Technical University of Munich

Chair of Transportation Systems Engineering

Master's Thesis

Route Choice Model Calibration for Multi-Modal Public Transport Network Using Visum's Headway-Based Assignment Procedure

Author:

Rajapakshe Mudiyanselage Thenuwan Thiranjaya Bandara Jayasinghe

Supervision:

Univ.-Prof. Dr. Constantinos Antoniou (TUM)

Dr. Andreas Rau (TUMCREATE)

Mentoring:

M.Sc. Moeid Quarashi (TUM)

Dr. Zain Ul Abedin (TUMCREATE)

May 28, 2020

Abstract

This research presents an automated calibration framework for macroscopic transit assignment models, focusing on calibrating the estimates of perceived journey time (PJT) parameters of the headway-based assignment. The calibration framework is based on a stochastic approximation algorithm, named Simultaneous Perturbation Stochastic Approximation (SPSA). SPSA makes the calibration computationally feasible for large scale networks as it requires only two objective function evaluations for gradient approximation. A modification is proposed to the algorithm such that at a given iteration, it chooses an objective function that produces the highest relative error difference with the perturbed parameter estimates. The measurements of the selected objective function are then used in the gradient approximation step. Since the selection of the objective function is made in a ‘dynamic’ fashion, this method is called as SPSA with dynamic objective function (SPSA-DOF). The proposed method and the standard implementation of SPSA with a single objective function (SPSA-SOF) are tested in a synthetic transit model and in a real-world transit network of Singapore. The calibrated parameter estimates with SPSA-DOF provide a better model fit than SPSA-SOF. For the Singapore network, SPSA-DOF can reduce the error of most of the simulation outputs related to passenger transfers at stops and passenger trips on different transit line routes. The final calibrated model of the Singapore network shows a better overall fit with the observed values in different aggregation levels. Notably, the model accurately represents total passenger boardings and passenger trips with zero transfers for bus and rail (LRT, MRT) modes. The total number of passenger transfers at transit stops also shows a better fit. Different transfer movement types, attached to a transit stop, also shows a good fit, but the results could be improved with more calibration runs. The overall results for the Singapore network show that, on average, passengers perceive transfer walk time nearly 3.6 times more than in-vehicle time, resulting in the average PJT for a public transport trip in Singapore to 44 minutes.

Acknowledgments

I would like to thank Dr. Andreas Rau, for providing the opportunity to work with TUMCRATE Rapid Road Transport (RRT) research team. This opportunity exposed me to new areas of the domain of urban transport planning. I would like to thank Prof. Dr. Constantinos Antoniou for supervising this thesis. His feedback greatly helped and motivated me to explore more and understand the concepts better.

I would like to thank Moeid Quarashi, my mentor from TUM, for helping me to understand the topic better and helping me to think in different directions. Next, I would like to thank Zain Ul Abedin, my mentor/colleague who helped me throughout my stay at TUMCREATE. His advice not only on the research but other areas in general, helped me to update as a person. Thank you for giving me the freedom to explore things and learn from mistakes.

I would also like to thank all the members of RRT research team for their support. Thank you, Swati and Xiaodong, for helping me to clarify issues related to transport modeling. Special thanks for Michael, Lucas, and Santiago – my friends from the master's program, for friendly chats, feedback on my presentations, and of course, for teaching me Rummikub.

To all my friends, thank you for your trust, generosity, and enlightening discussions on various topics.

To my parents, two sisters and brother. Thank you for helping me and having faith in me and for guiding me with the freedom to make decisions. Finally, a special thanks to my girlfriend Chathushka, for always checking on me and encouraging me.

This work was financially supported by the Singapore National Research Foundation (NRF) under its Campus for Research Excellence And Technological Enterprise (CREATE) program.

Declaration

I hereby confirm that this thesis was written independently by myself without the use of any sources beyond those cited, and all passages and ideas are taken from other sources are cited accordingly.

Singapore, May 28th, 2020

Table of Contents

List of Figures

List of Tables

1 Introduction

Transit systems play an important role in fulfilling the mobility needs of a large share of people, especially in dense urban areas. The implementation of a well-integrated transit system requires significant investments in terms of infrastructure and technology. Therefore, for sensible decision making, the use of transportation models is highly important as one can analyze the dynamics of transportation networks before making large investments on infrastructure. Transit assignment is one of the essential elements in transit modeling. In transit assignment, the passenger demand between origins and destinations are assigned via transit routes.

Passenger route choice in a transit network is a complex task because of the complexity of a transit network's supply and demand (Ortúzar & Willumsen, 2011). The complexity increases with the scale of the network. First, the nature of the transit system (supply) itself is complex, and it provides many options for a passenger to choose from to travel to his or her destination. Second, the passenger's decision-making process is complicated as passengers perceive the time components of a trip different to the actual time spent. Therefore, macroscopic transit assignment models (TAM) consider that passengers make travel decisions based on the perceived journey time (PJT). TAMs use a set of PJT parameters to model the behavior of the passengers. Different assignment outputs can be obtained by changing the coefficients of PJT parameters. The selection of the best parameter estimates that closely represent reality helps the decision-makers to make informed decisions. The process of selecting the best model parameters is called model calibration.

TAM calibration is an essential and challenging task. For a real-world transit network, transit assignment calibration can become challenging due to two reasons. The first reason is the lack of availability of the observed data. Traditionally, the data was collected using passenger surveys. Passenger surveys are a costly method to collect data. Therefore, passenger surveys cover only the sample of the passengers who use the network, and often the survey is carried out in peak periods at selected corridors. Therefore, the accuracy of the calibration is questionable (Zhu, Hu, & Huang, 2014). The development of automated data collection (ADC) methods (e.g., automated fare collection) made a technological leap forward. ADC methods made it possible to keep a record on all the passenger flows and made the observed data used for the calibration more accurate.

The second question, which is relevant to this thesis, is the methods used for calibrating the TAMs. Due to the complex interaction between model inputs and outputs, the transit assignment calibration problem cannot be expressed in an analytically soluble form. Therefore, calibration of TAMs is considered as optimization problems to minimize the error between observed values and simulated values. Traditionally, TAM calibration was done using a simple manual procedure (trial and error). The reason was the unavailability of data and the high cost of manual search techniques (Zhu et al., 2014). However, trial and error methods are tedious and cumbersome, and most importantly, it does not ensure finding the optimal set of values for the parameters (Parveen, Shalaby, & Wahba, 2007). Therefore, the recent studies conducted on this domain took advantage of a systematic calibration approach that relies on an optimization algorithm to reduce the error between simulated and observed values.

1.1 Motivation

The motivation for this thesis aligns with the recent focus of developing a systematic calibration approach for transit assignment calibration. Technologically advanced transit networks can passively collect valuable information about passenger movements in the transit network that can be used to make accurate TAMs. The public transport network in Singapore accounts for a higher share of smart card usage, which provides the opportunity to collect accurate data about passenger flows and represent in a transport model. These models, which show the ground truth, may not have the ability to forecast different scenarios (Liu, Zhou, & Rau, 2019) but can use as a source of observed data for TAM calibration. The calibration can be done systematically with the use of an optimization algorithm. This will ensure that the calibration can be done with minimum manual input and in a computationally efficient manner.

1.2 Goals and contributions

This thesis aims to develop an automated procedure for transit assignment calibration. The proposed method should be computationally feasible and should be able to perform with minimum manual interaction with the simulation model. The thesis contribution includes:

1. Develop an automated calibration methodology for PTV Visums' headway-based assignment procedure.

2. Apply the developed calibration framework on a real-world transit model of Singapore to calculate the PJT parameter estimates for the Singapore network.

1.3 Organization

The thesis is presented in six chapters. In chapter 2, the literature topics such as transit assignment, previous research work in transit assignment calibration, and stochastic approximation methods are discussed. Chapter 3 provides a detailed description of the proposed methodology to solve the transit assignment calibration problem. The methodology is then applied to a synthetic network and a real-world network of Singapore. The results of the calibration are presented in chapters 4 and 5. In chapter 6, the results are discussed, along with possible future research work.

2 Literature Review

This chapter reviews the literature about areas that are related to transit assignment model calibration, divided into six main sections. The first two sections describe transit assignment models (TAM), with a focus on headway-based assignment models and related passenger route choice methods. The third section describes data collection methods emphasizing different data collection methods and their usage in TAMs. The fourth section describes calibration in general and formulating a transit assignment calibration problem, describing its objectives, constraints, and parameters used for the calibration. The fifth section investigates the previous TAM calibration studies. The sixth and final section related to literature investigates stochastic approximation algorithms and their suitability for TAM calibration. The chapter ends with highlighting research gaps identified.

2.1 Transit assignment models

Travel demand modeling includes models that attempt to simulate passenger trips. Among the different modeling approaches available, the four-step model is one of the approaches which has been widely used in research. The fourth step of the four-step model, traffic assignment, is the procedure of loading trips to the network. In TAM, this is called in different names such as transit assignment, public transport assignment. In transit assignment, the passenger traffic between origins and destinations are assigned to the transit routes.

(Ortúzar & Willumsen, 2011) mentions that transit assignment, in general, is a difficult task than the private transport assignment as it involves more substantial computational requirements and simplified assumptions. Also, transit assignment is different from private transport assignment for many reasons. (Ortúzar & Willumsen, 2011) categorized the differences with six aspects:

1. From a transport supply point of view, the network for public transport is different. It has public transport services running between stops and link types (e.g., rail, road). Link capacity is related to the capacity of the transit unit (e.g., bus) and its frequency.
2. Transit route choice involves the movement of passengers. Passenger movement includes walking, waiting, and in-vehicle traveling. Passenger movement even can be further extended to shift between private transport and public transport (e.g., park and ride services). Therefore, passenger movement can get quite complex.

3. The fare structure of the public transport network has an impact on passenger route choice. Fare structure could get complicated with different types of tickets available for the passenger (e.g., fare vary with distance, flat fares, zonal fares, time-limited fares, seasonal tickets). The introduction of modern payment methods, such as smart cards or mobile payments, allows more complex fare structures.
4. The generalized cost function of passenger route choice involves time spent on different components of the journey (e.g., in-vehicle time, wait times, walk times, access/egress times, transfer penalties) each time components has a different weight. For example, a passenger might not perceive the minute spent on waiting as same as a minute spent on walking. Therefore, in transit assignment modeling, these different weights must be chosen carefully in order to reflect passenger behavior (refer section 2.2.2 for detailed information).
5. An interesting phenomenon called *'common lines problem'* arises when there are overlapping routes that run on common segments that share the same transit stops. Therefore, the passenger route choice may get complex, as the passengers may choose a 'set of paths' and let the transit mode that arrives first decide which of the paths the passengers will use.
6. The transit assignment method also varies with the frequency of the transit lines. If the transit lines in the network have relatively high frequencies, passengers do not tend to use or memorize the timetable. In contrast, when the frequencies are lower, passengers tend to check or memorize the timetable and arrive at the stop a few minutes before the departure. These two behaviors of the passengers have an impact on estimating the waiting times in transit modeling and require different assignment methods.

Based on the requirements, the model formulation can be different, which allows for the development of different models for transit assignment. One such difference stems from the question of whether passengers consider timetables when making their trips. Based on this question, two assignment model types can be formulated: 1. Headway-based (frequency-based) and 2. Timetable-based (schedule based).

2.2 Headway-based and timetable-based assignment

When the frequency on a transit service is high (e.g., urban metro service passing every 5 mins), the passengers tend to perceive the service in terms of headway between two subsequent departures. For such networks, the headway-based assignment method is suitable. On the other hand, if the arrival of the transit service is infrequent (e.g., intercity train service passing every 1 hour), passengers tend to perceive the service in terms of runs (e.g., departure at 11:00 am). For these kinds of networks, a timetable-based assignment is more suitable.

Each assignment procedure has its capabilities. Thus, the selection of the variant should be based on the actual need of the model design task and available data. Computational time should also be considered. Some of the advantages and disadvantages of each variant are summarized in Table 2-1.

