

ADAPTIVE SCHEDULING OF NON-EMERGENCY PATIENT
TRANSFERS WITH WORKLOAD BALANCING CONSTRAINTS:
A MIXED INTEGER PROGRAMMING APPROACH

by

Travis Foster

Submitted in partial fulfillment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
August 2019

© Copyright by Travis Foster, 2019

*This work is dedicated to my family, my partner and the paramedics
who serve communities worldwide.*

Table of Contents

List of Tables	v
List of Figures	vi
Abstract	viii
Acknowledgements	ix
Chapter 1 Introduction	1
1.1 Patient Transfers	1
1.2 Patient Transfers in Nova Scotia	2
1.3 The Patient Transfer Scheduling Process	3
1.4 Patient Transfer Data	6
1.5 Problem Statement	7
Chapter 2 Literature Review	9
2.1 Patient Transfer Review	9
2.2 Scheduling Models	9
2.3 Workload Balancing Review	11
2.4 Data-Driven Research	12
Chapter 3 Methodology	13
3.1 Advance Request Model	13
3.2 Same Day Request Model	16
Chapter 4 Data Analysis	22
4.1 Data Exploration	22
4.1.1 Data Characteristics	22
4.1.2 Descriptive Data Analysis	23
4.2 Travel Time Estimates	29
4.2.1 Missing Data	35
4.3 Service Time Models	36
4.3.1 Service Time Comparison	36
4.3.2 Pickup Service Time Model	39
4.3.3 Delivery Service Time Model	40

Chapter 5	Results	44
5.1	Experimental Design	44
5.2	Advance Model Experiments	45
5.3	Same Day Model Experiments	46
5.4	Value of Information Experiments	47
5.5	Same Day Strategy Experiments	48
5.6	Advance Model Results	50
5.7	Advance Model Comparison to Actual Schedules	52
5.8	Same Day Model Results	54
5.9	Value of Information Results	55
5.10	Same Day Strategy Results	55
Chapter 6	Conclusions	64
Appendix A	Electronic Supplements	69
Bibliography		70

List of Tables

Table 4.1	Pickup travel time metrics	33
Table 4.2	Delivery travel time metrics	34
Table 4.3	Percent of requests for top 10 locations	34
Table 4.4	Top 10 route travel statistics (minutes)	35
Table 4.5	Pickup service time metrics	37
Table 4.6	Delivery service time metrics	38
Table 4.7	Top 10 location service time means (minutes)	39
Table 4.8	Multiple linear regression model for pickup service time	41
Table 4.9	Multiple linear regression model for delivery service time	42
Table 5.1	Experiments summary	50
Table 5.2	Infeasible Advance Model vehicle experiments	51
Table 5.3	Time window experiments with MIP gap larger than 2%	52
Table 5.4	Advance Model computational results summary	53
Table 5.5	Advance Model travel time comparison to actual schedules	53

List of Figures

Figure 1.1	NSHA zones	4
Figure 1.2	Patient transfer scheduling process	6
Figure 3.1	Scheduling model framework	21
Figure 4.1	Patient transfer requests by day of week	24
Figure 4.2	Patient transfer requests by time of day	25
Figure 4.3	Patient transfer requests in the Central Region	26
Figure 4.4	Patient transfer requests assigned to PTUs	27
Figure 4.5	Patient transfer requests assigned to ambulances	28
Figure 4.6	Patient transfer requests by vehicle	29
Figure 4.7	Advance and Same Day patient transfer requests trend	30
Figure 4.8	Advance and Same Day patient transfer requests trend for PTUs	31
Figure 4.9	Pickup travel time distribution	32
Figure 4.10	Delivery travel time distribution	33
Figure 4.11	Pickup service time distribution	37
Figure 4.12	Delivery service time distribution	38
Figure 4.13	Scheduling model framework updated with the models derived from the data	43
Figure 5.1	Same Day Model experiments	47
Figure 5.2	Advance Model results for the number of vehicles	57
Figure 5.3	Advance Model results for workload balancing	57
Figure 5.4	Advance Model results for time windows	58
Figure 5.5	Advance Model computational time	58
Figure 5.6	Same Day Model hard time windows computational time	59
Figure 5.7	Same Day Model hard time windows travel time	59

Figure 5.8	Same Day Model hard time windows changed pickup times . . .	60
Figure 5.9	Same Day Model soft time windows computational time . . .	60
Figure 5.10	Same Day Model soft time windows travel time	61
Figure 5.11	Same Day Model soft time windows changed pickup times . . .	61
Figure 5.12	Perfect information vs. actual information	62
Figure 5.13	Total requests over the course of the day	62
Figure 5.14	Total travel time over the course of the day	63
Figure 5.15	Computational time required for simulated day	63

Abstract

Emergency Medical Services organizations are responsible for providing paramedic crews, vehicles and equipment to transfer patients from one location to another in emergency and non-emergency settings. In the non-emergency setting, they must solve difficult scheduling and assignment problems to ensure on-time arrival of patients and the efficient use of health care resources. The objective of this thesis is to develop a mathematical model that will assign Patient Transfer Units to non-emergency patient transfer requests and design a schedule that will minimize travel costs, balance workloads, and attempt to leave previously scheduled pickup times unchanged. This thesis also utilizes historical patient transfer data in the modelling process. The mathematical model provides decision support for the non-emergency patient transfer scheduling process.

Acknowledgements

Thank you to NSERC and MITACS for their funding which helped immensely as I progressed with this work. Thank you to the faculty of the Industrial Engineering department at Dalhousie University, especially my supervisors (Dr. VanBerkel and Dr. Venkatadri) for their guidance and allowing me to discuss ideas with them. Thank you to the Optimization Department at the Technical University of Delft (especially Dr. van Essen) for graciously welcoming me and providing a wonderful environment to work on my research.

Chapter 1

Introduction

This thesis examines the scheduling and assignment of non-emergency patient transfer requests and the use of historical patient transfer data in the scheduling process.

1.1 Patient Transfers

Non-emergency patient transfers are transportation requests for patients who are not able to transport themselves to receive non-emergency medical treatment. They are commonly referred to as patient transfers, which is the term that will be used for the remainder of this thesis. Patient transfers primarily take place between any of the following locations: a special care facility, a hospital or a personal residence. This on-demand transit system helps fill gaps in current transit infrastructure.

Patients require patient transfers for many reasons. They often cannot drive themselves to a medical appointment and may live in a rural area with limited or no public transportation. They may have special equipment and therefore require extra care. If they are in the care of a medical facility but require a treatment only offered at another location, patient transfer systems ensures the patient receives any required care during transit. Due to these reasons, patient transfer systems are depended upon by public health and safety systems [24].

In some jurisdictions of Canada, patient transfers are conducted by Emergency Medical Service (EMS) providers. They use paramedic crews with ambulances or similar vehicles for patient transfer operations. Increasing transfer volumes add pressure to EMS providers as they must also deal with rising emergency call volumes [38].

The patient transfer process varies across countries and jurisdictions, but the common steps include:

1. Driving an ambulance or similar vehicle to the pickup location.
2. Loading the patient into the vehicle while ensuring all patient care requirements are met.
3. Driving the vehicle to the delivery location.
4. Unloading the patient and bringing them to the correct clinic or residence. This includes transferring the responsibility of care to the receiving facility.

1.2 Patient Transfers in Nova Scotia

In Nova Scotia, Emergency Health Services (EHS) is the organization responsible for EMS and patient transfers. They dispatch crews of paramedics in ambulances and Patient Transfer Units (PTUs) to emergency incidents and patient transfer requests as required. PTUs are very similar to ambulances but do not have all of the same equipment and resources. Therefore, they are not equipped to respond to emergency calls and are used exclusively for patient transfers. Ambulances can and are used for both emergency calls and patient transfers. Emergency calls take priority for ambulances and there are strict rules in place for when an ambulance can help with patient transfer operations. Paramedic crews operating these vehicles in Nova Scotia communicate with the EHS Operations Communications Centre. The Communications Centre receives calls for emergencies and patient transfers and will assign a crew and relay information about the request to the crew, including where to travel to. This is done for both emergency calls and patient transfers.

The Nova Scotia Health Authority (NSHA) is responsible for the majority of health care operations in Nova Scotia. The NSHA has divided coverage into four zones as shown in figure 1.1. Every zone has its own dedicated set of paramedic crews and fleet of ambulances and PTUs. The majority of patient transfers in Nova Scotia take place in the Central Zone. The Central Zone includes the Halifax Regional Municipality (HRM), the largest municipality in the province. All four of the province's tertiary hospitals (The QEII Victoria General site, the QEII Halifax Infirmiry site, the Nova Scotia Hospital and the IWK Health Centre) are located in this region in

addition to eight other hospitals [41]. Patient transfers are often between these four tertiary hospitals, between one of the regional or community hospitals to the tertiary hospitals, or from a personal residence in the HRM to a tertiary hospital.

Many transfers originating in the other zones will involve one of the tertiary hospitals in the Central Zone as well. Due to the rural nature of Nova Scotia, many medical services are only located in the Central Zone. As such, patients in the other zones often require patient transfers to the Central Zone for medical care. The Central Zone also has the highest volume of emergency calls of all of the zones. The number of emergency and non-emergency calls have been increasing over the last few years. Therefore, not only is the demand for patient transfers increasing, but ambulances are busier than ever. This is discussed further in chapter 4.

Overcrowding in emergency departments resulting in ambulance offload delay [25] is another factor impacting ambulance availability. While the biggest impact from offload delay are emergency calls at risk of not receiving timely care, another consequence is that there are not enough resources (paramedic crews and vehicles) available to serve the patient transfer demand. As such, the ambulances in the Central Zone are often too busy to help with patient transfer operations. Emergency department overcrowding and long wait times may also be caused by non-emergency patients. If the department closes or is too overcrowded, there may be a surge of patient transfer requests [32]. The PTU crews in the Central Zone are often the only resources available for patient transfers and as such have a high utilization.

These factors have led to increased stress on the patient transfer system, especially in the Central Zone. In light of this information, this thesis focuses on the patient transfer system in the Central Zone.

1.3 The Patient Transfer Scheduling Process

The EHS Operations Communications Centre receives patient transfer requests via phone calls. Additional information may be faxed or emailed. Requests can be received from the pickup facility or from the patient themselves (in case of a pickup from



Figure 1.1: NSHA zones

a personal residence). The requester asks for a specific pickup time. The Communications Centre considers the volume of requests, the location of the pickup and PTU availability among other factors before deciding to agree to the requested pickup time. They may suggest a slightly modified (either earlier or later) time if the Communications Centre expects all of the PTUs to be busy at the requested time of pickup. If there is no acceptable time for both parties, the patient transfer may be delayed or cancelled, resulting in delayed service for the patient.

Once a pickup time has been agreed upon, the request information is entered into a database. This agreed upon pickup time may be changed once by EHS should the volume of requests increase and changing a pickup time will benefit the patient transfer system; it cannot be changed more than once. If the new time is not acceptable to the requesting party, a transfer may be delayed or cancelled, resulting in delayed service to the patient. EHS considers changing the pickup time a last resort and uses it only when the situation absolutely requires it (such as an unexpected delay during

another patient transfer). When the time for a pickup draws closer, the Communications Centre will dispatch a vehicle and crew to the pickup location.

From a scheduling and assignment perspective, patient transfers present a unique challenge to EMS providers and Operations Research (OR) scientists. Patient transfers are scheduled according to:

- The time the transfer request was placed. Some requests are placed the day prior or earlier while others are placed a few hours before the requested pickup time.
- The requested time of pickup.
- The estimated availability of PTUs and ambulances in the region at the requested time of pickup.
- Logistical issues surrounding the transfer such as equipment required and future pickups.
- Ensuring the work is distributed relatively evenly between the PTU paramedic crews so that no crew is overworked.
- The anticipated location of PTUs at the time of pickup. The Communications Centre attempts to schedule requests so that PTUs end up near the next request.

Requests that are placed the day prior (or earlier) to the requested pickup time are considered “advance notice requests” or simply Advance requests. Every evening, a preliminary schedule for the Advance requests that will occur the following day is created at the Communications Centre. This schedule is created by EHS Operations staff and uses all known information about the requests and domain expert knowledge. Requests that are placed on the same day as the requested pickup time are called “Same Day requests”. As Same Day requests are received, the Communications Centre adjusts the schedule as best as they can.

Except in extraordinary circumstances, patient transfer requests are treated as “first

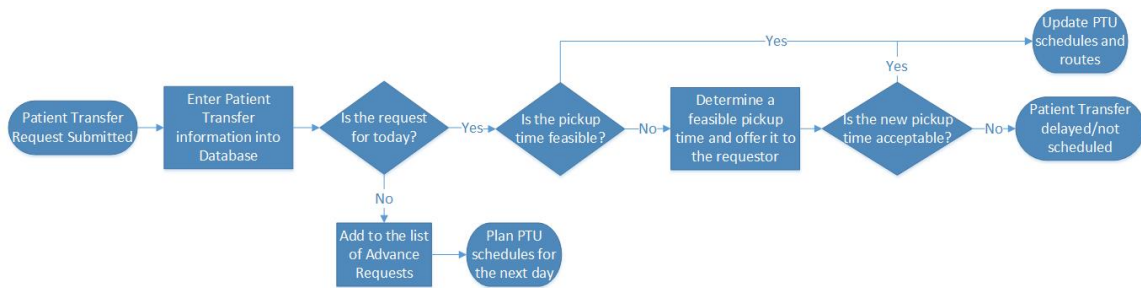


Figure 1.2: Patient transfer scheduling process

come, first serve.” If a Same Day request has a pickup time where the Communications Centre expects all of the PTUs will be busy with other requests (and they assume no ambulances will be available), the requester will be informed of the next available pickup time. Existing requests in the system will not be “bumped” or delayed for new patient transfer requests. The scheduling process can be seen in figure 1.2.

Within this scheduling and assignment procedure, there is an opportunity to improve scheduling efficiency and free Communications Centre Resources to help coordinate patient transfer operations in the other zones in Nova Scotia with the help of Operations Research. Scheduling models are one avenue to improve this process.

1.4 Patient Transfer Data

Similar to numerous organizations in the 21st century [18], EHS collects and stores data, including from its vehicles during operations. Information such as pickup location, destination, PTU identification, when a PTU arrives at the pickup location, how long is spent with the patient before departing for the patient transfer destination, special equipment required for the patient and more is captured for every patient transfer. The EHS Communications Centre receives near real time updates on the location and status of each vehicle.

Organizations are beginning to figure out ways to use their data as part of research projects and in tools to help with decision support. There is, however, still a gap between collecting the data and using it. In business and research, there is a need for

trained personnel to find ways to transfer the information found in data to decisions. The information collected by EHS is useful for evaluating their operations and analyzing the operation from a high-level. However, this data could be doing more to help with the schedule planning that takes place every evening and day, during the offline and online operational phases [15]. This thesis aims to make use of EHS data in the scheduling process and a scheduling model.

1.5 Problem Statement

This thesis focuses on scheduling patient transfer requests efficiently. Research on non-emergency patient transfers and vehicle routing and scheduling models are explored in order to understand the considerations for an intelligent scheduling model that takes the current EHS system into account. Recent research on workload balancing and using data driven methods in vehicle routing and scheduling models are also explored. The final model is constructed in a two stage approach. The first stage is a model that schedules Advance requests to minimize total travel time. The purpose of this model is to run it every evening with the Advance requests to create a preliminary schedule, similar to what EHS currently does. The second stage is a model that creates a new schedule and attempts to minimize the number of changed pickup times from the previous schedule (the PTU executing the transfer can change, however). This model would be run during the online operational phase of EHS Operations.

Statistical models were also explored to help inform the model. Past patient transfer data was explored and analyzed to help make modelling decisions. Travel and service time estimates are generated from models using the historical patient transfer request data. Where travel data is not available, Google Maps API and the googleway R package [4] are used to generate travel time estimates. These are fed to both steps of the scheduling model as inputs.

Experiments are performed on the model. Real requests from the historical patient transfer data are used to test how changing the parameters affects the run time and

results. Total travel time is also examined, and the difference between perfect information (all requests known) and actual information (splitting requests into Advance and Same Day requests) is analyzed. The difference in travel times between the step one model and the actual models are observed for how much improvement may be gained. A strategy for when to run the model in the second step is also tested and evaluated with the historical data.

The remainder of this thesis is organized as follows: Chapter 2 presents a literature review on patient transfers, scheduling models, workload balancing in vehicle routing problems and data-driven research in scheduling models. Chapter 3 presents the formulation of the two stage scheduling model. Chapter 4 presents descriptive data analysis of the patient transfer data and methods for estimating the travel time and service time parameters for the scheduling model. Chapter 5 presents the experiments performed on the final scheduling model and the results. Chapter 6 presents the conclusions of the thesis and ideas for future work.

Chapter 2

Literature Review

In this chapter, relevant literature for the thesis is reviewed. Topics covered include patient transfers, scheduling models applied to non-emergency patient transfers, workload balancing constraints applied in Vehicle Routing Problems (VRPs) and literature concerning uses of data used for non-emergency patient transfer systems.

2.1 Patient Transfer Review

Both inter-hospital and intra-hospital transfers have been studied, although focus has generally been on inter-hospital transfers. Kulshrestha and Singh [24] conducted a review of recent literature on emergency and non-emergency patient transfers. The literature has not settled on a strict definition of a non-emergency patient transfer [20, 38, 23], but possible reasons for a transfer may include specialty care, non-availability of beds or funding for medical treatment [45]. Pre-transfer stabilization has been noted as a key step in the patient transfer process [26]. Hains et al. [13] reviewed literature specifically for non-emergency transfers and found that efficiency and communication are key factors for high quality patient transfers. Even though non-emergency transfers are not “critically” ill, these patients often have the same needs as emergency patients [7].

2.2 Scheduling Models

Patient transfer systems are a common part of healthcare systems and as such, scheduling patient transfers has been studied extensively. These “Dial A Ride” transportation systems have been in use since the 1970s [33]. The first attempts at analytical investigations into scheduling a Dial A Ride system were also made in the 1970s [42]. Analytic solutions to these problems such as dynamic programming algorithms [35] were used to solve the “Dial A Ride Problem” (DARP) during the planning phase

of these transportation systems.

The DARP is a type of Vehicle Routing Problem (VRP) and part of the Traveling Salesman group of problems. The DARP will design vehicle routes and schedules for n requests divided among k vehicles. It is a difficult problem to solve because it is combinatorially difficult and a solution is required in real time. Recent reviews of the DARP in literature include Cordeau and Laporte [6] and Ho et al.[19]. Cordeau also introduced a formulation of the DARP with time windows (also known as the Pickup and Delivery Problem with Time Windows (PDPTW)) that focused on minimizing the total cost of the vehicle routes, and a branch and cut algorithm for solving it [5]. This algorithm proved feasible for small to medium sized problems. The DARP has many different design choices and these are reviewed, although focus is placed on research concerning the PDPTW as EHS must agree to pickup times with the patients.

The most common objective of the DARP is to minimize the total cost of the routes. The cost may be total travel time, total travel distance or driver working time. Some DARP objective functions focus on user inconvenience, such as total ride time and total wait time [19]. Guerriero et al [11] investigated a multi-objective DARP where both maximum total ride time and total waiting time were optimized. They solved the problem using a two-step approach of meta-heuristics and a set-partitioning formulation. Urrea et. al [43] used a weighted multi-objective function with several factors including travel time, ride time and route duration. Berg and Essen [44] looked at the problem of scheduling non-emergency patient transfers where resources are shared between non-emergency and emergency operations in order to minimize the impact on emergency operations.

It is often assumed in the DARP that the vehicles are located at a single depot; that is, every vehicle route starts and ends at the same location. Detti et al. [8] formulated a Mixed Integer Programming (MIP) model for a multi-depot DARP. They investigated Variable Neighborhood Search and Tabu Search as methods for solving the problem, and presented computational results on random instances generated

from a large real-world problem. Patients and vehicles both may be considered homogeneous or heterogeneous. Parragh [34] investigated a DARP with heterogeneous patients and vehicles and solved the problem with direct methods and heuristics.

