Master Thesis

Discrete Choice Modeling for Transfer Flow and Partial Multimodal Public Transportation OD Matrix Estimation

Lang, Glenn S. - *S2336502*
MSc Civil Engineering and Management

Faculty of Engineering Technology
August 15, 2024

tpg

UNIVERSITY OF TWENTE.

Colophon

Title	Discrete Choice Modeling for Transfer Flow and Partial Multimodal Public Transit OD Matrix Estimation
Location	Enschede and Geneva
Date	August 15, 2024

Author	G.S. Lang (Glenn)
Student Number	S2336502
Email	g.s.lang@student.utwente.nl
Institution	University of Twente (UT)
Faculty	Engineering Technology (ET)
Department	Transportation Engineering and Management
External Company	transports publics genevois (tpg)
Supervisor UT	A. Tirachini (Alejandro) *Associate Professor* Faculty of Engineering Technology
Supervisor UT	Prof.Dr.ir. E.C. van Berkum (Eric) *Professor* Faculty of Engineering Technology
Supervisor tpg	E. Tirindelli (Elisa) *Ingénieure Transports*
Supervisor tpg	D. Reck (Daniel) *Responsable Développement Mobilités*

Preface

From here on lies my Master Thesis, "*Discrete Choice Modeling for Transfer Flow and Partial Multimodal Public Transit OD Matrix Estimation*". This paper marks the conclusion of my Civil Engineering Master at the University of Twente and reports the findings of my research, which was jointly performed for the University of Twente and TPG. But before going into the finer details, I would like to thank all the people who supported me throughout this journey.

Firstly, I would like to thank my internal supervisors, Alejandro Tirachini and Eric van Berkum. You were continuously there to provide critical and helpful input, questioning my results, providing insights, and allowing me to think about my findings more in-depth. You were also always clear in your feedback and didn't hold back when you did not understand something, highlighting where I should improve and when I needed to rethink the communication of my methods or results.

Next, I would like to thank my external supervisors at TPG, Elisa Tirindelli and Daniel Reck. Elisa, thank you very much for all the tips, feedback, enthusiasm, and encouragement you showed me during our work together. You always brought a helpful opinion to the table, be it the subject matter, or general work and organizational advice, and I always appreciated it. Thank you, Daniel, although I couldn't hear your voice as often as I would have liked, your input was always helpful and brought a fresh perspective to the issues at hand, allowing me to view the problems from a different angle.

I would also like to thank my aunt and her partner, who let me stay in their home, from where I was able to go to the office in Geneva, allowing me to experience and utilize the public transportation network of the region and meet my supervisors and other colleagues at TPG.

Finally, I would like to thank my friends and family who were there to support me through this experience and showed me nothing but love and encouragement.

Glenn S. Lang
June 2024, Vieux-Ferrette

Summary

Public transportation is a key component of sustainable urban mobility and ensuring its efficiency and quality of service is pivotal for its success. To this end, analyzing travel demand and passenger flow through the system is an important task for public transportation providers, as they can identify where bottlenecks arise or the most important connections and propose countermeasures. The network-level Origin-Destination (OD) matrix is the primary tool that is used to enable this analysis, as it captures the origin and destination stop of a passenger, as well as the lines taken to arrive there. However, the process of estimating the network-level OD matrix is time-consuming, costly, and heavily data-reliant, with only the technological and methodological advancements in recent years allowing for effective and accurate estimation. This advancement in methodology came with the arrival of automatic data collection systems, such as automatic fare collection systems or GPS, which allowed public transportation providers to more effectively track individuals throughout their network. On the other hand, though these systems have become more widespread, not all of them are present in every public transportation system, and the estimation of the network-level OD matrix remains complex. In particular, one of the harder patterns to capture is the transfer flow at a station, where travel surveys, due to sample size, often do not capture the whole picture, leading to many structural zeros.

Two methodologies are proposed using Multinomial Logit Models, a travel sample, network data, boarding and alighting counts, and schedule data, with which the transfer flow patterns at a station may be estimated and analyzed, and subsequently utilized to construct the network-level OD matrix. The choice set of method 1 consists of all lines passing at the transfer station, to which a passenger would reasonably transfer, as well as exiting the network, indicating this is their final destination. The choice set of method 2 is then an expansion of that of method 1, where each line is disaggregated into the zones they stop in. To describe the transfer patterns, sets of 9 and 13 explanatory variables are defined, for methods 1 and 2, respectively, from which the tractable and statistically significant ones were singled out. These methods were applied to a chosen transfer station of the Geneva public transportation network, Lancy-Bachet. The results of this application show that the model fit is $\bar{\rho}^2 \geq 0.4$ for all models, which indicates that both methods can effectively explain the transfer and transfer-destination flow. Furthermore, it is shown that method 2 outperforms method 1, as it is less prone to errors.

In the case of both methods, the explanatory variables with the highest coefficient dominance and contribution to model fit are the Intercept, Transfer time, and Angle cost, with the addition of Travel time for method 2.

However, although the overall fit is shown to be excellent, both methods are subject to wrongly estimating low flows (< 10 passengers/hr), where errors are perceivable. Furthermore, though method 1 is shown to apply to Lancy-Bachet, it is found that at other stations of the Geneva network, the coefficient values lose interpretability. This behavior is not shared by method 2, where the calibrated coefficients are tractable.

It is thus concluded that, although the data requirement is higher, method 2 is more suitable for practical use in estimating travel demand and transfer flow, than method 1. In addition, it is found to more accurately explain transfer and travel behavior, and is recommended to be incorporated in network-level OD matrix estimation as required.

Table of Contents

1 Introduction

One of the key aspects of modern human life is our requirement to move between locations, be it when doing mandatory tasks, such as commuting to work, or for leisure outings. To enable this, humans make use of travel modes, such as a bike, a car, or public transportation, depending on factors like their physical capabilities or the distance required to be traveled.

Of these modes of transportation, the use of a personal car has been proven to be unsustainable and inefficient in mass transit, due to the resulting congestion (Lunke et al., 2021) and associated emissions, such as greenhouse gases (Lebrusán and Toutouh, 2020; National Geographic, 2019) and tire wear microplastic (Z. Luo et al., 2021). Thus, due to public transportation having the potential to reduce personal car use (Mohammed and Oke, 2023), it is a key player in sustainable (urban) mobility (Celikoglu and Cigizoglu, 2007; Zannat and Choudhury, 2019) and requires to provide a highly competitive service (Lunke et al., 2021). Furthermore, in the case an individual has no access to a car, public transportation enables this human need. This is why it is of high importance that the provided system run as effectively as possible, putting a major focus on the strategic, tactical, and operational planning of current public transportation management (Gkiotsalitis, 2022, Part III; Hussain et al., 2021).

Within transportation management, a key method to gain insight into network activity is the use of Origin-Destination (OD) matrices, which describe the flow of people between origin-destination pairs, also called demand, such as between city districts (zones) or stops. Thus, the estimation of OD matrices is oftentimes defined as the fundamental step in public transportation passenger-flow data analysis (Huang et al., 2020; Kong et al., 2023). More specifically, this designation is due to the OD matrix reflecting the spatial and temporal distribution of trips within a network (Blume et al., 2022). This enables transport planners to understand where, when, and how people travel, holding the potential to improve management levels (Kong et al., 2023) and allowing for the tailoring of services (Huang et al., 2020; Vanderwaart et al., 2017; Zannat and Choudhury, 2019).

When considering the benefits of OD matrices within the context of public transportation, two OD matrix types are considered, the route- and the network-level. For the former, only the demand between two locations on a single line, such as a bus, tram, or train line, is captured, whilst the latter describes the demand between two locations in the network, often considering multiple modes of transportation and routes, which allows for different levels of network optimization. At a route-level, the OD matrix may, for example, be used to decide whether a line requires an increased frequency during peak hours (Huang et al., 2020) or if certain stops can be removed from routes. On the other hand, a network-level OD matrix can be utilized to design new lines, extend existing ones (Vanderwaart, 2016), or synchronize schedules to allow for more streamlined transfers. Finally, pedestrian movements at stations can also be modeled using the transfer flows used for network-level OD matrix estimation, thus allowing for more optimal arrangements in platform assignments.

In addition to the general importance of OD matrices in public transportation planning, the rise of big data availability through automatic data collection (ADC) systems has been a major drive towards promoting and advancing research in the field (Cui, 2006; Mohammed and Oke, 2023). The collected data consists of three major types: (1) automatic passenger count data (APC), (2) automatic fare collection data (AFC),

and (3) automatic vehicle location data (AVL) (Lu et al., 2021; Zannat and Choudhury, 2019), though some recent research has been done with the intent to exploit new data sources, such as Wi-Fi data (Håkegård et al., 2018). Two of the main reasons for big data being a driving force in OD matrix estimation research are real-time estimation (Aguiléra et al., 2014), and its ability to capture the whole picture of passengers utilizing the transportation network, instead of having to rely on expensive and potentially unrepresentative surveys (Cui, 2006; Zannat and Choudhury, 2019).

On the other hand, although this data has become more readily available, not all public transport providers have implemented all these systems, and thus lack some of the data needed for the complete estimation procedure. Of the three mentioned data types, one of the more critical ones required for network OD matrix estimation is AFC data (Ji et al., 2015; Mohammed and Oke, 2023), and the estimation in the lack thereof has only been partially investigated in the literature (Blum et al., 2010; Zhang, 2008).

1.1 Problem context and data overview

Transports Public Genevois (TPG) is the public transport service provider of the Canton of Geneva, Switzerland. They are responsible for providing public transportation services to the 600′000 inhabitants of the region (World Population Review, 2023), in addition to the many commuters coming from the adjacent Canton of Vaud and France. Thus, they are required to manage the three levels of public transport planning (Gkiotsalitis, 2022). With this responsibility in mind, they desire to gain insight into the network-wide passenger demand by estimating the network-level OD matrix, so that they can efficiently improve their service.

To this end, from September 5th to October 14th, 2022, TPG initiated a (data-gathering) study. During this study, they gathered two types of data: web survey and GPS tracking data. Combined, there were about 26′000 respondents, who provided information on some of their performed trips. More specifically, the former asked participants where they entered the public transport network, which modes/lines they took, and where they exited the public transportation network. The latter was gathered through a smartphone application, which automatically recorded trip and trip-leg data, such as mode (bus, train, tram, walking), start and end stop, and, if applicable, public transportation line.

Additionally, a third type of data they gather is hourly aggregated APC data. To gather this data, TPG installed infrared counters at the entrances/exits of their vehicles. This data shows how many passengers boarded and alighted a line (bus/tram/train) at a given stop, within a given hour of the day.

Using the above data, Placiakis (2023) performed a study, where, for selected lines, line-level OD matrices were able to be constructed using either Iterative Proportional Fitting (IPF) or a Gravity Model, depending on the quality of the seed matrices. For each line, 4 OD matrices were defined according to the desires of TPG, where a distinction was made between week- and non-weekday, and peak and non-peak-hour. In this study, it was found that for Geneva, when a high-quality seed matrix is available, IPF should be preferred over the Gravity Model (Placiakis, 2023). Using this study, TPG went on to build route-level OD matrices for the whole network.

1.2 Research objective

Problem statement

As described in Section 1.1, TPG wants to estimate a network-level OD matrix using the two main types of public transportation data available, APC data and survey/GPS data. However, in the literature (see Section 2.2), unavailable Automatic Fare Collection (AFC) data is key to identifying transfers and expanding route-level OD matrices to the network level. Furthermore, network exits, which are a requirement for traditional OD matrix estimation techniques, are also dependent on AFC data to be reliably estimated. Finally, an issue lies with the survey data. Due to a low respondent rate, there is a potential for many transfers to be underrepresented, thus likely resulting in a false representation of transfer proportions and a false estimation of the transfer flow totals.

Research objective

As follows from the problem context, the key research objective of this study is to define and assess two methodologies, with which the transfer flow at a station or the partial network-level OD matrix can be estimated. More specifically, this study aims to provide a tool, with which to support public transportation network and transfer stop design decisions, as well as investigate the effects of transit system design, such as line frequencies and stop locations, on travel behavior.

Research questions

Considering these limitations and the objective, the questions below are formulated, where the first two aim to assess the applicability of the devised methodologies, while the third is a follow-up on the first two questions.

RQ1: Is the first devised methodology, which uses a Multinomial Logit model to estimate transfer proportions, applicable to estimate transfer flow totals for transfer stops within the bus-tram-train network of TPG, Geneva?

RQ2: Can the second devised methodology, which incorporates a Multinomial Logit model to infer destination, be effectively utilized to estimate (part of) the public transportation network OD matrix?

RQ3: Which of the devised methodologies is recommended to be used during the estimation of the network-level OD matrix of the Geneva public transportation network?

1.3 Thesis organization

The remainder of this Thesis consists of 9 sections. Section 2 is a literature review, that first describes the general approaches and requirements of traditional OD matrix estimation, followed by previously established methodologies utilized to estimate a network-level public transportation OD matrix. Following the literature review is the proposed methodology (Section 3), which describes the two methods developed and how they will be applied and compared to one another. Subsequently, Section 4 gives an overview of the chosen case study location within the Geneva public transportation network. Section 5 then shows the results for the calibration of the models developed for the two methods, after which Section 6 presents the

results for the transferability assessment, and Section 7 exhibits applications of both methodologies. Finally, Section 8 closes off this Thesis with conclusions, discussions, and recommendations for future research.

2 Literature Review

Though OD matrix estimation has been a research topic since the early 1970s (Mohammed and Oke, 2023), multimodal public transport OD matrix estimation has only been attempted since the late 1980s and, until the 2000s, with little discussion (Wong et al., 2005). Nowadays, the topic has been treated more explicitly, especially with the increased use of the aforementioned ADC systems (Fabbiani et al., 2018; Mohammed and Oke, 2023).

2.1 Traditional OD matrix estimation

A key component towards understanding public transport OD matrix estimation is traditional OD estimation techniques as the second/third stage of 4-stage transport modeling, which have been the focus of research for many years. Within these four stages, the first consists of determining trip attraction and production for the different zones of the study area based on a household travel survey (Blum et al., 2010; McNally, 2008; Ortúzar and Willumsen, 2011). During the second stage, a standard OD matrix estimation technique is then utilized depending on the available supplementary data and the conditions. Some of the more widely known and established methods, beyond public transportation, are the Gravity Model and Entropy Maximization (ME) (Blume et al., 2022; Gkiotsalitis, 2022, Part III; Mohammed and Oke, 2023), although more methods have been developed since. Next, during the third (modal split) stage, the OD matrix is subdivided between alternative mode choices, generally car, bike, and public transport, oftentimes using a discrete choice model (McNally, 2008; Ortúzar and Willumsen, 2011). Finally, the fourth stage is the route choice or traffic assignment stage, where, for each mode, OD flow is allocated to routes/links of the network (McNally, 2008; Ortúzar and Willumsen, 2011).

The resulting public transport OD matrix is then at a zonal level, whereas, with the rise of data availability, public transport providers often prefer working with stop-level OD matrices. This is due to the trip production and attraction points being stops belonging to the public transportation network and the people considered being strictly the ones utilizing the network (Blum et al., 2010). To determine these public transport OD matrices, a different process is utilized, with the general framework described in Section 2.2.

2.2 Public transportation network OD estimation

One of the first and more widely known frameworks, established for the estimation of stop-level public transportation network OD matrices that exploits the available big data, is the one developed by Cui (2006). This method uses three types of data and consists of three stages, including the data input stage, as depicted in Figure 1, which is a modern adaptation of the process by Mohammed and Oke (2023). Beyond the depicted three stages, this process depends on inferring three key components: trip origin, trip destination, and transfer flows (Mohammed and Oke, 2023).

2.2.1 3-stage OD matrix estimation

The first stage of the method by Cui (2006) is the data input stage. In this stage, three types of data are utilized, APC, AFC, and supplementary data, of which the former two have been mentioned previously (see Section 1) and are utilized as the main source for understanding passenger behavior. The latter category,

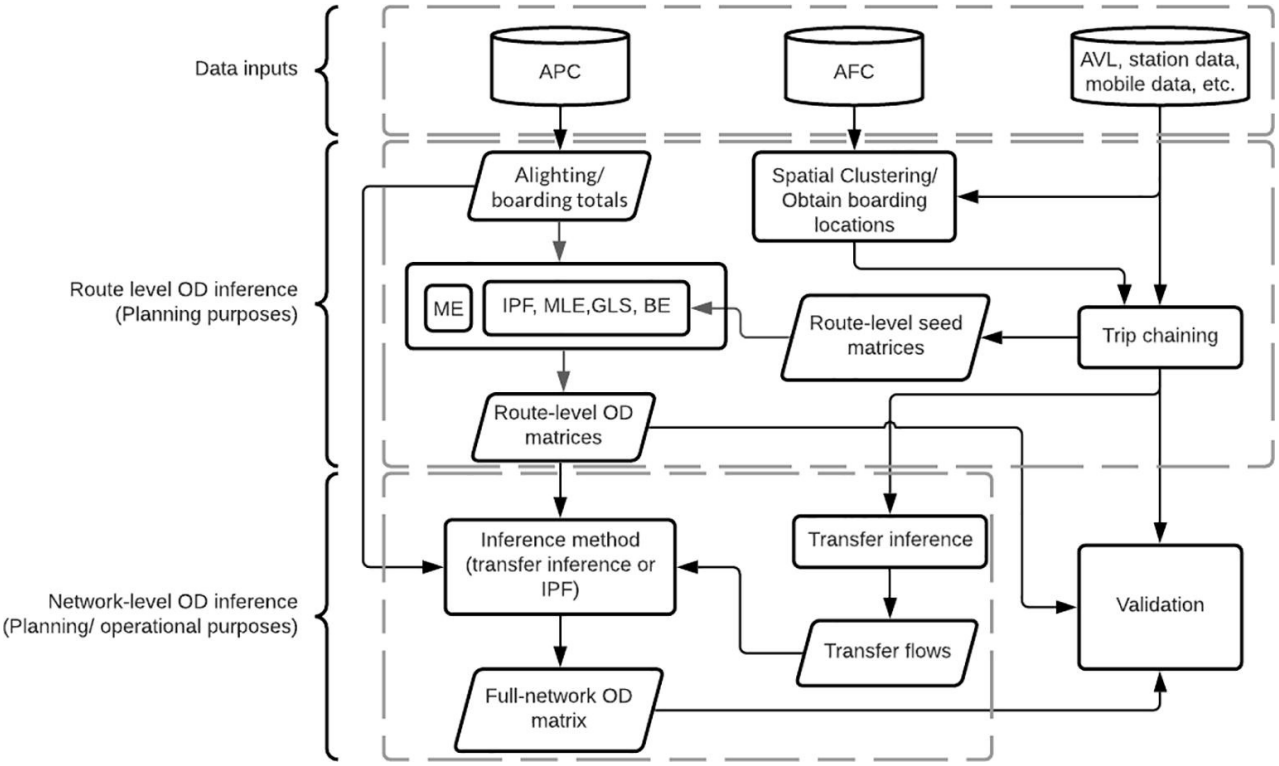


Figure 1: General network-level OD estimation process (source: Mohammed and Oke, 2023)

supplementary data, is called as such because it can take on many forms, and is not necessarily required, depending on the quality and completeness of the former two types. AVL and station (or stop location) data are mainly used to help with the inference of trip chaining (Cui, 2006; Mohammed and Oke, 2023), though the latter can also be utilized to build a travel cost matrix for the route-level OD matrix estimation (Placiakis, 2023).

During the second stage, the main aim is to utilize the data to build route-level OD matrices. To this end, Cui (2006) mentions two primary methods, Iterative Proportional Fitting (IPF) and Maximum Likelihood Estimation (MLE), although a wider variety of methods is available nowadays, such as Entropy Maximization (Gkiotsalitis, 2022, Part III; Mohammed and Oke, 2023), Bayesian Estimation (Blume et al., 2022), Generalized Least Squares (Mohammed and Oke, 2023), Compressed Sensing (Kumar et al., 2019), Fluid Modeling (Blum et al., 2010), or Neural Networks (Celikoglu and Cigizoglu, 2007). Though differing in their process, one of the shared similarities and requirements for these trip distribution methods is that trip production and/or attraction counts are present, in this case in the form of boarding and alighting counts, such that there are not infinite solutions and the methods converge.

In the third stage, the network-level OD matrix is derived by expanding the route-level OD matrices with transfer flows, inferred from AFC and supplementary data. To estimate the transfer flows, passenger trajectories are followed through the network by looking at where they boarded, alighted, and potentially reboarded (Mohammed and Oke, 2023). To then estimate whether a passenger has transferred or initiated a new trip, two subsequent trips (alighted and boarded) are said to be chained if the boarding occurred within a time threshold, and the boarding station is within a maximum distance of the alighting station. If these criteria

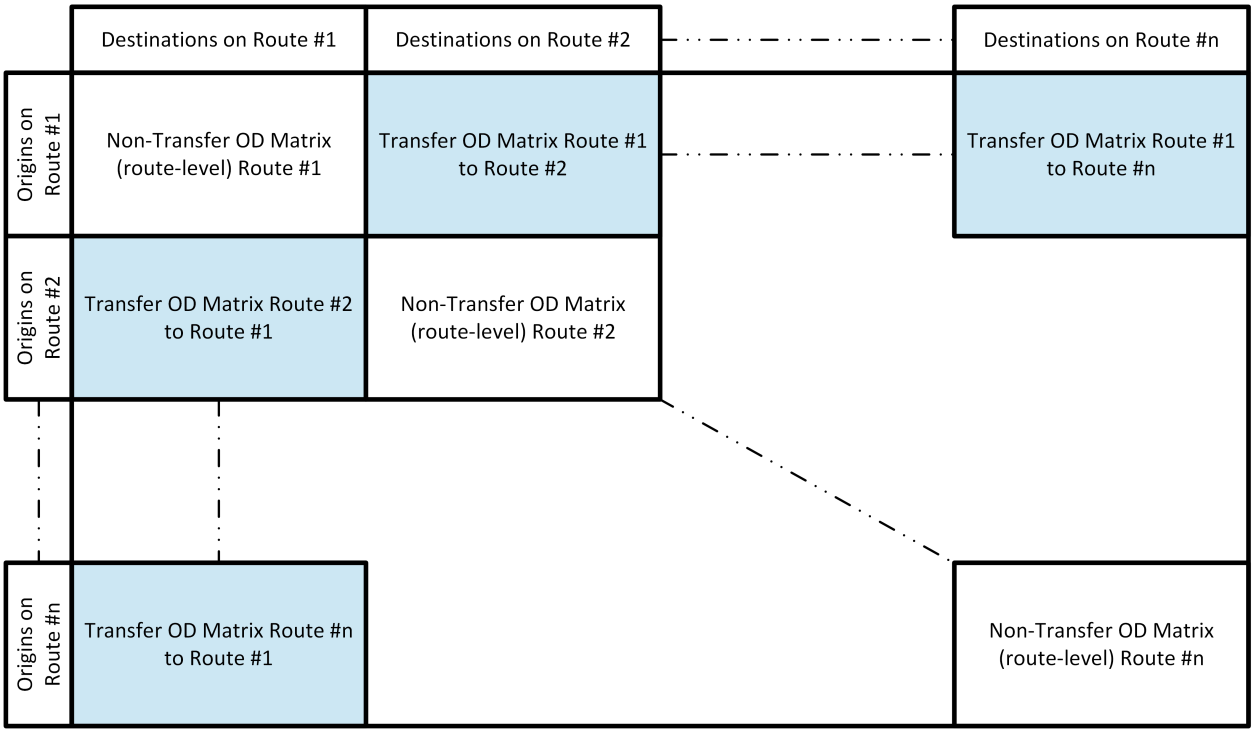


Figure 2: Network OD matrix (adapted from Cui, 2006)

are met, the passenger is assumed to have performed a transfer (Liu et al., 2021; Mohammed and Oke, 2023). Aggregating all these together for each combination of arriving and departing line is what then results in the transfer flows (Cui, 2006). Finally, when estimating transfer flows, Cui (2006) constrained his method to only consider the possibility of transferring once.

Using these two inputs, transfer origins and destinations need to be estimated using one of two methods, Proportional Distribution (PD) or Modified Iterative Proportional Fitting (MIPF) (see below). These are split into three steps, the first and second leg of the trip, and leg-linking. The first two steps aim to estimate the origin of the first leg and the destination of the second leg, respectively, whilst the third is self-explanatory (Cui, 2006; Zhang, 2008). Then, to compute the network-level OD matrix, first, the route-level OD matrices need to be adjusted through the removal of the transferring passengers. Secondly, utilizing an all-ones matrix in conjunction with the marginal transfer origins and destinations estimated prior and IPF, the transfer OD matrices can be computed (Cui, 2006). In this case, no seed matrix is utilized, as, according to Cui (2006), the transfer OD flows are usually low, thus resulting in a low-quality seed matrix. Combining the route-level and transfer OD matrices leads to the network OD matrix shown in Figure 2.

Proportional Distribution (PD)

PD has two requirements, the total transfer flow at the considered stop and the route-level OD matrix for the trip leg. Using these, at a stop s, when considering the transfer producing line a (first leg) and transfer receiving line b (second leg) with total transfer flow $t_{s;ab}$, the number of passenger with their origin at stop i of line a can be determined using Equation 1.

$$o_i = t_{s;ab} \frac{d_{a;is}}{\sum_{j \in R_{a;0s}} d_{a;js}} \tag{1}$$

Where $R_{a;0s}$ is the set of stops at which line a stops before stop s, $d_{a;ks}$ is the number of passengers on line a having their route-level origin as $k \in R_{a;0s}$ and destination as s.

The number of passengers with destination i of line b can then similarly be determined using Equation 2, where $R_{b;s}$ is the set of stops at which line b stops after stop s.

$$d_i = t_{s;ab} \frac{d_{b;si}}{\sum_{j \in R_{b;s}} d_{b;sj}} \tag{2}$$

Modified Iterative Proportional Fitting (MIPF)

MIPF can be employed if IPF is utilized to estimate the route-level OD matrix of the previous stage (Cui, 2006). If that is the case, pseudo-stops are added to the matrix estimation problem, which represent a transfer to a separate line. To this end, at a transfer stop, the number of boarding or alighting passengers, depending on the analyzed leg, needs to be adjusted according to the number of passengers transferring. The transferring passengers are thus removed from the boarding/alighting count where they originated and set as the boarding/alighting count of the pseudo-stops. An example of the OD matrix modification is shown in Figure 3. After the necessary pseudo-stops have been added and the APC counts are adjusted, standard IPF can be utilized to re-estimate the OD matrix.

	A	B	C	TOTAL
A	0	13	6	19
B	0	0	17	17
C	0	0	0	0
TOTAL	0	13	23	36

		A	B		C	TOTAL
			Exit	To 1	To 2	
A		0	10	1	2	6	19
B	Origin	0	0	0	0	12	12
	From 1	0	0	0	0	2	2
	From 2	0	0	0	0	3	3
C		0	0	0	0	0	0
TOTAL		0	10	1	2	23	36

Figure 3: Example of normal to modified route-level OD matrix with two possible transfer lines (1 and 2) at stop B

Application of Cui's method

Shortly after Cui (2006) developed his method, Zhang (2008) applied it to a subset of the Central Ohio Transit Authority (COTA) public transportation network, intending to assess the applicability of MIPF and PD in the third stage. In this study, three types of data were available and a subset of two bus lines was chosen to apply the methodology on. As can be noted by the data available, no AFC data could be utilized, meaning that the total transfer flows had to be determined differently. To this end, Zhang (2008) devised a two-step scheme, where first bus pairs are matched on whether or not a transfer is possible or likely. Subsequently, for each matched pair, an upper boundary is set on the possible total transfer flows based on the minimum between the alighting and boarding counts of trip leg 1 and 2, respectively. From this upper boundary, the total transfer flow is estimated using an arbitrary (predefined) percentage. To the knowledge of the author,

this two-step scheme is the only account in the current literature where transfer flow totals are estimated without the use of AFC data or supplementary household survey data. This is due to the non-triviality of the estimation when lacking AFC data, as mentioned by Ji et al. (2015), since APC data does not provide clear details on where passengers enter and exit the public transportation system.

2.2.2 Alternative methodologies

Four years after the development of the method by Cui (2006), Blum et al. (2010) developed a workflow towards estimating a network-level OD matrix. For their case study, APC and household survey data, where the latter is aggregated at a zonal level, were combined to build both zone- and stop-level network OD matrices by using a two-step process. The first step consists of determining route-level OD matrices using a Fluid model (Blum et al., 2010), after which in the second step, a minimization problem is optimized where the objective is the weighted sum of the Sum Squared Error (SSE) between three modeled and observed components. These components are the zonal OD matrix entries, the boarding counts, and the alighting counts. To compute the first component, a stop-level OD matrix is built using a hyperpath model (Blum et al., 2010; Mandl, 1980), which can be aggregated to a zonal OD matrix by exploiting the spatial position of each stop (Blum et al., 2010).

A more recent approach towards estimating network-level OD matrices was developed by Liu et al. (2021). In this study, three mathematical optimization models were developed, a quadratic integer program (QIP), a quadratic convex program (QCP), and a linear integer program (IP), where, for each stop, the objective is to minimize the L2 or L1 norm between the observed and modeled transfer proportions. These three models were then compared by employing a real-world scenario. To utilize the models, Liu et al. (2021) make use of APC data, AVL data, bus schedules, and transfer proportions, which are either directly estimated from observations (AFC data) or expert analysis. The results of this study showed that the IP performed equally or even better than the QPs.