Table 2-1 Comparison of headway-based and timetable-based assignment

Assignment variant	Advantages	Disadvantages
Headway-based	▪ Ideal for urban networks with short headways ▪ Can determine average loads on the lines ▪ Computationally efficient	▪ Lacks detailed model outputs (compared to timetable-based) ▪ Not suitable for transit networks with large headways
Timetable-based	▪ Can determine passenger loads of each run of the service ▪ Ideal for networks with long headways ▪ Coordination of timetable is possible	▪ Requires more detailed inputs ▪ Computationally expensive (compared to headway based)

Based on the comparison in Table 2-1, the headway-based transit assignment method is more suitable for an urban setting with shorter headways. The details of the headway-based transit assignment are discussed in the next subsections.

2.2.1 Headway based transit assignment

The headway-based assignment assumes that the passenger knows only the travel time and headway of a transit line route. The modeling principles remain the same, but the implementation may differ from simulator to simulator. The details described in this section are extracted from (PTV, 2019).

The headway-based assignment has three operational steps: 1. Headway calculation, 2. Route search and route choice, 3. Route loading. The first step defines the headway of a transit line route. PTV Visum provides three headway calculation methods. A brief description of each method is given below.

- *Constant from time profile attribute* – This is the simplest case used when the departure information is not essential due to dense headways of the network or timetable information is not available.
- *From mean headway according to the timetable* –The headway is calculated based on the number of departures for a given time slot. This approximation is acceptable if the network has shorter headways and sufficiently broad time intervals.
- *From mean wait time according to the timetable* – Here, the headway is calculated as double the expected wait time for the next departure of the line route.

In the second step, the possible routes between an OD pair are detected (route search) and allocated (route choice). The paths are assessed by their impedance, calculated based on passengers' perceived journey time (PJT). Moreover, the choice of boarding to a line and transfer between lines is based on the additional information a passenger has about the network. More about PJT and boarding decisions are discussed in 2.2.2 and 2.2.3, respectively. In the final step, the routes found on the second step are loaded from the OD matrix.

2.2.2 Perceived journey time

For various reasons, passengers perceive travel time different from the actual time spent. Many researches were done in this area and found evidence to support this statement. For example, (Beirão & Sarsfield Cabral, 2007) conducted a qualitative study to understand how passengers perceive public transport services in Portugal. One finding shows that passengers perceived waiting time too long and mentioned it as a barrier to using public transport. Similarly, passengers perceive the time spent on other components of public transport travel (e.g., walking time, boarding time) differently. (Meng, Rau, & Mahardhika, 2018) summarized findings of passengers' travel time perception based on the studies conducted in many countries. Therefore, in transport assignment modeling, passengers make travel decisions based on PJT and not based on actual time spent on the network.

PJT can be represented as a sum of below travel time components.

$$\mathrm{PJT} = \mathrm{c_{IVT}} \cdot \mathrm{IVT} + \mathrm{c_{AT}} \cdot \mathrm{AT} + \mathrm{c_{OWT}} \cdot \mathrm{OWT} + \mathrm{c_{WT}} \cdot \mathrm{WT} + \mathrm{c_{TWT}} \cdot \mathrm{TWT} + \mathrm{c_{ET}} \cdot \mathrm{ET} + \mathrm{TP} \cdot N\mathrm{oT} \tag{2.1}$$

Where:

IVT	In-vehicle time
AT	Access time
OWT	Origin wait time
WT	Walk time (transfer walk time)
TWT	Transfer wait time
ET	Egress time
NoT	Number of transfers
TP	Transfer penalty (minutes)
c	Coefficients/estimate for each travel time component

The estimate for each travel time component describes the perceived weight. Usually, c_{IVT} is set to 1 so that the perceived weight in other travel time components can be expressed as multiples of in-vehicle time. This representation is mainly used when onboard congestion is not considered in transit assignment modeling.

2.2.3 Passenger boarding strategies

In a transit network, we assume that a passenger behaves in a way that minimizes passenger's expected travel costs. In an urban transit network, passengers have many choices available to travel from origin to destination. Therefore, instead of relying on a specific path (e.g., shortest path), a passenger should follow a strategy. For example, it could be better to get onboard a slower line that is arriving first, which provides a direct connection to a destination than get on board a line that takes multiple transfers to reach the destination (this connection may have longer PJT). In this context, a strategy is a set of rules which allows the passenger to reach the destination starting from any node in the network (Spiess & Florian, 1989).

The strategy which the passenger chooses mostly depends on the information available to the passenger during the trip. For example, if a transit stop has dynamically updated information about the arrival times, passengers can make more informed decisions compared to a situation where a passenger has no such information at a stop.

One of the most common strategies used in transit assignment modeling is the ‘optimal strategies’ proposed by (Spiess & Florian, 1989). In optimal strategy, a passenger may choose to board the first arriving vehicle from a given line set, rather than waiting for a particular transit line. The journey will follow along different routes depending on the line which arrives first. For transit assignment modeling, the optimal strategy approach is suitable for networks with low headways. For a large transit network, the application of optimal strategy requires less computational time.

2.3 Data collection methods for transit assignment

Models require reliable data. Data about the transport supply and demand is essential for public transport modeling. The data collection methods can be divided into two main categories: 1. Manual-based methods, 2. Automated-based methods.

2.3.1 Manual data collection

Traditionally, manual data collection methods used to collect data regarding transit operations. (Ceder, 2007) identifies five main categories of manual transit data collection techniques, namely, 1. Point check, 2. Ride check, 3. Deadhead check, 4. Passenger survey, and 5. Population survey.

In a passenger survey, information related to passenger’s trip (e.g., origin-destination, access and egress modes, trip purpose) are collected. This information, in aggregation, can use to estimate the share for each transit route and to estimate the OD matrix. However, passenger surveys do not capture the movements of all the passengers of the network. Therefore, modelers need to apply expansion factors to estimate the OD matrix.

2.3.2 Automated data collection

In recent years, with the development of technology, automated data collection (ADC) techniques have become popular with the emergence of intelligent transport systems (ITS). The data generated from ITS enabled transit systems are a rich source for transit assignment modeling. (Gentile & Nökel, 2016) has identified four ITS systems, which so far has been used as a source of data collection for TAMs. A brief description of each method is given below:

- **Automatic Vehicle Location (AVL)** – AVL systems can track vehicles' positions. Therefore, this technology is used for security purposes and assess the real-time performance of a fleet against the schedule. For transit assignment modeling, data generated from AVL can be used to determine headways and calculate reliability indicators.
- **Automatic Passenger Counter (APC)** – APC systems provide counts of boarding and alighting passengers. This data can be used to calibrate and validate TAMs.
- **Traveler Information Systems (TIS)** – This technology is also known as Advanced Traveller Information Systems (ATIS). TIS includes journey planners and real-time information systems that are publicly available so that the passengers can make better-informed decisions to plan their trips.
- **Automated Fare Collection (AFC)** – The primary usage of this technology manages fare calculation and collection for passenger trips in a transit network. AFC omits the need for purchasing tickets for the passenger. Passengers can use a smart card or a bank card and travel through the network by simply tapping in and tapping out the card. Besides this primary task, AFC systems can keep track of valuable information on passenger behavior (e.g., origin and destination of the trips, transit line used, time of the day, and duration of the trip), which can be used for transit assignment modeling.

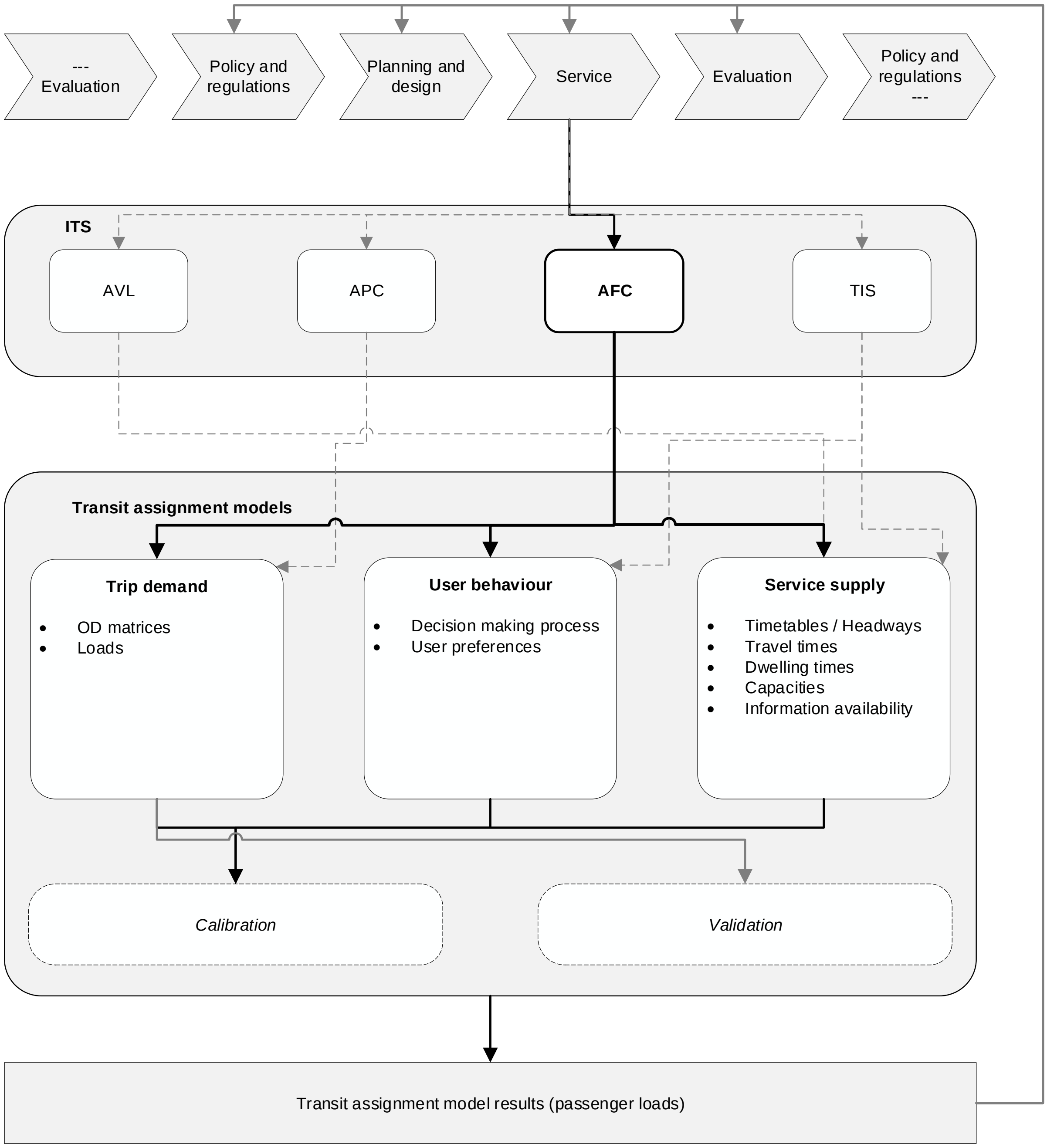

Figure 2-1 Use of ITS for transit assignment modeling (Source: Gentile & Nökel, 2016, p. 265)

Figure 2-1 shows how each ITS data collection system can assist in transit assignment modeling. The figure highlights the benefits of AFC for transit assignment modeling over the other data collection methods as it is capable of feeding data for trip demand estimation, understand user behavior and aspects related to transit supply. If the passengers have a higher share of usage of smart cards, the quality of the AFC data will be richer, and the accuracy of the TAMs will be higher. This rich data source can also be used in TAM calibration (Zhu, Hu, & Huang, 2014).

2.3.3 Comparison

Manual data collection methods do not provide complete data sets of a transit network. Moreover, the data collection cost is higher, and the validity of the data is uncertain. In contrast, automated data collection methods passively collect large volumes of data, which consist of different types of data about passenger behavior. Therefore, data collected from automated data collection methods are a useful resource for transit assignment modeling and calibration.

2.4 Calibration

Model calibration is a process of adjusting model parameters so that the model can be in close representation of the real system (Olstam & Tapani, 2011). Generally, model calibration is an iterative process so that the agreement between simulation outputs and the observed data (ground truth) can be improved in each iteration by learning from the previous iterations.

In relation to transportation modeling, model calibration follows the same steps mentioned above. However, based on the application, transport model calibration can take two main approaches: 1. Offline calibration, and 2. Online calibration.