The previous papers focused on deterministic DARPs, or problems where the information is known with certainty. However, certain information may be stochastic. Hyyti et al. [21] investigate a stochastic DARP model with a single vehicle. VRPs may also be dynamic; that is, new information may be received over time and fed into the model. Dynamic vehicle routing and dispatching has been researched since the 1970s. Psaraftis et al. [36] present a review on the topic. Ichoua, Gendreau and Potvin [22] used knowledge of future demands to plan vehicle dispatching in real time. Hill et al. [14] evaluated a dynamic DARP system using simulation. Markovi et al. [28] used a greedy insertion algorithm to solve the dynamic DARP in a real-world setting. The literature on the dynamic DARP primarily focuses on new requests triggering new vehicle routes. However, Beaudry et al. [2] presented a VRP model with other considerations for adjusting the vehicle routes such as lengthy patient delays.

2.3 Workload Balancing Review

Workload balancing in VRPs is often applied to keep workers happy and ensure no worker is overworked. Matl et al. [29] reviewed the recent literature on workload balancing in VRPs. Workload balancing can be applied in the objective function or as a constraint. The measure of workload balancing is commonly the total time of the routes but can also be demand served by each vehicle. The literature also defines two primary methods of measure workload balancing, either through min-max [3] or a range of acceptable workloads [1]. Mourgaya and Vanderbeck [30] used a workload balancing constraint on a cluster of delivery points in a VRP. Goodson [9] used a constraint for the range of the workloads in a specialized VRP for election days. Gror et al. [10] used constraints to enforce balanced workloads in a VRP.

2.4 Data-Driven Research

Using data in conjunction with VRPs is not new, but recent trends in data collection, computing power and statistical modelling techniques does open new avenues for organization to use their data, and for operations researchers to apply their skills and tools. Yalnda et al. [46] used historical data to estimate travel times of home care givers in a home health care problem. Markovi et al. [27] used machine learning algorithms during the planning phase of a new DARP system to estimate the capacity required. Saadi et al. [40] used various machine learning methods to forecast short-term (time interval of ten minutes) spatio-temporal demand for a ride-hailing service. Nazari et al. [31] used deep reinforcement learning to solve a VRP problem with instances sampled from a random distribution. While these examples are promising, there is still opportunity to use data and apply statistical and machine learning models to the DARP and VRPs, especially in real world case studies.

Chapter 3

Methodology

In this section, the approach and assumptions made when modelling the scheduling process with a DARP model are reviewed. This approach is used to mimic the actual scheduling process used by EHS. The model is conducted in two stages as described below.

- A preliminary schedule is created in the evening for all patient transfer requests submitted in advance that are scheduled for the following day.
- An updated schedule is created during the online operational phase where new routes are created with Advance and Same Day requests. The scheduled pickup times from the previous solution is included as an input and the model attempts to keep these pickup times unchanged.

The model assumes a single depot to act as the start and end point for every vehicle. It also assumes that the fleet of vehicles and the patients are homogeneous. In reality, patients do have different needs but this has little impact on the time required for a request. Every PTU is equipped with the same resources and therefore are able to handle all patient transfer requests.

3.1 Advance Request Model

The Advance request model is an adaptation of the deterministic, static DARP model in [5] which includes time windows. All of the non-linear constraints in the DARP model are linearized. The model creates a set of routes for k vehicles and n requests while minimizing the travel time across all routes. The vehicles represent the PTUs. The DARP model is based on of the three-index formulation found in Cordeau [5]. However, the Advance Model includes workload balancing constraints and variable shift times for the PTU crews. The time windows are used to ensure patients arrive

at their destination in a timely manner. Each request for transportation includes a pickup and a delivery.

Let n represent the number of patient transfer requests. The DARP is formulated on a directed graph $G = (V, A)$. Each request from n has a pickup node and a delivery node represented by i and $n + i$ respectively. These nodes must be visited in order and by the same vehicle. The pickup nodes are represented by the subset $P = (1, \dots, n)$ and the delivery nodes are represented by the subset $D = (n + 1, \dots, 2n)$. Therefore, $P \cup D$ represents every location that must be visited by a vehicle. The depot is represented by nodes 0 and $2n + 1$. These three indices make up the vertex set $V = (0, 1, \dots, n, n + 1, \dots, 2n, 2n + 1)$.

K represents the set of vehicles. Each vehicle $k \in K$ has a capacity of Q_k , a minimum shift start time of $Tmin_k$ and a maximum shift end time of $Tmax_k$. Each node $i \in V$ has a non-negative service time d_i and a load q_i such that $q_{n+i} = -q_i$. These values for the depot are such that $d_0 = d_{2n+1} = q_0 = q_{2n+1} = 0$. Each node $i \in V$ has a time window $[e_i, l_i]$ where e_i and l_i are the earliest and latest times that service may begin at node i . Each arc $(i, j) \in A$ has an associated travel time t_{ij} . This travel time also acts as the cost measure, which is discussed in section 4.1. Parameters wb^+ and wb^- are used as the maximum and minimum workload, respectively, for each vehicle.

The model has three types of decision variables:

- For every arc $(i, j) \in A$ and vehicle $k \in K$, x_{ij}^k is a binary decision variable that is 1 if vehicle k will traverse the arc from node i to node j and is 0 otherwise.
- For every node $i \in V$ and vehicle $k \in K$, u_i^k represents the start time of service by vehicle k at node i .
- For every node $i \in V$ and vehicle $k \in K$, f_i^k represents the number of patients in vehicle k after visiting node i .

The MIP formulation is (3.1) – (3.19):

$$Min \sum_{k \in K} \sum_{i \in V} \sum_{j \in V} t_{ij} x_{ij}^k \quad (3.1)$$

$$\sum_{k \in K} \sum_{j \in V} x_{ij}^k = 1 \quad \forall (i \in P) \quad (3.2)$$

$$\sum_{i \in V} x_{0i}^k = \sum_{i \in V} x_{i,2n+1}^k = 1 \quad \forall (k \in K) \quad (3.3)$$

$$\sum_{j \in V} x_{ij}^k - \sum_{j \in V} x_{n+i,j}^k = 0 \quad \forall (i \in P, k \in K) \quad (3.4)$$

$$\sum_{j \in V} x_{ji}^k - \sum_{j \in V} x_{ij}^k = 0 \quad \forall (i \in P \cup D, k \in K) \quad (3.5)$$

Equation (3.1) is the objective function where the total travel time of the routes is minimized. Constraints (3.2) ensures each request is served once. Constraints (3.3) ensures that every vehicle route begins and ends at the depot. Constraints (3.4) ensures the pickup and delivery nodes of a request are served by the same vehicle. Constraints (3.5) certifies that the vehicle that enters node i will also depart from node i .

$$u_j^k \geq u_i^k + d_i + t_{ij} - M(1 - x_{ij}^k) \quad \forall (i, j \in V, k \in K) \quad (3.6)$$

$$u_{i+n}^k \geq u_i^k + d_i + t_{i,i+n} - M(1 - x_{i,i+n}^k) \quad \forall (i \in P, k \in K) \quad (3.7)$$

$$u_{2n+1}^k \leq T max_k \quad \forall (k \in K) \quad (3.8)$$

$$u_0^k \geq T min_k \quad \forall (k \in K) \quad (3.9)$$

$$e_i \leq u_i^k \leq l_i \quad \forall (i \in V, k \in K) \quad (3.10)$$

Constraints (3.6) and (3.7) guarantee that service time is consistent among every node, and that the service time at node $i + n$ does not begin until after the service at node i is completed. Note that Constraints (3.7), valid inequalities, only work when the vehicle capacity is 1 and the patient delivery must be completed before the next patient is picked up, as applicable to the problem in this thesis. M is a large constant. Constraints (3.8) and (3.9) uphold the vehicle and crew shift start and end times and are similar to the constraints in the DARP model of [5] which place a limit on vehicle route length and user ride time. Constraints (3.10) ensures that a request is completed during its time window.

$$f_j^k \geq f_i^k + q_j - M(1 - x_{ij}^k) \quad \forall (i, j \in V, k \in K) \quad (3.11)$$

$$f_j^k \leq f_i^k + q_j + M(1 - x_{ij}^k) \quad \forall (i, j \in V, k \in K) \quad (3.12)$$

$$f_i^k \leq Q_k \quad \forall (i \in V, k \in K) \quad (3.13)$$

$$f_i^k \leq Q_k + q_i \quad \forall (i \in V, k \in K) \quad (3.14)$$

$$f_i^k \geq 0 \quad \forall (i \in V, k \in K) \quad (3.15)$$

$$f_i^k \geq q_i \quad \forall (i \in V, k \in K) \quad (3.16)$$

$$\sum_{i \in K} \sum_{j \in V} (t_{ij} + d_i) x_{ij}^k \leq wb^+ \quad \forall (k \in K) \quad (3.17)$$

$$\sum_{i \in K} \sum_{j \in V} (t_{ij} + d_i) x_{ij}^k \geq wb^- \quad \forall (k \in K) \quad (3.18)$$

$$x_{ij}^k \in (0, 1) \quad \forall (i \in V, j \in V, k \in K) \quad (3.19)$$

Constraints (3.11) - (3.16) ensures consistency for the passenger load in every vehicle and every node. Constraints (3.17) and (3.18) are the workload balancing constraints. Workload is measured as the sum of the travel time and service time along a vehicle's route.

3.2 Same Day Request Model

The Same Day Model is structured similarly to the Advance Model, but has a different objective. The Same Day Model is designed to be run during the online operational period [15] and attempts to schedule all Advance requests and (known at the time) Same Day requests while not changing any previously scheduled pickup times. More specifically, the model will minimize the number of significantly changed pickup times. "Significantly changed" will be defined later in this section. The idea behind the model is to use previously scheduled patient transfer pickups to influence the new schedule. An agreement on a pickup affects the resources available and service provided in the future. As described in section 1.2, EHS attempts to avoid changing any agreed upon pickup times. Delaying a pickup time may result in diminished medical service for the patient in the form of a delayed or cancelled appointment.

In addition to avoiding delayed pickups, the Same Day Model attempts to avoid

“advancing” pickup times as well. Advancing is defined as changing the pickup time to a significantly earlier time. Many transfer pickups are located in busy hospitals. Hospital staff must prepare a patient for pickup and will often plan around an expected pickup time. Arriving for a pickup early and unannounced in these situations is not beneficial for the hospital staff and the paramedics as the patient will not be ready for transport. Calling a hospital ahead of time may alleviate this problem but the hospital staff may be too busy to accommodate the early pickup request. A similar problem may occur at the delivery location. Staff may be unprepared or may not have a bed available to accommodate an earlier delivery, resulting in wasted time for the paramedic crew and the vehicle.

The number of patient transfer requests are represented by n , and each request has a pickup node and a delivery node represented by i and $n + i$ respectively. The indices for the pickup nodes, delivery nodes and the vehicle depot are the same as in section 3.1. The sets and parameters described in section 3.1 remain in the Same Day Model. The routing constraints (3.2) – (3.5), the service time constraints (3.6 – 3.10), the load constraints (3.11) – (3.18) and the workload balancing constraints (3.18) – (3.19) are unchanged.

The three decision variables from the Advance Request Model remain in the Same Day Model, but three new decision variables are introduced. The first two are listed below.

- For every node $i \in P$, Y_i is a binary decision variable that is 1 if the service start time at node i is delayed by more than α^+ minutes and is 0 otherwise.
- For every node $i \in P$, Z_i is a binary decision variable that is 1 if the service start time at node i is advanced by more than α^- minutes and is 0 otherwise.

The service start times, u_i^k , generated by the Advance Model for every Advance request pickup are used as an input to the Same Day Model. Only the service start times where vehicle k visited node i are kept. They are represented by a new parameter, w_i . For any requests that have a previously scheduled pickup time, w_i is the previously agreed upon pickup time. For any new requests, w_i is the requested pickup time.

The Same Day Model has four global parameters. They are as follows:

- α^+ is the amount of time a pickup may be delayed before it is considered significantly changed (late) at any pickup node $i \in P$.
- α^- is the amount of time a pickup may be advanced before it is considered significantly changed (early) at any pickup node $i \in P$.
- β^+ is the maximum amount of time a pickup may be delayed at any pickup node $i \in P$.
- β^- is the maximum amount of time a pickup may be advanced at any pickup node $i \in P$.

The objective function for the Same Day MIP Model is (3.20).

$$\text{Min} \sum_{i \in P} Y_i + \sum_{i \in P} Z_i \quad (3.20)$$

For every node $i \in P$, the third new decision variable, b_i , is used to measure the difference between the previous scheduled pickup time, w_i , and the new pickup time, u_i^k . The following constraint is introduced to measure this difference.

$$\sum_{j \in V} x_{ij}^k u_i^k - \sum_{j \in V} x_{ij}^k w_i = b_i \quad \forall (i \in P, k \in K) \quad (3.21)$$

Constraints (3.21) is non-linear, therefore $x_{ij}^k u_i^k$ is replaced with s_{ij}^k , a continuous decision variable with bounds of $[S_{low}, S_{high}]$. Constraints (3.21) is replaced by constraints (3.22-3.26).

$$\sum_{j \in V} s_{ij}^k - \sum_{j \in V} x_{ij}^k w_i = b_i \quad \forall (i \in P, k \in K) \quad (3.22)$$

$$s_{ij}^k \geq x_{ij}^k * S_{low} \quad \forall (i \in P, j \in V, k \in K) \quad (3.23)$$

$$s_{ij}^k \leq x_{ij}^k * S_{high} \quad \forall (i \in P, j \in V, k \in K) \quad (3.24)$$

$$s_{ij}^k \leq u_i^k - S_{low}(1 - x_{ij}^k) \quad \forall (i \in P, j \in V, k \in K) \quad (3.25)$$

$$s_{ij}^k \geq u_i^k - S_{high}(1 - x_{ij}^k) \quad \forall (i \in P, j \in V, k \in K) \quad (3.26)$$

S_{low} and S_{high} should be picked so that s_{ij}^k may only take a value between the earliest possible service start time and the latest possible service start time. In the Same Day Model, these are the earliest shift start time and the latest shift end time of the vehicles. This is shown in equations (3.27) and (3.28).

$$S_{low} = \min_{(k \in K)} (Tmin_k) \quad (3.27)$$

$$S_{high} = \max_{(k \in K)} (Tmax_k) \quad (3.28)$$

Two of the global parameters introduced earlier, β^+ and β^- , are used to impose a maximum amount of time a pickup time may be changed. Constraints (3.29) and (3.30) show this.

$$b_i \geq \beta^- \quad \forall (i \in P) \quad (3.29)$$

$$b_i \leq \beta^+ \quad \forall (i \in P) \quad (3.30)$$

Constraints (3.31)-(3.32) dictate the relationship between b_i and the new decision variables, Y_i and Z_i . If a pickup time has been delayed by greater than α^+ , than Y_i will be equal to 1. If a pickup time has been advanced by greater than α^- , than Z_i will be equal to 1.

$$G * Y_i \geq b_i - \alpha^+ \quad \forall (i \in P) \quad (3.31)$$

$$(-G) * Z_i \leq b_i + \alpha^- \quad \forall (i \in P) \quad (3.32)$$

$$Y_i \in (0, 1) \quad \forall (i \in P) \quad (3.33)$$

$$Z_i \in (0, 1) \quad \forall (i \in P) \quad (3.34)$$

G is a sufficiently large constant. It should be larger than the largest possible deviance between b_i and α^+ or α^- . This can be determined by equation (3.35).

$$G \geq \max(\beta^+ - \alpha^+, |\beta^- + \alpha^-|) \quad (3.35)$$

The relationship between the four new global parameters can be summarized in equation (3.36).

$$\beta^- \leq \alpha^- \leq 0 \leq \alpha^+ \leq \beta^+ \quad (3.36)$$

In the Same Day Model, α^+ and α^- act as *soft time windows*. The Same Day Model attempts to keep every pickup time within this soft time window. If necessary, a pickup time can be changed past the soft time windows but not past β^+ and β^- ; these act as *hard time windows*.

The schedule generated by the Same Day Model can also be used as an input to re-run the Same Day Model, replacing the input from the Advance Model. As discussed previously, when Same Day requests are received, the Same Day Model can be run with the previous solution service start times (u_i^k) as the w_i parameter. This process can be repeated as the Same Day model also generates service start times. The Same Day Model schedule can therefore replace the Advance Model schedule as an input and be rerun over the course of the online operational phase.

The Same Day Model is used to generate new vehicle routes while attempting to maintain the scheduled pickup times. The process of generating a preliminary schedule and using that schedule as an input to the Same Day Model is shown in figure 3.1.

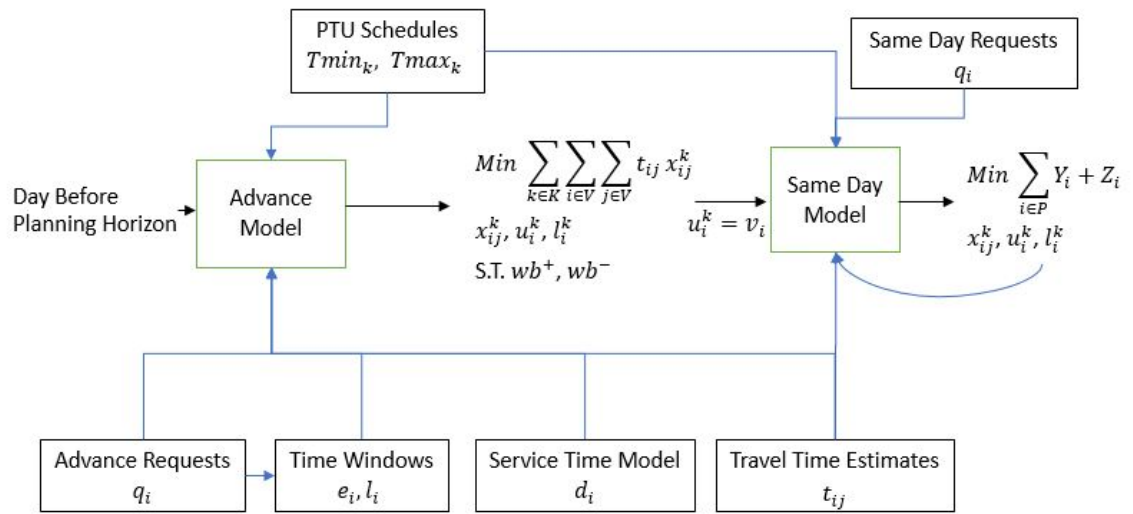


Figure 3.1: Scheduling model framework

Chapter 4

Data Analysis

In this chapter, analysis performed on patient transfer data provided by the EHS Operations team will be explored. This data is used to compute certain parameters in the Advance and Same Day Models. This analysis also helps with decisions for experiments to test the models with in chapter 5. Note that a lot of the charts do not have labelled y axis. This is to protect some information for EHS. The trends shown are what is important on the charts.

4.1 Data Exploration

4.1.1 Data Characteristics

Data for every patient transfer request in 2016 and 2017 in the Central Region of Nova Scotia was provided by the EHS Operations Team. The data includes requests that were not completed. Descriptive information such as which PTU was assigned to the transfer, pickup location and destination are included.

Sensors on the PTUs send continuous updates on the vehicle’s status. Transfer requests are assigned to a PTU by EHS Operation staff. This is done over the radio, but staff also flag the patient transfer in their computer system. When this event occurs, the PTU is considered “Enroute” to the pickup location. Once the PTU reaches the pickup location, the PTU status is changed to “At Scene”. When a PTU leaves a pickup location, the status changes to “Depart Scene”. Once it has arrived to the delivery location, the status becomes “At Destination”. The status is updated a final time to “Available” once the patient transfer is complete.

The system saves the timestamp where every status change occurs. This allows for metrics of interest to be calculated. These metrics are as follows.

- The Pickup Travel Time is estimated from the difference between the Enroute and At Scene timestamps.
- The Pickup Service Time is estimated from the difference between the At Scene and Depart Scene timestamps.
- The Delivery Travel Time is estimated from the difference between the Depart Scene and At Destination timestamps.
- The Delivery Service Time is estimated from the difference between the At Destination and Available timestamps.

These metrics are important to the models discussed in chapter 3 and are examined in sections 4.2 and 4.3.

4.1.2 Descriptive Data Analysis

Exploratory analysis of the patient transfer data is presented. First, the patient transfer demand is analyzed. Figure 4.1 shows the sum of all patient transfer demands by the day of the week. Labels 1 and 7 represent Sunday and Saturday, respectively. As patient transfer volume is low on the weekend, the weekday was focused upon.

