}^{SRVC} + T_{ri}^{INGRS} \leq \tau_i^{MAX}, \\ \tau_{rt}^{SRVC} + T_{ri}^{DGATE} + \omega' \geq \tau_i^{MIN}, \\ \mathcal{r}_r \in \mathcal{R}_t^{CAND}, \mathcal{f}_i \in \mathcal{F}_t^{CAND}\big\} \tag{7.46}$$

Where ω' , the adjusted maximum acceptable delay, takes into account the time that the passenger's relocation on the ground is in the direction of their destination. In this case, the ground transportation would shorten the flight time, and as a consequence, the extra time could be spent on reaching the starting point of the flight. ω' is defined in Equation (7.47).

$$\omega' = \omega \left(1 + \frac{v^{DRIVE}}{v^{AIR}}\right) \tag{7.47}$$

If a flight leg i cannot serve its intended request, it is not included in the candidate flight legs. In other words, if $(r, i) \notin \mathcal{A}_t^{ALCTTW}$ for $(r, i) \in \mathcal{A}_t^{INTND}$, flight leg i should be excluded from the set of candidate requests. Accordingly, $\dot{\mathcal{F}}_t^{CAND}$, defined in Equation (7.48), specifies the adjusted candidate flight legs after excluding the flight legs that could not possibly serve their intended requests.

$$\dot{\mathcal{F}}_t^{CAND} = \{\mathbb{f}_i \in \mathcal{F}_t^{CAND} : (r, i) \in \mathcal{A}_t^{ALCTTW} \cap \mathcal{A}_t^{INTND}\} \tag{7.48}$$

Furthermore, considering the availability time of the UAT aircraft and the time windows for the flight legs, the size of $\mathcal{A}_t^{INIT}$ and $\mathcal{A}_t^{SEQ}$ could be reduced. The arcs between the aircraft and flight legs could be limited to (k, i) tuples for which aircraft k could reach flight leg i before the latest acceptable time to serve flight leg i. Therefore, $\mathcal{A}_t^{INITTW}$ is calculated in Equation (7.49).

$$\mathcal{A}_t^{INITTW} = \left\{(k, i) \in \mathcal{A}_t^{INIT} : \mathbb{f}_i \in \dot{\mathcal{F}}_t^{CAND}, \; \tau_{kt}^{AVL} + T_{kit}^{0} \leq \tau_i^{MAX}\right\} \tag{7.49}$$

Similarly, for $(i, j) \in \mathcal{A}_t^{SEQ}$, an aircraft could potentially serve flight leg j after flight leg i if it conducted $\mathbb{f}_i$ by starting at the earliest allowable start time, performed the repositioning flight between $\mathbb{f}_i$ and $\mathbb{f}_j$, and reached the starting point of $\mathbb{f}_j$ before its latest allowable start time. As a result, $\mathcal{A}_t^{SEQTW}$ is defined in Equation (7.50).

$$\mathcal{A}_t^{SEQTW} = \{(i,j) \in \mathcal{A}_t^{SEQ}: \not{f}_i, \not{f}_j \in \dot{\mathcal{F}}_t^{CAND}, \\ \tau_i^{MIN} + T_i^{SRVREV} + T_{ij} \leq \tau_j^{MAX}\} \tag{7.50}$$

Consequently, $(k,i) \in \mathcal{A}_t^{INITTW}$ and $(i,j) \in \mathcal{A}_t^{SEQTW}$ would determine what flight legs are feasible to be served under the time constraints. $\ddot{\mathcal{F}}_t^{CAND}$, defined in Equation (7.51), denotes the candidate flight legs that are feasible to be served directly by an aircraft or followed by another flight leg given the time windows.

$$\ddot{\mathcal{F}}_t^{CAND} \\ = \{\not{f}_i: (k,i) \in \mathcal{A}_t^{INITTW}\} \\ \cup \{\not{f}_i: (j,i) \in \mathcal{A}_t^{SEQTW}\} \tag{7.51}$$

Furthermore, for $\not{f}_i \in \overline{\mathcal{F}_t^E}$, none of $\not{f}_j$ (for $\forall j \in \mathcal{N}_{it}^{SUCC}$) that must follow $\not{f}_i$ might be included in the adjusted candidate flight legs $\ddot{\mathcal{F}}_t^{CAND}$. As a result, Algorithm 1 seeks to modify $\tilde{\mathcal{N}}_{it}^{SUCC}$, $\overline{\tilde{\mathcal{F}}_t^E}$, and $\tilde{\mathcal{F}}_t^{CAND}$ accordingly. The algorithm repeatedly updates the succeeding flights of $\not{f}_i \in \overline{\mathcal{F}_t^E}$, and if the set of succeeding nodes of $\not{f}_i$ is empty, $\not{f}_i$ will be excluded from the set of candidate flight legs. The steps are repeated until no new flight leg is excluded from the set of candidate flight legs.

In conclusion, Equations (7.52) define $\tilde{\mathcal{N}}_t^{LEG}$. The symbol ~ suggests the modified variable after network reduction. Accordingly, $\tilde{\mathcal{A}}_t^{INIT}$ and $\tilde{\mathcal{A}}_t^{SEQ}$ are defined in Equations (7.53) and (7.54), respectively.

$$\tilde{\mathcal{N}}_t^{LEG} = \{i: \not{f}_i \in \tilde{\mathcal{F}}_t^{CAND}\} \tag{7.52}$$

$$\tilde{\mathcal{A}}_t^{INIT} = \{(k,i) \in \mathcal{A}_t^{INITTW}: i \in \tilde{\mathcal{N}}_t^{LEG}\} \tag{7.53}$$

$$\tilde{\mathcal{A}}_t^{SEQ} = \{(i,j) \in \mathcal{A}_t^{SEQTW}: i,j \in \tilde{\mathcal{N}}_t^{LEG}\} \tag{7.54}$$

Algorithm 1: Modified Candidate Flight Legs

Inputs: $\mathcal{A}_t^{SUCC}, \mathcal{A}_t^{SEQTW}, \overline{\mathcal{F}_t^E}, \ddot{\mathcal{F}}_t^{CAND}$

Outputs: $\tilde{\mathcal{F}}_t^{CAND}, \overline{\tilde{\mathcal{F}}_t^E}, \tilde{\mathcal{N}}_{it}^{SUCC}$

$CONTINUE \leftarrow TRUE$

$\tilde{\mathcal{F}}_t^{CAND} \leftarrow \ddot{\mathcal{F}}_t^{CAND}$

while $CONTINUE$ **do**:

$$\overline{\tilde{\mathcal{F}}_t^E} \leftarrow \tilde{\mathcal{F}}_t^{CAND} \cap \overline{\mathcal{F}_t^E} \tag{7.55}$$

for $\not{f}_i \in \overline{\tilde{\mathcal{F}}_t^E}$ **do**:

$$\tilde{\mathcal{N}}_{it}^{SUCC} \leftarrow \left\{j: (I,j) \in \mathcal{A}_t^{SUCC}, I = i, \not{f}_j \in \tilde{\mathcal{F}}_t^{CAND}\right\} \tag{7.56}$$

end for

$\hat{\mathcal{F}}_t^{SUCC} \leftarrow \{\not{f}_i \in \tilde{\mathcal{F}}_t^{CAND}: \tilde{\mathcal{N}}_{it}^{SUCC} = \emptyset\}$

if $\hat{\mathcal{F}}_t^{SUCC} \neq \emptyset$ **then:**

$$\tilde{\mathcal{F}}_t^{CAND} \leftarrow \tilde{\mathcal{F}}_t^{CAND} \backslash \hat{\mathcal{F}}_t^{SUCC} \tag{7.57}$$

else:

$CONTINUE \leftarrow FALSE$

end if

end

Consequently, $\tilde{\mathcal{N}}_t^{eVTOL}$ is defined in Equation (7.58). Furthermore, $\tilde{\mathcal{A}}_t^{ALCT}$ and $\tilde{\mathcal{N}}_t^{REQ}$ are defined in Equations (7.59) and (7.60), respectively. $\tilde{\mathcal{A}}_t^{ALCT}$ ensures that request r could be assigned to flight i only if flight leg i could be served given the time windows (i.e., $i \in \tilde{\mathcal{N}}_t^{LEG}$). $\tilde{\mathcal{N}}_t^{REQ}$ includes the requests that could be assigned to at least one flight leg, and therefore, $(r,i) \in \tilde{\mathcal{A}}_t^{ALCT}$.

$$\tilde{\mathcal{N}}_t^{eVTOL} = \{k: (k,i) \in \tilde{\mathcal{A}}_t^{INIT}\} \tag{7.58}$$

$$\tilde{\mathcal{A}}_t^{ALCT} = \left\{(r,i) \in \mathcal{A}_t^{ALCTTW}: i \in \tilde{\mathcal{N}}_t^{LEG}\right\} \tag{7.59}$$

$$\tilde{\mathcal{N}}_t^{REQ} = \{r: (r,i) \in \tilde{\mathcal{A}}_t^{ALCT}\} \tag{7.60}$$

In summary, to reduce the size of the problem and accelerate the solution time, the reduced network $\tilde{\mathcal{G}}_t = \left(\tilde{\mathcal{N}}_t, \tilde{\mathcal{A}}_t\right)$ would replace $\mathcal{G}_t = (\mathcal{N}_t, \mathcal{A}_t)$.

7.6 Concluding Remark

The network associated with UAT fleet operation could be modeled in two ways: (1) arc-based and (2) node-based. In the arc-based representation, each physical arc in the transportation network corresponds to an arc in the modeled network. However, in node-based representation, each task (e.g., transporting goods or people, traversing a street) is collapsed into one node.

The presence of time windows complicates the modeling choice. To incorporate the time windows in the arc-based network representation, the problem could be modeled in a time-expanded network. However, the time discretization makes the network flow models challenging due to the problem size. Another complicating factor in the modeling choice is the flexibility of the passengers towards pick-up and drop-off UAT pads. To incorporate this flexibility in a node-based representation, each combination of possible pick-up and drop-off pairs should be modeled as one node, leading to a significant increase in the problem size. Lastly, some attributes of the proposed UAT concept of operations (e.g., the sequence of flight legs given the requests they intend to serve) are easier to incorporate in a node-based representation.

Consequently, this chapter presents a node-based representation of the network associated with the UAT problem. The revenue-generating flight legs are nodes where the aircraft should be routed to. Additionally, the requests are represented as another set of nodes that should be allocated to the conducted revenue-generating flight legs. As a result, the network representing the UAT problem is modeled as a location-allocation-routing problem, where some of the candidate revenue-generating flight legs are performed (i.e., the location is open), the requests will be assigned to these flight legs (i.e., the allocation), and the aircraft are routed to serve the flight legs (i.e., routing).

Given the dynamic and on-demand nature of the problem and the ubiquitous network of UAT pads, the set of candidate requests and flight legs vary at each decision epoch. Therefore, this chapter first defines the relevant entities at each decision epoch. Before transforming the network, the metrics of the network, including the distances and times, are defined in detail. Subsequently, the UAT physical network is transformed into a node-based network. Finally, the network is modified to reduce the network size and lower the solution time.

Chapter 8 presents the formulation for the network representing the Capacitated Location-Allocation-Routing Problem with Time Windows and Short Repositioning Elimination (CLARPTW-SRE) defined in this chapter.

Chapter 8 Capacitated Location-Allocation-Routing Problem With Time Windows and Short Repositioning Elimination (CLARPTW-SRE): Formulation

8.1 Overview

The Capacitated Location-Allocation-Routing Problem with Time Windows and Short Repositioning Elimination (CLARPTW-SRE) is the building block of the decision-making policy of UAT fleet operation. At the beginning of decision epoch $e \in \mathcal{E}$, i.e., $t = \tau_{v_e^S}$, CLARPTW-SRE is applied to retrieve the routing and scheduling of the UAT aircraft and requests. This chapter represents the CLARPTW-SRE model solved at each decision epoch given the network $\mathcal{G}_t = (\mathcal{N}_t, \mathcal{A}_t)$.

8.2 Parameters

CLARPTW-SRE is defined over the network $\mathcal{G}_t = (\mathcal{N}_t, \mathcal{A}_t)$, with $\mathcal{N}_t$ representing the set of nodes and $\mathcal{A}_t$ representing the set of arcs as of time t. Let $\mathcal{N}_t^{eVTOL} \subset \mathcal{N}_t$ denote the set of nodes associated with the UAT aircraft that the UAT operator could dispatch at time t. $\mathcal{N}_t^{REQ} \subset \mathcal{N}_t$ is the set nodes associated with requests to (re)allocate to flight legs at time t, where $\mathcal{N}_t^{REQ} = \mathcal{N}_t^{UNASGN} \cup \mathcal{N}_t^{FLXSTRT} \cup \mathcal{N}_t^{FXDSTRT}$. $\mathcal{N}_t^{UNASGN} \subseteq \mathcal{N}_t^{REQ}$ presents the set of nodes associated with the unassigned requests. $\mathcal{N}_t^{FLXSTRT} \subseteq \mathcal{N}_t^{REQ}$ denotes the nodes related to the requests that were accepted in the previous decision epochs and, therefore, must be served, but their pick-up UAT pad is flexible. Similarly, $\mathcal{N}_t^{FXDSTRT} \subseteq \mathcal{N}_t^{REQ}$ denotes the nodes related to the requests that were accepted in the previous decision epochs and, therefore, must be served, but they have already left their origin and their pick-up UAT pad is fixed.

Furthermore, let $\mathcal{N}_t^{LEG} \subset \mathcal{N}_t$ denote the set of nodes associated with revenue-generating flight legs available at time t to serve the requests. $\overline{\mathcal{N}_t^E} \subseteq \mathcal{N}_t^{LEG}$defines a set of nodes associated with flight legs that do not end at the desired UAT pad of their intended request, implying that a flight leg should succeed $i \subseteq \overline{\mathcal{N}_t^E}$ to justify performing such flights. As a result, $\mathcal{N}_{it}^{SUCC} \subseteq \mathcal{N}_t^{LEG}$ is the set of nodes associated with succeeding flight legs of flight $i \subseteq \overline{\mathcal{N}_t^E}$ as of time t, suggesting that flight leg i cannot be served unless one of the flight legs $j \in \mathcal{N}_{it}^{SUCC}$ is served.

$\mathcal{A}_t$ is comprised of three subsets: the initial arcs from aircraft to flight legs ($\mathcal{A}_t^{INIT}$), the sequencing arcs between flight legs $\left(\mathcal{A}_t^{SEQ}\right)$, and the allocation arcs between requests and flight legs ($\mathcal{A}_t^{ALCT}$). The existence of arc $(k, i) \in \mathcal{A}_t^{INIT}$ between aircraft k and flight leg i suggests that flight leg i could potentially be served as the first flight on aircraft k's route starting from $\boldsymbol{L}_{kt}^{AVL}$, while the arc $(i, j) \in \mathcal{A}_t^{SEQ}$ between the nodes of flight leg i and flight leg j specifies that flight leg j could potentially be served after flight leg i. Furthermore, arc $(r, i) \in \mathcal{A}_t^{ALCT}$ between request r and flight leg i implies that request r could be served by flight leg i. Lastly, $\mathcal{A}_t^{INTND} \subseteq \mathcal{A}_t^{ALCT}$ denotes the set of (r, i) tuples where request r is the intended request of flight leg i.

Let τ_i^{MIN} and τ_i^{MAX}, defined in Equations (7.26) and (7.27), denote the earliest and latest start time of flight leg i, respectively. Additionally, τ_r^{DLN}, defined in Equation (5.2), denotes the latest time when the passenger group of request r must reach its destination. τ_{rt}^{SRVC} , defined in Euqation (5.6), represents the earliest time that the UAT operator could start serving request r, while τ_{kt}^{AVL} denotes the earliest time UAT aircraft k would be able to modify its future itinerary, and therefore, become available for service as of time t.

Furthermore, T_i^{SRVREV} is the flight service time of a revenue-generating flight leg i, i.e., the time it takes to serve the revenue-generating flight leg i from its start to its completion, T_{kit}^0

denotes the empty flight service time to reposition aircraft k from its first availability UAT pad as of time t to the starting point of flight i, and T_{ij} denotes the empty flight service time to reposition a UAT aircraft from the ending point of flight leg i to the starting point of flight leg j. Moreover, T_{ri}^{INGR} denotes the ingress duration and includes ground-based travel time and access time to the departure gate. Similarly, T_{ri}^{EGR} denotes the egress duration and includes the access time from the arrival gate to ground transportation, wait time for ground-based transportation (which is assumed zero), and ground-based travel time to the destination. Lastly, ω is the maximum acceptable delay compared to the desired trip time.

For the pricing, let α represent the revenue of providing the UAT service per mile per passenger, β be the operational cost of UAT aircraft per mile, and $\mathcal{C}$ denote the fixed cost of conducting a flight leg. Lastly, γ_1 denotes the cost of one relocation and γ_2 represents how exponentially worse off two relocations are compared to one.

8.3 Decision Variables

Following the notations in Bertsimas et al. [136], let y_{ki} represent a binary variable for $(k, i) \in \mathcal{A}_t^{INIT}$. Its value is 1 if flight leg i is the revenue-generating flight served by aircraft k immediately from its availability UAT pad as of time t. Additionally, x_{ij} for $(i, j) \in \mathcal{A}_t^{SEQ}$ is 1 when revenue-generating flight leg j is served immediately after revenue-generating flight leg i. To allocate the requests to the flight legs, z_{ri} for $(r, i) \in \mathcal{A}_t^{ALCT}$ is defined as a binary variable, which takes the value of 1 when request r is assigned to flight leg i, 0 otherwise. Furthermore, p_i for $i \in \mathcal{N}_t^{LEG}$ is a binary variable, where $p_i = 1$ implies that flight leg i will be conducted. Lastly, τ_i^{BOARD} for $i \in \mathcal{N}_t^{LEG}$ is the time revenue-generating flight i starts the boarding process.

It is worth noting that CLARPTW-SRE is solved sequentially over the planning horizon, and therefore, the decision variables, namely, y_{ki}, x_{ij}, z_{ri}, p_i, and τ_i^{BOARD}, have a temporal dimension. However, for notational simplicity, we drop the t index from the notations. The variables are depicted in Figure 8.1.

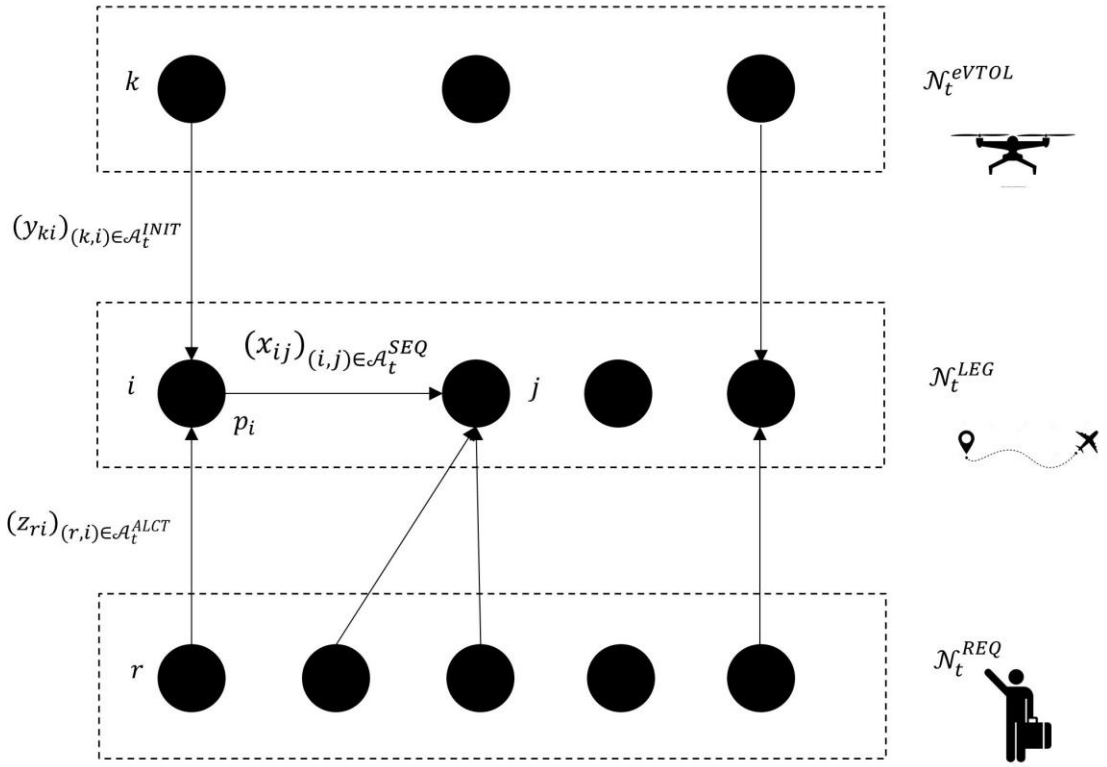

Figure 8.1 Depiction of the binary decision variables in the CLARPTW-SRE network

8.4 Objectives

In the UAT fleet operation problem, the UAT operator seeks to maximize the revenue and minimize its operating costs while providing an acceptable level of service. From the operator's perspective, the operating costs are associated with the number of flights and the mileage aircraft fly. On the other hand, relocations cause user inconvenience. In the following sections, we discuss the components of the objective function.

8.4.1 Revenue

In eVTOL operation, the ascending and descending of the aircraft are relatively costly, and therefore, trips with longer distances would have a lower cost per mile. Additionally, the UAT operation is per-seat, and therefore, serving more passengers translates to higher revenue.

As a result, we assume the revenue for serving r is proportional to the straight-line distance between the request's origin and destination (i.e., D_r^{OD}) and its group size (i.e., q_r). This assumption implies that when serving all requests is not feasible, requests that have a bigger group size or a longer distance are more profitable to be served. Let R_r denote the revenue earned by serving request r, as defined in Equation (8.1), where α denotes the revenue per passenger per mile.

$$R_r = \alpha q_r D_r^{OD} \tag{8.1}$$

8.4.2 Aerial Fixed and Variable Costs

The total aerial distance that the UAT aircraft travels consists of empty mileage and revenue-generating mileage. Additionally, there is a fixed cost associated with ascending and descending of the aircraft. In terms of the energy consumed, the cost per mile is higher when an aircraft is loaded compared to when it moves empty. As a result, one could assign different costs for empty and loaded mileage. Without loss of generality, we assume empty and loaded mileage have the same cost. Let β denote the operating cost per aerial mileage. Consequently, the cost of the revenue-generating flight leg i is $\mathcal{C} + \beta D_i^{LEG}$. Additionally, the cost of the empty flight leg from the first availability pad of aircraft k as of time t to the starting pad of flight leg i is $\mathcal{C} + \beta D_{kit}^0$, while the cost of the empty flight leg between the end of flight leg i and the start of flight leg j is $\mathcal{C} + \beta D_{ij}$.

More concisely, the costs of revenue-generating flights (i.e., locations in LARP) could be incorporated into the cost of preceding empty flight legs (i.e., arcs in LARP). Therefore:

$$\bar{C}^0_{kit} = \left(\mathbb{1}_{D^0_{kit}>0}\mathcal{C} + \beta D^0_{kit}\right) + \left(\mathcal{C} + \beta D^{LEG}_i\right) \tag{8.2}$$

$$\bar{C}_{ij} = \left(\mathbb{1}_{D_{ij}>0}\mathcal{C} + \beta D_{ij}\right) + \left(\mathcal{C} + \beta D^{LEG}_j\right) \tag{8.3}$$

Where $\mathbb{1}$ is the indicator function, implying the cost associated with the empty flight legs becomes zero when D^0_{ki} or D_{ij} is zero. $\bar{C}^0_{kit}$ defined in Equation (8.2) is the total cost of serving revenue-generating flight leg i as of time t, which includes the preceding empty flight leg from $\boldsymbol{L}^{AVL}_{kt}$ to $\boldsymbol{S}_i$. $\bar{C}_{ij}$, as defined in Equation (8.3), is the total cost of serving revenue-generating flight leg j, including the preceding empty flight leg from $\boldsymbol{E}_i$ to $\boldsymbol{S}_j$.

8.4.3 Relocation Cost

In a ubiquitous network, the origin and destination of a request coincide with the desired pick-up and drop-off UAT pads, respectively. A *relocation* (or *transfer*) occurs when a passenger is moved one on the ground to a UAT pad other than its desired. Let C^{RLC}_{ri}, defined in Equation (8.4), represent the total cost of relocating request r to take flight leg i. $\chi_{ri} \in \{0, 1, 2\}$ is the number of relocations required for request r to take flight leg i. The value of $\gamma_2 \geq 1$ determines how exponentially worse off one relocation is compared to two. For instance, Therefore, $\gamma_2 = 2$ implies that experiencing 2 relocations has a disutility 4 times worse than experiencing 1 relocation.

$$C^{RLC}_{ri} = \gamma_1 {\chi_{ri}}^{\gamma_2} \tag{8.4}$$

Let γ_1 denote the cost of one relocation and β represent the cost of aerial mileage. As a result, $\frac{\gamma_1}{\beta} = x$ suggests that 1 relocation is equivalent to x aerial miles. Therefore, a request would be relocated once if, as a consequence, the aerial mileage were reduced by more than x. On the other hand, $\frac{\gamma_1}{\beta}$ cannot be so high that it prevents the requests from relocation within Δ^{EMPTY}.

To elaborate, Figure 8.2 depicts the origin of request r with two candidate pick-up pads for a flight: aircraft k located at Δ^{EMPTY^-} of $\boldsymbol{O}_r$ and aircraft l placed at Δ^{EMPTY^+}. Since there is no repositioning legs within Δ^{EMPTY}-radius of $\boldsymbol{O}_r$, if the request is not relocated, a UAT aircraft from outside of Δ^{EMPTY}-radius of $\boldsymbol{O}_r$ should reposition to $\boldsymbol{O}_r$. To allow the relocation of the request within Δ^{EMPTY}, the cost of moving aircraft l to the origin of request r should be higher than the cost of relocating the request from its origin to aircraft k. Therefore, $\gamma_1 < \beta(1+\epsilon)\Delta^{EMPTY}$.

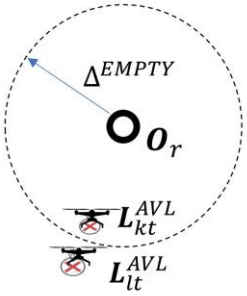

Figure 8.2 Comparison of aerial mileage and relocation cost for short repositioning elimination

Consequently, $\frac{\gamma_1}{\beta} = x < (1+\epsilon)\Delta^{EMPTY}$. Given $\epsilon = 0.1$ and $\beta = 1$, γ_1 should be smaller than $1.1\Delta^{EMPTY}$. For instance, for $\Delta^{EMPTY} = 0.5$ miles and 1 mile, $\gamma_1 < 0.55$ and 1.1, respectively, suggesting that a request would be relocated once if the aerial mileage (including the empty mileage) of the resulting itinerary could be decreased by as much as 0.55 or 1.1 miles.

8.5 CLARPTW-SRE Formulation

CLARPTW-SRE seeks to address the routing and scheduling problem and location problem simultaneously. For routing and scheduling, we employ the model with time constraints developed by Bertsimas et al. [136]. This formulation assumes that the network has no cycles and, therefore, could be applied to any dynamic vehicle routing problem with time windows that are smaller than the typical trip time. In the UAT operations, the sufficient condition to avoid cycles (i.e., $\tau_i^{MAX} - \tau_i^{MIN} < T_i^{SRVREV} + T_{ij}$) is satisfied.

The MIP formulation for CLARPTW-SRE given the state of the system at time t is presented as follows:

$$\max \sum_{r \in \mathcal{N}_t^{REQ}} R_r\Big(\sum\nolimits_{i \in \mathcal{N}_t^{LEG}:(r,i) \in \mathcal{A}_t^{ALCT}} z_{ri}\Big) - \left(\sum_{(k,i) \in \mathcal{A}_t^{INIT}} \bar{C}_{kit}^{0} y_{ki} + \sum_{(i,j) \in \mathcal{A}_t^{SEQ}} \bar{C}_{ij} x_{ij} \right) - \sum_{(r,i) \in \mathcal{A}_t^{ALCT}} C_{ri}^{RLC} z_{ri} \tag{8.5}$$

Subject to:

$$\sum\nolimits_{k \in \mathcal{N}_t^{eVTOL}:(k,i) \in \mathcal{A}_t^{INIT}} y_{ki} + \sum\nolimits_{j \in \mathcal{N}_t^{LEG}:(j,i) \in \mathcal{A}_t^{SEQ}} x_{ji} = p_i \qquad \forall i \in \mathcal{N}_t^{LEG} \tag{8.6}$$

$$\sum\nolimits_{i \in \mathcal{N}_t^{LEG}:(k,i) \in \mathcal{A}_t^{INIT}} y_{ki} \leq 1 \qquad \forall k \in \mathcal{N}_t^{eVTOL} \tag{8.7}$$

$$\sum\nolimits_{j \in \mathcal{N}_t^{LEG},(i,j) \in \mathcal{A}_t^{SEQ}} x_{ij} \leq p_i \qquad \forall i \in \mathcal{N}_t^{LEG} \tag{8.8}$$

$$\tau_i^{BOARD} \geq \tau_i^{MIN} + (\tau_{kt}^{AVL} + T_{kit}^{0} - \tau_i^{MIN})\, y_{ki} \qquad \forall (k,i) \in \mathcal{A}_t^{INIT} \tag{8.9}$$

$$\tau_j^{BOARD} - \tau_i^{BOARD} \geq \left(\tau_j^{MIN} - \tau_i^{MAX}\right) + \left(T_i^{SRVREV} + T_{ij} - \left(\tau_j^{MIN} - \tau_i^{MAX}\right)\right) x_{ij} \qquad \forall (i,j) \in \mathcal{A}_t^{SEQ} \qquad (8.10)$$

$$\tau_i^{MIN} \leq \tau_i^{BOARD} \leq \tau_i^{MAX} \qquad \forall i \in \mathcal{N}_t^{LEG} \qquad (8.11)$$

$$\sum\nolimits_{i \in \mathcal{N}_t^{LEG}:(r,i) \in \mathcal{A}_t^{ALCT}} z_{ri} \leq 1 \qquad \forall r \in \mathcal{N}_t^{UNASGN} \qquad (8.12)$$

$$\sum\nolimits_{i \in \mathcal{N}_t^{LEG}:(r,i) \in \mathcal{A}_t^{ALCT}} z_{ri} = 1 \qquad \forall r \in \mathcal{N}_t^{FLXSTRT} \cup \mathcal{N}_t^{FXDSTRT} \qquad (8.13)$$

$$z_{ri} \leq p_i \qquad \forall (r,i) \in \mathcal{A}_t^{ALCT} \qquad (8.14)$$

$$p_i \leq z_{ri} \qquad \forall (r,i) \in \mathcal{A}_t^{INTND} \qquad (8.15)$$

$$\sum\nolimits_{r \in \mathcal{N}_t^{REQ}:(r,i) \in \mathcal{A}_t^{ALCT}} q_r z_{ri} \leq Q \qquad \forall i \in \mathcal{N}_t^{LEG} \qquad (8.16)$$

$$\tau_i^{BOARD} \geq (\tau_{rt}^{SRVC} + T_{ri}^{INGR})\, z_{ri} \qquad \forall (r,i) \in \mathcal{A}_t^{ALCT} \qquad (8.17)$$

$$\tau_i^{BOARD} + T_i^{SRVREV} + T_{ri}^{EGR} \leq \tau_r^{DLN} + M(1 - z_{ri}) \qquad \forall (r,i) \in \mathcal{A}_t^{ALCT} \qquad (8.18)$$

$$p_i \leq \sum\nolimits_{j \in \mathcal{N}_{it}^{SUCC}} p_j \qquad \forall i \in \overline{\mathcal{N}_t^E} \qquad (8.19)$$

$$p_i \in \{0,1\} \qquad \forall i \in \mathcal{N}_t^{LEG} \qquad (8.20)$$

$$y_{ki} \in \{0,1\} \qquad \forall (k,i) \in \mathcal{A}_t^{INIT} \qquad (8.21)$$

$$x_{ij} \in \{0,1\} \qquad \forall (i,j) \in \mathcal{A}_t^{SEQ} \qquad (8.22)$$

$$z_{ri} \in \{0,1\} \qquad \forall (r,i) \in \mathcal{A}_t^{ALCT} \qquad (8.23)$$

$$\tau_i^{BOARD} \geq 0 \qquad \forall i \in \mathcal{N}_t^{LEG} \qquad (8.24)$$

Objective Function. The maximization function in Equation (8.5) is multi-objective and is formulated as the weighted sum of the objectives, namely, revenue, fixed and variable costs of serving the flight legs, and the disutility of the number of relocations.

Vehicle-routing. Equations (8.6)-(8.8) and (8.20)-(8.22) are the constraints that cover the routing. Equation (8.6) specifies that flight leg i is served if it is either the first flight on an aircraft route or it is served right after another flight leg. Equation (8.7) ensures that each aircraft serves at most one flight leg as the first flight leg on its route. Equation (8.8) suggests that a flight leg could be served right after flight leg i by an aircraft if flight leg i is served in the first place.

Scheduling the Flights. Equations (8.9)-(8.11) address the flight leg scheduling. Equation (8.9) ensures that if aircraft k had flight leg i as the first flight on its route starting from $\boldsymbol{L}_{kt}^{AVL}$, the boarding time of flight leg i would be at least $\tau_{kt}^{AVL} + T_{kit}^{0}$ given the state of the system at time t. Equation (8.10) enforces that if flight leg j were served immediately after flight leg i, the boarding time of flight leg j should be at least equal to the the boarding time of flight leg i plus the time for serving revenue-generating flight i plus the empty flight service time to reposition the aircraft from

the ending point of flight i to the starting point of flight j. Equation (8.11) enforces the time windows for the start of flight legs.

Request Allocation. Equations (8.12)-(8.16) cover the location-allocation part of the formulation and assign requests to flight legs. Equation (8.12) ensures that each unassigned request is assigned to a flight leg at most once, while Equation (8.13) covers the requests with flexible and fixed pick-up UAT pads. These requests must be served since they were accepted in previous decision epochs and, therefore, must be assigned to a flight leg exactly once. Equation (8.14) specifies that request r could be assigned to flight leg i only if flight leg i is served in the first place. Equation (8.15) assumes that a flight leg is conducted only if it serves its intended request. Lastly, Equation (8.16) enforces the capacity constraint of the UAT aircraft when assigning the requests to the flights.

Requests and Flights Synchronization. Equations (8.17) and (8.18) cover the synchronization between aerial and ground-based modes to serve the first and last mile of the trip. Equation (8.17) indicates that flight leg i must start its boarding process after all the requests assigned to it have reached the departure gate of flight leg i. Equation (8.18) ensures that if request r is assigned to flight leg i, the boarding time of flight leg i should be such that the passengers of request r arrives at their destination before the deadline (i.e., τ_r^{DLN} in Figure 5.1). M is a big number and should be customized for $(r, i) \in \mathcal{A}_t^{ALCT}$.

Succeeding Legs for Short Repositioning Elimination. Equation (8.19) explicitly specifies that if flight leg i ends at a UAT pad other than the desired UAT pad of its intended request (i.e., $i \in \overline{\mathcal{N}_t^E}$) to eliminate a short repositioning flight, one flight leg $j \in \mathcal{N}_{it}^{SUCC}$ starting at that UAT pad should be served to justify the relocation.

Decision Variables. Equations (8.20)-(8.23) are the binary constraints for the decision variables. Lastly, Equation (8.24) specifies that τ_i^{BOARD} is a positive real number.

8.6 Solution Method

8.6.1 Solver

The instances of CLARPTW-SRE are solved using the free academic license of Gurobi interface implemented in Python 3.7, gurobipy 9.1, and on a machine with 3.00GHz Intel® Xeon® CPU and 128 GB RAM.

Table 8.1 Parameters associated with optimization

Parameter	Symbol	Value	Unit
Re-optimization interval	Δ^{UPDATE}	1	minutes
Acceptable gap	-	5	percent
Initial time limit	-	30	seconds

Table 8.1 presents the parameters associated with the optimization. The problem is re-optimized every minute. The MIP gap is set to 5%. However, a smaller MIP gap is desirable if achieving it takes less than 30 seconds. Therefore, after an initial time limit of 30 seconds, the MIP gap is examined. If it is less than 5%, the Gurobi Optimizer stops. Otherwise, it continues until the termination criteria of the 5% MIP gap is reached.

8.6.2 Warm Start

At the beginning of decision epoch e, i.e., $t = \tau_{v_e^S}$, the itinerary and schedule of the aircraft could be used to provide a partial warm start to the MIP. To this end, let $Q_{kt}^{WREV} = \left\{q_{kt1}^{WREV}, \dots, q_{kt,n}^{WREV}, q_{kt,n+1}^{WREV} \dots, q_{kt|Q_{kt}^{WREV}|}^{WREV}\right\}$ denote the ordered list of revenue-generating flights that have not started as of $t = \tau_{v_e^S}$. Q_{kt}^{REV} is defined in Equation (8.25).

$$\mathcal{Q}_{kt}^{WREV} = \left\{ \mathscr{f}_i \in \mathcal{Q}_{kt} : H_i = 1, \varsigma_{it}^{LEG} = 1 \right\} \tag{8.25}$$

If $\mathcal{Q}_{kt}^{WREV} \neq \emptyset$ for $a_k \in \mathcal{K}$ at $t = \tau_{v_e^S}$ (i.e., there are some flights assigned to the aircraft), the values of $\hat{y}_{ki}$, $\hat{x}_{ij}$, $\hat{p}_i$ and $\hat{\tau}_i^{BOARD}$, where $\wedge$ represents the warm start values, are defined as follows:

I. Let $\hat{y}_{ki} = 1$ for $\mathscr{f}_i = q_{kt1}^{WREV}$ (i.e., the first revenue-generating on the aircraft itinerary that has not started).

II. Let $\hat{x}_{ij} = 1$ for $\mathscr{f}_i = q_{kt,n}^{WREV}$, $\mathscr{f}_j = q_{kt,n+1}^{WREV}$, and $n \in \{1, \dots, \left|\mathcal{Q}_{kt}^{WREV}\right| - 1\}$ if $\left|\mathcal{Q}_{kt}^{WREV}\right| \geq 2$.

III. Let $\hat{p}_i = 1$ and $\hat{\tau}_i^{BOARD} = \tau_{i\tau_{v_e^S}}^{STRT}$ for $\mathscr{f}_i \in \mathcal{Q}_{kt}^{WREV}$.

Additionally, for $\mathscr{r}_r \in \mathcal{R}_t^{FLXSTRT} \cup \mathcal{R}_t^{FXDSTRT}$ at $t = \tau_{v_e^S}$, the value of $\hat{z}_{ri}$ is defined as follows:

IV. Let $\hat{z}_{ri} = 1$ for $\mathscr{f}_i = \varphi_{rt}$, suggesting flight leg i is assigned to request r.

8.7 Outputs

As an acceptable solution for CLARPTW-SRE is retrieved at the end of decision epoch $e \in \mathcal{E}$, the dynamic decisions of the UAT problem are derived from the decision variables of the CLARPTW-SRE as follows:

I. $\bar{\mathcal{R}}_e^{ACCPT} = \left\{ \mathscr{r}_r : \sum_{i \in \mathcal{N}_t^{LEG} : (r,i) \in \mathcal{A}^{ALCT}} z_{ri} = 1, r \in \mathcal{N}_t^{REQ} \right\}$ for $t = \tau_{v_e^S}$ represents the set of accepted candidate requests during decision epoch e, and therefore, $\varsigma_{r\tau_{v_e^E}}^{REQ} = 1$ for $\mathscr{r}_r \in (\bar{\mathcal{R}}_e^{ACCPT} \cup \mathcal{R}_{\tau_{v_e^S}}^{UNASGN})$. Consequently, $\bar{\mathcal{R}}_e^{REJCT} = \mathcal{R}_{\tau_{v_e^S}}^{CAND} \backslash \bar{\mathcal{R}}_e^{ACCPT}$ and $\varsigma_{r\tau_{v_e^E}}^{REQ} = -1$ for $\mathscr{r}_r \in \bar{\mathcal{R}}_e^{REJCT}$.

II. $(z_{ri})_{(r,i)\in\mathcal{A}^{ALCT}_{\tau_{\nu_e^S}}} = 1$ implies $\varphi_{r\tau_{\nu_e^E}} = \not{f}_i$, suggesting flight leg i is assigned to request r as of time $\tau_{\nu_e^E}$.

III. $(\tau^{ORG}_{rt})_{r\in\mathcal{N}^{REQ}_{\tau_{\nu_e^S}}}$ for all accepted candidate requests by the end of decision epoch e (i.e., $r_r \in \bar{\mathcal{R}}^{ACCPT}_e$) is calculated as $\tau^{ORG}_{r\tau_{\nu_e^E}} = \sum_{i\in\mathcal{N}^{LEG}_{\tau_{\nu_e^S}}:(r,i)\in\mathcal{A}^{ALCT}_{\tau_{\nu_e^S}}} (\tau_i^{BOARD} - T_{ri}^{INGR}) z_{ri}$ as of time $\tau_{\nu_e^E}$.