Offline calibration models attempt to calibrate traffic conditions observed over multiple days, which are based on large archived databases of prior observations. Offline calibration models do not have specific computational requirements. On the other hand, online calibration is used when the objective of the calibration task requires to replicate short term variation of the demand. Therefore, in online calibration, the offline calibrated parameters need to be further calibrated in an ‘on the fly’ fashion.

TAMs are generally used for planning aspects. Therefore, TAMs require data collected over multiple days in order to represent the general behavior of the passengers. Therefore, an offline calibration approach suits the best for transit assignment calibration.

2.4.1 Transit assignment calibration

Objectives

The objective of transit assignment calibration is to closely represent the simulation model to the real-world data (ground truth). Due to the complex interaction between model parameters and simulation outputs, TAMs cannot be expressed in an analytically solvable form. Therefore, calibration of a TAM is considered as an optimization problem with the objective of minimizing an error between simulated and observed values.

More formally, the calibration problem can be represented as follows:

Let $\theta = (p_1, \cdots, \mathrm{p_n})$ be the parameters of the transit assignment model M, which needs to be calibrated. Let S be the simulated output of M for a given set of θ and R be the observed value for the equivalent simulated output. The difference between R and S can be calculated as an error which is represented in $f_E(R, S)$. The objective of the transit assignment calibration is to find optimal parameter estimates $\theta^\star$ such that:

$$\theta^\star = arg\,min\, f_E(R, S) \qquad (2.2)$$

$$subject\ to\ \ \theta \in \Theta$$

Where Θ is the domain of allowable values for θ.

The calibration results should be valid in different aggregation levels of the observed data. For example, a more aggregated level, the calibrated model should accurately represent the average journey times, the total number of passenger transfers of the observed network. In a less aggregated level, for example, the simulated model should accurately represent the passenger flows on each transit line route, and total passenger transfers made at each stop in the network.

As described in 2.3.2, the data collected using automated methods have more information on passenger trips compared to survey-based methods. Therefore, calibrating a TAM with smart card data provides the freedom to use different types of observed data in the calibration. At the same time, it demands more accuracy in the calibrated model such that all relevant types of observed data should be accurately matched. Moreover, the methodology used for calibration should be automated as much as possible. In this way, the calibration procedure can be applied with less manual work to a different network or the same network with updated data.

Calibration parameters

The coefficients of the PJT parameters explained in equation (2.1), including the transfer penalty (TP), are the dimensions used in the calibration. Based on the model setup, some variables may not require calibrating. For example, if the access links and egress links are not represented in the transit network, the parameters associated with those links (i.e., access time, egress time) do not need to be calibrated.

Constraints

All the parameter coefficients should hold non-negative values. More specifically, the wait time and walk time coefficients should hold a value greater than or equal to one, so that the concept of PJT remain valid. The range of values allowed for each parameter should be based on the limits given by the simulation software. If required, these limits can be further reduced by adding conditions.

2.5 Approaches to solving transit assignment calibration

There have been several approaches used to address the transit assignment calibration problem. The significant difference between these approaches is the way that the error minimization done described in equation (2.2). Based on this difference, two major categories can be identified: 1) Trial and error, and 2) Algorithm-based.

2.5.1 Trial and error

In this approach, the error minimization is done based on the knowledge and experience of the modeler. Therefore, the choice of the model parameter estimates is not entirely random. Based on the simulation results of the initial estimate, judgment is used to slightly modify the parameters until a reasonable model calibration is done. However, this approach is tedious and hard to replicate if the calibration needs to be done for another TAM.

(Farrol & Livshits, 1998), used a trial and error approach to solve the transit assignment calibration problem for Toronto Transit Commission, Canada. The data collected in 1996 from the ‘Transportation Tomorrow Survey’ provided the opportunity to calibrate the model parameters. The survey covered 5% of the household members of the greater Toronto area. Expansion factors were applied based on the census data to create the OD matrix for the transit assignment.

The calibration was conducted only for the morning peak by testing different sets of parameter coefficients to match the ridership statistics of the model output and the travel survey. These statistics include 1) volumes on key subway links, 2) total ridership on each mode (subway, bus, streetcar), 3) average number of transit boarding per trip, 4) ridership on each transit line. The initial parameters were taken from the 1986 model. The parameter coefficients were manually adjusted after each run to match the simulated values with the observed values.

(Fung, 2005) conducted a manual transit assignment calibration and validation for the metro network of Hong Kong. Three different models (1. headway-based assignment model, 2. dynamic schedule-based assignment model, and 3. Multi interval, multiclass schedule-based model) was used for the calibration. Passenger counts at line segments and passenger boarding and alighting at all station platforms were used as observed values, which were derived from AFC data. EMME/2 was used to run the simulations.

Four parameters were used for the calibration, and each parameter was allowed to take values from pre-defined boundaries. In this study, wait-time and walk-time coefficients could vary between 1.0-1.5 and 1.0-2.0 with 0.5 increments. Boarding penalty value could take values between 0 min and 4 mins with an increment of 1min. The in-vehicle time coefficient was fixed at 1 throughout the calibration. In addition, to limit the search space, the possible coefficient combinations were further limited by introducing a rule-based on logical passenger behavior (in-vehicle coefficient ≤ wait time coefficient ≤ walk time coefficient). For the headway-based assignment model, all possible combination of coefficients which are in align with criteria as mentioned above, was checked. The coefficient sets which provided the least absolute error were analyzed further by comparing link flow errors to obtain the best set of coefficients.

For the other two models (2 and 3), different sets of weights were checked, and the error between the simulated and observed values was taken from the passenger boarding and alighting counts. The sets of coefficients which provided the least RMSE (root mean square error) were checked for further analysis.

2.5.2 Algorithm-based

In this approach, stochastic approximation methods are used to minimize the error between observed and simulated values. Details about stochastic approximation methods are discussed in 2.6.

(Parveen, Shalaby, & Wahba, 2007) used an automated transit assignment calibration for Toronto Transit Network, Ontario, Canada. This study is an extension to work conducted by (Farrol & Livshits, 1998). This study is the first attempt for an algorithm-based systematic calibration approach for transit assignment. An overview of the automated calibration procedure used in this study is shown in Figure 2-2

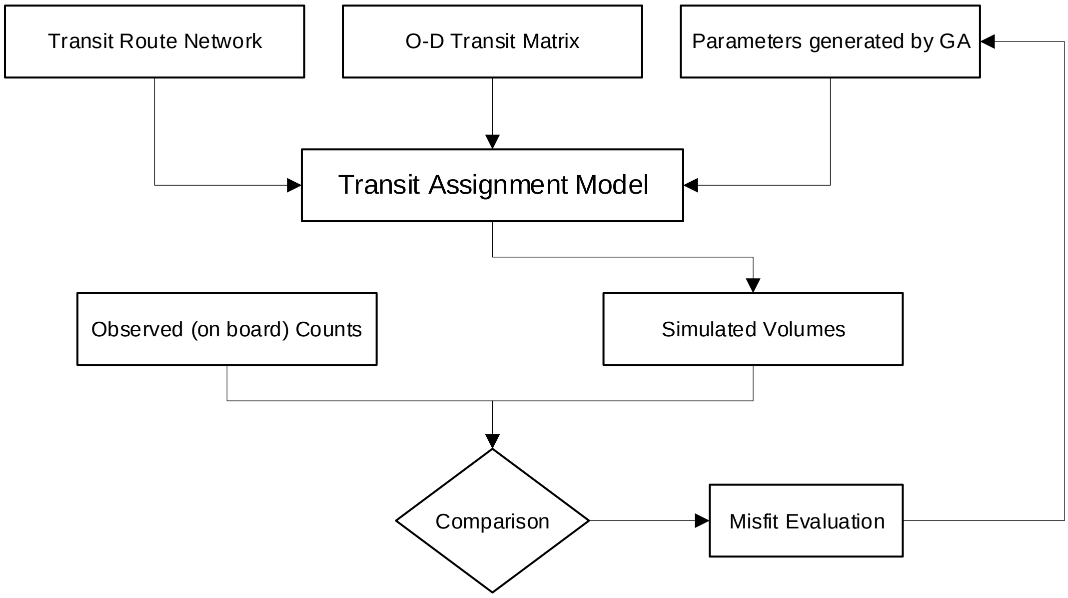

Figure 2-2 Automated calibration procedure for transit assignment (Source: Praveen et al., 2007)

This study also used the data from the Transportation Tomorrow Survey conducted in 1996. Transit assignment with 'optimal strategies' was done using EMME/2 as the simulator with an OD matrix created for morning peak hour (06:00 am – 08:59 am). The model had 463 zones, 241 transit routes, and 7394 stops.

The genetic algorithm was used as the optimization algorithm. The difference between simulated and observed link counts was used in the objective function. This study conducted experiments with two different error calculation methods 1.GRE (global relative error) and 2. MPRE (mean point relative error). The in-vehicle time parameter estimate was fixed at 1, and the rest of the parameters were bounded to limit the search space. Judgment was used to choose the final set of parameters in case if there were several solutions with similar objective function value (fitness value).

(Rydergren, 2013) also used an automated transit assignment calibration approach for the network of Stockholm. The network consisted of 1394 zones and 942 transit lines. These transit lines included bus, subway, ferry, and commuter train. The model did not include vehicle capacities or transit fares.

The observed values were based on two sources of data. The first source is from the Swedish national travel survey, based on 27000 telephone interviews, which was the complete data available for the region. The trips made on weekdays between 06:00 am, and 09:00 am were extracted (606 records). The second source was the data generated from transit trip planning software called OpenTripPlanner.

SPSA (Simultaneous Perturbation Stochastic Approximation) and Compass search algorithms were used as optimization algorithms. The hyperparameter values for each algorithm were found based on initial numerical tests. Transit assignment was done using PTV Visum simulator with four different model variants, where two models were in the class of 'optimal strategies', and two were in 'random departure' models.

Two objective functions were used in the study. One measure is a similar measure used in (Parveen et al., 2007), but based on the travel time of routes, named $\mathrm{MPR}E_T$ (T for time). The other measure is called $\mathrm{MIR}E_T$(model interval relative error).

It was found that random departure time model variants produced a better fit to the observed trips compared to the model variants in the class of 'optimal strategies'. Hence it was concluded in this study that more focus should be on finding the best model variant than finding the best set of coefficients for the parameters. It was also found that the performance of SPSA algorithm is efficient for $\mathrm{MPR}E_T$ and requires very less objective function evaluations. In contrast, the performance of the compass search algorithm was better with $\mathrm{MIR}E_T$, yet required a much larger number of objective function evaluations.

(Zhu et al., 2014) conducted a genetic algorithm-based approach to calibrate Urban Rail Transit (URT) assignment model using AFC data. A simplified test network was used for the study due to the computational barriers.

The objective function used for this study is somewhat different compared to other studies. The difference in the travel time distribution of the observed values and simulated values were used as the objective function. Three parameters were calibrated in this study. Namely, 1) standard deviation of normal distribution, 2) relative threshold of travel cost difference, and 3) absolute threshold of travel cost difference. It was found that the results from the calibrated model matched the observed values more closely compared to the parameters derived from manual travel surveys.

(Tavassoli, Mesbah, & Hickman, 2019), used the South-East Queensland network in Australia consists of bus, rail, and ferry modes. The trips made during am peak period (7:00 am – 9:00 am) used for the calibration. The simulations were done suing EMME/4 simulator with a headway-based assignment using 'optimal strategy' as passenger boarding strategy. This network also, like previous studies, did not consider passenger congestion on the transit network.

The observed values were based on AFC data (GoCard), which had around 82% of customer penetration during the period in which the study was conducted. This rich data set was used to estimate the OD matrix. Also, it helped to formulate an objective function with two terms. In the first term, %RMSE (percent root mean square error) was used to measure the difference between observed and simulated passenger flow on a given segment for a given mode. The second term, MAPE (mean absolute percentage error), was used to measure the difference between observed and simulated passenger trips in a given mode. The relative weight between these two error terms was fixed as 1 for this study.

In the calibrated model, higher dispersions were observed for the bus mode compared to rail and ferry modes. One reason for this could be the fact that bus passengers have more path choices compared to other modes. Also, another reason could be that the passengers are not necessarily choosing the optimal strategy boarding strategy approach.

A summary of all the previous studies mentioned above is summarized in Table 2-2.