Figure 4.2 shows the patient transfer demand by the time of day, with each bin representing one hour of the day. The data was graphed according to the agreed upon pickup time. Patient transfer volume begins increasing at 6.00 and rises sharply until 13.00. There is a steady decrease until 18.00 where demand remains relatively constant until the end of the day.

The time period that takes place between 6.00 and 19.00 during a week day is defined as the “peak demand period” and is the focus of the data analysis. Demand for the 24 month time period the data was collected from was plotted in figure 4.3. This chart shows the number of patient transfer requests in the Central Region with a trend line. The trend line shows a downward trend in patient transfers for the period of 2016 through 2017.

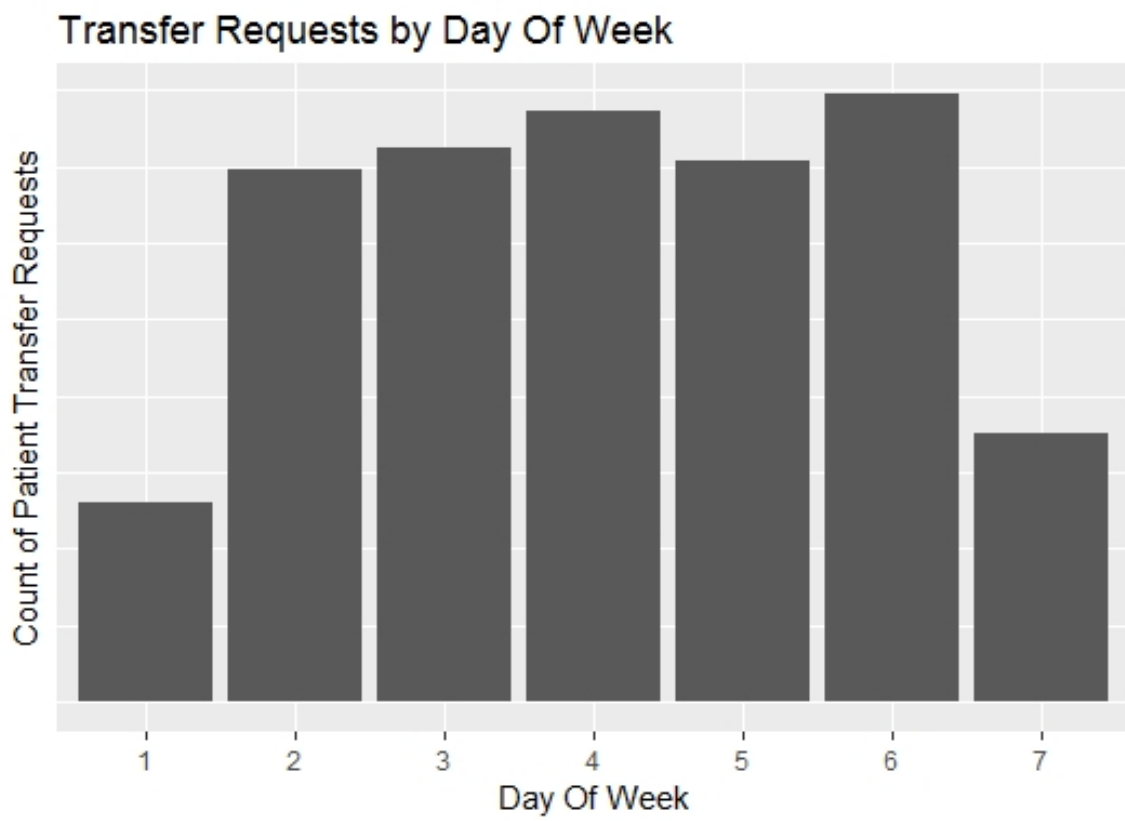


Figure 4.1: Patient transfer requests by day of week

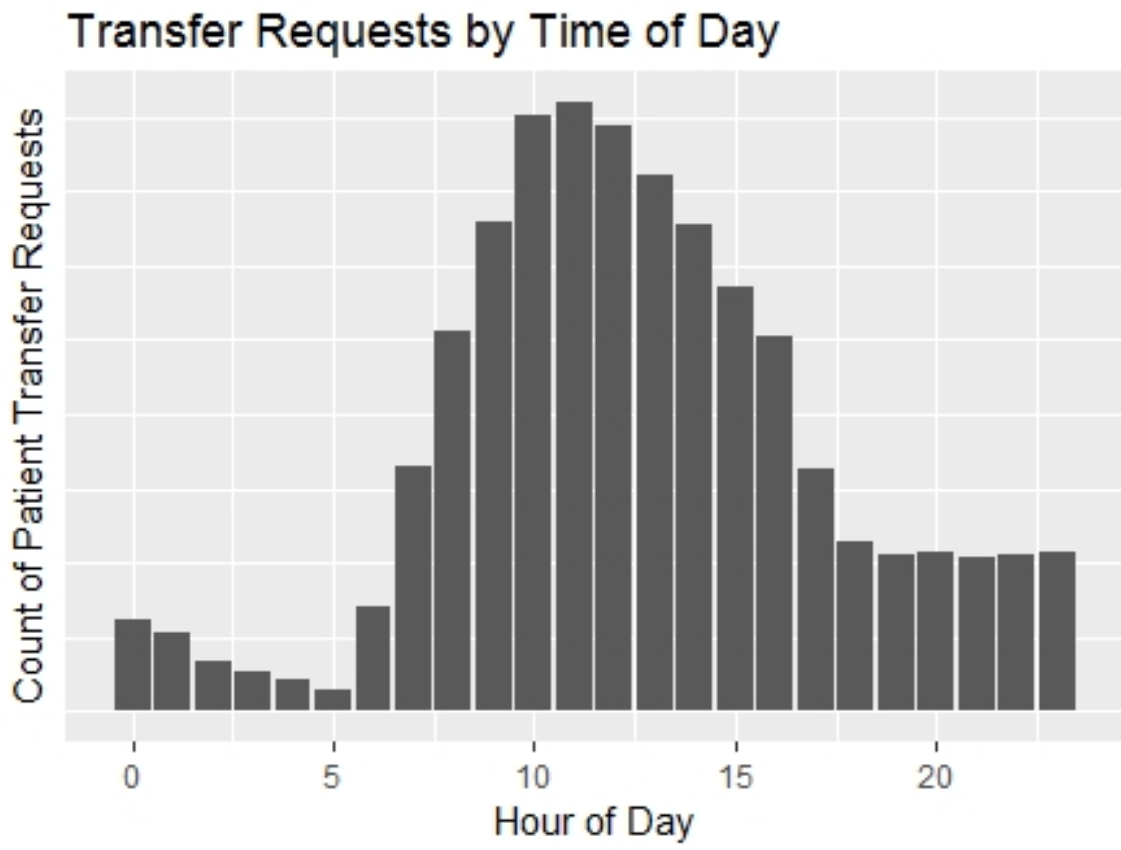


Figure 4.2: Patient transfer requests by time of day

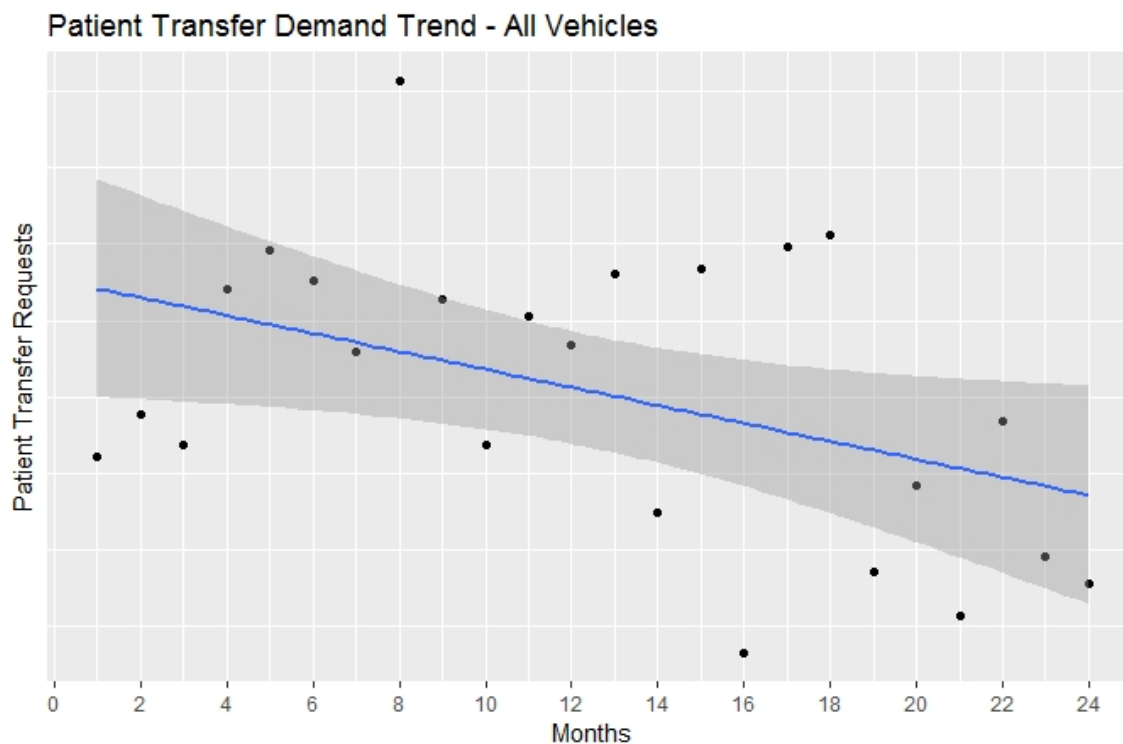


Figure 4.3: Patient transfer requests in the Central Region

The patient transfer demand was drilled into to determine any causes for the decrease in patient transfers. Figures 4.4 and 4.5 show the number of patient transfers that are assigned to PTUs and ambulances, respectively, in the Central Region per month. Figure 4.4 shows a clear upward trend in the data whereas figure 4.5 has a decreasing trend. This is evidence supporting the notion that ambulances are helping less with patient transfer operations (likely for reasons discussed in chapter 1), and the PTUs must make up for the lack of service previously provided by ambulances and their paramedic crews.

Demand was split into two categories: Advance and Same Day. These terms were previously used in chapter 3 and the same definitions apply here. Figure 4.7 shows the trend for Advance and Same Day requests across all vehicles. The charts show a small increasing trend for Advance Requests (with a large amount of variance) and a clear decreasing trend for Same Day requests.

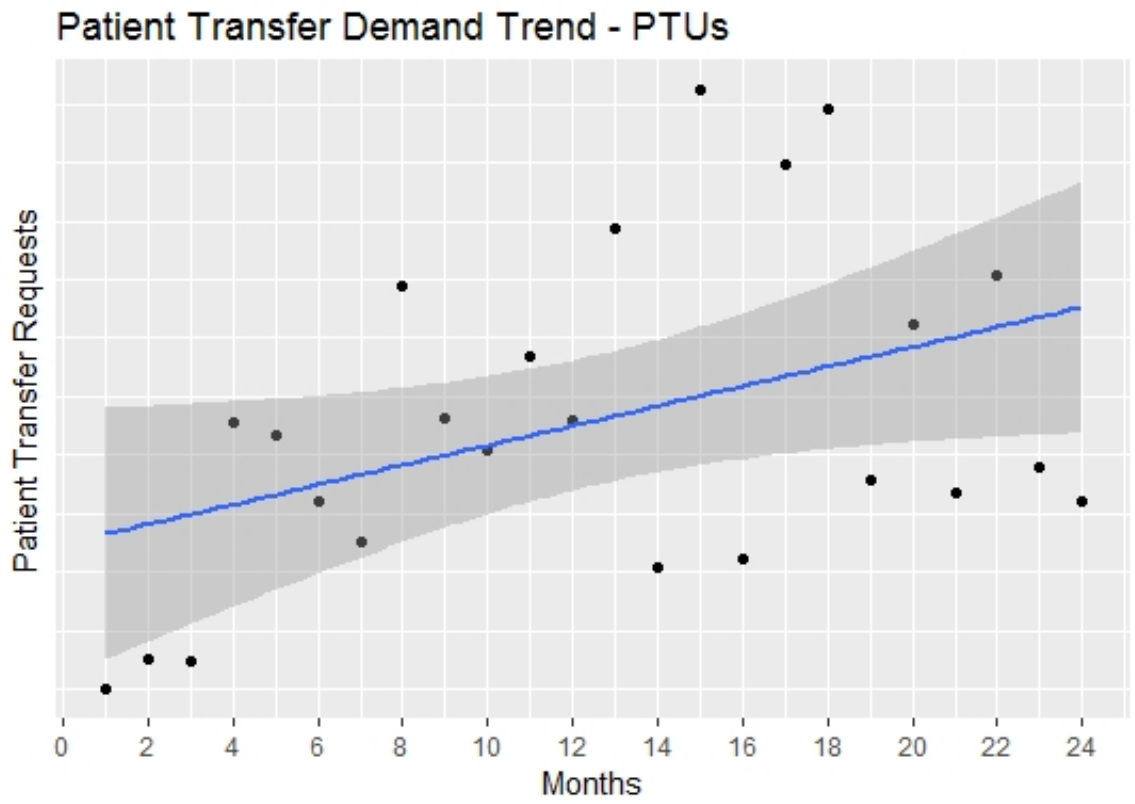


Figure 4.4: Patient transfer requests assigned to PTUs

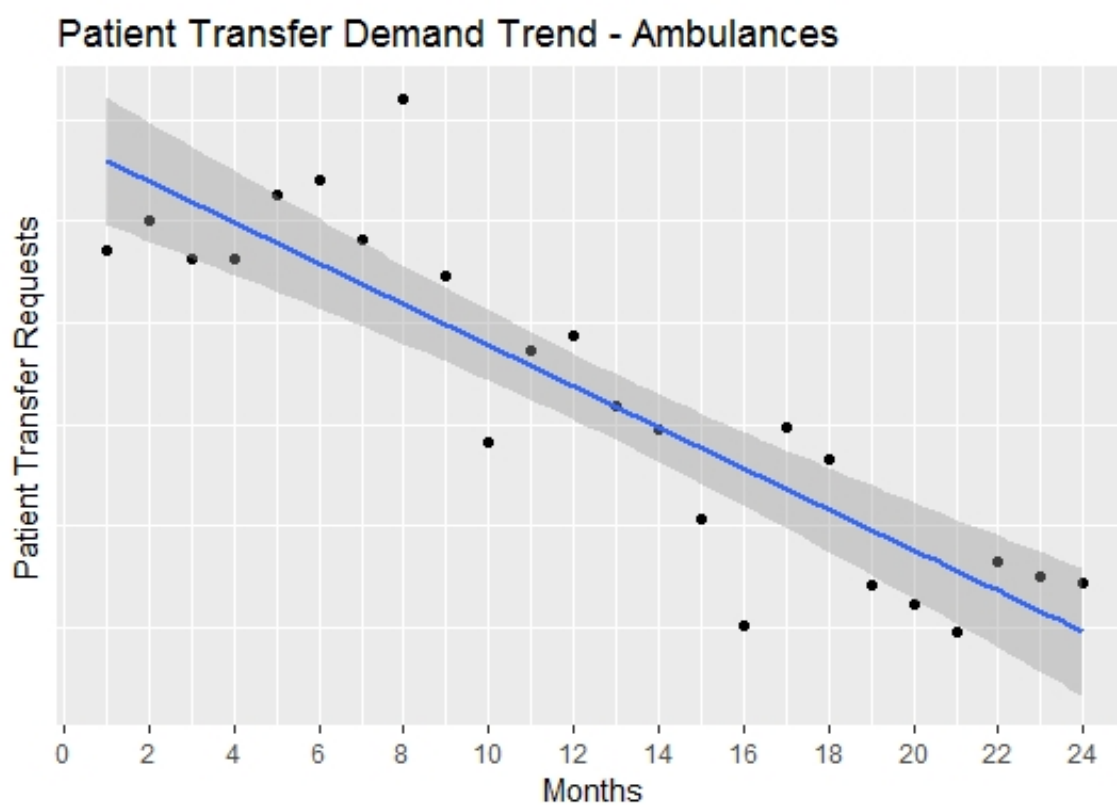


Figure 4.5: Patient transfer requests assigned to ambulances

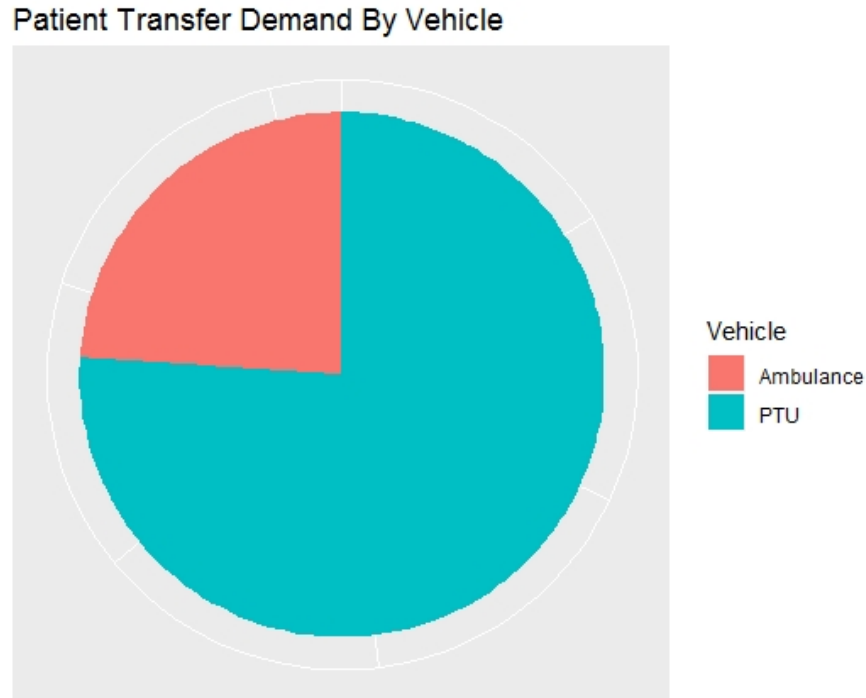


Figure 4.6: Patient transfer requests by vehicle

Digging into requests assigned to PTUs, figure 4.8 shows the trend for requests assigned to PTUs. Advance requests have significantly increased over the two year period. Same Day requests have declined slightly, but there is quite a bit of variance in the data. With more Advance requests being submitted, there is less capacity to schedule Same Day requests.