Finally, a public transportation OD matrix estimation approach that must be mentioned is a method utilizing entropy maximization which has been developed by Guex et al. (2023). This method makes use of the boarding and alighting counts in combination with entropy maximization to estimate the network-level public transportation OD matrix.

2.3 Choice modeling

Choice modeling, or discrete choice models (DM), are models that are built to determine choice probabilities for individuals based on rational decision-making (Ortúzar and Willumsen, 2011). They are more often used in the context of the modal split stage of traditional transport modeling, rather than for the purpose of OD matrix estimation (Clifton et al., 2016; Ortúzar and Willumsen, 2011), although it has been shown that a singly-constrained gravity model can be rewritten as a discrete choice model (Ortúzar and Willumsen, 2011). Furthermore, more recent studies have been performed to utilize choice models for the estimation of pedestrian OD matrices or for discrete route-choice modeling, as other factors, besides travel cost, have been deemed to influence such decisions (Clifton et al., 2016; Raveau et al., 2011).

To build a discrete choice model, two primary components are necessary, a finite set of choices (Q) and a set of choice attributes (H). These are utilized to define a utility function (U_{jq}) for each alternative $q \in Q$ and individual j considered, which is composed of a deterministic part (V_{jq}) and a random part (ϵ_{jq}) (Ortúzar and Willumsen, 2011). Please note that, for this research, the individual level is left out. The general form of the utility function for alternative $q \in Q$ is shown in Equation 3 and 4, where θ_i is a weight associated with attribute value η_i for $i \in H$, and with $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ being vectors of length $|H|$. More specifically, attributes are the characteristics of choice alternatives that make an option more or less attractive (increasing or decreasing its utility), such as travel time or ease of access.

$$U_q = V_q + \epsilon_q \tag{3}$$

$$V_q = \sum_{i \in H} \theta_i \eta_i \tag{4}$$

Additionally, it must be noted that attribute values may take on a non-linear form. Commonly used in those cases are the Box-Cox or Box-Tuckey transformations, with the Box-Tuckey transformation allowing for values smaller or equal to zero (Ortúzar and Willumsen, 2011). The form which this transformation can take on is described in Equation 5.

$$f(x) = \begin{cases} \ln(x+\mu), & \text{if } \tau = 0 \\ x, & \text{if } \tau = 1 \\ \frac{(x+\mu)^\tau - 1}{\tau}, & \text{otherwise} \end{cases} \tag{5}$$

Where x is the non-transformed attribute value, τ is the pre-defined transformation constant, and μ is a constant ensuring all values $(x+\mu) > 0$.

Using these components, the two main types of choice models utilized in practice are the Multinomial Logit (MNL) and the Multinomial Probit (MNP) models (Kropko, 2007; Ortúzar and Willumsen, 2011).

2.3.1 MNL and MNP models

The primary and simpler model of the two is the MNL model, which only makes use of the deterministic part of the utility function (V_q) for each alternative (Kropko, 2007; Ortúzar and Willumsen, 2011). Thus, the probability of selecting alternative $q \in Q$ using a MNL model is defined by Equation 6, where, according to Ortúzar and Willumsen (2011), β is generally set to 1, due to the inability to distinguish it from the attribute weights $\boldsymbol{\theta}$.

$$P_q = \frac{e^{\beta V_q}}{\sum_{j \in Q} e^{\beta V_j}} \tag{6}$$

Comparatively, the alternative MNP model has the advantage of being able to introduce random variation among individuals, though it becomes non-trivial to utilize due to requiring simulation to approximate the required parameters (McCulloch et al., 2000; Ortúzar and Willumsen, 2011). Furthermore, the subsequent determination of choice probabilities is highly complex and computationally intensive, as it requires an additional nested integration per alternative. This makes it unsuitable for use with large choice sets (Kropko, 2007).

2.3.2 Model calibration

For choice models, the parameter estimation generally consists of determining the utility function coefficient vector $\boldsymbol{\theta}$. To this end, when a random sample (of size n) of observed choices ($q_i \in Q$) is available, the general approach is to use Maximum Likelihood Estimation (MLE) (Ortúzar and Willumsen, 2011), where $\boldsymbol{\theta}$ is set to be the random variable.

In this case, the sample ($X_1 = q_1$, $X_2 = q_2$, ..., $X_n = q_n$), with probability density function $p_x(q_i;\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is unknown, has the likelihood described in Equation 7.

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} p_x(q_i;\boldsymbol{\theta}) \tag{7}$$

It is important to note that the Log-Likelihood, as outlined in Equation 8, is frequently employed because the resulting estimator is equivalent, and probability density functions often involve exponential functions (Larsen and Marx, 2012, Chapter 5).

$$l(\boldsymbol{\theta}) = \ln\left(L(\boldsymbol{\theta})\right) = \ln\left(\prod_{i=1}^{n} p_x(k_i;\boldsymbol{\theta})\right) \tag{8}$$

The estimator $\boldsymbol{\theta}_e$ (the vector of coefficient values) which maximizes l, can then be determined by setting the derivative of l to 0 and solving for $\boldsymbol{\theta}$, as shown in Equation 9 (Larsen and Marx, 2012, Chapter 5). Furthermore, to ensure it is a maximizer, it must be shown that the Hessian at $\boldsymbol{\theta}_e$ is negative definite (Larsen and Marx, 2012, Chapter 5; Sheffi, 1985, Chapter 2).

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta_i} = 0 \quad \forall i \in H \tag{9}$$

2.4 OD matrix estimation using discrete choice modeling

When utilizing a discrete choice model for OD matrix estimation (during the trip distribution step), it is designated as a disaggregate or random utility model (Cascetta et al., 2007; Zhu and Ye, 2018). Its use in this context allows for more flexibility by considering significant variables, thus potentially enabling better model fit than conventional methods. However, due to the nature of OD matrices and the generally large sizes of study areas, where there can be upwards of 1 million zones, and thus choice sets of that size, it is often not possible to consider it, as the computational requirement is too large. To overcome this issue, some methodologies, such as the one by Cascetta et al. (2007), have been developed to aggregate zones together and reduce the choice set.

In the context of public transportation OD matrix estimation, to the author's knowledge, choice modeling has only ever been utilized for route choice modeling, such as by Clifton et al. (2016), or to identify mode choice (Limtanakool et al., 2006; Wang et al., 2016). However, in the case of the latter, these mode choices are often part of a secondary step in traditional four-step modeling, or as a second level of a nested Logit model, where the car is also part of the equation, such as in the method developed by Wang et al. (2016).

To fill this gap, the author attempts to develop a method, which combines both destination, mode, and route choice, for the public transportation network OD matrix estimation of the Geneva network.

3 Methodology

To answer the research questions and finally fulfill the research objective, a methodology has been devised, where two approaches are proposed. The first aims to improve upon the strictly naive approach, where the transfer proportions are estimated directly from the survey, whilst the second method is an extension of the first. These two methods have been defined, as no approach has previously been clearly established to estimate the transfer flows when no AFC or household survey data is available. The lack of AFC data means that no clear counts are available to estimate the OD matrix traditionally, with the lack of household survey data ruling out the method established by Blum et al. (2010). Furthermore, previously established methods lack concrete investigation of their applicability, and, due to a low respondent rate for the survey, there is a high likelihood of non-structural zeros and low representativeness of the whole study area. The methods could thus be used to enrich the transfer flow and network-level OD matrix and more accurately represent the flow within the network. Finally, the devised methods could allow for the determination of all transfer flows based on a survey taken at a single location, which would substantially reduce the effort, time, and money required to estimate the model, thus enabling more frequent and accurate determining of the OD matrix.

The devised methods, of which a general outline is shown in Figure 4, build upon the same principle, to estimate a MNL model at a good survey location using network and route attributes, which can be transferred to other locations in the network. The first method, where the choice set is defined as lines an alighting passenger can transfer to at a given stop, aims to estimate transfer flows by including line information in the MNL model, such as transfer time and distance, attractiveness, and route-choice properties. Thus, being an intermediate step between line-level and network-level OD matrices, where after the transfer flows are estimated, the methods by Cui (2006) can be applied. The second method aims to combine the transfer flow and network-level OD estimation by integrating destination information into the MNL model. This is done by adding travel time as an attribute and altering the weighted potential. The choice set in this case is extended from lines one can transfer to, to the reachable destinations on those lines, where a destination is aggregated at a zonal level. Using this MNL model, the second leg of the trip can directly be inferred, after which the final destination stop can be distributed. In the case of both methods, it is assumed that no more than single transfer trips are possible, as non- and single-transfer trips make up more than 96% of them. Finally, it is important to note that both methods combine destination and route choice.

The first step of the research methodology consists of building and applying a framework with which to process the data such that it can be used for the rest of the research. Following the processing of the data, the feasible transfer sets, one for the morning and one for the evening peak, are to be estimated, which will be used by both methodologies. After this, the first transfer flow estimation method will be implemented, calibrated, and validated, followed by the implementation, calibration, and validation of the second, partial network OD matrix estimation method. Finally, both methods will be applied to different stops, where their performance will be assessed and compared.

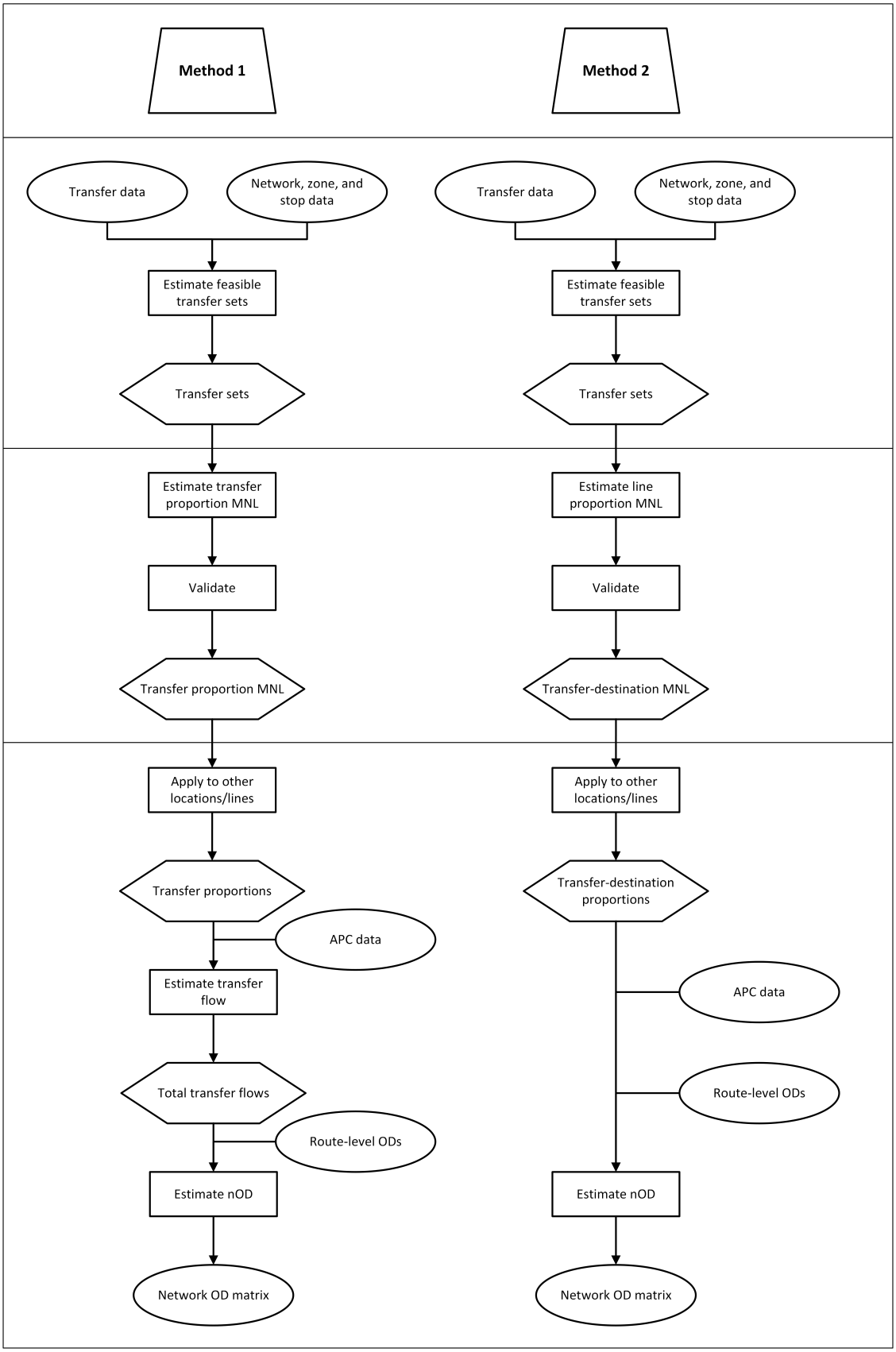


Figure 4: Outline of methods

3.1 Data preparation

The available and used data consists of 6 main types, which TPG has partially preprocessed. Of this data, only three have to be explicitly processed, with the other three, APC, line-level OD matrices, and stop-distances at the line level, only requiring explicit filtering for the location and time of interest. The three that require to be preprocessed are thus the survey data, the GTFS data, which holds the line schedules and stop locations, and the population census data.

Additionally, it must be noted that the APC flows will be rebalanced using the method devised by Guex et al. (2023), such that the boarding and alighting flows obey the consistency conditions. These entail that the total boardings and alightings are equal and that the alightings at the first stop and boardings at the last stop are 0 (Guex et al., 2023).

3.1.1 Survey data

The survey data, of which an excerpt is depicted in Table 1, is at an individual level and consists of single trip leg entries. As depicted, some leg entries are split into two or more possible options (lines), with an associated uniformly distributed weight. In these cases, it will be assumed that a weighted number of passengers performed said trip. The required output of this preparation is a transfer-exit seed matrix for each line and stop, i.e., whether a passenger exited the public transport network or transferred to another line.

Table 1: Excerpt of survey data

User_id	Leg_id	Date	Origin	Mode	Destination	Weight	Time	Line
20822	1	27/08/2022	Genève, Gautier	bus	Genève, gare Cornavin	0.5	23	25
20822	1	27/08/2022	Genève, Gautier	bus	Genève, gare Cornavin	0.5	23	1
20822	2	27/08/2022	Genève, gare Cornavin	tramway	Genève, Poterie	1	23	3

To process this data, it will be aggregated per (grouped) arrival stop, where for each line, the passengers arriving at the stop will be counted. From this information, the passengers will be separated based on the lines they arrive on, and the transfer seed matrix will be built.

3.1.2 GTFS

The GTFS data consists of multiple datasets for the entirety of the Swiss public transport network, each updated weekly for the considered study period (September 5th - October 15th, 2022), with the key datasets for this study, their association (shared fields), and their attributes shown in Figure 5. From this data, three outputs are required: (1) the geographical location of each stop in the study area, (2) the interarrival times between different lines at each stop, and (3) the stops that are grouped as transfer stops or are considered one. To produce these outputs, first, the data is filtered to only consist of TPG-relevant information based on the agency ID, which can be added to the other datasets by exploiting the relationships. After this, the first and third outputs can be extracted, which are just the filtered stop and transfer datasets.

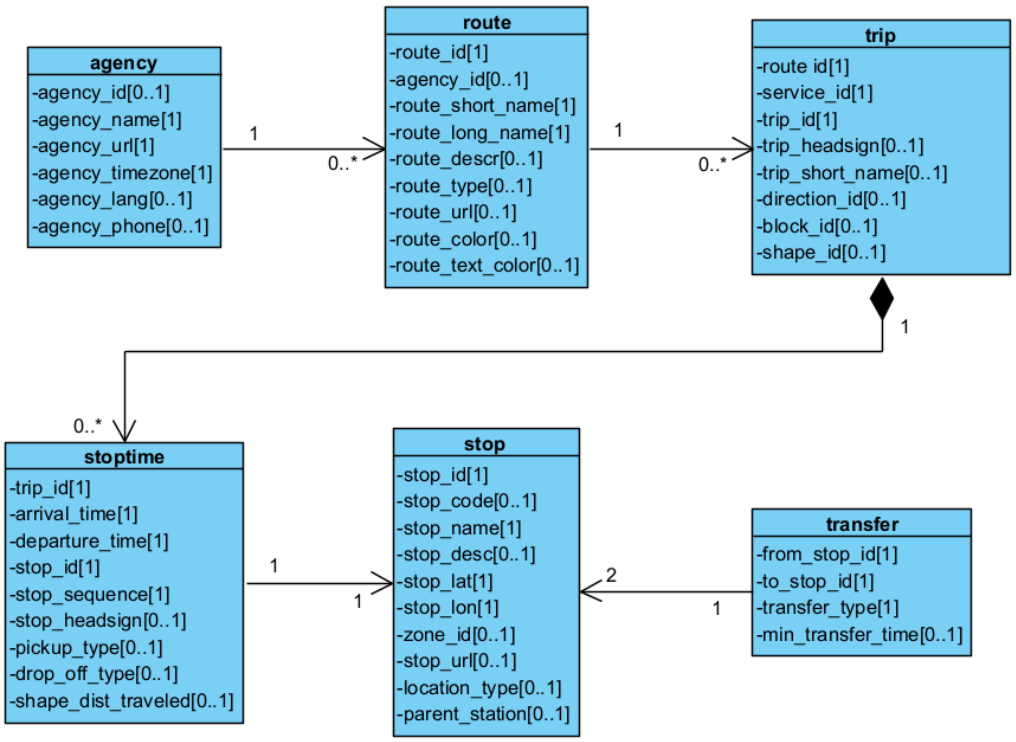


Figure 5: GTFS dataset relationships

To extract the second output, the *stoptime* dataset can be split at the stop level, including stops to which can be transferred, by considering the transfer dataset. For each line at a stop, the difference between its arrival time and the first departure of each other line can then be computed and used as the interarrival time.

3.1.3 Population census and employment data

The population census and employment data are at a zonal level, where each zone has an associated population and employment count. To make use of this data, each stop in the network needs to identify with one of the zones (sous-secteur or commune), from which it will inherit the corresponding value. To this end, the general workflow utilized by Blum et al. (2010) is followed, for which the geographical locations of both the stops and the zones are required. The former can be retrieved from the previous data preparation step described in Section 3.1.2, whilst the latter was retrieved from opendata.swiss (2023b) for the canton of Geneva, from data.gouv.fr (2022) for the french border region, and opendata.swiss (2023a) for the canton of Vaud (for more details on the data sources utilized, see Appendix A). To identify to which zone a stop belongs, two scenarios have to be considered. First is complete encapsulation, in this case, the stop is entirely in a zone and not potentially part of a neighboring zone, thus directly indicating a one-to-one relationship. The second case is when the stop is placed between two zones, and it is thus unclear which zone it belongs to. In these cases, Blum et al. (2010) assign these stops to multiple zones, which will also be done for this research, although the population count assigned to that stop will be an average of the zones it belongs to.

3.2 Feasible transfer set

The first step for both methodologies towards estimating the network-level OD matrix is the estimation of the feasible transfer set. In this case, the feasible transfer set is defined as the set of lines $j \in L_s$ (where L_s is the set of lines stopping at transfer stop s) to which line i can transfer to, i.e., the set of transfer-receiving routes $T_{s;i} \subset L_s$. Analogous, the transfer-receiving set $I_{s;j} \subset L_s$ is the set of lines $j \in L_s$ that can transfer to a line j. To determine these two sets, three parameters must be defined, the minimum transfer time (MnTT), the maximum transfer time (MxTT), and the maximum transfer distance (MTD), which are also utilized in standard trip chaining methods with AFC data (Liu et al., 2021; Mohammed and Oke, 2023). However, in the case of this study, the Geneva public transportation network has dedicated stops which are defined as transfer stations, thus, the MTD is not a requirement for matching lines. Finally, some rules concerning same-line transfers must also be defined.

In the context of trip chaining, the most often utilized MxTT is 30 minutes (Kumar et al., 2018; Liu et al., 2021), which will also be utilized for this case study. With regards to the MnTT, Kumar et al. (2018) state that a threshold of 1.7 minutes allows for high accuracy in the results, though for this case study, 2 minutes will be used as the base MnTT. In addition to this base, the threshold will be increased by an estimated walking time if an individual has to walk between the arriving and departing stops. To this end, the straight line distance between the two stops is utilized, with a walking speed of $1 m/s$ or $3.6 km/h$, which is slightly lower than the average (Bohannon, 1997; Bohannon and Williams Andrews, 2011; Montufar et al., 2007), as no consideration is made towards vertical distances or walking on the quay.

Finally, additional rules have been defined that restrict the choice set. The first is often used in the trip chaining context and stipulates that a passenger may not transfer to the same line, be it a following bus going in the same direction or going back where they came from (Liu et al., 2021). The second is that an individual may not transfer to a line that has served the same upstream stops or that will serve the same downstream stops as the transfer-producing line. The third rule is specific to the train lines of the network (LEX). Due to the nature of their restricted routes, with lines $L1$ to $L4$ and lines $L5$ and $L6$ serving very similar stops respectively, interrail transfers have been restricted. More specifically, if a passenger arrives on lines $L1$ to $L4$, the transfer set does not include those lines, with the same applying if one arrives on lines $L5$ or $L6$. The last rule is that one may not transfer to a line which has its final destination as the transfer stop.

As the APC data is in hourly aggregates, the transfer time between two lines must be aggregated as an hourly average, as described by Zhang (2008). Thus, at a stop (or transfer stop), all lines that have an average transfer time between the MnTT and the MxTT will be matched. More formally, the relations in Equations 10 and 11 are utilized, where the average transfer time between lines a and b at stop s is $t_{s;ab}$.

$$T_{s;i} = \left\{x \mid \text{MnTT} \leq t_{s;ix} \leq \text{MnTT}, \quad \forall i, x \in L_s\right\} \tag{10}$$

$$I_{s;j} = \left\{x \mid \text{MnTT} \leq t_{s;xj} \leq \text{MnTT}, \quad \forall x, j \in L_s\right\} \tag{11}$$

3.3 Method 1: Total transfer flow estimation

The first method consists of estimating the transfer flows per stop and line using an MNL model before using MIPF or PD to expand to a network-level OD matrix. First, the MNL must be built and calibrated at

a high-quality location, followed by a transferability assessment at a secondary (or multiple other) high-quality location. In this case, a high-quality location is a stop where there are a high number of respondents and a large percentage of transfers. Using the MNL model, transfer proportions can be estimated and subsequently expanded to transfer flows. Furthermore, if found to be applicable, the calibrated model can then be utilized at locations where the survey data is of lower quality and where transfer proportions are less reliable.

3.3.1 Transfer proportions

The transfer proportion MNL model determines the probabilities that a passenger alighting from a transfer-producing line l at transfer station s chooses to transfer to line j for all j in $T_{s;i} + \{\Pi\}$. In this case, the choice set $T_{s;i} + \{\Pi\}$ consists of all possible transfer-receiving lines $T_{s;i}$, as well as the network exit Π (no transfer).

To build the transfer proportion MNL model, nine choice attributes (explanatory variables) have been defined, which aim to capture factors that a passenger may deem important (Clifton et al., 2016; Raveau et al., 2011). The final model will be built from the most statistically significant and well-behaved (tractable) ones.

Explanatory variables
The first of the attributes is the transfer time between a transfer-producing line l and a transfer-receiving line x at a stop s ($t_{s;lx}$). It is defined as the average minimum time difference between the arrival of a trip for line l and the subsequent departure of line x during a considered time period (morning peak, evening peak, etc.). This attribute aims to capture the unwillingness of passengers to wait excessive amounts of time to transfer, as it contributes to an increased total travel time (Raveau et al., 2011).

The second attribute is the transfer distance ($d_{s;l}$). This attribute aims to capture how far people have to walk if they wish to transfer to a line x when arriving on line l. Due to time and data limitations, this distance is based on the Euclidean distance between the two stop locations and does not consider whether there are vertical obstacles that must be traveled.

The next two are indices, which aim to represent the attractiveness of a line as proxies and are based on aggregated measures of spatial accessibility with their roots in gravity modeling (Luo and Wang, 2003). The first of the two attractiveness indices is titled "Decaying Population Density Index" (or PI) and is defined in Equation 12, whilst the second is titled "Decaying Workplace Density Index" (or WI) and is defined in Equation 13.

$$\mathrm{PI}_{s;x} = \sum_{i \in R_{x;s}} g_i \delta_{x;si}^{-\beta_p} \tag{12}$$

$$\mathrm{WI}_{s;x} = \sum_{i \in R_{x;s}} w_i \delta_{x;si}^{-\beta_w} \tag{13}$$

Where PI_x is the Decaying Population Density Index of transfer-receiving line x, WI is the Decaying Workplace Density Index of transfer-receiving line x, $R_{x;s}$ is the set of stops visited by line x after the current stop s, g_i is the population density of the sous-secteur or commune stop i is in (in inhabitants per km^2), analogous to g_i, w_i is the workplace density (in workplaces per km^2), $\delta_{x;si}$ is the travel time (in seconds) to stop i from s using line x, and β is a decay factor.

These two indices are separated, as passengers move for different reasons depending on the time of day considered. More specifically, individuals commute to work during the morning, whilst in the evening, they go back home. This leads to a difference in the perceived utility of alternatives depending on the time of day, which is to be represented by the use of a different Density Index.

Additionally, as an alternative to the two previously defined indices, an index based on the downstream alighting counts of the transfer-receiving line has been defined. This index aims to be used as a direct measure of the attractiveness, rather than through a proxy attribute, in addition to simply exploiting the available data. Its definition is the same mathematically as for the other two proxies and is shown in Equation 14.

$$\mathrm{AI}_{s;x} = \sum_{i \in R_{x;s}} a_{i;x} \delta_{x;si}^{-\beta_a} \tag{14}$$

Where $\mathrm{AI}_{s;x}$ is the alighting index of transfer-receiving line x, $a_{i;x}$ is the number of passengers disembarking from line x at stop i, and β_a is a decay factor.

For the three described attractiveness descriptors, PI_x, WI_x, AI_x, it must be noted that the Box-Tuckey transformation may be applied, such that a better fit for the model may be found. Furthermore, as described by Ortúzar and Willumsen (2011), the transformation from the Gravity Model to the MNL model for OD estimation results in the number of trip destinations of an alternative in the utility function being Box-Cox transformed with $\tau = 0$.

The angle cost attribute is derived and adapted from the decision choice model defined by Raveau et al. (2011), where a passenger chooses the most direct path to their destination. In this case, a passenger is assumed to continue their journey in the same (or close to the same) direction their original trip was going to, while they would generally not go back whence they came (turning 180°). To this end, the arriving angle ($\varphi_{x;s}$) and departing angle ($\alpha_{x;s}$) of a line x at a stop s are defined in Equations 15 and 16, as well as shown in Figure 6.

$$\varphi_{x;s} = \frac{1}{|S_{x;s}|} \sum_{i \in S_{x;s}} \varphi_{si} \tag{15}$$

$$\alpha_{x;s} = \frac{1}{|R_{x;s}|} \sum_{i \in R_{x;s}} \alpha_{si} \tag{16}$$

Where $S_{x;s}$ is the set of stops line x visits before stop s, φ_{si} is the angle formed by a straight line from stop s to stop i and a vector pointing south, $R_{x;s}$ is the set of stops visited by line x after stop s, and α_{si} is the angle formed by a straight line from stop s to stop i and a vector pointing north. To construct these angles, the true locations of the stops have been utilized.

Using these two angles, the angle cost (AC_{lx}) between a transfer-producing line l and transfer-receiving line x is defined in Equation 17. The usage of the cosine ensures that an angle of 180° or −180° has a cost of 1, which decreases the closer the departing angle is to the arriving angle, where an angle of 0° has a cost of 0.

$$\mathrm{AC}_{s;lx} = 0.5 - \frac{\cos(\varphi_{l;s} - \alpha_{x;s})}{2} \tag{17}$$

The feasible transfer ratio ($\mathrm{TR}_{s;lx}$) is defined to complement the transfer time ($t_{s;lx}$) between two lines, as well as capture an effect that is not considered when aggregating the minimum transfer time. It is the

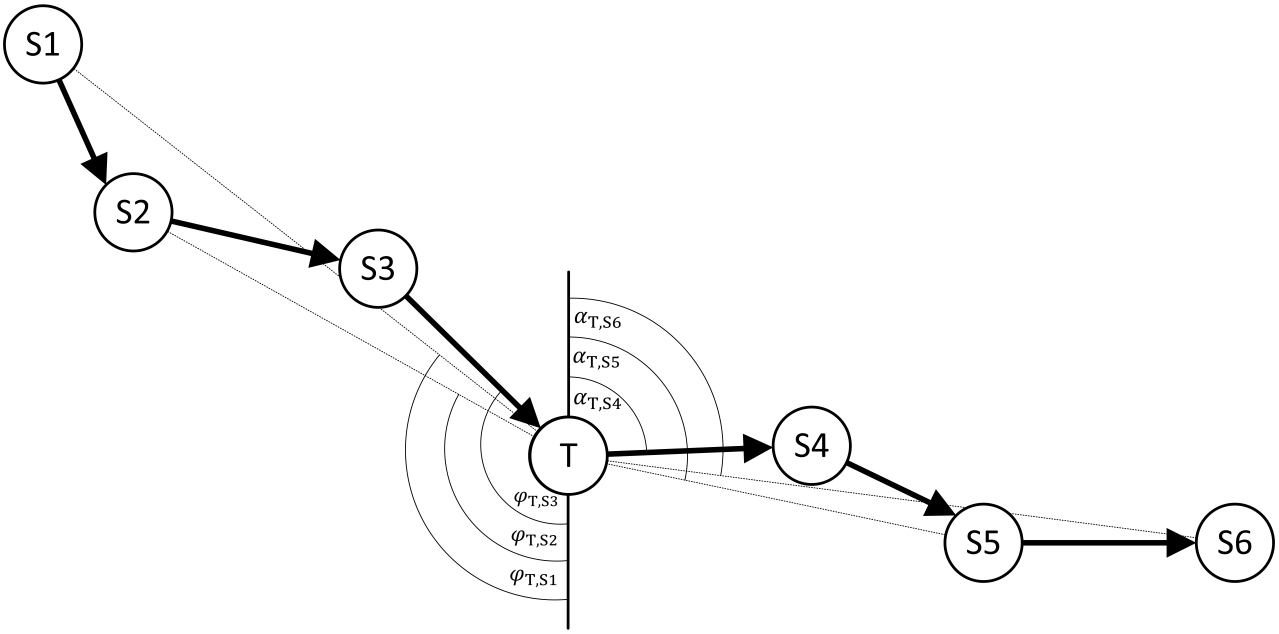


Figure 6: Example of arriving and departing angle of a line at stop T

ratio between the number of transfers possible and the total number of arriving trips, as is depicted by the example in Table 2. In this example, it is shown how the transfer ratio between line 1 going in direction 0 and line 2 going in direction 1 is computed at a certain stop. As depicted, during the considered period, line 1_0 arrives at the stop 3 times, whilst line 2_0 departs 5 times, which are both denoted by Ti. Based on the trips, their arrival and departure times, as well as the minimum and maximum transfer time (MnTT and MxTT), it can be determined whether or not an incoming trip can transfer to a departing trip, generating the feasible transfer grid.