IV. $(y_{ki})_{(k,i)\in\mathcal{A}^{INIT}_{\tau_{\nu_e^S}}}$ and $(x_{ij})_{(i,j)\in\mathcal{A}^{SEQ}_{\tau_{\nu_e^S}}}$ determine the order of revenue-generating flight legs for each aircraft, and therefore, partially defines $\left(\mathcal{Q}^{WAIT}_{k\tau_{\nu_e^E}}\right)_{k\in\mathcal{N}^{eVTOL}_{\tau_{\nu_e^S}}}$. $(y_{ki})_{(k,i)\in\mathcal{A}^{INIT}_{\tau_{\nu_e^S}}} = 1$ implies that flight leg i is the first revenue-generating flight served by aircraft k taking-off from $\boldsymbol{L}^{AVL}_{kt}$, and $(x_{ij})_{(i,j)\in\mathcal{A}^{SEQ}_{\tau_{\nu_e^S}}} = 1$ implies flight leg j is the first revenue-generating flight leg served after flight leg i. As a result, the ordered list $[i, j, \dots]$ for $k \in \mathcal{N}^{eVTOL}_{\tau_{\nu_e^S}}$ would define the revenue-generating flight legs assigned to UAT aircraft k. To fully specify $\left(\mathcal{Q}^{WAIT}_{k\tau_{\nu_e^E}}\right)_{k\in\mathcal{N}^{eVTOL}_{\tau_{\nu_e^S}}}$, the empty flight legs for repositioning the aircraft between two revenue-generating flight legs should be added. Let i and j be two consecutive revenue-generating flight legs for $i, j \in \mathcal{Q}^{WAIT}_{k\tau_{\nu_e^E}}$. Empty flight leg m would be added only if $\boldsymbol{E}_i \neq \boldsymbol{S}_j$, leading to the sequence of flights i, m, j on $\left(\mathcal{Q}^{WAIT}_{k\tau_{\nu_e^E}}\right)_{k\in\mathcal{N}^{eVTOL}_{\tau_{\nu_e^S}}}$. In this case, $\varsigma^{LEG}_{m\tau_{\nu_e^E}} = 1$ (i.e., waiting for service).

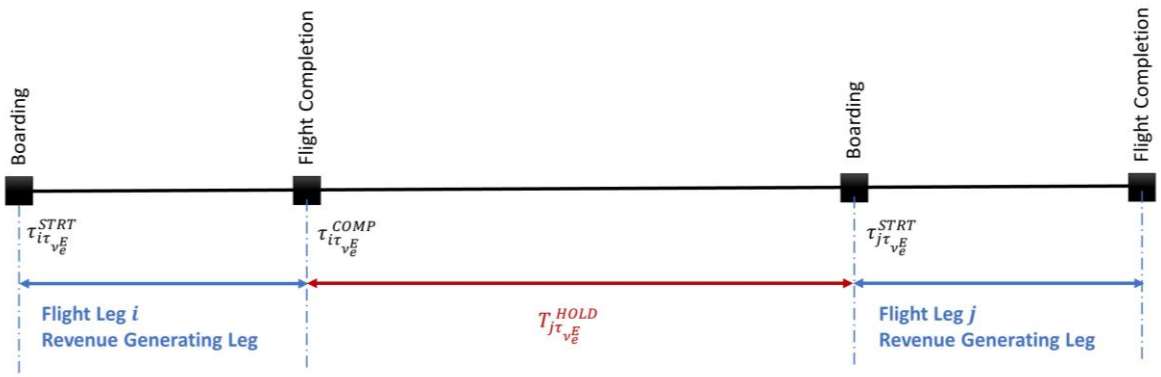

Figure 8.3 Holding time for two consecutive revenue-generating flight legs without any repositioning leg in between

It is worth noting that $(\tau_i^{BOARD})_{i\in\mathcal{N}^{LEG}_{\tau^{S}_{v^{E}_{e}}}}$ explicitly determines the start time of revenue-generating flight legs, and therefore, $\left(\tau^{STRT}_{i\tau^{E}_{v^{E}_{e}}}\right)_{i\in\mathcal{N}^{LEG}_{\tau^{S}_{v^{E}_{e}}}} = (\tau_i^{BOARD})_{i\in\mathcal{N}^{LEG}_{\tau^{S}_{v^{E}_{e}}}}$, while it only has implications for the start time of empty flight legs. Let us consider the two following cases:

I. $\not{f}_i$ and $\not{f}_j$ are two consecutive flight legs on $\left(\mathcal{Q}_{k\tau_{v^{E}_{e}}}\right)_{k\in\mathcal{N}^{eVTOL}_{\tau^{S}_{v^{E}_{e}}}}$, and are both revenue-generating, implying $\boldsymbol{E}_i = \boldsymbol{S}_j$. Therefore, $\tau^{STRT}_{i\tau^{E}_{v^{E}_{e}}} = \tau_i^{BOARD}$, $\tau^{STRT}_{j\tau^{E}_{v^{E}_{e}}} = \tau_j^{BOARD}$, and $\tau^{COMP}_{i\tau^{E}_{v^{E}_{e}}} = \tau_i^{BOARD} + T_i^{SRVREV}$. Since $\boldsymbol{E}_i = \boldsymbol{S}_j$, the aircraft will be held after completing $\not{f}_i$ and before starting $\not{f}_j$, and therefore, $T^{HOLD}_{j\tau^{E}_{v^{E}_{e}}} = \tau^{STRT}_{j\tau^{E}_{v^{E}_{e}}} - \tau^{COMP}_{i\tau^{E}_{v^{E}_{e}}}$. Figure 8.3 depicts $T^{HOLD}_{j\tau^{E}_{v^{E}_{e}}}$ for two consecutive revenue-generating flight legs without any repositioning leg in between.

II. $\not{f}_i, \not{f}_m$, and $\not{f}_j$ are three consecutive flight legs on $\left(\mathcal{Q}_{k\tau_{v^{E}_{e}}}\right)_{k\in\mathcal{K}}$, where $\not{f}_i$ and $\not{f}_j$ are revenue-generating while $\not{f}_m$ is an empty flight leg. Similar to case (I), $\tau^{STRT}_{i\tau^{E}_{v^{E}_{e}}} = \tau_i^{BOARD}$,

$\tau_{j\tau_{v_e^E}}^{STRT} = \tau_j^{BOARD}$, and $\tau_{i\tau_{v_e^E}}^{COMP} = \tau_i^{BOARD} + T_i^{SRVREV}$. The earliest start time of $\not{f}_m$would be immediately after completing $\not{f}_i$, i.e., $\tau_{i\tau_{v_e^E}}^{COMP}$, while the latest start time of $\not{f}_m$ would be $\tau_{j\tau_{v_e^E}}^{STRT} - T_m^{SRVEMP}$, which ensures $\not{f}_j$ starts on time. As a result, $\tau_{i\tau_{v_e^E}}^{STRT} + T_i^{SRVREV} \leq \tau_{m\tau_{v_e^E}}^{STRT} \leq \tau_{j\tau_{v_e^E}}^{STRT} - T_m^{SRVEMP}$. Let $T_{m\tau_{v_e^E}}^{STRTTW} = \tau_{j\tau_{v_e^E}}^{STRT} - \left(\tau_{i\tau_{v_e^E}}^{STRT} + T_i^{SRVREV} + T_m^{SRVEMP}\right)$ denote the time window available for a UAT aircraft to start the empty flight leg m after completing $\not{f}_i$ and before starting $\not{f}_j$. Figure 8.4 depicts $T_{j\tau_{v_e^E}}^{HOLD}$ for for two consecutive revenue-generating flight legs with a repositioning leg in between.

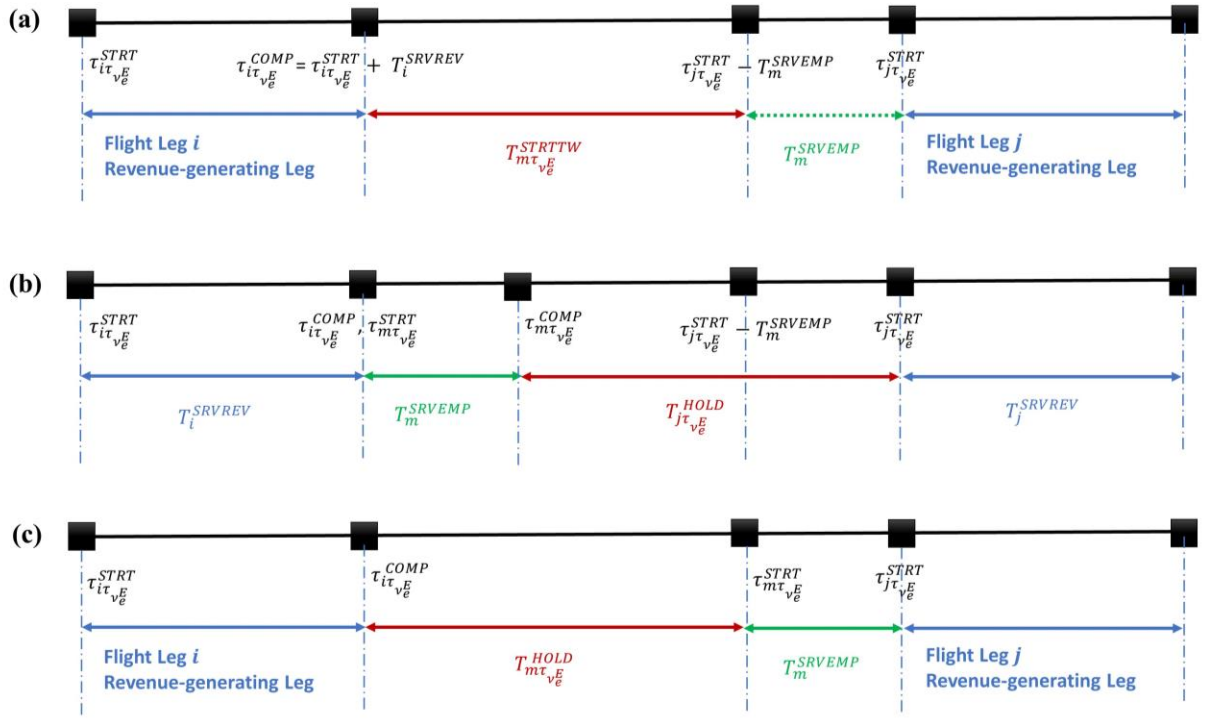

Figure 8.4 Holding time for two consecutive revenue-generating flight legs with a repositioning leg in between

If $\tau_{m\tau_{v_e^E}}^{STRT} = \tau_{i\tau_{v_e^E}}^{STRT} + T_i^{SRVREV}$, the aircraft will be immediately repositioned to, and subsequently held at the starting point of $\not{f}_j$ after completing $\not{f}_i$. On the other extreme, if

$\tau^{STRT}_{m\tau^{E}_{v^{E}_{e}}} = \tau^{STRT}_{j\tau^{E}_{v^{E}_{e}}} - T^{SRVEMP}_{m}$, the aircraft will be held at the ending point of $\not{f}_i$ as long as possible and then repositioned to the starting point of $\not{f}_j$. The former case (i.e., *reposition-first-hold-second*) is best suited for a system with stochastic travel times to minimize the probability of delays for $\not{f}_j$. However, the latter case (i.e., *hold-first-reposition-second*) is ideal for a highly reliable system where the aircraft would benefit from waiting in place until more information becomes known. In this research with deterministic travel time, we use the *hold-first-reposition-second* strategy. Therefore, $\tau^{STRT}_{m\tau^{E}_{v^{E}_{e}}} = \tau^{STRT}_{j\tau^{E}_{v^{E}_{e}}} - T^{SRVEMP}_{m}$ for $H_m = 0$ (i.e., $\not{f}_m$ is an empty flight leg) and $H_j = 1$ (i.e., $\not{f}_j$ is revenue-generating flight leg), $T^{HOLD}_{m\tau^{E}_{v^{E}_{e}}} = \tau^{STRT}_{m\tau^{E}_{v^{E}_{e}}} - \tau^{COMP}_{i\tau^{E}_{v^{E}_{e}}}$, and $T^{HOLD}_{j\tau^{E}_{v^{E}_{e}}} = 0$. However, in reality, a combination of both strategies (i.e., *hold-reposition-hold*) is more likely to be implemented.

8.8 Limitations

The limitations of CLARPTW-SRE formulation presented in this chapter are listed as follows:

- The objective function does not attempt to minimize the delays. Consequently, it is theoretically possible that the customers experience relatively high delays in a setting with lower demand.
- The formulation is vulnerable to numerical issues. Notably, under the settings that scheduling constraints are tight, numerical issues could make the problem infeasible.
- When requests are pooled together, among all the associated desired flight legs with relatively close pick-up and drop-off UAT pads, the one minimizing the objective

function should be selected in the optimal solution. Consequently, finding the optimal solution with a 0% gap is challenging when bigger problems are solved.

8.9 Concluding Remarks

This dissertation models the dynamic and stochastic problem of UAT fleet operation on a rolling horizon basis. A static and deterministic problem (i.e., snapshot problem) is solved at each decision epoch to help the UAT operator make the dynamic operational decisions, including acceptance and rejection of requests, routing and scheduling the aerial fleet, and assigning the requests to flight legs.

To achieve this goal, the snapshot problem is modeled as a Capacitated Location-Allocation-Routing Problem with Time Windows and Short Repositioning Elimination (CLARPTW-SRE), based on the node-based network presented in Chapter 7. The objective function of the MIP problem covers the revenue of the UAT operator as well as the cost associated with the empty and revenue-generating aerial mileage.

Additionally, the user inconvenience associated with serving the requests at UAT pads different from their desired ones is included as a relocation cost in the objective function. At the same time, the requests are expected to move within a short radius of their origin or destination to facilitate the elimination of short repositioning flight legs. Consequently, there is a trade-off between relocating the requests and short repositioning elimination. With a low relocation cost, the request would be relocated more often to reach a UAT aircraft nearby, while with a high relocation cost, an aircraft needs to fly from further distances to serve the request.

The CLARPTW-SRE is solved using the Gurobi interface in Python, gurobipy. To reduce the solution time, the current route and schedule of the aerial fleet at each decision epoch as a

feasible solution for the MIP problem. Ultimately, the outputs of CLARPTW-SRE are used to make dynamic operational decisions.

The CLARPTW-SRE embedded in the proposed dynamic solution framework for the UAT fleet operations is applied to a synthetic network and the Chicago network using a discrete-event simulation. Chapter 9 and Chapter 10 present the numerical results for the synthetic and Chicago network, respectively.

Chapter 9 Numerical Experiments: Synthetic Network

9.1 Overview

Even though some companies (e.g., Blade) currently offer air taxi services, on-demand and at-scale UAT operation using eVTOL technology is a conceptual system that does not exist. Simulation is a tool that is often used to model such systems when analytical models are hard to formulate. Simulations are less costly to implement, provide a better understanding of the system, and provide a tool for evaluating various strategies for the system's operation [129,137]. Therefore, simulation could be a valuable tool in a decision support system. *Discrete-event simulations (DES)* are well suited for modeling systems with complex queuing theory and resource allocation problems, whereas *agent-based models (ABS)* are helpful for modeling actions and interactions of autonomous individuals [130].

From other perspectives, simulation models could be classified as static vs. dynamic, deterministic vs. stochastic, or discrete vs. continuous [129]. *Stochastic discrete-event simulation* [129,137,138] provides a framework to model UAT fleet operation and evaluate various concepts of operations and policies. *Stochasticity* refers to the random processes (e.g., the arrival of customer's request for service) involved in the system, while *discrete-event* suggests that the measures of interest or *states* change at discrete points in time.

The discrete-event simulation framework includes a *sequencing* component, which specifies the order in which the *resources* (i.e., UAT aircraft) perform the required *tasks* (i.e., serving the customers). In the realm of simulation, the rule for determining this sequence is called *queuing discipline*. It could be as simple as a first-come-first-served (FCFS) rule, or it could be a

more complicated queuing rule derived from a mathematical programming model (see, for instance, [131,136]).

This section uses the DES to implement and evaluate the proposed dynamic solution framework for UAT fleet operations, and it further employs CLARPTW-SRE as the basis for the queuing discipline.

9.2 Experiment Design

9.2.1 Simulation Design

When operating UAT, consolidating the requests is only possible if requests are sufficiently close. Therefore, we generate the requests in clusters to better understand the impacts of the proposed consolidation scheme. Each cluster represents a town or suburb of a metropolitan area. The centroids are located on the vertices of a square with the edges of length δ. Consequently, the network has 12 OD pairs with an average Euclidean distance of 1.138δ.

Let Δt^{UPDATE} denote the interval between two decision epochs. If new requests arrive within Δt^{UPDATE}, the problem will be re-optimized to update the system's state. When request r arrives at time τ_r^{ARV} , its attributes are defined by the vector $\mathbb{A}_r^{REQ} = (\boldsymbol{O}_r, \boldsymbol{D}_r, \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_r^{DSRD}, q_r, \tau_r^{REQ})$. We assume the network is ubiquitous, and therefore, the desired pick-up and drop-off UAT pads of request r coincide with their origin and destination. The origin $\boldsymbol{O}_r$ and destination $\boldsymbol{D}_r$ of request r are randomly generated around the centroids using isotropic Gaussian distributions with the standard deviation of σ. Therefore, σ represents the spread of the demand around the centroids. The corresponding centroids of the request's origin and destination are randomly chosen from the four centroids. Δ^{OD} denotes the minimum distance between the origin

and destination of a request to qualify for a UAT trip. Consequently, the origin and destination of the requests are generated so that the distance between origin and destination exceeds Δ^{OD}.

The request arrival process is a Poisson process with the intensity of λ. Therefore, the interarrival times are exponentially distributed with the mean of $\mathcal{T}^{INT} = 1/\lambda$. τ_r^{REQ}, the requested service time for request *r*, is calculated as $\tau_r^{ARV} + T_r^{ADV}$, where T_r^{ADV} is randomly drawn from a uniform distribution with the mean of $\mathcal{T}^{ADV}/2$ and the range of $[0, \mathcal{T}^{ADV}]$. Lastly, we assume that each request has one passenger, and therefore, $q_r = 1$.

The static attributes of $a_k \in \mathcal{K}$ are represented by $\mathbb{A}_k^{eVTOL} = (Q_k, v_k^{AIR})$. We assume that the fleet of UAT aircraft is homogenous, and therefore, their capacity and speed are denoted by Q and v^{AIR}, respectively. Let Δ^{EMPTY} denote the minimum Euclidean distance to justify an empty repositioning flight leg.

Furthermore, the state of UAT aircraft at the beginning of the planning horizon (i.e., $t = 0$) is presented by $S_0^{eVTOL} = (\varsigma_{k0}^{eVTOL}, \tau_{k0}^{AVL}, \boldsymbol{L}_{k0}^{AVL}, \mathcal{Q}_{k0}, \mathfrak{E}_{k0}^{NDSRD})_{a_k \in \mathcal{K}}$. The initial location of $a_k \in \mathcal{K}$ at time $t = 0$ (i.e., $\boldsymbol{L}_{k0}^{AVL}$) is randomly generated around the centroids using isotropic Gaussian distributions with the standard deviation of σ. Furthermore, all the aircraft are idle and available at the beginning of the planning horizon, and there is no incomplete flight leg on their itinerary. In other words, $\tau_{k0}^{AVL} = 0$, $\varsigma_{k0}^{eVTOL} = 0$, and $\mathcal{Q}_{k0} = \emptyset$ for $a_k \in \mathcal{K}$. Lastly, $\mathfrak{E}_{k0}^{NDSRD} = 0$ for $a_k \in \mathcal{K}$.

9.2.2 Evaluation

In a stochastic process, a *sample path* is defined as the sequence of *sample realizations* of random variables. In other words, a sample path is a sequence of outcomes over time [132]. If Φ represents the set of all possible sample paths, let $\phi \in \Phi$ denotes a sample path.

Each sample path generates an independent *replication* of the system that is being simulated. Therefore, the outputs (i.e., performance measures) of simulations could be classified either as *within-replication* or *across-replication* data. The within-replication data is summarized to produce the across-replication outputs. Subsequently, the overall statistics are specified by using the across-replication outputs.

Across-replication outputs are i.i.d. They are independent because different random seeds are used for each replication, and they are identically distributed because the same model is run in each replication to produce the outputs. They also tend to follow a normal distribution. None of these might be true for within-replication data.

The input parameters of the simulation are called *(controllable) parameter setting*, or *(proposed) configuration* or *design* [139]. Let θ denote the input parameters (i.e., configuration) of the simulation. These parameters could be quantitative (e.g., number of aircraft to be used for serving the users' requests) or qualitative (the queuing discipline for serving the users). Let $L^{\theta}(\phi)$ denote the performance measure of the sample path ϕ as an across-replication output under configuration θ. Therefore, $J^{\theta} = \mathbb{E}\left[L^{\theta}(\phi)\right]$ shows the overall statistics of the simulation under configuration θ. Since the expectation cannot be computed, we choose a sample $\widehat{\Phi} \subseteq \Phi$ and then calculate $\hat{\jmath}^{\theta} = \sum_{\phi \in \widehat{\Phi}} \frac{L^{\theta}(\phi)}{|\widehat{\Phi}|}$.

The simulation for the synthetic network is designed as a steady-state simulation to eliminate the influence of the initial conditions. Each experiment in the base-case scenario and sensitivity analyses is replicated 30 and 20 times, respectively, where the stochasticity stems from the customer requests and the initial locations of the UAT aircraft.

9.2.3 Parameter Setting

In this section, we present the parameters used in the experiments. Table 9.1 summarizes the network design parameters. Let $\delta = 30$ miles, where δ is the length of the edges of the square. With the ground speed in the range of 20-30 mph, traversing one edge would take between 60-90 minutes on the ground. We set $\Delta^{OD} = 10$ miles. Finally, the standard deviation of the isotropic Gaussian distribution is set to 2 miles.

Table 9.1 Network parameters

Parameter	Symbol	Value	Unit
Number of centroids	-	4	-
Distance between centroids	δ	30	miles
Standard deviation of the Gaussian distributions	σ	2	miles
OD minimum distance	Δ^{OD}	10	miles

Table 9.2 summarizes the parameters associated with the ingress and egress of the passengers. Since the average walking speed of adults is 3-4 mph [140], the walking speed, v^{WALK}, is set to 3 mph. The passengers will walk for a maximum of 5 minutes to reach a UAT pad. As a result, the walking threshhold, Δ^{WALK}, is set to 0.25 miles.

Table 9.2 Parameters associated with the ingress and egress of the passengers

Parameter	Symbol	Value	Unit
Euclidean driving speed in the downtown area	v^{DRIVE}	20	mph
Maximum access distance	Δ^{ACCESS}	3	miles
Maximum walking distance	Δ^{WALK}	0.25	miles
Walking speed	v^{WALK}	3	mph
Elapsed time from ground transportation area to the departure gate of flight i for request r	T_{ri}^{DGATE}	3	minutes
Elapsed time from the arrival gate of flight i to the ground transportation area for request r	T_{ri}^{AGATE}	2	minutes

The speed of the vehicles on the ground over Euclidean distances (v^{DRIVE}) is set to 20 mph. We further assume that the maximum acceptable delay for UAT service, i.e., ω, in the synthetic network is 15 minutes. $v^{DRIVE} = 20$ mph and $\omega = 15$ minutes suggest that the maximum accessible distance on the ground is nearly 5 miles or around 2.5 miles from either the origin or destination in a symmetric case. Consequently, we set $\Delta^{ACCESS} = 3$ miles, implying that the access distance from either origin or destination is about 10% of the OD distance.

The time to reach the departure gate, including the security screening, (i.e., T_{ri}^{DGATE}) and the time to reach the ground transportation area after landing (i.e., T_{ri}^{AGATE}) are assumed to be identical for all passengers and independent of the UAT port or pad design. The values of T_{ri}^{DGATE} and T_{ri}^{AGATE} for $(r, i) \in \mathcal{A}_t^{ALCT}$ are set to 3 and 2 minutes, respectively.

Table 9.3 Parameters associated with flight operation

Parameter	Symbol	Value	Unit
Number of UAT aircraft	K	60	-
Minimum of Euclidean distance for repositioning flights	Δ^{EMPTY}	1	miles
Aerial speed	v^{AIR}	150	mph
Boarding duration	T^{BOARD}	3	minutes
Deboarding duration	$T^{DEBOARD}$	2	minutes
Departure clearance	$T^{TAKEOFF}$	0.5	minutes
Landing clearance	$T^{LANDING}$	0.5	minutes
Hover ascend	T^{ASCEND}	0.75	minutes
Hover descend	$T^{DESCEND}$	0.75	minutes
Detour factor	ϵ	0.1	-

The assumptions regarding the flight operation are presented in Table 9.3. The aerial fleet size is fixed with 60 UAT aircraft. Multiple original equipment manufacturers (OEMs) (including Joby Aviation [18] and Kitty Hawk [17]) have presented UAT aircraft designs with a cruise speed of at least 180 mph, as discussed in Table 2.2. Thus, we choose the aerial speed of 150 mph. The

detour factor of the aerial trip is assumed 0.1, which is in the range of 0.5-0.15 suggested in other UAM studies [8,141].

The minimum Euclidean distance to justify a UAT repositioning flight, Δ^{EMPTY}, is set to 1 mile. With the average driving speed of 20 mph, traversing within 1 mile of both origin and destination would take at most 6 minutes, which is well below the maximum acceptable delay of 15 minutes. This suggests that even though the aircraft within Δ^{EMPTY} cannot be repositioned to serve a request, the passengers have enough time to relocate on the ground to reach the aircraft.

The UAM market study commissioned by NASA [8] estimates the boarding duration and deboarding duration to be in the range of 3-5 and 2-3 minutes, respectively, while Porsche Consulting [6] estimates 3 minutes for boarding or deboarding. Accordingly, we set the boarding and deboarding duration to 3 and 2 minutes, respectively. Additionally, we assume 30 seconds before departure and after landing for clearance. We further assume it takes a UAT aircraft 45 seconds to ascend and 45 seconds to descend vertically.

Consequently, the turnaround time, including boarding, deboarding, and take-off and landing clearance, is minutes, which is 6 minutes, which is consistent with the turnaround time estimated by Joby Aviation [18] (see Table 2.6). The overhead time of serving a flight leg, either empty or revenue-generating, includes hover ascend and descend and ATC clearance for the take-off and after the landing, which amounts to 2.5 minutes. If the flight leg serves passengers, an additional 5 minutes will be added to the flight service time. As a result, the overhead time of empty and revenue-generating flight legs are 2.5 and 7.5 minutes, respectively.

Assuming the ground speed of 30 mph over long distances, an average trip would take 68.3 minutes with ground-based transportation. The parameters in Table 9.1, Table 9.2, and Table 9.3

translate to the mean aerial distance of 37.56 miles (i.e., 1.1× 34.14). Consequently, the average time for serving a revenue-generating flight leg is 22.5 minutes (i.e. $\frac{37.56}{150} \times 60 + 7.5$). As a result, the maximum service rate is 2.66 revenue-generating flights per hour per aircraft. This extreme value corresponds to cases where the request's arrival rate is so high, or the requests are so close to each other that the empty distance from the destination of one flight to the origin of the next flight origin is close to zero in expectation.

In queuing theory, the *traffic intensity* or *utilization* ρ is defined as $\frac{\lambda}{c\mu}$, where λ is the average arrival rate, μ is the average service rate, and c is the number of servers. Furthermore, let $r = \frac{\lambda}{\mu}$ represent the *offered load*. Utilization shows the fraction of time a server is busy, while offered load represents the average number of busy servers [142]. Without job rejection and flight-sharing (also called *partial batch* service [142]), ρ should be below 1 for the queue to not grow indefinitely, and therefore, for the system to be stable. With $c = K = 60$ and $\mu = 2.66$ flight per hour per aircraft, the system can accommodate $\lambda = 128$ requests for flight per hour (i.e., $T^{INT} =$ 28 seconds) to achieve 80% utilization ($\rho = 0.8$).

Furthermore, the average of 22.5 minutes for serving a revenue-generating flight would translate to the average trip time of 27.5 minutes (i.e., 22.5 + 3+ 2) for each passenger if there were no wait time for the aerial service, and the requests were served without any ground-based transportation. Considering a maximum delay of $\omega = 15$ minutes (in Figure 5.1), the minimum and maximum mean trip times for passengers are 27.5 and 42.5 minutes, respectively. These numbers correspond to travel time savings of 25.8 minutes (37.8%) to 40.8 minutes (59.7%) compared to driving on the ground, which almost meets the 40% travel time savings rule-based demand model suggested by Uber Elevate [23].

The CLARPTW-SRE is formulated as a multi-objective problem, and therefore, the objective function is a weighted sum of different objectives. Table 9.4 shows the parameters in the objective function. Since the monetary value of one objective to another is not well-defined, we treat these parameters as the weight of one objective relative to another. The ratio of revenue per passenger mile to cost per mile (i.e., α/β) of 2 implies that, roughly speaking, serving a request without pooling is profitable as long as the empty repositioning mileage is shorter than the OD distance of the request. Furthermore, in the base-case experiment, the weight of the number of relocations (γ_1) is set to zero, suggesting the disutility of the relocation from desired pick-up and drop-off UAT pads is not explicitly considered; however, a sensitivity analysis to the value of γ_1 is conducted. The reason is discussed in depth in Section 8.4.3.

Table 9.4 Parameters in the objective function

Parameter	Symbol	Value	Unit
Weight of revenue per passenger-mile of OD distance	α	2	-
Weight of cost per mile of flight	β	1	-
Weight of disutility of relocations from desired pads	γ_1	0	-
Exponential disutility of 1 relocation compared to 2	γ_2	2	-
Fixed cost of conducting a flight leg	$\mathcal{C}$	0	-

9.2.4 Planning Horizon

Steady-state simulations run over a long time, which is specified by the analyst, to eliminate the influence of the initial conditions. The goal of a steady-state simulation is to study the long-run or steady-state performance of the system. However, the statistics are biased during the warm-up period since the system starts empty and idle. Therefore, to achieve statistically meaningful performance measures, we need to determine the simulation runtime.

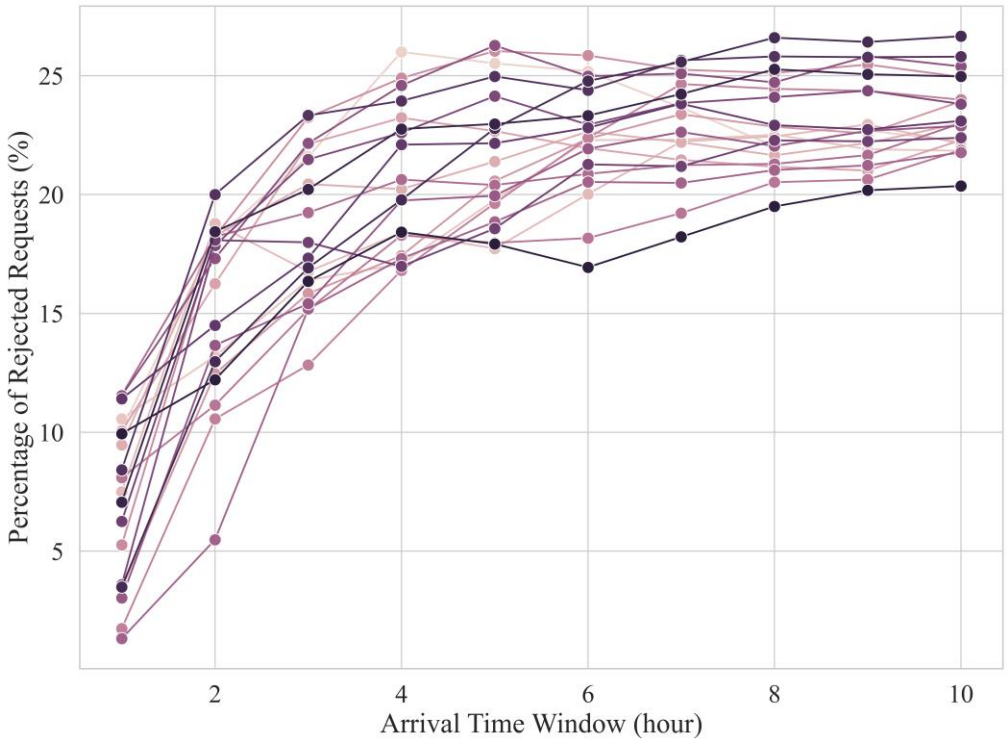

Figure 9.1 Percentage of rejected requests over arrival time windows of 1 to 10 hours for 20 replications and $Q = 1$

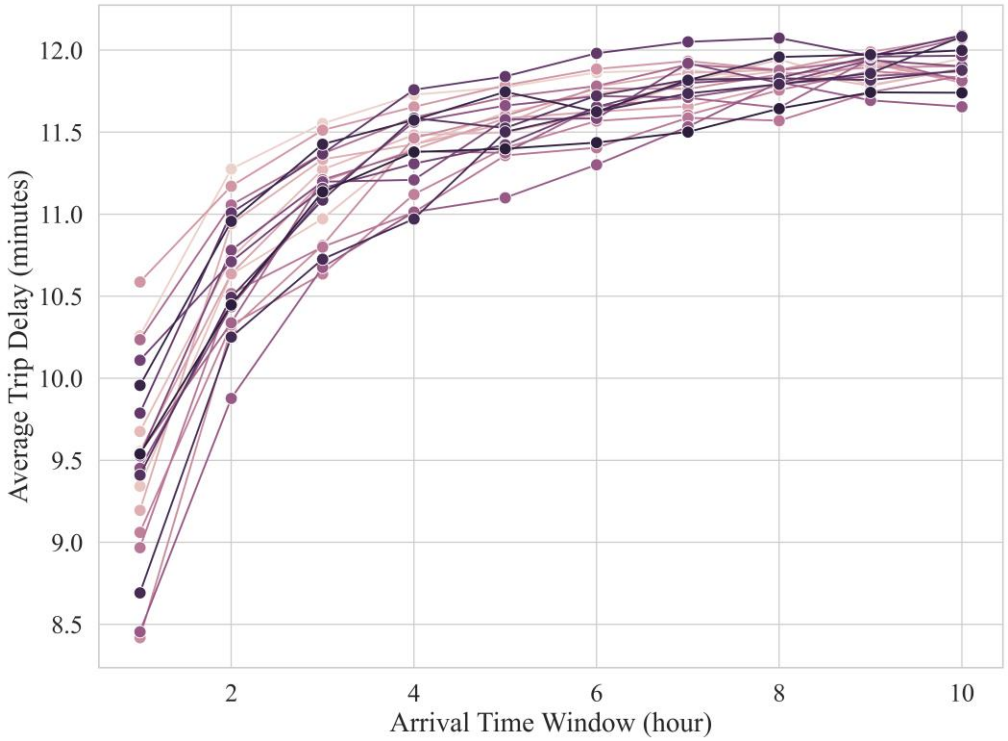

Figure 9.2 Average trip delay per request over arrival time windows of 1 to 10 hours for 20 replications and $Q = 1$

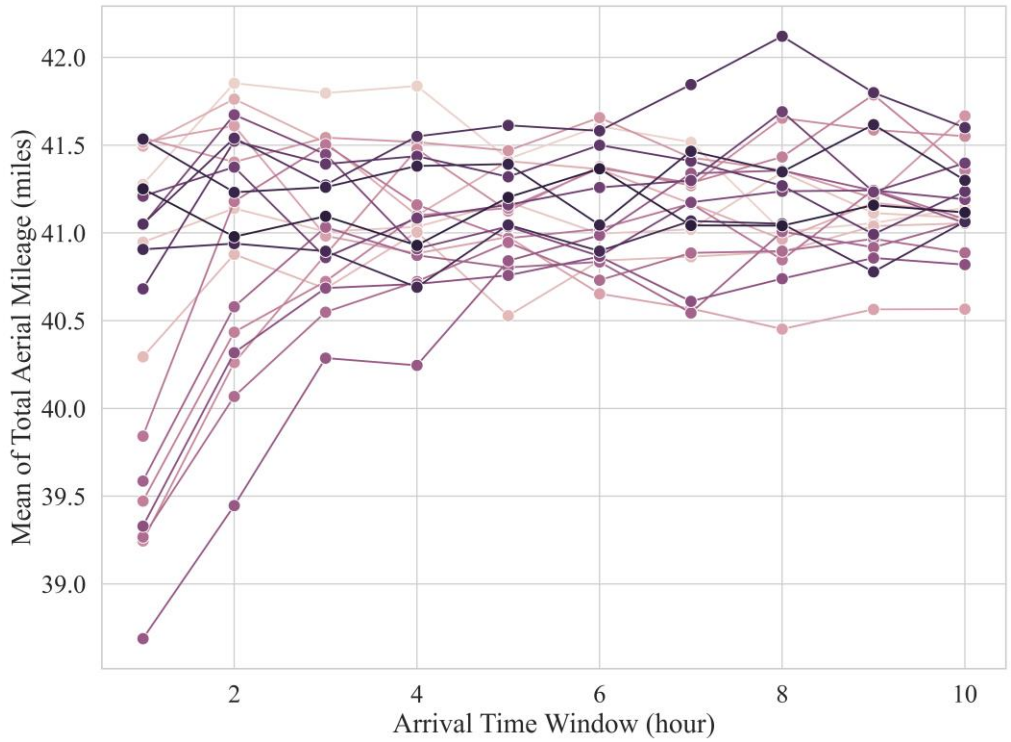

Figure 9.3 Average aerial mileage over arrival time windows of 1 to 10 hours for 20 replications and $Q = 1$

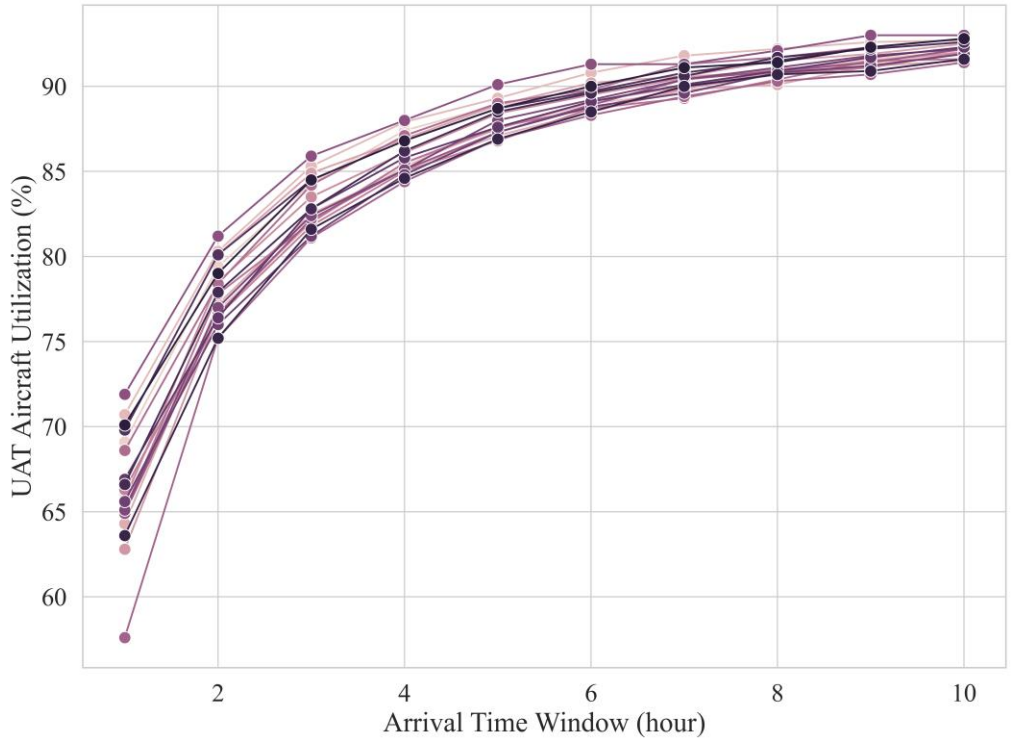

Figure 9.4 UAT aircraft utilization over arrival time windows of 1 to 10 hours for 20 replications and $Q = 1$

Using UAT aircraft with capacity 1, we run 30 replications of the problem over 10 hours, with increments of one hour. Figure 9.1 through Figure 9.4, respectively, illustrate the mean of the percentage of rejected requests, trip delay per served request, total aerial mileage (including empty and revenue-generating mileage) per served request, and UAT aircraft utilization. After reviewing these figures, we use an arrival time window of 8 hours (i.e., 480 minutes) to model a steady-state system in the experiments. In other words, $\mathfrak{T}^{STRTARV} = 0$ and $\mathfrak{T}^{ENDARV} = 480$, where $\mathfrak{T}^{STRTARV}$ and $\mathfrak{T}^{ENDARV}$ denote the start and end of request arrival period.

9.3 Numerical Results: Base-case Experiment

Table 9.5 summarizes the exogenous and design parameters for the base-case experiment. It is worth noting that ν^{DRIVE} is the driving speed along the Euclidean distances, and therefore, the actual speed of the vehicles on the ground is greater than or equal to 20 mph.

Table 9.5 Exogenous and design parameters associated with the base-case experiment

$\mathcal{T}^{INT}$ **(sec)**	σ **(mi)**	ν^{DRIVE} **(mph)**	ϵ	ν^{AIR} **(mph)**	ω **(min)**	$\mathcal{T}^{ADV}$ **(min)**	α/β	$(T^{BOARD}, T^{DEBOARD})$ **(min)**	$(T_{ri}^{DGATE}, T_{ri}^{AGATE})$ **(min)**
Exogenous Parameters				*Design Parameters*					
20	2	20	0.1	150	15	30	2	(3, 2)	(3, 2)

9.3.1 Problem Size and Solution Time

For each replication, we calculate the mean and standard deviation of the number of flight legs, requests, and arcs at each decision epoch. Subsequently, we estimate the mean of the mean and the standard deviation for all replications. Table 9.6 summarizes the estimated mean of the mean and standard deviation of the number of flight legs, requests, and arcs at each decision epoch.