4.2 Travel Time Estimates

PTU travel times are an important measure for the models presented in this thesis. Before running the models, estimates are required for a number of parameters, including t_{ij} . For a problem with $n = 50$ requests (EHS routinely serves 50 or more patient transfer requests over the course of day in the Central Region of Nova Scotia), there are $(2(n + 1))^2 = 10404$ t_{ij} values required. The patient transfer data will be used to estimate these values quickly, and an alternative solution for generating either missing or all t_{ij} values is also presented. Travel times during the peak hours as described in section 4.1.2 are only considered as traffic tends to be busiest during

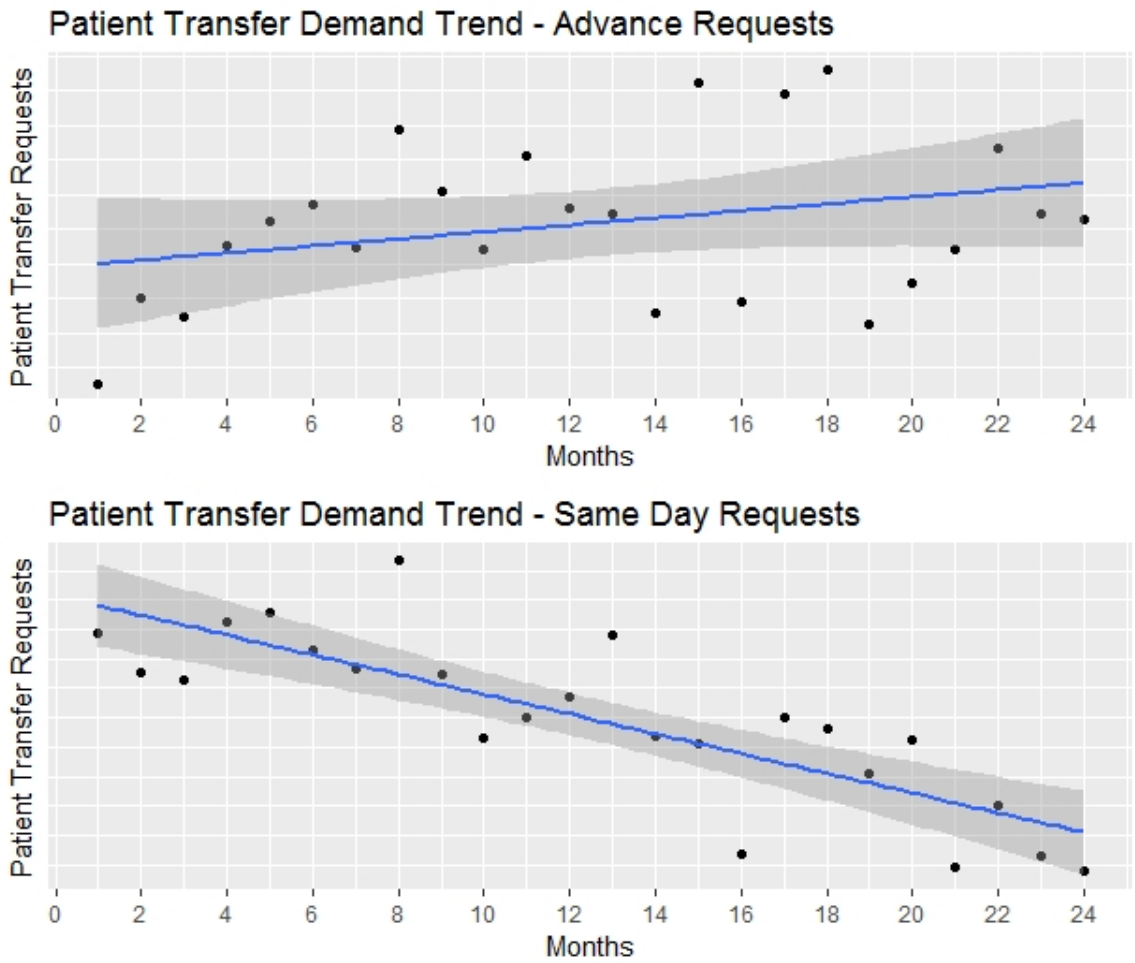


Figure 4.7: Advance and Same Day patient transfer requests trend

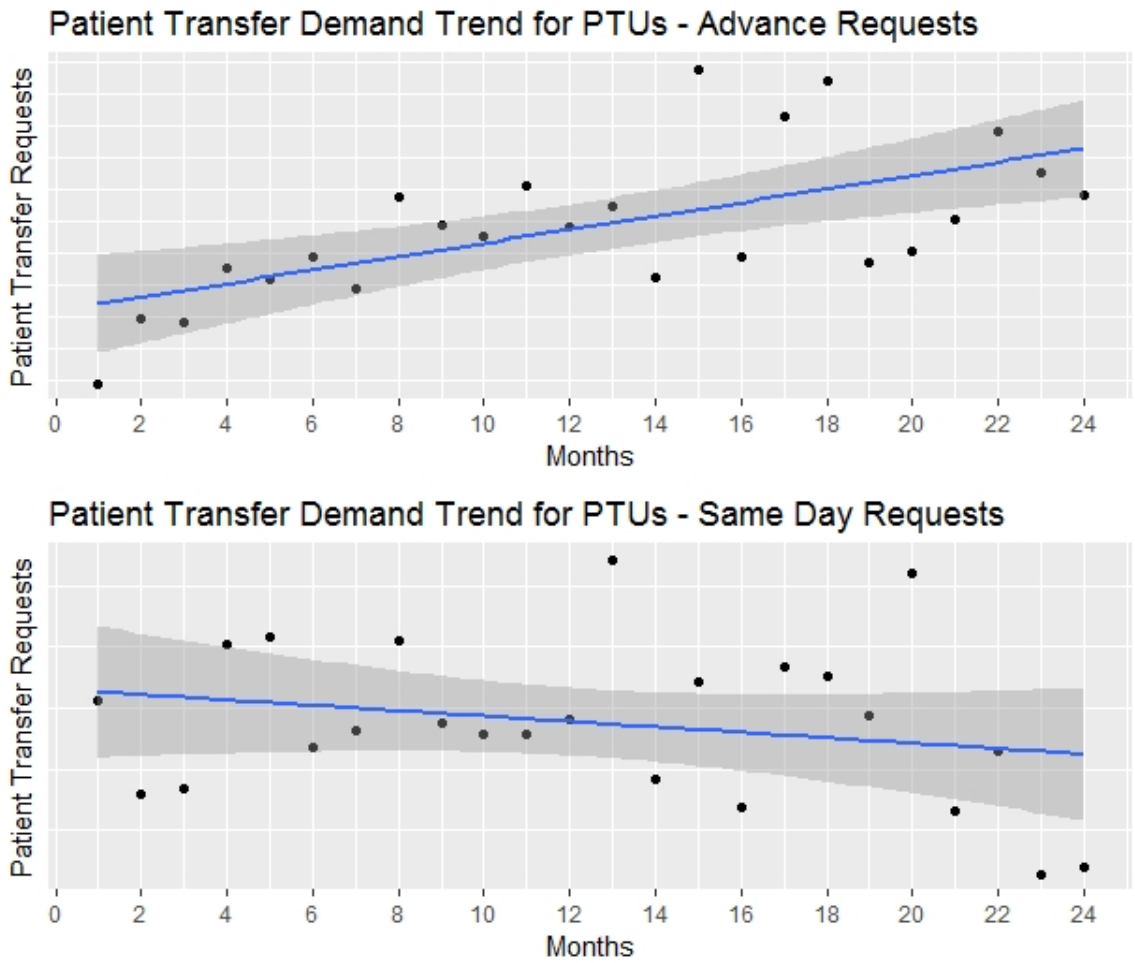


Figure 4.8: Advance and Same Day patient transfer requests trend for PTUs

this time.

It is important to note that, like many data sets, the patient transfer data set has errors and anomalies. The data was analyzed and cleaned prior to analysis. Pickup and delivery location names were standardized as some locations had different spellings and abbreviations. Outliers for travel and service times were identified. Most of these were the result of missing data and workarounds for the calculation were used. A small amount of outliers could not be corrected and had to be removed from the analysis. Several calculated columns were added to the data to help make analysis easier (such as month and day of the week).

First, pickup travel times are plotted in figure 4.9 and measures are shown in table 4.1.

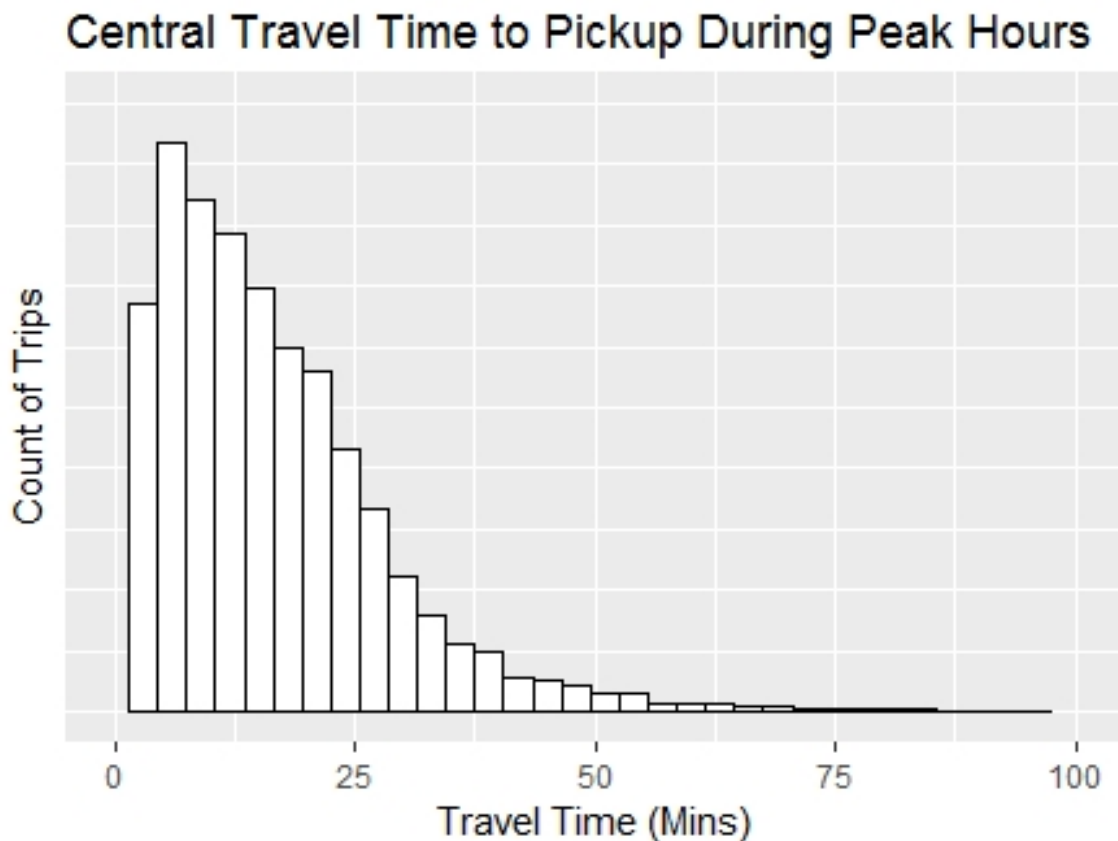


Figure 4.9: Pickup travel time distribution

Pickup Travel Time Metrics			
Mean	Median	Standard Deviation	Coefficient of Variation
12.24	9.25	12.88	1.05

Table 4.1: Pickup travel time metrics

The majority of pickups requires less than 12 minutes for travel, but there are a number of outliers with some pickups exceeding 90 minutes worth of travel. Next, delivery travel times are plotted in figure 4.10 and metrics are shown in table 4.2.

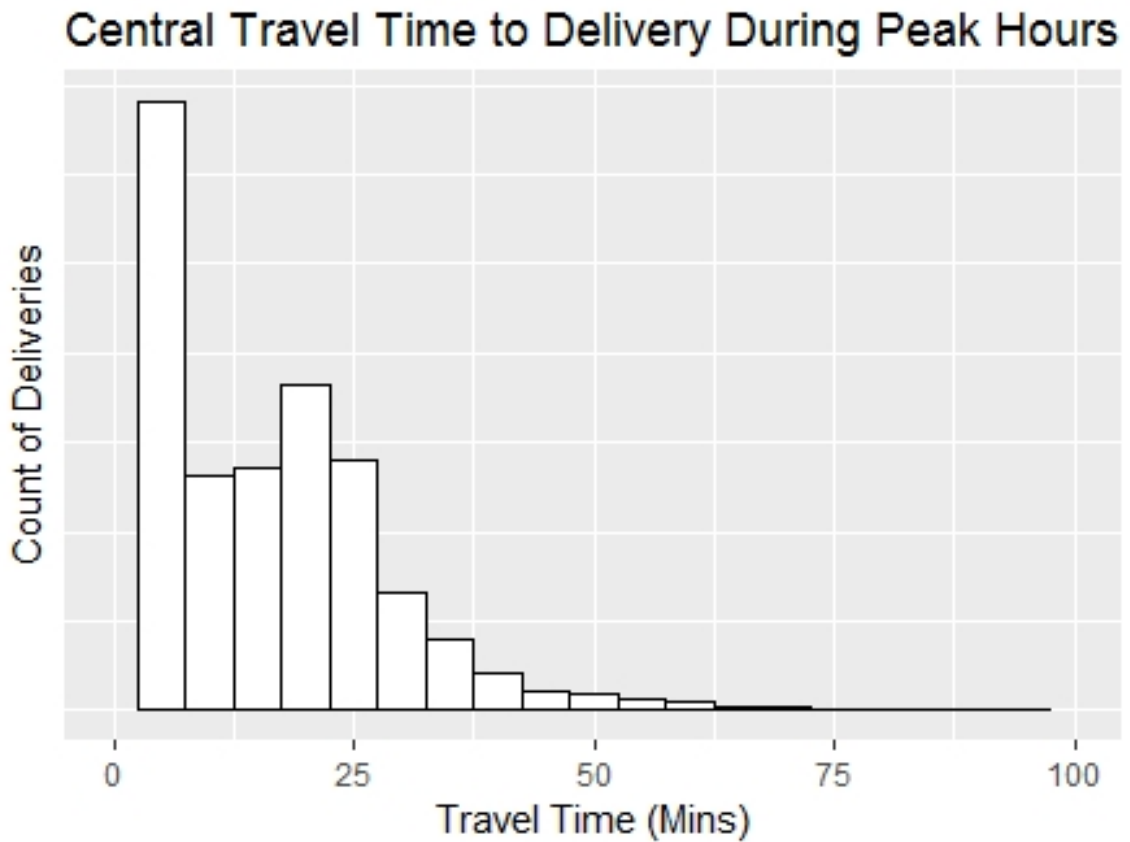


Figure 4.10: Delivery travel time distribution

There are two peaks in figure 4.10. The first occurs in the 0-5 minute bin of the histogram. This is because roughly 33% of the patient transfers in our data set are between the Halifax Infirmary and the Victoria General Hospital, the two largest hospitals in the Central Region. These two facilities are located very close to each

Delivery Travel Time Metrics			
Mean	Median	Standard Deviation	Coefficient of Variation
15.85	14.00	12.07	0.76

Table 4.2: Delivery travel time metrics

other, and the travel times show this; the mean and median travel times for this trip are 7 and 5 minutes, respectively.

The second peak occurs in the 15-20 minutes bin, and is the result of the remainder of the patient transfers in the data set. Similar to the pickup travel times, delivery travel times have outliers up to 100 minutes, although the median is 14 minutes.

From the patient transfer data, we can observe the Pareto principle in that almost 50% of requests involve one of the top two locations, and 84.9% involve one of the top 10 locations. With this data, estimates can be found for the majority of routes that the PTU crews will have to drive. The statistics for the top 10 locations are found in table 4.3.

Percent of Patient Transfers by Location		
Location	Percent of Requests	Cumulative Percent Sum
1	26.2	26.2
2	21.9	48.1
3	11.4	59.5
4	7.7	67.2
5	5.5	72.7
6	4.5	77.2
7	2.6	79.8
8	1.8	81.6
9	1.7	83.3
10	1.6	84.9

Table 4.3: Percent of requests for top 10 locations

For estimating travel times, a measure of central tendency such as the mean or median is chosen. To demonstrate the benefits of these two options, the mean and median travel times in minutes for the ten most populous routes from the patient transfer data are shown in table 4.4.

Top 10 Route Statistics		
Route	Mean	Median
1	7.5	5.1
2	22.6	21.0
3	23.0	21.0
4	7.2	5.0
5	19.9	16.0
6	26.7	25.6
7	50.3	48.9
8	19.3	14.4
9	35.1	29.4
10	6.4	5.0

Table 4.4: Top 10 route travel statistics (minutes)

Due to the outliers on the upper end of the travel time distributions (shown in figures 4.9 and 4.10), the travel time mean is consistently higher than the median for a given route. Using the mean is a “risk averse” method for estimating the travel time and is useful if delays are expected during the operating period.

While there are differences between the mean and median for pickup travel time and delivery travel time, they are treated the same for estimating. This means the same route for travelling to a pickup and travelling to a delivery will have the same estimated travel time. The reason is because the start and end points of a route are the primary factors that influence how long it takes to drive a route.

4.2.1 Missing Data

The travel time estimates for some routes are missing from the data; this means a PTU has never driven from a particular location to a particular destination (in the data set). For instance, there are many routes missing where a personal residence is either the pickup or delivery location. For these travel time estimates, a solution using the R programming language [37] and the googleway R package [4] is presented.

The googleway R package uses a user-provided Google API key to access Google Maps and retrieve data. One such use is to retrieve route distance and estimated

travel time for a route between two locations. This is done with the `google_distance` function. While two addresses can be used as inputs, sometimes exact addresses are unknown. However, EHS does store the latitude and longitude for every pickup and delivery location; `google_distance` can also use latitude and longitude as an input. For any routes with no travel data (and therefore no estimate for t_{ij}), a list of the routes can be inputted into the `google_distance` function to retrieve a travel time estimate. The code for this function is included in the `TravelTimeGenerator.R` file found on DalSpace.

It is recommended to use historical patient transfer data for the t_{ij} estimates where possible. Operating a PTU is not the same as operating a smaller vehicle such as a car. PTUs also have special equipment and may be transporting a patient, requiring extra care while driving. However, using data from Google Maps is an acceptable substitute in absence of patient transfer data.

4.3 Service Time Models

In this section, models to estimate service time with a patient are introduced. Service time, both in this chapter and in chapter 3, refers to the time spent with a patient at the pickup and delivery locations. It does not include any travel time. It can be thought of as steps 2 and 4 of the patient transfer process described in chapter 1.1. In the Advance and Same Day Models, every node $i \in P \cup D$ has a service time $d_i \geq 0$. Service time taking place at the pickup location, $i \in P$, is called pickup service time. Service time taking place at the delivery location, $i \in D$, is called delivery service time.

First, the pickup and delivery service time data are analyzed. The statistical differences are explored. Next, multiple linear models are explored to generate service time estimates. Only peak hours are considered again, as this is when hospitals are at their busiest.

4.3.1 Service Time Comparison

Pickup service time and delivery service time are analyzed and compared to determine if they require separate models. The distribution of pickup service time is shown in

figure 4.11 and measures are in table 4.5.

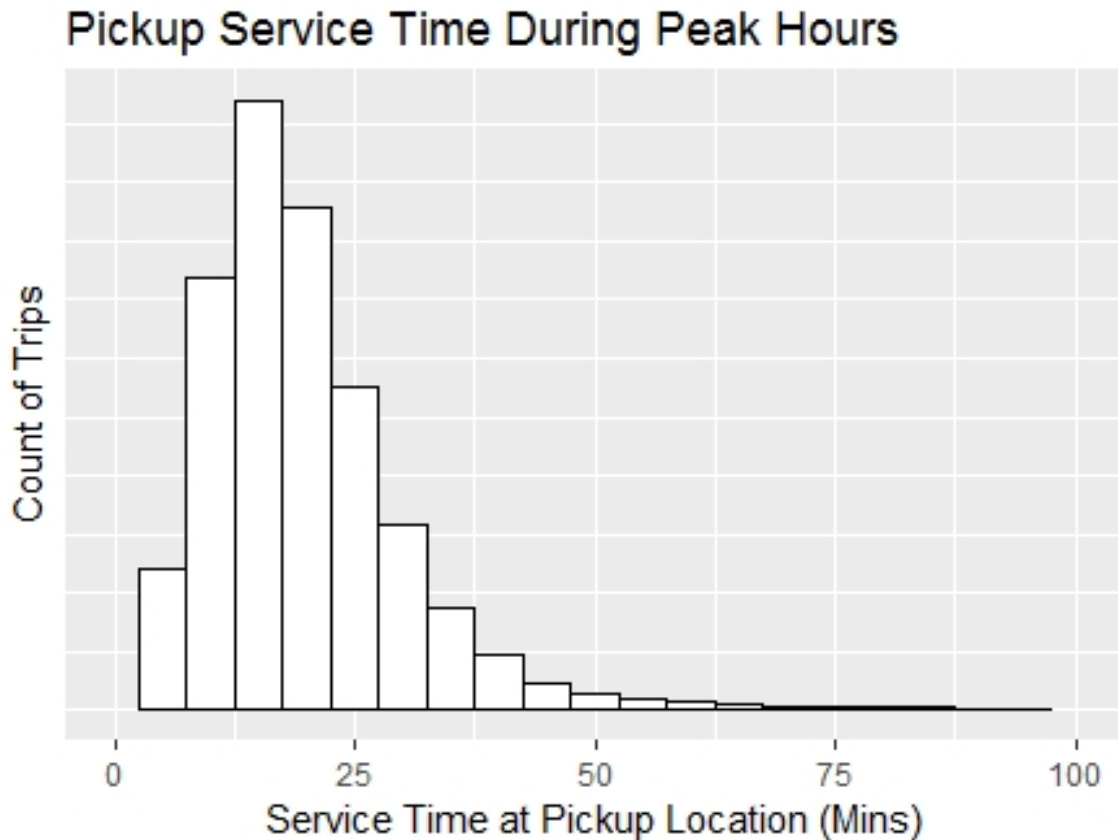


Figure 4.11: Pickup service time distribution

Pickup Service Time Metrics			
Mean	Median	Standard Deviation	Coefficient of Variation
18.49	17.00	12.11	0.65

Table 4.5: Pickup service time metrics

Pickup service time has a similar distribution shape to pickup travel time, with outliers as far as 100 minutes. Half the patient transfers have a pickup service time below 17 minutes. The delivery service time distribution and measures are shown in figure 4.10 and table 4.6. It also has outliers in the right tail. However, it also has a lower mean and median. A Welch's t test is used to test the hypothesis between the two group means (under the assumption of unknown and different population variances). Even though the data is non-normal, the t test should be sufficient due

to the very large number of samples (> 10000). The test statistic is 44.151 and the p-value is approximately 0. This confirms the alternative hypothesis that the pickup and delivery service time means are different.