Table 2: Feasible transfers between line 1_0 to line 2_1 as stop E (−: not feasible, X: feasible)

		Line 2_1				
		T1	T2	T3	T4	T5
Line 1_0	T1	−	X	X	−	−
	T2	−	−	−	X	X
	T3	−	−	−	−	−

The feasible transfer ratio between line 1_0 to 2_1 is thus $TR_{E;1_0\rightarrow 2_1} = 2/3$, as two of the three incoming trips can transfer.

The final two attributes are two dummy variables to represent the mode of the arriving line (train, tram, or bus). The first dummy (TRAIN) is set to 1 if the arriving line is a train or 0 otherwise, whilst the second (TRAM) follows the same behavior but for trams. This leads to the bus being the reference category.

In addition to the attributes, it must be noted that the utility function and corresponding coefficients of a transfer-receiving line will be separated by mode (bus, tram, train, and network exit). This is because individuals perceive modes differently on a qualitative level (Scherer, 2010; Scherer and Dziekan, 2012),

such as trains, which are often perceived as more reliable and comfortable than other modes of public transportation. This disaggregation also allows for further fine-tuning of the attraction values for exiting the network at the current stop and not transferring. Finally, this disaggregation allows for the network exit to only use the intercept as an attribute. This is due to the other explanatory variables having values that are either zero or tending to infinity, thus rendering them redundant or incompatible with the utility function. Thus their exclusion from the utility function is necessary for the proper calibration of the model.

Transfer sample

In addition to the utility function, the survey data for all lines (and directions) arriving at stop s is available in the form shown in Table 3. Each row denotes a transfer-producing line and each column a potential transfer-receiving line (including Π). It is important to note that, in this example, all lines at the stop are be depicted, with some destination lines possibly not being part of the feasible transfer sets of all origin lines.

Table 3: Example of transfer data

	1_0	1_1	2_0	2_1	3_0	3_1	Π
1_0	0	0	3	2	7	0	15
1_1	0	0	1	0	2	0	12
2_0	4	2	0	0	6	0	20
2_1	1	2	0	0	6	9	25
3_0	2	0	1	2	0	0	7
3_1	3	4	1	0	0	0	12

Using this utility function, the probability of transferring from line l to line q can be computed using Equation 18, for which the coefficient vector $\boldsymbol{\theta}$ can be estimated using MLE and the transfer data.

$$P_{lq} = \frac{e^{V_{lq}}}{\sum_{j \in Q} e^{V_{lj}}} \tag{18}$$

Model verification

After having estimated the MNL model at the first survey location using MLE, it first has to be verified at the current location. To this end, four performance indicators have been selected. The first indicator is McFadden's Pseudo-$\bar{R}^2$ (McFadden and Domencich, 1975; Mokhtarian, 2016; Ortúzar and Willumsen, 2011), also known as $\bar{\rho}^2$, which indicates the goodness-of-fit of a regression, where values from 0.2 onwards indicate excellent fit (McFadden, 1977).

The second indicator selected is the Mean Transfer Error (MTE), which is adapted from Guex et al. (2023) and represents the average percentage error in the estimated transfer flow. It is computed using Equation 19. In combination, the root-mean-square error (RMSE), computed using Equation 20 (Glen, n.d.; Shcherbakov et al., 2013), will be used to indicate the average flow over- and underestimation of the model.

$$\text{MTE} = \frac{\sum_{ij} |E_{ij} - O_{ij}|}{\sum_{ij} O_{ij}} \tag{19}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{ij} (E_{ij} - O_{ij})^2} \tag{20}$$

Where O_{ij} and E_{ij} are the observed and predicted transfer flow from line i to j, respectively, and n is the number of combinations ij.

The final performance indicator for the verification is the Pearson Correlation Coefficient (PCC), where transfer flows are lower than 5 (McClave and Sincich, 2018). This indicator aims to assess whether or not the predicted and observed transfer flows, below that threshold, are correlated. This is done, as about 95% of the flows are located below that threshold, thus being the flows that require the most attention. Furthermore, high transfer flow values may skew the results in favor of the model when they are accurately predicted.

Model transferability

After the verification and goodness of fit test, the transferability of the model will be determined by differentiating two types of transferability, coefficient and framework transferability. The former assesses whether or not the coefficients calibrated at a first location can be utilized at a second location, without requiring additional regressions. The second describes whether the methods developed are applicable in other locations, where the separate regressions result in interpretable model coefficients.

For the coefficient transferability, the statistical difference between different location coefficients has to be assessed. To this end, the model coefficients will be estimated at the validation locations using MLE, after which the statistical test devised by Cohen et al. (2003) and Paternoster et al. (1998) will be utilized. For this test, the null-Hypothesis (H_0) is that the two coefficients estimated from regressions performed at two different locations (β_1 and β_2), for the same attribute, are the same, i.e., $\beta_1 = \beta_2$. For a significance level α, this hypothesis is rejected if $|z| > z_{\alpha/2}$, where z is computed using Equation 21.

$$z = \frac{\beta_1 - \beta_2}{\sqrt{(\mathbf{SE}\beta_1)^2 + (\mathbf{SE}\beta_2)^2}} \tag{21}$$

Where $\mathbf{SE}\beta_1$ and $\mathbf{SE}\beta_2$ are the standard error of the two regression coefficients, respectively.

3.3.2 Transfer flows

Using the transfer proportions estimated in the previous step and APC data, the total transfer flows can be estimated per stop. To this end, first, the initial transfer totals can be computed per line $i \in L_s$ using Equation 22, where $a_{s;i}$ is the number of alighting passengers from line i at stop s. Additionally, it must be noted that the number of passengers exiting the network at this station can also be computed using Equation 22, only that $j = \Pi$ instead.

$$b_{s;ij} = P_{s;ij} a_{s;i} \tag{22}$$

After the initial transfer counts are computed, the estimated transfer-boarding totals ($B_{s;j}$) can be computed as the sum of all transfers to a line j, as shown in Equation 23.

$$B_{s;j} = \sum_{i \in I_{s;j}} b_{s;ij} \tag{23}$$

The estimated transfer-boarding totals can then be compared with the APC boarding counts $\omega_{s;j}$, where two cases are possible: (1) $B_{s;j} \leq \phi_1 \omega_{s;j}$ and (2) $B_{s;j} > \phi_1 \omega_{s;j}$ ($0 < \phi_1 \leq 1$ is a selected parameter which defines what percentage of the boarding counts the transfer flow can maximally make up). In the case of (1), the

difference $\Delta_{s;j} = \omega_{s;j} - B_{s;j}$ is set to be the number of passengers entering the public transportation network at stop s. In the case of (2), the transfer probabilities are adjusted, such that the computed boarding counts are within a margin of error (ϵ) of the actual boarding counts. In this case, this is achieved by uniformly reducing the transfer proportions on all transfer-producing lines for j and increasing the exiting proportion. Due to the reliability of the infrared data gathering system in Geneva, an error of $\epsilon = 0.02$ is utilized (Placiakis, 2023).

3.3.3 Network-level OD matrix

After the transfer flow OD matrix has been estimated, the route-level and total transfer flows can be expanded to the network level by applying either MIPF or PD, as described in Section 2.2.1.

3.4 Method 2: Combined estimation

The second method consists of a similar workflow, with the need to build, calibrate, and validate a MNL model, though the transfer flows are not explicitly estimated, with the destination zone (administrative district) being a part of the MNL model. The general workflow is the same as for method 1, first, the MNL model is estimated, after which its performance and transferability are assessed. The MNL can then be used to build a partial network OD matrix and OD matrix using the boarding and alighting counts.

3.4.1 Transfer-destination proportions

To build the MNL model, a similar setup as for method 1 will be utilized, though the utility function will be for each line-line-zone, i.e., per transfer-producing line (l), transfer-receiving line (x), and destination on the transfer-receiving line ($k \in Z_{s;x}$ where $Z_{s;x}$ is the set zones that line x stops at downstream of stop s). To this end, 11 attributes have been defined, of which a subset will be selected depending on the statistical significance, as well as their tractability.

Explanatory variables
The first set of five attributes are strictly transfer attributes and are as described earlier for method 1: The time between the arrival of l and x ($t_{s;lx}$), the travel distance to the departure stop of x ($d_{s;lx}$), the feasible transfer ratio, which denotes the number of trips arriving for l capable of transferring to line x ($TR_{s;lx}$), and the two arrival mode dummies, i.e., whether or not the transfer-producing mode is a train or a tram.

The next set of three attributes are adaptations of the three indices, PI, WI, and AI, which now just reflect the destination. For the population density (g_k) and workplace density (w_k), these attributes take on the value of the corresponding zone k. On the other hand, the number of alightings in a given zone k for line x (A_{xk}) is defined as the total number of alightings from line x, for all stops in zone k, as shown in Equation

$$A_{xk} = \sum_{i \in M_{xk}} a_{i;x} \tag{24}$$

Where $M_{xk} \subseteq R_{s;x}$ is the set of downstream stops which are served by line x that are in zone $k \in Z_{s;x}$ and $a_{i;x}$ is the number of passengers alighting from line x at stop i.

In addition to the modified proxies of attractiveness, due to the disaggregation of the choice set from lines to downstream stops, the angle cost has been modified. The departing angle $\alpha_{xk;s}$ now only considers the angle to the destination zone k, instead of the average angle to all downstream stops. Thus, at stop s, for a downstream destination zone k along a transfer-receiving line x, the departing angle is as defined in Equation 25.

$$\alpha_{xk;s} = \frac{1}{|M_{xk}|} \sum_{i \in M_{xk}} \alpha_{si} \tag{25}$$

Where $M_{xk} \subseteq R_{s;x}$ is the set of downstream stops which are served by line x that are in zone $k \in Z_{s;x}$ and α_{si} is the angle formed by a straight line from stop s to stop i and a vector pointing north. Identically to method 1, the true locations of the stops were used to construct these angles.

Using this new departing angle, the angle cost between an arriving line l and destination lk is as defined in Equation 26.

$$\mathrm{AC}_{s;lxk} = 0.5 - \frac{\cos(\varphi_{l;s} - \alpha_{xk;s})}{2} \tag{26}$$

Where $\varphi_{l;s}$ is the arrival angle of line l at stop s, as defined in Equation 15.

The last four attributes are newly defined ones specifically for this model. First is the travel time from stop s to destination k using line x ($\delta_{x;sk}$), which is defined as the average travel to all stops served by line x in a destination zone $k \in Z_{s;x}$. This attribute is a central piece of almost all transport models and is a pivotal factor in estimating human behavior, both in general OD-matrix estimation, where the travel time is often the cost used in Gravity models (Gkiotsalitis, 2022; Ortúzar and Willumsen, 2011), and in route choice modeling (Ortúzar and Willumsen, 2011; Prato, 2009; Raveau et al., 2011).

The next explanatory variable is the ratio between the number of upstream stops on the transfer-receiving line l which have a direct connection to destination zone k and the total number of upstream stops. This attribute is defined as the 'Direct Connection Ratio' (DCR) and aims to capture the tendency of public transportation users to prefer non-transfer connections to their destination over a connection with a transfer, even at the cost of increased travel time (Cui, 2006; Raveau et al., 2011).

The final two attributes are a pair of dummy attributes and are defined in Equation 27 and 28. The first one indicates whether or not the destination zone k was part of the upstream stops of transfer-producing line l, partially indicating a return to the origin, though not completely, as some zones may have been passed before the boarding of a passenger. The second one indicates whether the destination is in the set of downstream stops that transfer-producing line l serves, signifying that no transfer would be necessary to reach this destination and a passenger could just stay on the vehicle he is currently on.

$$D^{\text{passed}}_{s;lk} = \begin{cases} 1, & \text{if } k \in U_{s;l} \\ 0, & \text{otherwise} \end{cases} \tag{27}$$

$$D^{\text{downstream}}_{s;lk} = \begin{cases} 1, & \text{if } k \in Z_{s;l} \\ 0, & \text{otherwise} \end{cases} \tag{28}$$

Transfer-destination sample

In addition to the utility function and attributes, transfer-destination data is to be utilized. This (sample)

data from the survey shows, at a stop s, for a given line-direction l_d, how many passengers transferred to line x and went to destination k, as shown in the example Table 4.

Table 4: Example of transfer-destination data

line	1_0		1_1		2_0		2_1		Π
zone	a	b	c	d	e	f	g	h	
1_0	0	0	0	0	2	1	2	0	8
1_1	0	0	0	0	1	0	1	1	10
2_0	3	1	2	0	0	0	0	0	14
2_1	1	0	1	1	0	0	0	0	10

The probability of transferring to a given destination k on line q at stop s can be computed using Equation 29, for which $\boldsymbol{\theta}$ must be estimated using MLE and the transfer-destination data.

$$P_{lqk} = \frac{e^{V_{lqk}}}{\sum_{j \in Q} \sum_{r \in R_{q;s}} e^{V_{ljr}}} \tag{29}$$

Model verification and transferability

For the model verification, the same four indicators that were chosen for method 1 will also be utilized ($\bar{\rho}^2$, MTE, RMSE, and PCC). But, in addition to comparing the observed and expected transfer-destination flows, the resulting flows will be aggregated to transfer flows from line-direction to line-direction, allowing a more direct comparison to estimates of method 1.

Following the verification, the transferability of the model will also be determined in the same way as for method 1.

3.4.2 Network-level OD matrix

The next phase of method 2, following the computation of the transfer-destination proportions, is building the network OD matrix by utilizing the APC data and the route-level matrices. For this, the first step consists of computing the initial transfer-destination counts, i.e., the initial number of alighting passengers going to transfer-destination jk for line $l \in L_s$. This can be done using Equation 30, where $P_{s;ljk}$ is the previously computed probability and $a_{s;l}$ is the number of passengers alighting from line l at stop s. Once again, the initial number of passengers exiting the system can be computed by setting $jk = \Pi$.

$$b_{s;ljk} = P_{s;ljk} a_{s;l} \tag{30}$$

Using the initial transfer-destination flow $b_{s;ljk}$, it is aggregated to the estimated transfer-boarding totals per line j using Equation 31 and compared to the actual boarding counts of line j at stop s ($\omega_{s;j}$).

$$B_{s;j} = \sum_{i \in I_{s;j}} \sum_{k \in Z_{s;j}} b_{s;ijk} \tag{31}$$

Where $I_{s;x}$ is the set of lines capable of transferring to line j at stop s (the transfer-receiving set) and $Z_{s;j}$ is the set of zones line j stops at downstream of stop s.

Analog to method 1, the same two cases are present during the comparison ((1) $B_{s;j} \leq \phi_1 \omega_{s;j}$ and (2) $B_{s;j} > \phi_1 \omega_{s;j}$), for which the appropriate rebalancing will be applied.

Following this first check, the estimated total transfer-destination flow from stop s to destination k on line j ($D_{j;sk}$) can be compared to the alighting count A_{jk} by aggregating the transfer-destination flow using Equation 32.

$$D_{j;sk} = \sum_{i \in I_{s;j}} b_{s;ljk} \tag{32}$$

Where $I_{s;x}$ is the set of lines capable of transferring to line j at stop s.

Analog to the first rebalancing, two cases must be considered: (1) $D_{j;sk} \leq \phi_2 A_{jk}$ and (2) $D_{j;sk} > \phi_2 A_{jk}$ ($0 < \phi_2 \leq 1$ is a selected hyperparamter which defines what percentage of the alighting count the transfer-destination flow can maximally make up). In the case of (1), no more adjustments are required. In the case of (2), the transfer-destination proportions for the transfer-producing lines are uniformly reduced and the exiting proportion is increased, such that the computed OD counts are within a margin of error ϵ.

Second-to-last, per line j the transfer-destination flow can be distributed among the stops in destination zone k using proportional distribution, shown in Equation 33.

$$f_{s;ljz} = \frac{a_{z;j}}{\sum_{x \in M_{jk}} a_{x;j}} b_{s;ljk} \tag{33}$$

Where $f_{s;ljz}$ is the passenger flow transferring at stop s from line l to line j with their destination as stop z, $a_{z;j}$ is the alighting count of line j at stop z, and M_{jk} is the set of stops served by line j in zone k.

Finally, the route-level marginals can be adjusted, after which MIPF or proportional distribution can be used to determine the first leg of the trip, followed by leg-linking, leading to the network-level OD matrix.

3.5 Coefficient importance

To gain further insight into the method application and gain an understanding of how the explanatory variables impact the utility and models, the coefficient importance will be investigated. To this end, the method described by Luchman (2014) to determine general coefficient dominance will be utilized. This method uses the differences in the $\bar{\rho}^2$ value to investigate the contribution of a certain explanatory variable to the model fit.

To utilize this method, first, MNL model regressions using all explanatory variable combinations ($2^N - 1$ where N is the number of explanatory variables) must be performed, for which the final $\bar{\rho}^2$ value will be recorded. In the case of a model using three explanatory variables, this results in seven explanatory variable combinations ($\{X_1\}$, $\{X_2\}$, $\{X_3\}$, $\{X_1,X_2\}$, $\{X_1,X_3\}$, $\{X_2,X_3\}$, $\{X_1,X_2,X_3\}$), and thus seven results.

Using the regression results, the coefficient dominance of explanatory variable k can be computed using Equation 34. Comparing this value between all explanatory variables allows one to build a ranking of the most impactful ones.

$$C_{X_k} = \frac{1}{N} \sum_{i=1}^{N} C_{X_k}^i \tag{34}$$

Where N is the number of explanatory variables and $C^i_{X_k}$ is the average marginal $\bar{\rho}^2$ contribution of variable X_k at order i. To compute the average marginal contribution, Equation 35 is utilized, where the upper limit and denominator are the binomial coefficient.

$$C^i_{X_k} = \sum_{j=1}^{\binom{N-1}{i-1}} \frac{\bar{\rho}^2_{X_k S_j} - \bar{\rho}^2_{S_j}}{\binom{N-1}{i-1}} \tag{35}$$

Where i is the order of the regression or the number of explanatory variables utilized for a given regression and X_k is the variable of interest. Furthermore, S_j is the jth set of length $i-1$ of explanatory variable combinations where X_k is excluded. For example, in the case of $k = 1$ and $i = 3$, $S_1 = \{X_2, X_3\}$.

3.6 Comparison of methods

To investigate which method is most suitable, they will be compared qualitatively and, to a certain extent, quantitatively. To this end, first, the differences between the two methods are highlighted, after which the points they will be compared on and the corresponding methodology is described.

3.6.1 Differences between methods

To summarize, the two proposed alternatives differ in the utility function, where the indices acting as proxies for the attractiveness of the transfer-receiving line are disaggregated and assigned per line-zone pair. Furthermore, additional explanatory variables have been introduced which are destination-zone specific. From this, it follows that the choice set at stop s for transfer-producing line i is replaced by the zones visited by the feasible transfer lines of set $T_{s;i}$ after stop s.

3.6.2 Method comparison

Due to a lack of ground truth, the models will primarily be compared at a qualitative level, for which two general criteria have been established. The first is the scalability of the models: Public transportation networks can be vast and complex, with many destinations, as well as lines and modes to serve them. This entails that, a model that is harder to scale can not be utilized as frequently and easily.

The second criterion is "data necessity", which aims to capture how much data is required to utilize the models reliably. This considers both the quality of the survey data, as well as of the size of the sample necessary to determine the models reliably.

In addition to this qualitative analysis, the methods will also be quantitatively compared in terms of their prediction capability using the aggregated transfer-destination flows of method 2 and comparing these to the results of method 1.

Finally, the coefficient and framework transferability of both methods will be compared.

4 Case Study: Lancy-Bachet

As stated previously, the model calibration will be done utilizing a single stop of the Geneva public transportation network. For this purpose, a transfer station with a high-quality and representative survey response must be chosen. In this case, the chosen stop is 'Lancy-Bachet', which is a major transfer station of the network, where all three available modes (bus, tram, and train) meet. Furthermore, it is not as complex as one of the larger stations, such as the main train station of Geneva, "Gare Cornavin", which may make it more generic and representative of the other stations of the network.

4.1 Overview

The chosen stop is located in the southeastern part of Geneva, as shown in Figure 7, where there are a total of 13 lines passing, thus 26 directions. Of these 13 lines, four are train lines (LEX), two are tram lines (lines 12 and 18), and the rest are bus lines. From the survey for this stop (transfer sample), it can be observed that, during the morning peak hour (6:00 - 9:59), about 49% of the arriving public transportation users transfer at this station, where the two lines with the greatest number of arrivals are line D_1 (line D-direction 1: from Saint-Julien-en-Genevois, FR), and line 62_1 (from Collonges-sous-Salève, FR), with 25% and 10% of the arrivals, respectively. Of all transfers, most go to line 12_0 (line 12-direction 0: towards the Geneva city center), attracting 33% of the total transfers, with the most attractive transfer destinations on that line being "Carouge, Ancienne" and "Carouge, Marché", where 26% of the transferring passengers go. These

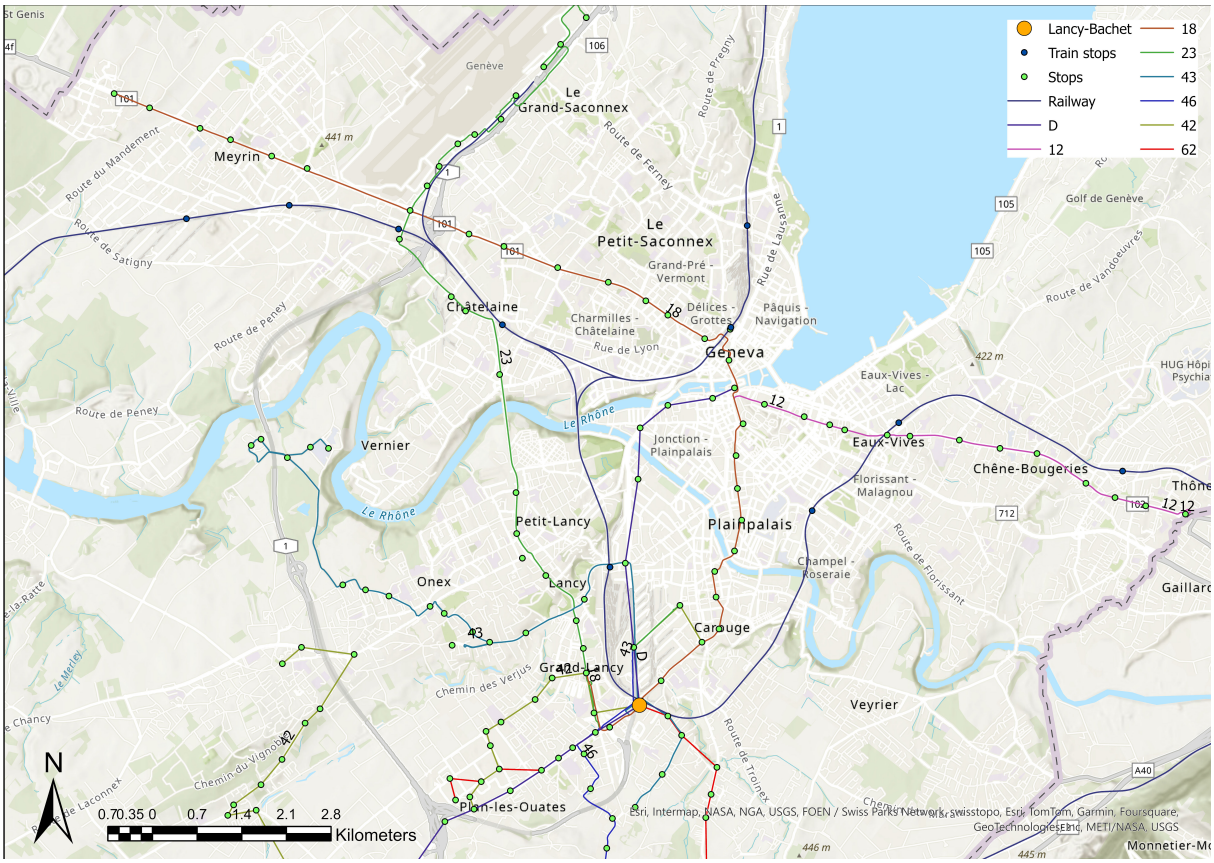


Figure 7: Lines passing at Lancy-Bachet

two stops are the third and fourth stops after Lancy-Bachet, respectively. Additionally, although line 12_0 is the most popular transfer line overall, the most popular transfer destinations are line D_0, stops "Plans-les-Ouates, Vélodrome", "Plans-les-Ouates, ZIPLO", and "Plans-les-Ouates, Galaises", which attract 10% of all transferring passengers.

Finally, it must be noted that the density of the sample transfer matrix for the morning peak is at 28% for line-direction pairs, and at 6% for line-direction to line-direction-zone pairs.

During the evening peak (16:00 - 18:59), according to the survey, about 45% of public transportation users transfer, with 24% of all arrivals being with line 12_1 (from the Geneva city center) and 11% with line D_1 (from Collonges-sous-Salève). For transferring passengers, most go to line D_0 and line 12_0, with 35% and 12% of all transfers being to those lines.

Additionally, something that is of note, is that the representation of the train lines in the survey is quite muted. At Lancy-Bachet, a total of 574 passengers per hour arrive during the morning peak using the train, which, according to the APC data, makes up 55% of the arriving morning traffic at the station. On the other hand, only 37% of the survey arrivals during that time frame arrive using a train. This is further reflected in the departure counts of the survey data, where 28% are departing using the train, whilst the APC data indicates that 55% of the departing passengers board a train. This low representativeness is also present during the evening peak, where 20% of the arrivals and departures recorded in the survey are with the train, whereas the APC data indicates that 67% of arrivals and 64% of departures are with the train.

4.2 Attribute overview

To gain a better understanding of the data, some of the attribute values and their distribution with respect to the number of observed transfers at Lancy-Bachet are investigated.

Figure 8 shows the relationship between the average transfer time from a transfer-producing line to a transfer-receiving line and the number of observed transfers. As shown, for Lancy-Bachet, the relationship between transfer time and number of transfers is inversely proportional for transfers to trams and buses, with the decline being more pronounced for trams. Though, in this case, the maximum transfer time observed is strongly reduced, as the trams have a high frequency (One tram every 5 minutes). On the other hand, this relationship is much less pronounced when the mode of the transfer-receiving line is a train. In this case, there is a larger number of outliers around the 15-minute mark, which is also the average transfer time for all trains, with a slight reduction in the number of transfers with the next bins. Furthermore, something that must be noted is that there are no transfers for connections with lower transfer times (between 4 and 8 minutes).

For the angle cost (AC), shown in Figure 9, it can be seen that the expected relationship is present, where an increasing angle cost, indicating a return towards the origin, leads to a reduced number of transfers. On the other hand, something that must be noted is that there is a slight hump towards the end of the angle costs range, where the destination lines go almost directly back towards the origin. This hump is more prevalent for trains and trams, as there are more than just outliers, indicated by the larger interquartile range. Furthermore, compared to buses, the trams and trains also have more transfers in the middle of the range.