Table 9.6 Estimated mean of the average number of flight legs, requests, and arcs at each decision epoch

	Flight Legs			Requests				Arcs		
Q	$\widetilde{\mathcal{N}}_t^{\mathrm{LEG}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{DSRD}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{CNCT}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{REQ}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{UNASGN}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{FLXSTRT}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{FXDSTRT}}$	$\widetilde{\mathcal{A}}_t^{INIT}$	$\widetilde{\mathcal{A}}_t^{SEQ}$	$\widetilde{\mathcal{A}}_t^{ALCT}$
1	237	70	167	73	3	62	8	1977	1175	381
2	234	69	165	78	3	62	13	2161	1059	386
3	230	68	162	78	3	61	14	2188	1018	379
4	229	68	161	78	3	61	14	2192	1008	377

Table 9.7 shows the average simulation time over 30 replications and worst MIP gap, worst MIP solution time, and worst decision time over all decision epochs for 30 replications. The decision time, i.e., T_e^{EPOCH}, includes the network transformation and reduction in addition to the MIP solution time. Therefore, the decision time accounts for the total time required by the operator to make a decision and is greater than the MIP solution time.

Table 9.7 Average simulation time for 30 replications and worst MIP gap, worst MIP solution time, and worst decision time over all decision epochs for 30 replications

Aircraft Capacity (Q)	Average Simulation Time (minutes)	Worst MIP Solution Time (seconds)	Worst Decision Time (seconds)	Worst MIP Gap (%)
1	11.4	0.8	3.4	0
2	12.0	3.3	5.5	0
3	11.6	1.2	4.1	0
4	11.6	1.2	4.8	0

As shown in Table 9.7, using the proposed solution framework, simulating the base-case scenario over the planning horizon of 8 hours would take, on average, nearly 12 minutes. The worst MIP gap of all decision epochs over all the replications is 0, implying that all the MIPs are solved to optimality. The worst MIP solution time for $Q = 1$ is 0.8 seconds, while the worst

decision time is 3.4 seconds. The worst-case decision time over all capacities is around 5.5 seconds, which is well below the 1-minute re-optimization interval, i.e., Δ^{UPDATE}.

9.3.2 User Experience

Table 9.8 through Table 9.12 present the performance measures associated with the requests and the provided level of service. Table 9.8 summarizes the estimated mean of performance measures associated with the UAT aircraft average load factor over four capacity levels. The average number of passengers per flight increases by 0.3 for $Q = 2$ compared to $Q = 1$. However, the average number of passengers per flight does not change significantly for $Q = 3$ and $Q = 4$, implying that for the designed experiment, the additional capacity is not conducive to pooling more requests in one flight. The percentage of flights with 3 or 4 passengers is relatively low. Additionally, in all cases with $Q > 1$, the majority of flights (i.e., nearly 70%) are conducted without air pooling.

Table 9.8 Estimated mean of performance measures related to UAT aircraft load factor over for 30 replications over $Q = 1, 2, 3,$ and 4

Q	Average Load Factor (%)	Average Number of Requests per Flight	Percentage of Flights with 1 Request	Percentage of Flights with 2 Requests	Percentage of Flights with 3 Requests	Percentage of Flights with 4 Requests
1	100	1	100	0	0	0
2	65.7	1.3	68.7	31.3	0	0
3	45.6	1.4	69.5	24.4	6.1	0
4	34.3	1.4	69.6	24.3	5.3	0.8

Table 9.9 summarizes the performance measures associated with served and rejected requests. Increasing the capacity from $Q = 1$ to $Q = 2$ leads to a reduction in the rate of rejection by almost 17%, from 23.0% to 5.8%, while the further increase in the capacity to $Q = 4$ results in

a slight reduction in the percentage of rejected requests, implying that the benefits of air pooling beyond $Q = 2$ is not significant for the given experiment.

Table 9.9 Estimated mean of performance measures associated with served and rejected requests

Q	Number of Served Requests	Percentage of Rejected Requests (%)	Average OD Distance of Served Requests (miles)	Average OD Distance of Rejected Requests (miles)
1	1111	23.0	34.5	33.5
2	1359	5.8	34.4	32.4
3	1384	4.1	34.4	31.8
4	1387	3.9	34.4	31.6

As shown in Table 9.9, the OD distance of served requests is higher than the OD distance of rejected requests. The difference stems from the structure of the objective function, where the earnings are based on the distance between the origin and destination of the requests, i.e., D_r^{OD}. When serving all the requests is not possible, requests with higher D_r^{OD} are more profitable to serve.

Table 9.10 summarizes the estimated mean of average trip delay, percentage of trip delay, and total trip time per request over 30 replications. An average trip would take approximately 38-40 minutes, and its value is not very sensitive to capacity increase.

Table 9.10 The estimated mean average trip delay, average of trip delay to total trip time in percentage, and average trip time over four levels of aircraft capacity for 30 replications

Q	Average Trip Delay (minutes)	Average of Percentage of Trip Delay to Total Trip Time (%)	Average Trip Time (minutes)
1	11.8	29.2	39.5
2	10.9	27.4	38.6
3	10.8	27.0	38.4
4	10.7	26.9	38.3

The average trip delay decreases for $Q \geq 2$ compared to $Q = 1$, since the system becomes less busy. Nonetheless, the higher capacity implies more potential delays resulting from demand consolidation and longer ground transportation. Moreover, the trip delays are constrained but not minimized in the objective function. That being said, air pooling has reduced the average delay per request by around 1 minute while reducing the rejection rate by at least 17%. The average percentage of trip delay to total travel time is in the range of 26.9 to 29.2%.

Given the similar performance of UAT aircraft with capacities 2, 3, and 4 in the designed experiment, we limit the rest of our analysis to $Q = 1$ and 2.

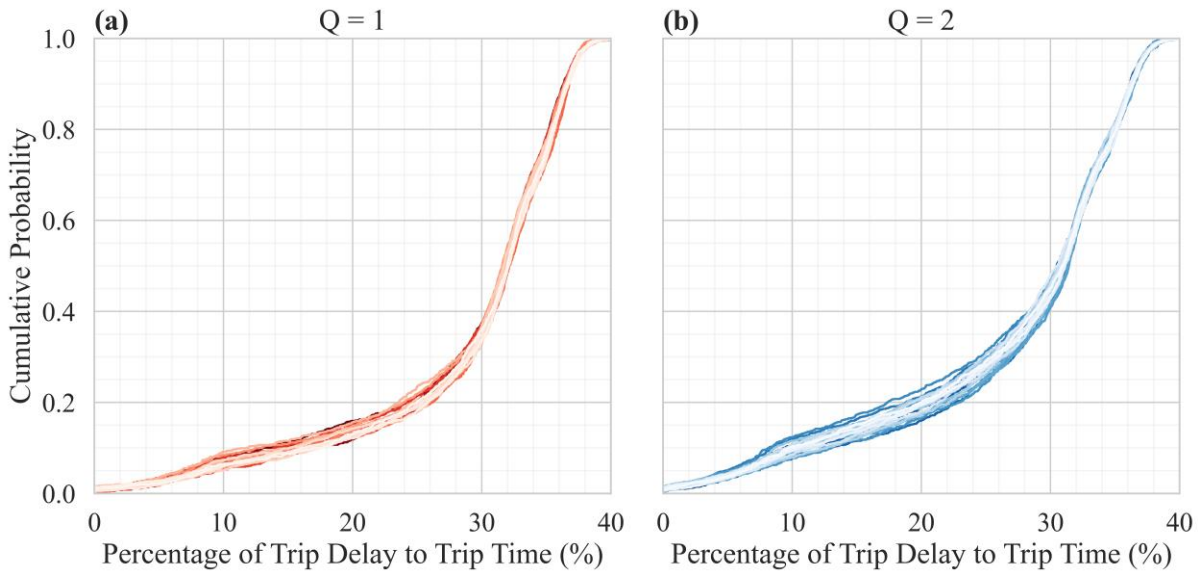

Figure 9.5 Empirical Cumulative Distribution Function (eCDF) of the percentage of trip delay to trip time over 30 replications for Q of (a) 1 and (b) 2

Figure 9.5 illustrates the empirical Cumulative Distribution Function (eCDF) of the percentage of trip delay to total trip time for $Q = 1$ and 2. The eCDF of each replication is shown as one graph. The plots show two slopes, a milder one up to the value of 25% on the x-axis, followed by a sharper slope for the range of 30-40%, suggesting that the majority of the requests (i.e., 60%) experience a percentage of trip delay between 30 and 40% of their total trip time.

Additionally, Figure 9.5 depicts how eCDFs slightly shift from $Q = 1$ to $Q = 2$. For instance, while 15% of the requests have a trip delay percentage smaller than 20% with $Q = 1$, 20% of the request have similar delays with $Q = 2$.

Table 9.11 compares the performance measures associated with the user experience for CLARPTW-SRE and CLARPTW, the formulations with and without short repositioning elimination, respectively. With the elimination of short repositioning flight legs, the percentage of rejected requests decreases around 2%, with a 0.4-minute decrease in the average trip delay.

Table 9.11 Comparison of performance measures associated with the user experience for CLARPTW-SRE and CLARPTW

Q	Rejected Requests (%)	Average Trip Delay (minutes)	Average Percentage of Trip Delay to Total Trip Time (%)	Average Trip Time (minutes)
CLARPTW-SRE				
1	23.0	11.8	29.2	39.5
2	5.8	10.9	27.4	38.6
CLARPTW				
1	25.4	12.3	30.1	39.9
2	8.0	11.3	28.0	38.9

Table 9.12 summarizes the performance measures related to ground-based legs and the relocations. It is worth mentioning that passengers are relocated within Δ^{ACCESS} of their origin or destinaton for air pooling while they are relocated within Δ^{EMPTY} of desired UAT pads to eliminate short repositioning legs. Therefore, with $Q = 1$, the relocations occur only to eliminate repositioning legs conducted within the 1-mile radius (i.e., $\Delta^{EMPTY} = 1$ mile) of the request's desired pick-up and drop-off UAT pads (which coincide with its origin and destination, respectively), while with $Q = 2$, relocations could stem from demand consolidations within the 3-mile radius (i.e., $\Delta^{ACCESS} = 3$ miles) of the request's origin or destination. Consequently, more

relocations and longer ground transportation are expected as the capacity increases from $Q = 1$ to $Q = 2$.

Table 9.12 Estimated mean of performance measures related to relocations and ground-based legs of the request trip

Q	Average Number of Relocations	Percentage of Requests with 0 Relocation (%)	Percentage of Requests with 1 Relocation (%)	Percentage of Requests with 2 Relocations (%)	Average Ground-based Distance (mi)	Average Ground-based Travel Time (min)
1	0.33	69.4	28.1	2.4	0.22	0.72
2	0.76	49.9	24.4	25.7	1.04	3.17

Table 9.12 verifies that the average of ground-based legs increases from 0.22 miles to 1.04 miles per request as Q increases. Corresppindingly, the average ground-based travel time increases 0.72 minutes to 3.17 minutes per request. Furthermore, the average number of relocations per served request increases as the capacity increases. For $Q = 1$, 70% of the requests have no relocation, while this number reduces to nearly 50% for $Q = 2$. Additionally, the values of ground-based distance and time in Table 9.12 suggest that the speed of ground-based transportation is 18.3 and 19.7 mph for $Q = 1$ and 2, respectively. These values are between the speed of the two modes of ground transportation: driving with $v^{DRIVE} = 20$ mph and walking with $v^{WALK} = 3$ mph.

Figure 9.6 provides more detailed information on the distribution of ground-based travel distance for $Q = 1$ and $Q = 2$. The two slopes on the plots are due to the distance criteria of 1 mile and 3 miles for short repositioning flight legs and demand consolidation, respectively. Since only 2.45% of the requests experience 2 relocations with $Q = 1$, nearly all the requests have an access distance smaller than 1 mile. With $Q = 2$, around 20% of the requests have a ground-based trip distance greater than 3 miles.

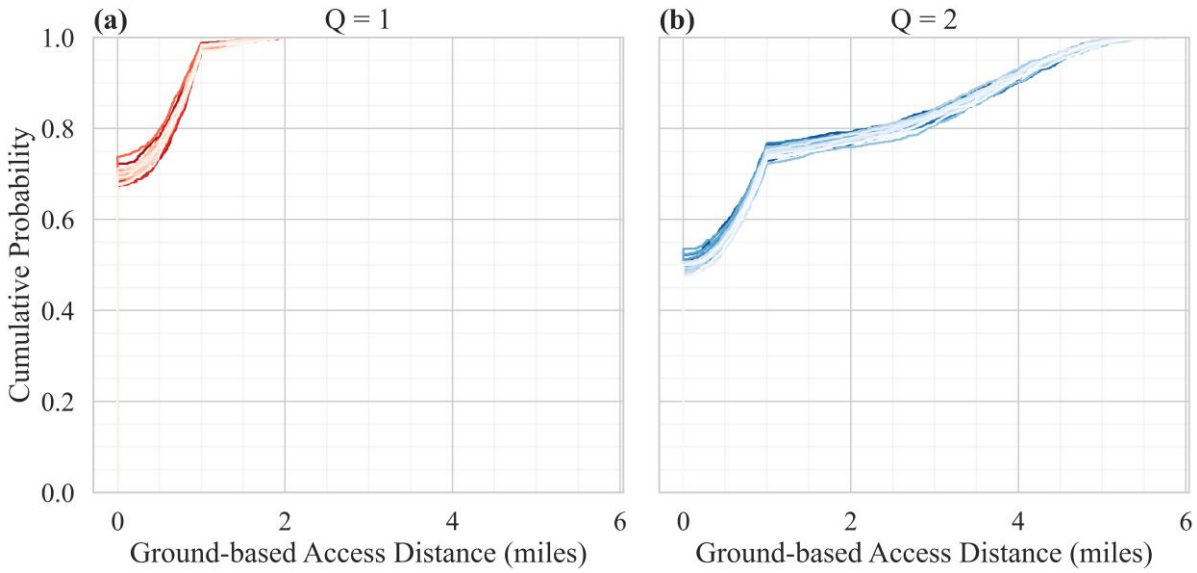

Figure 9.6 Empirical Cumulative Distribution Function (eCDF) of ground-based travel distance over 30 replications for Q of (a) 1 and (b) 2

9.3.3 UAT Operator Costs

This section reviews the performance measures associated with the operator costs, namely, the empty and revenue-generating aerial mileage. Table 9.13 summarizes the performance measures associated with revenue flights, revenue mileage, and total mileage (i.e., the summation of revenue mileage and empty mileage). The mean of revenue mileage for $Q = 1$ is 37.9 miles, which is slightly over the theoretical 37.6 miles since requests with longer OD distances are more likely to be served. The mean of revenue mileage approaches the theoretical mean as the capacity increases. One reason is that the number of rejected requests decreases with increased capacity. Therefore, serving requests with higher D_r^{OD} is not prioritized. Second, when two requests are pooled together, the shorter revenue flight leg is more likely to be selected. Similarly, the mean of total aerial mileage per revenue flight decreases as Q increases. The mean of total aerial mileage per served request sharply declines from 41.3 miles for $Q = 1$ to 31.2 miles for $Q = 2$. The ratio

of revenue flights to served requests, which is the inverse of the average passenger load of the UAT aircraft, is 76% for $Q = 2$.

Table 9.13 The performance measures associated with revenue flights, revenue mileage, and total mileage for $Q = 1$ and 2

Q	Mean of Revenue Mileage (mi)	Mean of Total Aerial Mileage per Revenue Flight (mi)	Mean of Total Aerial Mileage per Served Request (mi)	Ratio of Revenue Flights to Served Requests (%)
1	37.9	41.3	41.3	100
2	37.6	41.0	31.3	76.2

Table 9.14 Estimated mean of performance measures associated with empty repositioning flight legs for CLARPTW-SRE and CLARPTW

Q	Mean of Empty Flight Mileage (mi)	Minimum of Empty Flight Mileage (mi)	Maximum of Empty Flight Mileage (mi)	Percentage of Empty Flight Legs within Δ^{EMPTY} (%)	Percentage of Empty Flights to Revenue Flights (%)	Percenatge of Empty to Revenue Mileage (%)
CLARPTW-SRE						
1	4.86	1.1	44.18	0	69.98	8.97
2	5.16	1.1	44.47	0	67.16	9.23
CLARPTW						
1	3.68	0.06	42.11	18.7	100.35	9.74
2	3.83	0.05	45.82	21.2	100.54	10.24

Table 9.14 highlights the importance of the short repositioning elimination (SRE) scheme in a ubiquitous network, and it presents the performance measures associated with empty repositioning flight legs for CLARPTW-SRE and CLARPTW, models with and without short repositioning elimination, respectively. Without SRE, 18.7 and 21.2% of the flight legs, for $Q = 1$ and 2, respectively, are within distances shorter than $\Delta^{EMPTY} = 1$ mile. For CLARPTW-SRE, the minimum of empty flight mileage is 1.1 ($= 1.1 \times 1 = (1 + \epsilon) \times \Delta^{EMPTY}$) miles, which is consistent with the detour factor of 0.1 for the aerial legs and $\Delta^{EMPTY} = 1$ mile. However, for

CLARPTW, the empty flight distances could get as low as 0.05 miles. The maximum of empty flight mileage corresponds to the legs conducted by traversing from one vertice of the square to another diagonally.

Table 9.14 shows that the length of empty flights is about 3.7 miles on average without SRE. However, eliminating the short repositioning legs would increase the average empty mileage to almost 5 miles. Nonetheless, the percentage of empty to revenue mileage is around 1% lower for CLARPTW-SRE compared to CLARPTW. Furthermore, the percentage of empty to revenue flights is around 70% for CLARPTW-SRE, suggesting that relocation of the passengers has led to a reduction of nearly 30% in the empty flight legs. Lastly, for CLARPTW, the percentage of empty to revenue flights is slightly over 100%. The reason is that the almost zero empty distances and flight times would cause the aircraft to conduct two consecutive empty flight legs in rare cases, possibly due to time window constraints.

Table 9.15 evaluates the performance measure associated with the connecting legs, which are conducted to eliminate the short repositioning flight legs. For $Q = 1$, 30% of the revenue flights are connecting legs, 18.4% of which have the desired pick-up UAT pad, while 73.6% have the desired drop-off UAT pad. Only 8.0% of the connecting flight legs have undesired pick-up and drop-off pads, implying that the passengers would experience two relocations to take these flights.

Table 9.15 Estimated mean of performance measures associated with connecting flight legs

Q	Percentage of Connecting Flights to Revenue Flights (%)	Percentage of Connecting Flights with the Desired Pick-up UAT Pad (%)	Percentage of Connecting Flights with Desired Drop-off UAT Pad (%)	Percentage of Connecting Flights with Undesired Pick-up and Drop-off UAT Pads (%)
1	30.6	18.4	73.6	8.0
2	34.5	17.2	74.8	8.0

Table 9.16 estimates the performance measures associated with aerial service time for 30 replications. The results suggest that serving a revenue and empty flight with $Q = 1$ would take, on average, around 22.7 and 4.4 minutes, respectively. The total aerial service time, i.e., the total time an aircraft is in use, per served revenue flight is around 25.8 minutes. Air pooling reduces the total aerial service time per served request by 24%, from 25.8 minutes with $Q = 1$ to 19.5 minutes with $Q = 2$. Morover, the aircraft utilization decreases from 91% to 85% for $Q = 2$ compared to $Q = 1$ despite serving more requests.

Table 9.16 Estimated mean of performance measures associated with aerial service time

Q	Mean of Aerial Service Time of Revenue Flight (min)	Mean of Aerial Service Time of Empty Flight (min)	Mean of Total Aerial Service Time per Revenue Flight (min)	Mean of Total Aerial Service Time per Served Request (min)	Aircraft Utilization (%)
1	22.7	4.4	25.8	25.8	91.1
2	22.5	4.6	25.6	19.5	85.0

9.3.4 UAT Operator Revenue

Table 9.17 Estimated mean of performance measures associated with passenger revenue

Q	Total Aerial Mileage (mi)	Available Seat Mile (ASM)	Total Passenger Revenue (PR/α)	Passenger Revenue per Available Seat Mile (PRASM/α)
1	45,877	45,877	38,366	0.836
2	42,492	84,985	46,759	0.550

Table 9.17 presents performance measures associated with passenger revenue, where α denotes the revenue per passenger mile. As previously discussed in Table 9.9, the rate of rejected requests decreases by 17% from $Q = 1$ to 2. Nonetheless, the total aerial milage decreases around 7% by increasing Q from 1 to 2. The total passenger revenue (PR) increases nearly 22% with air pooling as more passengers are served. However, passenger revenue per available seat mile (PRASM) decreases from 0.836α to 0.550α, suggesting low shared flights.

9.4 Numerical Results: Sensitivity Analyses

Even though some companies currently offer air taxi service, on-demand and at-scale UAT operations using eVTOL technology is a conceptual system that does not exist. Consequently, there are uncertainties around the parameters used in the base-case experiment. Factors such as price per minute of flight, the density of UAT pads, aerial cruising speed, and the ground-based access time could significantly impact the UAT market [6]. For instance, Grandl et al. [6] suggest that increasing the ground-based transportation from 5 to 15 minutes for the first and last mile of the trip would decrease the UAT market by 20%, while reducing the cruising speed from 200 km per hour (124 mph) to 70 km per hour (43 mph) would lead to a 30% reduction in the UAT market.

As a result, in this section, we design and conduct 11 experiments to study the sensitivity of the performance measures to 4 exogenous and 7 design parameters. The four exogenous parameters include the request arrival intensity ($\mathcal{T}^{INT}$), request spread (σ), driving speed (over euclidean distance) (v^{DRIVE}), and the detour factor of the aerial path (ϵ). The seven design parameters are aerial speed (v^{AIR}), maximum acceptable delay (ω), maximum advance reservation window ($\mathcal{T}^{ADV}$), the ratio of revenue per passenger mile to cost per mile (α/β), the ratio of relocation cost to cost per mile (γ_1/β), boarding and deboarding duration ($T^{BOARD}, T^{DEBOARD}$), and departure gate ingress time and arrival gate egress time ($T_{ri}^{DGATE}, T_{ri}^{AGATE}$).

Table 9.18 summarizes the parameters in each experiment, where the bold numbers show the value in the base-case experiment. Each experiment is replicated 20 times, where the stochasticity is associated with requests arrival time and OD pair in addition to the initial locations of the UAT aircraft.

Table 9.18 Values of exogenous and design parameters used in the sensitivity experiments

	$\mathcal{T}^{INT}$ (sec)	σ (mi)	v^{DRIVE} (mph)	ϵ	v^{AIR} (mph)	ω (min)	$\mathcal{T}^{ADV}$ (min)	α/β	γ_1/β	(T^{BOARD}, $T^{DEBOARD}$) (min)	(T_{ri}^{DGATE}, T_{ri}^{AGATE}) (min)
Exogenous Parameters											
E1	10, 15, **20**, 25, 30	2	20	0.1	150	15	30	2	0	(3,2)	(3,2)
E2	20	1, **2**, 3, 4	20	0.1	150	15	30	2	0	(3,2)	(3,2)
E3	20	2	10, **20**, 30	0.1	150	15	30	2	0	(3,2)	(3,2)
E4	20	2	20	0.05, **0.1**, 0.15, 0.2	150	15	30	2	0	(3,2)	(3,2)
Design Parameters											
E5	20	2	20	0.1	100, 125, **150**, 175	15	30	2	0	(3,2)	(3,2)
E6	20	2	20	0.1	150	5, 10, **15**, 20	30	2	0	(3,2)	(3,2)
E7	20	2	20	0.1	150	15	1, 5, 10, 20, **30**, 40, 60	2	0	(3,2)	(3,2)
E8	20	2	20	0.1	150	15	30	1.2, 1.5, **2.0**, 2.5, 3.0	0	(3,2)	(3,2)
E9	20	2	20	0.1	150	15	30	2	**0**, 1, 2, 5, 10	(3,2)	(3,2)
E10	20	2	20	0.1	150	15	30	2	0	(2,1), **(3,2)**, (5,3), (8,5)	(3,2)
E11	20	2	20	0.1	150	15	30	2	0	(3,2)	**(3,2)**, (5,4), (10,8)

9.4.1 Request Intensity ($\mathcal{T}^{INT}$)

With requests arriving more frequently, more opportunities are present for demand consolidation since more requests will have adjacent origins and destinations. In experiment E1, six levels of demand are examined [136]: $\mathcal{T}^{INT}$ = 10, 15, 20, 25, 30, and 40. With $\mathcal{T}^{INT}$ = 10 seconds, the demand is so high that more than 50% of requests are rejected. With $\mathcal{T}^{INT}$ = 20 seconds (base case), the demand roughly meets the supply. And with $\mathcal{T}^{INT}$ = 40 seconds, the demand is so low that aircraft utilization is around 50%. Table 9.19 presents the estimated mean and standard error of the mean (SEM) of the number of requests over the planning horizon for various request interarrival times ($\boldsymbol{\mathcal{T}^{INT}}$).

Table 9.19 Estimated mean and standard error of the mean (SEM) of number of requests over the planning horizon for various request interarrival times ($\boldsymbol{\mathcal{T}^{INT}}$)

$\mathcal{T}^{INT}$ (second)	λ (requests/hour)	Number of Requests over 8 Hours	
		Mean	SEM
10	360	2,877	11.1
15	240	1,925	9.4
20	180	1,444	8.6
25	144	1,158	7.8
30	120	962	8.3
40	90	719	7.0

Table 9.20 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for request intensity ($\mathcal{T}^{INT}$) of 10, 15, 20, 25, 30, and 40 seconds with Q = 1 and 2, while Figure 9.7 depicts the sensitivity of select performance measures to $\mathcal{T}^{INT}$. For Q = 2, Figure 9.7(c) shows that decreasing $\mathcal{T}^{INT}$ from 40 seconds (where the utilization is about 50%) to 10 seconds (where more than 40% of the requests are rejected) would increase the average load factor by nearly 12%. Even with the low demand intensity of $\mathcal{T}^{INT}$ = 40 seconds, the average load factor is 60%, suggesting some flights could be shared.

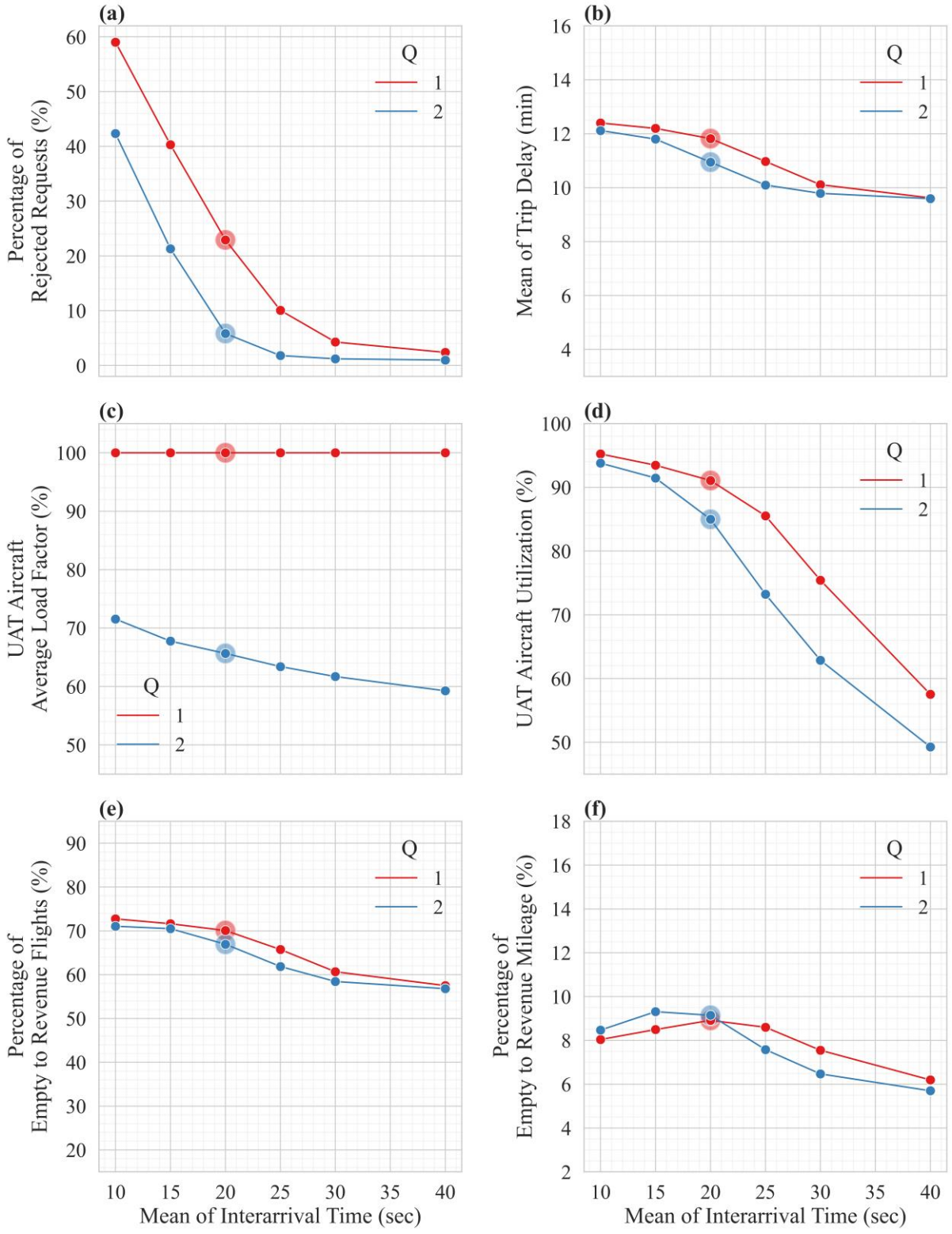

Figure 9.7 Sensitivity of performance measures to request intensity ($\mathcal{T}^{INT}$) of 10, 15, 20, 25, 30, and 40 seconds for aircraft with capacities of 1 and 2

Table 9.20 Impacts of request intensity ($\mathcal{T}^{INT}$) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
$\mathcal{T}^{INT}$ (seconds)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
10	95.2	40.9	72.7	8.0	100	40.7	0.843	59.0	12.4	30.4	0.6	0.29
15	93.5	41.1	71.6	8.5	100	39.6	0.840	40.3	12.2	30.0	0.7	0.31
20	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
25	85.6	41.1	65.7	8.6	100	35.9	0.840	10.1	11.0	27.3	0.8	0.38
30	75.6	40.6	60.7	7.6	100	31.7	0.848	4.3	10.1	25.3	0.9	0.43
40	58.8	40.0	57.5	6.2	100	24.1	0.859	2.4	9.6	24.2	1.0	0.46
$Q = 2$												
10	93.8	28.6	71.0	8.5	71.5	57.1	0.603	42.3	12.1	30.0	3.6	0.82
15	91.5	30.4	70.5	9.3	67.8	52.2	0.567	21.3	11.8	29.3	3.3	0.77
20	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
25	73.4	31.8	61.9	7.6	63.4	39.0	0.540	1.8	10.1	25.5	3.0	0.76
30	63.6	32.4	58.4	6.5	61.7	32.6	0.531	1.2	9.8	24.7	2.9	0.75
40	51.1	33.4	56.8	5.7	59.3	24.4	0.513	1.0	9.6	24.2	2.6	0.70

Figure 9.7(b) suggests that the average trip delay per passenger decreases by 2 minutes for lower utilization rates. However, even with the utilization of 50-60% for $\mathcal{T}^{INT} = 40$ seconds, the average delay is 10 minutes. The reason is that the delay is not explicitly minimized in the objective function.

Furthermore, the percentage of empty to revenue flights is higher when the utilization is high. Figure 9.7(e) depicts that the percentage of empty to revenue flights decreases by more than 10% by increasing $\mathcal{T}^{INT}$ from 10 seconds to 40 seconds. Figure 9.7(f) suggests that the percentage of empty to revenue mileage is lower for the high and low intensity of the requests compared to the medium request intensity.

9.4.2 Request Spread (σ)

Table 9.21 Impacts of request spread (σ) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				**Revenue**			**User Experience**				
σ (miles)	**Aircraft Utilization (%)**	**Mean of Total Aerial Mileage per Served Request (mi)**	**Percentage of Empty to Revenue Flights (%)**	**Percentage of Empty to Revenue Mileage (%)**	**Average Load Factor (%)**	**PR/α (thousands of miles)**	**PRASM/α**	**Percentage of Rejected Requests (%)**	**Average Trip Delay (min)**	**Mean of Percentage of Trip Delay (%)**	**Average Ground Travel Time (min)**	**Average Number of Relocations**
$Q = 1$												
1	90.8	39.4	36.2	4.6	100	40.6	0.874	18.3	11.2	28.0	1.6	0.73
2	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
3	90.7	42.7	83.6	12.4	100	37.0	0.810	26.0	12.3	30.1	0.4	0.18
4	90.6	44.2	90.0	15.5	100	36.2	0.788	28.1	12.5	30.5	0.2	0.11
$Q = 2$												
1	59.6	20.6	19.9	2.8	91.6	49.2	0.831	0.4	9.9	25.5	5.6	1.41
2	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
3	89.8	39.0	83.8	12.7	54.8	40.0	0.443	19.8	12.0	29.6	1.3	0.34
4	90.3	42.7	89.8	15.6	51.8	37.3	0.407	25.7	12.4	30.3	0.6	0.18

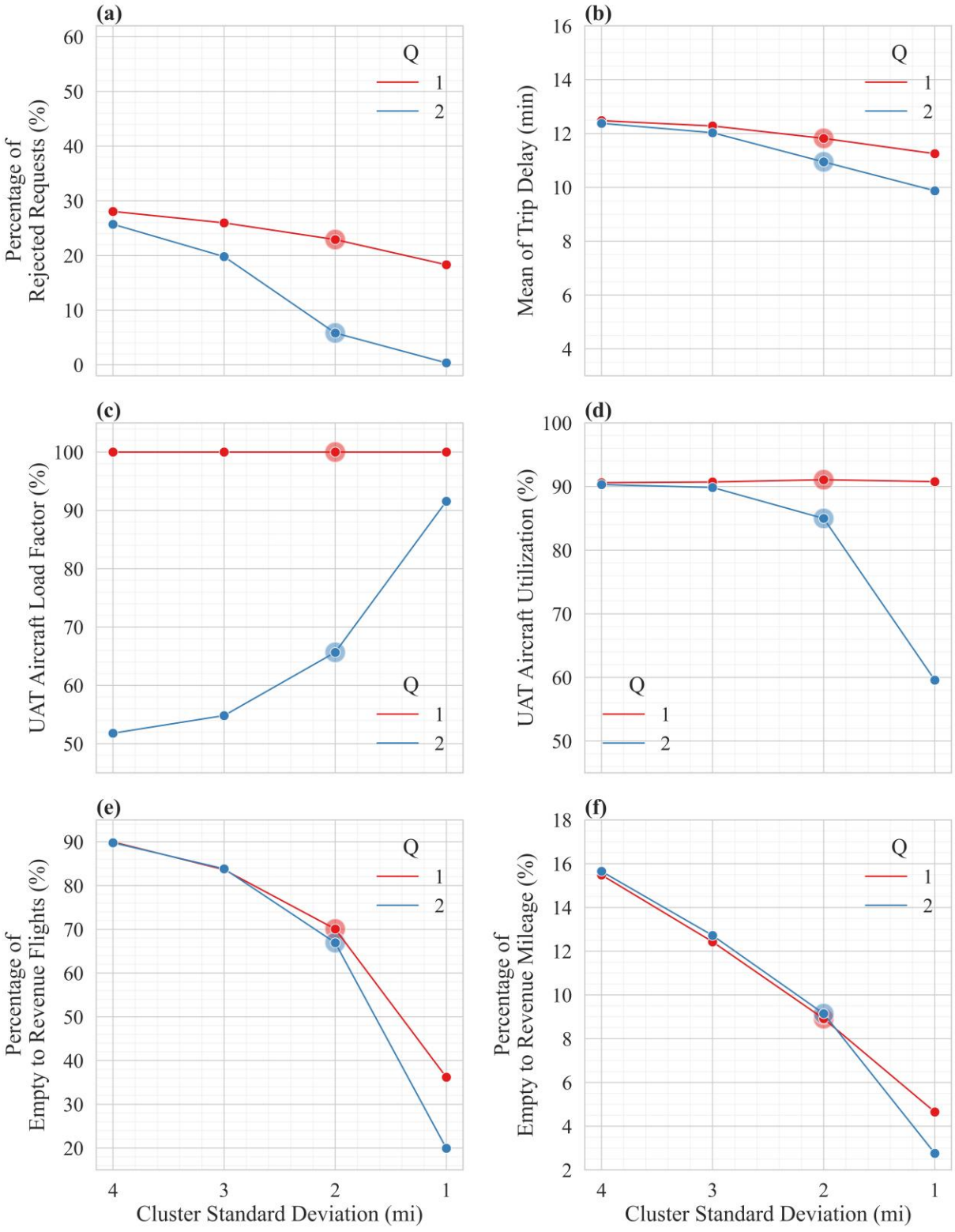

Figure 9.8 Sensitivity of performance measures to request spread (σ) of 1, 2, 3, and 4 miles for aircraft with capacities of 1 and 2

The closer the requests are spread around the centroid, and therefore, the lower is σ, the higher is the chance of eliminating the short empty distances and consolidating the demand. Table 9.21 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for request spread (σ) of 1, 2, 3, and 4 miles with $Q = 1$ and 2, while Figure 9.8 depicts the sensitivity of select performance measures to σ.

In this experiment, the ground speed of 20 mph and maximum acceptable delay of 15 minutes suggests that the requests could be relocated within an approximate radius of 5 miles. However, the access distance is capped at 3 miles (i.e., 10% of one edge). Consequently, as shown in Figure 9.8(c), at $\sigma = 3$ and 4 miles, the load factor is slightly over 50%, implying that only a few flights are pooled. At $\sigma = 1$ mile, the load factor increase to more than 90%.

Figure 9.8(a) shows that decreasing σ from 4 miles to 1 reduces the rejection of the request by nearly 10% for $Q = 1$. However, the noticeable benefit is seen for the capacity of 2, where the percentage of requests gets close to zero, resulting from the high number of shared flights. Moreover, low request spread and demand consolidation lead to a 2-minute reduction in the average trip delay per passenger and aircraft utilization of 60% at $\sigma = 1$ and $Q = 2$.

Figure 9.8(e) and (f) depict the percentage of empty to revenue flight and empty to revenue mileage, respectively. The elimination of short repositioning flight legs depends on the distance between the destination of one request to the origin of another. As a result, lower σ would provide higher opportunities for the elimination of short repositioning legs. The percentage of empty to revenue flights could vary between 20 and 90%, while the percentage of empty to revenue mileage could vary between 2.8% and 16%, depending on the spread.

9.4.3 Driving Speed (v^{DRIVE})

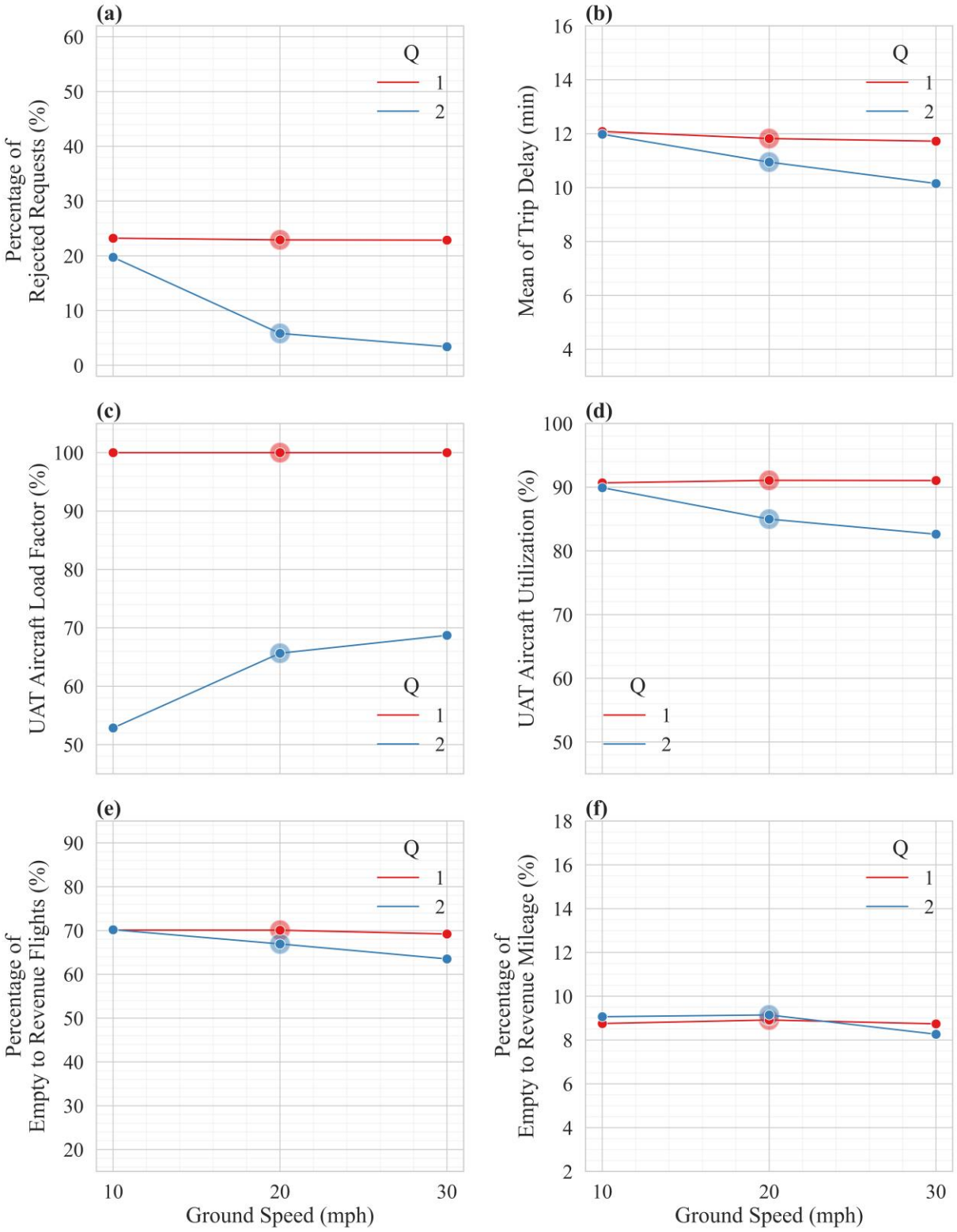

Figure 9.9 Sensitivity of performance measures to ground speed (v^{DRIVE}) of 10, 20, and 30 mph for aircraft with capacities of 1 and 2

Table 9.22 Impacts of ground speed (v^{DRIVE}) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
v^{DRIVE} (mph)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
10	90.7	41.1	70.1	8.8	100	38.2	0.838	23.2	12.1	29.8	1.3	0.32
20	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
30	91.0	41.1	69.2	8.7	100	38.4	0.838	22.9	11.7	28.9	0.5	0.34
$Q = 2$												
10	89.9	39.0	70.2	9.1	52.9	39.9	0.442	19.7	12.0	29.6	1.9	0.41
20	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
30	82.6	29.5	63.5	8.3	68.7	47.9	0.582	3.4	10.2	25.7	2.5	0.85

Ground speed would significantly impact access time and, therefore, relocation of the passengers on the ground to eliminate short repositioning legs or to consolidate the demand. Table 9.22 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for the ground speed (v^{DRIVE}) of 10, 20, and 30 with $Q = 1$ and 2, while Figure 9.9 depicts the sensitivity of select performance measures to v^{DRIVE}.