Table 2-2 Summary of previous transit assignment calibration work

Study	Calibration procedure	Transit data used for the objective function	Calibrated parameters
Farrol & Livshits, 1998	Trial and error	▪ Volumes on key subway links ▪ Total ridership on transit modes ▪ Avg. no. of transit boarding per trip ▪ Ridership on each transit line	▪ access time ▪ wait time ▪ transfer penalty
Fung, 2005	Trial and error	▪ Link flow errors ▪ Passenger boarding and alighting counts	▪ in-vehicle time ▪ wait time ▪ walk time ▪ boarding penalty
Parveen et al., 2007	Automated - Genetic algorithm	▪ Link counts	▪ in-vehicle time ▪ waiting time ▪ auxiliary time ▪ boarding time
Rydergren, 2013	Automated - SPSA algorithm, compass search algorithm	▪ Passenger travel time	▪ in-vehicle time ▪ access time ▪ egress time ▪ transfer walk time ▪ transfer wait time ▪ transfer penalty
Zhu et al., 2014	Automated – Genetic algorithm	▪ Travel time distribution	▪ standard deviation of normal distribution ▪ relative threshold of travel cost difference ▪ absolute threshold of travel cost difference
Tavassoli et al., 2019	Automated - PSO algorithm	▪ Passenger flow on a given segment for a given node ▪ Passenger trips on a given mode	▪ in-vehicle time ▪ wait time factor ▪ spread factor ▪ boarding time ▪ waiting time ▪ auxiliary transit time ▪ boarding time

2.6 Stochastic approximation

Stochastic approximation (SA) algorithms, also called stochastic search and optimization algorithms is an iterative type of optimization algorithms. Let Θ be the domain of allowable values for a vector θ. Here, θ represents the collection of adjustable values (e.g., estimates of the parameters used for calibration). SA algorithms aim to find the values of $\theta \in \Theta$ which minimize a scalar-valued objective function $L(\theta)$. At each iteration k, the SA algorithms produce new estimates, $\hat{\theta}_k$. As $k \rightarrow \infty$, $\hat{\theta}_k$ will converge to an optimal solution θ^{*} (James C. Spall, 2003).

The relationship between θ and $L(\theta)$ can be complex, and often it cannot be expressed in a mathematical form for the practical applications. For this reason, the simulation model considered as a black box that only allows the evaluation of the objective function for a given set of values in θ (Amaran, Sahinidis, Sharda, & Bury, 2016).

In the context of SA, the problems and algorithms should have at least one of the following properties (James C. Spall, 2003):

- There is random noise in the measurements of the objective function or gradient.
- As the algorithm iterates towards the solution, random choices are made to navigate in the search domain.

SA algorithms, in general, can be classified into four categories: 1. Direct search methods. 2. Gradient-based, 3. Annealing methods, and 4. Evolutionary computing. (Wolpert & Macready, 1997) suggests that no optimization algorithm can always be more efficient than other algorithms. Therefore, the researcher must consider the practical implications when choosing a SA algorithm.

One such practical consideration is that a transit assignment for a larger network usually requires considerable computational time to execute. Therefore, if a SA algorithm requires a large number of objective function evaluations, the computational time required for the calibration task becomes untenable.

2.6.1 Gradient-free form of stochastic approximation

For practical problems, it is often difficult or impossible to obtain the gradient information. Therefore, some SA algorithms approximate the gradient based on the measurements of the objective function $L(\theta)$. Therefore, these algorithms do not require the full knowledge of the relationship between inputs and outputs of the simulation.

The oldest method for gradient approximation is called Finite Difference Stochastic Approximation (FDSA) proposed by (Kiefer & Wolfowitz, 1952). In FDSA, small changes to the elements of θ are done one at a time, and for each change, the value of the objective function is measured to approximate the gradient. FDSA becomes inefficient at higher-dimensional problems as the required number of objective function evaluations grows directly with the dimensions. Due to this reason, the computational time required for a calibration task with FDSA can become untenable, as described in 2.6.

Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm provides a solution to the above-mentioned drawback of FDSA. The main benefit of SPSA is a reduction in the number of objective function measurements required for the gradient approximation. SPSA requires only two measurements of the objective function irrespective of the dimensions of the optimization problem. For a p-dimensional problem, FDSA requires 2p objective function evaluations for gradient approximation. For the same problem SPSA only requires two evaluations of the objective function, making SPSA p-folds faster than FDSA.

The number of iterations required for SPSA is always more than FDSA. However, under general conditions, SPSA and FDSA achieve the same level of statistical accuracy with a faster computational runtime (James C. Spall, 2003). Therefore, SPSA is a better choice to be used in transit assignment calibration as it provides advantages both in solution quality and computational time. However, the basic form of SPSA at higher dimensional calibration problems shows convergence issues (Kostic, Gentile, & Antoniou, 2017). This problem, in relation to transport modeling, has been studied intensely in dynamic demand calibration problems where SPSA algorithm has been used as the SA algorithm. Solutions, such as W-SPSA (Antoniou, Azevedo, Lu, Pereira, & Ben-Akiva, 2015), PC-SPSA (Qurashi, Ma, Chaniotakis, & Antoniou, 2020) have been proposed to overcome some shortcomings of the basic form of SPSA for higher-dimensional problems. For TAM model calibration, so far, only the basic form of SPSA has been applied (Rydergren, 2013).

2.7 Research gaps

From the review of literature, the following research gaps are summarized:

1. For TAM calibration, SPSA algorithm has been used in a study where the observed values were obtained from survey data. SPSA algorithm has not been tested for the TAMs where the observed data were collected with automatic data collection methods (e.g., AFC)
2. The results of the calibrated models were mostly limited to the simulator outputs used in the objective function. However, TAMs have a large number of simulator outputs. The impact of the chosen objective function on the simulator outputs, which are not a part of the objective function, also needs to be checked. This aspect has not been fully explored in previous studies.

2.8 Summary

In this chapter, a brief overview of TAMs and its complexity in terms of transit supply and demand was explained. For urban networks with lower headways, headway-based TAMs are more suitable. Since passengers make decisions based on the perceived journey time (PJT), the PJT parameters need to be calibrated to represent the passenger flows in the transit assignment model accurately. Next, the data collection methods for TAMs were explained. It was identified that the technological advancements with automated data collection methods, especially automated fare collection methods, help to accurately calibrate TAMs as it provides accurate data about observed passenger flows. The previous approaches in solving transit assignment calibration were then explained. Recent work is focusing more on automating the calibration process with the help of stochastic search and optimization algorithms. SPSA is one of such algorithms that requires a fixed (two) number of objective function evaluations regardless of the number of parameters being calibrated. Thus, SPSA is a more suitable algorithm to be used in calibrating large-scale TAMs where transit assignment is computationally expensive.

3 Methodology

This chapter is divided into three subsections. The first subsection describes the steps with SPSA algorithm and the approach to implementing SPSA for a transit assignment calibration problem. The next subsection describes the formulation of the objective function with details of goodness of fit measure. Finally, the method used to evaluate the calibration results are explained.

A schematic overview of the proposed methodology is illustrated in Figure 3-1. The inputs section describes the types of data required (i.e., models, hyperparameters) to implement the calibration methodology. The methodological components section provides an overview of the implementation of the calibration methodology. SPSA-DOF, highlighted in gray color, is the proposed calibration algorithm of this thesis is described in 3.1.2.

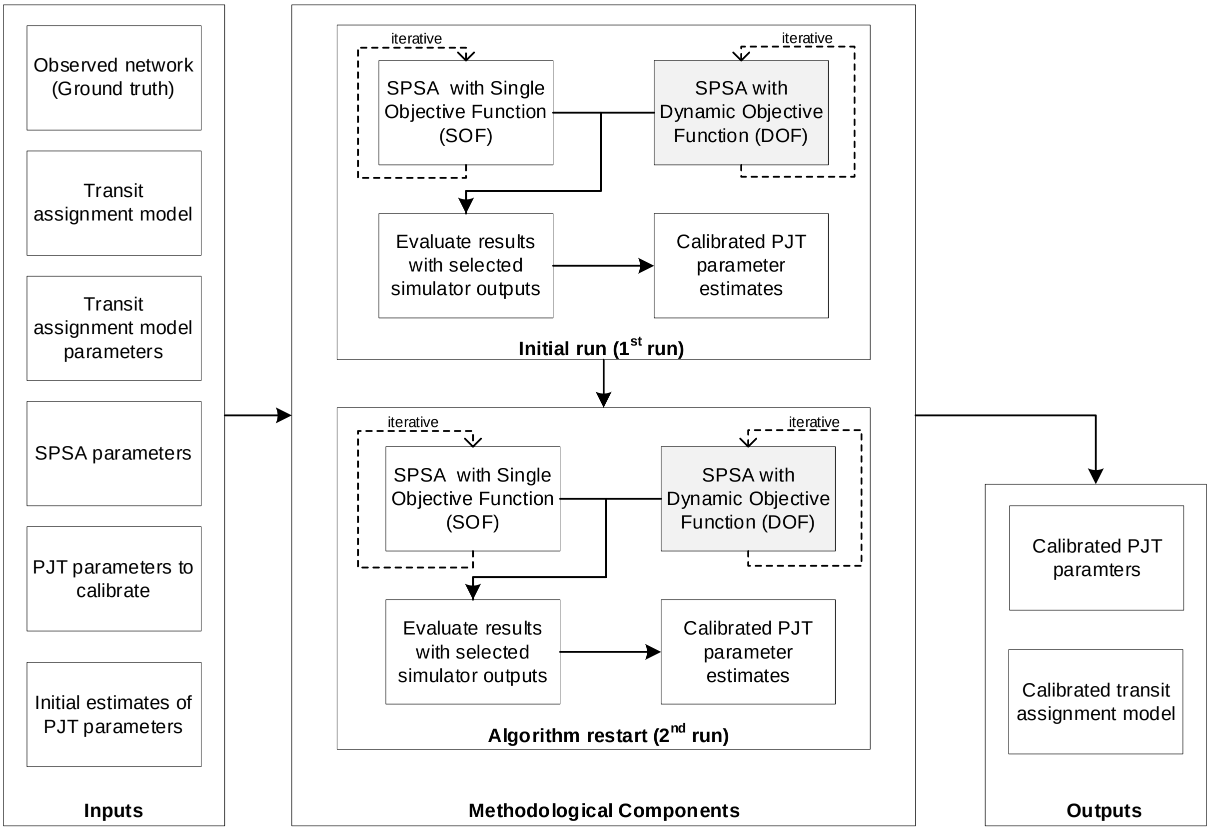

Figure 3-1 Schematic overview of the proposed transit assignment calibration methodology

This method is compared with the standard implementation of SPSA (referred to as SPSA-SOF in this thesis). The algorithm is restarted after a certain number of iterations with the aim of attaining better results. Then, the results from the calibration are evaluated to select the best set of PJT parameter estimates for a given network. These

PJT parameters and final calibrated models are the outputs of the implementation of the calibration algorithm.

3.1 Simultaneous Perturbation Stochastic Approximation

SPSA algorithm, proposed by (J. C. Spall, 1998), comes from the family of stochastic approximation (SA) methods. More specifically, SPSA can be used as a gradient-free SA, where the optimization is done only with the measurements (possibly noisy) of the objective function. This algorithm can provide significant efficiency gains when a problem has many variables to be optimized.

There are five hyperparameters in the algorithm. These are:

c, γ	To specify the gain sequence c_k, where k is the iteration number
a, A, α	To specify the gain sequence a_k
a_k	The gain sequence which governs the magnitude of minimization, where $a_k = \frac{1}{(1+k+A)^{\alpha}}$
c_k	The gain sequence which governs the magnitude of perturbation where $c_k = \frac{c}{(k+1)^{\gamma}}$

The chosen values for a, c, A, α, γ should satisfy the below criteria for the convergence of the algorithm:

$$a_k, c_k > 0 \qquad\qquad a_k, c_k \to 0 \text{ as } k \to \infty$$

$$\sum_{k=1}^{\infty} a_k = \infty \qquad\qquad \sum_{k=1}^{\infty} \frac{{a_k}^2}{{c_k}^2} < \infty$$

The selection of the hyperparameters, which meet the above criteria, can be made by the guidelines provided in (J. C. Spall, 1998) and (James C. Spall, 2003).

3.1.1 Major Steps with SPSA

There are six major steps involved in implementing the algorithm (J. C. Spall, 1998).