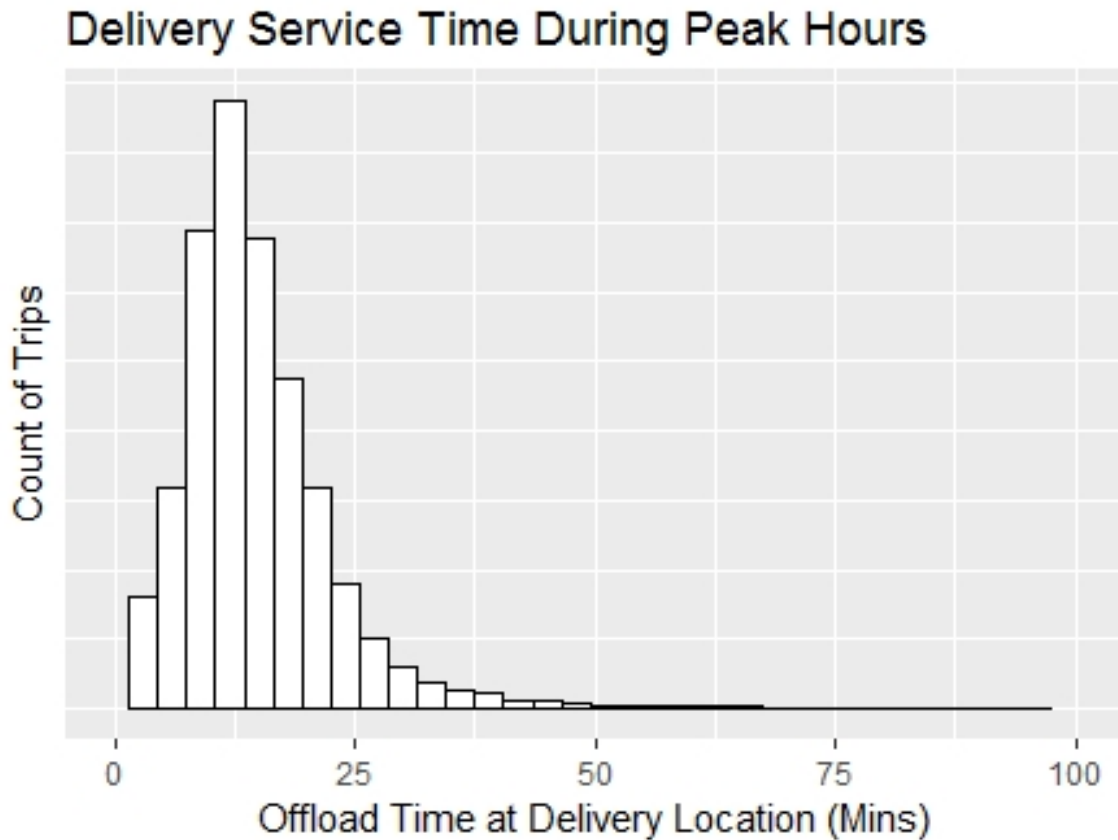


Figure 4.12: Delivery service time distribution

Delivery Service Time Metrics			
Mean	Median	Standard Deviation	Coefficient of Variation
13.86	12.78	9.93	0.72

Table 4.6: Delivery service time metrics

Perhaps service time is more dependent on the location, and not whether the service performed in a pickup or a delivery. Certain locations are represented in the data much more often than others (shown in table 4.3). To investigate, the mean pickup and delivery service times for the top ten patient transfer locations are shown in table 4.7. The table provides evidence that while the location probably does affect the

service time length, pickup and delivery service times are likely to be two distinct populations. Nine of the top ten locations have a pickup service time mean greater than the delivery service time mean. Because of this evidence, pickup and delivery service time are treated as two independent populations.

Top 10 Location Service Time Means		
Location	Pickup Service Mean	Delivery Service Mean
1	21.76	16.79
2	17.72	14.84
3	19.98	15.26
4	14.77	10.85
5	15.58	11.19
6	18.24	11.40
7	16.10	12.96
8	17.74	12.50
9	8.68	8.44
10	18.38	28.34

Table 4.7: Top 10 location service time means (minutes)

4.3.2 Pickup Service Time Model

In this section, a multiple linear regression model is investigated for pickup service time. Several predictors are investigated for their effects on patient transfer service time. The first predictor to investigate is the pickup location. As shown in section 4.3.1, where the pickup takes place likely affects the pickup service time. However, there are 100 distinct pickup locations in the patient transfer data set. 64 of these locations have less than 10 pickups. Therefore, only locations with greater than 10 pickups were included in the linear regression. 10 locations were kept in the final model (F statistic = 123.2).

The day of the week is added to the linear model as a categorical variable. Saturday and Sunday are omitted due to the low patient transfer demand on the weekend. The day of the week has a F statistic of 26.38 (p-value of approximately zero) and is therefore significant in the model even with the presence of the 10 pickup locations.

Hour of the day is added, also as a categorical variable. The hours included are

6.00 to 18.00, matching the peak demand period established in section 4.1.1. The F statistic for hour of the day is 8.25 (p-value of approximately zero) when added to the linear model. Day of the week and pickup location remain significant.

A new column indicating whether the patient is bariatric or not was added to the data set. This is based on a code used by EHS. This “bariatric flag” was added to the model. It has an F statistic of 35.24 (p-value of approximately zero) and is included in the linear regression model. All previous predictors remain significant.

Next, the destination was investigated for effects on pickup service time. The same approach as the pickup location was taken for the destination. Only four destination location were shown to have a statistically significant effect on the pickup service time, and they are included in the linear model. The final model, including predictor and F-statistic, is found in table 4.8. The R^2 value is 0.09 and the root mean square error (RMSE) is 10.93.

Several other predictors were tested in addition to the ones above, but were found to not have a significant effect on the pickup service time. These include whether the patient requires special equipment, whether special precautions for the paramedics are required, whether the request is an Advance or Same Day request, the type of facility the pickup is located in, and several pickup and destination locations.

4.3.3 Delivery Service Time Model

The process from section 4.3.2 was repeated for the delivery service time model. The distribution of delivery locations is similar to pickup locations. There are 109 distinct delivery locations in the data set, but 72 of these locations have less than 10 pickups. The remaining 37 locations were included in a linear regression. 16 of these locations were kept in the final model (F statistic of 60.5).

Day of the week (F statistic of 3.93) and hour of the day (F statistic of 7.02) are included again. The bariatric flag is included as well (F statistic of 37.54). Unlike the pickup service time model, whether the request was an Advance or Same Day request

Predictor	Type	F Statistic
Day of Week	Categorical	26.64
Hour	Categorical	12.24
Pickup Location 1	Binary	948.79
Pickup Location 2	Binary	222.89
Pickup Location 3	Binary	134.33
Pickup Location 4	Binary	193.48
Pickup Location 5	Binary	23.81
Pickup Location 6	Binary	22.16
Pickup Location 7	Binary	11.30
Pickup Location 8	Binary	13.25
Pickup Location 9	Binary	12.69
Pickup Location 10	Binary	32.88
Bariatric	Binary	52.17
Destination 1	Binary	18.72
Destination 2	Binary	7.75
Destination 3	Binary	30.16
Destination 4	Binary	16.40

Table 4.8: Multiple linear regression model for pickup service time

does affect the delivery service time (F statistic of 31.06). The type of facility is also statistically significant (F statistic of 34.53). The final delivery service model is in table 4.9.

Other predictors were tested and found not to be statistically significant. These included whether the patient requires special equipment, whether special precautions for the paramedics are required and several pickup and delivery locations.

The framework of the scheduling model shown in figure 4.13 has been updated to include the information provided from the data and the googleway package.

Predictor	Type	F Statistic
Day of Week	Categorical	3.96
Hour	Categorical	7.04
Delivery Location 1	Binary	687.63
Delivery Location 2	Binary	66.50
Delivery Location 3	Binary	111.50
Delivery Location 4	Binary	193.48
Delivery Location 5	Binary	521.42
Delivery Location 6	Binary	11.04
Delivery Location 7	Binary	14.58
Delivery Location 8	Binary	6.58
Delivery Location 9	Binary	14.05
Delivery Location 10	Binary	24.40
Delivery Location 11	Binary	8.50
Delivery Location 12	Binary	23.38
Delivery Location 13	Binary	281.66
Delivery Location 14	Binary	24.83
Delivery Location 15	Binary	13.28
Delivery Location 16	Binary	12.47
Bariatric	Binary	38.04
Advance or Same Day	Binary	31.42
Location Type	Categorical	34.53

Table 4.9: Multiple linear regression model for delivery service time

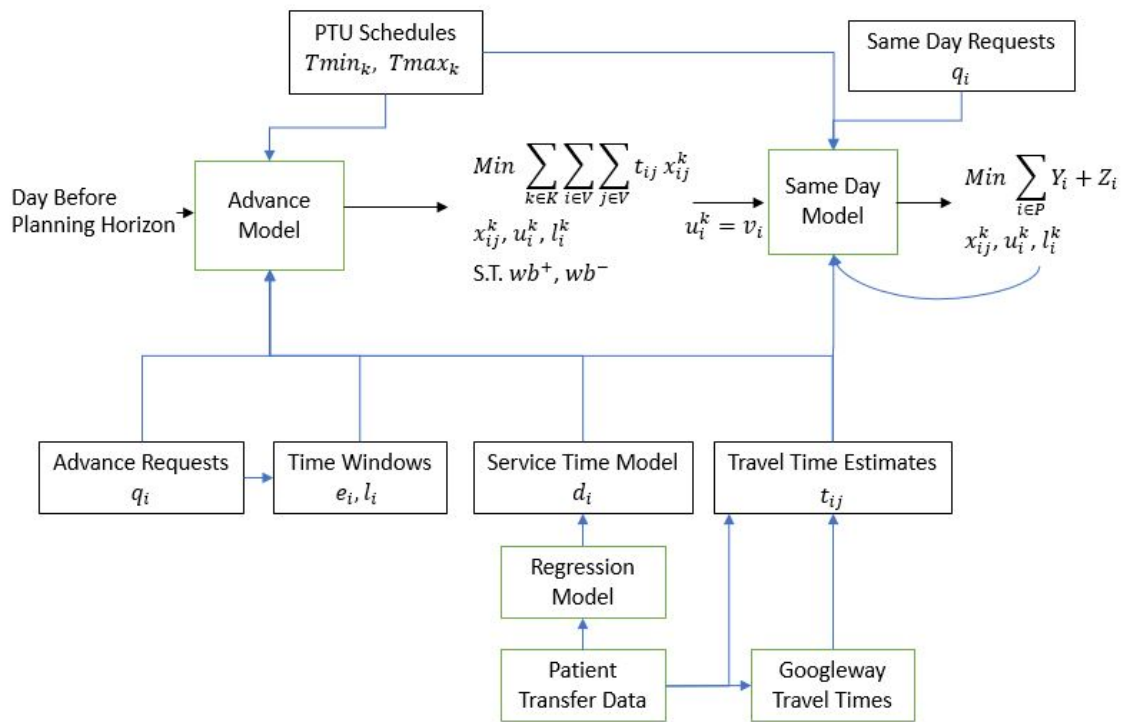


Figure 4.13: Scheduling model framework updated with the models derived from the data

Chapter 5

Results

In this chapter, experiments are designed for the Advance Model, Same Day Model and the two used in conjunction. Afterwards, the results are analyzed. The Advance and Same Day Models have experiments conducted to explore several parameters. The quality of the schedules generated by the Advance Model are compared to the actual schedules (with some limitations) as well. The value of information is explored next, comparing results from scenarios where every patient transfer request is known in advance and the actual scenario where some requests are received during the same day as the requested pickup time. The final experiment simulates an actual operational day as new requests are received using the models.

The computational time to solve the DARP models, the total travel time and the number of changed pickup times are the variables of interest. The input files containing the parameters that are used by the model are created with the patient transfer data. The R programming language [37] and RStudio [39] are used to extract the information from the data into the input files. The DARP Advance and Same Day Models are constructed with the Pyomo package [16] [17] and written in the Python programming language. The problems are solved using the Gurobi Optimizer [12] on a server with an Intel 4114 CPU, 2.20 GHz processor with 20 cores and 64 GB of RAM. A maximum solve time of six hours was placed on every experiment. Vehicle capacity, Q_k , was kept to one to match the capacity of the PTUs.

5.1 Experimental Design

The design of the experiments is reviewed in this section. First, experiments testing parameters in the Advance Model are reviewed. Next, experiments testing the effect of parameters in the Same Day Model are reviewed. The experiments that put a value on perfect information are reviewed next. The last set of experiments simulates

a single day where the Advance Model is used to generate the preliminary schedule, and the Same Day Model is run several times over the course of the day as new patient transfer requests are received.

5.2 Advance Model Experiments

The Advance Model is tested with different inputs for certain parameters. The patient transfer data is used for the experiments. Data from nine different days is used; they are chosen to encompass a wide range of the number of patient requests. The smallest experiment has 25 patient transfer requests and the largest has 64 requests; this is the largest number of patient requests served by Central Region PTUs in a single day in the data set. Every request is treated as an Advance request. The parameters of interest are the number of vehicles, the length of the time windows and the use of the workload balancing constraints. The experiments are described below.

- The vehicles are tested for values of 5, 6, 7, and 8. These numbers are chosen because they are close to the actual number of PTUs used in the Central Region during the peak operating hours by EHS. They can provide evidence of how increasing or downsizing the fleet of PTUs may affect operations.
- The time window length is applied to both sides of the pickup time. Therefore, PTUs may arrive a little early or a little late. This decision was made to balance early and late pickups. As discussed in section 3.2, arriving early is not necessarily a positive outcome for EHS. The time window lengths chosen for testing are 10, 12, 15 and 20 minutes.
- The workload balancing constraints are turned on (with a wide range between wb^+ and wb^-) and off. This is because when there are no workload balancing constraints, the Advance Model will attempt to use as few vehicles as possible to serve every request. In addition, a low wb^+ value will have the side effect of limiting the number of requests that can be served by the vehicles.

After these experiments are carried out, the Advance model is tested for practical use and the quality of the schedules generated by the Advance Model for selected days in the patient transfer data where the number of Same Day requests is minimized. For

the practical use experiments, the Advance Model is tested on 40 randomly selected days.

5.3 Same Day Model Experiments

For the Same Day Model, every day in the historical data is assigned to a category. The number of Advance and Same Day requests are summed for each day and sorted into three categories: Low, Medium and High. If the number of requests for a day is in the 33rd percentile or lower, it is placed into the Low Category. If the number of requests is in the 67th percentile or higher, it is placed into the High category. Otherwise, the day is placed into the Medium category. This is done for both Advance and Same Day requests. Every day is then placed into a category according to the matrix in figure 5.1. A random day is chosen from the following categories (highlighted in grey in figure 5.1) for the Same Day Model experiments:

- Low Same Day, High Advance
- Medium
- Medium Same Day, High Advance
- High Same Day, Low Advance
- High Same Day, Medium Advance
- High

The Low and Low-Medium categories are ignored as they are less interesting. In the Same Day Model experiments, different values of β^+ , β^- , α^+ and α^- are investigated while the number of PTUs is kept constant and workload balancing constraints are used. For simplicity, $\beta^- = -\beta^+$ and $\alpha^- = -\alpha^+$. For each experiment, the first step is to create a schedule with the Advance Model and Advance requests so that the pickup times may be used as the w_i parameter in the Same Day Model.

		Advance Requests		
		Low	Medium	High
Same Day Requests	Low	Low	Low Same Day Medium Advance	Low Same Day High Advance
	Medium	Medium Same Day Low Advance	Medium	Medium Same Day High Advance
	High	High Same Day Low Advance	High Same Day Medium Advance	High

Figure 5.1: Same Day Model experiments

5.4 Value of Information Experiments

The next set of experiments attempt to put a value on information. 19 days describing a wide range of total patient transfer requests were selected for the experiments. For every day, the scheduling problem is solved with two different approaches. First, all of the Advance requests are used as inputs to the Advance Model and solved with 7 vehicles, 10 minute time windows and workload balancing. The pickup times generated by this solution, the Advance requests and the Same Day requests then are submitted as inputs to the Same Day Model and solved.

The second approach treats every request as an Advance request, and the problem is solved in the Advance Model. In this scenario, the Advance Model has 7 vehicles and workload balancing, and the time window is the smallest value the problem will solve with from 10, 12, 15 and 20 minutes. The sum of the travel time from both methods are compared to each other.

5.5 Same Day Strategy Experiments

The final experiment tests how the models perform while being used throughout the online operational phase. It will simulate how EHS or a similar organization would use the Advance and Same Day Models throughout a simulated day. A ‘typical’ day with a total number of requests close to the daily mean for our sample data ($n = 48$) and a relatively even split between Advance and Same Day requests is chosen from the patient transfer data. Our chosen day has 49 requests with 24 Advance requests and 25 Same Day requests. A preliminary schedule with the 24 Advance requests is created with the Advance Model (similar to EHS creating a preliminary schedule in the evening for the following day). The Advance Model is run with 7 vehicles, 10 minute time windows and workload balancing constraints.

In the patient transfer data, there is a column that lists the time EHS picked up the phone for each request. With this, it is known exactly when each request was placed with EHS. This information is used to devise a possible strategy for the Same Day Model and to conduct this experiment. There are many possible strategies for choosing when to run the Same Day Model throughout the day. One option is to run the model every time a Same Day request is received. Another option is to run the Same Day Model once a threshold is reached. This threshold can be either once a certain number of requests have been received, or once a certain amount of time has passed.

For this experiment, using the Same Day Model is tested with a threshold of 30 minutes. This means the Same Day Model is run every half an hour during the simulated day. This is done to balance efficiency and practicality. Waiting a long time or until a large number of new, Same Day requests have been received may not be feasible as some requests have a short duration between the time the request is placed and the requested pickup time. Running it for every request or every 5 minutes may be inefficient if the new schedule has to be scrapped because a new call immediately came in after the model finished solving. Running the Same Day Model every 30 minutes is done to balance these two extremes. The Same Day Model is run with 7 vehicles, hard time windows of 20 minutes, soft time windows of 10 minutes and workload balancing.

One final consideration for the simulated day are pickups that take place before the time the Same Day Model is run. For instance, consider a pickup that takes place at node $i = 1$ with a pickup time of 6.15. In the previous schedule, vehicle $k = 1$ is to drive to node $i = 1$ from the depot for this pickup. When the Same Day Model is run at six thirty, vehicle $k = 1$ will already be at node $i = 1$ (or will have left, transporting the patient to the destination). This pickup would obviously not be changed and should be reflected in the new schedule. Certain x_{ij}^k are set to 1 according to the following rule in order to account for this possibility.

If x_{ij}^k (where $j \in P$) is equal to 1 in the previous schedule and the service start time at node j by vehicle k , u_j^k , is earlier than the present time when the Same Day Model is to be run, x_{ij}^k is set to 1. In the example given above, $x_{01}^1 = 1$ for the Same Day Model run at six thirty. There is no need to account for the nodes in the delivery set ($j \in D$) as the routing constraints will ensure vehicle k will visit the delivery node $j + n$.