Figure 9.9 illustrates that for $Q = 1$, the performance measures are not sensitive to the driving speed. The reason is that with $Q = 1$, the ground transportation is within $\Delta^{EMPTY} = 1$ mile (or a total of 2 miles for both the origin and destination), while $v^{DRIVE} = 10$ mph and $\omega = 15$ minutes provide a 2.5-mile accessible radius on the ground.

Figure 9.9(c) shows that with the ground speed of 10 mph over euclidean distances, the average load factor is slightly over 50% for $Q = 2$, implying that decreasing the ground speed from 20 mph to 10 mph will lead to nearly no demand consolidation since relocating the passengers on the ground is so slow that there cannot be moved in time while satisfying the maximum delay of $\omega = 15$ minutes. The average load factor could be increased to nearly 70% with a driving speed of 30 mph, highlighting the importance of fast and reliable ground-based transportation in the success of the proposed UAT concept of operations with demand consolidation. As a result of higher speed and demand consolidation, the rejection rate would decrease from 23.2% to 3.4%, and the aircraft utilization would decrease by around 8% for $Q = 2$.

As seen in Figure 9.9(e) and Figure 9.9(f), the percentage of empty to revenue flights decreases nearly 7% by increasing v^{DRIVE} from 10 mph to 30 mph for $Q = 2$, however, the percentage of empty to revenue mileage is not noticeably sensitive to the driving speed.

Lastly, Table 9.22 shows that, with $Q = 2$, the average ground travel time is 1.9 minutes for v^{DRIVE} =10 mph and 2.5 minutes for v^{DRIVE} =30 mph, implying that the requests could be relocated to further distances on the ground to take a shared flight. Moreover, the average number of transfers doubles as the ground speed increases from 10 mph to 30 mph under $Q = 2$.

9.4.4 Detour Factor (ϵ)

The detour factor specifies how the flight distance deviates from the straight-line distance between the starting and ending point, and therefore, it directly impacts the flight service time. With a lower detour factor, the service time would decrease, and consequently, more requests could be served. Table 9.23 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for the detour factor (ϵ) of 0.05, 0.1, 0.15, and 0.2 with $Q = 1$ and 2, while Figure 9.10 depicts the sensitivity of select performance measures to ϵ.

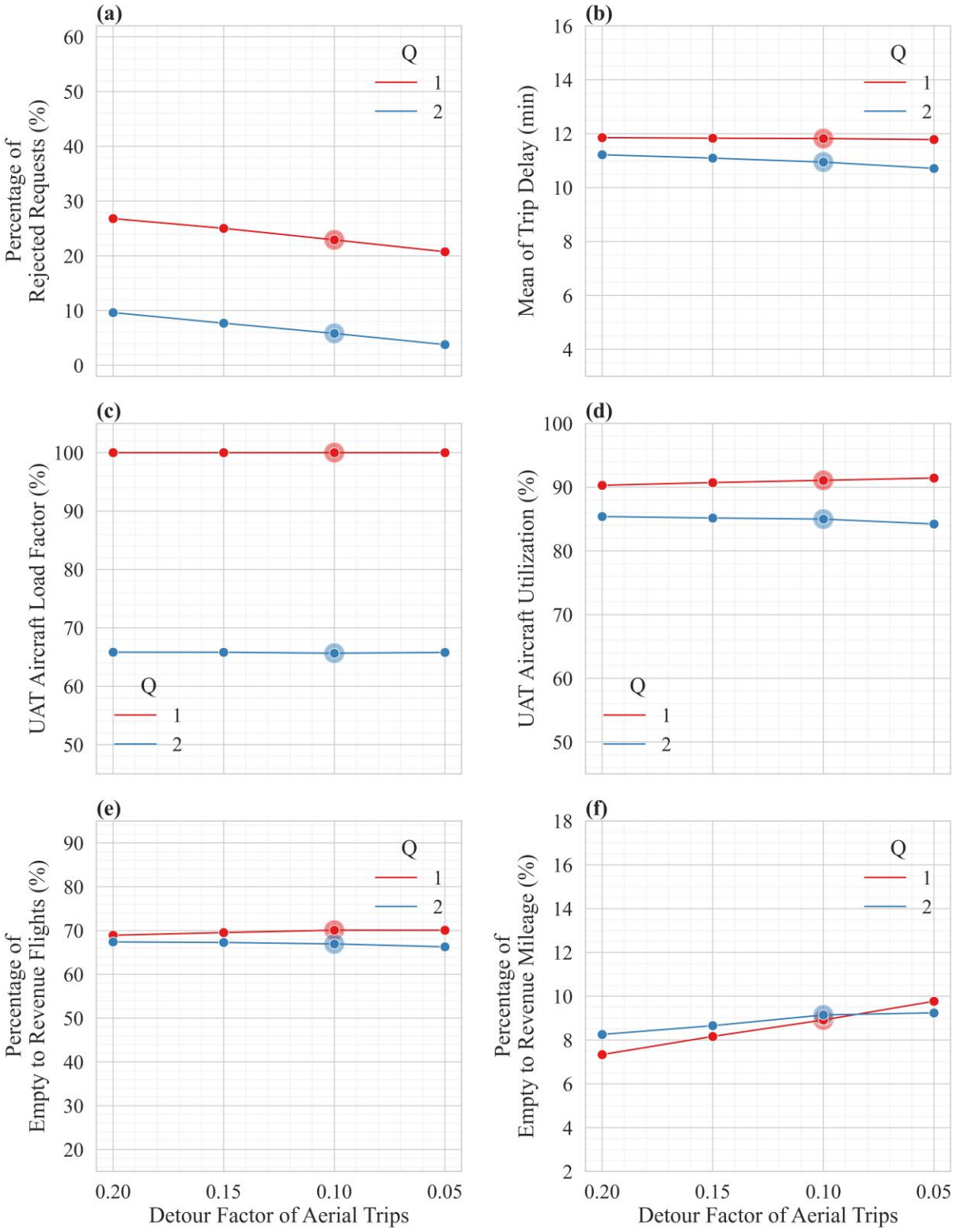

Figure 9.10 Sensitivity of performance measures to detour factor (ϵ) of 0.2, 0.15, 0.1, and 0.05 for aircraft with capacities of 1 and 2

Table 9.23 Impacts of detour factor (ϵ) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
ϵ	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
0.05	91.4	39.6	70.1	9.8	100	39.4	0.870	20.7	11.8	29.7	0.7	0.33
0.10	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
0.15	90.7	42.8	69.5	8.2	100	37.3	0.806	25.0	11.8	28.7	0.7	0.33
0.20	90.3	44.2	68.9	7.3	100	36.4	0.778	26.8	11.9	28.3	0.7	0.34
$Q = 2$												
0.05	84.2	29.7	66.3	9.2	65.8	47.7	0.578	3.8	10.7	27.4	3.2	0.77
0.10	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
0.15	85.2	32.4	67.3	8.7	65.8	45.8	0.530	7.7	11.1	27.2	3.2	0.76
0.20	85.4	33.7	67.4	8.3	65.8	44.9	0.510	9.6	11.2	27.1	3.2	0.76

Table 9.23 shows that increasing the detour factor from 0.05 to 0.2 could increase the average total aerial mileage by at least 4 miles. Correspondingly, the percentage of empty to revenue mileage decreases, which seems to be due to high rejection rates at $\epsilon = 0.2$. Additiontionally, as seen in Figure 9.10(a), the percentage of rejected requests decreases by nearly 6% for both $Q = 1$ and 2. Nonetheless, the detour factor does not seem to have a significant impact on ground-based transportation, demand consolidation, and elimination of short repositioning legs.

9.4.5 Aerial Speed (v^{AIR})

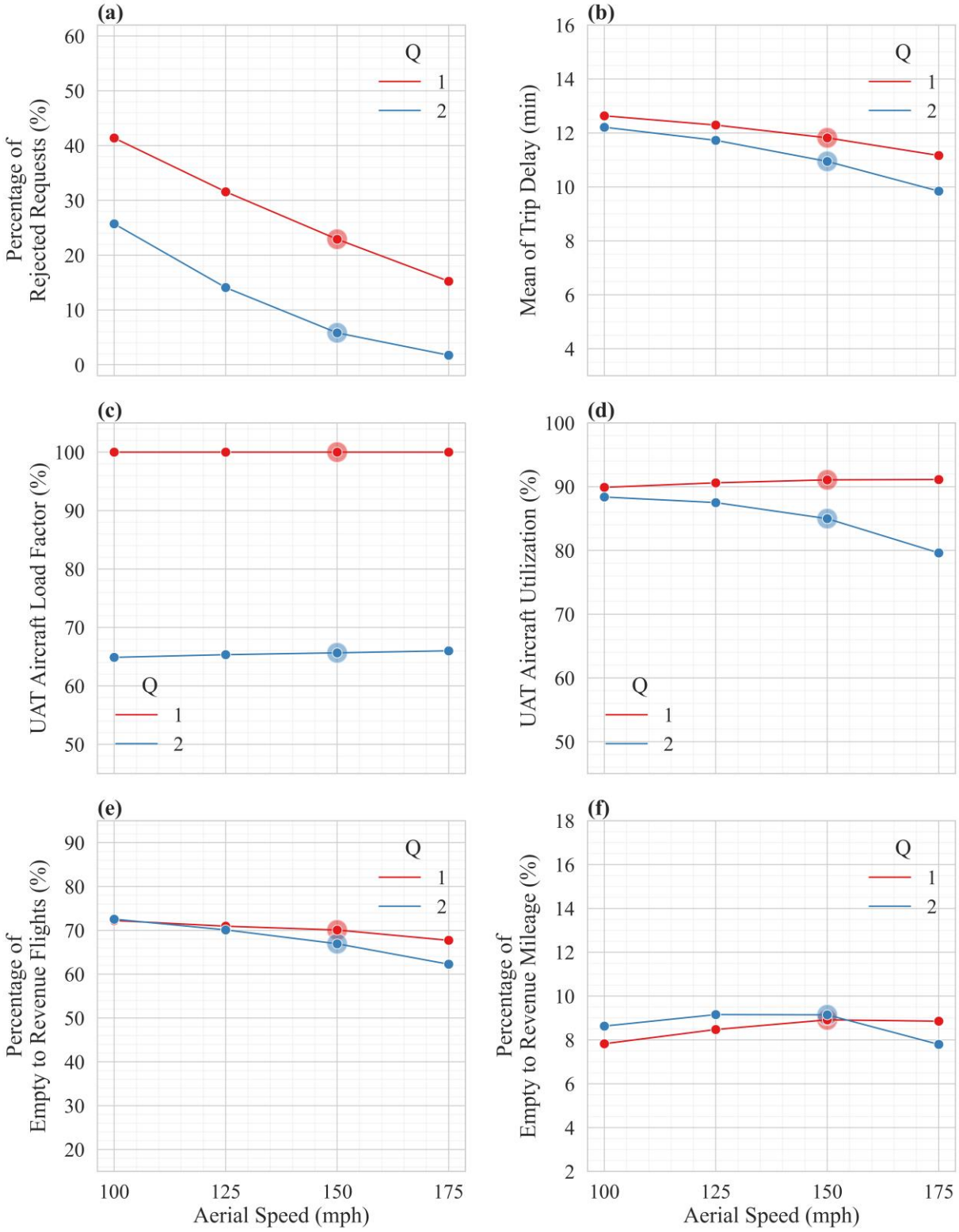

Figure 9.11 Sensitivity of performance measures to aerial speed (v^{AIR}) of 100, 125, 150, and 175 mph for aircraft with capacities of 1 and 2

Table 9.24 Impacts of aerial speed (v^{AIR}) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
v^{AIR} (mph)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
100	89.9	40.7	72.2	7.8	100	29.1	0.845	41.4	12.6	26.1	0.6	0.29
125	90.6	41.0	70.9	8.5	100	34.0	0.840	31.6	12.3	28.1	0.7	0.31
150	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
175	91.1	41.1	67.7	8.9	100	42.2	0.838	15.3	11.2	29.4	0.8	0.36
$Q = 2$												
100	88.4	31.5	72.6	8.6	64.9	36.9	0.546	25.7	12.2	25.4	2.9	0.68
125	87.5	31.4	70.1	9.2	65.3	42.6	0.548	14.1	11.7	27.1	3.0	0.72
150	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
175	79.6	30.6	62.3	7.8	66.0	48.7	0.561	1.8	9.8	26.4	3.3	0.81

Table 9.24 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for aerial speed (v^{AIR}) of 100, 125, 150, and 175 mph with $Q = 1$ and 2, while Figure 9.11 depicts the sensitivity of select performance measures to the aerial speed. Aerial speed impacts the cruise time and, therefore, the service time. Consequently, a fleet with a higher aerial speed would serve more requests than a fleet with the same size but lower aerial speed. Increasing the aerial speed from 100 mph to 175 mph would decrease the rejection rate from 41.4% to 15.3% for $Q = 1$ and from 25.7% to 1.8% for $Q = 2$.

Higher aerial speed leads to a slightly higher average number of relocations and ground-based travel time for both $Q = 1$ and 2 since serving more requests would provide more opportunities for demand consolidation. Increasing the aerial speed from 100 mph to 175 mph leads to a 1% increase in the average load factor. Therefore, even though more requests are served with higher aerial speed, the percentage of shared flights is uninfluenced.

Two factors would impact the percentage of empty to revenue flights, depicted in Figure 9.11(e). First, in a busy system with high utilization, the operator may have to reposition an aircraft from farther locations to meet the time constraints, while lower utilization suggests that more aircraft and flights would be available without violating the acceptable delay of the requests. Second, serving more requests means serving more flight legs, and therefore, more alternatives are available within the origin and destination of a request to eliminate short repositioning flight legs. Consequently, the percentage of empty to revenue flights decreases by nearly 5% for $Q = 1$ in Figure 9.11(e), even though the utilization has not decreased. For $Q = 2$, the percentage of empty to revenue flights decreases by 10%, given the decrease in rejected requests and utilization.

9.4.6 Maximum Acceptable Delay (*ω*)

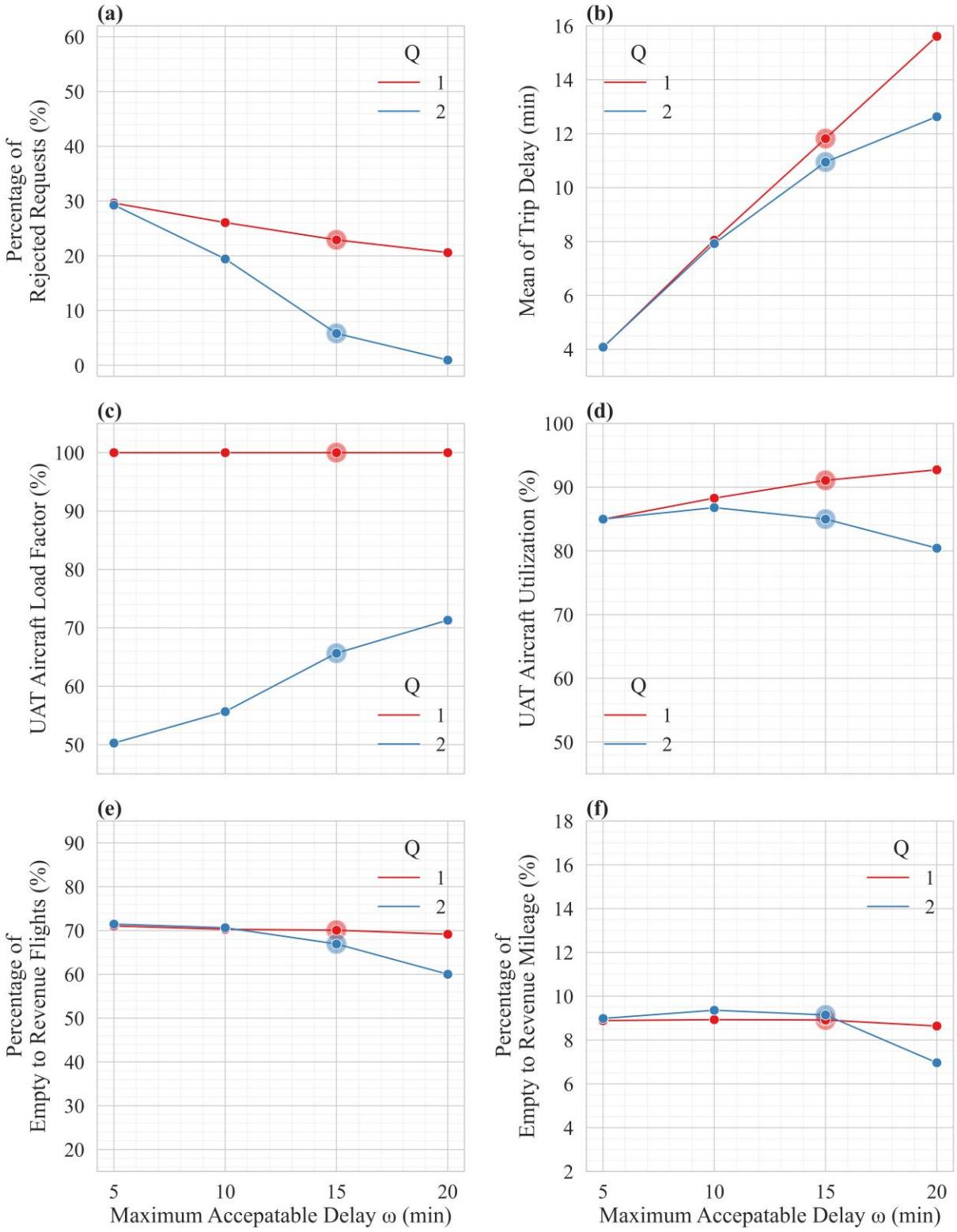

Figure 9.12 Sensitivity of performance measures to maximum acceptable delay (*ω*) of 5, 10, 15, and 20 minutes for aircraft with capacities of 1 and 2

Table 9.25 Impacts of maximum acceptable delay (ω) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
ω (minutes)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
5	85.0	41.3	71.0	8.9	100	35.1	0.836	29.6	4.1	12.7	0.7	0.30
10	88.3	41.2	70.3	8.9	100	36.8	0.836	26.1	8.1	22.2	0.7	0.32
15	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
20	92.7	41.0	69.1	8.6	100	39.5	0.839	20.6	15.6	34.9	0.8	0.34
$Q = 2$												
5	85.0	41.1	71.5	9.0	50.3	35.3	0.420	29.3	4.1	12.8	0.7	0.31
10	86.8	37.1	70.6	9.4	55.7	40.1	0.464	19.4	7.9	21.9	1.4	0.49
15	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
20	80.4	28.0	60.0	7.0	71.3	49.0	0.612	1.0	12.6	29.6	4.1	0.92

Table 9.25 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for maximum acceptable delay (ω) of 5, 10, 15, and 20 minutes for $Q = 1$ and 2, while Figure 9.12 depicts the sensitivity of select performance measures to ω. Increasing the maximum acceptable delay provides the UAT operator with more time to move the UAT aircraft to serve the requests. Additionally, the requests could be relocated to further distances on the ground in time to eliminate the short empty distances or consolidate the demand. As seen in Table 9.25, as ω increases, the average ground travel time and average number of relocations

increases. However, the increase is more significant for $Q = 2$ compared to $Q = 1$, since with the former the requests could be relocated within a 3-mile radius (i.e., Δ^{ACCESS}), while with the latter the requests could move with a 1-mile radius (i.e., Δ^{EMPTY}).

Figure 9.12(a) shows that with $\omega = 5$ minutes, 30% of the requests get rejected. However, increasing ω to 20 minutes in combination with demand consolidation (i.e., $Q = 2$) would decrease the rejection rate to 1%. Figure 9.12(c) shows that the average load factor for $\omega = 5$ minutes and $Q = 2$ is almost 50%, suggesting that the maximum accpetable delay of 5 minutes is too short for the requests to be relocated on the ground for demand consolidation. As ω increases to 20 minutes, the average load factor increases to 70%.

Figure 9.12(d) illustrates the utilization of the UAT aircraft. With $Q = 1$, as ω increases, more aircraft could be dispatched in time to serve the increasing number of requests. However, with $Q = 2$, having a maximum acceptable delay of 15 and 20 minutes prompts more air pooling. Consequently, the system becomes less busy, and the aircraft utilization for $\omega = 20$ minutes decreases to 80%.

Figure 9.12(b) demonstrates that the average trip delay decreases with increasing the maximum acceptable delay. While the relationship between average trip delay and the maximum acceptable delay is linear for $Q = 1$, for $Q = 2$, the rate of increase in average trip delay is decreasing, which could be attributed to the decreasing utilization.

Figure 9.12(e) shows that, with the decreasing utilization for $Q = 2$, the percentage of empty to revenue flights decreases by 11.5%, from 71.5% at $\omega = 10$ to 60% at $\omega = 20$. Correspondingly, the percentage of empty to revenue mileage decreases by 2.4% in Figure 9.12(f).

9.4.7 Maximum Reservation Time Window ($\mathcal{T}^{ADV}$)

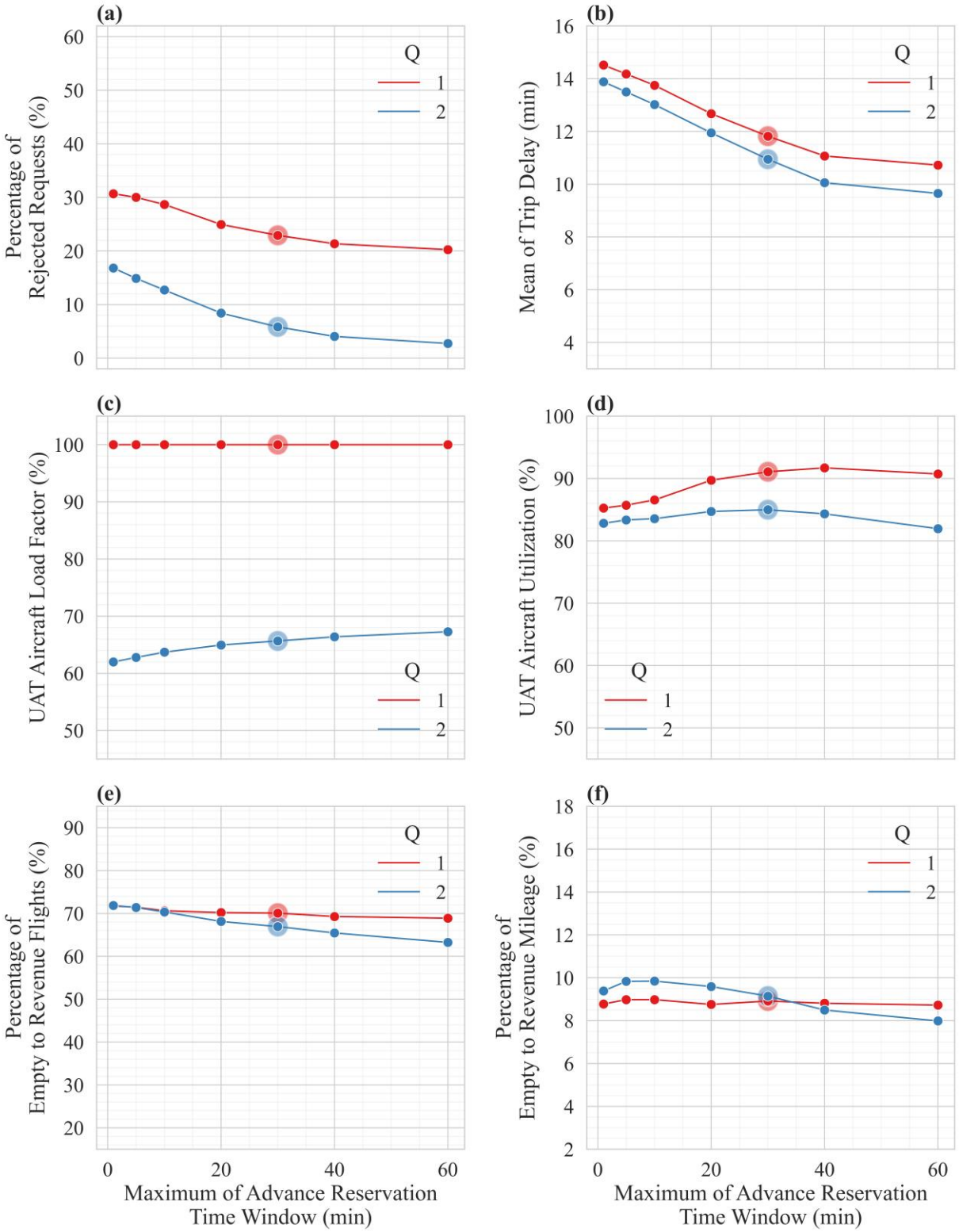

Figure 9.13 Sensitivity of performance measures to the maximum of the reservation time window ($\mathcal{T}^{ADV}$) of 1, 5, 10, 20, 30, 40, and 60 minutes for aircraft with capacities of 1 and 2

Table 9.26 Impacts maximum of advance reservation time window ($\mathcal{T}^{ADV}$) on performance measures associated with operator's cost and revenue and user inconvenience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
$\mathcal{T}^{ADV}$ (minutes)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
1	85.2	41.5	71.7	8.8	100	34.7	0.836	30.7	14.5	34.3	0.6	0.28
5	85.7	41.5	71.5	9.0	100	35.0	0.835	30.0	14.2	33.7	0.6	0.29
10	86.6	41.4	70.6	9.0	100	35.6	0.835	28.7	13.8	32.9	0.7	0.30
20	89.7	41.2	70.2	8.8	100	37.4	0.838	24.9	12.7	30.9	0.7	0.32
30	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
40	91.7	41.1	69.3	8.8	100	39.1	0.838	21.3	11.1	27.7	0.8	0.35
60	90.7	41.1	68.9	8.7	100	39.6	0.839	20.2	10.7	26.9	0.8	0.35
$Q = 2$												
1	82.8	33.5	71.9	9.4	62.0	41.6	0.516	16.8	13.9	33.2	2.5	0.62
5	83.4	33.1	71.4	9.8	62.8	42.5	0.521	14.9	13.5	32.5	2.6	0.64
10	83.6	32.5	70.4	9.8	63.7	43.4	0.530	12.7	13.0	31.6	2.8	0.67
20	84.7	31.7	68.1	9.6	65.0	45.5	0.542	8.4	11.9	29.5	3.0	0.72
30	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
40	84.3	30.6	65.5	8.5	66.4	47.6	0.561	4.1	10.1	25.5	3.3	0.80
60	82.0	30.0	63.3	8.0	67.3	48.2	0.571	2.7	9.7	24.6	3.5	0.85

Table 9.26 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for the maximum of the reservation time window ($\mathcal{T}^{ADV}$) of 1, 5, 10, 20, 30, 40, and 60 minutes for aircraft with $Q = 1$ and 2, while Figure 9.13 depicts the

sensitivity of select performance measures to $\mathcal{T}^{ADV}$. Knowing the requests ahead of their desired service time provides the UAT operator with more opportunities for efficient routing of the aircraft and demand consolidation. Consequently, as shown in Table 9.26, increasing $\mathcal{T}^{ADV}$ results in higher average ground travel time and number of relocations per request.

Figure 9.13(a) shows that increasing $\mathcal{T}^{ADV}$ from 1 minute to 60 minutes decreases the percentage of rejected requests by nearly 10% for $Q = 1$ and 14% for $Q = 2$. However, the benefits are diminishing, showing no noticeable improvements beyond the maximum of 40-minute advance notice. Figure 9.13(c) depicts the average load factor increases by 5% under $Q = 2$.

Figure 9.13(e) shows that the percentage of empty to revenue flight decreases by nearly 2.8% and 8.6% for $Q = 1$ and $Q = 2$, respectively. Figure 9.13(f) depicts that the percentage of empty to revenue mileage does not go under a noticeable change for $Q = 1$, while it varies 1.8% for $Q = 2$.

9.4.8 Ratio of Revenue per Passenger Mile to Cost per Mile (α/β)

Table 9.27 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for the ratio of revenue per passenger mile (α) to cost per mile (β) of 1.2, 1.5, 2, 2.5, 3, and 4 for aircraft with $Q = 1$ and 2, while Figure 9.14 depicts the sensitivity of select performance measures to α/β.

Without passenger pooling, serving requests are profitable as long as $\alpha q_r D_r^{OD} \gtrapprox 1.1\beta(D_r^{OD} + \overline{D}_r^0)$, where D_r^{OD} is the OD distance of the request r and is approximately equal to the revenue mileage for serving request r, $q_r = 1$ is the group size of the request, $\overline{D}_r^0$ denotes the empty repositioning mileage for serving request r, and 1.1 converts the geodesic distances on the ground to the aerial distance by taking into account the 0.1 detour factor.

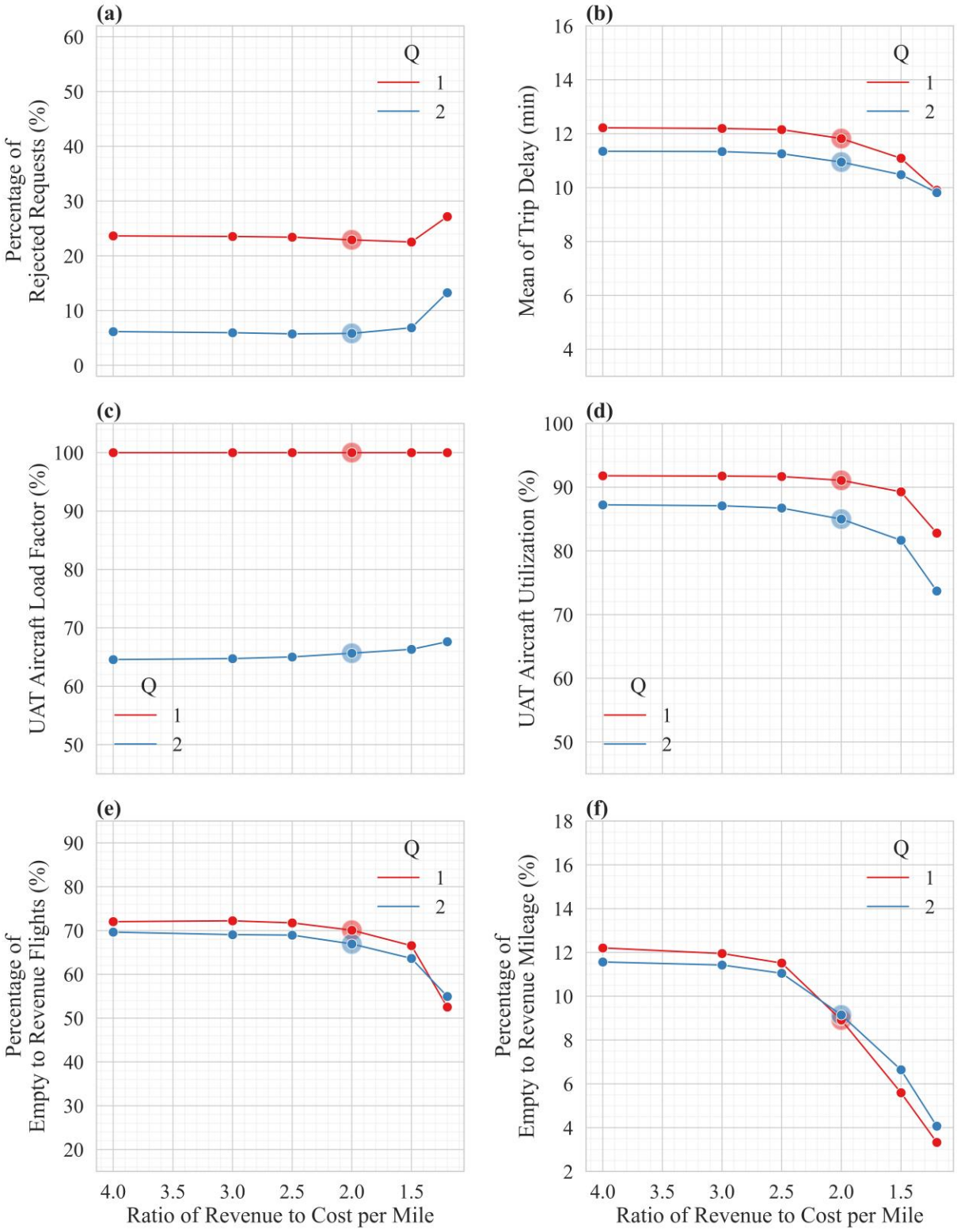

Figure 9.14 Sensitivity of performance measures to the ratio of revenue per passenger mile to cost per mile (α/β) of 1.2, 1.5, 2, 2.5, 3, and 4 for aircraft with capacities of 1 and 2

Table 9.27 Impacts of the ratio of revenue per passenger mile to cost per mile (α/β) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
α/β	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
1.2	82.8	39.4	52.5	3.3	100	36.6	0.884	27.2	9.9	24.9	1.1	0.52
1.5	89.3	39.7	66.5	5.6	100	38.3	0.863	22.5	11.1	27.6	0.8	0.37
2.0	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
2.5	91.7	42.0	71.7	11.5	100	37.9	0.817	23.4	12.2	30.0	0.7	0.31
3.0	91.7	42.1	72.2	12.0	100	37.8	0.814	23.5	12.2	30.1	0.7	0.30
4.0	91.8	42.2	72.0	12.2	100	37.8	0.812	23.6	12.2	30.1	0.7	0.31
$Q = 2$												
1.2	73.7	29.0	54.9	4.1	67.6	43.3	0.595	13.3	9.8	24.8	3.7	0.89
1.5	81.7	30.0	63.6	6.6	66.3	46.1	0.570	6.9	10.5	26.4	3.3	0.80
2.0	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
2.5	86.7	32.0	69.0	11.0	65.0	46.6	0.536	5.7	11.3	28.1	3.1	0.73
3.0	87.1	32.2	69.1	11.4	64.7	46.5	0.532	6.0	11.3	28.3	3.0	0.72
4.0	87.2	32.3	69.6	11.6	64.6	46.4	0.529	6.2	11.3	28.3	3.0	0.72

Consequently, $\alpha/\beta = 1.2$ implies that serving a request without passenger pooling is profitable if $0.09D_r^{OD} \gtrapprox \bar{D}_r^0$. Considering the configuration of the synthetic network with the demand generated around four centroids on the vertices of a square with a 30-mile edge, the value of $\alpha/\beta = 1.2$ would prevent the repositioning of the aircraft from one vertex to another to serve a

request. Therefore, the served requests are closer to each other under $\alpha/\beta = 1.2$, which provides more opportunities for short legs elimination and demand consolidation. As α/β increases, serving requests with higher empty to revenue mileage become profitable.

Table 9.27 verifies that with an increase in α/β the percentage of empty to revenue flights and empty to revenue mileage increases, while the average ground travel time, the number of relocations, and the load factor (for $Q = 2$) decreases. Additionally, the percentage of the rejected request is highest at α/β =1.2, with 27.2% for $Q = 1$ and 13.3% for Q =2. Even though the percentage of rejected requests decreases as α/β increases from 1.2, it reaches its lowest value at α/β =1.5 and 2.5 for $Q = 1$ and 2, respectively.

Furthermore, Figure 9.14(f) shows that for $\frac{\alpha}{\beta} \geq 2.5$, the percentage of empty to revenue mileage is around 12%. However, for $\frac{\alpha}{\beta} < 2.5$, this value could significantly decrease to 3.3%.

9.4.9 Ratio of Relocation Cost to Cost per Mile (γ_1/β)

Table 9.28 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for the ratio of relocation cost to cost per mile (γ_1/β) of 0, 1, 2, 5, and 10 for aircraft with $Q = 1$ and 2, while Figure 9.15Figure 9.16 depicts the sensitivity of select performance measures to γ_1/β. The value of γ_2 is set to 2 in this experiment. The results show that increasing γ_1/β from 0 to 1 increases the percentage of empty to revenue flights by nearly 4 to 5%, while at $\gamma_1/\beta = 2$, the percentage of empty to revenue flights increases to 80%. The other performance measures are almost insensitive to increasing γ_1/β from 0 to 2.

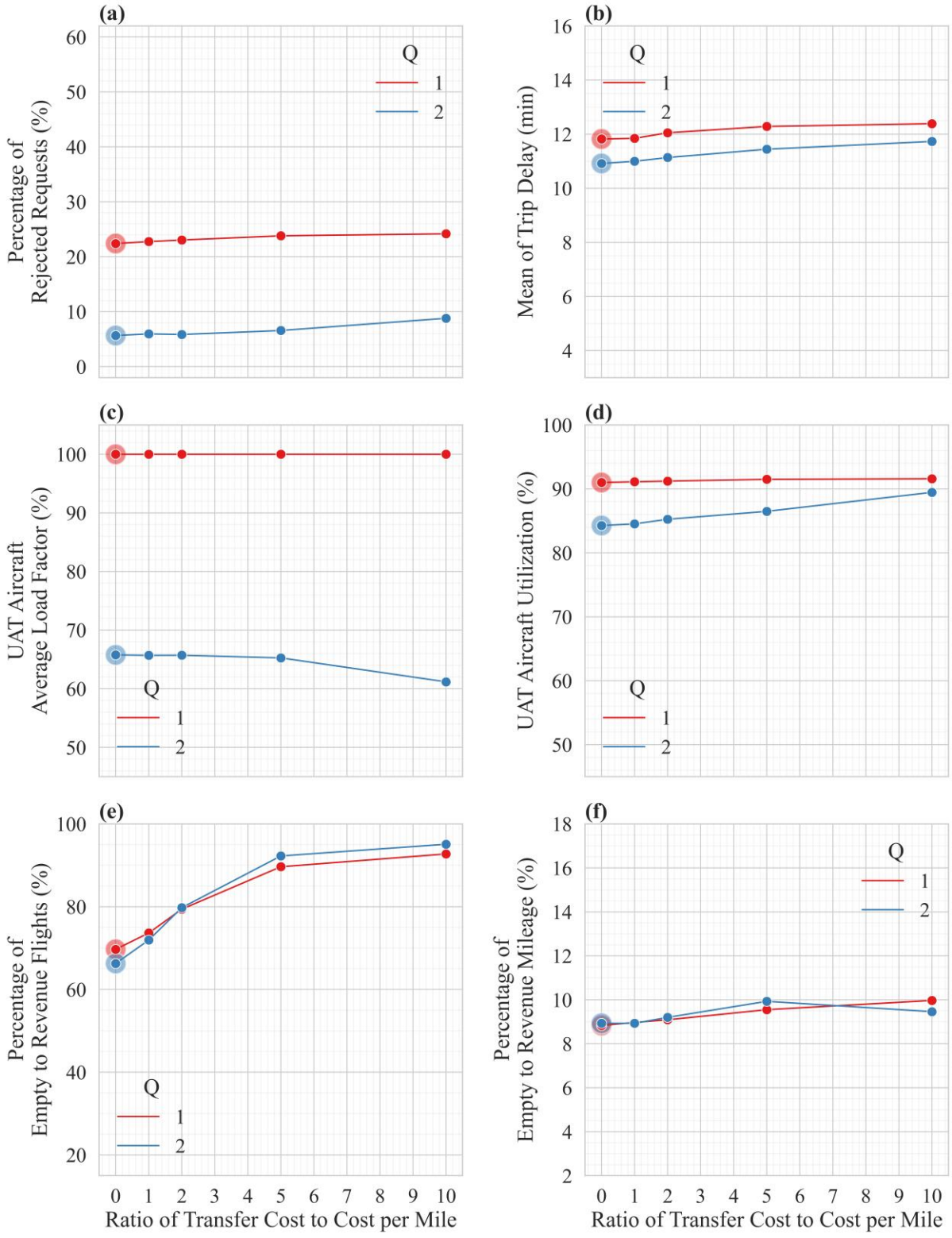

Figure 9.15 Sensitivity of performance measures to the ratio of relocation cost to cost per mile (γ_1/β) of 0, 1, 2, 5, and 10 for aircraft with capacities of 1 and 2

Table 9.28 Impacts of the ratio of the relocation cost to cost per mile (γ_1/β) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				**Revenue**			**User Experience**				
γ_1/β	**Aircraft Utilization (%)**	**Mean of Total Aerial Mileage per Served Request (mi)**	**Percentage of Empty to Revenue Flights (%)**	**Percentage of Empty to Revenue Mileage (%)**	**Average Load Factor (%)**	**PR/α (thousands of miles)**	**PRASM/α**	**Percentage of Rejected Requests (%)**	**Average Trip Delay (min)**	**Mean of Percentage of Trip Delay (%)**	**Average Ground Travel Time (min)**	**Average Number of Relocations**
$Q = 1$												
0	91.0	41.1	69.7	8.8	100	38.2	0.8	22.4	11.8	29.2	0.7	0.33
1	91.1	41.2	73.6	9.0	100	38.1	0.8	22.8	11.8	29.2	0.6	0.28
2	91.2	41.2	79.4	9.1	100	37.9	0.8	23.0	12.1	29.7	0.5	0.21
5	91.5	41.4	89.7	9.6	100	37.5	0.8	23.8	12.3	30.1	0.2	0.11
10	91.6	41.6	92.8	10.0	100	37.3	0.8	24.2	12.4	30.3	0.2	0.08
$Q = 2$												
0	84.3	31.0	66.3	8.9	65.8	46.4	0.6	5.7	10.9	27.4	3.2	0.77
1	84.5	31.1	71.9	8.9	65.7	46.2	0.6	6.0	11.0	27.5	3.1	0.71
2	85.2	31.2	79.8	9.2	65.7	46.2	0.6	5.9	11.1	27.8	2.9	0.64
5	86.5	31.6	92.3	9.9	65.3	45.9	0.5	6.6	11.4	28.4	2.6	0.53
10	89.5	33.4	95.1	9.5	61.2	44.8	0.5	8.8	11.7	29.0	2.0	0.41

Table 9.29 shows how the percentage of requests with 0, 1, and 2 relocations varies with γ_1/β. As a result of increasing γ_1/β from 0 to 1, the percentage of requests with 2 relocations becomes 0. At $\gamma_1/\beta = 10$, 92.4% of the requests have 0 relocation, impacting the average load factor. Additionally, the percentage of connecting flight legs decreases to around 7% since the cost is high that relocating the passengers to eliminate the short repositioning legs is not beneficial.