Step 1 **Initialization and coefficient selection:** Set counter index to 0 ($k = 0$). Set the initial estimate $\hat{\theta}_0$ (with *p*-dimensions). Set hyperparameters (a, c, A, α, γ) based on the criteria mentioned above to calculate the gain sequences.

Step 2 **Generation of the simultaneous perturbation vector:** Generate a Monte Carlo *p*-dimensional random perturbation vector Δ_k. Each element of the Δ_k vector is independently generated from a zero-mean probability distribution (Bernoulli, segmented uniform, U-Shaped). Effective and theoretically valid distribution is to use a Bernoulli ± 1 distribution with a probability of 0.5 for each outcome.

Step 3 **Objective function evaluation:** Perturbate the current estimate $\hat{\theta}_k$ and calculate two sets of intermediate estimates $(\hat{\theta}_k + c_k \Delta_k), (\hat{\theta}_k - c_k \Delta_k)$. Use these intermediate estimates (simply written: θ_k^+ , θ_k^-) to obtain two measurements of the objective function (y^+, y^-)

Step 4 **Gradient approximation:** Calculate the unknown gradient, $g(\hat{\theta}_k)$ according to equation (3.1).

$$\hat{g}_k(\hat{\theta}_k) = \frac{y(\hat{\theta}_k + c_{k\Delta}) - y\,(\hat{\theta}_k - c_k)}{2c_k} \begin{bmatrix} \Delta_{k1}^{-1} \\ \Delta_{k2}^{-1} \\ \vdots \\ \Delta_{kp}^{-1} \end{bmatrix} \tag{3.1}$$

Step 5 **Update θ estimate:** Use $\hat{g}_k(\hat{\theta}_k)$ with gain sequence a_k to update $\hat{\theta}_k$ to a new value $\hat{\theta}_{k+1}$ according to equation (3.2). In this step, constraints can be applied (if relevant) to keep the estimated values within a given range.

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k) \tag{3.2}$$

Step 6 **Iteration or termination:** Return to step 2 and increment k to $k + 1$. Terminate the algorithm if there is a small change between several successive iterations or maximum allowed number of iterations has been reached.

3.1.2 Algorithm to implement SPSA for transit assignment calibration

Implementation of SPSA on transit assignment calibration is done with some modifications to the basic SPSA algorithm. A standard approach to implement the calibration procedure is presented in Figure 3-2. The proposed method from this thesis is presented in Figure 3-3, and the modifications proposed to the algorithm are highlighted in gray. A step by step guideline to implement the proposed calibration workflow is described below. Most of the steps are common for both implementations, the differences, where applicable, are described separately. The implementation of the proposed calibration algorithm is done in Python 2.7 with Visum-COM API.

1. Initial estimate (θ_0) (with p-dimensions) can be set based on the parameter coefficients from a previously calibrated model. When this information is not available, the coefficients can be approximated by considering the rational behavior of the passengers and within the allowed value range in the simulator. Some general guidelines are summarized below:
    a. The in-vehicle time parameter coefficient can be fixed at 1 when onboard congestion is not considered in the transit assignment model (TAM). This will help to reduce the calibration problem by one dimension. The other parameters (wait times, walk times, access/ egress times) can have an initial estimate greater than 1. The transfer penalty can be set to the default values used in the models. The usual transfer penalties used in the headway-based assignment is 5 minutes or 10 minutes.
    b. The simulator provides lower and upper bounds for the parameter coefficients. For example, in PTV Visum, all the PJT parameters except the transfer penalty can vary between 0 and 9.9. The transfer penalty can take a maximum value of 1440 minutes (one day).
2. SPSA hyperparameters are set based on the guidelines provides by (J. C. Spall, 1998) and (James C. Spall, 2003). However, multiple tests are required in a small synthetic with low computational time to fine-tune these hyperparameters.
3. Evaluate a_k and c_k based on the gain sequence calculation equations provided in 3.1. These gain sequences are recalculated at the beginning of each iteration as k increases.
4. Generate a Monte Carlo *p*-dimensional random perturbation vector Δ_k with Bernoulli ± 1 distribution with a probability of 0.5 for each outcome. Perturb the estimate $\hat{\theta}_k$ such that it generates two intermediate estimates θ_k^+, θ_k^-. Make sure that the new intermediate estimates are within the allowed value range. In case if a perturbed estimate for a parameter goes out of the allowed value range,

replace the value with the respective parameter coefficient of the current estimate.

5. Objective function:
   a. For standard implementation – A single objective function (OF) can be used. This objective function may have multiple error calculation terms.
   b. For proposed implementation - Use two objective functions (OF_1, OF_2), which represents conflicting passenger behaviors. For example, one objective function could be the error of the direct trips (zero transfers) made by passengers on transit lines. The conflicting passenger behavior to the direct trips is the error of trips made with transfers or the error of the transfers made at transit stop locations, which can be chosen as the second objective function.
6. Evaluate objective function at iteration k:
   a. For standard implementation - Run the transit assignment with θ_k^+ and θ_k^- and evaluate the objective function, that will result in two objective function measurements (y_{OF}^+, y_{OF}^-)
   b. For proposed implementation - Run the transit assignment with θ_k^+ and θ_k^- and evaluate both objective functions separately, that will result in four objective function measurements ($y_{OF_1}^+$, $y_{OF_1}^-$, $y_{OF_2}^+$, $y_{OF_2}^-$).
7. Only applicable for proposed implementation: Select the objective function that provides the highest relative difference. Here the relative difference (RD) for a given objective function is defined as follows:

$$\mathrm{RD_{OF_i}} = \frac{y_{OF_i}^+ - y_{OF_i}^-}{0.5 \cdot \left(y_{OF_i}^+ + y_{OF_i}^-\right)} \tag{3.3}$$

   Based on the above calculation, the objective function which provides the highest RD is chosen as the objective function to be used in the current (k[th]) iteration OF_k for the gradient approximation (step 8).

$$\mathrm{OF_k} = \begin{cases} OF_1, & if\ \mathrm{RD_{OF_1}} \geq \mathrm{RD_{OF_2}} \\ OF_2, & otherwise \end{cases} \tag{3.4}$$

8. Approximate the gradient $\hat{\mathrm{g}}_\mathrm{k}(\hat{\theta}_\mathrm{k})$ based on the objective function measurements of OF_k as per equation (3.1).