Therefore, the simulated day experiment can be summarized as follows:

1. Run the Advance Model for all 24 Advance requests ‘the night before’ the simulated day.
2. Run the Same Day Model at 6.00 if any new requests were received overnight. This is the beginning of the online operational phase for EHS. The Advance Model scheduled pickup times are used as w_i in the Same Day Model.
3. Run the Same Day Model every 30 minutes until 19.00. Run only if any new requests are received in the previous 30 minutes. The scheduled pickup times from the most recent schedule are used as w_i . If u_j^k where $j \in P$ is less than the current time, set $x_{ij}^k = 1$ for $j \in P$ if x_{ij}^k was equal to 1 in the previous schedule.

The feasibility and practicality of this strategy will be evaluated through the computational time to run the models and the sum of the travel time as compared to an optimal solution generated by the Advance Model for all 49 patient transfer requests.

Experiment	Purpose
Advance Model Parameter Testing	Test different inputs for number of vehicles, time windows and workload balancing
Advance Model Comparison to Actual	Compare travel time of Advance Model routes to actual routes
Same Day Model Parameter Testing	Test different inputs for soft and hard time windows
Value of Information	Compare travel time of routes from actual Advance/Same Day split and scenario where all requests are known in advance
Same Day Strategy	Simulate using the Advance and Same Day models over the span of one operational day

Table 5.1: Experiments summary

5.6 Advance Model Results

The first set of Advance Model experiments tested different values for the number of vehicles (k in the model). Values of 5, 6, 7 and 8 were used for k . Workload balancing constraints were used and the time windows were kept at 10 minutes. 9 days were chosen to test that encompassed a wide range of patient transfer requests. In total, 36 experiments were performed. All were solved to optimality, with the exception of the experiments where a status of “infeasible” was returned by the solver. The results are found in figure 5.2. The computational time and the sum of the vehicle travel time are charted against the number of patient transfer requests. The symbol represents the number of vehicles used in the experiment. As expected, the computational time increased exponentially as the number of requests increased. The total travel time increased linearly. The computational time chart shows that increasing the value of k slightly increases the time required to optimally solve the DARP; the 8 vehicle problems require the longest amount of time. Changing the value of k appears to have little effect on the total travel time. The larger problems ($n > 50$) show slightly less total travel time for the 8 vehicle problems, however.

Some of the test problems were unable to solve and a status of “infeasible” was returned. Problems with $k = 5$ and $k = 6$ quickly became infeasible as the number of patient transfer requests grew. The infeasible experiments are found in table 5.2.

Number of Requests	Vehicles
35	5
39	5
39	6
45	5
50	5
50	6
50	7
56	5
56	6
60	5
60	6
64	5
64	6
64	7
64	8

Table 5.2: Infeasible Advance Model vehicle experiments

For the workload balancing experiments, the previous 36 experiments were repeated without the workload balancing constraints and solved to optimality. The results are in figure 5.3. There is almost no difference in the total travel time. Experiments with no workload balancing constraints appear to take slightly longer. On average, these experiments took 6% longer than their partner experiment with workload balancing. The same experiments that were infeasible in the vehicle experiments remained infeasible for the workload balancing experiments.

The final set of experiments on the Advance Model tested time window lengths of 12, 15 and 20 minutes (in addition to the 10 minute time windows tested in the previous experiments). These experiments were solved to a MIP gap of 2% as some problems reached the six hour time limit imposed on the solver. The experiments used workload balancing, but only used values of 7 and 8 for k due to lower numbers being infeasible for larger problem sizes. The results are found in figure 5.4.

The results show that a larger time window greatly increases the computational time required to solve the DARP. The time window length appears to have minimal impact on the total travel time, however. The main function of increasing the time

windows is to make larger problems become more likely to solve, as some of the previous problems that were infeasible solve under larger time windows (including the 64 request problem, which solved with 12 and 15 minutes time windows). Some of these problems reached the six hour time limit, but returned solutions that did not reach the 2% MIP gap. These problems were attempted with a 5% MIP gap. 5% is considered an acceptable gap because even in a large DARP problem for EHS, the total travel time for all vehicles is typically around 1200 minutes. 5% spread across 7 vehicles is roughly 8.5 minutes. A single patient transfer typically takes 60 minutes to complete; therefore, 8.5 minutes will not have a large impact on the final solution. Some problems still reached the six hour time limit. These experiments and their MIP gaps are listed in table 5.3. Every one of these problems solved to within a 5%

Number of Requests	Vehicles	Time Window	Travel Time	MIP Gap
39	7	20	711	12.24%
39	8	15	733	4.91%
39	8	20	727	12.24%
45	7	20	1009	4.96%
45	8	20	1016	5.91%
50	7	20	1069	4.96%
50	8	20	1072	4.94%
56	7	20	1177	16.31%
56	8	20	1159	19.41%
60	7	15	1127	3.82%
60	8	15	1132	15.02%
64	8	12	1488	6.05%
64	8	15	1448	24.45%

Table 5.3: Time window experiments with MIP gap larger than 2%

optimality gap with smaller time windows, except for the 64 request problem (which only generated solutions for the 12 and 15 minute time window problem with 8 vehicles). If an experiment did not solve to the 5% MIP gap in six hours, experiments with larger time windows were not performed in order to save time.

5.7 Advance Model Comparison to Actual Schedules

EHS typically receives 15 to 45 Advance requests per day in the HRM. The Advance Model is tested on randomly selected actual problems from the historical data. The

	Requests	Computational Time (Seconds)
Max	49	3281
Min	19	2
Mean	33	319

Table 5.4: Advance Model computational results summary

Advance Requests	Model Travel Time	Actual Travel Time	Actual Completed Requests	Travel Time Difference	Same Day Requests
24	645	747	22	102	8
26	771	833	26	62	8
32	847	1244	32	397	9
34	744	1048	34	304	11
35	768	993	35	225	10
37	890	933	36	43	8
39	920	1211	37	291	10
39	740	1191	38	451	11
43	1289	1271	41	-18	12
309	7614	9471	301	1857	

Table 5.5: Advance Model travel time comparison to actual schedules

computational times are plotted on a logarithmic scale in figure 5.5. Every problem tested solved in under one hour of run time. Since EHS would run the model overnight, this shows that the model can be solved in a short time frame where the results will be of use to EHS. Table 5.4 shows a summary of the tests including the maximum, minimum and mean number of requests and time (measured in seconds) required to solve the DARP.

The total travel time from the Advance Model is compared against the actual recorded travel time from the data for Advance requests. However, it is not possible to obtain the original schedules with only the Advance requests. Instead, the data contains the final schedules with Advance and Same Day requests. Therefore, the 9 days where Same Day requests make up the smallest percentage of all requests are selected from the patient transfer data. While this is not a completely fair comparison, focusing on the days with the least impact from the Same Day requests is the closest possible to a fair comparison. The results are found in Table 5.5. The final row in the table shows the sum of the columns. For these Advance Model problems, there is a total

reduction in travel time of 1857 minutes, or 206 minutes per day. This is a 19.6% improvement on the actual travel time spent on the Advance requests. In addition, the PTUs only completed 301 Advance requests vs. the 309 that were planned for as 8 requests were cancelled. Cancellations can happen for a number of reasons, but it is most likely these requests were cancelled because the patient was not ready for transport. The average travel time spent per request as per the model is 24.6 minutes vs. 31.5 minutes in the data, an improvement of 21.7%.

5.8 Same Day Model Results

As discussed in section 5.1, the hard and soft time windows are tested in the Same Day Model. Six days from the patient transfer data are chosen corresponding to figure 5.1. Two levels are tested for both time windows. 20 and 30 minutes are used for the hard time windows (β^- and β^+). Soft time windows of 5 and 10 minutes are used (α^- and α^+). This resulted in 24 total experiments. The results are in figures 5.6, 5.7 and 5.8. The figures list the category of how many Advance and Same Day requests were in the experiment on the y and x axis, respectively. The variables of interest are the computational time, the total travel time of the vehicles and the number of changed pickup times (the objective function of the Same Day Model). The size of the points indicates how large the variable is, and the colour indicates the hard time window length.

Figure 5.6 shows the computational time. While the High categories have larger computational times, there does not appear to be a large difference between the hard time windows. Figure 5.7 shows the travel time. There is no noticeable difference between the hard time windows within each experiment again. Figure 5.8 shows the changed pickup times. The main takeaway from this chart is the medium experiment had no changed pickup times, where any experiment with a high number of Advance or Same Day requests required multiple changed pickup times.

The results for the same experiments compared to the soft time windows are shown in figures 5.9, 5.10, 5.11. These figures follow the same format as the figures describing the results for the hard time windows. Figure 5.9 shows the computational time. It

appears that the smaller time window has longer computational times for every high Advance request experiment and the medium Advance with high Same Day requests experiment. The rest show no difference. Figure 5.10 shows the travel time. There does not appear to be any difference within the experiments for this measure. Figure 5.11 shows the changed pickup times, and every experiment except the medium experiment has a higher number of changed pickup times for the smaller soft time window.

5.9 Value of Information Results

The third set of experiments tested several randomly chosen days from the patient transfer data to determine how knowing all requests in advance affects the potential solutions to the DARP. The experiments consist of a wide range of realistic number of patient transfer requests, from 38 to 56 requests. Figure 5.12 shows the results. Clearly, the ‘Actual Advance/Same Day Split’ travel times are higher than every experiment where all requests were known in advance. Knowing all requests in advance saved a total of 4724 minutes of travel time across the 19 experiments; an average of 248.6 minutes per day. This is 19.4% of the total travel time for the ‘Actual Advance/Same Day Split’ scenarios. This is worth roughly the equivalent of half a PTU shift (just over four hours).

5.10 Same Day Strategy Results

The final set of experiments simulates the online operational phase for a single day. The Advance Model is run on the 24 Advance requests. Those pickup times are fed as the first inputs to the Same Day Model, and the Same Day Model is run every 30 minutes if any new Same Day requests are made. Figure 5.13 shows how the number of requests come in. The x axis is the time of day in hours. The requests steadily come in until the last request is received by five thirty. The first point at zero on the x axis represents the 24 Advance requests.

At the end of the day, no pickup times were significantly changed from the original schedule. Figure 5.14 shows how the total travel time for the vehicles changes over

the course of the day. Certain times, the new solution had less travel time than the previous solution. This occurred at 10.00 and 11.30. This is likely because there are multiple potential solutions for the objective function, and the quality of the total travel time will vary between the solutions.

The final total travel time is 1357 minutes. From the patient transfer data, the final actual travel time is 1662 minutes. The final schedule generated by the model shows an improvement of 18.4% over the actual schedule. All of the requests were fed into the Advance Model to determine the optimal solution for travel time, and the travel time is 1085 minutes, or a 20% improvement. Figure 5.15 shows the computational time required for each time the Same Day Model was run. The point at zero on the x axis is the time required for the 24 Advance requests with the Advance Model. It required the longest time at 20 seconds. None of the Same Day Model runs required longer than 10 seconds to solve. This is likely because it is being forced to use part of the previous solution, which reduces the solution space.