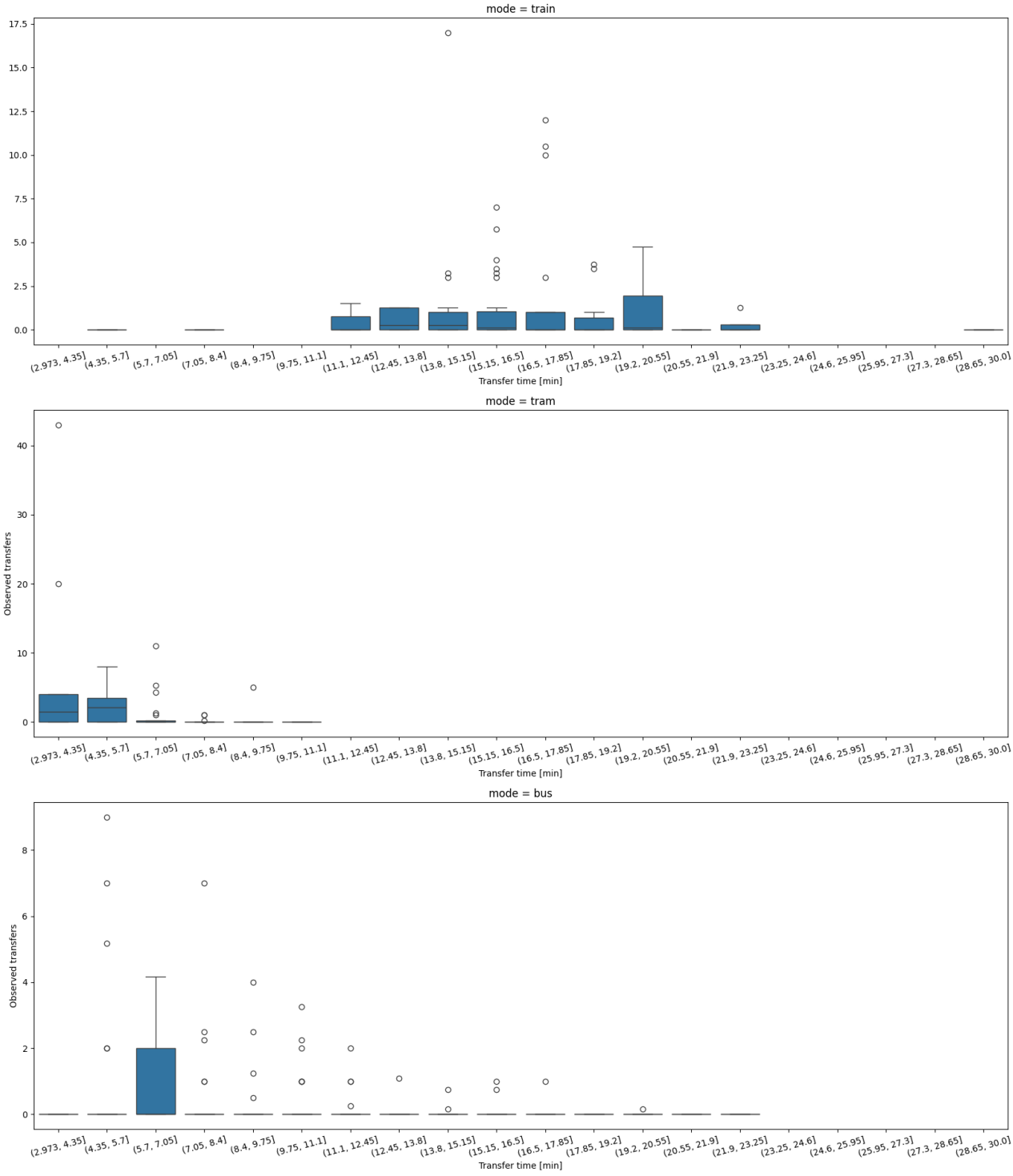


Figure 8: Transfer time (x-axis) VS number of observed transfers (y-axis) at Lancy-Bachet (morning peak)

4.3 Scaled transfer flows

As an extension of the transfer sample and to allow for as close a comparison to the real flows as possible, the transfer sample will be scaled up using the weights TPG generated when they estimated the route-level OD matrices using IPF. These weights are available per line-direction, for each stop-to-stop connection that the line serves and will be applied to the individuals in the transfer sample, thus scaling it. After this initial

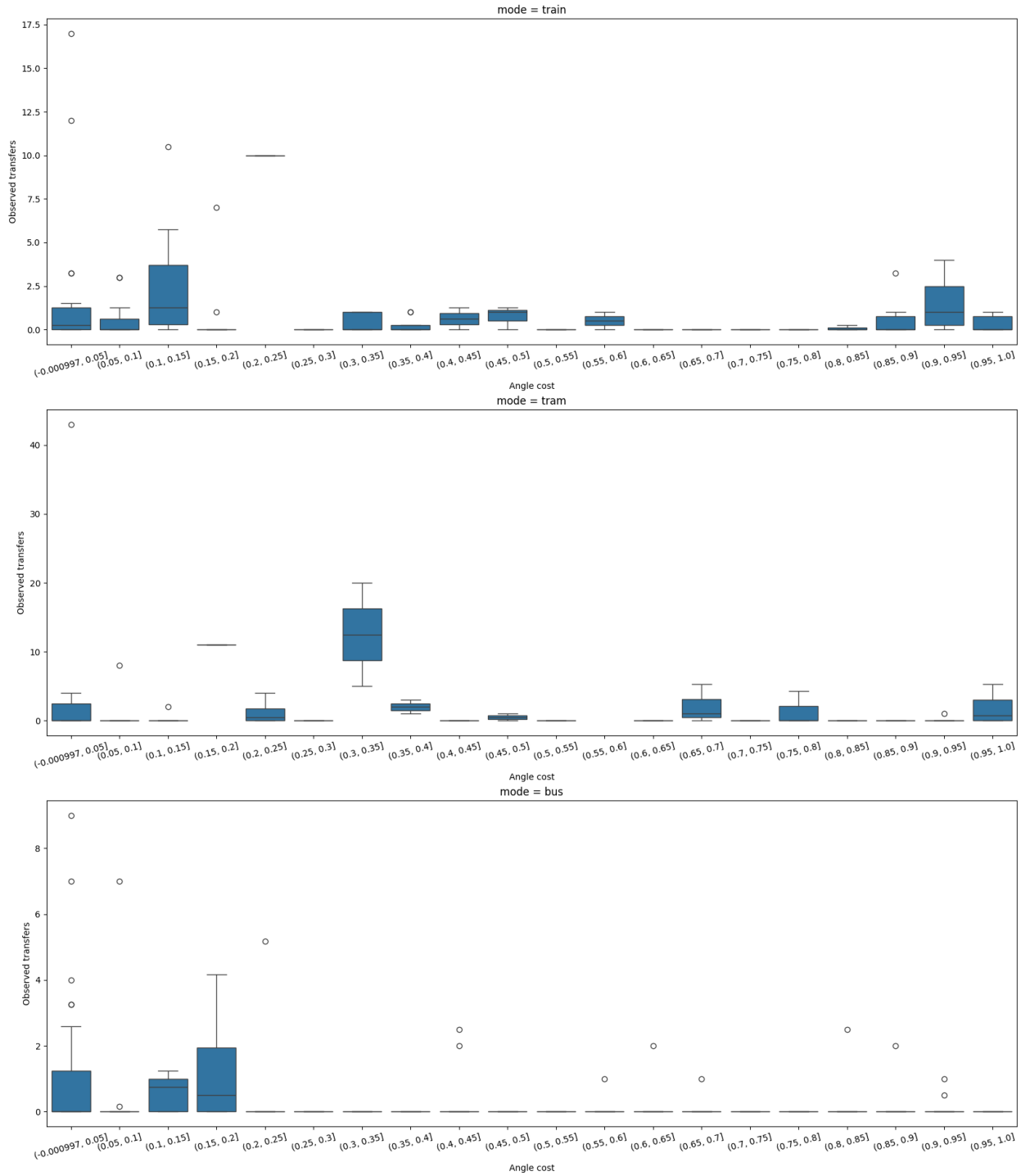


Figure 9: Angle (x-axis) VS number of observed transfers (y-axis) at Lancy-Bachet (morning peak)

scaling, the flows are rebalanced using the alighting counts. Due to the nature of this scaling, it must be noted that zero flows will remain as such, which means that the captured picture may not be complete.

These scaled flows will first be used as additional validation of the models by estimating the transfer flow using the calibrated coefficients and assessing the performance indicators. And secondly, they will also be used to calibrate the models directly. The results will then be assessed and compared to the models calibrated directly on the transfer sample.

5 Estimation Results

5.1 Method 1

Table 5 shows the coefficient value, standard error, and p-value for the most performant, significant, and interpretable models for the morning peak (6:00 - 9:59). Model 1 uses the Alighting Index ($AI_{s;q}$) as the attractiveness proxy variable and model 2 uses the Workplace Density Index ($WI_{s;q}$). In both cases, the Box-Tuckey transformation was applied to the indices, with $\tau = 0$ for model 1 and $\tau = 1$ for model 2. Furthermore, β was set to 1 for model 1 and 1.1 for model 2, as it led to the model with the best fit (according to the chosen performance indicators), though the differences were often minor (see Section 5.1.4 for more details).

Table 5: Method 1 estimation results - Lancy-Bachet (morning peak)

		Model 1			Model 2		
Attribute	Mode	Coefficient	Std-Error	P-value	Coefficient	Std-Error	P-value
Intercept	Π	2.927	0.561	0	1.561	0.454	0.001
Transfer time	train	-0.63	1.025	0.539	-2.203	0.946	0.02
	tram	-14.21	2.396	0	-16.648	2.336	0
	bus	-5.016	1.457	0.001	-8.823	1.187	0
Alighting index	train	1.149	0.261	0	-	-	-
	tram	3.565	0.420	0	-	-	-
	bus	2.4	0.509	0	-	-	-
Workplace index	train	-	-	-	15.616	7.652	0.041
	tram	-	-	-	21.557	2.894	0
	bus	-	-	-	17.745	7.304	0.015
Angle cost	train	-1.91	0.438	0	-2.048	0.436	0
	tram	-2.083	0.499	0	-2.173	0.514	0
	bus	-2.753	0.463	0	-3.112	0.484	0
From train	tram	0.809	0.451	0.073	0.995	0.462	0.031
	bus	0.414	0.339	0.222	0.826	0.318	0.009
From tram	train	0.553	0.341	0.105	0.599	0.342	0.08
	bus	0.651	0.399	0.103	0.999	0.392	0.011

As shown from the coefficient of the intercept for Π (network exit), there is a base increase in the utility of that option, which is to be expected, as on average, about 50% of the arriving passengers exit the network. This effect is reflected in both models.

When observing the model 1 coefficients for the transfer time, it can be seen that a large difference is present between the coefficient value for trams compared to the coefficient value for trains (-14.21 VS -0.63), but is shown to be statistically insignificant. As highlighted before (see Section 4.2), this behavior is as expected, due to the generally lower frequency of trains at the station, thus leading to a similar transfer time across the board, usually around 15 minutes. Furthermore, this is also reflected in the coefficients for the second model, albeit in a different way. The transfer time coefficients for trams and trains show greater magnitudes (absolute values) of -16.648 and -2.203, respectively, with the difference in magnitude still apparent. Although, in this case, the latter explanatory variable is statistically significant at the $\alpha = 0.05$ level. This difference is likely a reflection of the contribution made by the chosen indices (alighting VS workplace density), as the latter is a less direct representation of attractiveness, with trains generally stopping in

workplace-dense neighborhoods, thus bolstering their index value, in spite of not necessarily being truly attractive.

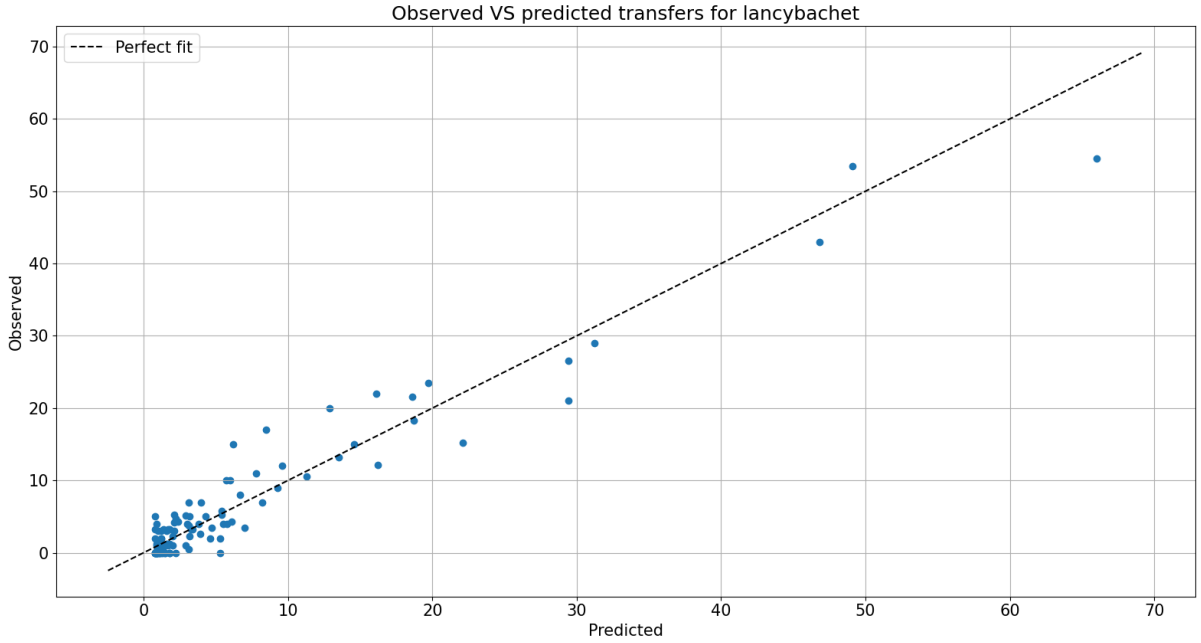
The coefficient value for buses, on the other hand, is consistent in both models, being statistically significant in both cases, but with the magnitude of the model 2 coefficient value being greater. This difference reinforces that the alighting and workplace indices both capture attractiveness differently, with AI being a more direct representative. This disparity is possibly also due to the peculiarity of tram lines 12_0 and 18_0. Though they both share the same first stops, line 12_0 is much more attractive, which is reflected in the alighting counts but not the employment density of downstream zones. This leads to line 18_0 having a greater Workplace Index, thus having the transfer time compensate for the increased utility.

For both models, the angle cost coefficients behave as expected and are statistically significant at the $\alpha = 0.05$ level. Furthermore, the coefficients of model 1 are statistically the same as the coefficients of model 2, indicating that the impact of the angle cost is of considerable importance to the models, as well as consistent, irrespective of the used attractiveness proxy. Additionally, it must be noted that the coefficients follow the expected pattern, as described in section 4.2. First, the relationship between the number of transfers to a bus and the angle cost follows the most predictable pattern. This is reflected in the bus coefficient, where it has the highest magnitude out of the three destination modes. Secondly, for the other two modes, both coefficients are slightly lower, which are reflections of some transfers still occurring at a higher angle cost, as shown in Section 4.2.

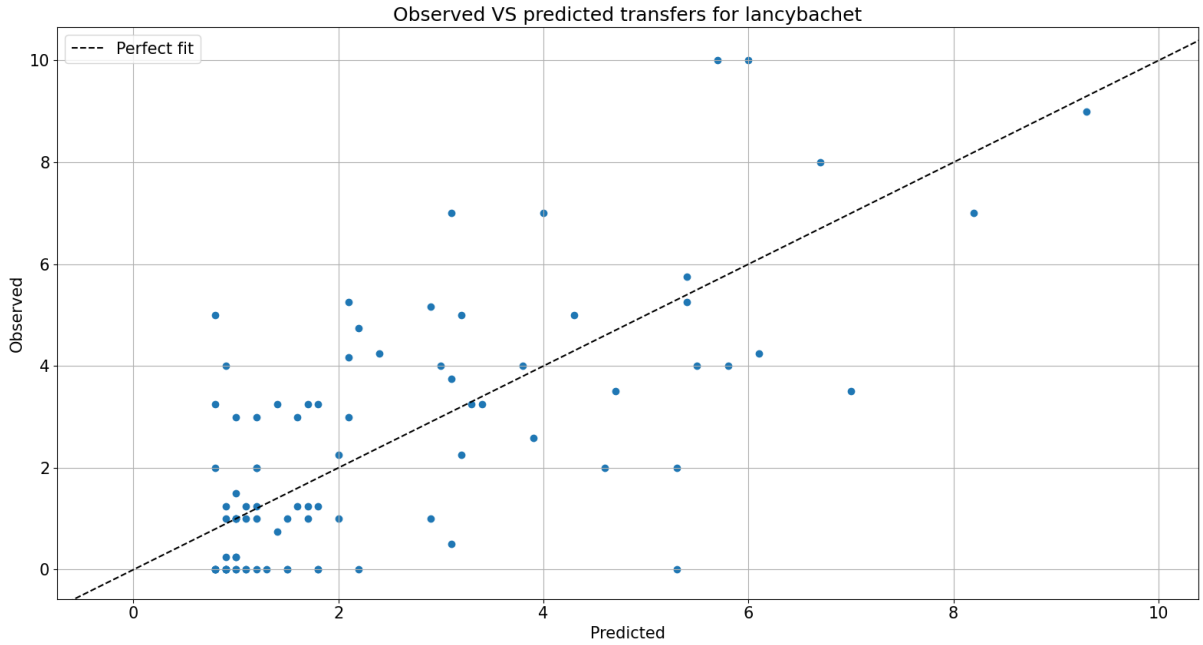
The two final attributes shown are the two dummy variables. As depicted, when arriving on a tram or a train, one is more likely to transfer to a line of another mode, instead of to a line of the same mode. This is due to the build-up of the Lancy-Bachet station and the LEX train network. First, as mentioned in Section 4.1, two tram lines are arriving at Lancy-Bachet. The trams going towards the city center (direction 0) both depart in the same direction and share the same first stops, before splitting up. This strong overlap would mean that there is no reason to transfer to the other tram, as one would instead have boarded it prior. Furthermore, due to the high frequency of both (arriving about once every 5 minutes), there is virtually no difference between boarding one or the other, if one's destination is within the first half of the serviced stops, which holds 80% of the route-OD flows going from Lancy-Bachet for line 12_0 or 85% for line 18_0.

Finally, as depicted in Table 5, both dummy variables are statistically insignificant for model 1, but they were kept, as their absence negatively affected model fit by reducing $\bar{\rho}^2$ to 0.431 and the PCC indicator to 0.701 and increasing the MTE and RMSE slightly.

Next to the coefficients, the computed performance indicators for the models are depicted in Table 6. Furthermore, to visualize the model performance, Figure 10 shows the observed and predicted transfer flow per transfer-producing and transfer-receiving line combination (each point represents a different combination). As depicted, both morning peak models are shown to have an excellent fit ($\bar{\rho}^2$), with the value being greater than 0.4 in both cases and the model using AI showing a better fit (difference of 0.01). Coupled with the interpretable coefficients, this indicates that the model is generally capable of describing transfer flows using the chosen attributes. This is supported by the MTE, which sits at around 0.3 for both models, as well as the RMSE, where the former indicates that the average over-/under-estimation of transfer flows is at about 30%. For larger flows, this RMSE (over-/under-estimation of 1.4 passengers-flow) is acceptable,



(a) Predicted VS observed transfer flows



(b) Predicted VS observed transfer flows (< 10)

Figure 10: Comparison of predicted (x-axis) and observed (y-axis) transfer flows at Lancy-Bachet using the transfer sample and model 1 (points where the transfer flow is less than 0.75 are not shown)

but for lower flows (less than five transfers), which constitute the majority, it is somewhat large, as the expected flows can easily be doubled or tripled (from 1), as shown in Figure 10b. Although, considering the PCC for those flows, the relationship shows a strong correlation, indicating that the model is capable of representing the relationship between the flows with a high degree of accuracy, though with the over and underestimation of some flows, which will be investigated in further detail in Section 5.1.1.

Table 6: Method 1 performance indicators - Lancy-Bachet (morning peak)

	Model 1 (AI)	Model 2 (WI)
τ	0	1
β	1	1.1
$\bar{\rho}^2$	0.441	0.433
MTE	0.293	0.31
RMSE	1.425	1.441
PCC (<5)	0.751	0.708

Finally using the methodology by Luchman (2014) and described in section 3.5, the coefficient importance for the best performing model (morning peak - model 1) can be determined. Due to the number of required regressions increasing exponentially with the number of explanatory variables ($2^{14} - 1$ regressions), the respective dominance will not consider the application per mode, but strictly per explanatory variable. The corresponding results are shown in Table 7. As depicted, the most important variables are the "Transfer time" and the "Intercept". For the latter, this follows from the number of non-transfers being at about 50%, meaning that simply increasing the utility of this case is very impactful.

Table 7: Coefficient dominance Method 1 - model 1

Intercept	Transfer time	Alighting index	Angle cost	From train	From tram
0.115	0.155	0.04	0.074	0.049	0.007

5.1.1 Transfer flow analysis

One of the first issues to address, which is shared by both models, is the underestimation of the transfer flow from line 62_1 to line 12_0, where it is estimated that only 12 or 13 passengers, instead of 20 transfer. Comparatively, although the sample shows zero transfer flow from line 62_1 to line 18_0, it predicts that some passengers will transfer to that line. This means that the passengers from 12_0 are allocated to the wrong transfer destination, although they are predicted to go generally in the correct direction. This disparity in the transfer flow to lines 12_0 and line 18_0, though they share the same frequency in the morning peak and their first stops overlap, indicates a quirk of the network, which is reflected in the alighting counts, but not in the workplace density values, as these are line-agnostic. For model 2, this means that, although lines 18_0 and 12_0 have a similar Workplace Index, inducing a similar utility, users of the network may perceive them differently. A possible reason line 12_0 is more attractive than line 18_0 during the morning peak is that it starts at Lancy-Bachet, thus being already present when the passengers transfer and inducing a sense of reliability. Additionally, although the disparity in the Workplace Density Index explains the miscategorization for model 2, it does not for model 1. In this case, another factor comes into play, the angle cost. Due to the departing trajectories of the two tram lines, these have vastly different angles, and the corresponding angle cost is very different as well. Line 12_0 has a conflicting angle to the arrival angle of line 62_1, thus leading to a high angle cost (0.33), which reduces the utility of this line for a transferring passenger. On the other hand, the departing angle of line 18_0 is almost straight on from line 62_1's arrival angle, leading to a lower angle cost (0.08) and increasing the utility.

The second issue to look at in more detail is the transfers from line D_0, which is the largest contributor of arrivals in the sample for Lancy-Bachet, of which 70% do transfer. The model tends to underestimate this large number of transfers, as, for other lines, these are generally around 50% instead. In addition to the average transfer rate being smaller, this overestimation also indicates a mismatch between the actual and modeled utility of transferring.

Another issue to mention is the potential that the angle cost (AC) has to strongly deter transfers from passengers who boarded shortly before the transfer station. This is particularly prevalent for lines such as 12_1, which have many stops before Lancy-Bachet and go a non-linear parkour. In this case, there is an underestimation of the flow from line 12_1 to line LEX $L2_1$, where four transfers are observed, with the transferring passengers boarding at "Carouge, Rondeau" and "Carouge. Ancienne", two and three stops before Lancy-Bachet. But, due to the long route of line 12_1, the arrival angle is strongly skewed in favor of the stops before these, inducing a high angle cost and reducing the utility of this alternative, resulting in low estimated transfer flows.

This effect is also reflected for transfers from line LEX $L3_1$ to line 12_0, which fall prey to the same issue, although this connection is the fastest to go to their desired destination.

5.1.2 Application to scaled flows

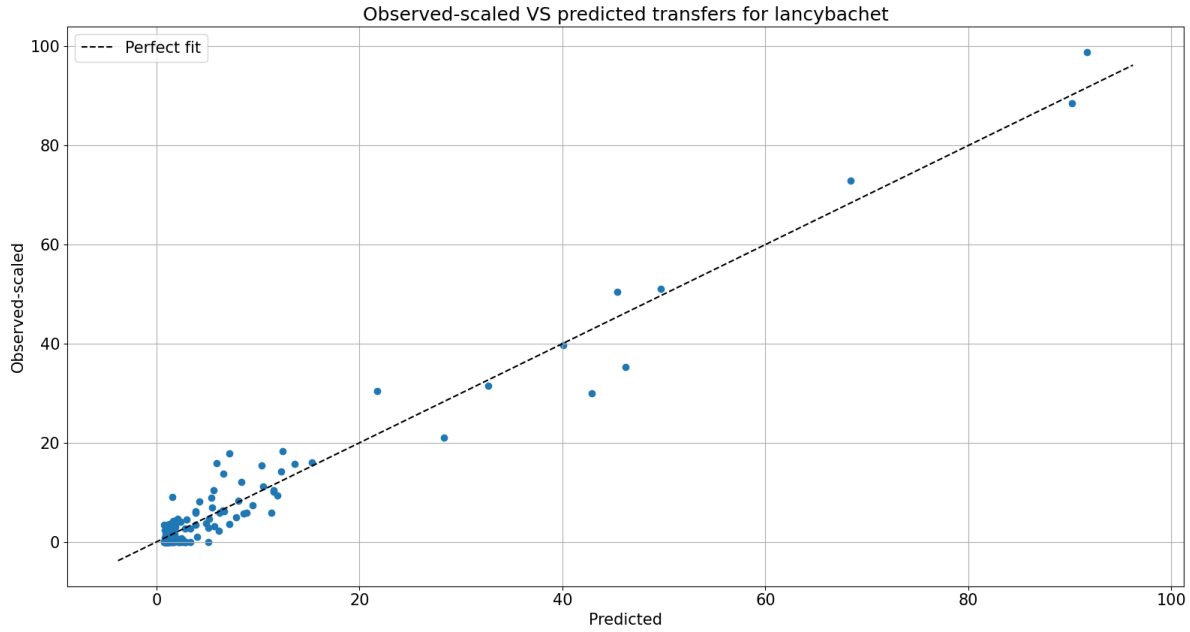
In addition to applying the model to the transfer sample to gain an initial insight into how it predicted the sample flows, the model was also applied to the scaled hourly arrival and transfer flow at Lancy-Bachet (see Section 4.3 for more information).

The performance indicators for the model application are shown in Table 8, and Figure 11 depicts the predicted transfer/network exit flow using model 1, compared to the observed-scaled transfer/network exit flow. For model 1, which uses the Alighting Index, the performance is very similar to the transfer sample performance, with the MTE being lower than for the transfer sample by 0.03 or improving by 10% relative to the transfer sample value. Furthermore, though the RMSE is slightly worse, by 0.4 or about 24%, it is still within a reasonable margin, especially considering the higher total flow, which increased from 714 to 1018. The final performance indicator, PCC, dropped by 15%, but is still strongly correlated.

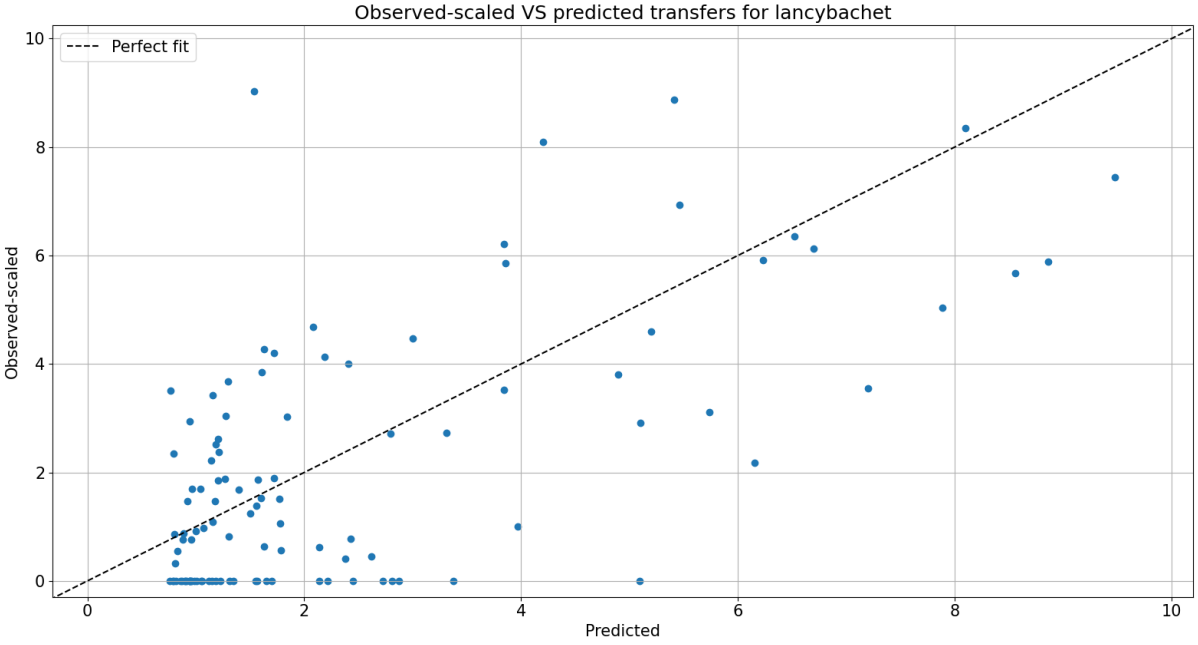
Table 8: Method 1 performance indicators - Lancy-Bachet scaled

	Morning peak	
	Model 1 (AI)	Model 2 (WI)
MTE	0.259	0.287
RMSE	1.768	1.794
PCC (<5)	0.602	0.566

This drop in performance for the PCC indicator is due to an increased estimation of transfer flows where none were observed, as is shown in Figure 11b. As mentioned in Section 4.3 zero flows will remain zero flows due to the nature of this scaling. On the other hand, as the sample is scaled according to the route-OD weights, there will be an increase in incoming flow (arriving flow) for some lines, thus increasing the estimated number of transfers, and resulting in a higher error, or in this case, a lower PCC.