Table 9.29 Impacts of the ratio of the relocation cost to cost per mile (γ_1/β) on performance measures associated with relocations

γ_1/β	Percentage of Requests with 0 Relocation (%)	Percentage of Requests with 1 Relocation (%)	Percentage of Requests with 2 Relocations (%)	Mean of Empty Flight Mileage (mi)	Percentage of Connecting Flights to Revenue Flights (%)
$Q = 1$					
0	69.3	28.1	2.6	4.8	30.7
1	72.4	27.6	0.0	4.6	27.6
2	78.7	21.3	0.0	4.3	21.3
5	89.4	10.6	0.0	4.0	10.6
10	92.4	7.6	0.0	4.1	7.6
$Q = 2$					
0	49.2	25.0	25.8	5.0	35.3
1	52.7	23.9	23.3	4.7	30.7
2	58.7	18.4	22.9	4.3	22.9
5	69.1	8.6	22.3	4.0	9.8
10	76.4	6.4	17.2	3.7	6.6

9.4.10 Boarding and Deboarding Duration ($T^{BOARD}, T^{DEBOARD}$)

Table 9.30 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for boarding and deboarding duration ($T^{BOARD}, T^{DEBOARD}$) of (2,1), (3,2), (5,3), and (8,5) for aircraft with $Q = 1$ and 2, while Figure 9.16 depicts the sensitivity of select performance measures to T^{BOARD} and $T^{DEBOARD}$. Decreasing the boarding and deboarding time would decrease the flight service time. As seen in Figure 9.16 and Table 9.30, with ($T^{BOARD}, T^{DEBOARD}$) = (8,5) the rejection rate is 41.3% and 25.5% for $Q = 1$ and 2, respectively. However, with ($T^{BOARD}, T^{DEBOARD}$) = (2,1), the rejection rate reduces to 16.6% and 2.5% for $Q = 1$ and 2, respectively.

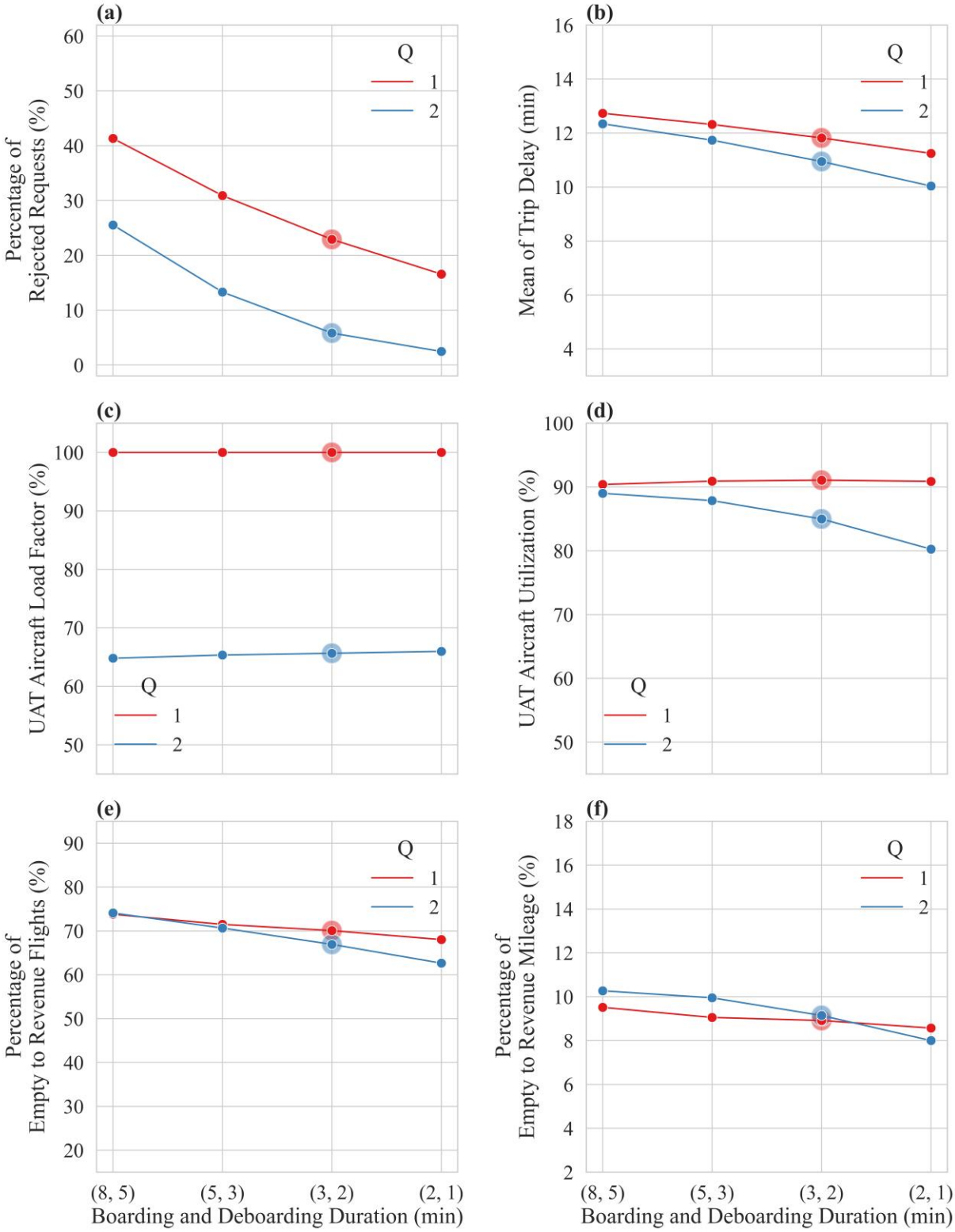

Figure 9.16 Sensitivity of performance measures to boarding and deboarding duration (T^{BOARD}, $T^{DEBOARD}$) of (2, 1), (3, 2), (5, 3), and (8, 5) minutes for aircraft with capacities of 1 and 2

Table 9.30 Impacts of boarding and deboarding duration ($T^{BOARD}, T^{DEBOARD}$) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
$(T^{BOARD}, T^{DEBOARD})$ (min)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
(2,1)	90.9	41.0	68.0	8.6	100	41.5	0.840	16.6	11.2	29.5	0.8	0.36
(3,2)	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
(5,3)	90.9	41.3	71.5	9.1	100	34.4	0.835	30.9	12.3	28.1	0.7	0.31
(8,5)	90.4	41.5	73.8	9.5	100	29.2	0.831	41.3	12.7	25.9	0.6	0.28
$Q = 2$												
(2,1)	80.3	30.7	62.7	8.0	66.0	48.4	0.560	2.5	10.0	26.8	3.3	0.81
(3,2)	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
(5,3)	87.9	31.6	70.7	10.0	65.4	43.0	0.543	13.3	11.7	27.1	3.0	0.72
(8,5)	89.0	32.1	74.1	10.3	64.8	37.1	0.537	25.5	12.3	25.3	2.8	0.67

As the boarding and deboarding duration decreases, more requests are served, and aircraft utilization decreases. Consequently, there is a slight increase in the average load factor under $Q = 2$ (1.2%), average ground travel time, and the average number of relocations. Furthermore, Figure 9.16(e) depicts the percentage of empty to revenue flights decreases by 5.8% and 11.4% for $Q = 1$ and $Q = 2$, respectively, as $(T^{BOARD}, T^{DEBOARD})$ decreases from (8,5) to (2,1), while Figure 9.16(f) demonstrates a decrease of 0.9% and 2.3% for $Q = 1$ and $Q = 2$, respectively.

9.4.11 Arrival and Departure Gate Access Time ($T_{ri}^{DGATE}, T_{ri}^{AGATE}$)

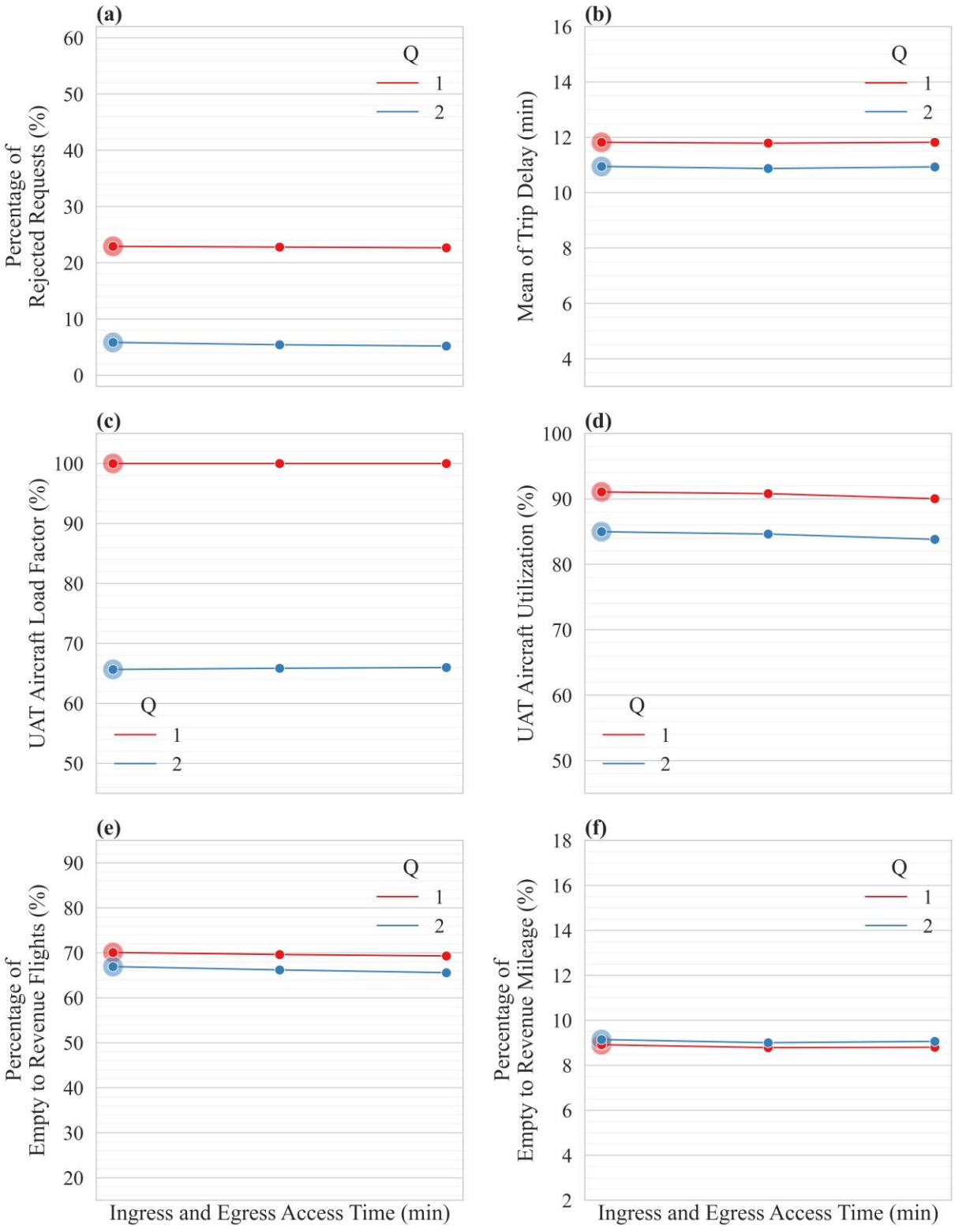

Figure 9.17 Sensitivity of performance measures to arrival and departure gate access duration ($T_{ri}^{DGATE}, T_{ri}^{AGATE}$) of (3, 2), (5, 4), and (10, 8) minutes for aircraft with capacities of 1 and 2

Table 9.31 Impacts of arrival and departure gate access duration ($T_{ri}^{DGATE}, T_{ri}^{AGATE}$) on performance measures associated with operator's cost and revenue and user experience for $Q = 1$ and 2

	Cost				Revenue			User Experience				
($T_{ri}^{DGATE}, T_{ri}^{AGATE}$) (min)	Aircraft Utilization (%)	Mean of Total Aerial Mileage per Served Request (mi)	Percentage of Empty to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Average Load Factor (%)	PR/α (thousands of miles)	PRASM/α (1/seat-mile)	Percentage of Rejected Requests (%)	Average Trip Delay (min)	Mean of Percentage of Trip Delay (%)	Average Ground Travel Time (min)	Average Number of Relocations
$Q = 1$												
(3,2)	91.1	41.2	70.1	8.9	100	38.4	0.837	22.9	11.8	29.2	0.7	0.33
(5,4)	90.8	41.1	69.6	8.8	100	38.4	0.838	22.8	11.8	26.5	0.7	0.33
(10,8)	90.0	41.1	69.3	8.8	100	38.5	0.838	22.7	11.8	22.1	0.7	0.33
$Q = 2$												
(3,2)	85.0	31.2	66.9	9.1	65.7	46.7	0.551	5.8	10.9	27.4	3.2	0.76
(5,3)	84.6	31.1	66.2	9.0	65.9	46.9	0.553	5.4	10.9	24.7	3.2	0.77
(8,5)	83.8	31.0	65.6	9.1	66.0	47.0	0.554	5.2	10.9	20.6	3.2	0.77

Table 9.31 summarizes the performance measures associated with UAT operator revenue and costs as well as user experience for arrival and departure gate access duration ($T_{ri}^{DGATE}, T_{ri}^{AGATE}$) of (3, 2), (5, 4), and (10, 8) for aircraft with $Q = 1$ and 2, while Figure 9.17 depicts the sensitivity of select performance measures to T_{ri}^{DGATE} and T_{ri}^{AGATE}. As the results show the performance measures are insensitive to the value of T_{ri}^{DGATE} and T_{ri}^{AGATE}. However, these values impact the total trip time and, therefore, the UAT demand.

9.5 Concluding Remarks

Urban air taxi (UAT) is the envisioned use case of passenger Urban Air Mobility (UAM) in its mature state. Given the ubiquitous operations of UAT, pooling the passengers and increasing the aircraft load factor is deemed as a critical step in the success of UAT operations. However, the absence of a dominant eVTOL aircraft technology and UAT operator feeds the uncertainty around UAT. To this end, we examine the impacts of various exogenous and design parameters on-demand consolidation using a dynamic solution framework through an event-based discrete-event simulation.

The runtime of the proposed policy suggests that CLARPTW-SRE could be employed in real time to address the UAT fleet operation problem. However, the maximum accessible distance (Δ^{ACCESS}), the minimum of repositioning distance (Δ^{EMPTY}), the maximum acceptable delay (ω), and the maximum of the reservation time window ($\mathcal{T}^{ADV}$) significantly impact the problem size and runtime. For narrow time windows, the routing would be limited to 1-3 flight legs, which could be solved quickly. The shorter is Δ^{EMPTY}, the fewer is the number of connecting flight legs. Finally, limiting Δ^{ACCESS} and ω would limit the number of requests that should be evaluated for assignment to different flight legs.

To provide an acceptable level of service and meaningful travel time savings that warrant the choice of UAT, the operator guarantees to limit the delay incurred due to wait times and relocations. The results show that providing service with short delays while relocating passengers on the ground hinges on fast and reliable ground-based transportation. For the synthetic network used in this study, increasing the driving speed from 10 mph to 20 mph results in a 14% increase in the average load factor. However, achieving the ground speed of 20 mph over short distances might be challenging, particularly in downtown areas of a densely populated city. A study on travel

time variability in Chicago using the data of transportation network providers (TNPs) shows that the OD pairs with high travel time variability are primarily located in the downtown area and have an average distance of nearly 3 miles [143]. Using the TNPs data in Chicago, the average speed over Euclidean distances smaller than 3 miles is 9 mph.

Another significant factor in demand consolidation is the spread of the demand. For the given experiment with the driving speed of 20 mph and the maximum delay of 15 minutes, reducing the standard deviation of the Gaussian distribution of requests around the centroids from 2 miles to 1 results in a 25% increase in average load factor. Closely spread demand would result in the average load factor of 90%, which is well beyond the range of 50%-80% estimated in [8,141]. Nonetheless, ground speed and demand spread, as the highly influential factors in demand consolidation, are exogenous information and are primarily beyond the control of the UAT operator. However, special attention should be given to these factors when selecting the passenger UAM market, particularly in the initial stages of the operation. Moreover, placing the UAM ports in locations that could provide a short and reliable ground access time to a dense and closely spread demand is another challenge facing the passenger-carrying UAM operations in the early stages. Locating the UAM ports near highways or a high capacity public transit system could provide a high access speed or a high demand density; however, whether they can provide both is highly dependent on the specific UAM market under consideration.

Among the design parameters, aerial speed is an influential factor in reducing the rate of rejection requests. However, it has minimal impacts on the demand consolidation and average load factor. The results suggest that a similar rejection rate could be achieved whether using high-speed aircraft with no demand consolidation or low-speed aircraft with demand consolidation,

highlighting the value of the air pooling concept in developing and selecting the aircraft technology.

Increasing the reservation time window and maximum allowed delay decreases rejection rates and increases the average load factor. For the synthetic network in this study, the analysis suggests no noticeable improvements beyond the maximum of 40-minute advance notice. That being said, when the maximum acceptable delay is long enough to allow the UAT operator to move the customers on the ground for demand consolidation and the UAT aircraft in the network to serve them, the UAT operator could immediately serve the requests with no advance notice required. Consequently, the maximum acceptable delay has a noticeable impact on the average load factor. However, the maximum acceptable delay cannot be increased to the point that it diminishes the travel time savings.

In the CLARPTW-SRE formulation, the passenger delay is modeled as a soft constraint, and therefore, the passenger's delay is not explicitly minimized. Consequently, even with aircraft utilization of 40%, the average delay is about 10 minutes per passenger. Furthermore, the results highlight how increasing the maximum of reservation time window from 1 minute to 60 minutes would decrease the average delay by almost 4 minutes, a 26% reduction in delay.

Lastly, it is worth noting that the sensitivity analyses are conducted with a fixed demand, which is not dependent on the service time. However, factors such as the detour factor, aerial speed, and access time directly impact the service time and trip time savings, which changes the UAT demand. Nonetheless, given the long distances covered in the synthetic network, the passengers would benefit from choosing UAT even under worst-case scenarios.

Chapter 10 Numerical Experiments: Chicago Network

10.1 Overview

This chapter implements the proposed operational policy with CLARPTW-SRE in the Chicago network. The UAT demand for this experiment is estimated based on the Transportation Network Providers (TNP) data in Chicago using a simplified rule-based mode choice model. Subsequently, the experiments are designed by randomly selecting a specified fraction of the eligible TNP demand over 5 replications. An in-depth analysis is provided for one day of UAT operation in Chicago, and a more concise analysis is provided for a week.

In the following sections, we first present the mode choice model used to estimate UAT demand. We subsequently introduce the TNP data for Chicago. Afterward, the experiment design is discussed, and ultimately, the numerical results for the one day and one week of UAT operations are reviewed.

10.2 Mode Choice Model

We use a simplified rule-based mode choice model to estimate the UAT demand in this dissertation. Multiple market studies project that the minimum distance for passenger UAM trips would be around 10 miles [9,19]. Meanwhile, Booz Allen Hamilton's market study [8] finds no significant demand for mandatory (i.e., work-related) trips that take less than 30 minutes on the ground. Furthermore, they assert that most of the UAM demand is captured for trips that are at least 45 minutes on the ground. Consequently, we qualify a trip for the UAT mode when the following rules are satisfied:

1. The distance between the origin and destination of the trip is at least 10 miles (i.e., $\Delta^{OD} = 10$ miles);

2. The trip would take at least 45 minutes using ground-based transportation.

Furthermore, Uber Elevate [23] assumes that a UAM trip should be at least 40% faster than the corresponding ground-based trip, while Porsche Consulting [6] suggests UAM needs to offer at least 20% travel time savings to be competitive with other modes. Therefore, we compare the travel time savings for the UAT demand in the experiments against these numbers.

Ultimately, a fraction of the qualified TNPs demand would use UAT as their mode of travel. In other words, the probability of choosing UAT is $\eta < 1$.

10.3 Chicago TNP Data

Trip data of Transportation Network Provider (TNP) (also known as ridesharing companies) for Chicago [144] includes more than 128.7M trips from 2018-11-01 to 2019-12-31. The dataset includes pick-up and drop-off census tract and their coordinates, trip travel time in seconds, trip distance in miles, and trip start and end time, among other attributes.

Chicago has approximately 800 census tracts, ranging from about 89,000 square feet (i.e., 0.003 square miles) to eight square miles [145]. These census tracts have a population of about 1200 households, or 2000-4000 people [146]. For privacy reasons, the available data is anonymized by projecting the origin and destination of trips to the centroid of their corresponding census tracts in addition to rounding the start and end time of the trips to its nearest 15 minutes. Therefore, the pick-up and drop-off locations and times cannot be known with precision beyond a 15-minute time interval and an 89,000 square foot (approximately 8,270 square meters) area [145]. Pick-up and drop-off census tracts are often left blank if the location is outside of Chicago [144]. Additionally, if either the pickup or drop-off census tract of a trip has fewer than three trips over the 15 minutes period, both census tracts are left blank in the dataset [147].

Trips with missing values for pickup and drop-off census tracts and coordinates, start and end time, trip time, and trip distance are excluded, which reduces the number of trips to nearly 86.4M. To cover the trips that might benefit from UAT, we consider trips that are at least 10 miles [9] and take at least 45 minutes on the ground [8]. As a result, the number of trips reduces to nearly 1.99M.

Like many real-world datasets, Chicago TNP data should be cleaned up to remove erroneous and missing inputs. As an example of invalid entries, 106 trips have zero travel distance. After filtering out the trips based on unreasonable trip distance, trip time, travel speed, and the difference between trip distance and geodesic distance, the number of trips drops to nearly 1.94M.

Table 10.1 summarizes travel time, trip distance, Geodesic distance, and speed of the qualified TNP demand for UAT service. The mean travel speed is 20.0 mph with a mean travel time of 57.6 minutes, respectively. Figure 10.1 depicts the travel time distribution over 4 ranges of geodesic distances. Similarly, Figure 10.2 and Figure 10.3 illustrate the travel speed and geodesic speed (i.e., speed over geodesic distance).

Table 10.1 Summary of qualified Chicago TNP trips for UAT service

	Geodesic Distance (mi)	Trip Distance (mi)	Travel Time (min)	Travel Speed (mph)
Min	10.0	10.1	45.0	5.6
25th percentile	13.5	16.8	49.3	16.9
Median	15.3	18.1	54.7	19.5
Mean	14.8	18.8	57.6	20.0
75th percentile	15.8	19.8	63.0	22.4
Max	28.8	41.6	120.0	51.2

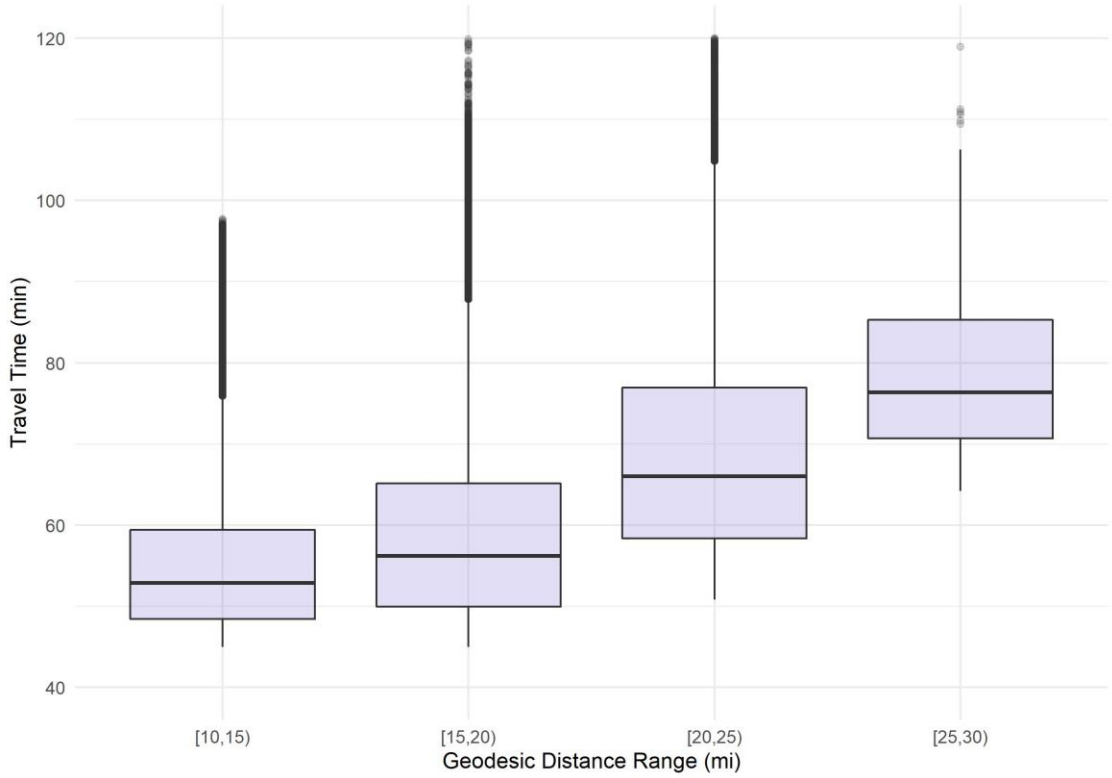

Figure 10.1 Boxplots of travel time over geodesic distance ranges

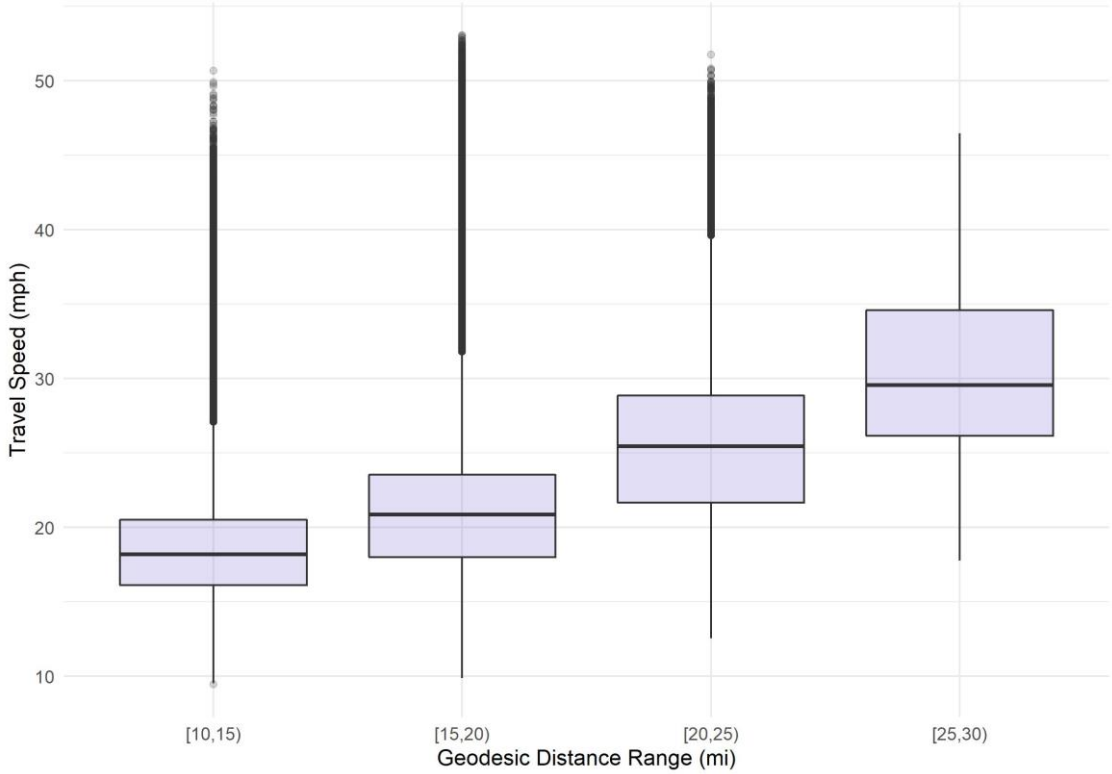

Figure 10.2 Boxplots of travel speed over geodesic distance ranges

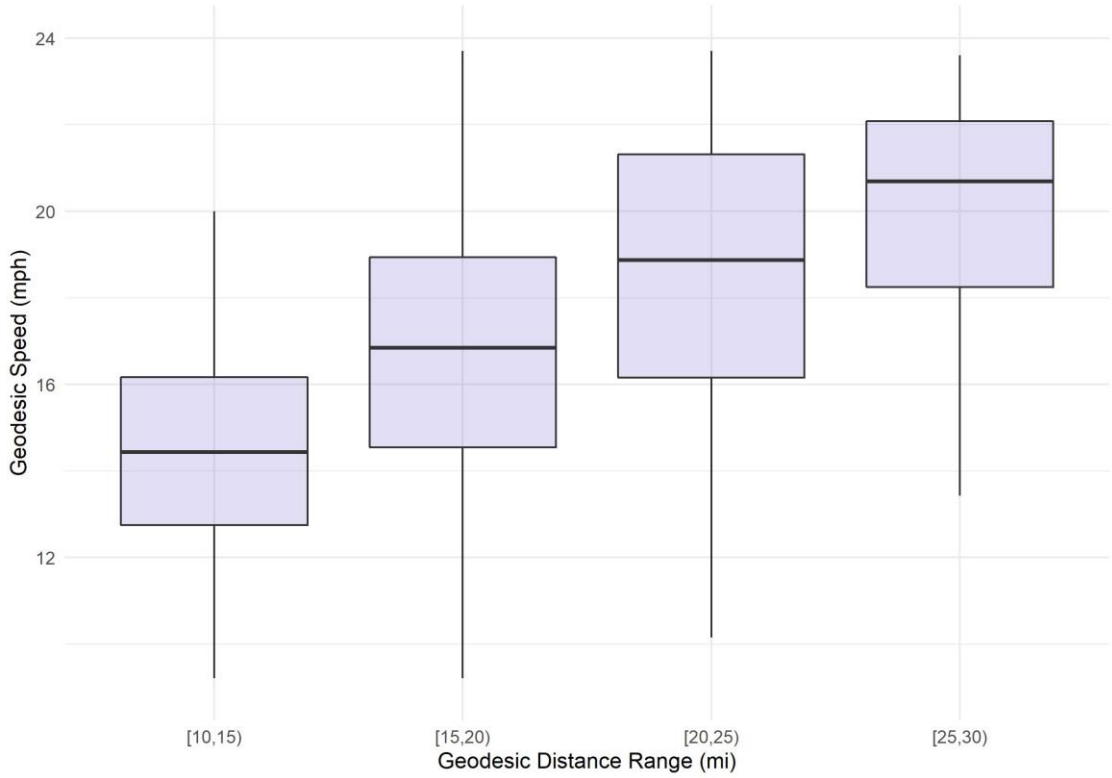

Figure 10.3 Boxplots of geodesic speed over geodesic distance ranges

Figure 10.4 illustrates the spatial distribution of the qualified TNP demand for UAT service. The grey polygons outline 801 census tracts in Chicago, while the lines show the trips between origin and destination census tracts. Thicker lines imply a higher number of trips. The majority of trips are between O'Hare International Airport, Midway International Airport, and downtown Chicago. Five and one census tracts, respectively, are the origins and destinations of 50% of the qualified TNP demand, while 50 and 25 census tracts account for 80% of the origins and destinations of qualified TNP demand.

Given the parameters assumed in Section 10.4.2, the desired trip time, where the requests are served immediately without ground transportation, and subsequently, the maximum travel time savings, are calculated for the TNP trips. Table 10.2 shows that for the nearly 1.94M TNP trips, the mean of maximum travel time saving is 37.7 minutes with a minimum of 24.7 minutes. The

mean of maximum travel time saving percentage is 66.5%, which is higher than 20% suggested by Porsche Consulting [6] and 40% suggested by Uber Elevate [23].

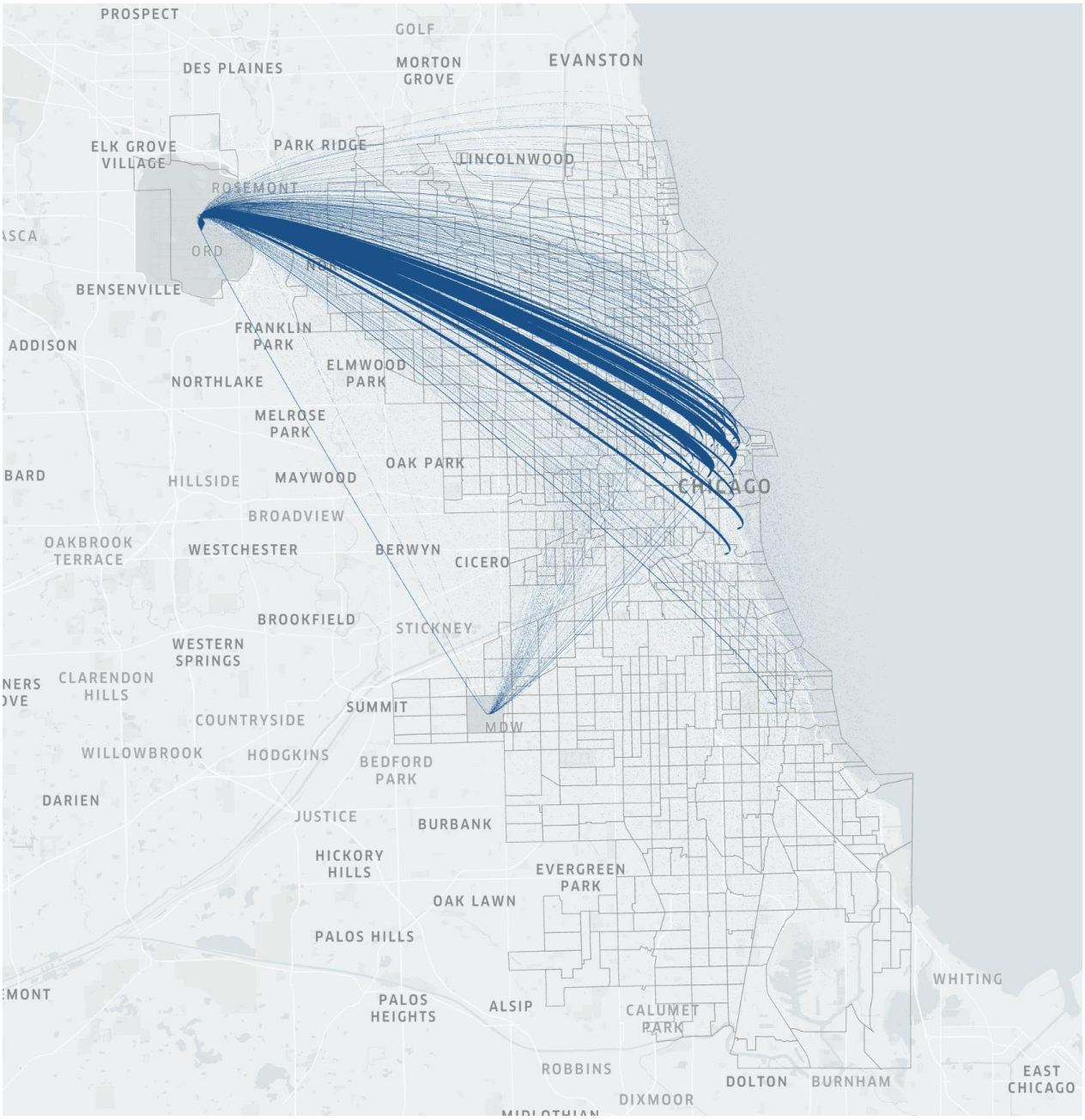

Figure 10.4 Spatial distribution of the qualified TNP demand for UAT service

The maximum of total travel time savings is nearly 1.25M hours over 14 months. The minimum wage for the City of Chicago, effective July 1st 2019 to July 1st 2020, is $13.00 per hour

[148]. Therefore, the value of maximum travel time savings is estimated at USD 16.2M under an unconstrained (i.e., best-case) scenario.

Table 10.2 Summary of maximum travel time saving for Chicago TNP trips for UAT service

	Travel Time Saving (min)	Travel Time Saving Percentage (%)
Min	24.7	54.9
25th percentile	30.5	61.9
Median	35.9	65.5
Mean	37.7	66.5
75th percentile	43.9	69.8
Max	98.6	82.6

Figure 10.5 and Figure 10.6 depict the distribution of travel time saving and travel time saving percentage under the best-case scenario. As expected, the mean of travel time saving increases with the distance between the origin and destination.

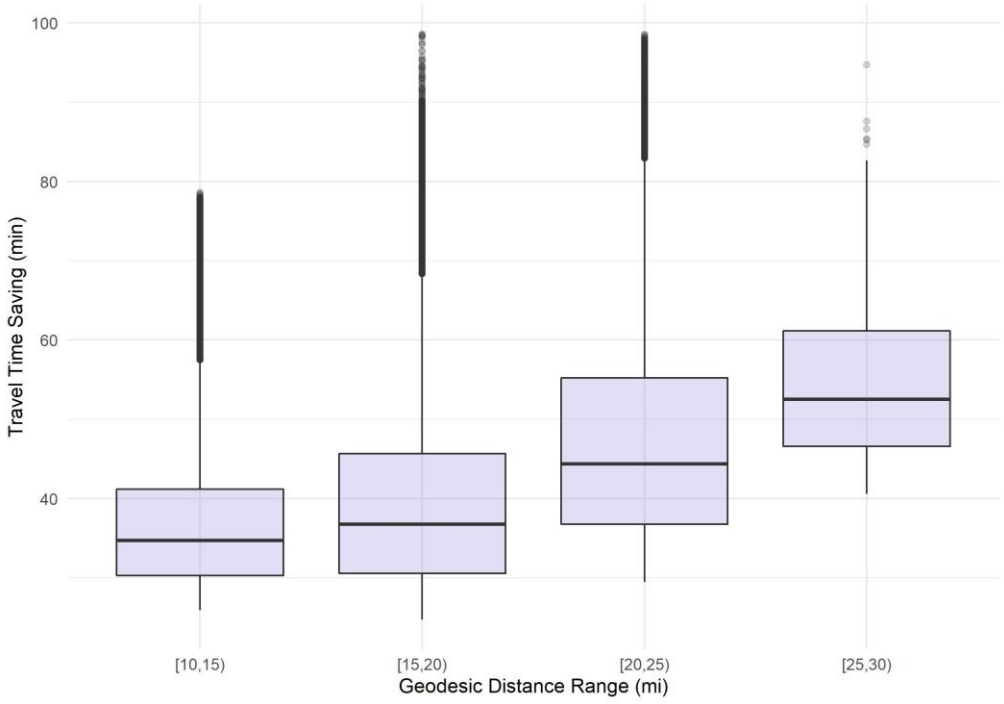

Figure 10.5 Maximum travel time savings over geodesic distance range

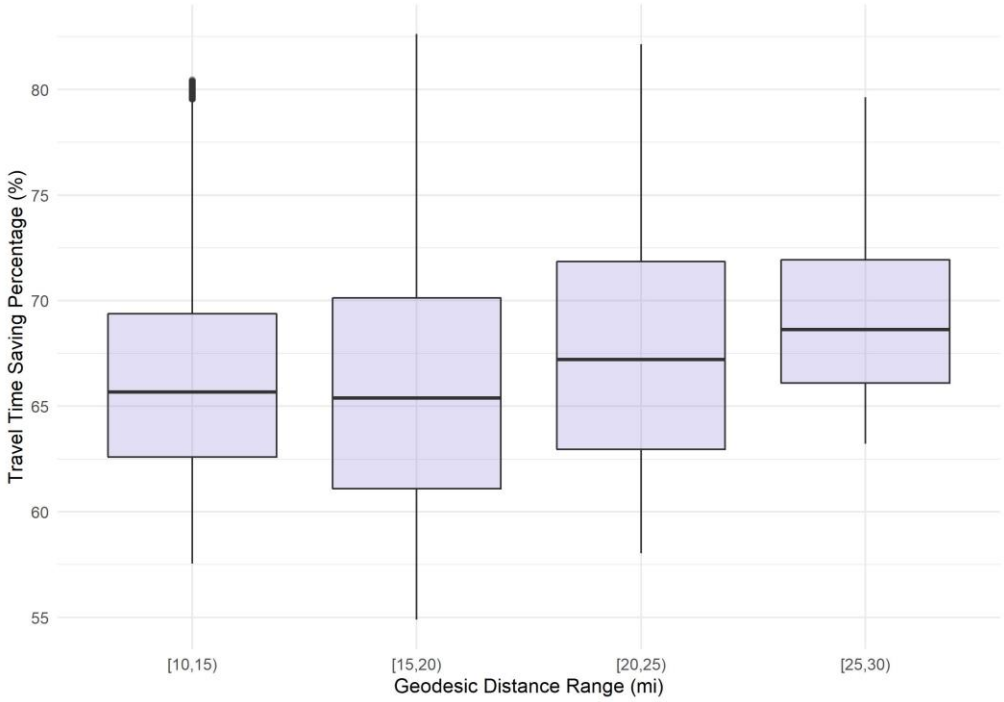

Figure 10.6 Maximum travel time savings percentage over geodesic distance range

10.4 Experiment Design

10.4.1 Simulation Design

The simulation for UAT operation in Chicago is *terminating* (also referred to as *transient* or *non-stationary*). In terminating simulations, the initial state of the system at time 0 is well specified, and the simulation ends either with stopping event v^{END} or at stopping time $\mathfrak{T}^{ENDSIM}$. Let $\mathfrak{T}^{STRTARV} = \min_{r \in \mathcal{R}^{ARV}} \tau_r^{ARV}$ and $\mathfrak{T}^{ENDARV} = \max_{r \in \mathcal{R}^{ARV}} \tau_r^{ARV}$ denote the start and end of the request arrival period. Subsequently, the simulation starts with the start of the request arrival period (i.e., $\mathfrak{T}^{STRTSIM} = \mathfrak{T}^{STRTARV}$) and ends when all the requests are served. Each experiment is replicated five times, where the stochasticity stems from the customer requests and the initial locations of the UAT aircraft.