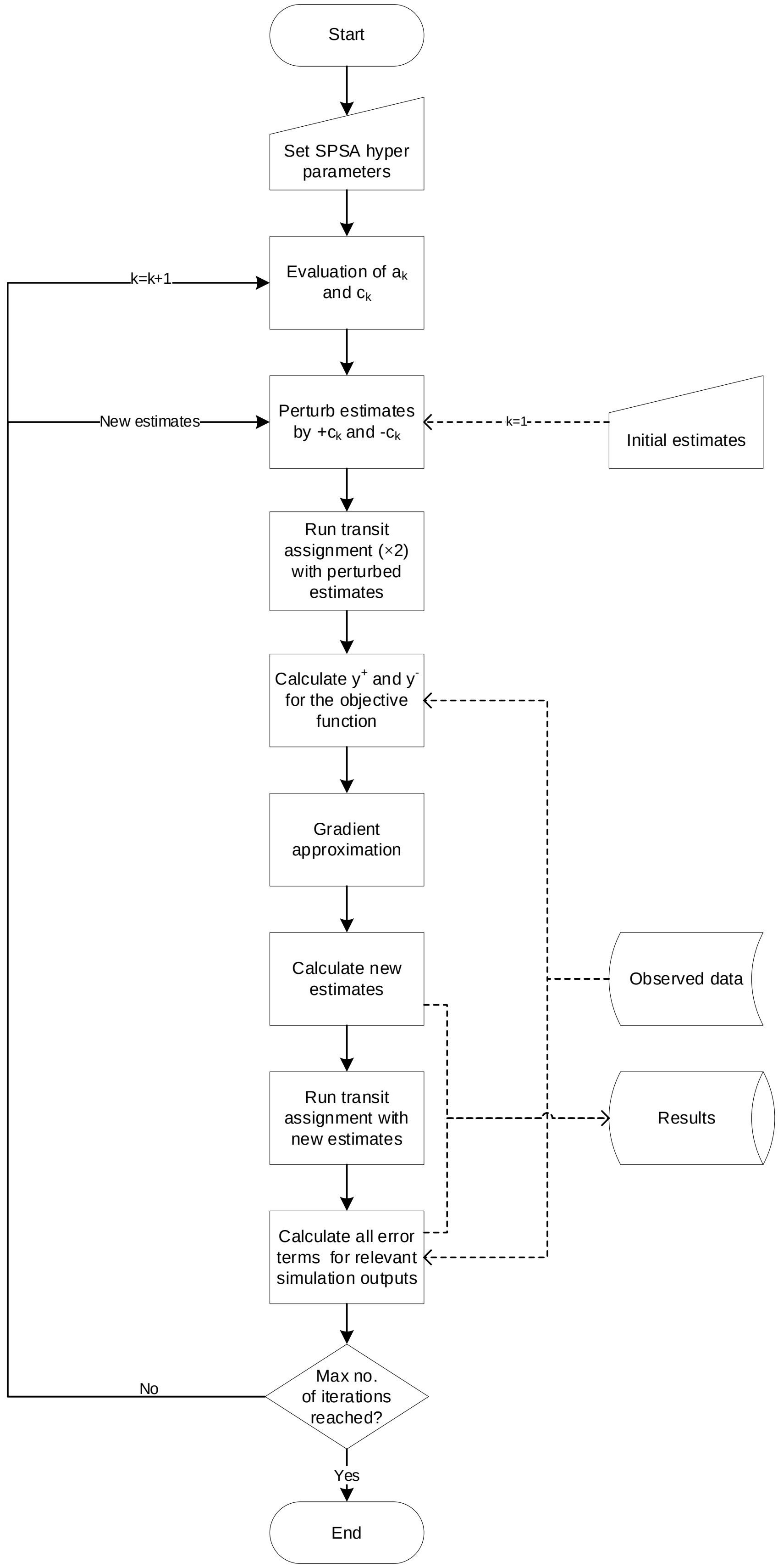

Figure 3-2 Workflow for transit assignment model calibration: SPSA-SOF

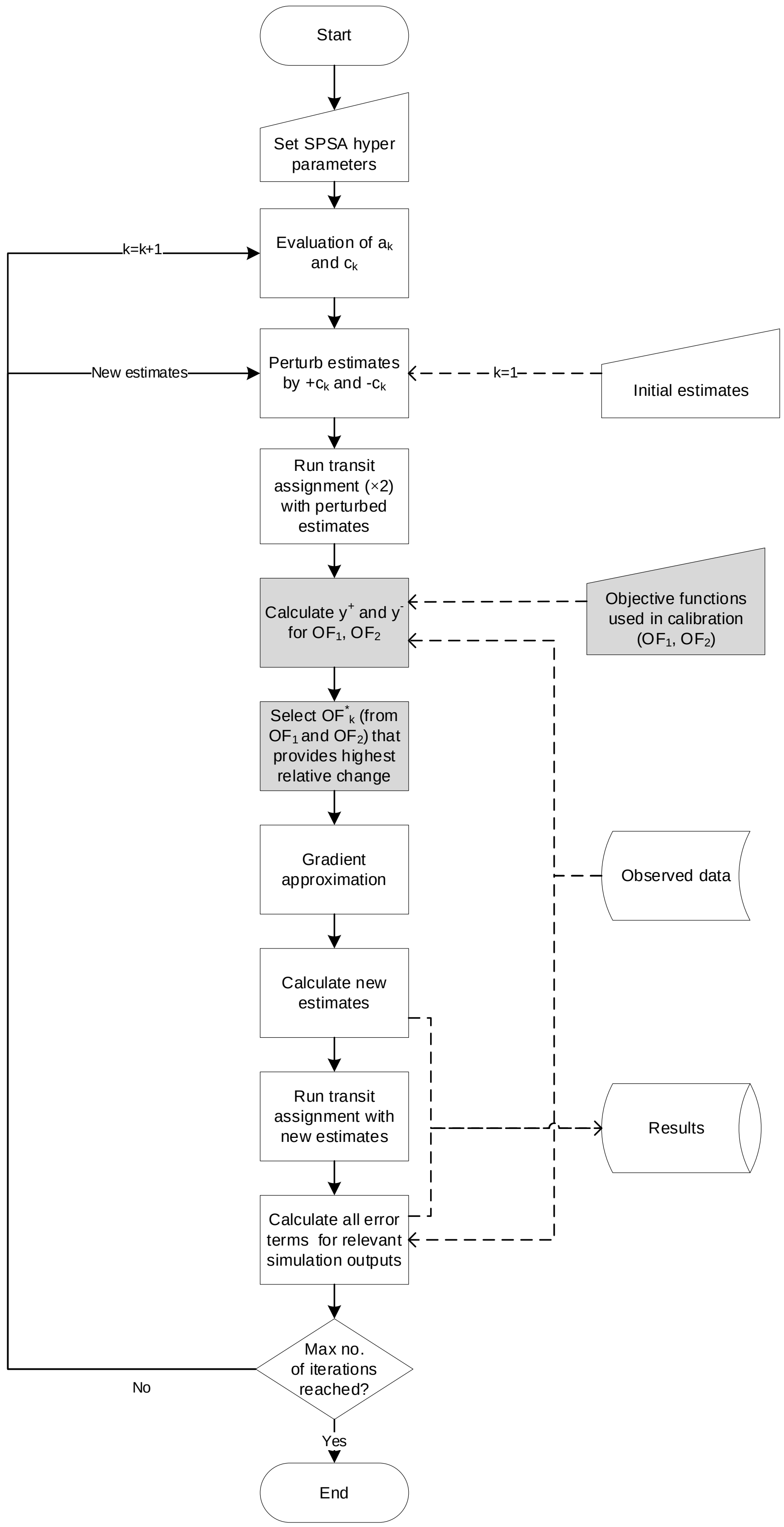

Figure 3-3 Workflow for transit assignment model calibration: SPSA-DOF

9. Calculate new estimates $\hat{\theta}_{k+1}$ as per equation (3.2), subject to the allowed value range. If a coefficient for a particular parameter goes beyond the constraints, it is replaced by the coefficient of the same parameter that provided the lowest error (best estimate). In the first iteration, the best estimate is considered as the initial guess.
10. Run the transit assignment with new estimates and calculate all the error terms for simulation outputs, which are in the interest of the calibration problem. A list of simulation outputs that can be used for the error calculation is given in Appendix-B.
11. Terminate the algorithm after the maximum number of iterations reached.

The implementation of SPSA with the proposed modifications (steps 5, 6, and 7) will be called SPSA with Dynamic Objective Function (SPSA-DOF). The term ‘dynamic’ is used to represent the behavior of the objective function selection. The pseudocode for SPSA-DOF implementation is given in Appendix-A.

Both standard implementation and SPSA-DOF are implemented for a given problem, and their performances are compared. In this thesis, the standard implementation, for the purpose of comparison, is named as SPSA-SOF (single objective function).

3.1.3 Algorithm restart

Usually, in the first half of the iterations, SPSA is able to reduce the error with a steep descent. However, towards the latter parts of the iterations, the error reduction is almost insignificant. As a solution, the algorithm can be restarted (second run) after a certain number of iterations (Kostic et al., 2017). A new set of initial estimates are chosen to restart the algorithm. This new set of estimates is the best solution from the previous calibration run. The number of iterations to run in the first run and the second run is decided based on the error reduction pattern and computational time required for the calibration.

3.2 Objective function

3.2.1 Introduction

The objective function used in this study represents two different types of passenger movements. The first type is the passenger flows on a transit line or a transit line route. Here, a transit line represents a dedicated transit service. A transit line route represents a dedicated transit service and its direction of operation. A typical transit network can have lines routes with two directions. Line route with one direction is also possible when the line route operates in a loop, or the service operates unidirectionally. The second type is the passenger transfers at a transit stop.

These two types of movements are somewhat conflicting in nature. If the selected objective function tries to minimize the error of the simulated and observed trips on a transit line, it might not reduce the error of the number of transfers passengers made at stops (and vice versa). Therefore, a more realistic objective function should contain both types of movements.

3.2.2 Goodness of fit measure

Goodness of fit (GoF) measures outputs a scalar value, which describes the magnitude of the error between the observed and simulated values. Different types of GoF measures quantify the error from different views (Papathanasopoulou & Antoniou, 2015). The GoF measure, normalized root mean square error (RMSN), is used in this study.

Equation (3.5) provides the mathematical representation of RMSN error.

$$\mathrm{RMSN} = \frac{\sqrt{n \cdot \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}}{\sum_{i=1}^{n} y_i} \tag{3.5}$$

Where y_i is the observed value and $\hat{y}_i$ is the simulated value of the corresponding observed value y_i. n presents the total number of values, and i is from the set $[1,2,\ldots,n]$.

3.2.3 General structure of the objective function

Based on the introduction and selection of GoF measure, the objective function to be evaluated at each iteration can be formulated as follows:

$$\alpha \cdot \frac{\sqrt{n \cdot \sum_{i=1}^{n} (\hat{y}_l - y_l)^2}}{\sum_{i=1}^{n} y_l} + \beta \cdot \frac{\sqrt{m \cdot \sum_{i=1}^{m} (\hat{y}_s - y_s)^2}}{\sum_{i=1}^{m} y_s} \tag{3.6}$$

Where:

$y_l, \hat{y}_l$	Observed and simulated passenger flows on transit lines (n lines)
$y_s, \hat{y}_s$	Observed and simulated passenger transfers at stops (m stops)
α, β	Weights with possible values of 0 and 1

For SPSA-DOF, according to the implementation described in 3.1.2 – step 7, the selection of one objective function according to equation (3.3) can be represented with values of α and β:

- $\alpha = 1, \beta = 0$: if the first objective function is selected
- $\alpha = 0, \beta = 1$: if the second objective function is selected

The standard implementation of SPSA (SPSA-SOF), can be explained as the case where $\alpha = 1$ and $\beta = 1$. Values of α, β remain unchanged at each iteration.

3.3 Evaluation of results

As per equation (2.2), the objective of the calibration is to find the coefficients of the PJT parameters that produce the least possible error between the simulated and observed values. This can be done by taking the parameters that provide the lowest RMSN value for the objective function used for the calibration. However, this method could become biased, as the other simulator outputs relevant to the transit assignment is not checked. This has been identified as a gap in the literature review (see 2.7).

At each iteration, a transit assignment is run with the new set of estimates. Therefore, it is possible to evaluate the RMSN error with all relevant simulator outputs without additional computational effort. A short description of the selected simulator outputs is given in Table 3-1. A detailed description of each simulator output is given in Appendix-B.

Table 3-1 Selected simulator outputs for solution evaluation

Simulator output	Short name
1. Passenger boarding for a transit line route	boardings
2. Passenger trips on transit line routes with no transfers	0 transfers
3. Passenger trips on transit line routes with one transfer	1 transfer
4. Passenger trips on transit line routes with two transfers	2 transfers
5. Passenger trips on transit line routes more than two transfers	> 2 transfers
6. Passengers making a transfer at the same stop	direct
7. Passengers alight at this stop and walk to another stop to make a transfer	alight walk
8. Passengers boarding at this stop after walking from another stop	walk board
9. Total number of passengers transferring (direct + alight walk + walk board) at a stop	total

The best set of estimates is selected considering the overall reduction of RMSN error for the simulator outputs mentioned above.

Once the best set of estimates is selected, the quality of the simulation can be evaluated using scatter plots. In principle, a ‘perfect’ calibration should align all the scatter points along the $45°$ line, but often that is not the case. Moreover, it is hard to visually inspect a scatter plot and evaluate the quality of the solution. The solution evaluation criteria for this study are established based on the theoretical conclusions of (Mesplé, Troussellier, Casellas, & Legendre, 1996) and practical implementation for a transit assignment calibration by (Parveen et al., 2007) and (Tavassoli et al., 2019). The steps are as follows:

1. First, a simple linear regression is done between the observed and simulated values. The quality of the solution is better if the slope is closer to 1, and the intercept is closer to 0.
2. Next, the r-squared (r^2) value is calculated between the observed and simulated values. The quality of the solution is better with higher r^2 values, as a higher proportion of the variance for the dependent variable is explained by the independent variable.

3.4 Summary

This chapter explained the different components involved in the proposed methodology for a systematic and automated calibration of a transit assignment model. A new method is proposed, named SPSA-DOF, where the selection of the objective function for a given iteration is made based on the relative difference of the error that an objective function produces by the perturbed parameter estimates. RMSN error is proposed as the error calculation method between the simulated and observed values. The proposed methodology is compared with the standard implementation of SPSA (SPSA-SOF), where a single objective function is evaluated throughout the iterations. In the next two chapters, the implementation of the proposed methodology on a synthetic network and a real network is presented.

4 Implementation of Proposed Calibration on a Synthetic Network

In this chapter, the methodology explained in chapter 3, is implemented in a synthetic transit network. First, the network setup is explained briefly with a description of generating the observed passenger flows. Then, the steps of implementing the calibration algorithms (SPSA-SOF, SPSA-DOF) are explained. Finally, the results of the implementation are presented, and the calibrated models are compared with each other.

4.1 Introduction

The synthetic network developed as a testbed to implement the calibration algorithm. The computational time required to execute a transit assignment is very low (less than a second). Therefore, multiple tests can be run with less computational time, which helps to understand the behavior of the calibration algorithm with a transit assignment model (TAM).

4.1.1 Transit network

The synthetic network is created in PTV Visum. The created network consists of 11 zones, four bi-directional transit lines (8 transit line routes), 11 stops, and 22 stop points. A synthetic OD matrix is used for transit assignment. The demand is assigned over a one-hour time period (7:00 am – 07:59 am). The synthetic network is able to simulate all the travel time components mentioned in equation (2.1).

Since this is a small network, some intentional adjustments were made to add some complexity to passenger movements. First, the transit lines were arranged in such a way that transfers are encouraged. Second, for a given stop, the walking times within the same stop area are increased than the walking times between two stop areas. This arrangement is a bit unrealistic compared to real transit networks but makes passenger transfers more complex by encouraging the transfers between stop areas. An overview of the synthetic network is given in Figure 4-1.

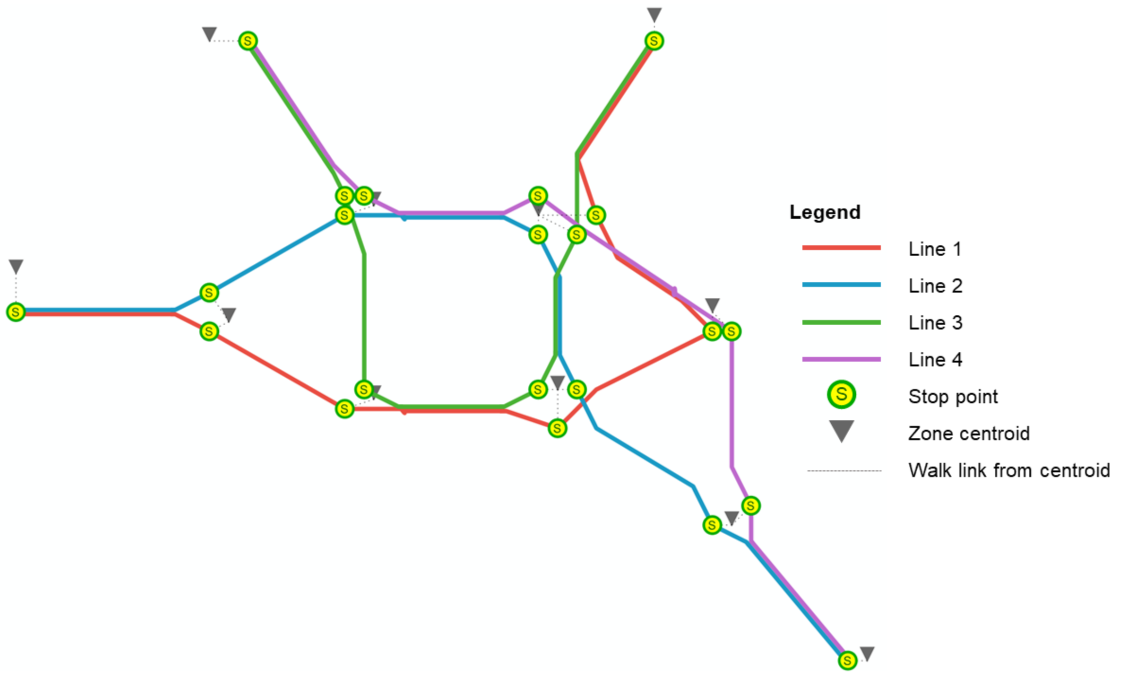

Figure 4-1 Synthetic network

4.1.2 Observed values

Since there are no observed values for the synthetic network, observed values were generated with a specific set of parameter coefficients.