(a) Predicted VS observed transfer flows



(b) Predicted VS observed transfer flows (< 10)

Figure 11: Comparison of predicted (x-axis) and observed-scaled (y-axis) transfer flows at Lancy-Bachet using the transfer sample and model 1 (points where the transfer flow is less than 0.75 are not shown)

Other than the low flows being more spread out along the x-axis, model 1 is capable of predicting the distribution of the high transfer flows well, as depicted in Figure 11a. Noticeable errors, though, are the overestimation of network exits (between observed counts of 20 and 40), for line LEX $L2_1$ and line D_1. This overestimation for D_1 is then reflected in the underestimation of transfers to line LEX $L4_0$.

Furthermore, the model underestimates the number of transfers from line 62_1 to 12_0, which is a behavior highlighted in Section 5.1.1. These are instead assigned to line 18_0, to which the transfer flow from line 62_1 is still zero.

For model 2, the same pattern as for model 1 is found, where the MTE is slightly better, and the RMSE and PCC are slightly worse. However the improvements are less significant, and the drop in the PCC is more significant. Furthermore, once again, model 1 performs slightly worse than model 2 across the board, with only the RMSE being almost equal.

With regards to the transfer flow, a noticeable pattern in Figure 11a is the three dots which are vertically above one another between an observed-scaled flow of 15 and 20. The two most prominent points are transfers from line D_1 to line LEX $L4_0$, as well as network exits from line D_0, which both have been estimated at around 6.8. In the case of the former, most non-transfers are redirected to either line 12_0 or the network exit, leading to a larger discrepancy. The reason the utility of LEX $L4_0$ is perceived as lesser is due to two factors, first, it has a relatively high average transfer time of around 15 minutes, though fairly standard for trains, which is coupled with a comparatively very low alighting index. These two factors would normally result in an even lower utility, but the angle cost is almost 0, thus improving the final utility and still making the destination line a competitive alternative.

For the latter, the low estimation of network exits for line D_0 is due to the overestimation of transfers to line 12_0 and zero-flow errors. The high utility of line 12_0, from the perspective of line D_0, is due to its alighting index and the angle cost between the two, thus attracting a large number of transfers. Furthermore, due to the generally low usage of line D_0, the sample is very sparse with only three of the alternative choices being non-zero. This leads to the model assigning flows to destinations where no observations were made, contributing to the high error.

5.1.3 Scaled flow regression

In addition to applying the regressed model coefficients to the scaled arrivals, estimating the transfer flows, and comparing them to the observed-scaled transfer flow as done in section 5.1.2, the model coefficients are also estimated using the scaled flows. This led to the results depicted in Table 9, which shows the estimation results using the Alighting Index and the Workplace Density Index models, models 1 and 2, respectively.

Though the coefficients of the non- and scaled regression are statistically the same at the $\alpha = 0.05$ level, there are some noticeable differences. These include an increased transfer time coefficient value, differences in the two indices, and the "From train" dummy variable.

With regards to the strong increase in the coefficient value of the "From train" dummy variable, though this might indicate an increase in the share of transfers from a LEX-line to a tram or bus in the scaled sample, that is not fully the case. The ratio for the number of transfers to trams dropped from 12 to 10% and increased by 1% for buses. Instead, this increase is in part a reflection of the magnitude increases for the negative coefficient values, the transfer time and the angle cost. As the choice sets and corresponding attribute values are the same, a necessary balance needs to be established, which is achieved by this variable and the Alighting Index.

Table 9: Method 1 scaled estimation results - Lancy-Bachet (morning peak)

		Model 1			Model 2		
Attribute	Mode	Coefficient	Std-Error	P-value	Coefficient	Std-Error	P-value
Intercept	Π	3.027	0.511	0	1.305	0.423	0.002
Transfer time	train	-0.715	0.962	0.457	-3.11	0.965	0.001
	tram	-17.873	2.339	0	-20.965	2.285	0
	bus	-5.576	1.280	0	-9.815	1.046	0
Alighting index	train	1.257	0.279	0	-	-	-
	tram	4.13	0.405	0	-	-	-
	bus	2.499	0.425	0	-	-	-
Workplace index	train	-	-	-	20.639	8.478	0.015
	tram	-	-	-	24.382	2.804	0
	bus	-	-	-	12.799	6.498	0.049
Angle cost	train	-2.157	0.439	0	-2.375	0.445	0
	tram	-2.938	0.509	0	-3.04	0.529	0
	bus	-3.049	0.401	0	-3.709	0.432	0
From train	tram	1.558	0.467	0.001	1.784	0.482	0
	bus	0.558	0.334	0.095	1.056	0.323	0.001
From tram	train	0.276	0.317	0.384	0.35	0.321	0.276
	bus	0.464	0.406	0.253	0.861	0.406	0.034

Beyond the estimation results, as can be seen in Table 10, the utilization of the scaled flows leads to better-fitting models, according to the $\bar{\rho}^2$, which, for both models 1 and 2, is at 0.532 and 0.525 respectively. On the other hand, the MTE, RMSE, and PCC performance indicators show slightly worse performance than when applying the sample regressed models to the scaled flows (shown in section 5.1.2), except for the PCC indicator of model 2.

Table 10: Method 1 performance indicators - Lancy-Bachet scaled regression

	Morning peak	
	Model 1 (AI)	Model 2 (WI)
τ	0	1
β	1	1.1
$\bar{\rho}^2$	0.532	0.525
MTE	0.27	0.292
RMSE	1.786	1.849
PCC (<5)	0.617	0.649

5.1.4 Sensitivity of τ and β

To describe the utility of an alternative, three indices have been defined, as described in Section 3.3.1, for which the Box-Tuckey transformation may also be applied. These indices and the Box-Tuckey transformation are influenced by two parameters, the decay factor β and the transformation constant τ, for which the sensitivity of the model performance indicators was investigated.

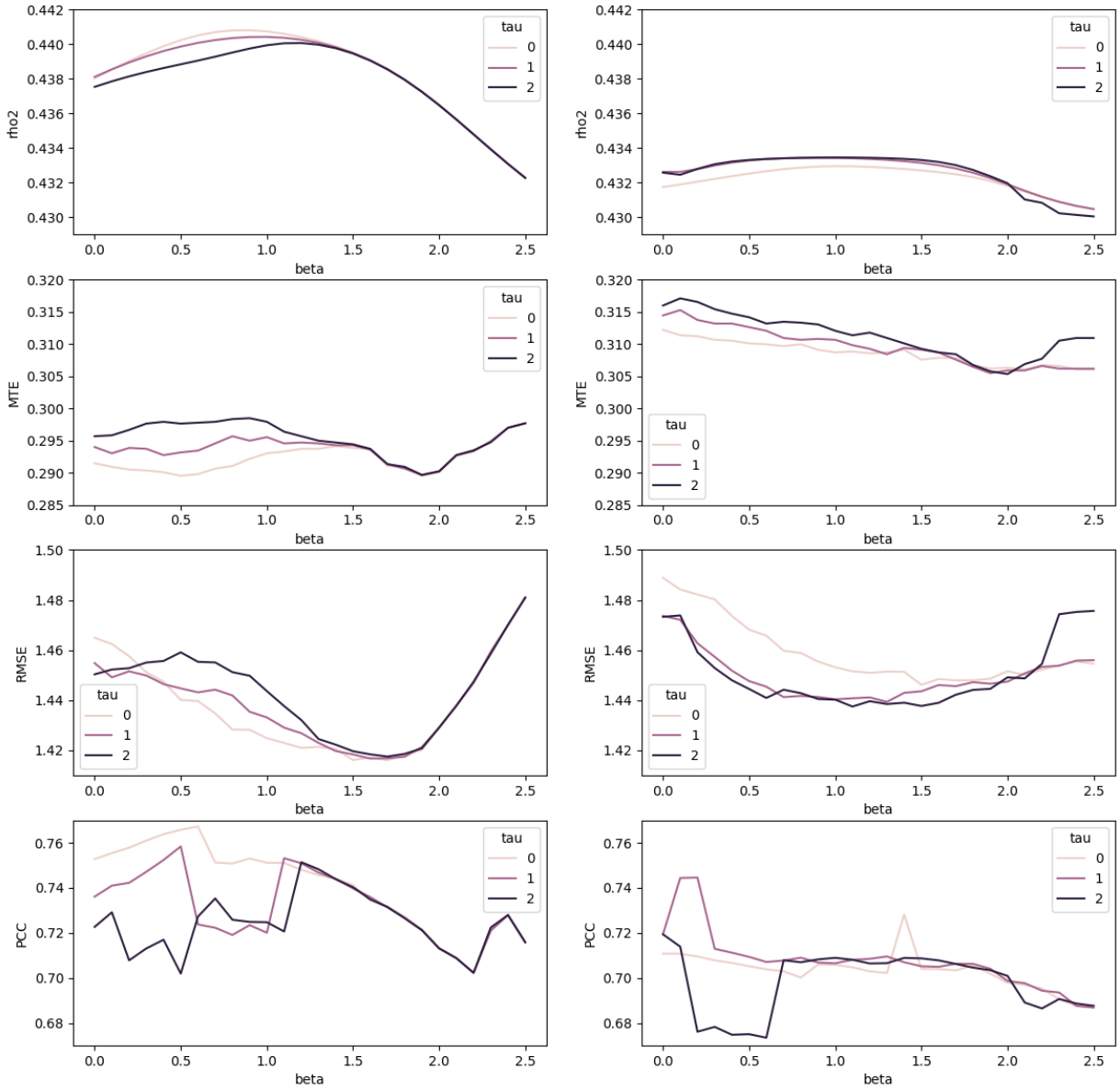


Figure 12: Sensitivity of performance indicators from the morning peak based on used Index, τ, and β (left: model 1, and right: model 2)

Figure 12 shows the four performance indicator values at different levels of β, which ranged between 0.0 and 2.5, and for τ in $\{0,1,2\}$. Overall, for both models, the range of the $\bar{\rho}^2$, the MTE, and the RMSE indicators is restrained, with $\bar{\rho}^2$ ranging from 0.429 to 0.442, MTE from 0.285 to 0.32, and RMSE from 1.4 to 1.5. This indicates that the overall model fit and general transfer flow error are, to a certain extent, insensitive to the changes of both τ and β.

For model 1, it can be seen that irrespective of the value τ takes on, all performance indicators have the same behavior, where their curves merge after β passes a certain threshold. This threshold is about 1.3, 1.6, 1.9, and 1.3, for each performance indicator, respectively. This indicates, that, in spite of a transformation being applied to the AI, it is inconsequential in the face of the values taken on by the variable. The reason for

this is the definition of the AI, where the travel time has a large influence on the value. This travel time value is then exponentiated using β, reducing the variable values to be almost zero for all choice alternatives, which are possibly further wrongly approximated due to floating point errors.

This behavior is not shared by model 2, where the τ-curves stay mostly separate throughout the complete range of β, though they follow very similar patterns. This difference in behavior is due to the magnitude of the alightings and the employment density. The highest workplace density value is $47'519\,\text{EMP}/km^2$, whilst the highest number of disembarkments during the morning peak is $605\text{p}/hr$. With the travel time ranging from 1 minute to a couple of hours, the divisor in the equation is thus often as large or larger than the dividend, leading to a greater reduction of the range of the values and trivializing their differences, reducing the value diversity and thus the response of the model. These differences are also the reason for model 1 being more sensitive to changes in β than model 2, where for the latter, besides the PCC, the performance indicators have a much more constrained range.

Finally, for the PCC, the models show less predictable sensitivity, with all six curves being more erratic than for the other indicators. This is due to it being computed on a subset of the points, where minor alterations in transfer flow can be quite large compared to the magnitude of the observed transfer flow, thus inducing higher sensitivity and sporadicity. Additionally, this sporadicity is also an indication that the models put less weight on these lower flows compared to the larger flows.

5.2 Method 2

Table 11 shows the results of the regression for the most performant, significant, and tractable models in the morning peak (6:00 - 9:59). Model 1 makes use of the line-destination alightings (A_{xk}) and model 2 uses the destination employment density (w_k) as the attractiveness attribute, with both being transformed using Box-Tuckey and $\tau = 0$. As depicted, they both make use of different subsets of the explanatory variables, with the main difference being the dummy variables they utilize. The first model makes use of the "From tram", "Passed", and "Served downstream" dummy variables, whilst the second one makes additional use of the "From train" variable, but drops the "Passed" variable. These changes in the subset are due to the traceability and statistical significance of the variables, which will be discussed in more detail later.

The first aspect to note is the positivity of the intersect. Analogous to the first method (line-to-line model), the intercept is associated with the network exit and represents the base utility compared to any transfer alternative. However, its magnitude is much greater for method 2, at 2.927 VS 4.96 and 1.561 VS 8.82, which reflects the increased number of choice alternatives due to the disaggregation to line-zones. Each additional choice alternative has its own utility, which intensifies competition for the network exit, necessitating compensation to maintain the same transfer-to-non-transfer ratio as in method 1.

Similar to method 1, the transfer time coefficient is significantly higher when the destination mode is a tram compared to the other modes. However, unlike method 1 - model 1, the transfer time for trains is statistically significant.

For the travel time, it can be seen that it is statistically significant for both models, with the tram coefficient having the highest magnitude for both. This pattern is reflected in the route-level OD matrices, where it can be seen that the first 5 stops after Lancy-Bachet hold 50% of the total flow from this stop. Furthermore, as

Table 11: Method 2 estimation results - Lancy-Bachet (morning peak)

		Model 1			Model 2		
Attribute	Mode	Coefficient	Std-Error	P-value	Coefficient	Std-Error	P-value
Intercept	Π	4.96	0.472	0	8.82	0.831	0
Transfer time	train	-2.221	1.088	0.041	-2.458	1.481	0.097
	tram	-12.777	2.073	0	-21.179	2.745	0
	bus	-4.445	1.159	0	-6.573	1.21	0
Travel time	train	-10.951	1.575	0	-5.287	1.741	0.002
	tram	-21.753	2.207	0	-11.794	1.517	0
	bus	-9.402	2.133	0	-8.089	2.401	0.001
Alighting	train	4.955	0.594	0	-	-	-
	tram	10.323	0.945	0	-	-	-
	bus	5.328	0.466	0	-	-	-
Workplace dens	train	-	-	-	8.688	1.063	0
	tram	-	-	-	11.873	1.1	0
	bus	-	-	-	9.026	1.024	0
Angle cost	train	-2.002	0.480	0	-2.371	0.485	0
	tram	-1.164	0.316	0	-1.916	0.395	0
	bus	-2.425	0.488	0	-3.286	0.494	0
From train	tram	-	-	-	0.792	0.362	0.028
	bus	-	-	-	0.717	0.327	0.028
From tram	train	1.192	0.397	0.003	1.339	0.406	0.001
	bus	0.826	0.365	0.024	1.668	0.421	0
Passed	tram	-1.109	0.355	0.002	-	-	-
Served Downs.	ALL	-1.876	0.508	0	-1.453	0.509	0.004
DCR	train	-1.424	0.542	0.009	-0.981	0.527	0.063
	tram	-1.197	0.555	0.031	-	-	-
	bus	-2.617	0.941	0.005	-2.788	0.903	0.002

mentioned in Section 4.1 and demonstrated in Section 4.2, most transfers are made to the first or second stops on the tram routes. This pattern is also observed in train and bus lines, where many transfers don't extend beyond the first or second stops on their respective transfer-receiving lines. However, this effect is less pronounced, which is reflected in the coefficients. Additionally, due to the nature of tram lines, they experience a high number of alightings at downstream locations and pass through areas with a high density of workplaces, which are located beyond the first or second stops. As a result, a stronger deterrent is needed for these destinations compared to trains or buses.

The next variables, the number of alightings for model 1 and the workplace density for model 2, are statistically significant and depict similar behavior. In both models, the tram coefficient is the highest, followed by the bus, with the train coefficient being the lowest. However, the difference between these mode coefficients varies depending on the model. For model 1, the gap between the tram coefficient and the other modes is particularly large, whilst for model 2, they are closer together. This disparity in the coefficient values is partially a consequence of the corresponding attribute values, but also of the transfer patterns at Lancy-Bachet. As mentioned in Section 4.1, line 12_0 has some of the most popular Lancy-Bachet transfer destinations, leading to a higher attribute value for trams. Comparatively, the trains serve stops with very high alighting counts and workplace densities which are also quickly reached, such as "Genève Champel",

"Gare Cornavin", or "Genève, Eaux-vives". But, although these are quite attractive transfer destinations, there are not as many transfers to these destinations as for destinations on line 12_0, possibly due to the less efficient transfer-to-train route compared to direct connections. Consequently, the attractiveness attribute values are disproportionate, leading to an overall lower coefficient value.

For model 1, the reason the "From train" dummy is not utilized is due to its statistical insignificance. This is different compared to method 1 model 1, where, though also statistically insignificant at the $\alpha = 0.05$ level, the inclusion of the variable had a minor impact on the model fit. This improved model fit may simply be due to an additional variable. However, this disparity in the contribution is possibly due to the disaggregation of the choice set(s), which signifies that, from the perspective of a train passenger, a tram, a bus, or exiting the network are described well enough with the other variables. The reason the dummy of method 2 is less significant than that of method 1 is that the travel time and direct alighting count are better descriptors of the utility of a choice alternative, in combination with the addition of the other new variables.

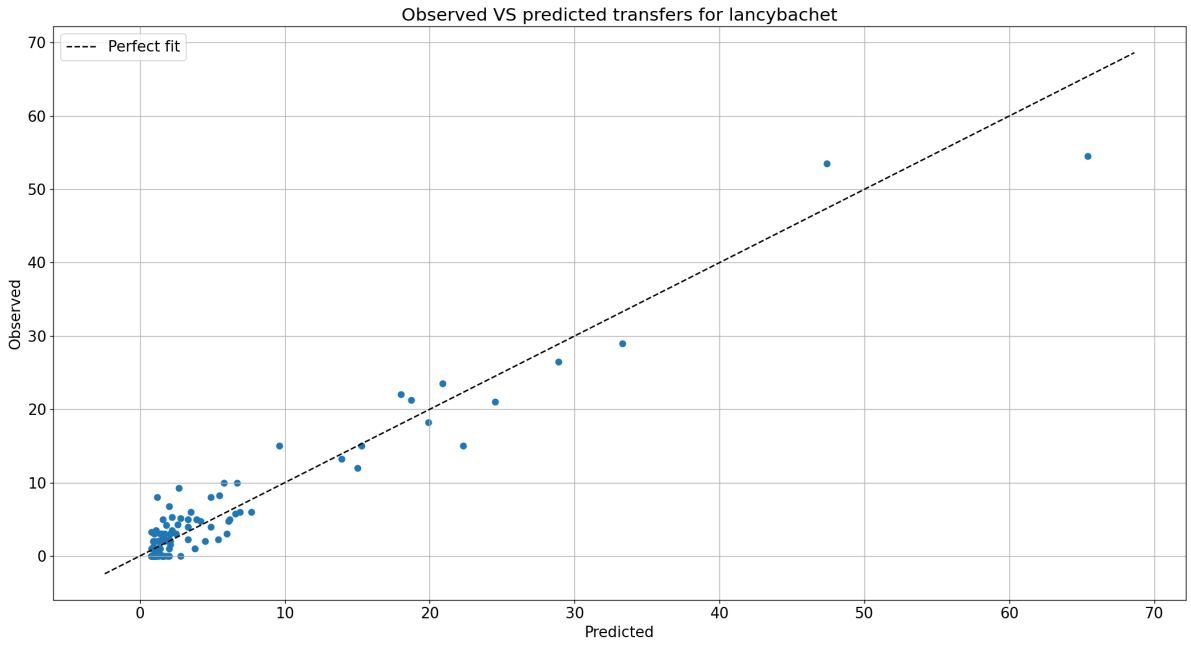
On the other hand, model 2 makes use of the "From train" variable, which is statistically significant, and for which the coefficients are statistically the same as in method 1. This indicates that the utilized variables, in particular the workplace density, show an insignificant contribution to distinguishing the utility of choice alternatives from the perspective of a passenger arriving by train. This is partly because trams and buses often have destinations with a high number of transfers from trains, where workplace density is low, such as for the transfers to line D_0, stops "Plans-Les-Ouates, Arare", "Plans-Les-Ouates, Vélodrome", and "Plans-Les-Ouates, Galaise", or to line 12_0, stops "Carouge, Ancienne" and "Carouge, Marché". Additionally, the intercept for exiting the network (Π) is much greater, in combination with the magnitude of the transfer time coefficient being much greater for model 2 than model 1, possibly inducing greater attraction for "From Train" transfers.

The "Passed" dummy variable is only used by model 1 for the tram as a destination mode. This is due to it being statistically insignificant for the other modes, as well as for model 2 entirely. In the former case, the coefficient value is negative, indicating, as expected, that destination zones on the upstream part of the transfer-producing route have lesser utility and thus less demand. On the other hand, it has been removed from the other modes and model 2 in part due to it overlapping with the angle cost, resulting in statistical insignificance.

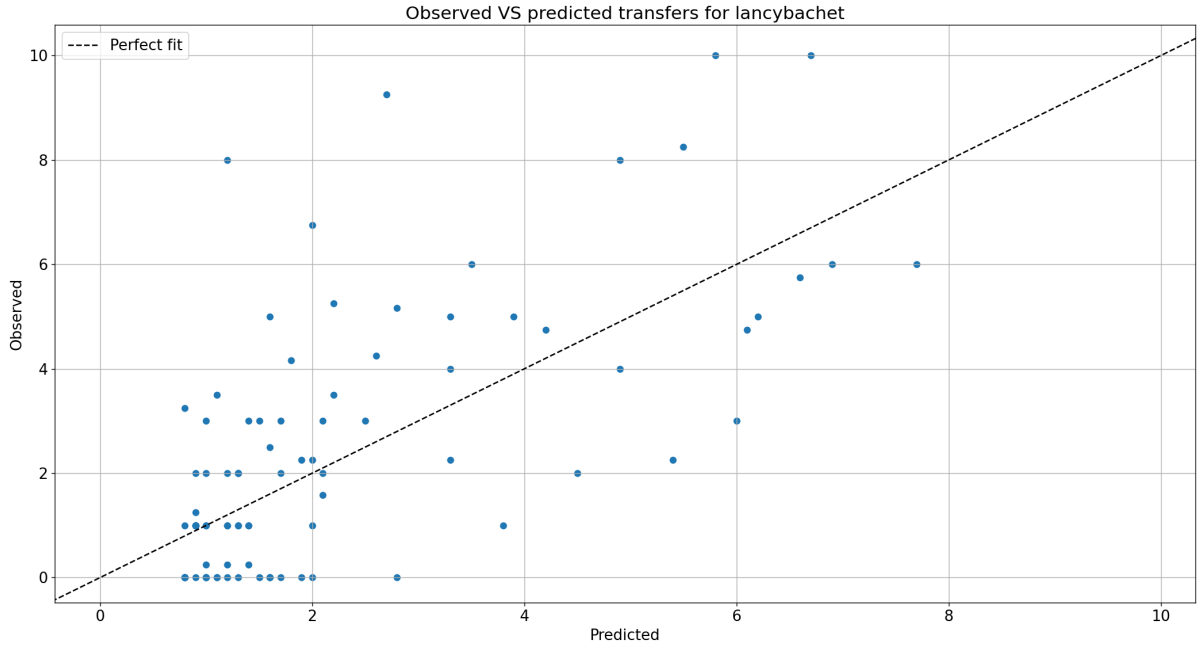
The "Served Downstream" dummy variable, which is grouped for all modes (except Π), has a negative coefficient value and is statistically significant, meaning that people would rather continue on their current mode than transfer to another. This tendency is also reflected in the "Direct Connection Ratio" (DCR) variable, where all coefficients are statistically significant, except for model 2, tram. In particular for buses, the DCR has a notably large coefficient value, meaning that people do not want to transfer to a bus if they don't have to. This may be due to the perception they have of the bus as a mode of transportation, such as slower or less reliable due to being stuck in traffic jams. Compared to trams, the bus has a much lower frequency of passing, requiring much longer waiting times if transferring to one. Furthermore, compared to the other two modes, this one is often perceived as less reliable, particularly compared to trains.

Figure 13 shows the predicted (expected) and observed transfer-destination flow at Lancy-Bachet, where each point represents an arrival line-direction + departure line-direction-zone combination. As depicted,

the general form of the predictions is as desired, with them being close to the perfect fit line, especially from an observed flow greater than 10. However, within the observed flow range of 5 to 10, there are several points with the same observed flow but differing expected flows, along with a point that is significantly farther to the left (with an observed flow of 8). These will be discussed more in detail in Section 5.2.1.



(a) Predicted VS observed transfer-destination flows



(b) Predicted VS observed transfer-destination flows (< 10)

Figure 13: Comparison of predicted (x-axis) and observed (y-axis) transfer-destination flows at Lancy-Bachet using the transfer sample and model 1 (points where the transfer flow is less than 0.75 are not shown)

Next-to-last, Table 12 shows the value τ was set to, as well as the performance indicators. First, are those generated when comparing the expected transfer-destination flow and the observed flows. And second, denoted by "Agg", are those generated when comparing the aggregation to the transfer flow (line-direction to line-direction).

Table 12: Method 2 performance indicators - Lancy-Bachet (morning peak)

	Model 1 (A)	Model 2 (WI)
τ	0	0
$\bar{\rho}^2$	0.528	0.508
MTE	0.389	0.413
RMSE	0.457	0.488
PCC (<5)	0.7	0.639
MTE (Agg)	0.271	0.308
RMSE (Agg)	1.181	1.302
PCC (<5, Agg)	0.741	0.697

As depicted, both models have excellent fit, with $\bar{\rho}^2$ being 0.528 and 0.508, respectively. Furthermore, the non-aggregated performance indicators are within the expected margin. As displayed, the MTE is around 0.4 for both models, meaning that the average transfer-destination error is at 40%. This follows from the large number of zero flows (93%), where many errors are present by construction, due to the inability of a MNL model to predict zero-flows, which are also a cause of the relatively high RMSE.

The aggregated indicators show particularly high performance. As depicted, the MTE and the RMSE are lower for both of the models when compared to their method 1 counterparts. However, for some, there is little to no difference, such as for the MTE of model 2. On the other hand, the PCC is slightly lower, indicating that this performance increase came at a loss for the lower flows. This is in part supported by Figure 13b, where many observed zero flows are estimated to be greater than zero, further reinforcing the point that the lower density results in a greater number of errors by the model.

Finally, the coefficient importance of method 2 - model 1 was also investigated. This led to the results shown in Table 13. As depicted, analog to method 1, the coefficients with the highest dominance are once again the "Transfer time" and the "Intercept", though a close third is the "Travel time". The first two follow for the same reasons as their importance in method 1. For the additional "Travel time", this behavior is also expected, as it is often depicted as one of the most important variables in route and OD modeling.

Table 13: Coefficient dominance method 2 - model 1

Intercept	Transfer time	Travel time	Alighting count	Angle cost	From tram	Served downs.	DCR
0.154	0.16	0.102	0.03	0.054	0.004	0.003	0.019

5.2.1 Transfer-destination flow analysis

As mentioned prior and shown in Figure 13, the model wrongly predicts certain flows. A first noticeable point is the aforementioned large gap between the observed flow of 8 and the predicted flow of 1, which

is for line D_1 to destination "Carouge, Armes" on line 12_0. The main reason for this lower estimation is the low alighting count at this destination, as well as a higher angle cost, compared to other high-demand destinations, inducing a lower utility despite a low travel time. This is the same pitfall as the angle cost issues addressed in Section 5.1.1, where, despite this being the fastest, and possibly only connection to a destination, the angle cost is moderately high, as the trip directions do not align perfectly.

Another larger error is the transfers from line D_1 to LEX $L4_0$, where they are underestimated by 6.5 (from 9.2 to 2.7). The underestimated destination, in this case, is "Genève-Eaux-Vives", which is 2 stops downstream. Although this stop has adequate explanatory variable values, high alighting count, low travel time and angle cost, and is not served by line D_1, the predicted flow is very low. This is due to the transfer sample. As described in Section 4.1, the representativeness of the train as a mode of transportation is quite low in the sample, leading to a low observed number of transfers to and fro, and decreasing the modeled utility.

Furthermore, another error that must be addressed is the general overestimation of network exits from lines LEX $L2_1$ and D_1, which leads to the largest errors of the model (overestimation of ~10 passengers in both cases). For line LEX $L2_1$, this is mainly due to the fact that the average number of network exits per train line is around 70%, whereas the number of network exits for line LEX $L2_1$ is 43%. This means that the coefficient values have adapted to better explain the behavior of the other lines and consider LEX $L2_1$ an anomaly.

For line D_1, the error is, in part, a consequence of the aforementioned underestimation of transfer flow to line LEX $L4_0$, which is redirected to the network exit instead. In particular, the underestimation of the transfer-destination flow to LEX $L4_0$-"Genève-Eaux-Vives" has the most impact.

Finally, concerning zero-flow errors, many are made to lines 12_0, 18_0, and LEX $L1_1$, which go through workplace-dense neighborhoods and where alighting counts can be very high. However, these destinations are further away, and thus are not destinations of transfer-destination flow. This is in part taken care of by the travel time, but not necessarily to a sufficient degree, resulting in some greater residue.