The trip start time in Chicago TNP data is anonymized by rounding it to the nearest 15-minute intervals. Consequently, we estimate the original trip start time by using the uniform

distribution with the range of [0, 15]. The requested service time for request r, τ_r^{REQ}, is the same as the estimated trip start time. Let $\mathfrak{T}^{STRTREQ}$ and $\mathfrak{T}^{ENDREQ}$ denote the start and end of the requested time for service, which are specified by the UAT operator. To generate the requests, η% of the qualified UAT demand with $\tau_r^{REQ} \in [\mathfrak{T}^{STRTREQ}, \mathfrak{T}^{ENDREQ}]$ are randomly drawn.

The arrival time of request r, i.e., τ_r^{ARV}, is calculated as $\tau_r^{REQ} - T_r^{ADV}$, where T_r^{ADV} is randomly drawn from a uniform distribution with the mean of $\mathcal{T}^{ADV}/2$ and the range of $[0, \mathcal{T}^{ADV}]$. When request r arrives at time τ_r^{ARV}, its attributes are defined by the vector $\mathbb{A}_r^{REQ} = (\boldsymbol{O}_r, \boldsymbol{D}_r, \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_r^{DSRD}, q_r, \tau_r^{REQ})$, The origin, $\boldsymbol{O}_r$, and destination, $\boldsymbol{D}_r$, of r_r are the same as the centroid of their corresponding census tracts. We assume the network is nearly ubiquitous, with almost 800 UAT pads. As a result, the desired pick-up and drop-off UAT pads of request r coincide with their origin and destination. In other words, $\boldsymbol{S}_r^{DSRD} = \boldsymbol{O}_r$ and $\boldsymbol{E}_r^{DSRD} = \boldsymbol{D}_r$. Lastly, it is assumed that the group size of each request is 1.

Additionally, the fleet of UAT aircraft is homogenous, and therefore, their capacity and speed are denoted by Q and v^{AIR}, respectively. The state of UAT aircraft is presented by $S_t^{eVTOL} = (\varsigma_{kt}^{eVTOL}, \tau_{kt}^{AVL}, \boldsymbol{L}_{kt}^{AVL}, \mathcal{Q}_{kt})_{a_k \in \mathcal{K}}$ at $t = \mathfrak{T}^{STRTSIM}$(i.e., the beginning of the planning horizon). The initial location of $a_k \in \mathcal{K}$ (i.e., $\boldsymbol{L}_{kt}^{AVL}$ at time $t = \mathfrak{T}^{STRTSIM}$) is randomly drawn from the centroids of census tracts using a categorical distribution, a discrete probability distribution with the weights equal to the probability of a centroid being chosen as a destination from TNP data. Furthermore, all the aircraft are idle and available at the beginning of the planning horizon, and there is no incomplete flight leg on their itinerary. In other words, $\tau_{kt}^{AVL} = \mathfrak{T}^{STRTSIM}$, $\varsigma_{kt}^{eVTOL} = 0$, and $\mathcal{Q}_{kt} = \emptyset$ for $t = \mathfrak{T}^{STRTSIM}$.

10.4.2 Parameter Setting

In this section, we present the parameters used in the experiments. Table 10.3 summarizes the design parameters associated with UAT operations. $\alpha/\beta = 5$ in Chicago network implies that serving every request is profitable, and the requests will be rejected only if they are not feasible to be served. $\mathcal{T}^{ADV}$ is 30 minutes, suggesting that the requests for UAT service, at the earliest, could be placed 30 minutes ahead of the desired time. The maximum delay due to schedule delay and ground-based transportation is limited to 10 minutes, i.e., $\omega = 10$ minutes.

Table 10.3 Design parameters associated with UAT operation

Parameter	Symbol	Value	Unit
Maximum of the reservation time window	$\mathcal{T}^{ADV}$	30	minutes
Maximum acceptable delay	ω	10	minutes
Ratio of revenue per passenger mile to cost per mile	α/β	5	-
Ratio of the relocation cost to cost per mile	γ_1/β	0	-
Start of requested time for service	$\mathfrak{T}^{STRTREQ}$	6:00	AM
End of requested time for service	$\mathfrak{T}^{ENDREQ}$	6:59	PM

Table 10.4 summarizes the parameters associated with the ingress and egress of the customers. The travel times between the UAT pads, which are located at the centroid of the census tracts, are calculated using the Chicago TNPs dataset. The data shows that the average speed over Euclidean distances smaller than 2 miles is 8.3 mph. Consequently, for the missing values of travel times, we choose the driving speed, i.e., v^{DRIVE}, of 8.0 mph. We further assume that the maximum acceptable delay for the Chicago UAT service, i.e., ω, is 10 minutes. $v^{DRIVE} = 8.0$ mph and $\omega = 10$ minutes suggest that the maximum accessible distance on the ground from the origin or destination of the request is nearly 1.3 miles. Consequently, we set $\Delta^{ACCESS} = 2$ miles. The remaining parameters are similar to the ones specified in Section 9.2.3.

Table 10.4 Parameters associated with the ingress and egress of the customers

Parameter	Symbol	Value	Unit
Maximum access distance	Δ^{ACCESS}	2	miles
Euclidean driving speed in the downtown area	v^{DRIVE}	8	mph
Maximum walking distance	Δ^{WALK}	0.25	miles
Walking speed	v^{WALK}	3	mph
Duration from ground transportation area to the departure gate	T_{ri}^{DGATE}	3	minutes
Duration from arrival gate to the ground transportation area	T_{ri}^{AGATE}	2	minutes

Table 10.5 Parameters associated with flight operation

Parameter	Symbol	Value	Unit
Number of UAT aircraft	K	100	-
Minimum of Euclidean distance for repositioning flight	Δ^{EMPTY}	0.5	miles
Aerial speed	v^{AIR}	150	mph
Boarding duration	T^{BOARD}	3	minutes
Deboarding duration	$T^{DEBOARD}$	2	minutes
Departure clearance	$T^{TAKEOFF}$	0.5	minutes
Landing clearance	$T^{LANDING}$	0.5	minutes
Hover ascend	T^{ASCEND}	0.75	minutes
Hover descend	$T^{DESCEND}$	0.75	minutes
Detour factor	ϵ	0.1	-

The assumptions regarding the flight operation are presented in Table 10.5. The UAT operator has an aerial fleet with the size of $K = 100$. The minimum Euclidean distance to justify a UAT repositioning flight, i.e., Δ^{EMPTY}, is set to 0.5 miles. With the average driving speed of 8 mph, ground-based transportation within 0.5 miles of both the origin and destination would take at most 7.5 minutes, which is below the maximum acceptable delay of 10 minutes. This implies that while empty repositioning within 0.5 miles of a desired UAT pad is not allowed, the passengers could access these pads on the ground within the acceptable delay. The remaining parameters are similar to the ones specified in Section 9.2.3.

Consequently, the turnaround time, including boarding, deboarding, and take-off and landing clearance, is 6 minutes, consistent with the turnaround time estimated by Joby Aviation [18] (see Table 2.6). The overhead time of serving a flight leg, either empty or revenue-generating, includes hover ascend and descend and ATC clearance for the take-off and after the landing, which amounts to 2.5 minutes. If the flight leg serves passengers, an additional 5 minutes will be added to the flight service time. As a result, the overhead time of empty and revenue-generating flight legs are 2.5 and 7.5 minutes, respectively.

As shown in Table 10.1, the average OD distance and travel time of qualified TNP data are 14.8 miles and 57.6 minutes, respectively. Therefore, the average aerial distance is 16.3 miles (i.e., 1.1× 14.8). Consequently, the average time for serving a revenue-generating flight leg is 14.0 minutes (i.e. $\frac{16.3}{150} \times 60 + 7.5$).

As a result, the maximum service rate is 4.28 revenue-generating flights per hour per aircraft. This extreme value corresponds to cases where the request's arrival rate is so high, or the requests are so close to each other that the empty distance from the destination of one flight to the origin of the next flight origin is close to zero in expectation. Without job rejection and flight-sharing, ρ should be below 1 for the queue to not grow indefinitely, and therefore, for the system to be stable. With $c = K = 100$ and $\mu = 4.28$ flights per hour per aircraft, the system can accommodate $\lambda = 342.4$ requests for flight per hour (i.e., $T^{INT} = 10.5$ seconds) to achieve 80% utilization ($\rho = 0.8$).

Furthermore, the average of 14.0 minutes for serving a revenue-generating flight would translate to the average trip time of 19.0 minutes (i.e., 14.0 + 3+ 2) for each passenger if there were no wait time for the aerial service, and the requests were served without any ground-based

transportation. Considering a maximum delay of $\omega = 10$ minutes (in Figure 5.1), the minimum and maximum mean trip times for passengers are 19.0 and 29.0 minutes, respectively. These numbers correspond to average travel time savings of 27.7 minutes (51.1%) to 37.7 minutes (66.5%) compared to driving on the ground. Consequently, the average travel time saving is well above the 40% criterion suggested by Uber Elevate for the UAT service mode choice model [23].

10.4.3 UAT Demand

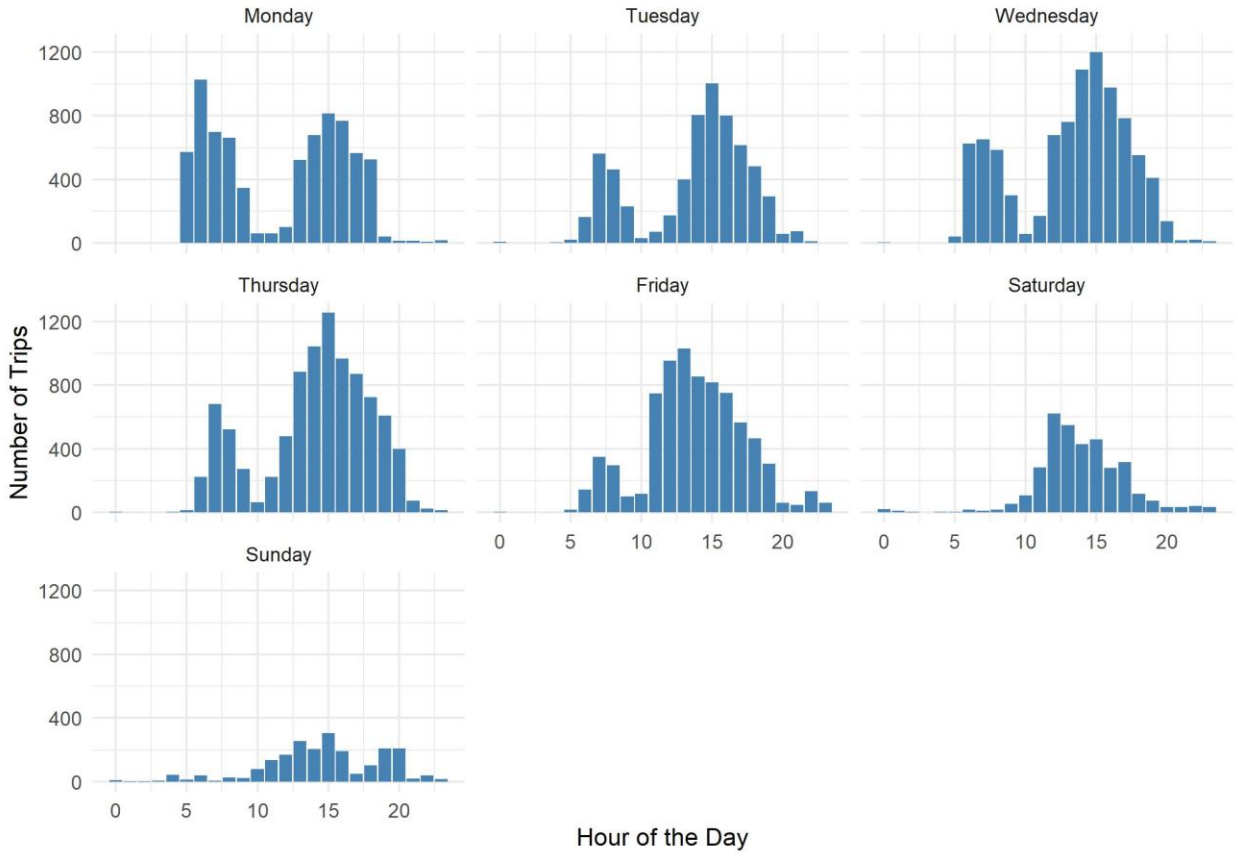

Figure 10.7 Number of trips per trip start hour between Sep. 23, 2019 and Sep. 29, 2019

To examine the UAT service in a more realistic setting, we use the qualified TNP demand for UAT service over one week, from Monday, Sep. 23 to Sunday, Sep. 29, 2019[6]. Figure 10.7

[6] A curious reader may wonder why we did not use the most recent demand from 2020. The reason is that the world basically stopped for a period of time in 2020 due to the COVID-19 pandemic, which in turn, had a significant impact on ridesharing operations and demand.

demonstrates the temporal demand during the week of Sep. 23rd. We further assume that the UAT service offers rides to requests that arrive between 6:00 and 18:59 since about 90% of the requests arrive within this period. Over this period, the average number of requests per hour is 451. However, during evening peak hours, the demand could reach 1250 requests per hour, corresponding to the mean interarrival time of approximately 3 seconds.

Table 10.6 Average hourly qualified TNP demand over weekday and weekend of the week of Sep. 23rd, 2019 from 6:00 AM to 6:59 PM

	Trip Start Hour												
	6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Weekday													
Average Demand (hr^{-1})	431	584	498	248	68	246	466	705	886	1020	870	696	543
Interarrival Time (sec)	8.3	6.2	7.2	14.5	52.9	14.6	7.7	5.1	4.1	3.5	4.1	5.2	6.6
Weekend													
Average Demand (hr^{-1})	27.5	7.5	22.5	44	98	206	394	391	313	378	229	177	108
Interarrival Time (sec)	130.9	480	160	82.8	36.9	17.5	9.1	9.2	11.5	9.5	15.8	20.3	33.3

Table 10.6 summarizes the average hourly demand over the weekday and weekend of Sep. 23rd. We use $\eta = 40\%$ and 60% of the qualified TNP demand for the weekly experiment, and we randomly select the requests given the desired fraction of demand. Table 10.7 presents the total demand for the given fraction of qualified TNP demand for the week of Sep. 23rd, 2019 from 6:00 AM to 6:59 PM. We further examine the UAT operation for Monday, Sep 23rd, 2019, with 60% of the qualified UAT demand. The corresponding hourly demand and interarrival time are summarized in Table 10.8. The minimum and maximum of the interarrival time for Sep. 23rd is 5.9 and 95 seconds, respectively, corresponding to 6 AM and 10 AM.

Table 10.7 Qualified TNP demand for the week of Sep. 23rd, 2019 from 6:00 AM to 6:59 PM

	Qualified TNP Demand		
Day of Week	**100%**	**60%**	**40%**
Monday	6,810	4,086	2,724
Tuesday	5,774	3,464	2,310
Wednesday	8,250	4,950	3,300
Thursday	8,152	4,891	3,261
Friday	7,320	4,392	2,928
Saturday	3,184	1,910	1,274
Sunday	1,601	961	64

Table 10.8 Hourly UAT demand and interarrival time (estimated as $\eta = 60\%$ of the qualified Chicago TNPs demand) for Monday, September 23rd, 2019

	Trip Requested Hour												
	6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Hourly Demand	608	410	396	208	38	39	61	314	404	486	467	344	311
Interarrival Time (sec)	5.9	8.8	9.1	17.3	95.2	92.3	58.8	11.5	8.9	7.4	7.7	10.5	11.6

10.5 Case Study: One Day

The following sections examine the performance measures associated with UAT operation on Monday, September 23rd, 2019, with $\eta = 60\%$ of the qualified TNP demand and four UAT aircraft capacity, i.e., $Q = 1$, 2, 3, and 4. More specifically, we review the runtime and gap to evaluate if the solution framework could be implemented in real time. Furthermore, we present the performance measures associated with UAT operator revenue and costs in addition to user inconvenience and trip delay. Lastly, we review the travel time savings resulting from choosing UAT over ground-based ridesharing service.

10.5.1 Problem Size, Runtime, and Gap

For each replication, we calculate the mean and standard deviation of the number of flight legs, requests, and arcs at each decision epoch. Subsequently, we estimate the mean of the mean and the standard deviation for all replications. Table 10.9 summarizes the estimated mean of the mean and standard deviation of the number of flight legs, requests, and arcs at each decision epoch.

Table 10.9 Estimated mean of the mean and standard deviation of the number of flight legs, requests, and arcs at each decision epoch

	Flight Legs			Requests				Arcs		
Q	$\widetilde{\mathcal{N}}_t^{\mathrm{LEG}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{DSRD}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{CNCT}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{REQ}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{UNASGN}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{FLXSTRT}}$	$\widetilde{\mathcal{N}}_t^{\mathrm{FXDSTRT}}$	$\widetilde{\mathcal{A}}_t^{INIT}$	$\widetilde{\mathcal{A}}_t^{SEQ}$	$\widetilde{\mathcal{A}}_t^{ALCT}$
Estimated Mean of Mean Values (of Each Replication)										
1	232	125	107	130	6	106	18	9,609	6,232	2,600
2	216	111	105	131	6	93	32	10,769	4,612	2,302
3	214	109	105	132	6	92	34	11,298	4,408	2,319
4	211	109	102	132	6	92	34	11,546	4,339	2,313
Estimated Mean of Standard Deviation (of Each Replication)										
1	123	56	72	58	3	49	9	4,570	4,312	1,678
2	122	52	74	64	3	45	20	5,454	3,360	1,759
3	123	52	75	67	3	45	23	5,831	3,286	1,868
4	122	52	74	67	3	45	23	5,993	3,245	1,878

Figure 10.8(a)-(d) depict the emipirical Cumulitavie Density Function (eCDF) of the MIP gap of the solution at the end of the decision epochs for $Q = 1$ to $Q = 4$, respectively. Each replication is illustrated by one graph. While for $Q = 1$ nearly all the solutions have a gap under 3%, for $Q = 2$, 3, and 4, the gap is under 2% for approximately all the solutions. Moreover, for $Q = 1$, around 85% of MIPs are solved to optimality, while this number decreases to 40% and 20% for $Q = 2$ and $Q = 4$, respectively.

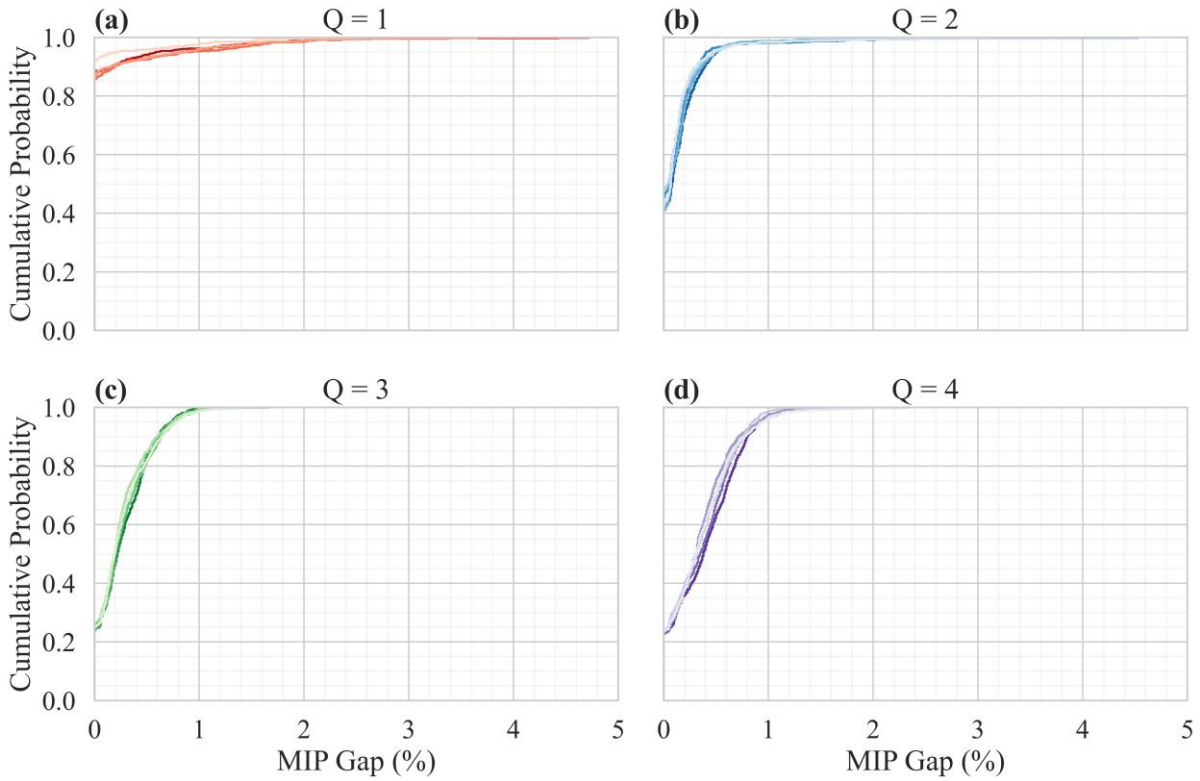

Figure 10.8 eCDF of MIP Gap for 5 replications and *Q* of (a) 1, (b) 2, (c) 3, and (d) 4

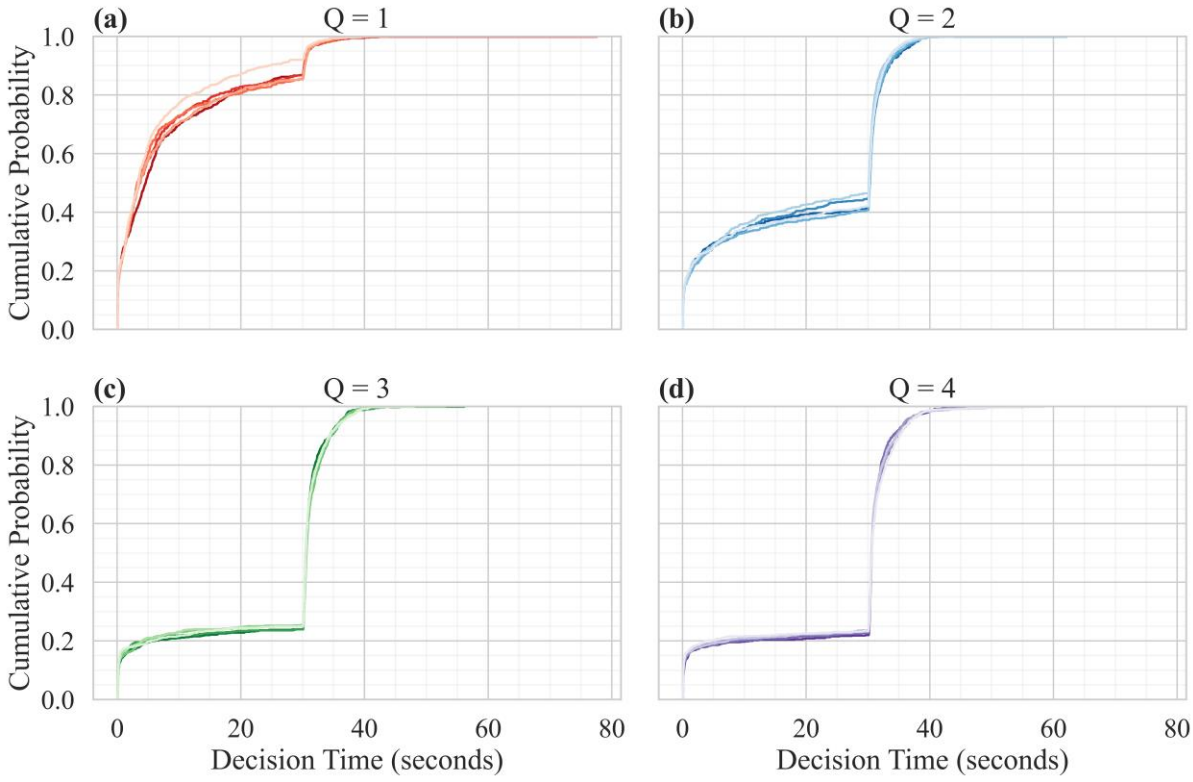

Figure 10.9 eCDF of decision time in seconds for 5 replications and *Q* of (a) 1, (b) 2, (c) 3, and (d) 4

Figure 10.9(a)-(d) depict the eCDF of corresponding decision time, i.e., T_e^{EPOCH}, associated with the decision epochs for $Q = 1$ to $Q = 4$, respectively. The nearly vertical slopes in the plots seen around the 30-second mark are due to the initial time limit of 30 seconds. The longer vertical slopes for $Q \geq 2$ compared to $Q = 1$ suggest that more problems are terminated in these cases at 30 seconds. Figure 10.9 suggests that the duration of almost all decision epochs is below the 1-minute re-optimization interval (i.e., Δ^{UPDATE}).

Table 10.10 presents the average simulation time over an 11-hour planning horizon for the 5 replications. Additionally, the worst cases of MIP solution time, decision time, and MIP gap over all decision epochs and replications are reported. Table 10.10 shows that the worst case of the MIP solution time is below the 1-minute re-optimization interval (i.e., Δ^{UPDATE}) for $Q = 3$ and 4 and almost 1 minute for $Q = 2$. The worst decision time, i.e., $\max T_e^{EPOCH}$, which is copmrised of constructing the network and MIP solution time, is under 60 seconds for $Q = 3$, while for $Q = 2$ and 4, the decision time slightly violates the 1-minute threshold. Therefore, given the assumed demand size and parameter settings, the solution framework has the potential for real-time implementation.

Table 10.10 Average simulation time for 5 replications and worst MIP solution time, worst decision time, and worst MIP gap over all decision epochs and 5 replications

Aircraft Capacity (Q)	Average Simulation Time (minutes)	Worst MIP Solution Time (seconds)	Worst Decision Time (seconds)	Worst MIP Gap (%)
1	136.3	77.6	80.6	4.7
2	276.9	62.2	62.8	4.5
3	327.8	56.1	56.9	1.7
4	333.1	58.3	64.2	2.3

10.5.2 UAT Operator Costs

The aerial mileage, either revenue-generating or deadhead, is associated with a cost. A high rate of air pooling would reduce the revenue-generating mileage. Moreover, UAT aircraft utilization indicates how long the aircraft are busy, either serving empty or revenue-generating flight leg. Consequently, the following sections present the performance measures associated with the revenue mileage, total aerial mileage, air pooling, utilization, and service time.

10.5.2.1 Utilization

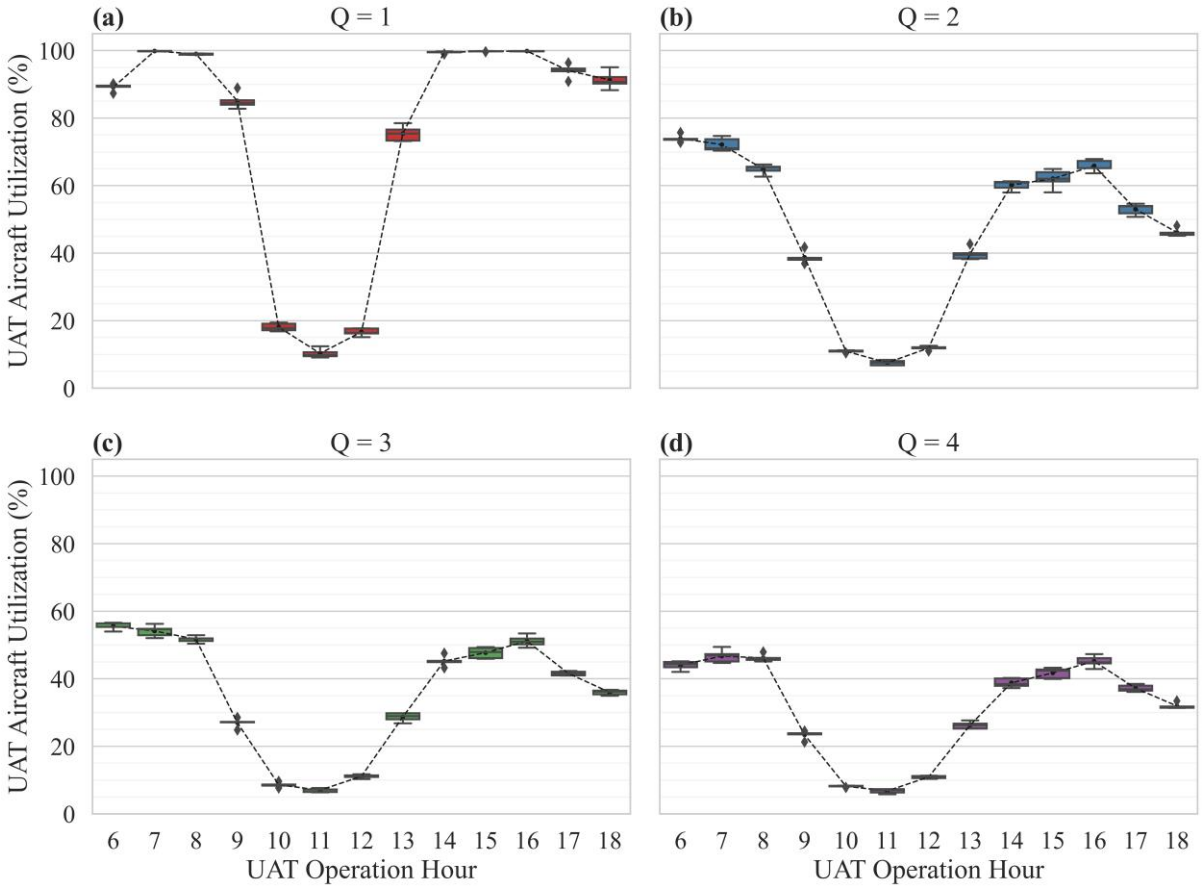

Figure 10.10 Distribution of UAT aircraft hourly utilization over 5 replications with Q of (a) 1, (b) 2, (c) 3, and (d) 4

Figure 10.10 depicts the distribution of UAT aircraft hourly utilization for $Q = 1$ to 4. The utilization is under 20% over off-peak hours, i.e., 10 AM to 12 PM. With $Q = 1$, the aircraft

utilization is almost 100% during the morning and evening peak hours. However, with $Q = 2$, the peak-hour utilization is less than 80%, while with $Q = 3$ and 4, it becomes less than 60%. Table 10.11 specifies that the UAT aircraft utilization over the planning horizon reduces from 73.1% for $Q = 1$ to 32.1% for $Q = 4$.

Table 10.11 UAT aircraft utilization over the planning horizon

Q	Aircraft Utilization (%)
1	73.0
2	46.6
3	35.9
4	31.8

10.5.2.2 Empty Mileage

Table 10.12 presents the performance measures associated with empty repositioning flight legs for CLARPTW-SRE and CLARPTW, models with and without short repositioning elimination, respectively. Without short repositioning elimination (SRE), between 7.1 and 9.5% of the flight legs are over short distances of shorter than 0.5 miles. For CLARPTW, the minimum of empty flight mileage is 0.15 miles. For CLARPTW-SRE, the minimum of empty flight mileage is 0.55 ($= 1.1 \times 0.5$) miles, which is consistent with the detour factor of 0.1 for the aerial legs and the radius of 0.5 miles for short repositioning elimination. The empty flight legs could be as long as 29 miles, which is the maximum of geodesic distance in the Chigaco network (see Table 10.1).

Moreover, for CLARPTW-SRE, the mean of empty flight legs is 10.3, 8.6, 7.0, and 6.1 miles for $Q = 1$ to 4, respectively. There are two possible explanations for the reduction in empty flight mileage as the UAT aircraft capacity increases. First, during air pooling, when multiple requests are consolidated and served by one flight leg, the corresponding revenue-generating flight leg is chosen so that, ceteris paribus, it minimizes the revenue-generating and empty mileage. With

$Q = 1$, the passengers could be relocated within $\Delta^{EMPTY} = 0.5$ to eliminate short repositioning legs, while with $Q \geq 2$, the passengers practically could be moved on the ground within $\Delta^{ACCESS} =$ 2 miles to reach another flight leg or UAT aircraft. As a result, given the 2-mile radius for access distance and detour factor of 0.1, the empty aerial mileage could decrease by a maximum of 2.2 miles. Second, higher capacity leads to lower utilization. Therefore, the UAT aircraft could be utilized more efficiently, which in turn could reduce the empty repositioning mileage.

Table 10.12 Estimated mean of performance measures associated with empty repositioning flight legs

Q	Mean of Empty Flight Mileage (mi)	Minimum of Empty Flight Mileage (mi)	Maximum of Empty Flight Mileage (mi)	Percentage of Empty Flights to Revenue Flights (%)	Percentage of Empty to Revenue Mileage (%)	Percentage of Empty Flight Legs within Δ^{EMPTY} (%)
CLARPTW-SRE						
1	10.35	0.56	28.33	47.0	29.9	0
2	8.70	0.55	27.17	40.0	22.0	0
3	7.11	0.55	27.37	39.7	17.9	0
4	6.24	0.55	28.29	41.8	16.6	0
CLARPTW						
1	9.29	0.15	28.77	52.9	30.2	7.1
2	7.58	0.15	27.97	46.8	22.3	9.1
3	6.19	0.15	26.48	46.5	18.2	9.5
4	5.62	0.16	26.58	48.7	17.4	9.1

Furthermore, Table 10.12 suggests that for CLARPTW-SRE, the percentage of empty to revenue mileage reduces from 30% for $Q = 1$ to 16% for $Q = 4$, while the percentage of empty to revenue-generating flight legs is in the range of 40 to 47.

A connecting flight leg is shorter than the combination of original empty and revenue flight legs. Furthermore, they offer shorter service time resulting from fewer flights and the consequent overhead time. Therefore, when a connecting request is within reach of passengers, it is more likely

to be included in the optimal solution. Table 10.13 evaluates the performance measure associated with the connecting legs, which are conducted to eliminate the short repositioning flight legs. For Q =1, nearly 15% of the revenue flights are connecting legs. For $Q \geq 2$, connecting flight legs are about 21-23% of all flight legs since there is a higher chance of reaching a connecting flight leg when requests are relocated within $\Delta^{ACCESS} = 2$ miles for $Q \geq 2$ compared to $\Delta^{EMPTY} = 0.5$ miles for $Q = 1$.

Table 10.13 Estimated mean of performance measures associated with connecting flight legs

Q	Percentage of Connecting Flights to Revenue Flights (%)	Percentage of Connecting Flights with the Desired Origin (%)	Percentage of Connecting Flights with the Desired Destination (%)	Percentage of Connecting Flights with Undesired Origin and Destination (%)
1	14.6	14.5	85.5	0.0
2	23.4	29.8	70.2	0.0
3	23.1	25.6	74.4	0.0
4	21.0	25.3	74.7	0.1

10.5.2.3 Revenue Mileage

Table 10.14 summarizes the performance measures associated with revenue flights, revenue mileage, and total mileage (i.e., the summation of revenue mileage and empty mileage). The mean of revenue-generating flight legs is in the range of 15.7 to 16.2 miles over various aircraft capacities. As the capacity increases, the mean of revenue mileage decreases slightly. The reasons are twofold. First, under $Q = 1$, not all the requests could be served. As a result, the shorter trips, which are associated with less revenue, are more likely to get rejected. Second, during air pooling, when one flight leg gets selected to serve multiple requests, a flight leg resulting in a lower mileage has a higher chance of being selected.

Furthermore, with $Q = 1$, serving one revenue-flight would require an average of 21.1 miles, while for $Q = 4$, this value is reduced to 18.3 miles, primarily due to the decrease in empty

mileage. Moreover, air pooling leads to a shorter aerial mileage per served request. Consequently, the mean of total aerial mileage for serving one request reduces nearly 66%, from 21.1 miles for $Q = 1$ to 7.0 miles for $Q = 4$.

Table 10.14 The performance measures associated with revenue flights, revenue mileage, and total mileage for $Q = 1, 2, 3,$ and 4

Q	Mean of Revenue Mileage (mi)	Mean of Total Aerial Mileage per Revenue Flight (mi)	Mean of Total Aerial Mileage per Served Request (mi)	Ratio of Revenue Flights to Served Requests (%)
1	16.2	21.1	21.1	100
2	15.8	19.3	10.8	56.2
3	15.8	18.6	8.1	43.4
4	15.7	18.3	7.0	38.1

10.5.2.4 Air Pooling

Table 10.15 specifies the estimated mean of performance measures related to UAT aircraft load factor over 5 replications. The average load factor is 89%, 77%, 65% for $Q = 2$, 3, and 4, respectively. With $Q \geq 2$, only 22-30% of the flight legs serve only one request. Moreover, nearly 78%, 57%, and 40% of the flight legs are filled to capacity with $Q = 2$, 3, and 4, respectively. Antcliff et al. [100] project that, in the long term, over 85% of the flights of a 2-seat aircraft have 2 passengers, which is comparable to 78% for UAT service using Chicago TNP demand.

Table 10.15 Estimated mean of performance measures related to UAT aircraft load factor over for 5 replications over $Q = 1, 2, 3,$ and 4

Q	Average Load Factor (%)	Average Number of Requests per Flight	Percentage of Flights with 1 Request	Percentage of Flights with 2 Requests	Percentage of Flights with 3 Requests	Percentage of Flights with 4 Requests
1	100	1.0	100	0	0	0
2	89.0	1.8	22.0	78.0	0	0
3	76.8	2.3	26.6	16.5	56.9	0
4	65.4	2.6	30.6	16.7	13.1	39.6

10.5.2.5 Service Time

Table 10.16 presents the estimated mean of the performance measures associated with aerial service time over 5 replications. The results suggest that serving a revenue flight leg on average would take around 14.0 minutes, regardless of the aircraft capacity. However, the service time of empty flight legs decreases from 6.6 minutes with $Q = 1$ to 4.9 minutes with $Q = 4$, due to the reduction in empty mileage as shown in Table 10.12. The total aerial service time, i.e., the total time an aircraft is in use, per served revenue flight is around 16-17 minutes. Taking advantage of air pooling reduces the total aerial service time per served request by 64%, from 17.1 minutes with $Q = 1$ to 6.1 minutes with $Q = 4$.

Table 10.16 Estimated mean of performance measures associated with aerial service time

Q	Mean of Aerial Service Time of Revenue Flight (min)	Mean of Aerial Service Time of Empty Flight (min)	Mean of Total Aerial Service Time per Revenue Flight (min)	Mean of Total Aerial Service Time per Served Request (min)
1	14.0	6.6	17.1	17.1
2	13.8	6.0	16.2	9.1
3	13.8	5.3	15.9	6.9
4	13.8	5.0	15.9	6.0

10.5.2.6 Estimated Cost per Mile (β)

Table 10.17 summarizes the estimated value of cost per mile, i.e., β, for the various projected unit of costs. The values range from $1.15 per mile for long-term operations estimated by Uber Elevate [28] to $12.7 per mile of piloted operation projected by Booz Allen Hamilton [8].

10.5.3 UAT Operator Revenue and Estimated Price

To evaluate the UAT operator revenue, we use three measures. Available seat miles (ASM), total passenger revenue (PR), and passenger revenue per available seat mile (PRASM). ASM represents the revenue-generating capacity of the fleet, while PRASM represents the

passenger revenue earned per seat (either empty or full) miles flown by an aircraft. Since the network is ubiquitous and all the flown mileage could potentially earn revenue, we calculate the ASM for all the conducted flights, either empty or revenue-generating. To exclude the impacts of the value of α (i.e., revenue per mile), we present PR and PRASM per α, i.e., PR/α and PRASM/α, respectively.