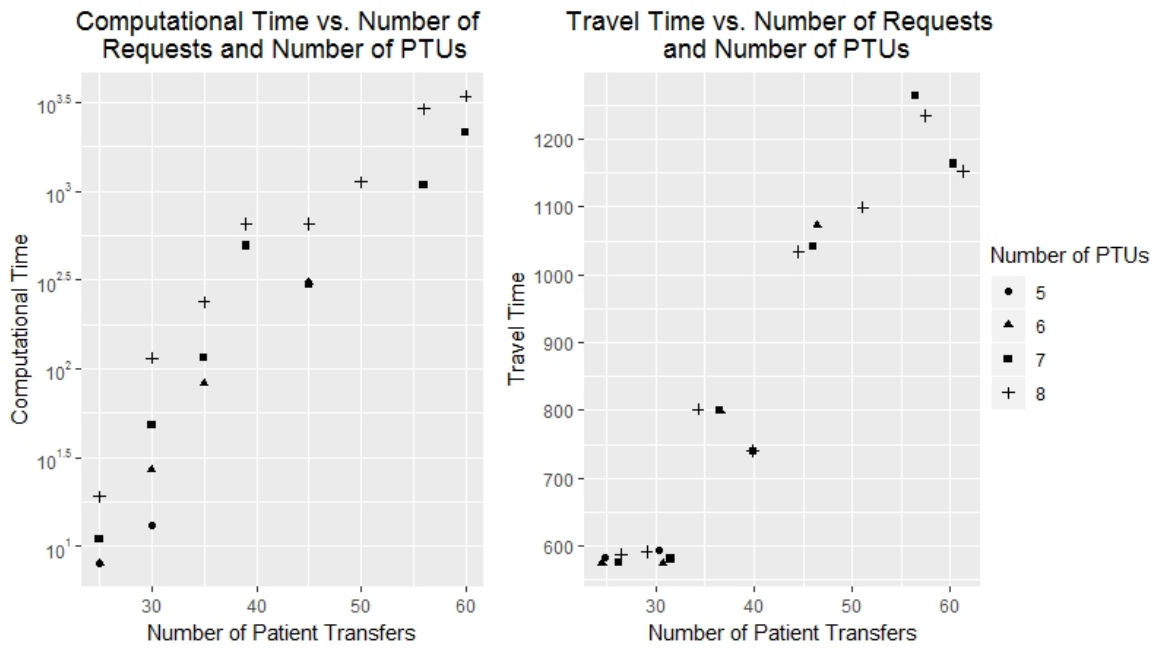


Figure 5.2: Advance Model results for the number of vehicles

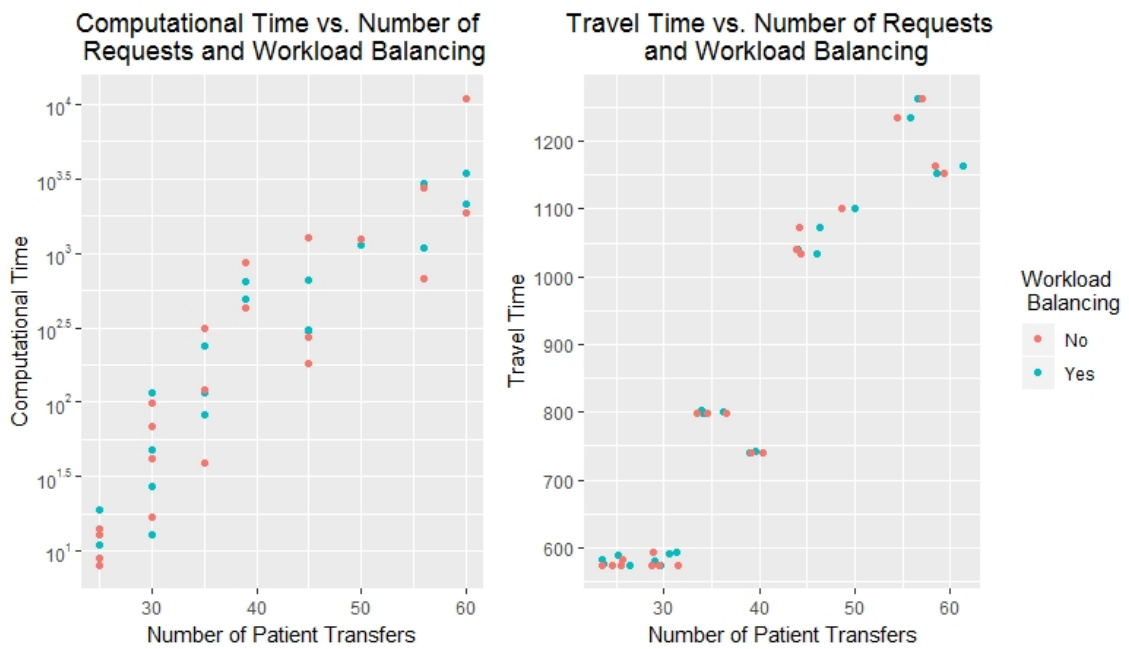


Figure 5.3: Advance Model results for workload balancing

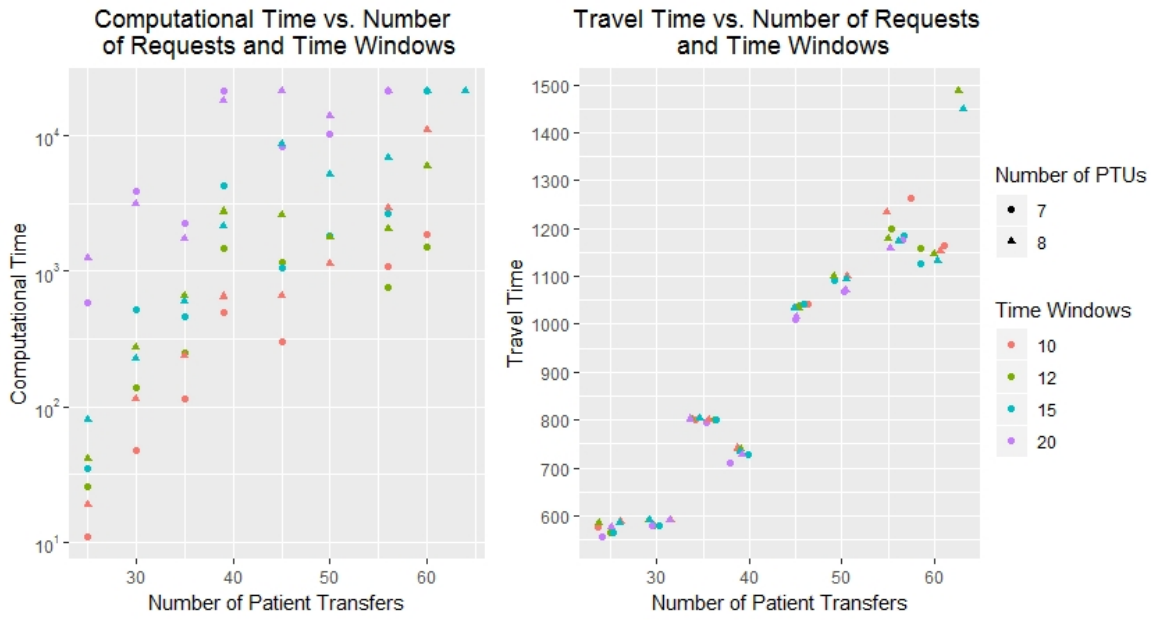


Figure 5.4: Advance Model results for time windows

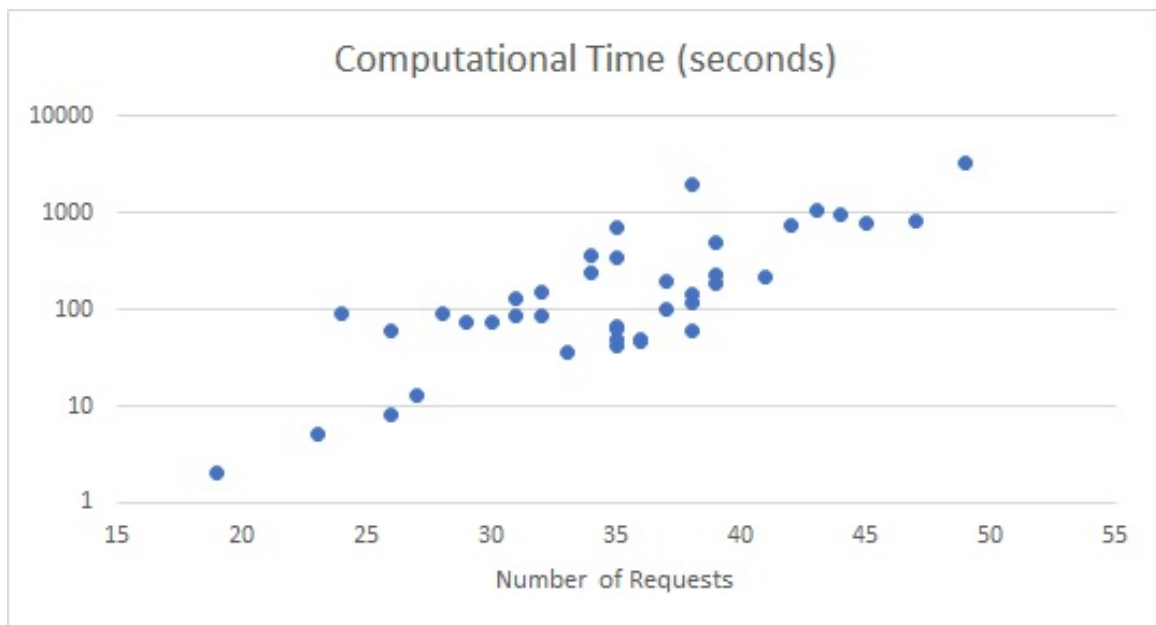


Figure 5.5: Advance Model computational time

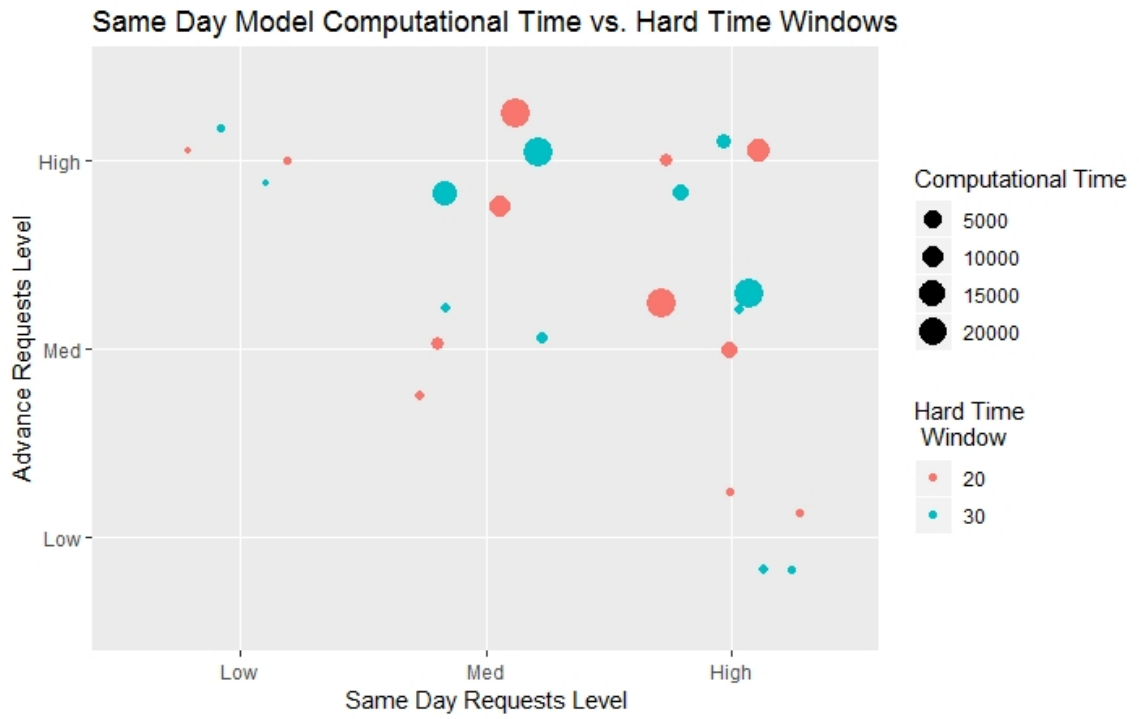


Figure 5.6: Same Day Model hard time windows computational time



Figure 5.7: Same Day Model hard time windows travel time

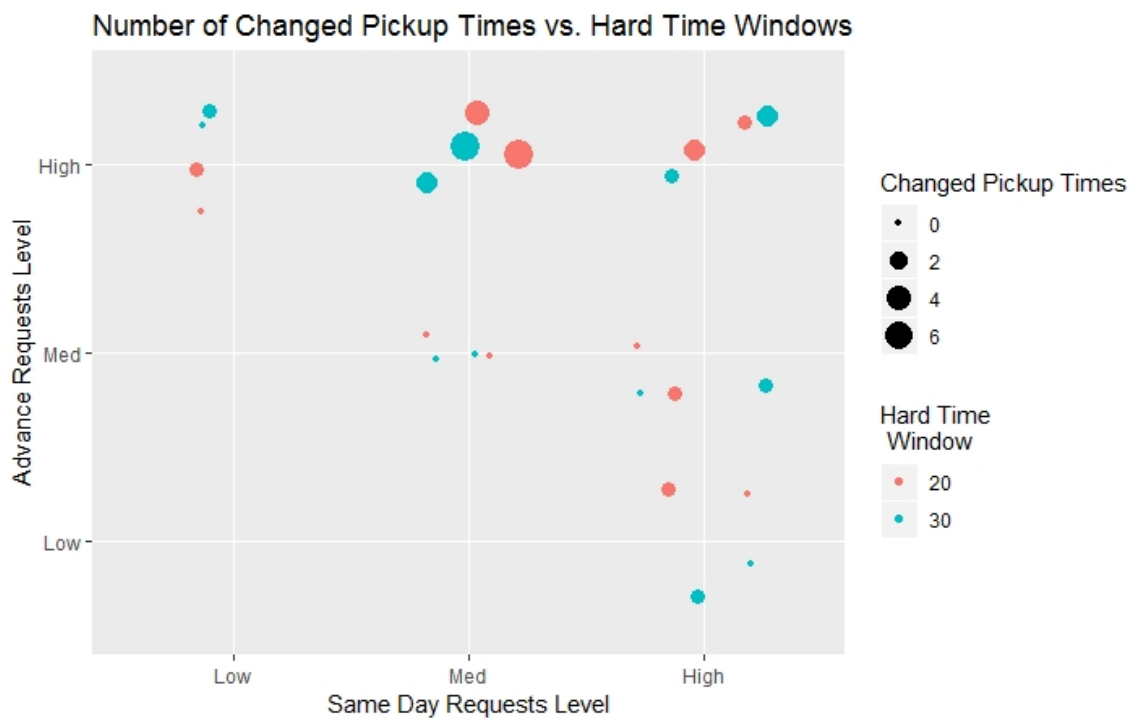


Figure 5.8: Same Day Model hard time windows changed pickup times

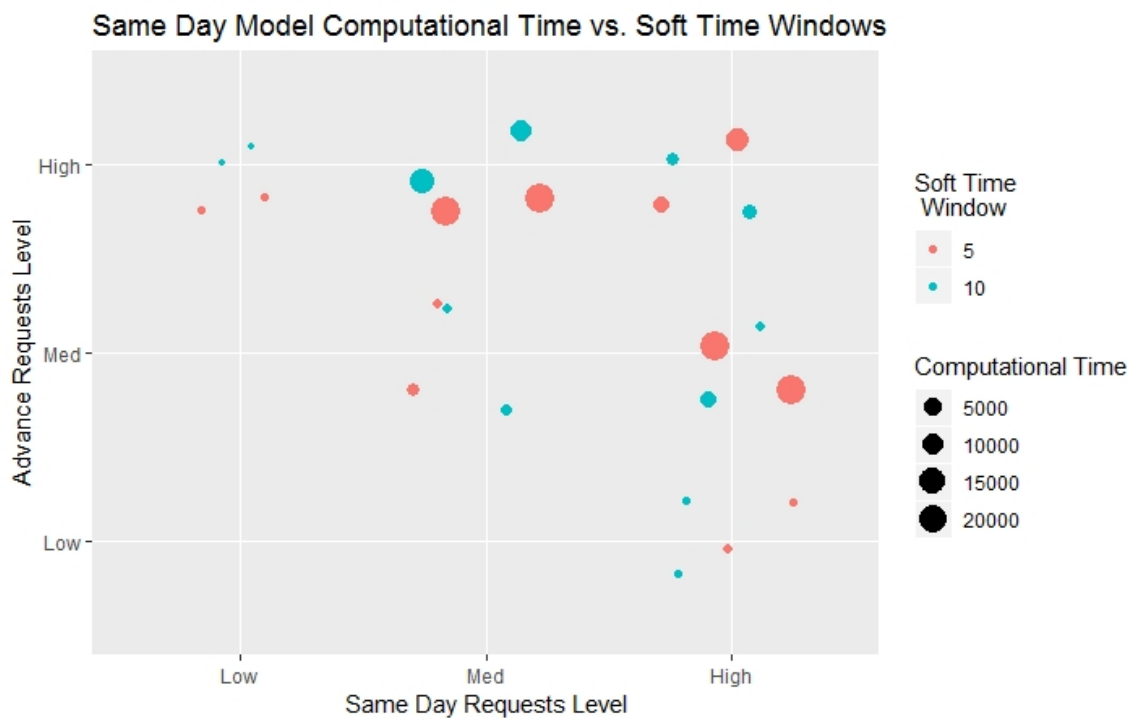


Figure 5.9: Same Day Model soft time windows computational time

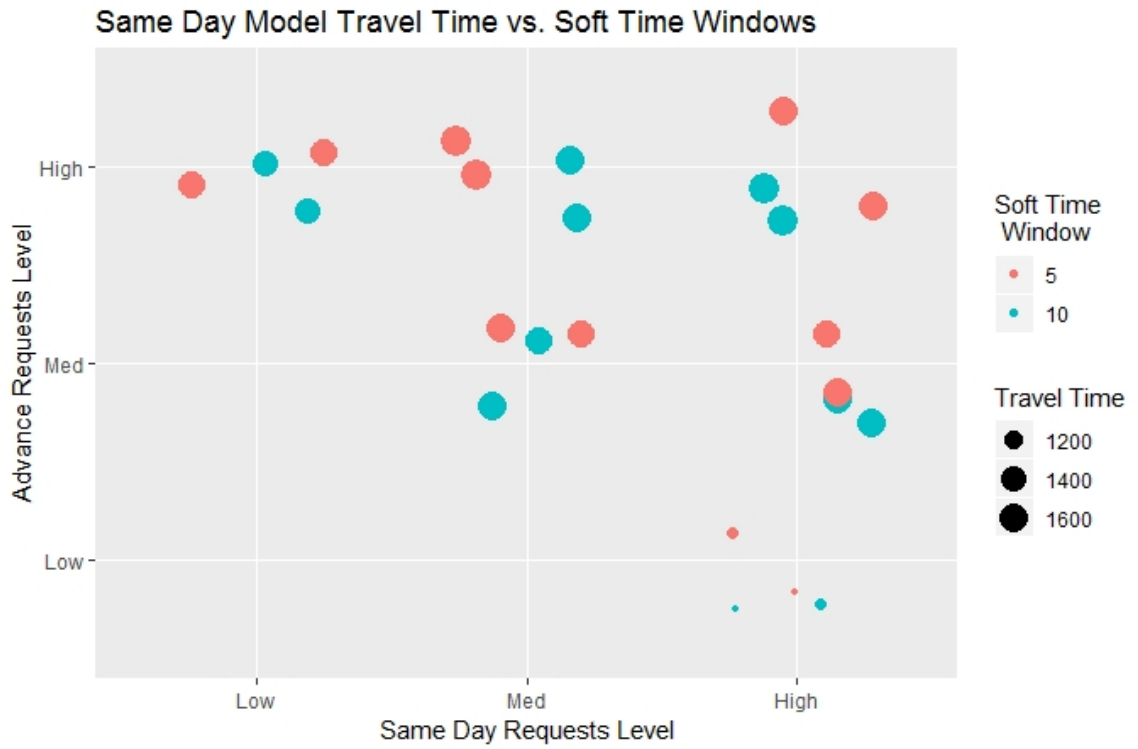


Figure 5.10: Same Day Model soft time windows travel time

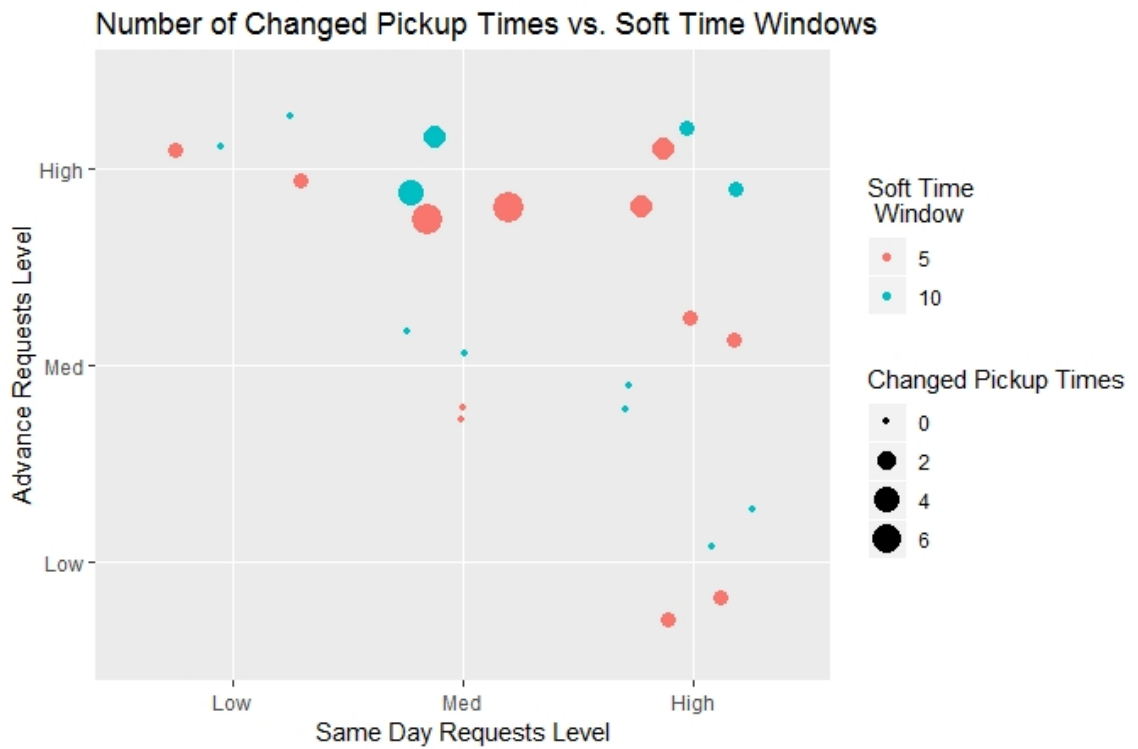


Figure 5.11: Same Day Model soft time windows changed pickup times

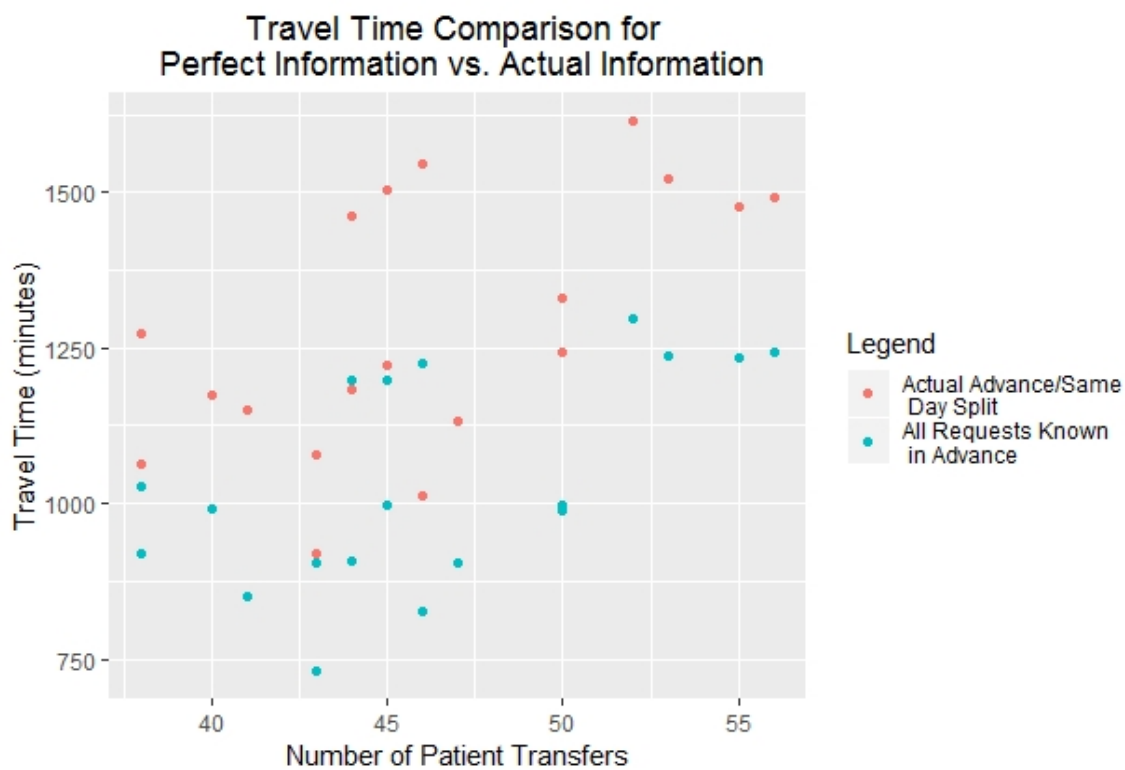


Figure 5.12: Perfect information vs. actual information

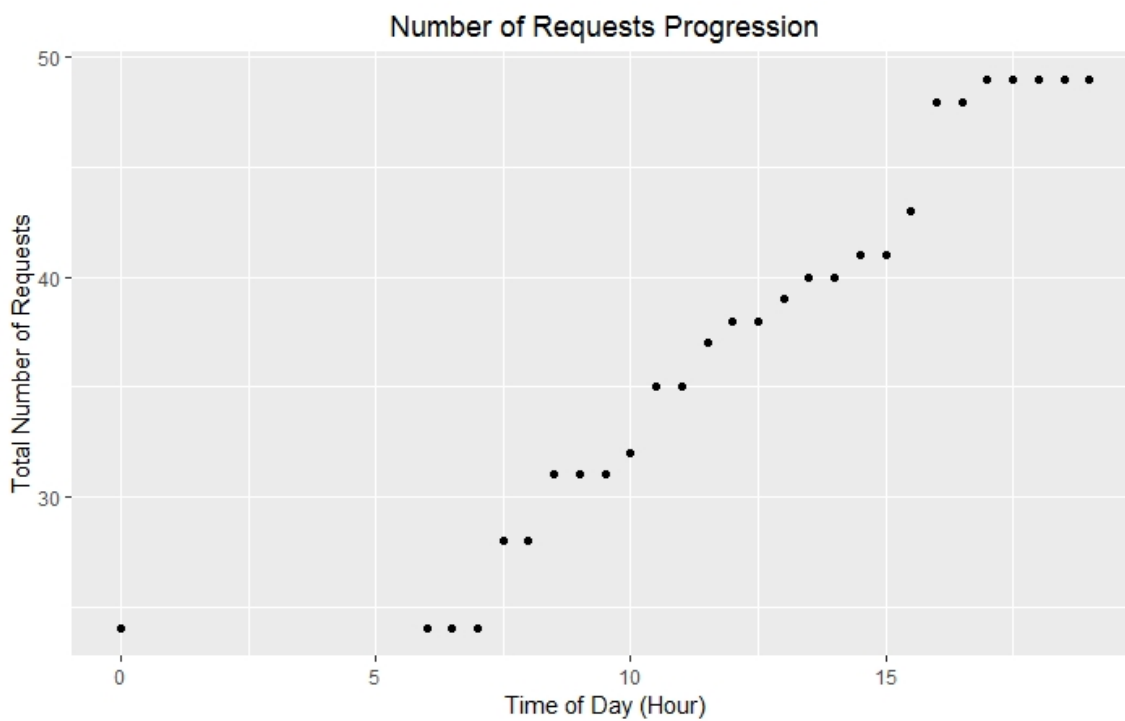


Figure 5.13: Total requests over the course of the day

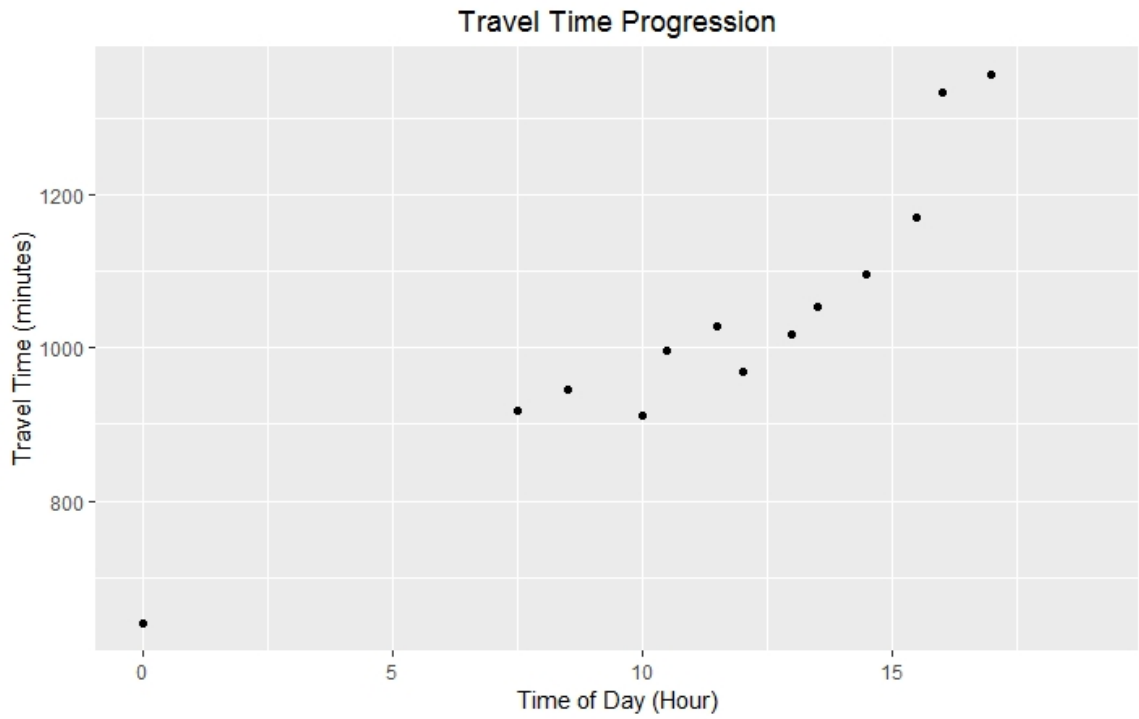


Figure 5.14: Total travel time over the course of the day

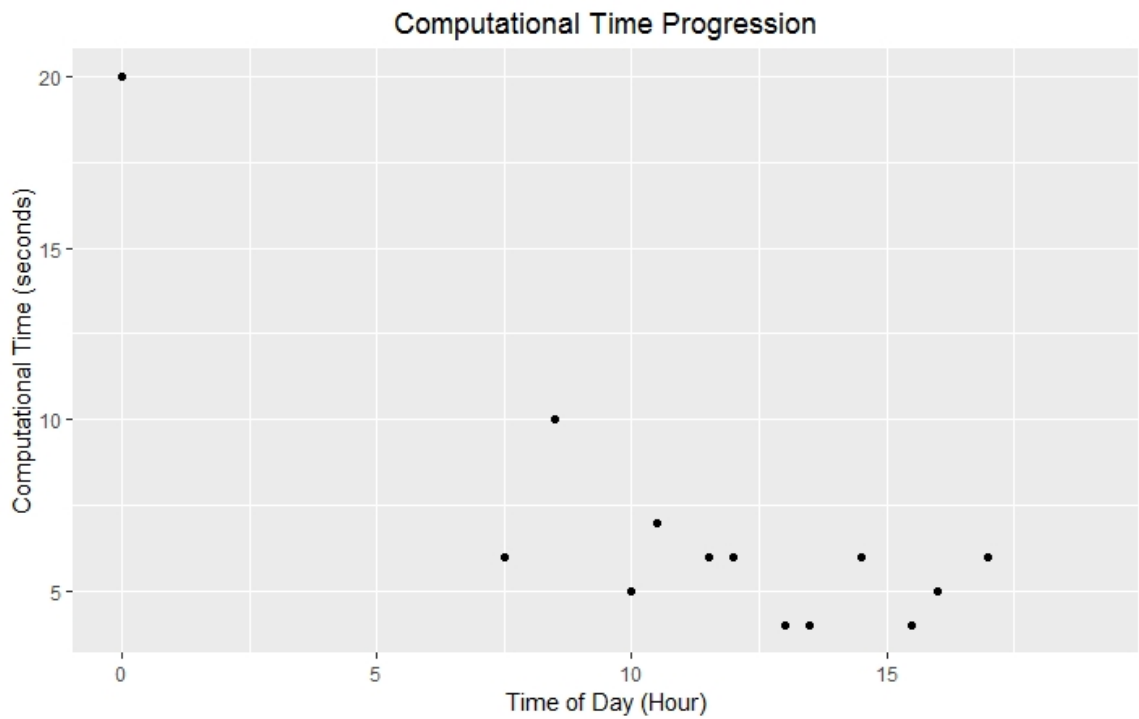


Figure 5.15: Computational time required for simulated day

Chapter 6

Conclusions

In this thesis, a non-emergency patient transportation system in Nova Scotia, Canada was explored. The scheduling and assignment process in place was examined. Patient transfer requests can be received before the requested pickup date; these requests are considered ‘Advance’ requests. They may also be received on the same day as the requested pickup date; these are called ‘Same Day’ requests. Schedulers typically take requests in a first come, first served manner and must balance current operations against new requests.

Literature on VRPs, including the DARP, was reviewed to understand how to model a scheduling and assignment problem for patient transfers. This included workload balancing constraints as EHS does not want to overload the paramedic crews. Research using data-driven models for VRPs was reviewed to understand how operations researchers are making use of the growing amount of available data as part of their research. A two-part scheduling model based on the DARP was created and tested. The first model is called the Advance Model and is designed to schedule all Advance requests. The objective function minimizes the total travel time of the vehicles. It generates a planned schedule for EHS to update as new requests are submitted.

The second model is called the Same Day Model, and it will generate a new schedule when Same Day requests have been submitted. The pickup times from the previous schedule (whether generated by the Advance Model or the Same Day Model) are used as an input. It minimizes the number of significantly changed pickup times from the previous schedule. A pickup time is ‘significantly changed’ if the pickup time is changed by greater than a certain number of minutes that is user specified.

Both the Advance and Same Day Models are influenced by two years worth of patient

transfer data provided by EHS, the organization responsible for patient transfers in Nova Scotia. Travel time estimates for the routes in both models are generated from the data. Missing data is filled in with estimates generated from the Google Maps API and the `googleway` R package. Service times at the pickup and delivery locations are estimated with a linear regression model developed with the patient transfer data.

Experiments were performed on the two models. For the Advance Model, the number of vehicles influenced the computational time required to solve the model slightly, while the number of requests greatly influenced the computational time. The time window length appears to be the main driver of the computational time; they should be kept as small as possible to ensure practical use of the Advance Model. The Advance Model was also shown to deliver better solutions in terms of travel times when compared (with some limitations) to the actual schedules of the PTUs. Workload balancing appears to slightly reduce the computational time required for the model, likely because it limits the solution space. None of the variables investigated affected the total travel time of the solutions. The experiments also demonstrated the feasibility of EHS using the Advance Model. With the ability to leave the model to solve overnight, EHS can potentially solve even large problems.

For the Same Day Model, the hard and soft time windows were examined for six experiments classified by the number of Advance and Same Day requests in the problem. The categories are High, Medium and Low. Unsurprisingly, the High category experiments all required longer computational time. The length of the hard time window did not appear to impact the computational time, the travel time or the number of changed pickup times. However, smaller soft time windows had a longer computational time in four of the six experiments, including every High category experiment. The smaller time window also resulted in a higher number of changed pickup times in every experiment except the Medium experiment, an expected consequence as the soft time window directly influences the objective function in the Same Day Model.

The value of information was shown to be quite valuable to scheduling patient transfers. Across 19 experiments, knowing every patient transfer request in advance when compared to the actual scenario was worth 19.4% of the time the vehicles spend travelling, or 248.6 minutes per day. This is the rough equivalent of the time required for another six patient transfers per day.

The final experiment simulated an operational day where the Same Day Model is used repeatedly. The Advance Model generates a schedule with all Advance requests, and the Same Day Model is rerun every 30 minutes with new requests (if new requests are received). Completed pickups are forced to match the previous schedule every time the Same Day Model is run. This process for using the Advance and Same Day Models is shown to be effective as the Same Day Model solves in ten seconds or less due to the low number of ‘unsolved’ requests in each iteration. One shortcoming of the strategy used is if a request requires an immediate pickup, waiting until the (somewhat arbitrary) 30 minute timeline to pass is impractical. In the experiment, there was one such call that asked for an immediate pickup. A better strategy is likely a combination of a threshold signal for running the model (either by the number of minutes passed or the number of new requests received) and running the model immediately when required. This final experiment also showed a 20% improvement in the total travel time if every request was known in advance, further demonstrating the value of perfect information for scheduling.

While the approach and methods described in this thesis are shown to be adequate for the patient transfer scheduling and assignment problem, there are shortcomings that can be improved upon in future work. The Advance Model is directly solved with the Gurobi solver, but other cheaper solvers may be of better use to health organizations, especially if finances are a concern. Different heuristics may be of use to solving the Advance Model as well.

The Same Day Model focuses on not changing the pickup times but does not take into consideration the objective of the Advance Model, the total travel time. There are

likely multiple optimal solutions for the Same Day Model as currently constructed, especially when there are a low number of total requests. These solutions may have different travel times (as shown in figure 5.14). Travel time could be incorporated into the objective function, and different solve methods such as heuristics may be used to solve this multi-objective function. Column generation may be one avenue to investigate for solving both models. Determining the effects of weights on these objectives is also of value. Both models are deterministic; stochastic methods should be investigated as travel times and service times have an element of randomness to them. Both models deal with the expected range of patient transfer requests well however.

One strategy for using the models during the online operation phase was considered in this thesis. The 30 minute threshold strategy has limits, and is not appropriate for any requests where the requested pickup time is sooner than the next scheduled run of the Same Day Model. Different strategies should be investigated. For instance, run the Same Day Model once a number of new requests have been received or run the model for every new request. In addition, different triggers for running the Same Day Model should be considered. For instance, when a PTU incurs a significant delay during a patient transfer, EHS may want to rerun the Same Day Model to determine the impact on operations. This would allow them to give future pickups earlier notice about a late pickup. The models can be used to show the effects on operations from changes to the system, such as additional capacity (more PTUs, paramedics) or new, high volume medical facilities.

One other consideration not explored in this thesis is the effect of the Advance Model schedule on the Same Day Model schedule, specifically how parameters such as the workload balancing parameters (wb^+ and wb^-) impact the future schedules. For example, a reasonable question to ask could be whether a wide range of workload balancing parameters would impact the quality of the updated schedule generated by the Same Day Model. For instance, if all of the Advance requests are located in one part of the municipality, it may make sense to assign all of these requests to one or two PTUs