5.2.2 Application to scaled flows

Applying the coefficients to the scaled transfer sample led to the results shown in Table 14. As depicted, the indicator values are similar to the non-scaled ones, with the non-aggregated MTE, being slightly better by (0.042 and 0.045, respectively), and the remaining indicators being worse. In particular, the non-aggregated PCC is significantly worse for model 1, dropping from 0.7 to 0.5, which, though to a lesser degree, is also reflected in the non-aggregated model 2 performance, as well as the aggregated PCC. This strong drop is, analog to method 1, due to the same number of zero-flows being present, even though the number of arriving passengers has increased. And, this effect is more perceivable due to the density of the sample transfer-destination flow matrix being so low.

An additional point that must be mentioned, is that, as shown in Table 14, model 2 performs better than model 1.

The largest prediction errors, shown in Figure 14, are for the network exits of lines LEX $L2_1$ (towards Coppet) and D_1 (towards the city center), where these are generally overestimated. This is the same behavior as for

Table 14: Method 2 performance indicators - Lancy-Bachet scaled

	Morning peak	
	Model 1 (AI)	Model 2 (WI)
MTE	0.347	0.368
RMSE	0.657	0.652
PCC (<5)	0.541	0.585
MTE (Agg)	0.272	0.284
RMSE (Agg)	1.734	1.75
PCC (<5, Agg)	0.503	0.582

the non-scaled results, though, in this case, the difference is amplified, which is due to the increased flow. Although, it must be noted that the percentage overestimation has stayed the same.

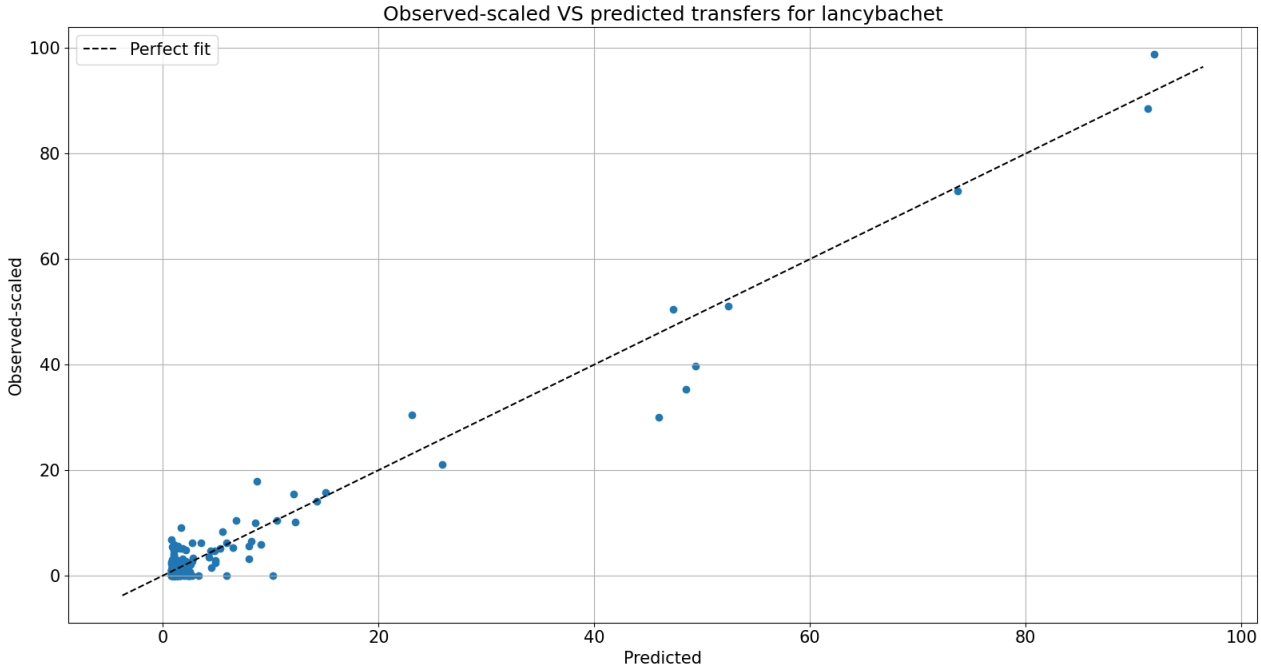


Figure 14: Predicted (x-axis) VS observed-scaled (y-axis) transfer-destination flows at Lancy-Bachet using transfer sample and model 1 (points where the transfer flow is less than 0.75 are not shown)

Another prediction error to note is the strong overestimation of transfers from line D_1 to line 12_0 stop "Grand-Lancy, De Stael", where according to the sample, no passengers transfer. This large error is due to very low attribute values for the transfer time, travel time, and angle cost. And is coupled with a fairly high alighting count, increasing the utility of this alternative substantially. This behavior is also evident in the number of transfers from line D_1 to line 18_0 at the "Grand-Lancy, De Stael" stop, where a significant number of passengers were expected, yet none were observed. One possible reason for not observing any such passengers is that some upstream stops on line D_1 have a partial direct connection to "Grand-Lancy, De Stael", about 40%, which is why the model still predicts some flow.

Finally, another error that must be highlighted is the underestimation of the transfer-destination flow from line LEX $L1_0$ (Geneva to Annemasse) to line 12_0 (towards the city center) stop "Grand-Lancy, De Stael". Due

to the arrival angle of line LEX $L1_0$, a very high angle cost is present (approx. 1), inducing an exceptionally low utility and presenting this destination as particularly undesirable.

5.2.3 Scaled flow regression

Using the scaled flows directly to regress the model coefficients for the morning peak yields the coefficients presented in Table 15. In the alighting model (model 1), though the magnitude of the coefficients has increased, all coefficients except for "Alightings - train" are statistically significant at $\alpha = 0.05$, similar to those calibrated using the non-scaled transfer sample. On the other hand, for model 2, where the increase in magnitude is also present, six coefficients are not the same, with the differences in these cases being very significant.

Table 15: Method 2 scaled estimation results - Lancy-Bachet (morning peak)

		Model 1			Model 2		
Attribute	Mode	Coefficient	Std-Error	P-value	Coefficient	Std-Error	P-value
Intercept	Π	5.547	0.501	0	12.376	0.980	0
Transfer time	train	-4.178	1.293	0.001	-5.89	1.851	0.001
	tram	-16.197	2.075	0	-22.777	2.818	0
	bus	-5.564	1.218	0	-9.973	1.288	0
Travel time	train	-13.422	2.362	0	-2.984	2.252	0.185
	tram	-28.348	2.761	0	-22.397	2.052	0
	bus	-11.952	2.368	0	-13.757	3.004	0
Alighting	train	7.03	0.734	0	-	-	-
	tram	12.351	1.139	0	-	-	-
	bus	6.556	0.439	0	-	-	-
Workplace dens	train	-	-	-	14.609	1.328	0
	tram	-	-	-	17.181	1.335	0
	bus	-	-	-	13.863	1.274	0
Angle cost	train	-2.083	0.500	0	-2.545	0.511	0
	tram	-1.194	0.263	0	-1.661	0.358	0
	bus	-2.387	0.453	0	-3.925	0.458	0
From train	tram	-	-	-	0.708	0.335	0.035
	bus	-	-	-	2.031	0.424	0
From tram	train	0.582	0.387	0.133	0.853	0.401	0.033
	bus	0.402	0.362	0.266	2.329	0.518	0
Passed	tram	-0.84	0.269	0.002	-	-	-
Served Downs.	ALL	-2.187	0.713	0.002	-1.557	0.716	0.03
DCR	train	-1.156	0.532	0.03	-0.814	0.511	0.111
	bus	-2.132	0.836	0.011	-3.068	0.792	0

Of these, the coefficient with the largest difference is the Travel time (tram) coefficient, which is twice as large as that of the non-scaled regression. This increased coefficient value is due to a change in the spread of the transfer-destination flow on trams, depending on the travel time. Scaling the transfer sample results in a higher peak at low travel times and a narrower tail at higher values compared to the non-scaled transfer sample. This leads to higher coefficient magnitudes in the models compared to the non-scaled regression, with the effect being more pronounced in model 2. This behavior indicates that model 1 and the alighting

counts are less prone to such effects and that the attractiveness explanatory variables and travel time have some overlapping responsibilities.

With regards to the performance indicators of the scaled regression, shown in Table 16, it can be seen that, similar to the scaled regression of method 1, both models have a very high $\bar{\rho}^2$ value, showing excellent fit. On the other hand, the other performance indicators have not improved much from the result shown in Section 5.2.2, though contrary to method 1, there is some improvement.

Table 16: Method 2 performance indicators - Lancy-Bachet scaled regression

	Morning peak	
	Model 1 (AI)	Model 2 (WI)
τ	0	0
$\bar{\rho}^2$	0.652	0.632
MTE	0.347	0.389
RMSE	0.651	0.694
PCC (<5)	0.53	0.573
MTE (Agg)	0.283	0.316
RMSE (Agg)	1.831	1.937
PCC (<5, Agg)	0.625	0.567

5.2.4 Sensitivity of τ

Method 2 uses a single hyperparameter, τ, which is the shape parameter that describes the application of the Box-Tuckey transformation, and for which the model response was investigated.

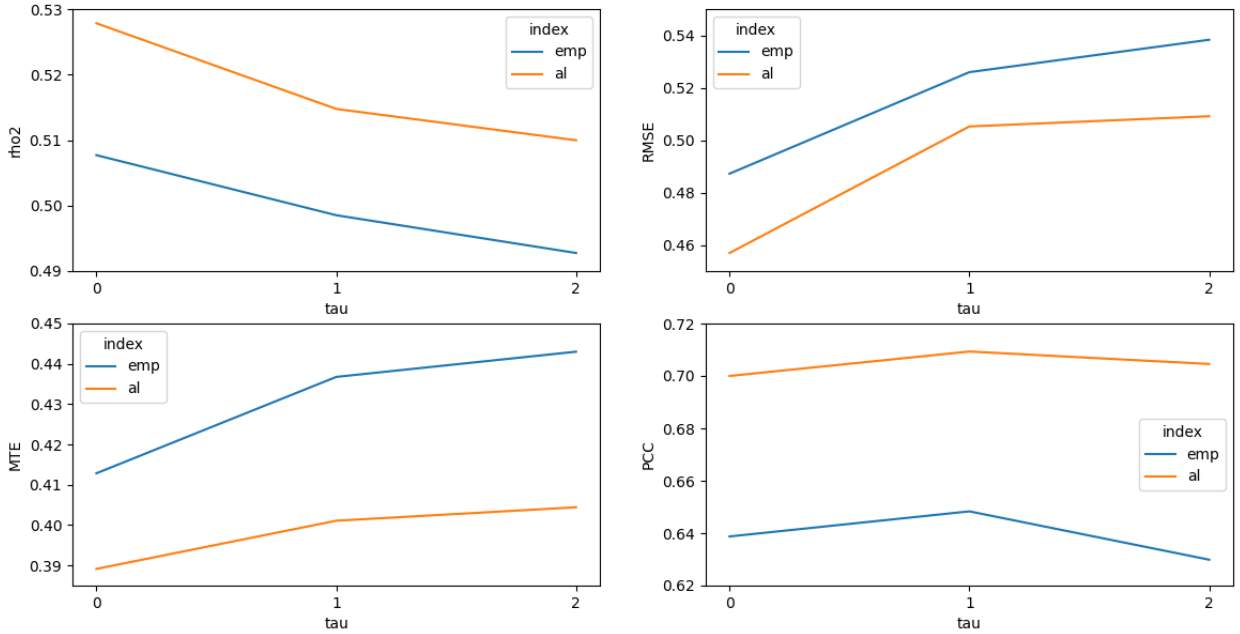


Figure 15: Sensitivity of performance indicator based on used attractiveness variable and τ (left: model 1, and right: model 2)

Figure 15 shows the performance indicator values with varying τ for the morning peak models using the alighting counts and the workplace density. As can be seen, irrespective of the indicator and the value of τ, the model using the alighting counts performs better than the workplace model. On the other hand, both models are equally insensitive to changes in τ, with only minor differences for the MTE and for the PCC. Additionally, as depicted, the first three performance indicators, $\bar{\rho}^2$, MTE, and RMSE all follow the same pattern, where $\tau = 0$ has the best performance, irrespective of whether the workplace density index or the alighting index is used, followed by a larger drop in performance when τ increases to 1, and a smaller drop in performance from 1 to 2.

These drops in performance are most likely due to the nature of both explanatory variables. With the linearity and squaring of the attribute value ($\tau \in \{1,2\}$), differences between alighting/workplace density counts are exemplified, thus making the coefficient value more influential on the utility of an alternative.

Regarding the PCC, it does not follow the same pattern as the other indicators. There is a drop in performance when τ increases from 0 to 1, but performance improves again when τ rises to 2. This behavior is similar to the sensitivity of the models developed for method 1 and indicates that the PCC is not a direct reflection of the general model performance. Furthermore, as mentioned in section 5.1.4, it also indicates and reinforces the notion that the calibration methodology puts a lower weight on these low flows.

5.3 Evening results

Due to time constraints, the evening-peak models were not investigated in detail. The results of the calibration for method 1, using the Alighting Index and the Population Density Index (AI and PI) are shown in Table 17. For both of these models, τ was set to 0, and $\beta = 1.3$ and $\beta = 1.7$ for models 1 and 3, respectively.

Table 17: Method 1 estimation results - Lancy-Bachet (evening peak)

		Model 1			Model 3		
Attribute	Mode	Coefficient	Std-Error	P-value	Coefficient	Std-Error	P-value
Intercept	Π	5.217	0.345	0	5.106	0.621	0
Transfer time	train	3.015	0.759	0	-	-	-
	tram	-4.981	2.114	0.018	-15.517	3.091	0
	bus	-	-	-	-7.671	1.031	0
Alighting index	train	4.221	1.080	0	-	-	-
	tram	12.615	1.308	0	-	-	-
	bus	13.370	1.091	0	-	-	-
Population index	train	-	-	-	6.028	1.426	0
	tram	-	-	-	8.86	1.311	0
	bus	-	-	-	7.932	0.950	0
Angle cost	train	-0.789	0.335	0.018	-0.938	0.348	0.007
	tram	-0.83	0.375	0.027	-0.482	0.370	0.193
	bus	-1.267	0.252	0	-1.392	0.252	0
From train	bus	-	-	-	0.753	0.301	0.12
From tram	bus	0.863	0.300	0.004	1.099	0.299	0

As can be seen, a noticeable characteristic of the estimation results is the transfer time variable. First, for model 1, the "Transfer time - train" attribute is positive, and for model 3, using PI, it is not used. Concerning

model 3, the transfer time attribute is not statistically significant at the $\alpha = 0.05$ level and was removed from the model. On the other hand, for model 1, although the coefficient is positive and thus contrary to the expectation that lower transfer times would be associated with higher utility, it was retained because it accurately reflects the data. As shown in Figure 16, the number of transfers increases with a rising transfer time, instead of decreasing. This behavior may indicate that the average transfer to a train is subject to long waiting times (20 minutes) and that commuters heading back home have no other choice but to wait longer times. On the other hand, this may also indicate a flaw in the definition of the variable, where the transfer time is an average. The possibility, in this case, is that the connection where many individuals transfer is subject to a shorter transfer time. But on average, this connection is poorly served. Furthermore, this poorer service may restrict travel options during the evening peak.

The second noticeable characteristic of the transfer time variable is for model 1, for the bus as a destination mode. As depicted, it is absent from the model, as it was removed because the coefficient value was not tractable and statistically insignificant.

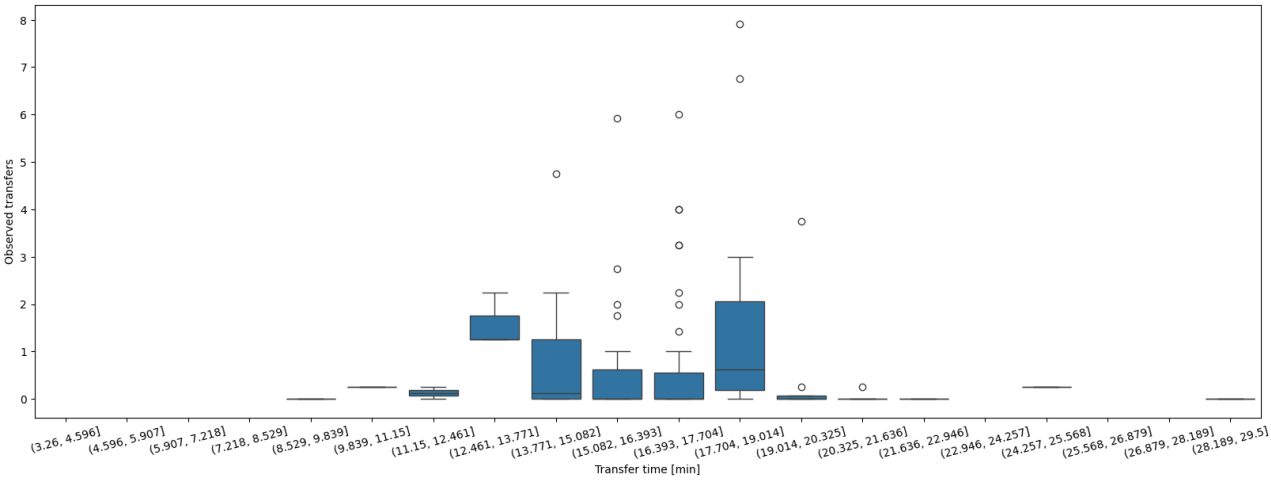


Figure 16: Observed number of transfers compared to the transfer time (for trains)

The other explanatory variables exhibit the expected trend and signs, similar to the morning peak. However, there is a notable difference in the Alighting index, where the coefficient for the bus is higher than that for the tram. This discrepancy reflects the station and city layout. Passengers typically arrive from the city center by tram, where there are many workplaces, and then transfer to a bus that services the residential areas. Additionally, the high coefficient for the tram is due to the high population density in the city center.

Table 18 shows the calibration results for method 2 during the evening peak, where τ was set to zero for two models. Model 1 uses the Alighting index and model 2 uses the population density. As can be seen for both models, the first noticeable difference to the morning peak estimation results is that the Intercept has not grown as much as during the morning peak when compared to method 1. And, for model 3, the value has even shrunk. This reduced intercept value is very unexpected, as the nature of the transfer sets would generally result in the contrary. However, this behavior can be attributed to the low and statistically insignificant Population density coefficient values. The utility of the intercept aligns with these low coefficients, as it represents the baseline increase in the utility value of the network exit. In other words, it doesn't need to demonstrate higher utility because the alternatives appear less favorable.

Table 18: Method 2 estimation results - Lancy-Bachet (evening peak)

		Model 1			Model 3		
Attribute	Mode	Coefficient	Std-Error	P-value	Coefficient	Std-Error	P-value
Intercept	Π	6.945	0.260	0	2.243	0.422	0
Transfer time	tram	-4.521	2.306	0	-4.166	2.987	0.163
	bus	-	-	0	-5.223	0.767	0
Travel time	train	-8.065	1.196	0	-8.964	1.087	0
	tram	-17.805	2.395	0	-8.45	1.468	0
	bus	-7.593	1.212	0	-5.612	0.966	0
Alighting	train	5.245	0.396	0	-	-	-
	tram	8.371	0.929	0	-	-	-
	bus	6.747	0.373	0	-	-	-
Population dens	train	-	-	-	-0.806	0.513	0.116
	tram	-	-	-	0.094	0.741	0.899
	bus	-	-	-	0.617	0.444	0.164
Angle cost	train	-0.524	0.422	0.215	-0.644	0.363	0.076
	bus	-1.324	0.231	0	-1.844	0.252	0
From train	tram	-	-	-	-0.976	0.386	0.011
From tram	train	0.447	0.356	0.209	-	-	-
	bus	0.804	0.185	0	-0.976	0.182	0
Passed	tram	-	-	-	-1.86	0.779	0.017
Served Downs.	ALL	-	-	-	-0.677	0.337	0.045
DCR	train	-2.108	0.543	0	-1.580	0.511	0.002
	tram	-2.566	0.769	0.001	-2.073	0.795	0.009
	bus	-2.867	0.456	0	-2.611	0.476	0

As mentioned previously, model 3 has a statistically insignificant attractiveness explanatory variable. This means that it is not able to adequately capture the transfer behavior of passengers, making this model unsuitable for use.

In addition to the estimation results, Table 19 depicts the performance indicators results of the models for the evening peak. As shown, the performance is substantially worse than during the morning, with only the $\bar{\rho}^2$ indicating excellent fit. In particular, method 2 - model 3 shows the worst performance. This is a reflection of the poor estimation results. As described previously, the Population density is incapable of capturing how attractive a destination is, thus improperly representing the utility of each transfer-destination and resulting in very poor performance.

Beyond that, as can be seen, the performance of model 1, irrespective of the method, is overall very similar, with the performance indicators being close to one another. The most noticeable difference is for the RMSE, where that of method 2 is smaller by 0.167. However, the other performance indicators are all within ± 0.01.

Furthermore, as depicted in Figure 17, the general prediction capability of method 2 - model 1 follows the perfect fit. There tends to be a greater spread around the line than for the morning peak, but the expected relation for flows > 10 is present.

Table 19: Method 1 and method 2 performance indicators - Lancy-Bachet (evening peak)

	Method 1		Method 2	
	Model 1 (AI)	Model 3 (PI)	Model 1 (AI)	Model 3 (PI)
τ	0	0	0	0
β	1.3	1.7	-	-
$\bar{\rho}^2$	0.419	0.414	0.470	0.411
MTE	-	-	0.468	0.511
RMSE	-	-	0.548	0.646
PCC (<5)	-	-	0.574	0.459
MTE (Agg)	0.377	0.399	0.389	0.44
RMSE (Agg)	1.785	1.859	1.618	2.117
PCC (<5, Agg)	0.57	0.657	0.561	0.528

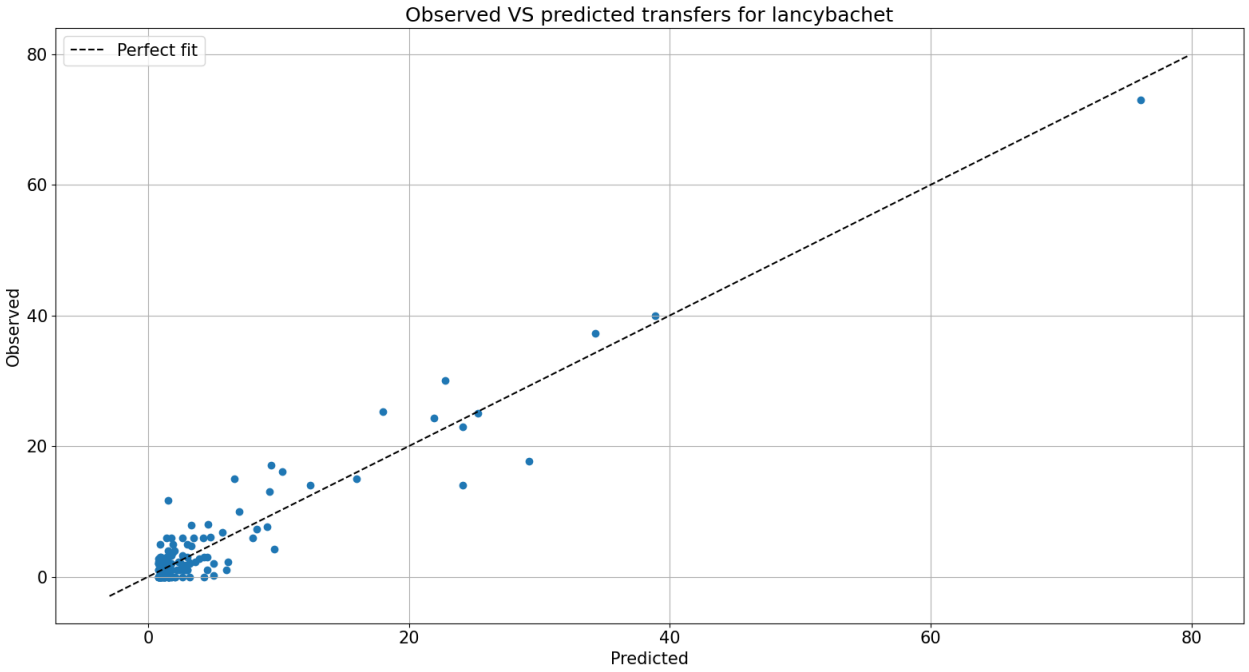


Figure 17: Predicted (x-axis) VS observed (y-axis) transfer flow at Lancy-Bachet using transfer sample and method 2 - model 1 during the evening peak (points where the transfer flow is less than 0.75 are not shown)

6 Model transferability

The final results consider the transferability of the models, i.e., the utilization of the models and coefficients calibrated at Lancy-Bachet at other stops in Geneva. In this case, two types of transferability are distinguished: (1) the transferability of the calibrated coefficients and (2) the transferability of the framework. To this end, the best models (subset of chosen explanatory variables) of methods 1 and 2 were regressed at 10 additional transfer stations in Geneva, which had the highest survey representativeness.

6.1 Coefficient transferability

6.1.1 Method 1

Overall, the coefficients of method 1 - model 1 (using the alighting index) and calibrated at Lancy-Bachet were found non-transferable. Table 20 shows the coefficient comparison for Gare Cornavin, the main train station of Geneva.

At the $\alpha = 0.05$ level (two-sided 95% confidence level), the null hypothesis that the coefficients for Lancy-Bachet are equal is rejected if the absolute z-value is greater than 1.96. As depicted in Table 20, for Gare Cornavin, eight of the coefficients are statistically the same and six are not. Of the coefficient values that are not statistically the same, five are significantly different, with the confidence level in the rejection being almost 100% for the alighting index for trams.

The first coefficients that are statistically different from one another are the transfer time for trains and the transfer time for trams. For the former, instead of being statistically insignificant and close to zero, the coefficient value is now strongly negative and statistically significant. This is partly due to the larger diversity in the stations served by the trains. Now, instead of having four train lines serving virtually the same locations, there are five lines serving two vastly different regions. Furthermore, due to the size of Gare Cornavin, a larger number of alternative routes are available, which reduces the necessity to transfer to a train. This is reflected in the meager number of transfers to trains, sitting at 20% of all transfers, which is 5% lower than at Lancy-Bachet.

With regards to the transfer time to trams, the magnitude of the coefficient value is much lower for Gare Cornavin than for Lancy-Bachet. This is in part due to the lower competitiveness between the passing tram lines. At Lancy-Bachet only two tram lines are passing. They go in virtually the same direction, have identical frequencies, and thus have very similar transfer times. However, one is vastly more popular than the other. On the other hand, at Cornavin, there are three tram lines that don't compete for the same market and have larger differences in their transfer times, leading to a lower necessity to extenuate these differences.

The coefficients that are the least similar are the alighting index coefficients. In particular for trams, for which the values have the largest statistical difference, and, instead of having the largest value, it is now almost 0. Furthermore, the coefficient value for trains is the highest, instead of the lowest. One of the reasons for this shift is that, although the demand for a tram transfer is much higher, with 40% of all transfers being to a tram, the alighting counts at the closest stops for the trams are much greater. This leads to significantly higher alighting index values and a larger gap to the other modes. This means that the coefficient value does not need to be as large to reflect this perceived utility.

Table 20: Method 1 coefficient comparison - Lancy-Bachet VS Gare Cornavin (morning peak)

		Lancy-Bachet		Gare Cornavin		Statistical test	
Attribute	Mode	Coefficient	Std-Error	Coefficient	Std-Error	z-value	< 1.96
Intercept	Π	2.927	0.561	3.007	0.164	0.136	✓
Transfer time	train	-0.63	1.025	-4.63	0.441	3.584	*X*
	tram	-14.21	2.396	-1.63	0.444	5.163	*X*
	bus	-5.016	1.457	-2.755	0.542	1.454	✓
Alighting index	train	1.149	0.261	2.372	0.222	3.570	*X*
	tram	3.565	0.420	0.395	0.094	7.366	*X*
	bus	2.4	0.509	-0.234	0.231	4.713	*X*
Angle cost	train	-1.91	0.438	-0.987	0.263	1.806	✓
	tram	-2.083	0.499	-0.766	0.138	2.545	*X*
	bus	-2.753	0.463	-1.884	0.159	1.776	✓
From train	tram	0.809	0.451	0.25	0.104	1.209	✓
	bus	0.414	0.339	0.285	0.135	0.355	✓
From tram	train	0.553	0.341	-0.146	0.157	1.863	✓
	bus	0.651	0.399	0.235	0.129	0.992	✓

With regards to the alighting index for buses, as can be seen, the coefficient value is now negative, instead of positive. Although, it must be noted that it is statistically insignificant, with a P-value of 0.311. This suggests that, unlike in Lancy-Bachet and the other two modes, transferring passengers do not consider the attractiveness of a bus line. However, it is more likely that the alighting index is not representative of the attractiveness of a bus line at Gare Cornavin.

Finally, except for some, the Gare Cornavin coefficients have a smaller absolute value. This is in part due to the larger choice set, which is also the reason the Intercept coefficient is slightly higher, as well as the larger number of non-transferring passengers, since, at Gare Cornavin, 62% exit the public transportation network.

6.1.2 Method 2

Similar to method 1, the model coefficients of method 2 are also not transferrable to different stations. Table 21 shows the coefficient comparison for Gare Cornavin.