Table 10.17 Estimated value of cost per mile for the various projected units of cost

Company	Unit of Cost	Value	Capacity	Load Factor (%)	β ($/mi)
McKinsey & Company [10]	cost per seat-mile	0.5-2.5	4	-	2-10
Uber Elevate [28]	cost per passenger-mile	5.73, 1.84, 0.44	4	65.6	15.0, 4.8, 1.15
Booz Allen Hamilton [8]	cost per passenger-mile	9.5[1]	1	100	9.5
Booz Allen Hamilton [8]	cost per passenger-mile	7.0[1]	2	89.1	12.5
Booz Allen Hamilton [8]	cost per passenger-mile	5.5[1]	3	76.9	12.7
Booz Allen Hamilton [8]	cost per passenger-mile	4.75[1]	4	65.6	12.5
Porsche Consulting [6]	cost per mile	2.9	-	-	2.9
Joby Aviation [5]	cost per available seat-mile	0.95	4	-	3.8
Joby Aviation [66]	cost per available seat-mile	0.64[2]	4	-	2.56

Notes: [1] Estimated based on piloted operations, [2] 22 cents per available seat-mile cost of pilots is deducted for autonomous operations

Table 10.18 presents performance measures associated with passenger revenue, where α denotes the revenue (i.e., price) per mile. The total aerial milage decreases around 61% from $Q = 1$ to 4, regardless of the 14% decrease in the rejection rate (see Table 10.19). The total passenger revenue increases by nearly 16.6% with air pooling as more passengers are served. However, passenger revenue per available seat mile (PRASM) decreases from 0.701α to 0.526α for $Q = 1$

to $Q = 4$. Joby Aviation [66] estimates the value of 3 for α with $Q = 4$, resulting to PRASM = \$1.57 based on our analysis compared to PRASM = \$1.73 projected by Joby.

Table 10.18 Estimated mean of performance measures associated with passenger revenue

Q	Total Aerial Mileage (mi)	Available Seat Mile (ASM)	Total Passenger Revenue (PR/α)	Passenger Revenue per Available Seat Mile (PRASM/α)
1	73,516	73,516	51,508	0.701
2	43,845	87,691	59,501	0.679
3	32,952	98,855	60,132	0.608
4	28,554	114,217	60,138	0.527

Porsche Consulting [6] assumes a price between \$8 and \$18 per minute for the on-demand air taxi service. Given the total flight time of 8 minutes (i.e., service time minus 6 minutes of turnaround time), UAT operation in Chicago would cost between \$64 and \$144. On the other hand, Booz Allen Hamilton's market study sussgets a passenger price of nearly \$2.50-\$2.85 per mile, leading to \$45-\$51 given average total mileage of 18 miles for $Q = 4$. These estimated price ranges per trip are well beyond the \$25 per trip for at-scale operation estimated by McKinsey & Company [10].

10.5.4 User Experience and Level of Service

When using UAT as a mode of transportation, the customer request could get rejected, the accepted passengers may incur a delay, and they might have to take a flight from a location that is not their desired pick-up or drop-off UAT pad, all causing an inconvenience to the user. Furthermore, the primary advantage of UAT compared to TNP is the travel time savings, particularly over longer distances. Consequently, the following sections present the performance measures associated with user experience.

10.5.4.1 Rejections and Trip Delay

Table 10.19 summarizes the performance measures associated with served and rejected requests. The percentage of rejected requests decreases from 14.6% to 1.1% by increasing Q from 1 to 2, and it becomes almost 0% for $Q \geq 3$. Furthermore, the revenue structure of the objective function favors longer trips. Comparing average OD distances for served and rejected requests verifies that rejected requests have shorter OD distances.

Table 10.19 Estimated mean of performance measures associated with served and rejected requests

Q	Number of Served Requests	Percentage of Rejected Requests (%)	Average OD Distance of Served Requests (miles)	Average OD Distance of Rejected Requests (miles)
1	3,485	14.7	14.8	14.3
2	4,041	1.1	14.7	14.0
3	4,087	0.0	14.7	15.0
4	4,087	0.0	14.7	N\A

Table 10.20 Estimated mean of averages of trip delay, percentage of trip delay, and total trip time per request over 5 replications

Q	Average Trip Delay (minutes)	Average of Percentage of Trip Delay to Total Travel Time (%)	Average Trip Time (minutes)
1	7.7	27.6	26.7
2	6.7	24.8	25.7
3	6.9	25.5	25.9
4	6.9	25.5	25.9

Table 10.20 summarizes the estimated mean of averages of trip delay, percentage of trip delay, and total trip time per request over 5 replications. The average trip delay decrease with $Q \geq 2$ compared to $Q = 1$, since the system becomes less busy. Nonetheless, the higher capacity implies more potential delays resulting from demand consolidation and longer ground transportation. The trip delay is in the range of 25-28% of the total trip time. Lastly, an average

trip would take approximately 26 minutes in the network and is not very sensitive to the capacity since the maximum acceptable trip delay is constrained, and the delay is not explicitly minimized in the objective function.

Figure 10.11 provides more detail on the temporal distribution of the rejected requests. Most requests are rejected from 6 AM to 7 AM, which coincides with the highest hourly demand of the planning day, as seen in Table 10.8. However, the initial location of the UAT aircraft could be a contributing factor as well.

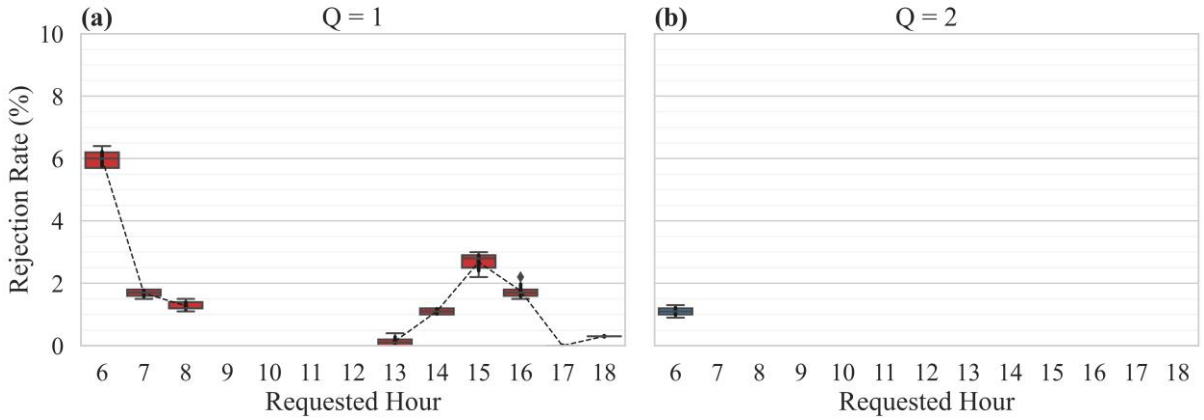

Figure 10.11 Temporal distribution of rejected requests for Q of (a) 1 and (b) 2

Figure 10.12 depicts the eCDF of the percentage of trip delay to trip time for 5 replications. The plots are similar for $Q = 2, 3,$ and 4, while they differ from $Q = 1$. The maximum of trip delay percentage is approximately 36%. With $Q = 1$, nearly 35% of the requests have a trip delay smaller than 30% of the total trip time, while for $Q \geq 2$, around 50% of the requests have a similar trip delay percentage.

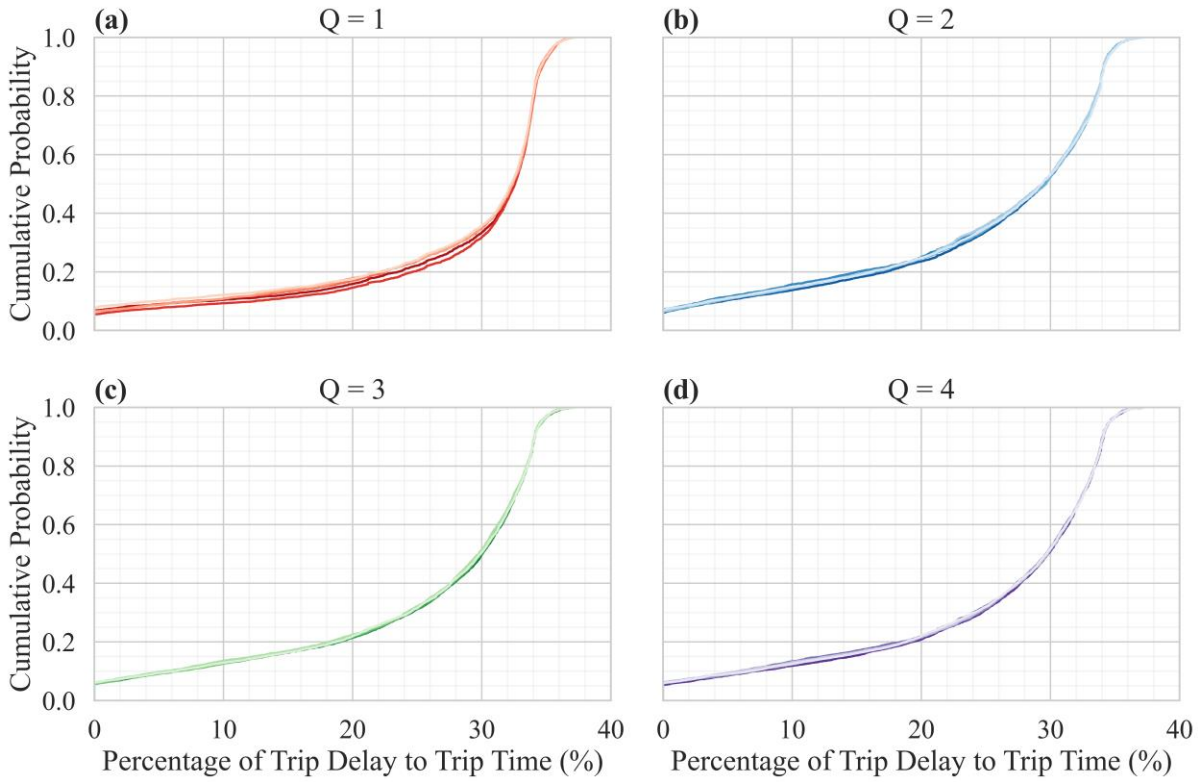

Figure 10.12 Empirical Cumulative Distribution Function (eCDF) of the percentage of trip delay to trip time for 5 replications with Q of (a) 1, (b) 2, (c) 3, and (d) 4

10.5.4.2 Ground Transportation

Table 10.21 summarizes the estimated mean of performance measures related to ground-based legs of the passenger trip. With $Q = 1$, the average ground-based distance and time are 0.06 mi and 0.8 minutes, while with $Q \geq 2$, the average ground-based distance and time are nearly 0.5 mi and 4 minutes, respectively. The minimum length of the non-zero ground-based legs is 0.14 miles, which is below the walking threshold. The maximum of ground leg distance is 0.5 mi for $Q = 1$ and nearly 2 mi for $Q \geq 2$. Since the radius associated with short leg elimination and demand consolidation is 0.5 and 2 mi, respectively, the maximum values of the ground-based legs verify that the passenger trips include a maximum of one ground-based leg.

Table 10.21 Estimated mean of performance measures related to ground-based legs of the request trip

Q	Average Ground Legs Distance (miles)	Minimum Non-zero Ground Legs Distance (miles)	Maximum Ground Legs Distance (miles)	Average Ground Legs Travel Time (minutes)	Minimum Non-zero Ground Legs Travel Time (minutes)	Maximum Ground Legs Travel Time (minutes)
1	0.06	0.14	0.50	0.79	2.80	7.50
2	0.42	0.14	1.95	3.86	2.80	10.70
3	0.48	0.14	1.94	4.38	2.80	10.68
4	0.47	0.14	1.95	4.34	2.80	10.67

Figure 10.13 and Figure 10.14 provide more detailed information on the distribution of ground travel distance and travel time, respectively. Figure 10.14(a) depicts that with $Q = 1$ nearly 85% of the passenger trips have no ground legs, while around 10 percent experience a total ground-based travel time between 4 to 6 minutes. Figure 10.14(c) and (d) show that with $Q \geq 3$, 40% of the requests have no ground-based legs, while the remaining 60% experience a travel time between 4 to 10 minutes.

10.5.4.3 Relocation

Table 10.22 summarizes the estimated mean of the average number of relocations per request and the percentage of customer requests experiencing 0, 1, and 2 relocations. Table 10.22 outlines that given the demand pattern and parameters setting assumed in the experiment design, no request would have more than 1 relocation. Moreover, the percentage of requests with 1 relocation is approximately 15%, 53%, and 60% for $Q = 1$, $Q = 2$, and $Q \geq 3$, respectively.

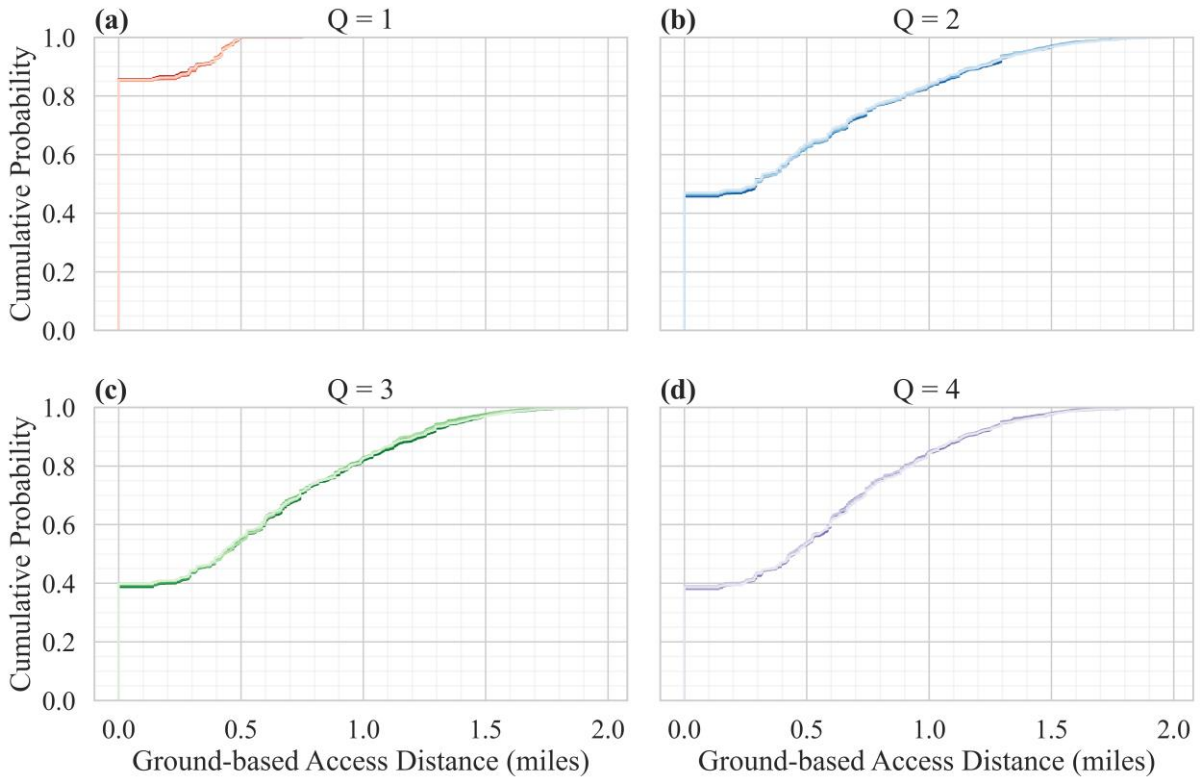

Figure 10.13 eCDF of ground legs distance in miles for Q of (a) 1, (b) 2, (c) 3, and (d) 4

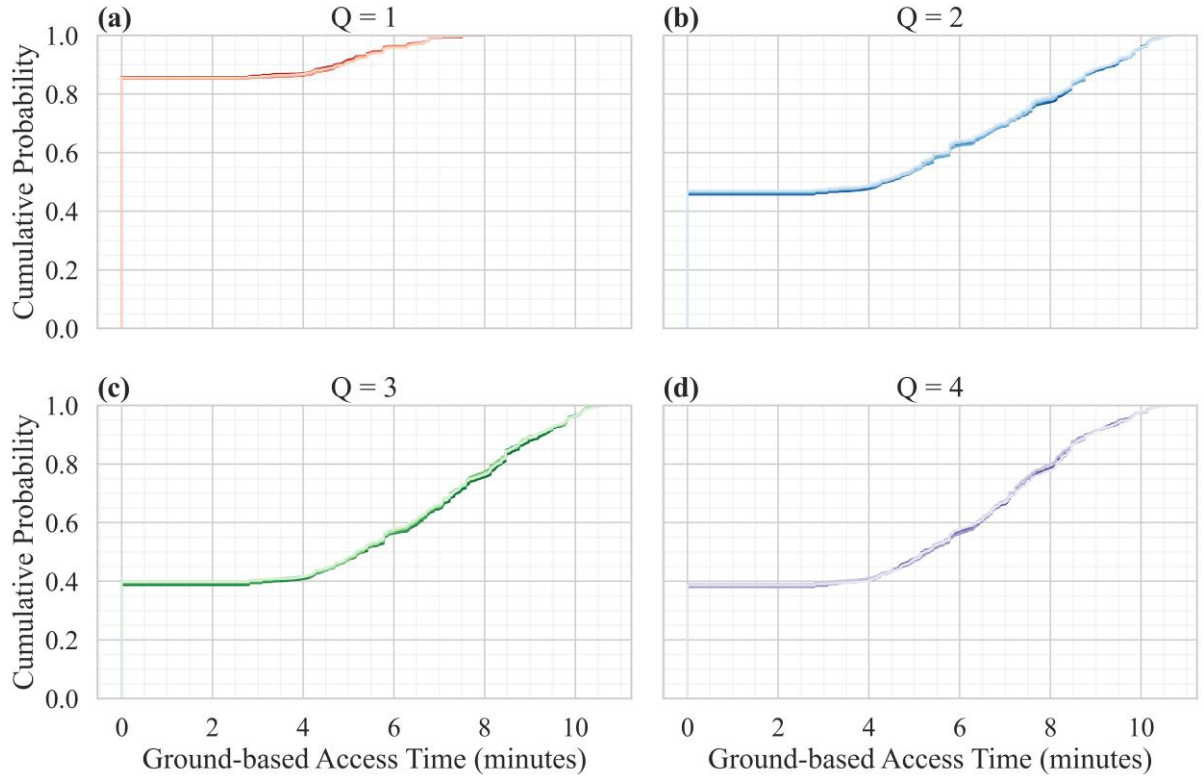

Figure 10.14 eCDF of ground legs travel time in minutes for Q of (a) 1, (b) 2, (c) 3, and (d) 4

Table 10.22 Estimated mean of performance measures related to relocations and ground-based legs of the request trip

Q	Average Number of Relocations	Percentage of Requests with 0 Relocation (%)	Percentage of Requests with 1 Relocation (%)	Percentage of Requests with 2 Relocations (%)
1	0.15	85.26	14.74	0
2	0.53	46.83	53.17	0
3	0.61	39.28	60.71	0
4	0.61	38.58	61.42	0

10.5.4.4 Travel Time Savings

Figure 10.16 outlines the eCDF of travel time savings compared to ground-based TNP trips for $Q = 1$ to 4 and 5 replications. The plots are very similar over various Q. The travel tme savings for nearly 10% of the requests is less than 20 minutes, 70% between 20 and 40 minutes, and the remaining 20% between 40 and 60 minutes. Correspondingly, Figure 10.15 depicts the percentage of trip time savings over 5 replications. Nearly 95% of the requests have a trip time saving between 40% and 70%.

10.6 Case Study: One Week

This section examines the performance measures associated with UAT operation on the week of September 23rd, 2019, with $\eta = 40\%$ and 60% of the qualified TNP demand with $Q = 2$. Table 10.23 summarizes the selected performance measures. For $\eta = 40\%$, the percentage of empty to revenue legs varies between 33% and 53%, while the percentage of empty to revenue mileage varies between 12% to 31%, depending on the demand pattern and intensity. Additionally, the average load factor is estimated between 82% and 89%, except for Sunday, with an average load factor of 74%.

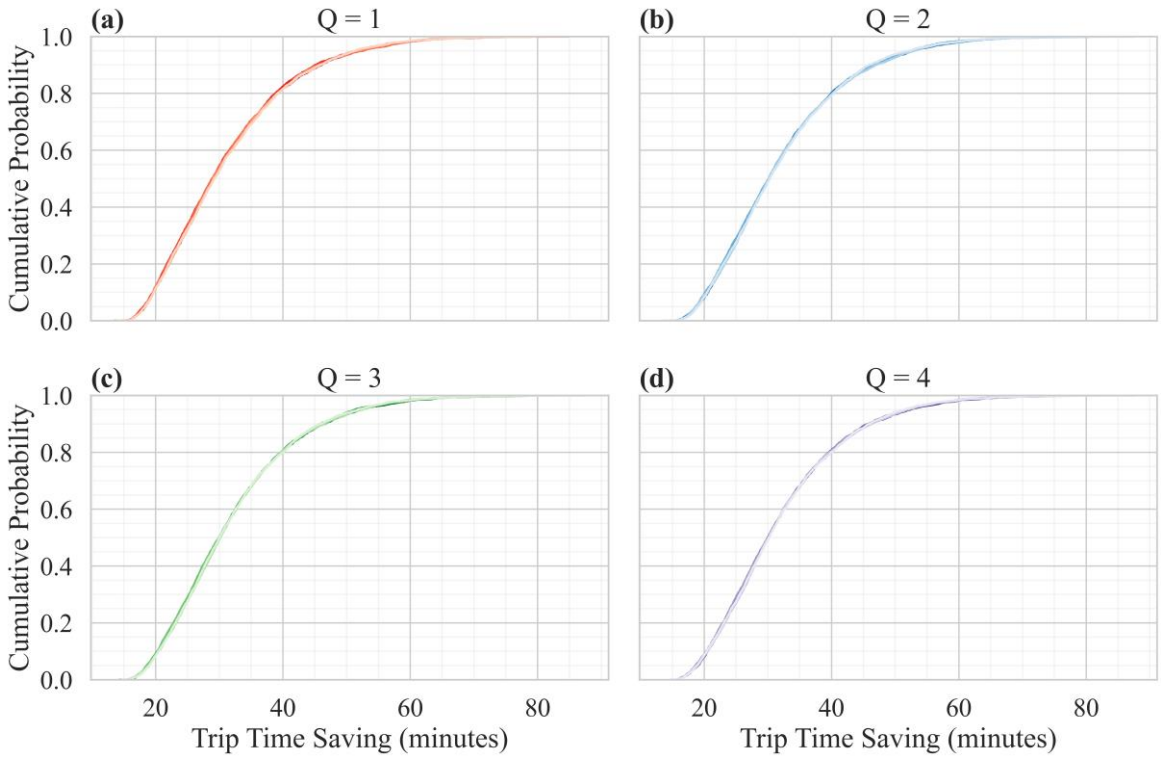

Figure 10.16 eCDF of the trip time saving in minutes for Q of (a) 1, (b) 2, (c) 3, and (d) 4

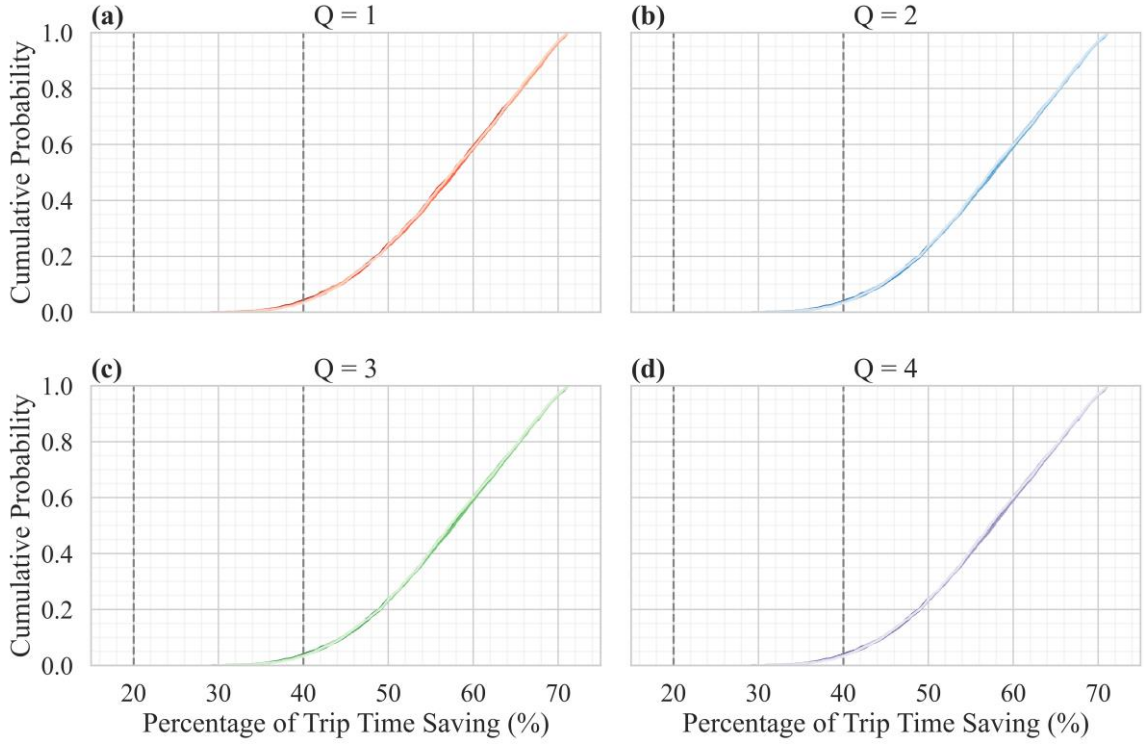

Figure 10.15 eCDF of the percentage of trip time saving for Q of (a) 1, (b) 2, (c) 3, and (d) 4

Table 10.23 Performance measures associated with the operator's cost and revenue and user experience for $Q = 2$

	Cost				Revenue				User Experience						
Day of Week	**Aircraft Utilization (%)**	**Total Aerial Mileage per Served Request (mi)**	**Percentage of Empty to Revenue Flights (%)**	**Percentage of Empty to Revenue Mileage (%)**	**Average Load Factor (%)**	**Total Number of Served Requests**	**PR/α (thousands of miles)**	**PRASM/α**	**Percentage of Rejected Requests (%)**	**Average Trip Delay (min)**	**Mean of Percentage of Trip Delay (%)**	**Average Ground Travel Time (min)**	**Average Number of Relocations**	**Average Trip Time Saving (min)**	**Average Percentage of Trip Time Saving (%)**
$\eta = 40\%$															
Mon	32	10.8	39	18	87	2,724	40.1	0.683	0	6.3	23	3.8	0.52	32.6	58
Tues	26	10.3	31	9	85	2,310	34.5	0.728	0	6.3	23	3.6	0.50	34.7	60
Wed	39	11.2	39	22	88	3,301	49.4	0.668	0	6.5	24	3.8	0.52	34.1	59
Thu	39	11.5	43	24	86	3,261	48.3	0.646	0	6.5	24	3.8	0.52	36.4	61
Fri	38	12.3	50	27	81	2,929	43.1	0.598	0	6.2	23	3.4	0.48	40.8	65
Sat	19	12.6	48	24	80	1,274	19.2	0.599	0	6.1	23	3.2	0.46	28.3	54
Sun	11	12.7	41	9	70	641	9.7	0.594	0	5.6	21	2.7	0.42	26.1	52
$\eta = 60\%$															
Mon	47	10.8	40	22	89	4,041	59.5	0.679	1.1	6.7	25	3.9	0.53	32.2	57
Tues	39	10.4	33	12	87	3,464	51.7	0.718	0	6.7	25	3.6	0.5	34.4	59
Wed	56	11.2	39	24	89	4,873	73.0	0.671	1.6	6.9	25	3.8	0.52	33.6	58
Thu	55	11.2	42	25	88	4,784	70.9	0.659	2.2	7.0	25	3.9	0.53	35.8	60
Fri	55	12.3	52	30	84	4,317	63.4	0.599	1.7	7.0	25	3.5	0.48	40.0	64
Sat	27	12.9	53	31	82	1,910	28.7	0.584	0	6.4	23	3.4	0.48	28.1	53
Sun	15	12.7	45	15	74	961	14.5	0.597	0	5.6	21	2.8	0.43	26.1	52

Assuming a price per passenger mile (i.e., α) of USD 3, the maximum passenger revenue over an 11-hour planning horizon with a fleet of 100 UAT aircraft serving 60% of Chicago TNP demand is approximately USD 219K. The maximum of PRASM/α is, interestingly, for Tuesday, which has the lowest empty to revenue mileage while having a relatively high average load factor.

The average delay is between 5.7 to 7 minutes, corresponding to nearly 25% of the total trip time of the passengers. The average ground-based travel time is between 2.8 and 3.9 minutes. The lowest and highest average trip time savings are for Sunday and Friday, respectively, with 26 minutes (or 52%) and 40 minutes (64%).

Comparing the performance measures for $\eta = 40\%$ and $\eta = 60\%$ shows many similarities. However, the aircraft utilization, percentage of empty to revenue flights, and percentage of empty to revenue mileage are noticeably different under the two scenarios.

10.7 Limitations

The numerical results presented in this section are based on the Chicago TNP demand. While Chicago TNP demand provides a tool for assessing the dynamic solution framework for UAT fleet operations in a real-world setting, it imposes the following limitations:

- The demand model for UAT was a simplified rule-based model, using a fraction of TNP trips longer than 10 miles that would take more than 45 minutes on the ground.
- The qualified TNP demand has a specific pattern that leads to high network efficiency. The majority of trips are between O'Hare International Airport, Midway International Airport, and downtown Chicago. Five and one census tracts, respectively, are the origins and destinations of 50% of the qualified TNP demand, while 50 and 25 census tracts account for 80% of the origins and destinations of the qualified TNP demand.

- The TNP demand is generated at the centroid of the census tracts and, therefore, has been consolidated in some levels. However, the census tracts could be as small as 0.003 square miles.
- The group size of the requests is assumed to be 1.

Additionally, without the loss of generality, the wait time for ground-based service is assumed to be zero at the origin. However, the results of the synthetic network analyses show that performance measures, including the percentage of rejected requests and average load factor, are insensitive to a fixed increase in access time. Nonetheless, the ground-based wait time could impact the travel time savings and, therefore, the UAT demand.

10.8 Concluding Remarks

The UAM OpsCon for passenger-carrying operations commissioned by NASA [29] projects 10s of UAT pads/ports for the intermediate state and 100s for the mature state. The Chicago network consists of nearly 800 census tracts, each corresponding to one UAT pad or port. The size of census tracts varies from 0.003 to 8 square miles, while Antcliff et al. [100] suggest 0.94 and 1.4 square miles per pad for the metropolitan and urban areas, respectively.

The dynamic solution framework for UAT fleet operation is implemented using a fixed fleet of 100 UAT aircraft and 60% of Chicago TNP demand for Monday, Sep. 23rd, 2019, with $Q = 1$, 2, 3, and 4, and the week of Sep. 23rd-30th, 2019, with $Q = 2$. Porsche Consulting [6] estimates that the modal split for a megacity like São Paulo is 9% and 65% for taxi and private cars, respectively. They further estimate that the number of UAT aircraft to replace these services is 30 and 820 UAT aircraft, respectively. However, they do not offer any estimates for the ride-hailing services used in this analysis. Nonetheless, the results of our analysis suggest that serving

60% of qualified Chicago TNP demand with 100 UAT aircraft with 1 passenger seat would result in a rejection rate of 14.6%, while the aircraft utilization is almost 100% during the morning and evening peak hours.

UAM Operational Concept (OpsCon) commissioned by NASA [29] projects 10s of UAT aircraft would fly simultaneously under the Intra-Metro Air Shuttle intermediate state mission, a value that increases to 1,000s for the Ubiquitous Air Taxi mature state mission. Even though this dissertation studies the UAT service, the qualified Chicago TNP demand is well-suited for Intra-Metro Air Shuttle. That being said, the required 100 UAT aircraft to serve 60% of qualified Chicago TNP demand is way below 1,000s for the Ubiquitous Air Taxi.

The results for Monday show that the average number of passengers per flight (with a passenger group size of 1) is 1.8, 2.3, and 2.6 for capacities of 2, 3, and 4. For the one-week analysis, the average load factor varies between 74% to 89%, suggesting the average passenger load of 1.5 to 1.8. In comparison, conventional air taxi services such as DayJet and SATSAir report average passenger loads of 1.3 to 1.7. [100], while UAM market study by Crown Consulting assumes 1 passenger per ride for Air Taxi operation [9].

Booz Allen Hamilton's UAM market study [8] suggests that high network efficiency could increase the UAM demand by more than 200% compared to the base scenario. The network efficiency parameters include utilization (7 hours/day vs. 4 hours/day), load factor (80% vs. 65%), and deadend trips (20% vs 37.5%). The aircraft utilization could get to 56% for the two-seater aircraft over a planning horizon of approximately 11 hours, which implies the utilization of nearly 6 hours/day. The average load factor for 6 days out of the 7 days in the study is more than 80%.

Antcliff et al. [100] posit that door-to-door travel time compared to ground-based transportation is improved by 2.0 to 3.6 times, depending on the cruise speed (120 mph vs. 200 mph) and the density of the pads. Using the actual trip time of the passengers in Chicago, our analysis shows the travel time savings of 52% to 64%, suggesting that the ground-based travel time is 2.3 to 2.8 times of the corresponding UAT service.

Moreover, the average trip time for passengers is about 26 minutes, which is 2.6 times higher than the 10-minute door-to-door trip time for Air Taxi operation in a ubiquitous network, put forward by Crown Consulting UAM market study [9].

Chapter 11 Conclusion

11.1 Summary and Contributions

In recent years, with the vision of eco-friendly autonomous aircraft equipped with electric propulsion and efficient batteries with short charging or swapping time, the interest in air transportation has resurfaced. Benefitting from this revolutionary aircraft technology, the Advanced Air Mobility (AAM) [7] initiative is pursuing to transfer cargo and passengers between urban, local, regional, and intraregional areas, while the Urban Air Mobility (UAM) market focuses on carrying passengers and goods within metropolitan areas [7-9].

Urban Air Taxi (UAT) is the use case of passenger-carrying UAM at its mature state, which offers a ubiquitous (nearly) on-demand per-seat service that moves passengers in urban or suburban areas using groundbreaking aircraft. As of February 2020, 110 passenger-carrying AAM city projects were in progress worldwide [6], and passenger UAM is projected to grow at a compound annual growth rate (CAGR) of 35% by 2035, with 2025 as the starting year [6]. Additionally, in the first half of 2020, USD 907 million was invested in UAM start-ups, nearly 20 times higher than the entire of 2016 [19].

Motivated by this rapid growth, the immense interest in passenger-carrying UAM, and the ensuing travel time savings for the users, this dissertation focuses on the stochastic and dynamic problem of the UAT fleet operation. A UAM market study [8] argues that high network efficiency, including high aircraft utilization and load factor, could increase the UAM demand by more than 200% compared to the base scenario. Furthermore, ride-sharing economics is projected to be one of the three critical steps towards lowering costs [28]. Meanwhile, the UAM research community maintains that more advanced passenger pooling and aircraft dispatching models are needed, and

it calls for algorithms that could be implemented online or use a rolling horizon framework to address the uncertainties encountered in the UAM operation [34,36].

Consequently, the main contributions of this dissertation include outlining the concept of operations for UAT services, defining the UAT problem, and the development and application of a dynamic solution framework to address the stochastic and dynamic problem of UAT fleet operation. Hence, this dissertation provides the UAT operator with a decision-making tool to achieve higher network efficiency. Cities aiming to start UAT operations in the near future could immediately benefit from this solution framework for the UAT operation. Nonetheless, this solution framework sets a benchmark for all use cases of passenger UAM for other city projects in the planning process.

Since a dominant player in the UAM market has yet to emerge, many uncertainties surround the UAM operations. Consequently, the relevant components of UAM and their envisioned characteristics are reviewed first (Chapter 2). Subsequently, the UAT concept of operations, which involves a ubiquitous service with air pooling and elimination of short repositioning flights, is accordingly outlined (Chapter 4). The UAT operation involves numerous components and events, many of which are irrelevant to the problem of UAT fleet operation. As a result, the entities required for modeling the UAT fleet operation are specified (Chapter 5).

Subsequently, a dynamic solution framework with sequential decision-making on a rolling horizon basis is proposed to address the UAT fleet operation problem (Chapter 6). A static and deterministic problem (i.e., snapshot problem) is solved at each decision epoch to help the UAT operator make the dynamic operational decisions, including acceptance and rejection of requests, routing and scheduling the aerial fleet, and assigning the requests to flights. To achieve this goal, the snapshot problem is modeled as a Capacitated Location-Allocation-Routing Problem with

Time Windows and Short Repositioning Elimination (CLARPTW-SRE). Ultimately, the node-based network representation for CLARPTW-SRE is specified (Chapter 7), and the MIP formulation is presented (Chapter 8).

Given the ubiquitous nature of the UAT service, pooling the passengers and increasing the aircraft load factor is deemed a critical step in the success of UAT operations. However, the absence of a dominant eVTOL aircraft technology and UAT operator feeds the uncertainty around UAT. As a result, the impacts of various exogenous and design parameters on demand consolidation are examined using comprehensive sensitivity analyses in a synthetic network (Chapter 9). Furthermore, the dynamic solution framework is implemented using a fixed fleet of UAT aircraft and Chicago Transportation Network Providers (TNPs) demand (Chapter 10). Augmenting the devised UAT operational strategy with real-world data would validate the network efficiency assumptions (e.g., the average load factor and utilization) made by many UAM market studies and offer estimates of the said parameters for future studies.

Ultimately, this research provides a tool for researchers to examine various concepts of operations and evaluate different operational strategies such as sharing or pricing schemes. The outcomes of such studies are valuable for the players from the industry as well as the regulators.

11.2 Limitations and Future Research Areas

The stochasticity in UAT fleet operation involves travel times and demand. The analysis in this dissertation highlights that the success of air pooling hinges on reliable ground-based transportation and the synchronization between aerial and ground-based modes. While this research considers deterministic travel times, incorporating stochasticity of aerial and ground-based wait times and travel times in the solution framework has a high priority for modeling real-

world UAT operations. Another direction for future research involves incorporating the forecasted demand in the decision-making process.

Moreover, the minimum required distance for short repositioning flight legs and the minimum of acceptable delay significantly impact the size of the optimization problem and, consequently, the potentials of CLARPTW-SRE for real-time application. The instances studied in this research are solved using commercial software. However, existing heuristic methods could be tailored, or new heuristic solution methods could be developed to solve the problem in a reasonable time for at-scale UAT operations involving 1000s of aircraft, wider time windows, and longer minimum distances for short repositioning legs.

This research uses a rule-based demand model for UAT, which is dependent on the distance between the origin and destination of each request and the corresponding ground-based travel time. Future studies could explore incorporating a pricing scheme and a more elaborate demand model for UAT operations.

Another research area is to include air traffic control, aerial congestion, and the availability of UAT pads in the UAT model and the decision-making process. As the technology of electric vertical take-off and landing (eVTOL) advances and more details of battery charging or swapping and maintenance requirements become available, these events could also be included in the model associated with UAT fleet operation.

Finally, seamless synchronization between the aerial and ground-based mode has a pivotal role in air pooling, lowering the operational costs, user satisfaction, and successful UAT operation. Integrating the proposed dynamic solution framework for UAT operations with a dynamic model for routing and assigning the vehicles on the ground is another area for future studies.

References

[1] Popular Mechanics. *The Future That Never Was: Pictures from the Past*. https://www.popularmechanics.com/flight/g462/future-that-never-was-next-gen-tech-concepts/. Accessed 10/05/2020.

[2] Vascik, P. D. *Systems-level analysis of On Demand Mobility for aviation*. Master's Thesis, Massachusetts Institute of Technology, 2017.

[3] Espinoza, D., et al. *Per-seat, on-demand air transportation part I: Problem description and an integer multicommodity flow model*. Transportation Science, Vol. 42, No. 3, 2008, pp. 263-278.

[4] Moore, M. D., and K. H. Goodrich. *High speed mobility through on-demand aviation*. In 2013 Aviation Technology, Integration, and Operations Conference, 2013. p. 4373.

[5] Joby & Reinvent. *Commercializing Aerial Ridesharing [Presentation]*. 2021.

[6] Grandl, G., et al. *The Future of Vertical Mobility: Sizing the market for passenger, inspection, and goods services until 2035*. Porsche Consulting, 2018.

[7] FAA NextGen Office. *UAM Concept of Operations (ConOps) Version 1.0*. 2020.

[8] Goyal, R., et al. *Urban Air Mobility (UAM) Market Study (Final Report)*, Booz-Allen and Hamilton Inc., 2018.

[9] Hasan, S. *Urban Air Mobility (UAM) Market Study (Technical Report)*, Crown Consulting, Inc., 2019.

[10] Johnston, T., R. Riedel, and S. Sahdev. *To take off, flying vehicles first need places to land*. McKinsey Center for Future Mobility, 2020.

[11] *Uber Elevate*. https://www.uber.com/us/en/elevate/. Accessed 07/27/2019.

[12] *Airbus Helicopters partners with Blade to boost Urban Air Mobility business*. https://www.airbus.com/newsroom/press-releases/en/2018/02/airbus-helicopters-partners-with-blade-to-boost-urban-air-mobili.html. Accessed 07/27/2019.

[13] *Larry Page-backed Kitty Hawk partners with Boeing on flying car development*. https://www.theverge.com/2019/6/26/18759514/kitty-hawk-cora-boeing-next-partnership-flying-cars-taxis-electric. Accessed 07/27/2019.

[14] *Human-Centred Urban Air Mobility Concept*. https://www.hyundai.com/au/en/why-hyundai/concept-cars/urban-air-mobility. Accessed 5/11/2021.

[15] *Lilium*. https://lilium.com/. Accessed 02/11/2021.

[16] *Volocopter*. https://www.volocopter.com/en/. Accessed 02/11/2021.