(by turning the workload balancing constraints off) and save the rest of the PTUs for the Same Day requests as they appear around the city. If the Advance requests are scattered throughout the municipality, EHS may want workload balancing to ensure the PTUs are distributed evenly over the city so that any possible future requests are covered.

The Advance and Same Day Models have been shown to be feasible for online operational use. The Advance Model is based on of a DARP model and has workload balancing constraints added. The Same Day Model is also based on a DARP model but attempts to keep previously scheduled pickup times unchanged. This is based on the idea of using past decisions to influence future outcomes and capture the consequences to future operations when assigning limited resources in the present. Both models make use of patient transfer data as simple statistical models are used to estimate important parameters in the models. The data also helped influence the time period the models are used for and provided the data used for the experiments. These case studies provide insight to how the DARP works when used with real life scenarios and how EMS organizations can use these scheduling models to improve their operations.

Appendix A

Electronic Supplements

The file `TravelTimeGenerator.R`, found on DalSpace, contains the code to generate travel time estimates as described in section 4.2.1.

Bibliography

- [1] Raúl Baños, Julio Ortega, Consolación Gil, Antonio L. Márquez, and Francisco de Toro. A hybrid meta-heuristic for multi-objective vehicle routing problems with time windows. *Computers & Industrial Engineering*, 65(2):286 – 296, 2013.
- [2] Alexandre Beaudry, Gilbert Laporte, Teresa Melo, and Stefan Nickel. Dynamic transportation of patients in hospitals. *OR Spectrum*, 32(1):77–107, Jan 2010.
- [3] Fred Blakeley, Burçin Bozkaya, Buyang Cao, Wolfgang Hall, and Joseph Knolmayer. Optimizing periodic maintenance operations for schindler elevator corporation. *Interfaces*, 33(1):67–79, 2003.
- [4] David Cooley. *googleway: Accesses Google Maps APIs to Retrieve Data and Plot Maps*, 2018. R package version 2.7.1.
- [5] Jean-Francois Cordeau. A branch-and-cut algorithm for the dial-a-ride problem. *Operations Research*, 54(3):573–586, 2006.
- [6] Jean-Francois Cordeau and Gilbert Laporte. The dial-a-ride problem: models and algorithms. *Annals of Operations Research*, 153(1):29–46, 2007.
- [7] Conor Deasy and Iomhar O’Sullivan. Transfer of patients—from the spoke to the hub. *Irish medical journal*, 100:538–9, 01 2012.
- [8] Paolo Detti, Francesco Papalini, and Garazi Zabalo Manrique de Lara. A multi-depot dial-a-ride problem with heterogeneous vehicles and compatibility constraints in healthcare. *Omega*, 70:1–14, 2017.
- [9] Justin C. Goodson. Election day routing of rapid response attorneys. *INFOR: Information Systems and Operational Research*, 52(1):1–9, 2014.
- [10] Chris Groër, Bruce Golden, and Edward Wasil. The balanced billing cycle vehicle routing problem. *Networks*, 54(4):243–254, 12 2009.
- [11] Francesca Guerriero, Ferdinando Pezzell, Ornell Pisacane, and Luigi Trollini. Multi-objective optimization in dial-a-ride public transportation. *Transportation Research Procedia*, 3:299–308, 2014.
- [12] LLC Gurobi Optimization. Gurobi optimizer reference manual, 2019.
- [13] Isla M. Hains, Anne Marks, Andrew Georgiou, and Johanna I. Westbrook. Non-emergency patient transport: What are the quality and safety issues? a systematic review. *International Journal for Quality in Health Care*, 23(1):68–75, 2011.

- [14] Carl H. Häll, Jan T. Lundgren, and Stefan Voß. Evaluating the performance of a dial-a-ride service using simulation. *Public Transport*, 7(2):139–157, Jul 2015.
- [15] Elias W. Hans, Mark van Houdenhoven, and P.J.H. Hulshof. *A Framework for Healthcare Planning and Control*, pages 303–320. International Series in Operations Research & Management Science. Springer, 2012.
- [16] William E. Hart, Carl D. Laird, Jean-Paul Watson, David L. Woodruff, Gabriel A. Hackebeil, Bethany L. Nicholson, and John D. Sirola. *Pyomo—optimization modeling in python*, volume 67. Springer Science & Business Media, second edition, 2017.
- [17] William E Hart, Jean-Paul Watson, and David L Woodruff. Pyomo: modeling and solving mathematical programs in python. *Mathematical Programming Computation*, 3(3):219–260, 2011.
- [18] Nicolaus Henke, Jacques Bughin, Michael Chui, Tamim Saleh, Bill Wiseman, and Guru Sethupathy. The age of analytics: Competing in a data-driven world. Technical report, McKinsey Global Institute, Dec 2016.
- [19] Sin C. Ho, W.Y. Szeto, Yong-Hong Kuo, Janny M.Y. Leunge, Matthew Petering, and Terence W.H. Tou. A survey of dial-a-ride problems: Literature review and recent developments. *Transportation Research Part B: Methodological*, 111:395–421, 2018.
- [20] Chris Huggins and David Shugg. Non-emergency patient transport in victoria: An overview. *Australasian Journal of Paramedicine*, 6(4), 2015.
- [21] E. Hyttiä, S. Aalto, A. Penttinen, and R. Sulonen. A stochastic model for a vehicle in a dial-a-ride system. *Operations Research Letters*, 38(5):432 – 435, 2010.
- [22] Soumia Ichoua, Michel Gendreau, and Jean-Yves Potvin. Exploiting knowledge about future demands for real-time vehicle dispatching. *Transportation Science*, 40(2):211–225, 2006.
- [23] Jude Kornelsen, Kevin McCartney, Lana Newton, Emma Butt, and Marieka Sax. Rural patient transport and transfer: Findings from a realist review. Technical report, Applied Policy Research Unit, Nov 2016.
- [24] Ashish Kulshrestha and Jasveer Singh. Inter-hospital and intra-hospital patient transfer: Recent concepts. *Indian Journal of Anaesthesia*, 60(7):451–457, 2016.
- [25] Mengyu Li, Peter Vanberkel, and Alix J. E Carter. A review on ambulance offload delay literature. *Health Care Management Science*, Jul 2018.
- [26] C. L. Gwinnett M. J. G. Dunn and A. J. Gray. Critical care in the emergency department: patient transfer. *Emergency medicine journal*, 24(1):40–44, 2013.

- [27] Nikola Marković, Myungseob (Edward) Kim, and Paul Schonfeld. Statistical and machine learning approach for planning dial-a-ride systems. *Transportation Research Part A: Policy and Practice*, 89(C):41–55, 2016.
- [28] Nikola Marković, Rahul Nair, Paul Schonfeld, Elise Miller-Hooks, and Matthew Mohebbi. Optimizing dial-a-ride services in maryland: Benefits of computerized routing and scheduling. *Transportation Research Part C: Emerging Technologies*, 55:156 – 165, 2015. Engineering and Applied Sciences Optimization (OPT-i) - Professor Matthew G. Karlaftis Memorial Issue.
- [29] Piotr Matl, Richard F. Hartl, and Thibaut Vidal. Workload equity in vehicle routing problems: A survey and analysis. *Transportation Science*, 52:239–260, 2017.
- [30] M. Mourgaya and F. Vanderbeck. Column generation based heuristic for tactical planning in multi-period vehicle routing. *European Journal of Operational Research*, 183(3):1028 – 1041, 2007.
- [31] Mohammad Reza Nazari, Afshin Oroojlooy, Lawrence V. Snyder, and Martin Takáč. Deep reinforcement learning for solving the vehicle routing problem. *CoRR*, abs/1802.04240, 2018.
- [32] Nasim Nezamoddini and Mohammad T. Khasawneh. Modeling and optimization of resources in multi-emergency department settings with patient transfer. *Operations Research for Health Care*, 10:23–34, 2016.
- [33] Philip Oxley. Dial-a-ride: a review. *Transportation Planning and Technology*, 6(3):141–148, 1980.
- [34] Sophie N. Parragh. Introducing heterogeneous users and vehicles into models and algorithms for the dial-a-ride problem. *Transportation Research Part C: Emerging Technologies*, 19(5):912 – 930, 2011. Freight Transportation and Logistics (selected papers from ODYSSEUS 2009 - the 4th International Workshop on Freight Transportation and Logistics).
- [35] Harilaos N. Psaraftis. A dynamic programming solution to the single vehicle many-to-many immediate request dial-a-ride problem. *Transportation Science*, 14(2):130–155, 1980.
- [36] Harilaos N. Psaraftis, Min Wen, and Christos A. Kontovas. Dynamic vehicle routing problems: Three decades and counting. *Networks*, 67(1):3–31, 2016.
- [37] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.
- [38] Victoria Robinson, Vivek Goel, Russell D. MacDonald, and Doug Manuel. Inter-facility patient transfers in ontario: Do you know what your local ambulance is being used for? *Healthcare Policy*, 4(3):53–66, 2009.

- [39] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2016.
- [40] Ismail Saadi, Melvin Wong, Bilal Farooq, Jacques Teller, and Mario Cools. An investigation into machine learning approaches for forecasting spatio-temporal demand in ride-hailing service. *CoRR*, abs/1703.02433, 2017.
- [41] Open Data Nova Scotia. Hospitals: The geographic locations of all hospitals in nova scotia by their civic address, 2019.
- [42] David Stein. Scheduling dial-a-ride transportation systems. *Transportation Science*, 12(3):232–249, 1978.
- [43] Enrique Urrea, Claudio Cubillos, and Daniel Cabrera-Paniagua. A hyperheuristic for the dial-a-ride problem with time windows. *Mathematical Problems in Engineering*, 2015:12, 2015.
- [44] Pieter van den Berg and Theresia Van Essen. Scheduling non-urgent patient transportation while maximizing emergency coverage. *Transportation Science*, 01 2019.
- [45] Jason Wagner, Theodore J. Iwashyna, and Jeremy M. Kahn. Reasons underlying interhospital transfers to an academic medical intensive care unit. *Journal of critical care*, 28(2):202–208, 2013.
- [46] Semih Yalçındağ, Andrea Matta, Evren Şahin, and J. George Shanthikumar. The patient assignment problem in home health care: using a data-driven method to estimate the travel times of care givers. In *Flexible Services and Manufacturing Journal*, pages 41–55, 2016.