As depicted, 11 of the 20 coefficients are not transferrable, with the highest confidence level in the rejection being for the "Travel time - tram" explanatory variable and the most similar coefficient being the "Direct connection ratio - tram". In the case of the former, the magnitude of the coefficient value is much smaller, indicating that transferring passengers travel further to their final destination than at Lancy-Bachet.

The second largest difference between coefficients is for the network exit Intercept, where the coefficient value is about 1.7 times as large as for Lancy-Bachet. This is, in part, due to two factors. First is the increased proportion of network exits. As mentioned in Section 6.1.1, the percentage of passengers exiting the network is 12% higher, meaning that Π has to have a greater base attractiveness. Furthermore, because of the central location of Gare Cornavin, the choice set is about 100% larger, meaning that exiting the network has much more competition, further increasing the necessity for a greater intercept coefficient.

Finally, similarly to method 1, the tendency is, with some exceptions, for the magnitude of the coefficient values to be more attenuated. As mentioned previously, this is partly due to the increased number of choice alternatives in the choice set.

Table 21: Method 2 coefficient comparison - Lancy-Bachet VS Gare Cornavin (morning peak)

		Lancy-Bachet		Gare Cornavin		Statistical test	
Attribute	Mode	Coefficient	Std-Error	Coefficient	Std-Error	z-value	< 1.96
Intercept	Π	4.96	0.472	8.428	0.214	6.691	*X*
Transfer time	train	-2.221	1.088	-2.75	0.531	0.436	✓
	tram	-12.777	2.073	-2.1	0.495	5.010	*X*
	bus	-4.445	1.159	-1.644	0.511	2.211	*X*
Travel time	train	-10.951	1.575	-6.598	0.958	2.361	*X*
	tram	-21.753	2.207	-3.412	0.706	7.915	*X*
	bus	-9.402	2.133	-2.016	0.370	3.412	*X*
Alighting	train	4.955	0.594	7.675	0.342	3.969	*X*
	tram	10.323	0.945	6.449	0.282	3.928	*X*
	bus	5.328	0.466	5.93	0.305	1.081	✓
Angle cost	train	-2.002	0.480	-1.567	0.242	0.810	✓
	tram	-1.164	0.316	-1.432	0.143	0.774	✓
	bus	-2.425	0.488	-1.603	0.142	1.617	✓
From tram	train	1.192	0.397	0.224	0.157	2.269	*X*
	bus	0.826	0.365	0.166	0.104	1.740	✓
Passed	tram	-1.109	0.355	0.059	0.260	2.654	*X*
Served Downs.	ALL	-1.876	0.508	-0.344	0.151	2.892	*X*
DCR	train	-1.424	0.542	-1.04	0.403	0.570	✓
	tram	-1.197	0.555	-1.19	0.234	0.013	✓
	bus	-2.617	0.941	-1.187	0.220	1.480	✓

6.2 Framework Transferability

When considering framework transferability, i.e., the option to apply the defined decision choice model and explanatory variables to another transfer station, method 2 strongly outperforms method 1. For the latter, of the 10 additional locations selected, only 6 models have interpretable coefficient values and are thus fully tractable. On the other hand, for method 2, all 10 locations have fully tractable coefficients. This difference stems from the defined choice set, as well as from the corresponding more complete and disaggregated explanatory variable set utilized. This allows for a more accurate interpretation of the utility of the choice alternatives, thus enabling the model to be calibrated more aptly and adapt to station differences.

In the case of method 1, some of the coefficients, in particular for the transfer time and the alighting index, have opposite signs, such as at "Genève-Eaux-Vives", "Lancy-Pont-Rouge", or the above "Gare Cornavin", where the two former are small to medium-sized transfer stations, allowing for all modes of transport. This indicates that, although the framework of method 1 captures the correct interpretation of the coefficients at Lancy-Bachet, its applicability and trackability at this station are only a coincidence.

To gain further insight into framework transferability, the performance indicators of both Lancy-Bachet and Cornavin are shown in Table 22. As depicted, the use of the methods at Gare Cornavin led to, overall,

worse performance indicators. Though, this is to be expected. Due to the greater flow at Gare Cornavin, with the observed transfer flow going up to 200, and the overall number of observations (3500 VS 700) and choice sets being greater, the RMSE is also expected to increase. Furthermore, due to the increased number of observed zero-flows, the PCC is reduced.

Table 22: Method 1 and method 2 performance indicators (alighting model) - Lancy-Bachet and Gare Cornavin

	Lancy-Bachet		Gare Cornavin	
	Method 1	Method 2	Method 1	Method 2
τ	0	0	0	0
β	1	-	1	-
$\bar{\rho}^2$	0.441	0.528	0.486	0.564
MTE	-	0.389	-	0.403
RMSE	-	0.457	-	0.845
PCC (<5)	-	0.7	-	0.503
MTE (Agg)	0.293	0.271	0.353	0.334
RMSE (Agg)	1.425	1.181	3.241	2.837
PCC (<5, Agg)	0.751	0.741	0.525	0.611

On the other hand, the MTE is in a similar range, being around 0.4 for the non-aggregated method 2 and around 0.05 greater than when applied for Lancy-Bachet. This indicates that, although the absolute error is slightly greater, the relative error has not changed substantially, meaning that the methods still perform well.

Finally, as can be seen, method 2 outperforms method 1 in all aspects when applied to Gare Cornavin.

7 Model application

Utilizing the models, the transfer-flow matrix at the transfer station can be generated using the workflow described in Section 3.3.2 and in Section 3.4.2, for methods 1 and 2, respectively. In the case of this application, the two parameters that define the maximum transfer flow and transfer-destination flow threshold are $\phi_1 = 0.8$ and $\phi_2 = 0.5$, respectively. This means that, of the total boarding count of a line, a maximum of 80% can be transfer flow, while, of the alighting count at a downstream destination, a minimum of 50% is non-transfer flow.

Applying the calibrated models and the workflow showed that, for both methods 1 and 2 (morning peak alighting model), 38.5% of passengers alighting at Lancy-Bachet transfer. During the morning peak, there are 1054 alightings per hour, resulting in an exit flow of 654 passengers per hour and a total transfer flow of 400 passengers per hour. Furthermore, of the 1163 total boardings at Lancy-Bachet, 35% are transfers and 65% are starting their journey here.

7.1 Demand estimation

As mentioned, the transfer flow at Lancy-Bachet is estimated to be around 400 passengers per hour. Using Larsen and Marx (2012) and the calibration results, the 95% confidence interval on the number of passengers per hour can be constructed as $[-10.26\%, +14.35\%]$ for method 1 and $[-7.68\%, +12.81\%]$ for method 2. From these results, part of the total travel demand for the morning peak can be estimated in number of trips per hour.

The partial travel demand can directly be estimated from the number of transferring trips and the total alighting count at a transfer station using Equation 36.

$$\Gamma_s = A_s - \xi_s \tag{36}$$

Where Γ_s is the partial travel demand at stop s, A_s is the total number of alightings at stop s across all lines, and ξ_s is the total transfer flow at stop s.

In the case of Lancy-Bachet, there are about 400 arrivals that are transferring, and thus the demand is 654 trips/hour, for which the 95% confidence interval is $[596.58, 695.03]$ for method 1 and $[602.76, 684.73]$ for method 2.

Using the same principle at all destinations in the network and taking the sum of the results, the total travel demand of the public transportation network can be estimated. In addition, the confidence interval for the transfer flow of each station can be utilized to compute the confidence interval for the total demand.

In addition to the total demand, the average number of stages per trip passing through a stop s can be determined by dividing the number of arrivals at a transfer station by the partial demand, as shown in Equation 37. In the case of Lancy-Bachet, there are 1.4 stages per trip.

$$\text{SPT}_s = \frac{\Gamma_s + 2(A_s - \Gamma_s)}{A_s} = 2 - \frac{\Gamma_s}{A_s} \tag{37}$$

Where SPT_s is the number of stages per trip at stop s, A_s is the total number of alightings at stop s across all lines, and Γ_s is the partial travel demand at stop s.

7.2 Pedestrian flow visualization

The transfer station Lancy-Bachet consists of 4 main stops: the train station, stop A-F, which is a combination of the tram stops and stops for some buses, and stops G and H, which are for some additional bus lines. The internal transfer station pedestrian flow can be determined using the stop locations and the estimated transfer and transfer-destination flows. This is done by aggregating the transfer flow per location pair of the transfer station, resulting in the flow shown in Figure 18, which depicts the hourly stop-to-stop pedestrian flow during the morning peak, estimated using method 1 (alighting model).

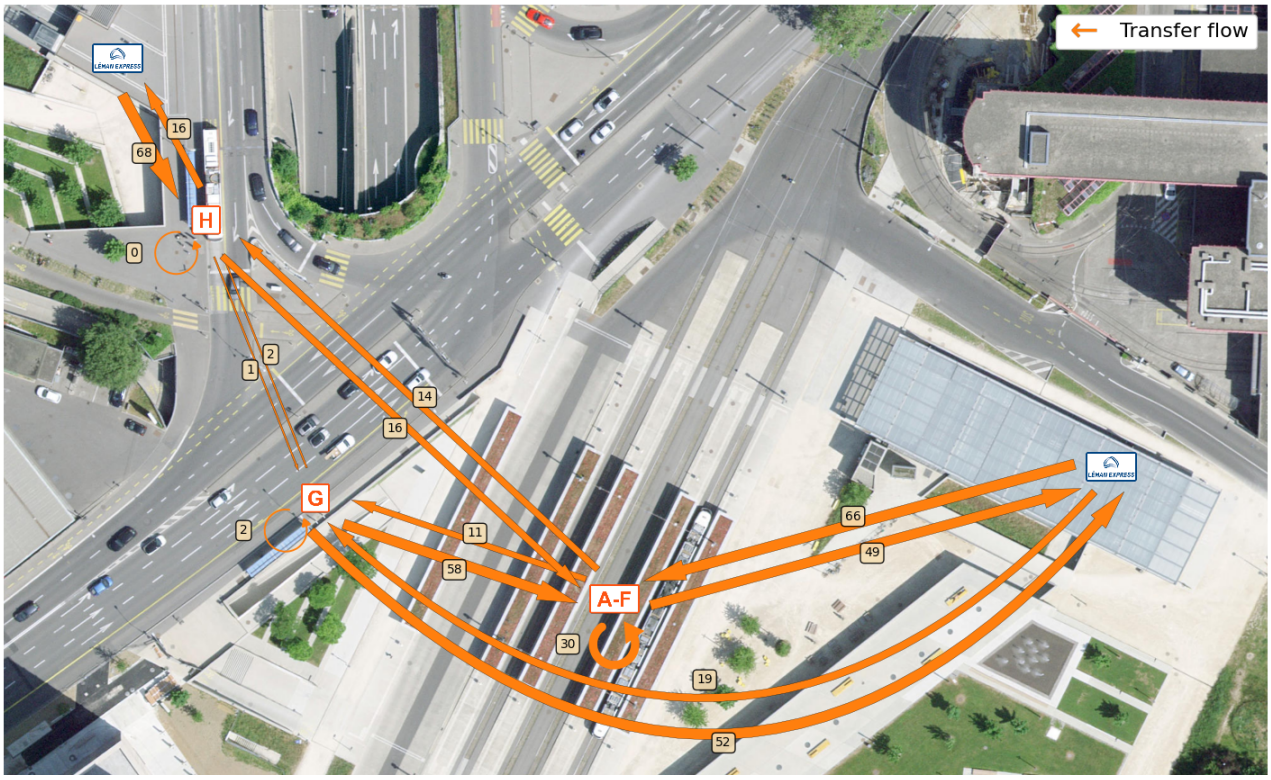


Figure 18: Expected hourly stop-to-stop pedestrian flow in the morning peak at Lancy-Bachet (method 1 - model 1)

As depicted, according to method 1 - model 1, the most pedestrian flow during the morning peak is from the train station to stop H, where line D_0 stops, and from the train station to stop A-F, where the trams stop, but in particular line 12_0. With this in mind, it must be noted that the most popular stop is A-F, with a total incoming flow of 170 pedestrians/hour. This follows from one of the most attractive transfer lines during the morning peak being line 12_0, with 136 transfer-boardings per hour, as well stop A-F also being where line 18 stops.

When utilizing method 2 - model 1, the results shown in Figure 19 are achieved, where it can be seen that the flow is very similar to the results of method 1. However, notable differences are that the flow to stop A-F is greater for all origin stops, including the same-stop transfer flow, and that the transfer flow to the train station is lower, irrespective of the origin stop. This indicates that transfers to trains are modeled as being less popular in method 2, and instead, some additional transfers are distributed elsewhere.

Finally, similar to method 1, stop A-F still attracts the most pedestrian transfer flow with 183 pedestrians/hour, indicating that both methods capture the attractiveness of line 12_0. However, for method 2, the number of transfers to line 12_0 is 130 passengers/hour, which is lower than for method 1. Instead, the transfers to lines 18_0 and 18_1 are higher.

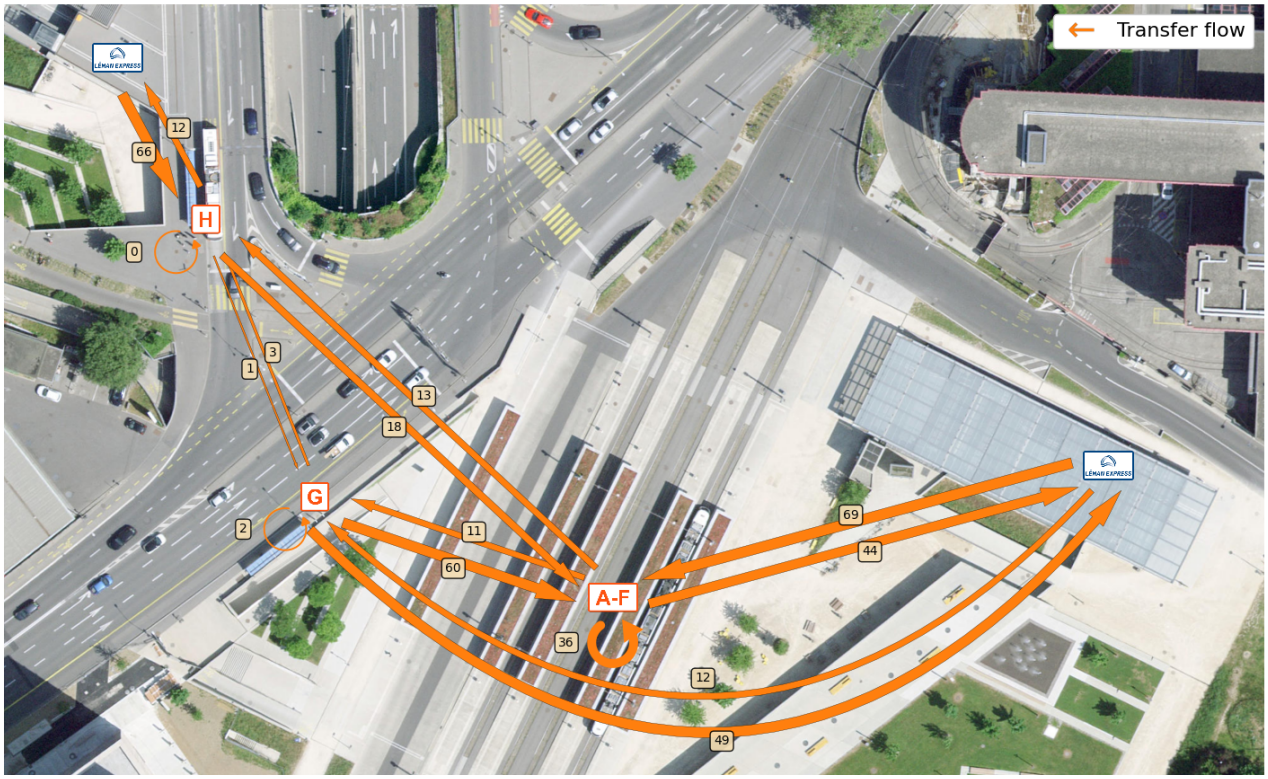


Figure 19: Expected hourly stop-to-stop pedestrian flow in the morning peak at Lancy-Bachet (method 2 - model 1)

7.3 OD matrix estimation

One of the main applications of the models is to function as an intermediary step in the estimation of the network-level OD matrix. In the following application, a partial network-level OD matrix is built by considering only the lines passing at Lancy-Bachet and singling it out as the only transfer station. The step-by-step workflow for both methods is described in Appendix C and MIPF was used where required (see Section 2.2.1). For the morning peak, this resulted in a total of 425 OD matrices, one at the route level for each line stopping at Lancy-Bachet and one for each feasible transfer combination, known as transfer-flow-only matrices (Cui, 2006). This complete assortment of OD matrices can then be combined into a single one, though that is not necessarily suitable for utilization.

As mentioned, a component of the network OD matrix are the transfer-flow-only matrices (an example of which can be seen in Appendix B). These generally consist of very low flows (< 0.01), as the 400 transferring passengers are distributed across $50'000$ possible origin-destination combinations.

Thus, to complement these shortcomings and highlight the most important connections, Cui (2006) recommends showing the demand lines between individual stops in the network. The results of which are shown in Figure 20 and Figure 21, for methods 1 and 2, respectively. Figure 20 shows the 6 connections with

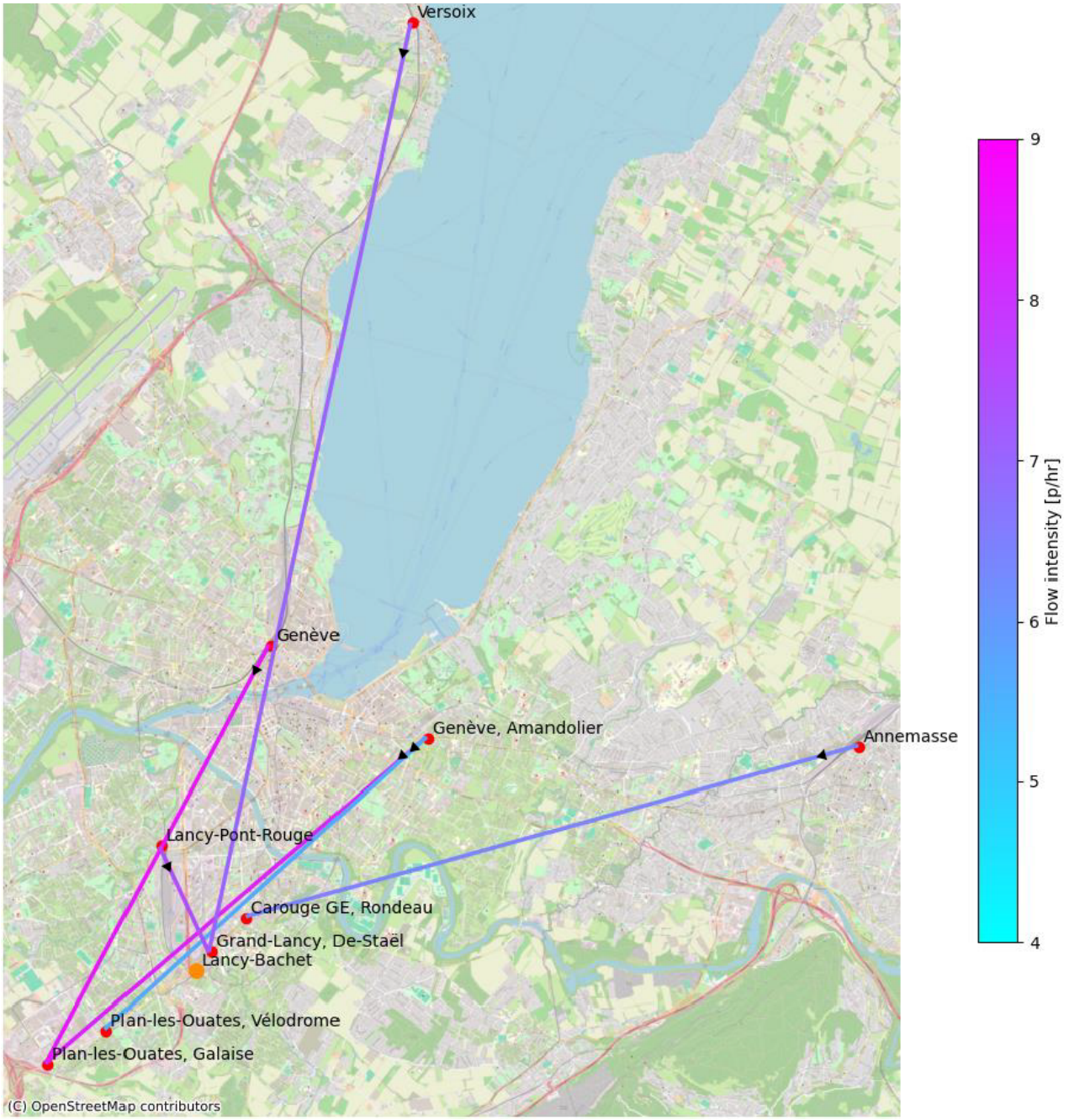


Figure 20: Top 6 stop-to-stop flows transferring at Lancy-Bachet during the morning peak using method 1 - alighting model

the greatest passenger-per-hour flow (greater than a flow of 5). As depicted, the largest flows, at about 8.5 passengers per hour, are from "Genève" (Gare Cornavin) and "Genève, Amandolier" to "Plans-Les-Ouates, Galaise", which are both larger stations in Geneva.

To compare, the results using method 2 show the 7 greatest flows, but this time, the minimum flow is 4. As depicted, the two greatest flows, at about 6 passengers per hour, go from "Genève, Amandolier" to "Plans-Les-Ouates, Galaise" and "Plans-Les-Ouates, Vélodrome". This result is quite dissimilar from that of method 1, even when one of the stops is the same. This indicates that, though both models perform similarly

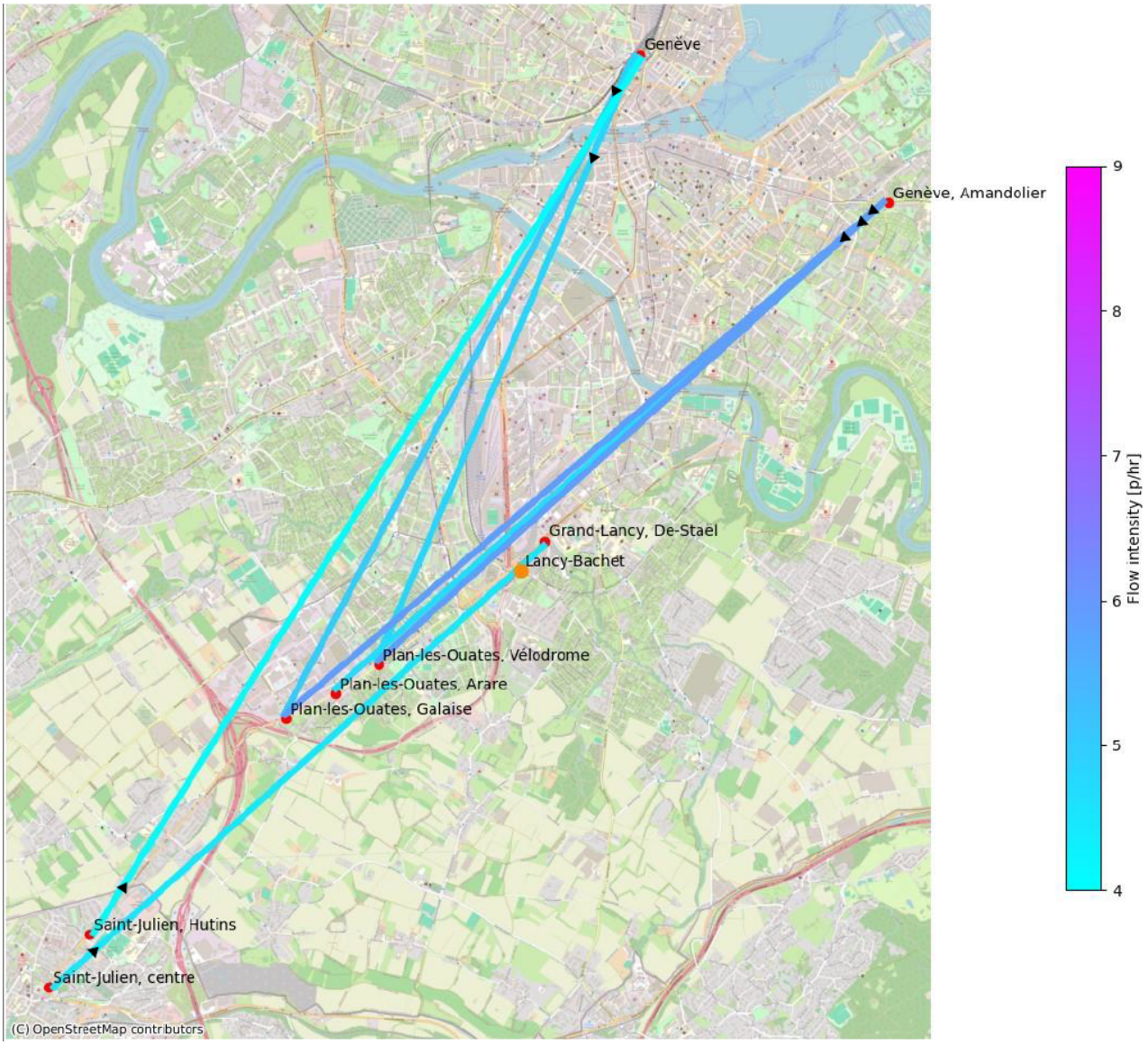


Figure 21: Top 7 stop-to-stop flows transferring at Lancy-Bachet during the morning peak using method 2 - alighting model

at the transfer flow estimation step, their differences propagate and lead to greater differences in the final OD matrix. More explicitly, these large differences are due to the differences in generating the modified route-level OD matrices. In the case of method 1, the route-level matrix for a given line has, in addition to the stops it serves, for both origin and destination, the feasible transfer-producing and transfer-receiving lines for the transfer station of interest. To then generate the final route-level OD matrix, the seed matrix from the transfer sample is inflated using (M)IPF so that the total counts match the modified boarding and alighting counts. Generally, this seed matrix is very sparse, in particular, the transfer-specific part has many structural zeros. On the other hand, for method 2, only the number of destinations increases with the number of feasible transfer-receiving lines, which leads to a lower number of zeros in the seed matrix.

8 Discussions, Conclusions, and Recommendations

8.1 Discussions

8.1.1 Low flow performance

One of the important points to discuss is mentioned multiple times throughout the results and present for all models: the dwindling prediction accuracy of the model for low transfer flows. As shown, the strength of the models lies in their capability to predict flows greater than 10, where their accuracy is well established. These flows consist primarily of the network exit flows and the high-demand flows at the transfer station.

Due to this lower accuracy in predicting low-demand flows, the applicability of the model is limited. One important consideration is the area of application—specifically, the city center versus the suburbs, where transfer stations in the latter are smaller and likely experience low traffic. This has two implications: first, the size of the transfer sample, and second, the accuracy of predicted flows. Given the reduced accuracy of low transfer flow predictions and the overall lower flow at suburban transfer stations, the number of relevant errors may exceed acceptable levels. Another issue may arise due to the size of the sample, where fewer individuals and a diversity of choices may inhibit adequate convergence of the model coefficients.

Finally, this issue of low-flow performance inhibits the usage of the models when the user desires to fully capture the picture, and not just focus on the high-transfer flow connections.

8.1.2 Model transferability

Another issue to address is the non-transferability of the model coefficients. As shown in section 6.1, it is not possible to use the coefficients calibrated at Lancy-Bachet at another transfer station of the network, as their respective regressions converge to values that are statistically significantly different from the ones for Lancy-Bachet. This results in a great loss of intended model functionality. Instead of being able to calibrate a model and using it to estimate the transfer flows/partial network OD matrices of other stops, a model has to be calibrated for every station in the network. This increases the burden on the necessary data quality and required computational time, in particular for large networks with many transfer stations. Furthermore, the models now fulfill only three primary roles. The first is as a data enrichment role, where they may adjust flows and complete the sample, i.e., set zero-flows to non-zero. The second is within predictive modeling, where the models may be utilized to reflect a change in transfer demand depending on changes in the parameters, such as line frequency or additions of new lines. And the third is to gain insight into passenger transfer flow characteristics and understand how different stations differ. Finally, though the models are shown to be non-transferable, it might be the case that, when greatly lacking sample data, a decision is made to apply the already regressed model coefficients to another location. However, in this case, a thorough investigation of the station build-up and sample transfer patterns must be made to assess the similarity between the stations. This would be a costly investigation and the gains achieved in flow accuracy may not be sufficient to warrant it.

8.1.3 Utilization of scaled flows

One of the methods utilized to investigate the applicability of the methods was to use scaled transfer flows as the sample data for the MNL model regression. As shown, though the results are very similar to those using the non-scaled sample, it was found that the coefficients tend to be more radicalized, i.e., the magnitude of the coefficient value increases. Furthermore, although the weights used to scale the flows were available in this case, this might not always be the case, as Gravity Modeling or another method might be utilized to estimate the route-level OD matrices. Another issue that should be mentioned is the zero-flows. Due to the nature of the scaling, these remain zero-flow, even though it might be the case that a transfer or more could have been observed, thus possibly skewing the model results and forcing the coefficients to be less representative.