[17] *Kitty Hawk*. https://kittyhawk.aero/. Accessed 02/14/2021.

[18] *Joby Aviation*. https://www.jobyaviation.com/. Accessed 02/11/2021.

[19] Roland Berger. *Urban Air Mobility | USD 90 billion of potential: How to capture a share of the passenger drone market*. 2020.

[20] Baur, S., et al. *Urban air mobility: The rise of a new mode of transportation*. Roland Berger Focus, 2018.

[21] *United States Census Bureau*. https://data.census.gov/. Accessed 07/27/2019.

[22] *Mega Commuting in the U.S.* https://www.census.gov/library/working-papers/2013/demo/SEHSD-WP2013-03.html.

[23] Holden, J., and N. Goel. *Fast-forwarding to a future of on-demand urban air transportation*. San Francisco, CA, 2016.

[24] *Though Accidents Are Rare, Crash Highlights Perils of Flying Copters Over City*. https://www.nytimes.com/2019/06/10/nyregion/nyc-helicopter-crash.html. Accessed 07/24/2019.

[25] *Uber shows off its latest concept for air taxis; FAA chief hits hard on safety issue*. https://www.geekwire.com/2019/uber-shows-off-latest-concept-air-taxis-faa-chief-hits-hard-safety-issue/. Accessed 07/23/2019.

[26] *That Nuisance in the Sky*. https://www.nytimes.com/2014/03/23/nyregion/taking-on-noisy-new-york-city-tourist-helicopters.html. Accessed 07/24/2019.

[27] Valente, F. *Frost & Sullivan Presents the Evolving Urban Air Mobility Landscape Up to 2040*. https://ww2.frost.com/news/press-releases/frost-sullivan-presents-the-evolving-urban-air-mobility-landscape-up-to-2040/. Accessed 02/17/2021.

[28] Holden, J. *Uber Keynote: Scaling Uber AIr*. Presented at 2nd Annual Uber Elevate Summit, Los Angeles, CA, 2018.

[29] Price, G., et al. *Urban Air Mobility Operational Concept (OpsCon) Passenger-Carrying Operations*. 2020.

[30] Boeing NeXt. *Flight Path for the Future of the Mobility*.

[31] Volocopter Inc. *Pioneering the Urban Air Taxi Revolution v1.0*. 2019.

[32] EHang. *The future of transportation: white paper on urban air mobility systems*. EHang, 2020. p. 47.

[33] Patterson, M. D., K. R. Antcliff, and L. W. Kohlman. *A proposed approach to studying urban air mobility missions including an initial exploration of mission requirements*. 2018.

[34] Straubinger, A., et al. *An overview of current research and developments in urban air mobility–Setting the scene for UAM introduction*. Journal of Air Transport Management, Vol. 87, 2020, p. 101852.

[35] Rajendran, S., and S. Srinivas. *Air taxi service for urban mobility: a critical review of recent developments, future challenges, and opportunities*. Transportation Research Part E: Logistics and Transportation Review, Vol. 143, 2020, p. 102090.

[36] Garrow, L. A., B. German, and C. Leonard. *Urban air mobility: A comprehensive review and comparative analysis with autonomous and electric ground transportation*. Working paper, Georgia Institute of Technology, 2020.

[37] *Stop calling them flying cars*. https://www.theverge.com/2017/4/25/15426178/flying-car-uber-larry-page-vtol-elevate. Accessed 07/27/2019.

[38] Al Haddad, C., et al. *Factors affecting the adoption and use of urban air mobility*. Transportation Research Part A: policy and practice, Vol. 132, 2020, pp. 696-712.

[39] Fu, M., R. Rothfeld, and C. Antoniou. *Exploring preferences for transportation modes in an urban air mobility environment: Munich case study*. Transportation Research Record, Vol. 2673, No. 10, 2019, pp. 427-442.

[40] Shihab, S. A. M., et al. *By Schedule or On Demand? A Hybrid Operation Concept for Urban Air Mobility*.In AIAA Aviation 2019 Forum, 2019. p. 3522.

[41] Johnson, W. C. *UAM Coordination and Assessment Team (UCAT) NASA UAM for ENRI Technical Interchange Meeting*. 2019.

[42] *Joby Aviation Welcomes New $75M Investment from Uber as it Acquires Uber Elevate and Expands Partnership*. https://www.jobyaviation.com/news/joby-aviation-welcomes-new-75m-investment-from-uber-as-it-acquires-uber-elevate-and-expands-partnership/. Accessed 02/20/2021.

[43] *Airbus Urban Air Mobility*. https://www.airbus.com/innovation/zero-emission/urban-air-mobility.html. Accessed 02/14/2021.

[44] *Uber Copter to Offer Flights From Lower Manhattan to J.F.K.* https://www.nytimes.com/2019/06/05/travel/uber-helicopter-nyc-jfk.html. Accessed 07/23/2019.

[45] *BLADE Urban Air Mobility, Inc.* https://www.blade.com/. Accessed 5/17/2021.

[46] *EHang*. https://www.ehang.com/uam/. Accessed 02/14/2021.

[47] *CityAirbus*. https://www.airbus.com/innovation/zero-emission/urban-air-mobility/cityairbus.html. Accessed 02/13/2021.

[48] *PAV – Passenger Air Vehicle: Redefining the future of safe on-demand air travel and transportation*. https://www.aurora.aero/pav-evtol-passenger-air-vehicle/.

[49] *Hyundai and Uber Announce Aerial Ridesharing Partnership, Release New Full-Scale Air Taxi Model at CES*. Accessed 5/11/2021.

[50] Rapino, M. A., and A. K. Fields. *Mega commuters in the U.S.: time and distance in defining the long commute using the American community survey*. Presented at Association for Public Policy Analysis and Management, 2013.

[51] Uber Elevate. *Uber eVTOL Vehicle Requirements and Missions*. https://s3.amazonaws.com/uber-static/elevate/Summary+Mission+and+Requirements.pdf. Accessed 02/21/2021.

[52] *Wisk*. https://wisk.aero/. Accessed 11/06/2021.

[53] Moore, M. D. *Personal air vehicles: a rural/regional and intra-urban on-demand transportation system*. Journal of the American Institute of Aeronautics and Astronautics (AIAA), Vol. 2646, 2003.

[54] Hill, B. P., et al. *UAM Vision Concept of Operations (ConOps) UAM Maturity Level (UML) 4 Version 1.0*. 2020.

[55] Antcliff, K. *Silicon valley early adopter CONOPs and market study*.In Transformative vertical flight workshop. NASA, Washington, 2015.

[56] Cheyno, E. *Vertiport design and operations*.In 4th Annual Meeting and Technical Display, 1967. p. 891.

[57] Holmes, B., et al. *NASA strategic framework for on-demand air mobility*. NASA Contractor Report NNL13AA08B, National Institute of Aerospace, Hampton, VA, 2017.

[58] Vascik, P. D., and R. J. Hansman. *Development of vertiport capacity envelopes and analysis of their sensitivity to topological and operational factors*.In AIAA SciTech 2019 Forum, 2019. p. 0526.

[59] Fagerholt, K., B. Foss, and O. Horgen. *A decision support model for establishing an air taxi service: a case study*. Journal of The Operational Research Society, Vol. 60, No. 9, 2009, pp. 1173-1182.

[60] Bonnefoy, P. A. *Simulating air taxi networks*.In Proceedings of the 2005 Winter Simulation Conference, IEEE, 2005. p. 10 pp.

[61] Boyd, J. A., et al. *Investigation of future air taxi services using discrete-event simulation*.In 2006 IEEE Systems and Information Engineering Design Symposium, IEEE, 2006. pp. 25-30.

[62] Hyland, M. F. *Real-Time Operation of Shared-Use Autonomous Vehicle Mobility Services: Modeling, Optimization, Simulation, and Analysis*. Ph.D. Dissertation, Northwestern University, 2018.

[63] Hicks, R., et al. *Bombardier flexjet significantly improves its fractional aircraft ownership operations*. Interfaces, Vol. 35, No. 1, 2005, pp. 49-60.

[64] Yao, Y., et al. *Strategic planning in fractional aircraft ownership programs*. European Journal of Operational Research, Vol. 189, No. 2, 2008, pp. 526-539.

[65] Skyway Air Taxi. *Air Taxi Cost*. https://skywayairtaxi.com/air-taxi-cost-pricing/. Accessed 5/18/2021.

[66] Joby Aviation. *Joby Oct'21 Corporate Deck [Presentation]*. 2021.

[67] Pelli, U., and R. Riedel. *Flying-cab drivers wanted*. McKinsey Center for Future Mobility, 2020.

[68] Cordeau, J.-F., et al. *Transportation on demand*. Handbooks in operations research and management science, Vol. 14, 2007, pp. 429-466.

[69] Pillac, V., et al. *A review of dynamic vehicle routing problems*. European Journal of Operational Research, Vol. 225, No. 1, 2013, pp. 1-11.

[70] Powell, W. B., P. Jaillet, and A. Odoni. *Stochastic and dynamic networks and routing*. Handbooks in operations research and management science, Vol. 8, 1995, pp. 141-295.

[71] Bektas, T., P. P. Repoussis, and C. D. Tarantilis. *Chapter 11: dynamic vehicle routing problems*. In Vehicle Routing: Problems, Methods, and Applications, Second Edition, SIAM, 2014. pp. 299-347.

[72] Berbeglia, G., J.-F. Cordeau, and G. Laporte. *Dynamic pickup and delivery problems*. European Journal of Operational Research, Vol. 202, No. 1, 2010, pp. 8-15.

[73] Regan, A. C., H. S. Mahmassani, and P. Jaillet. *Dynamic decision making for commercial fleet operations using real-time information*. Transportation Research Record, Vol. 1537, No. 1, 1996, pp. 91-97.

[74] ---. *Evaluation of dynamic fleet management systems: Simulation framework*. Transportation Research Record, Vol. 1645, No. 1, 1998, pp. 176-184.

[75] Yang, J., P. Jaillet, and H. S. Mahmassani. *On-line algorithms for truck fleet assignment and scheduling under real-time information*. Transportation Research Record, Vol. 1667, No. 1, 1999, pp. 107-113.

[76] Mahmassani, H. S., Y. Kim, and P. Jaillet. *Local optimization approaches to solve dynamic commercial fleet management problems*. Transportation Research Record, Vol. 1733, No. 1, 2000, pp. 71-79.

[77] Kim, Y., H. S. Mahmassani, and P. Jaillet. *Dynamic truckload truck routing and scheduling in oversaturated demand situations*. Transportation Research Record, Vol. 1783, No. 1, 2002, pp. 66-71.

[78] Yang, J., P. Jaillet, and H. Mahmassani. *Real-time multivehicle truckload pickup and delivery problems*. Transportation Science, Vol. 38, No. 2, 2004, pp. 135-148.

[79] Smilowitz, K. *Multi-resource routing with flexible tasks: an application in drayage operations*. Iie Transactions, Vol. 38, No. 7, 2006, pp. 577-590.

[80] Francis, P., G. Zhang, and K. Smilowitz. *Improved modeling and solution methods for the multi-resource routing problem*. European Journal of Operational Research, Vol. 180, No. 3, 2007, pp. 1045-1059.

[81] Ho, S. C., et al. *A survey of dial-a-ride problems: Literature review and recent developments*. Transportation Research Part B: Methodological, Vol. 111, 2018, pp. 395-421.

[82] Cordeau, J.-F., and G. Laporte. *The dial-a-ride problem: models and algorithms*. Annals of Operations Research, Vol. 153, No. 1, 2007, pp. 29-46.

[83] Engineer, F. G., G. L. Nemhauser, and M. W. Savelsbergh. *Dynamic programming-based column generation on time-expanded networks: Application to the dial-a-flight problem*. INFORMS Journal on Computing, Vol. 23, No. 1, 2011, pp. 105-119.

[84] Campbell, I. M. D. *Construction heuristics for the airline taxi problem*. Ph.D. Dissertaton, University of the Witwatersrand, Johannesburg, 2013.

[85] Reddy, D. T. *Agent based simulation of the dial-a-flight problem*. Master's Thesis, University of the Witwatersrand, Johannesburg, 2018.

[86] Martin, C., D. Jones, and P. Keskinocak. *Optimizing on-demand aircraft schedules for fractional aircraft operators*. Interfaces, Vol. 33, No. 5, 2003, pp. 22-35.

[87] Keskinocak, P., and S. Tayur. *Scheduling of time-shared jet aircraft*. Transportation Science, Vol. 32, No. 3, 1998, pp. 277-294.

[88] Yang, W., et al. *Aircraft and crew scheduling for fractional ownership programs*. Annals of Operations Research, Vol. 159, No. 1, 2008, pp. 415-431.

[89] Munari, P. *Mathematical modeling in the airline industry: optimizing aircraft assignment for on-demand air transport*. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, Vol. 5, No. 1, 2017.

[90] Munari, P., and A. Alvarez. *Aircraft routing for on-demand air transportation with service upgrade and maintenance events: Compact model and case study*. Journal of Air Transport Management, Vol. 75, 2019, pp. 75-84.

[91] Sun, X., S. Wandelt, and E. Stumpf. *Competitiveness of on-demand air taxis regarding door-to-door travel time: A race through Europe*. Transportation Research Part E: Logistics and Transportation Review, Vol. 119, 2018, pp. 1-18.

[92] Rajendran, S., and J. Zack. *Insights on strategic air taxi network infrastructure locations using an iterative constrained clustering approach*. Transportation Research Part E: Logistics and Transportation Review, Vol. 128, 2019, pp. 470-505.

[93] *Aeronautics and Space, 14 C.F.R.*

[94] *Aeronautics and Space, 14 C.F.R. §121*. 2021.

[95] *Aeronautics and Space, 14 C.F.R. §135*. 2021.

[96] U.S. Department of Transportation (DOT). *Charter Flights*. https://www.transportation.gov/individuals/aviation-consumer-protection/charter-flights#Single_Entity_Charters. Accessed 5/18/2021.

[97] Van der Zwan, F., K. Wils, and S. Ghijs. *Development of an aircraft routing system for an air taxi operator*. In Aeronautics and Astronautics, IntechOpen, 2011.

[98] CJI TEAM. *Linear Air Leads U.S. Air Taxi Revolution*. https://www.corporatejetinvestor.com/articles/linear-air-leads-us-air-taxi-revolution-722/#:~:text=Flying%20taxi%20cabs&text=Like%20with%20private%20jet%20charter,well%20served%20by%20commercial%20airlines. Accessed 5/18/2021.

[99] National Business Aviation Association (NBAA). *Guide to Sell Charter by Seat: Balancing Innovation with Regulations*. https://nbaa.org/wp-content/uploads/2018/02/guide-to-selling-charter-by-the-seat.pdf. Accessed 5/18/2021.

[100] Antcliff, K. R., M. D. Moore, and K. H. Goodrich. *Silicon Valley as an Early Adopter for On-Demand Civil VTOL Operations*. In 16th AIAA Aviation Technology, Integration, and Operations Conference, 2016. p. 3466.

[101] Rothfeld, R., et al. *Initial analysis of urban air mobility's transport performance in Sioux Falls*. In 2018 Aviation Technology, Integration, and Operations Conference, 2018. p. 2886.

[102] Rajendran, S., and J. Shulman. *Study of emerging air taxi network operation using discrete-event systems simulation approach*. Journal of Air Transport Management, Vol. 87, 2020, p. 101857.

[103] Ale-Ahmad, H., and H. Mahmassani. *A Dynamic Solution Framework for Urban Air Taxi Fleet Operation with Application to Chicago Network*. Working Paper.

[104] Baik, H., et al. *Forecasting model for air taxi, commercial airline, and automobile demand in the United States*. Transportation Research Record, Vol. 2052, No. 1, 2008, pp. 9-20.

[105] Kreimeier, M., E. Stumpf, and D. Gottschalk. *Economical assessment of air mobility on demand concepts with focus on Germany*. In 16th AIAA Aviation Technology, Integration, and Operations Conference, 2016. p. 3304.

[106] Binder, R., et al. *If You Fly It, Will Commuters Come? A Survey to Model Demand for eVTOL Urban Air Trips*. In 2018 Aviation Technology, Integration, and Operations Conference, 2018. p. 2882.

[107] Balac, M., et al. *Demand estimation for aerial vehicles in urban settings*. IEEE Intelligent Transportation Systems Magazine, Vol. 11, No. 3, 2019, pp. 105-116.

[108] Garrow, L. A., et al. *Commuting in the age of the Jetsons: a market segmentation analysis of autonomous ground vehicles and air taxis in five large US cities*. In AIAA AVIATION 2020 FORUM, 2020. p. 3258.

[109] Boddupalli, S.-S., L. A. Garrow, and B. J. German. *Mode Choice Modeling for an Electric Vertical Take-off and Landing (eVTOL) Air Taxi Commuting Service in Five Large U.S. Cities*. Working paper, Georgia Institute of Technology, Atlanta, GA, 2020.

[110] Keysan, G. *Tactical and operational planning for per-seat, on-demand air transportation*. Georgia Institute of Technology, 2009.

[111] Rath, S., and J. Y. Chow. *Air Taxi Skyport Location Problem for Airport Access*. In 99th Annual Meeting of the Transportation Research Board, Washington D.C., 2019.

[112] Fadhil, D. N. *A GIS-based analysis for selecting ground infrastructure locations for urban air mobility*. inlangen]. Master's Thesis, Technical University of Munich, 2018.

[113] Rothfeld, R., et al. *Agent-based simulation of urban air mobility*. In 2018 Modeling and Simulation Technologies Conference, 2018. p. 3891.

[114] Espinoza, D., et al. *Per-seat, on-demand air transportation part II: Parallel local search*. Transportation Science, Vol. 42, No. 3, 2008, pp. 279-291.

[115] La Foy, T. L. *On-demand air transportation flight scheduling*. 2013.

[116] Lee, D. W., et al. *A traffic engineering model for air taxi services*. Transportation Research Part E: Logistics and Transportation Review, Vol. 44, No. 6, 2008, pp. 1139-1161.

[117] Drexl, M. *Synchronization in vehicle routing—a survey of VRPs with multiple synchronization constraints*. Transportation Science, Vol. 46, No. 3, 2012, pp. 297-316.

[118] Nagy, G., and S. Salhi. *Location-routing: Issues, models and methods*. European Journal of Operational Research, Vol. 177, No. 2, 2007, pp. 649-672.

[119] Prodhon, C., and C. Prins. *A survey of recent research on location-routing problems*. European Journal of Operational Research, Vol. 238, No. 1, 2014, pp. 1-17.

[120] Jacobsen, S. K., and O. B. Madsen. *A comparative study of heuristics for a two-level routing-location problem*. European Journal of Operational Research, Vol. 5, No. 6, 1980, pp. 378-387.

[121] Hoff, A., and A. Løkketangen. *A tabu search approach for milk collection in western Norway using trucks and trailers*. 2008.

[122] Semet, F., and E. Taillard. *Solving real-life vehicle routing problems efficiently using tabu search*. Annals of Operations Research, Vol. 41, No. 4, 1993, pp. 469-488.

[123] Villegas, J. G., et al. *GRASP/VND and multi-start evolutionary local search for the single truck and trailer routing problem with satellite depots*. Engineering Applications of Artificial Intelligence, Vol. 23, No. 5, 2010, pp. 780-794.

[124] Gonzalez-Feliu, J. *The multi-echelon location-routing problem: Concepts and methods for tactical and operational planning*. Technical report, Laboratoire d'Economie des Transports, Lyon, France, 2010.

[125] Drexl, M. *Applications of the vehicle routing problem with trailers and transshipments*. European Journal of Operational Research, Vol. 227, No. 2, 2013, pp. 275-283.

[126] Park, J., and B.-I. Kim. *The school bus routing problem: A review*. European Journal of Operational Research, Vol. 202, No. 2, 2010, pp. 311-319.

[127] Chapleau, L., J.-A. Ferland, and J.-M. Rousseau. *Clustering for routing in densely populated areas*. European Journal of Operational Research, Vol. 20, No. 1, 1985, pp. 48-57.

[128] Carlsson, J. G., and S. Song. *Coordinated logistics with a truck and a drone*. Management Science, Vol. 64, No. 9, 2017, pp. 4052-4069.

[129] Banks, J., et al. *Discrete-event system simulation*. Prentice Hall, Upper Saddle River, N.J., 2010.

[130] Siebers, P.-O., et al. *Discrete-event simulation is dead, long live agent-based simulation!* Journal of Simulation, Vol. 4, No. 3, 2010, pp. 204-210.

[131] Hyland, M., and H. S. Mahmassani. *Dynamic autonomous vehicle fleet operations: Optimization-based strategies to assign AVs to immediate traveler demand requests*. Transportation Research Part C: Emerging Technologies, Vol. 92, 2018, pp. 278-297.

[132] Powell, W. B. *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley & Sons, 2007.

[133] Puterman, M. L. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[134] Corberán, Á., et al. *Arc routing problems: A review of the past, present, and future*. Networks, Vol. 77, No. 1, 2021, pp. 88-115.

[135] Boland, N., et al. *The price of discretizing time: a study in service network design*. EURO Journal on Transportation and Logistics, Vol. 8, No. 2, 2019, pp. 195-216.

[136] Bertsimas, D., P. Jaillet, and S. Martin. *Online vehicle routing: The edge of optimization in large-scale applications*. Operations Research, Vol. 67, No. 1, 2019, pp. 143-162.

[137] Fishman, G. S. *Discrete-event simulation: modeling, programming, and analysis*. 2001.

[138] Nelson, B. *Foundations and methods of stochastic simulation: a first course*. Springer Science & Business Media, 2013.

[139] Fu, M. C. *Optimization for simulation: Theory vs. practice*. INFORMS Journal on Computing, Vol. 14, No. 3, 2002, pp. 192-215.

[140] Cronkleton, E. *What Is the Average Walking Speed of an Adult?* https://www.healthline.com/health/exercise-fitness/average-walking-speed#takeaway. Accessed 11/30/2021.

[141] Goyal, R., et al. *Advanced Air Mobility: Demand Analysis and Market Potential of the Airport Shuttle and Air Taxi Markets*. Sustainability, Vol. 13, No. 13, 2021, p. 7421.

[142] Shortle, J. F., et al. *Fundamentals of queueing theory*. John Wiley & Sons, 2018.

[143] Ale-Ahmad, H., Y. Chen, and H. S. Mahmassani. *Travel Time Variability and Congestion Assessment for Origin–Destination Clusters through the Experience of Mobility Companies*. Transportation Research Record, Vol. 2674, No. 12, 2020, pp. 103-117.

[144] *Chicago Transportation Network Provider Trips*. https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p. Accessed 07/26/2019.

[145] *How Chicago protects privacy in TNP and taxi open data*. http://dev.cityofchicago.org/open%20data/data%20portal/2019/04/12/tnp-taxi-privacy.html. Accessed 05/31/2021.

[146] The University of Chicago Library. *Chicago Census Maps*. https://guides.lib.uchicago.edu/c.php?g=633403&p=4428901. Accessed 11/29/2021.

[147] Open Data Portal Team. *Census tract rules for taxi and TNP datasets*. http://dev.cityofchicago.org/open%20data/data%20portal/2019/07/29/taxi-tnp-low-count-explanation.html. Accessed 5/31/2021.

[148] City of Chicago. *Minimum Wage Update*. https://www.chicago.gov/city/en/depts/bacp/provdrs/enforce/alerts/2019/may/minimumwageupdate.html#:~:text=May%2028%2C%202019,information%20go%20to%20Minimum%20Wage. Accessed 5/31/2021.

Appendix A. Notations

SYMBOL	DEFINITION
Decision Epoch	
e	Decision epoch index
E	Number of decision epochs
$\mathcal{E}$	Set of decision epoch indices, $\mathcal{E} = \{1, 2, \dots, e, \dots, E\}$
v_e^S	The event associated with the start of decision epoch e
v_e^E	The event associated with the end of decision epoch e
$\tau_{v_e^S}$	The start time of decision epoch e
$\tau_{v_e^E}$	The end time of decision epoch e
T_e^{EPOCH}	Length of decision epoch e; $T_e^{EPOCH} = \tau_{v_e^E} - \tau_{v_e^S}$
Δt^{UPDATE}	The re-optimization interval; $\Delta t^{UPDATE} = \tau_{v_e^S} - \tau_{v_{e-1}^S}$
UAT Aircraft	
a_k	UAT aircraft k
K	Aerial fleet size
$\mathcal{K}$	Set of available aircraft; $\mathcal{K} = \{a_1, a_2, \dots, a_k, \dots, a_K\}$.
Q_k	Capacity of aircraft k
Q	Capacity of the aerial fleet in a homogeneous fleet
v_k^{AIR}	Speed of UAT aircraft k (mph)
v^{AIR}	Speed of UAT aircraft in a homogeneous fleet (mph)
τ_{kt}^{AVL}	The earliest time the subsequent itinerary of UAT aircraft k could be modified as of time t.
$\boldsymbol{L}_{kt}^{AVL}$	Location of UAT aircraft k at τ_{kt}^{AVL}
$\mathcal{Q}_{kt}$	Ordered list of non-completed flight legs assigned to UAT aircraft k as of time t
$\mathcal{Q}_{kt}^{WAIT}$	Ordered list of flight legs assigned to UAT aircraft k that have not started as of time t
$\mathcal{Q}_{kt}^{WREV}$	Ordered list of flight legs assigned to UAT aircraft k that have not started as of time t

$\mathfrak{E}_{kt}^{NDSRD}$	Binary variable, which is 1 if $\boldsymbol{L}_{kt}^{AVL}$ is the drop-off UAT pad of the passengers, but not their desired one, 0 in any other case.
$\mathbb{A}_k^{eVTOL}$	Static attributes of aircraft k, $\mathbb{A}_k^{eVTOL} = (Q_k, v_k^{AIR})$
S_t^{eVTOL}	State of the UAT aircraft at time t; $S_t^{eVTOL} = (\varsigma_{kt}^{eVTOL}, \tau_{kt}^{AVL}, \boldsymbol{L}_{kt}^{AVL}, \mathcal{Q}_{kt})_{a_k \in \mathcal{K}}$.
ς_{kt}^{eVTOL}	Status of the UAT aircraft k at time t

Requests

$\mathscr{r}_r$	Request r
$\mathscr{r}_i^{INTND}$	The intended request of flight leg i
$\boldsymbol{O}_r$	Origin of request r
$\boldsymbol{D}_r$	Destination of request r
$\boldsymbol{S}_r^{DSRD}$	Desired pick-up UAT pad of request r
$\boldsymbol{E}_r^{DSRD}$	Desired drop-off UAT pad of request r
q_r	Group size of request r
τ_r^{ARV}	Arival time of request r
τ_r^{REQ}	Requested time of service for request r
T_r^{ADV}	Reservation time window of request r; $T_r^{ADV} = \tau_r^{REQ} - \tau_r^{ARV}$.
τ_r^{DLN}	The latest time the passenger group of request r could reach their destination
T_r^{DSRD}	Minimum trip time corresponding to the desired flight leg of request r
τ_{rt}^{SRVC}	The earliest time the passenger group of request r could start the service and leave their origin
$\mathbb{A}_r^{REQ}$	Static attributes of request r; $\mathbb{A}_r^{REQ} = (\boldsymbol{O}_r, \boldsymbol{D}_r, \boldsymbol{S}_r^{DSRD}, \boldsymbol{E}_r^{DSRD}, q_r, \tau_r^{REQ})$,
S_t^{REQ}	State of all the requests that have been placed by time t
S_{rt}^{REQ}	State of request r as of time t
ς_{rt}^{REQ}	Status of request r as of time t

Flight Legs

$\mathscr{f}_i$	Flight leg i
$\mathscr{f}_r^{DSRD}$	Desired flight leg of request r
$\mathscr{f}_{kt}^{CRNT}$	Flight leg in service by aircraft k as of time t

$\boldsymbol{S}_i$	Starting UAT pad of flight i
$\boldsymbol{E}_i$	Ending UAT pad of flight i
H_i	A binary variable denoting whether flight leg i is empty (0) or revenue-generating (1)
τ_i^{MIN}	The earliest time that flight leg i could be served
τ_i^{MAX}	The latest time that flight leg i could be served
τ_{it}^{STRT}	Scheduled start time of flight leg i as of time t
τ_{it}^{COMP}	Scheduled completion time of flight leg i as of time t
$\mathbb{A}_i^{LEG}$	Static attributes of flight leg i, $\mathbb{A}_i^{LEG} = (\boldsymbol{S}_i, \boldsymbol{E}_i, H_i)$
$\mathbb{A}_i^{REVLEG}$	Additional static attributes of revenue-generating flight leg i; $\mathbb{A}_i^{REVLEG} = \left(\mathscr{r}_i^{INTND}, \tau_i^{MIN}, \tau_i^{MAX}\right)$
S_i^{LEG}	State of flight leg i as of time t
ς_{it}^{LEG}	Status of flight leg i as of time t
$\mathbb{F}_i(\boldsymbol{S}_i, \boldsymbol{E}_i, \mathscr{r}_i^{INTND})$	Function that defines $\mathscr{f}_i$ such that it starts at $\boldsymbol{S}_i$ and ends at $\boldsymbol{E}_i$, with intended request $\mathscr{r}_i^{INTND}$
$\mathbb{S}(\mathscr{f})$	Function that returns the starting point of $\mathscr{f}$

Sets of Requests

$\mathcal{R}_t^{CAND}$	Set of candidate requests as of time t
$\mathcal{R}_t^{UNASGN}$	Set of unassigned requests as of time t
$\mathcal{R}_t^{FLXSTRT}$	Set of requests with a flexible pick-up UAT pad as of time t
$\mathcal{R}_t^{FXDSTRT}$	Set of requests with a fixed pick-up UAT pad as of time t
$\bar{\mathcal{R}}_e^{ACCPT}$	Set of accepted candidate requests during decision epoch e
$\bar{\mathcal{R}}_e^{REJCT}$	Set of rejected candidate requests during decision epoch e

Sets of Flight Legs

$\mathcal{F}_{rt}^{KE}$	Set of flight legs that start at a first availability UAT pad of a UAT aircraft and end at the desired drop-off UAT pad of a candidate request
$\mathcal{F}_{rt}^{KS}$	Set of flight legs that start at a first availability UAT pad of a UAT aircraft and end at the desired pick-up UAT pad of a candidate request
$\mathcal{F}_{rt}^{SS}$	Set of flight legs that start at the desired pick-up UAT pad of a candidate request aircraft and end at the desired pick-up UAT pad of a candidate request

$\mathcal{F}_{rt}^{EE}$	Set of flight legs that start at the desired drop-off UAT pad of a candidate request aircraft and end at the desired drop-off UAT pad of a candidate request
$\mathcal{F}_{rt}^{ES}$	Set of flight legs that start at the desired drop-off UAT pad of a candidate request aircraft and end at the desired pick-up UAT pad of a candidate request
$\mathcal{F}_{t}^{CAND}$	Set of candidate flight legs at time t
$\mathcal{F}_{t}^{DSRD}$	Set of feasible desired flight legs as of time t
$\mathcal{F}_{t}^{CNCT}$	Set of feasible connecting flight legs of time t
$\mathcal{F}_{it}^{SUCC}$	Set of candidate flight legs that
$\mathcal{F}_{t}^{S}$	Set of candidate flight legs that start at the desired UAT pad of their intended request
$\overline{\mathcal{F}_{t}^{S}}$	Set of candidate flight legs that do *not* start at the desired UAT pad of their intended request
$\mathcal{F}_{t}^{E}$	Set of candidate flight legs that end at the desired UAT pad of their intended request
$\overline{\mathcal{F}_{t}^{E}}$	Set of candidate flight legs that do *not* end at the desired UAT pad of their intended request
$\overline{\tilde{\mathcal{F}}_{t}^{E}}$	Set of candidate flight legs that do *not* end at the desired UAT pad of their intended request for the reduced network $\tilde{\mathcal{G}}_t$ as of time t

Distances

$dist(p, q)$	Distance (as the crow flies) between point p and point q in the space
D_{r}^{OD}	Distance (as the crow flies) between the origin and destination of request r
D_{i}^{LEG}	Aerial distance of flight leg i
D_{kit}^{0}	Aerial distance between the first availability UAT pad of aircraft k as of time t and the starting UAT pad of flight leg i
D_{ij}	Aerial distance between the ending UAT pad of flight leg i and the starting UAT pad of flight leg j

Access Times

W_{rit}^{INGRS}	Wait time as of time t for ground-based transportation from the origin of request r to the starting UAT pad of flight leg i
W_{ri}^{EGRS}	Wait time as of time t for ground-based transportation from the origin of request r to the starting UAT pad of flight leg i

T_{ri}^{INGRS}	Duration of ingress of request r to flight leg i
T_{ri}^{EGRS}	Duration of egress of request r from flight leg i
T_{ri}^{INBND}	Time to reach the starting UAT pad of flight leg i from the origin of request r using walking or driving
T_{ri}^{OUTBND}	Time to reach the destination of request r from ending UAT pad of flight leg i using walking or driving
T_{ri}^{DGATE}	Elapsed time between arriving at the UAT port on the ground and reaching the departure gate of flight i for request r
T_{ri}^{AGATE}	Elapsed time for reaching the area of ground transportation from the arrival gate flight leg i for request r

Aerial Times

T_i^{FLIGHT}	Flight time of flight leg i, including ascending, descending, and cruising
T_i^{CRUISE}	Cruise time of flight leg i
T^{ASCEND}	Time elapsed in ascending for an eVTOL aircraft
$T^{DESCEND}$	Time elapsed in descending for an eVTOL aircraft
$T^{TAKEOFF}$	Time required for ATC clearance before take-off
$T^{LANDING}$	Time required for clearance after landing
T^{BOARD}	Duration of boarding passengers
$T^{DEBOARD}$	Duration of deboarding passengers
T_i^{SRVEMP}	Service time of empty flight leg i
T_i^{SRVREV}	Service time of revenue-generating flight leg i
T_{kit}^{0}	Time required to get from the first availability UAT pad of aircraft k as of time t to the starting UAT pad of flight leg i
T_{ij}	Time required to get from the ending UAT pad of flight leg i to the starting UAT pad of flight leg j
T_{mt}^{STRTTW}	Time window for starting empty flight leg m as of time t

Nodes

$\mathcal{N}_t$	Nodes of the network $\mathcal{G}_t$ as of time t
$\mathcal{N}_t^{eVTOL}$	Set of nodes associated with the UAT aircraft in the network $\mathcal{G}_t$ as of time t

$\widetilde{\mathcal{N}}_t^{eVTOL}$	Set of nodes associated with the UAT aircraft in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\mathcal{N}_t^{LEG}$	Set of nodes associated with the candidate flight legs in the network $\mathcal{G}_t$ as of time t
$\widetilde{\mathcal{N}}_t^{LEG}$	Set of nodes associated with the candidate flight legs in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\mathcal{N}_t^{REQ}$	Set of nodes associated with the candidate requests in the network $\mathcal{G}_t$ as of time t
$\widetilde{\mathcal{N}}_t^{REQ}$	Set of nodes associated with the candidate requests in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\overline{\mathcal{N}_t^E}$	Set of nodes associated with candidate flight legs that do *not* end at the desired UAT pad of their intended request
$\mathcal{N}_{it}^{SUCC}$	Set of nodes associated with succeeding flight legs of flight $i \subseteq \overline{\mathcal{N}_t^E}$ as of time t
$\widetilde{\mathcal{N}}_{it}^{SUCC}$	Set of nodes associated with succeeding flight legs of flight $i \subseteq \overline{\mathcal{N}_t^E}$ in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\mathcal{N}_t^{FLXSTRT}$	Set of nodes associated with requests with flexible pick-up UAT pad
$\mathcal{N}_t^{FXDSTRT}$	Set of nodes associated with requests with fixed pick-up UAT pad

Arcs

$\mathcal{A}_t$	Arc of the network $\mathcal{G}_t$ as of time t
$\mathcal{A}_t^{INIT}$	*Initial arcs*, which connect aircraft to revenue-generating flight legs, in the network $\mathcal{G}_t$ as of time t
$\mathcal{A}_t^{INITTW}$	Initial arcs, which connect aircraft to revenue-generating flight legs, after considering the time windows as of time t
$\tilde{\mathcal{A}}_t^{INIT}$	Initial arcs, which connect aircraft to revenue-generating flight legs, in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\mathcal{A}_t^{SEQ}$	*Sequencing arcs,* which connect two revenue-generating flight legs in the network $\mathcal{G}_t$ as of time t
$\mathcal{A}_t^{SEQTW}$	Sequencing arcs, which connect two revenue-generating flight legs, after considering the time windows as of time t
$\tilde{\mathcal{A}}_t^{SEQ}$	Sequencing arcs, which connect two revenue-generating flight legs in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\mathcal{A}_t^{ALCT}$	*Allocation arcs,* which allocate requests to revenue-generating flight legs in the network $\mathcal{G}_t$ as of time t

$\tilde{\mathcal{A}}_t^{ALCT}$	Allocation arcs, which allocate requests to revenue-generating flight legs in the reduced network $\tilde{\mathcal{G}}_t$ as of time t
$\mathcal{A}_t^{INTND}$	Set of (r, i), where $r_r = r_i^{INTND}$; $\mathcal{A}_t^{INTND} \subseteq \mathcal{A}_t^{ALCT}$
${\mathcal{A}^0}_t^{FREE}$	Subset of $\mathcal{A}_t^{INIT}$ with no constraints on the aircraft or flight legs
$\mathcal{A}_t^{FREE}$	Subset of $\mathcal{A}_t^{SEQ}$ with no constraints on the flight legs
${\mathcal{A}^0}_t^{PREC}$	Subset of $\mathcal{A}_t^{INIT}$ with preceding constraints
$\mathcal{A}_t^{PREC}$	Subset of $\mathcal{A}_t^{SEQ}$ with preceding constraints
${\mathcal{A}^0}_t^{SUCC}$	Subset of $\mathcal{A}_t^{INIT}$ with preceding constraints
$\mathcal{A}_t^{SUCC}$	Subset of $\mathcal{A}_t^{SEQ}$ with succeeding constraints

Set of Flights for Network Reduction

$\dot{\mathcal{F}}_t^{CAND}$	Adjusted candidate flight legs after excluding the flight legs that could not possibly serve their intended requests
$\ddot{\mathcal{F}}_t^{CAND}$	Adjusted candidate flight legs that are feasible to be served directly by an aircraft or followed by another flight leg given the time windows
$\hat{\mathcal{F}}_t^{SUCC}$	Connecting candidate flight legs with undesired destination UAT pad for which the set of succeeding flight legs becomes empty

Parameters

v^{WALK}	Speed of walking (mph)
v^{DRIVE}	Speed of driving (mph)
Δ^{EMPTY}	The minimum of Euclidean distance between two points to qualify for the empty repositioning flight (miles)
Δ^{OD}	The minimum distance between origin and destination of a request to qualify for a UAT trip (miles)
Δ^{WALK}	The maximum of walking distance (miles)
Δ^{ACCESS}	The radius of the accessible area around the origin or destination of a request for air pooling (miles)
ϵ	Detour factor of the aerial trip
$\mathcal{T}^{ADV}$	The maximum of the advance reservation time window, $T_r^{ADV} \in [0, \mathcal{T}^{ADV}]$
$\mathcal{T}^{INT}$	The mean of interarrival times (seconds)

Simulation

$\mathfrak{T}^{STRTARV}$	Start of arrival time
$\mathfrak{T}^{ENDARV}$	End of arrival time

$\mathfrak{T}^{STRTREQ}$	Start of request time for service
$\mathfrak{T}^{ENDREQ}$	End of request time for service
$\mathfrak{T}^{STRTSIM}$	Start of simulation
$\mathfrak{T}^{ENDSIM}$	End of simulation

Objective Function

$\mathcal{C}$	Fixed operational cost of revenue or empty flight legs
α	Revenue per passenger per mile
β	Operational cost per mile
γ_1	Weight of the relocation in the objective function
γ_2	Exponent of the number of relocations in the objective function
R_r	Earned revenue by serving request r
$\bar{C}^0_{kit}$	Total cost of serving revenue-generating flight leg i as of time t, which includes the preceding empty flight leg from $\boldsymbol{L}^{AVL}_{kt}$ to $\boldsymbol{S}_i$.
$\bar{C}_{ij}$	Total cost of serving revenue-generating flight leg j, including the preceding empty flight leg from $\boldsymbol{E}_i$ to $\boldsymbol{S}_j$.
C^{RLC}_{ri}	Total cost of relocating request r to take flight leg i
χ_{ri}	Number of relocations required for request r to take flight leg i; $\chi_{ri} \in \{0, 1, 2\}$.

Decision Variables

p_i	p_i for $i \in \mathcal{N}^{LEG}_t$ is 1 if flight leg i will be conducted
y_{ki}	y_{ki} for $(k, i) \in \mathcal{A}^{INIT}_t$ is 1 if flight leg i is the revenue-generating flight served by aircraft k immediately from its availability UAT pad as of time t; $y_{ki} \in \{0,1\}$
x_{ij}	x_{ij} for $(i, j) \in \mathcal{A}^{SEQ}_t$ is 1 when revenue-generating flight leg j is served immediately after revenue-generating flight leg i; $x_{ij} \in \{0,1\}$
z_{ri}	z_{ri} for $(r, i) \in \mathcal{A}^{ALCT}_t$ is 1 when request r is assigned to flight leg i
τ^{BOARD}_i	τ^{BOARD}_i for $i \in \mathcal{N}^{LEG}_t$ is the time revenue-generating flight i starts the boarding process

Other Notations

ω	Maximum acceptable delay
ω'	Adjusted maximum acceptable delay

q_{ktn}	n^{th} leg on $\mathcal{Q}_{kt}$
Ω_{ri}^{MIN}	Minimum delay incurred when r_r takes f_i instead of $f_{i^*} = f_r^{DSRD}$