Table 4-1 PJT parameter coefficients used to generate observed values

Parameter	Coefficient
In-vehicle time (IVT)	1.0
Access time (AT)	2.0
Origin wait time (OWT)	2.0
Transfer walk time (WT)	1.5
Transfer wait time (TWT)	3.0
Egress time (ET)	2.0
Transfer penalty (TP)	5 min

The coefficients were chosen such that the wait (OWT, TWT) and walk (AT, WT, ET) components of PJT are greater than one. TP is set to 5 mins, which is a standard value used in practice. The headway calculation method was set to 'Mean headway according to timetable', and 'Ignore path share' was set to 0.001. The transit assignment was run with these parameters. The simulated values related to passenger transfers at stop points and passenger flows on transit line routes were exported as comma-separated values (CSV) files.

4.2 Implementation of the calibration algorithm

4.2.1 Experimental parameters and setup

The experimental parameters used for the synthetic network are summarized in table Table 4-2. All the parameters remain unchanged across all experiments.

Table 4-2 Parameters used for the synthetic network calibration

Type	Parameter	Value / setting (remarks)
Simulator	Headway calculation	Mean headway according to timetable
	Assignment time interval	1 hour
	Boarding strategy	Optimal strategies
	Ignore path, if share	0.001
Initial guess	In-vehicle time (IVT)	1.0 (fixed)
	Access time (AT)	7.855
	Origin wait time (OWT)	7.479
	Transfer walk time (WT)	3.254
	Transfer wait time (TWT)	1.067
	Egress time (ET)	2.045
	Transfer penalty (TP)	6.562 minutes
SPSA	α	0.602
	γ	0.101
	a	7.250
	c	1.072
	A	30
	Max iterations	300

Initial estimates were chosen in a way such that it gives a higher RMSN error. IVT parameter was not calibrated and fixed at 1.00. Since the computational time required for one transit assignment is very less, the calibration algorithm was run for 300 iterations. The algorithm was restarted with the best estimates in the first run as the initial estimates for the second run. A higher a value is chosen in the SPSA hyperparameters so that the new estimates are more dependent on the gradient information.

The overview of the experimental setup for the synthetic network is given in Figure 4-2. The SPSA-SOF and SPSA-DOF is implemented for the first run with the initial guesses given in Table 4-2. The best estimates from the fist calibration are chosen by evaluating the change of the RMSN error for a selected set of simulator outputs.

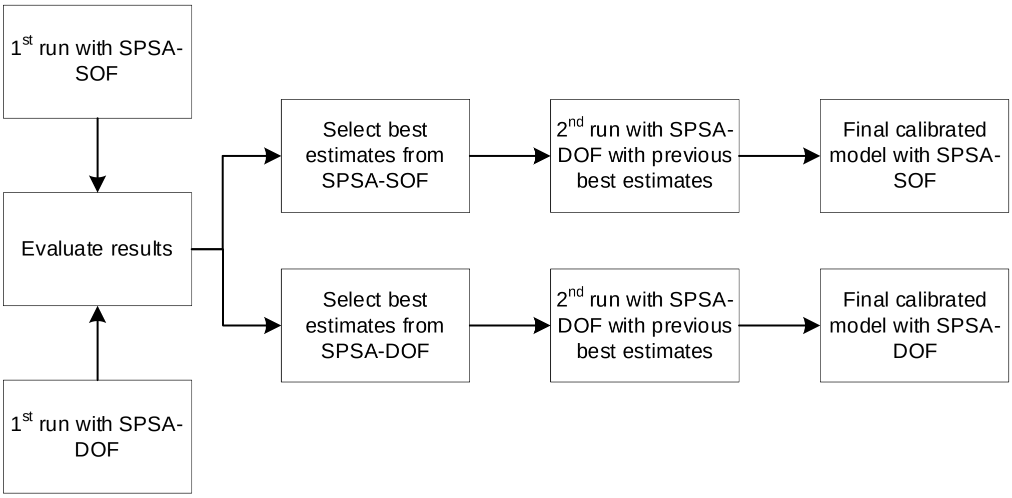

Figure 4-2 Experimental setup - synthetic network

The parameter estimate set that produced the lowest RMSN error was chosen as the best estimate. These estimates were used as the initial estimate for the second run. The same procedure was applied to choose the best estimates (final calibrated model parameter coefficients) from the second run.

4.2.2 Objective functions

A summary of the objective functions used for the calibration of the Synthetic network is given in table Table 4-3.

Table 4-3 Objective functions used for the calibration of the synthetic network

Objective function (OF)	Description
OF_1	RMSN error of simulated and observed passenger trips on transit line routes with zero transfers
OF_2	RMSN error of the total passenger transfers at stop points

Same objective functions were used in both SPSA-SOF and SPSA-DOF and implemented as described in 3.2.3.

4.3 Results

The results obtained from the implementation of SPSA-SOF and SPSA-DOF summarized in this subsection. First, the change of RMSN error with selected simulator outputs are presented. The best set of parameter coefficients are chosen, and the transit assignment is executed with those selected parameter coefficients, resulting in two calibrated TAMs. The accuracy of the calibrated models is then compared with the model outputs for the initial guesses.

The change of RMSN error with each iteration is presented in Figure 4-3. A comparison is made for each simulator output with SPSA-SOF and SPSA-DOF. The results from the first 300 iterations and the second 300 iterations are plotted together. The reason is for simplicity and in order to get a better understanding of the error reduction patterns.

In general, SPSA-DOF can reduce RMSN error compared to SPSA-SOF. A significant reduction of RMSN error can be seen with the restart of the algorithm (from iterations 301-600). In terms of passenger trips on transit line routes (subplots a.2 to c.2), the highest reduction of RMSN can be seen in passenger trips with one transfer. A similar reduction for passenger transfers at stop points (subplots d.2 to f.2) with a higher reduction in with passenger walk board transfer type. This reduction also affects error reduction in total passenger transfers. However, the error reduction with passenger transfers is wiggly. One of the possible reasons for this behavior is the limited number of transfer options available for the passengers.

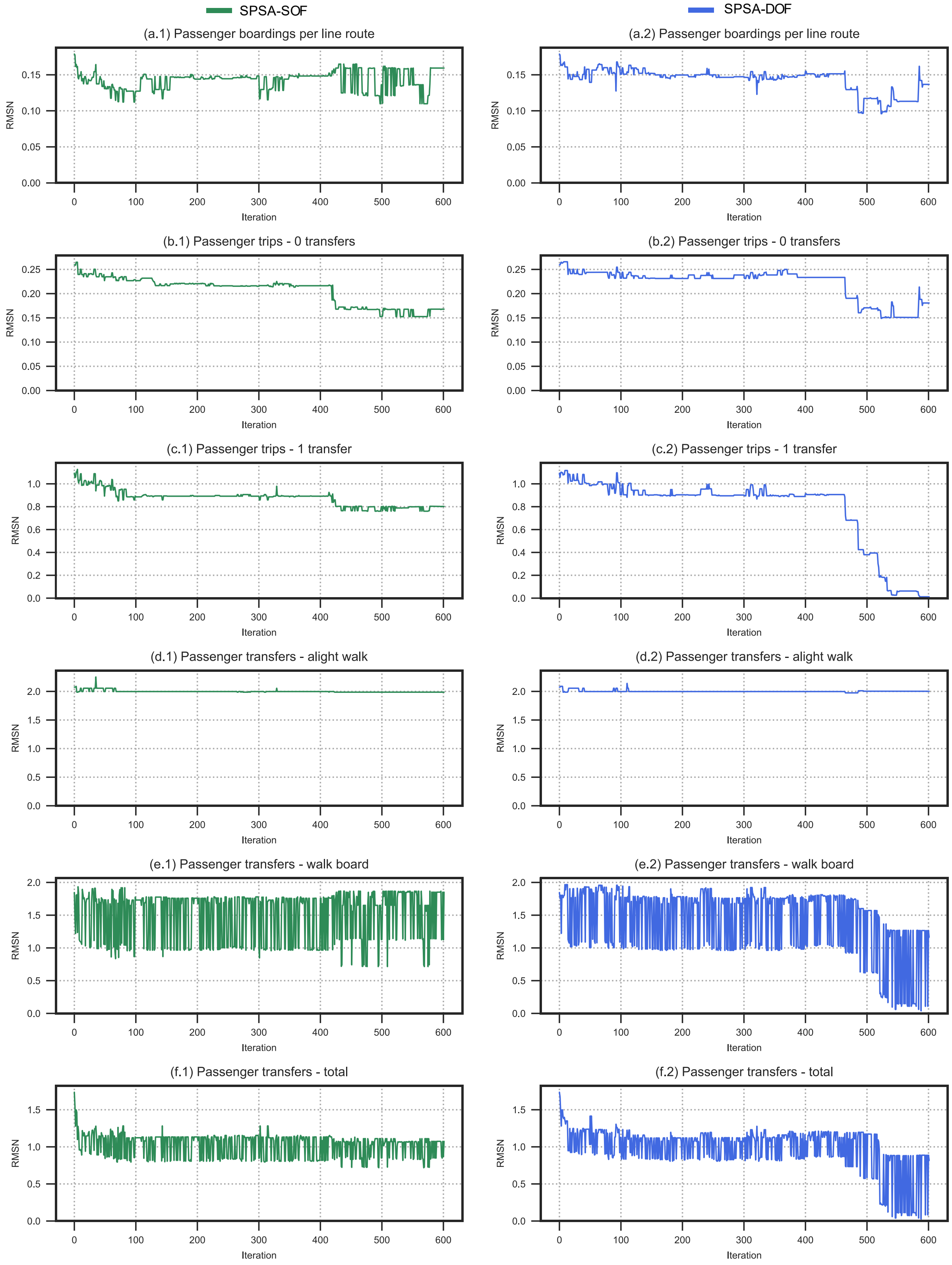

Figure 4-3 Comparison SPSA-SOF and SPSA-DOF: Change of RMSN error

The final set of calibrated parameters from the implementation of SPSA-SOF and SPSA-DOF is presented in Table 4-4. The best set of parameters was obtained with SPSA-DOF implementation, highlighted in the table.

Table 4-4 Summary of PJT calibrated values for the synthetic network

PJT parameter	Desired solution (θ_{obs})	Initial guess	Best estimate – 1st 300 iterations		Best estimate – 2nd 300 iterations	
			SPSA-SOF	SPSA-DOF	SPSA-SOF	SPSA-DOF
IVT	1.0	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
AT	2.0	7.855	4.021	7.159	1.460	5.373
OWT	2.0	7.479	7.678	7.175	7.521	3.178
WT	1.5	3.254	1.031	1.258	1.036	2.051
TWT	3.0	1.067	1.836	2.333	1.281	3.085
ET	2.0	2.045	3.309	4.779	3.313	5.356
TP	5.0	6.562	8.583	8.353	7.465	8.411
Best estimate at Iteration:			*144*	*164*	*268*	*245*

Most of the calibrated parameter coefficients are not comparable with the coefficients used to generate the observed data (desired solution). However, the calibrated models provide a better fit to the observed models. A comparison of the calibrated models with the initial values is provided in Figure 4-4 Figure 4-5.

Figure 4-4 shows a comparison among the model outputs with the initial guess, SPSA-SOF, and SPSA-DOF for passenger trips on transit line routes. In general, both calibrated models provide a better fit compared to the model outputs with the initial guess as calibrated models have a slope closer to 1, lower intercept and a higher r^2 values. A comparison between the calibrated models also reveals that the calibrated model outputs with SPSA-DOF provide a better fit than the calibrated model outputs with SPSA-SOF. For example, passenger trips with one transfer (c.3) show a near-perfect calibration as all the points lie on the $45°$ line with a r^2 value almost equal to one.