8.1.4 Origin information

Tied into the model performance is the general lack of origin information considered in the models. Nonetheless aspects of the origin are considered in some of the explanatory variables, like the angle cost for both methods 1 and 2, and the "passed" and "direct connection ratio" (DCR) for method 2. However, in all three of these cases, only an aggregate of the possible passenger origins is considered. For instance, the angle cost is the average angle between all origin stops and the transfer station which defines the arrival angle. On the other hand, this aggregation may also be beneficial, such as in the case that a transfer connection with an obtuse angle is the fastest. In such an example, the angle cost between the transfer-producing and the transfer-receiving line might be low, but the last two stops before the transfer station, which would utilize this connection the most, might be at an obtuse angle to the departing angle of the transfer-receiving line. In a non-aggregated scenario, this obtuse angle would lead to a high angle cost, thus possibly reducing the expected transfer flow, even if this connection is highly desirable, due to its speed.

The "passed" variable represents whether or not the origin line has passed that stop, but it may have been passed very early on in the route, such that a transfer connection might still be a faster option. Finally, for the DCR, there might be cases where irrelevant upstream stops are given too much weight compared to stops closer to the transfer station, skewing the DCR in favor of stops that would have more impact on the actual transfer at the station.

A final point of discussion is the lack of consideration for multi-transfers, which is in part due to the lack of origin information. However, as mentioned, these only comprise a smaller part of the sample, at 4% of all passengers, thus, though neglected, their impact on the overall scheme is minor.

8.1.5 Further discussion of key results

Explanatory variables

To make use of the MNL models, a wide variety of explanatory variables were utilized and defined. These can be split into two main categories: (1) adapted from literature and (2) original ones. With regards to category (1), which is composed of the transfer time, the travel time, and the angle cost, it was shown that even though the utilized model and context differed, their expected response and contribution were met.

As a consequence, the notion that public transportation users value these aspects of the service provided and how they react to potential changes is reinforced.

Category (2) consists of the three indices, their zonal counterparts, the feasible transfer ratio, the direct connection ratio, and the four dummy attributes (origin mode, passed, and served downstream). In this case, some of them were shown to be statistically insignificant, such as the feasible transfer ratio, or only contributing minutely to the model fit, as shown through the coefficient dominance results and the P-value. On the other hand, the indices and the zonal counterparts were shown to generally be adequate representations of the attraction of a line or destination. However, it must be mentioned that method 2, with the zonal explanatory variable, performs better than its counterpart.

Finally, the direct connection ratio and the dummy variables were also shown to contribute to a certain extent. In particular reinforcing the notion that generally, public transport users tend to prefer not transferring when possible, and are rational in their decision-making.

Demand estimation
Highlighted as one of the applications of the methods, total travel demand estimation is a key result of the methodologies. To the knowledge of the author, though this is an implicit result of network-level OD matrix estimation, it has never been highlighted in previous research. In this case, the models allow to estimate not just the total demand, but also the confidence interval of the estimated demand. This enables the public transport planner to gain insight into the use of the system and more actively ensure desired service levels.

In addition to the demand, the demand estimation also allows the public transport provider to identify the number of stages per trip going through a transfer station. This would allow them to compare the stations in their network and identify where potential bottlenecks occur.

8.2 Summary and conclusions

During this research, two methodologies were devised to address network-level OD matrix estimation. Both require survey data, network build-up, and public transportation service data (schedules) to run two Multinomial Logit model regressions. The first MNL model, associated with method 1, aims to describe the probability of transferring to a given line, while the second one describes the probability of transferring and going to a destination on that line. To assess the suitability of the methods, they were applied to a transfer station of the public transportation network of Geneva, Lancy-Bachet, after which the framework was applied to additional stops to assess the transferability, among which is the station Gare Cornavin.

Overall, both methods perform well during the morning peak, describing the transfer flow and transfer destination flow with satisfactory accuracy. However, they both share the same issues, i.e. the flows that are greater than 15 are aptly described, whereas the lower flows are prone to larger relative errors. Furthermore, when being applied to Lancy-Bachet, method 2 generally outperforms method 1, though only by a small margin. Finally, using the available alighting data as an explanatory variable was shown to lead to a more accurate estimation of the flow, instead of using the workplace or population density as a proxy.

Beyond the usage of alighting data over employment or population density data, it was found that, of the designed explanatory variables, only a subset was found to be statistically significant for explaining transfers.

For method 1, this subset consists of the "Intercept", the "Transfer time", the "Angle cost", the "Alighting index", and the "From train" and "From tram" dummies. For method 2, this set is expanded with the "Travel time", the "Direct connection ratio", and the "Served downstream" dummy, and the "From train" dummy is removed.

When assessing the transferability of the models (coefficient and framework), it is shown that, for either method, it is not possible to use the coefficients calibrated at a first location for a second location, as the transfer station characteristics are too influential. Furthermore, method 1 was shown to have unstable coefficient values, with some flipping from negative to positive or vice-versa. This is a behavior not shared by method 2, indicating that only its framework is transferable.

Thus, to conclude and answer the three research questions:

Is the first devised methodology, which uses a Multinomial Logit model to estimate transfer proportions, applicable to estimate transfer flow totals for transfer stops within the bus-tram-train network of TPG, Geneva?
Although method 1 was applicable to Lancy-Bachet, it was shown through the model transferability assessment, that the framework and methodology are not useable at different locations in the Geneva public transportation network, as the model is no longer reliable. Thus, it can be concluded that method 1 is only suitable for use in the transfer station that was used for estimation.

Can the second devised methodology, which incorporates a Multinomial Logit model to infer destination, be effectively utilized to estimate (part of) the public transportation network OD matrix?
Similar to method 1 in the application at Lancy-Bachet, method 2 and its corresponding models were capable of estimating the transfer-destination flow (partial network OD matrix) accurately for flows greater than 15. This result, in conjunction with the capability of the framework to be applied in other locations of the network, results in the conclusion that this method is suitable to effectively estimate parts of the network-level OD matrix.

Which of the devised methodologies is recommended to be used during the estimation of the network-level OD matrix of the Geneva public transportation network?
Following from the results of the two prior questions, the conclusion is that the second method is recommended for use in the complete network-level OD matrix estimation of Geneva.

8.3 Recommendations

In the following, recommendations for future research are described, ordered by relevance in the view of the author.

Firstly, as remarked previously, due to data availability and time constraints, no real ground-truth comparison was able to be performed. Thus, the first recommendation is to find a location where the real total transfer and transfer-destination flows are available, then calibrate a model, and apply it to that location. Through this, a more concrete picture of the model's applicability can be drawn, and its flaws can also be more accurately established, interpreted, and rectified. In particular for the lower transfer flows, which make up the largest part of the flows, and appear to be somewhat random and thus inaccurately represented by the model.

In combination with the prior recommendation, in the case that the methods are deemed to be suitable for use, a second recommendation follows. This is to construct the complete network OD matrices with the results of both methods using the steps outlined in Appendix C. These should then be compared to a ground truth. Using this comparison, a final assessment of the applicability of the complete methodologies can be made.

A recommendation which follows from the discussion is to try to integrate origin information into the models. This would allow for easier fine-tuning and potentially higher predictive accuracy. To this end, the arrival flow, either scaled or from the sample, can be split based on the route-level OD matrix, enabling the addition of multiple new explanatory variables or disaggregation of previously established ones. Among the modified ones would be the angle cost, as well as the "passed"-dummy and "direct connection ratio", which would be able to provide a clearer picture of the route-choice part of the choice model. Potential new explanatory variables would be the travel time to the considered transfer station, or an expansion/addition to the disaggregated direct connection ratio, which could include whether or not an alternative single-transfer route exists. On the other hand, this further disaggregation would result in a great increase in the sparsity transfer-flow matrix and require even greater survey quality, which may be infeasible. Furthermore, the disaggregation further increases the complexity of multi-transfer trips, and further investigation is required to consider their inclusion or non-consideration.

A further recommendation is to investigate the impact of schedule reliability on the model sensitivity. As mentioned, the realized schedule was no longer available, although it may have had a direct impact on the transfer choice of an individual. For example, if a line with a shorter transfer and travel time arrives so far behind schedule, that an alternative line can be taken instead, a user may choose to board this alternative. This is an effect that is not represented in the models, and should thus be investigated. Furthermore, this would also allow for more direct passenger behavior analysis concerning reactive route flexibility.

Another potential direction for future research could involve using the models to estimate the transfer-to-non-transfer ratio for each line. These results could then help identify trip origins and destinations at the network level through boarding and alighting counts. The defined OD counts, combined with survey data, could be used to directly estimate the network-level OD matrix using methods like Maximum Likelihood Estimation or Iterative Proportional Fitting.

Finally, instead of using a Multinomial Logit Model, an alternative discrete choice model can be utilized. More specifically, the Multinomial Probit or Mixed Multinomial Logit Model could be utilized to include random utility variations in the choice alternative perception for each passenger, possibly allowing the model to more aptly describe bias and randomness of peoples choices.

9 References

Aguiléra, V., Allio, S., Benezech, V., Combes, F., & Milion, C. (2014). Using cell phone data to measure quality of service and passenger flows of paris transit system [Special Issue with Selected Papers from Transport Research Arena]. *Transportation Research Part C: Emerging Technologies*, *43*, 198–211. https://doi.org/https://doi.org/10.1016/j.trc.2013.11.007

Blum, J. J., Sridhar, A., & Mathew, T. V. (2010). Origin–destination matrix generation from boarding–alighting and household survey data. *Transportation Research Record*, *2183*(1), 1–8. https://doi.org/10.3141/2183-01

Blume, S. O., Corman, F., & Sansavini, G. (2022). Bayesian origin-destination estimation in networked transit systems using nodal in- and outflow counts. *Transportation Research Part B: Methodological*, *161*, 60–94. https://doi.org/https://doi.org/10.1016/j.trb.2022.04.006

Bohannon, R. W. (1997). Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants. *Age and Ageing*, *26*(1), 15–19. https://doi.org/10.1093/ageing/26.1.15

Bohannon, R. W., & Williams Andrews, A. (2011). Normal walking speed: A descriptive meta-analysis. *Physiotherapy*, *97*(3), 182–189. https://doi.org/https://doi.org/10.1016/j.physio.2010.12.004

Cascetta, E., Pagliara, F., & Papola, A. (2007). Alternative approaches to trip distribution modelling: A retrospective review and suggestions for combining different approaches. *Papers in Regional Science*, *86*(4), 597–620. https://doi.org/https://doi.org/10.1111/j.1435-5957.2007.00135.x

Celikoglu, H. B., & Cigizoglu, H. K. (2007). Public transportation trip flow modeling with generalized regression neural networks. *Advances in Engineering Software*, *38*(2), 71–79. https://doi.org/https://doi.org/10.1016/j.advengsoft.2006.08.003

Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016). Development of destination choice models for pedestrian travel. *Transportation Research Part A: Policy and Practice*, *94*, 255–265. https://doi.org/https://doi.org/10.1016/j.tra.2016.09.017

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd). Lawrence Erlbaum Associates Publishers.

Cui, A. (2006). *Bus passenger origin-destination matrix estimation using automated data collection systems* [Master's thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/37970

data.gouv.fr. (2022, January). *Découpage administratif communal français issu d'openstreetmap* [Accessed: 2023-02-27]. https://www.data.gouv.fr/fr/datasets/decoupage-administratif-communal-francais-issu-d-openstreetmap/

ÉTAT DE VAUD. (2023, December). *Population résidante permanente, par commune, au 31 décembre, vaud* [Accessed: 2023-02-27]. https://www.vd.ch/etat-droit-finances/statistique/portrait-du-canton-et-portrait-des-communes-vaudoises

Fabbiani, E., Nesmachnow, S., Toutouh, J., Tchernykh, A., Avetisyan, A., & Radchenko, G. (2018). Analysis of mobility patterns for public transportation and bus stops relocation. *Programming and Computer Software*, *44*(6), 508–525. https://doi.org/10.1134/S0361768819010031

Gkiotsalitis, K. (2022). *Public transport optimization*. Springer International Publishing. https://doi.org/10.1007/978-3-031-12444-0

Glen, S. (n.d.). "NRMSE" [Accessed: 2023-01-09]. https://www.statisticshowto.com/nrmse/

Guex, G., Loup, R., & Bavaud, F. (2023). Estimation of flow trajectories in a multi-lines transportation network. https://doi.org/10.21203/rs.3.rs-2637944/v1

Håkegård, J. E., Myrvoll, T. A., & Skoglund, T. R. (2018). Statistical modelling for estimation of od matrices for public transport using wi-fi and apc data. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 1005–1010. https://doi.org/10.1109/ITSC.2018.8570009

Huang, D., Yu, J., Shen, S., Li, Z., Zhao, L., & Gong, C. (2020). A method for bus od matrix estimation using multisource data. *Journal of Advanced Transportation*, *2020*, 5740521. https://doi.org/10.1155/2020/5740521

Hussain, E., Bhaskar, A., & Chung, E. (2021). Transit od matrix estimation using smartcard data: Recent developments and future research challenges. *Transportation Research Part C: Emerging Technologies*, *125*, 103044. https://doi.org/https://doi.org/10.1016/j.trc.2021.103044

Insee. (2024, February). *Base du dossier complet* [Accessed: 2023-02-27]. https://www.insee.fr/fr/statistiques/5359146

Ji, Y., Mishalani, R. G., & McCord, M. R. (2015). Transit passenger origin–destination flow estimation: Efficiently combining onboard survey and large automatic passenger count datasets [Big Data in Transportation and Traffic Engineering]. *Transportation Research Part C: Emerging Technologies*, *58*, 178–192. https://doi.org/https://doi.org/10.1016/j.trc.2015.04.021

Kong, C., Guo, T., & He, L. (2023). Research on od estimation of public transit passenger flow based on multi-source data. In W. Wang, J. Wu, X. Jiang, R. Li, & H. Zhang (Eds.), *Green transportation and low carbon mobility safety* (pp. 589–603). Springer Nature Singapore.

Kropko, J. (2007). *Choosing between multinomial logit and multinomial probit models for analysis of unordered choice data* [Master's thesis, The University of North Carolina at Chapel Hill]. https://www.proquest.com/openview/f3ece03f523bf4f69966f24fe3df2487/1?pq-origsite=gscholar&cbl=18750

Kumar, P., Khani, A., & Davis, G. A. (2019). Transit route origin–destination matrix estimation using compressed sensing. *Transportation Research Record*, *2673*(10), 164–174. https://doi.org/10.1177/0361198119845896

Kumar, P., Khani, A., & He, Q. (2018). A robust method for estimating transit passenger trajectories using automated data. *Transportation Research Part C: Emerging Technologies*, *95*, 731–747. https://doi.org/https://doi.org/10.1016/j.trc.2018.08.006

Larsen, R. J., & Marx, M. L. (2012). *An introduction to mathematical statistics and its applications* (5th). Pearson.

Lebrusán, I., & Toutouh, J. (2020). Assessing the environmental impact of car restrictions policies: Madrid central case. In S. Nesmachnow & L. Hernández Callejo (Eds.), *Smart cities* (pp. 9–24). Springer International Publishing.

Limtanakool, N., Dijst, M., & Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography*, *14*(5), 327–341. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2005.06.004

Liu, X., Van Hentenryck, P., & Zhao, X. (2021). Optimization models for estimating transit network origin–destination flows with big transit data. *Journal of Big Data Analytics in Transportation*, *3*(3), 247–262. https://doi.org/10.1007/s42421-021-00050-3

Lu, K., Liu, J., Zhou, X., & Han, B. (2021). A review of big data applications in urban transit systems. *IEEE Transactions on Intelligent Transportation Systems*, *22*(5), 2535–2552. https://doi.org/10.1109/TITS.2020.2973365

Luchman, J. N. (2014). Relative importance analysis with multicategory dependent variables:: An extension and review of best practices. *Organizational Research Methods*, *17*(4), 452–471. https://doi.org/10.1177/1094428114544509

Lunke, E. B., Fearnley, N., & Aarhaug, J. (2021). Public transport competitiveness vs. the car: Impact of relative journey time and service attributes. *Research in Transportation Economics*, *90*, 101098. https://doi.org/https://doi.org/10.1016/j.retrec.2021.101098

Luo & Wang. (2003). Measures of spatial accessibility to healthcare in a GIS environment: Synthesis and a case study in chicago region. *Environ Plann B Plann Des*, *30*(6), 865–884.

Luo, Z., Zhou, X., Su, Y., Wang, H., Yu, R., Zhou, S., Xu, E. G., & Xing, B. (2021). Environmental occurrence, fate, impact, and potential solution of tire microplastics: Similarities and differences with tire wear particles. *Science of The Total Environment*, *795*, 148902. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.148902

Mandl, C. E. (1980). Evaluation and optimization of urban public transportation networks. *European Journal of Operational Research*, *5*(6), 396–404. https://doi.org/https://doi.org/10.1016/0377-2217(80)90126-5

McClave, J., & Sincich, T. (2018). Simple linear regression. In *Statistics* (pp. 612–680). Pearson.

McCulloch, R. E., Polson, N. G., & Rossi, P. E. (2000). A bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics*, *99*(1), 173–193. https://doi.org/https://doi.org/10.1016/S0304-4076(00)00034-8

McFadden, D. (1977). Quantitative methods for analyzing travel behaviour of individuals: Some recent developments. *Cowles Foundation Discussion Papers*, (707). https://elischolar.library.yale.edu/cowles-discussion-paper-series/707

McFadden, D., & Domencich, T. (1975). *Urban travel demand: A behavioral analysis*. North-Holland Publishing Co. https://eml.berkeley.edu/~mcfadden/travel.html

McNally, M. G. (2008). The Four Step Model. *UC Irvine: Center for Activity System Analysis*. https://escholarship.org/uc/item/0r75311t

Mohammed, M., & Oke, J. (2023). Origin-destination inference in public transportation systems: A comprehensive review. *International Journal of Transportation Science and Technology*, *12*(1), 315–328. https://doi.org/https://doi.org/10.1016/j.ijtst.2022.03.002

Mokhtarian, P. L. (2016). Discrete choice models' ρ^2: A reintroduction to an old friend [Standalone technical contributions in choice modelling]. *Journal of Choice Modelling*, *21*, 60–65. https://doi.org/https://doi.org/10.1016/j.jocm.2016.02.001

Montufar, J., Arango, J., Porter, M., & Nakagawa, S. (2007). Pedestrians' normal walking speed and speed when crossing a street. *Transportation Research Record*, *2002*(1), 90–97. https://doi.org/10.3141/2002-12

National Geographic. (2019, September). *The environmental impacts of cars, explained* [Accessed: 2023-12-13]. https://www.nationalgeographic.com/environment/article/environmental-impact

OCSTAT. (2024). *Atlas statistique du canton de Genève et de la région transfrontalière* [Accessed: 2024-02-27]. https://statistique.ge.ch/atlas/index.php#c=home

opendata.swiss. (2023a, December). *Administrative units switzerland - inspire - reference system wgs84 web mercator - format application/x-gmz* [Accessed: 2024-02-27]. https://opendata.swiss/fr/dataset/administrative-units-switzerland-inspire/resource/f41b18af-2207-43fe-9b9e-c14e2509da7b

opendata.swiss. (2023b, December). *Decoupage en sous-secteurs statistiques (girec)* [Accessed: 2024-01-09]. https://opendata.swiss/de/dataset/decoupage-en-sous-secteurs-statistiques-girec

Ortúzar, J., & Willumsen, L. G. (2011). *Modelling transport* (4th). John wiley & sons.

Paternoster, R., BRAME, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for equality of regression coefficients. *Criminology*, *36*, 859–866. https://doi.org/10.1111/j.1745-9125.1998.tb01268.x

Placiakis, J. (2023, July). *Single-route od matrix estimation using automatic passenger counting data: A case study in geneva*. http://essay.utwente.nl/95834/

Prato, C. G. (2009). Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling*, *2*(1), 65–100. https://doi.org/https://doi.org/10.1016/S1755-5345(13)70005-8

Raveau, S., Muñoz, J. C., & de Grange, L. (2011). A topological route choice model for metro. *Transportation Research Part A: Policy and Practice*, *45*(2), 138–147. https://doi.org/https://doi.org/10.1016/j.tra.2010.12.004

République et Canton de Genève. (2023). *Statistiques cantonales* [Accessed: 2024-01-09]. https://statistique.ge.ch/domaines/01/01_01/tableaux.asp#20

Scherer, M. (2010). Is light rail more attractive to users than bus transit?: Arguments based on cognition and rational choice. *Transportation Research Record*, *2144*(1), 11–19. https://doi.org/10.3141/2144-02

Scherer, M., & Dziekan, K. (2012). Bus or rail: An approach to explain the psychological rail factor. *Journal of Public Transportation*, *15*(1), 75–93. https://doi.org/https://doi.org/10.5038/2375-0901.15.1.5

Shcherbakov, M. V., Brebels, A., Tyukov, A., Janovsky, T., & Anatol, V. (2013). A survey of forecast error measures. *World Applied Sciences Journal*, *24*, 171–176.

Sheffi, Y. (1985). *Urban transportation networks: Equilibrium analysis with mathematical programming methods*. Prentice-Hall, Inc. https://sheffi.mit.edu/book/urban-transportation-networks

SITG. (2024). *Le territoire genevois à la carte* [Accessed: 2024-02-27]. https://map.sitg.ch/app/

Statistique Vaud. (2023). *Atlas statistique du canton de vaud* [Accessed: 2023-02-27]. https://cartostat.vd.ch/#c=indicator&i=pop_natio.pop_ch_etr&s=2023&view=map1

Territoires, O. D. (2023). *Emploi et dynamique du marché du travail* [Accessed: 2024-02-27]. https://www.observatoire-des-territoires.gouv.fr/nombre-demplois-au-lieu-de-travail

Vanderwaart, C. (2016). *Planning transit networks with origin, destination, and interchange inference* [Master's thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/103820

Vanderwaart, C., Attanucci, J. P., & Salvucci, F. P. (2017). Applications of inferred origins, destinations, and interchanges in bus service planning. *Transportation Research Record*, *2652*(1), 70–77. https://doi.org/10.3141/2652-08

Wang, Y., Wu, B., Dong, Z., & Ye, X. (2016). A joint modeling analysis of passengers' intercity travel destination and mode choices in yangtze river delta megaregion of china. *Mathematical Problems in Engineering*, *2016*(1), 5293210.

Wong, K., Wong, S., Tong, C., Lam, W., Lo, H., Yang, H., & Lo, H. (2005). Estimation of origin-destination matrices for a multimodal public transit network. *Journal of Advanced Transportation*, *39*(2), 139–168. https://doi.org/https://doi.org/10.1002/atr.5670390203

World Population Review. (2023). *Geneva population 2023* [Accessed: 2023-12-13]. https://worldpopulationreview.com/world-cities/geneva-population

Zannat, K. E., & Choudhury, C. F. (2019). Emerging big data sources for public transport planning: A systematic review on current state of art and future research directions. *Journal of the Indian Institute of Science*, *99*(4), 601–619. https://doi.org/10.1007/s41745-019-00125-9

Zhang, Q. (2008). *Od flow estimation for a two-route bus transit network using apc data: Empirical application and investigation* [Master's thesis, The Ohio State University]. https://etd.ohiolink.edu/acprod/odb_etd/etd/r/1501/10?clear=10&p10_accession_num=osu1228331301

Zhu, J., & Ye, X. (2018). Development of destination choice model with pairwise district-level constants using taxi gps data. *Transportation Research Part C: Emerging Technologies*, *93*, 410–424. https://doi.org/https://doi.org/10.1016/j.trc.2018.06.016

A Data sources

Due to the extent of the study, many open-source population, workplace, and shape data had to be retrieved. The utilized sources are described in the following:

- France
  - Population: Insee (2024) - https://www.insee.fr/fr/statistiques/5359146
  - Workplace: Territoires (2023) - https://www.observatoire-des-territoires.gouv.fr/nombre-demplois-au-lieu-de-travail
  - Shape data: data.gouv.fr (2022) - https://www.data.gouv.fr/fr/datasets/decoupage-administratif-communal-francais-issu-d-openstreetmap/
- Canton of Vaud
  - Population: ÉTAT DE VAUD (2023) - https://www.vd.ch/etat-droit-finances/statistique/portrait-du-canton-et-portrait-des-communes-vaudoises
  - Workplace: Statistique Vaud (2023) - https://cartostat.vd.ch/#c=indicator&i=pop_natio.pop_ch_etr&s=2023&view=map1
  - Shape data: opendata.swiss (2023a) - https://opendata.swiss/fr/dataset/administrative-units-switzerland-inspire/resource/f41b18af-2207-43fe-9b9e-c14e2509da7b
- Canton of Geneva
  - Population: République et Canton de Genève (2023) - https://statistique.ge.ch/domaines/01/01_01/tableaux.asp#20
  - Workplace 1: OCSTAT (2024) - https://statistique.ge.ch/atlas/index.php#c=home
  - Workplace 2: SITG (2024) - https://map.sitg.ch/app/
  - Shape data: opendata.swiss (2023b) - https://opendata.swiss/de/dataset/decoupage-en-sous-secteurs-statistiques-girec

B Transfer-flow-only matrices

Table 23 and Table 24 show resulting transfer-flow-only matrices, for method 1 and method 2, respectively, which describe where the originating flow on line LEX $L1_1$ goes when transferring to line D_1.

Table 23: Transfer-flow-only matrix example (method 1) - Origin on line LEX $L1_1$ and destination on line D_1

	Carouge GE, Stade de Genève	Lancy-Pont-Rouge, gare/Etoile	Genève, Queue-d'Arve	Genève, Jonction	Genève, Palladium	Genève, Stand	Genève, Bel-Air
Evian-Les-Bains	0.011	0.01	0.016	0.023	0.008	0.017	0.011
Thonon-les-Bains	0.017	0.016	0.025	0.036	0.013	0.027	0.017
Perrignier	0.029	0.026	0.042	0.059	0.022	0.045	0.029
Bons-En-Chablais	0.02	0.018	0.029	0.04	0.015	0.031	0.02
Machilly	0.018	0.016	0.026	0.036	0.014	0.027	0.018
Annemasse	0.012	0.011	0.017	0.025	0.009	0.019	0.012
Chêne-Bourg	0.005	0.005	0.008	0.011	0.004	0.008	0.005
Genève-Eaux-Vives	0.019	0.017	0.027	0.038	0.014	0.029	0.019
Genève-Champel	0.023	0.021	0.033	0.047	0.017	0.036	0.023

Table 24: Transfer-flow-only matrix example (method 2) - Origin on line LEX $L1_1$ and destination on line D_1

	Carouge GE, Stade de Genève	Lancy-Pont-Rouge, gare/Etoile	Genève, Queue-d'Arve	Genève, Jonction	Genève, Palladium	Genève, Stand	Genève, Bel-Air
Evian-Les-Bains	0.02	0.003	0.011	0.018	0.007	0.01	0.01
Thonon-les-Bains	0.032	0.005	0.017	0.028	0.011	0.016	0.016
Perrignier	0.052	0.008	0.029	0.047	0.019	0.027	0.027
Bons-En-Chablais	0.036	0.006	0.02	0.032	0.013	0.019	0.018
Machilly	0.033	0.005	0.018	0.029	0.012	0.017	0.017
Annemasse	0.022	0.003	0.012	0.02	0.008	0.011	0.011
Chêne-Bourg	0.01	0.001	0.005	0.009	0.003	0.005	0.005
Genève-Eaux-Vives	0.034	0.005	0.018	0.03	0.012	0.017	0.017
Genève-Champel	0.042	0.006	0.023	0.037	0.015	0.021	0.021

C Network-level OD matrix workflow

C.1 Method 1

1. Identify all transfer stations in the network
2. Generate a MNL for all transfer stations
3. Estimate the transfer flows for all transfer stations
4. Generate modified seed matrices per line (see Table 25)
   - **A** Add all feasible transfer-producing and transfer-receiving lines as possible origins and destinations
   - **B** Extract travel sample information
   - **C** If desired or necessary, inflate structural zeros with a marginal value
5. Adapt boarding and alighting counts at transfer stations with transfer flows
6. Apply IPF to line-level seed matrices and generate final line-level OD matrices
7. Link legs per transfer station and line-line combination

Table 25: Example of modified route-level OD matrix for method 1 with two possible transfer lines (1 and 2) at stop B

		A	B			C	TOTAL
			Exit	To 1	To 2		
A		0	10	1	2	6	19
	Origin	0	0	0	0	12	12
B	From 1	0	0	0	0	2	2
	From 2	0	0	0	0	3	3
C		0	0	0	0	0	0
TOTAL		0	10	1	2	23	36

C.2 Method 2

1. Identify all transfer stations in the network
2. Generate a MNL for all transfer stations
3. Estimate the transfer-destination flows for all transfer stations
4. Generate modified seed matrices per line (see Table 27)
   - **A** Add all feasible transfer-receiving lines as possible destinations
   - **B** Extract travel sample information
   - **C** If desired or necessary, inflate structural zeros with a marginal value

5. Adapt boarding and alighting counts at transfer stations with transfer flows
6. Adapt alighting counts at downstream stops with transfer-destination flows
7. Apply IPF to line-level seed matrices and generate final line-level OD matrices
8. Link legs per transfer station and line-line combination

Table 27: Example of modified route-level OD matrix for method 2 with two possible transfer lines (1 and 2) at stop B

	A	B			C	TOTAL
		Exit	To 1	To 2		
A	0	10	1	2	6	19
B	0	0	0	0	12	12
C	0	0	0	0	0	0
TOTAL	0	10	1	2	18	31