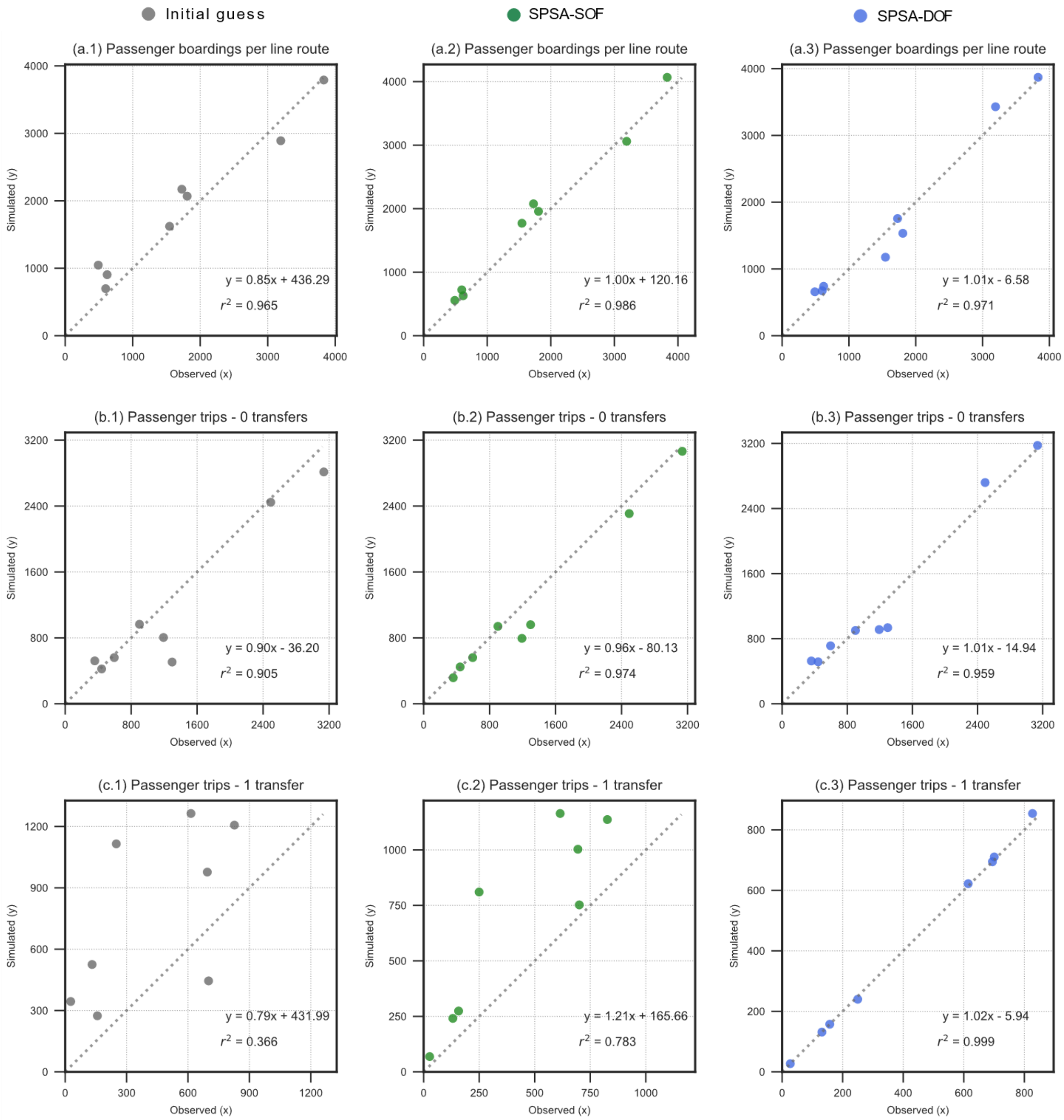

Figure 4-4 Comparison of the calibrated models with the initial guess: passenger trips on transit line routes

Figure 4-5 shows a comparison among the model outputs with initial guess, SPSA-SOF, and SPSA-DOF for passenger transfers at stop points. The calibrated models, in general, provides a better fit than the model simulated with the initial guess. A comparison between SPSA-SOF and SPSA-DOF shows that SPSA-DOF provides a better fit. However, the r^2 values are lower compared to SPSA-SOF as the simulated and observed values between a couple of stop points show a larger difference.

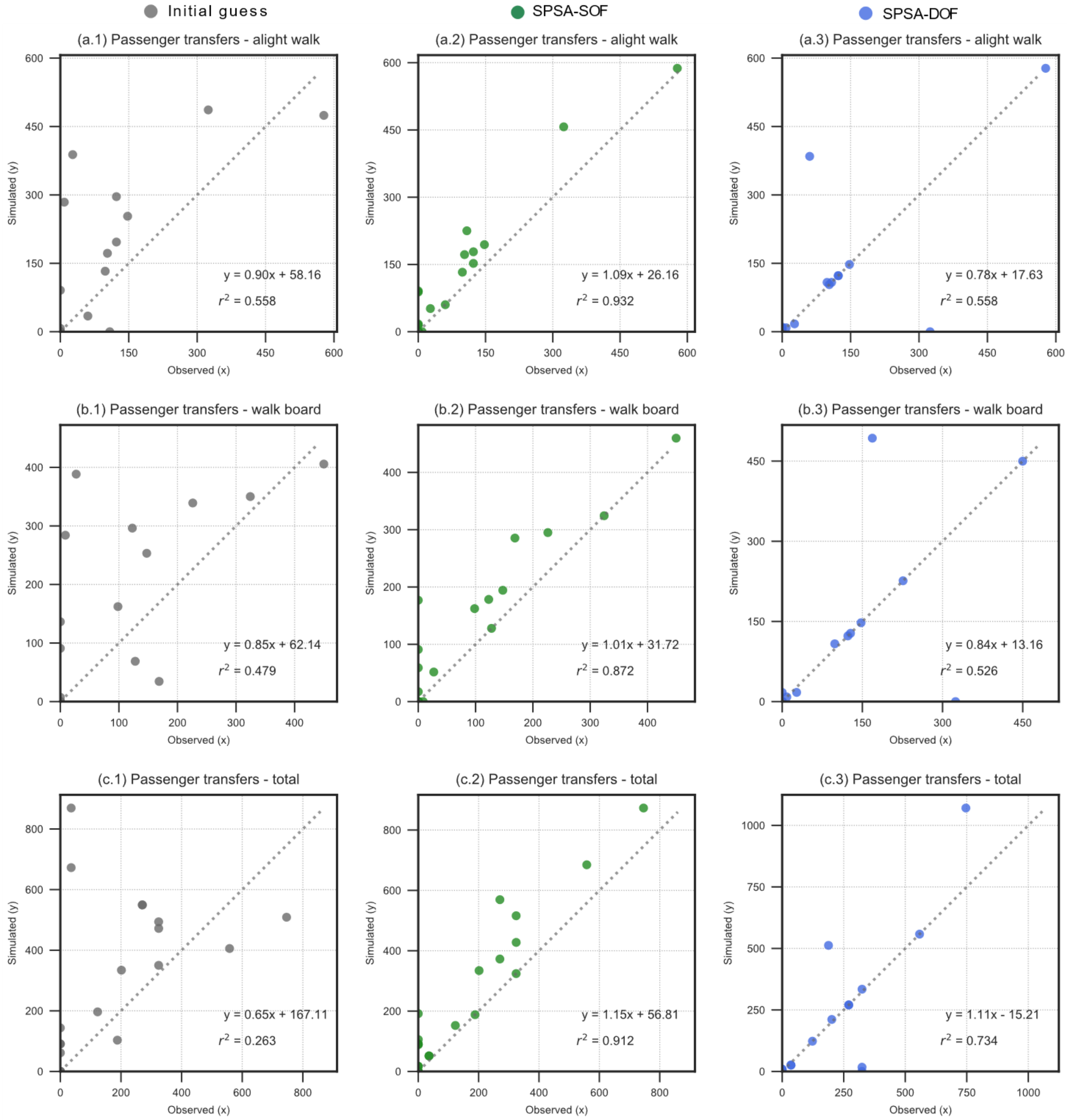

Figure 4-5 Comparison of the calibrated models with the initial guess: passenger transfers at stop points

4.4 Summary

This chapter focused on the implementation of the developed methodology on a synthetic transit network. The results achieved from the proposed method, SPSA-DOF, outperform the standard implementation (SPSA-SOF). Even though both algorithms were not able to reach the global optimum (desired solution in Table 4-4), the calibrated model parameters, especially with SPSA-DOF provides an accurate fit to the observed values. Due to the small scale of the network, there could exist multiple solutions for PJT parameter estimates that produce similar types of passenger flows, which is in close representation to the observed data. This could be one of the reasons for the calibrated parameter estimates are not comparable with the desired solution. In the next chapter, the proposed calibration methodology is tested on a large-scale transit assignment model of Singapore.

5 Implementation of Proposed Calibration on the Transit Network of Singapore

The methodology explained in chapter 3, is applied to the transit network of Singapore. TUMCREATE has AFC data-driven, macroscopic transit models modeled in PTV Visum. The availability of these models helped to apply the proposed methodology for a real-world, large scale transit model. In this chapter, first, a brief introduction of the public transport network of Singapore is given with details about the transit lines and ticketing system. Next, the transit assignment models (TAM) used for the calibration, along with some methods used for data preparation, are briefly explained. After that, the experimental setup is described in detail, along with a method used to reduce the computational time required for the calibration. Finally, the results are presented along with a comparison with the observed data and non-calibrated models.

5.1 Introduction

Singapore is a small island country with a population of 5.7 million and a population density of 7,866 people per km^2 (Department of Statistics Singapore, 2020). Singapore has relatively low car ownership of 101 cars per 1,000 people, which makes the transit system the backbone for mobility needs. Singapore is among the top-ranked cities for public transport with an 86% satisfaction rate for the overall situation of public transport (McKinsey & Company, 2018)

The aim of Singapore is to achieve 75% of trips during morning and evening peak hours to be made by using public transport modes. As of 2013, 63% of all trips during peak hours were made by public transport (LTA, 2013). From these trips, the bus network accounts for an average daily ridership of 3.6 million trips, followed by the Mass Rapid Transit (MRT) with 2.6 million trips and Light Rail Transit (LRT) with 132,000 trips.

5.1.1 Transit network

The transit network of Singapore consists of three primary modes: 1. MRT, 2. LRT, and 3. Bus network. Figure 5-1 shows an overview of the transit network as of 2013. The network consists of four MRT lines, four LRT lines, and over 300 bus lines.

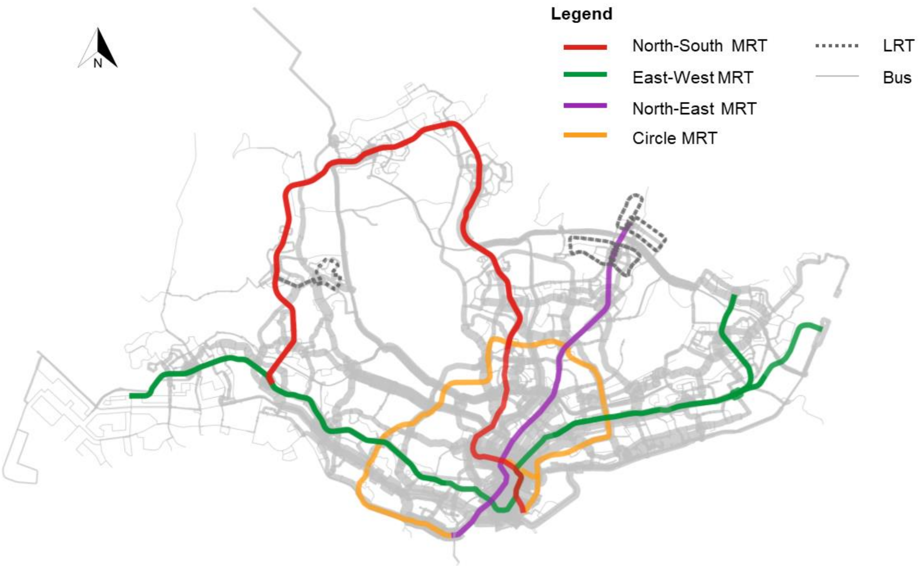

Figure 5-1 Transit network of Singapore as of 2013

There are different bus service types. Except for the basic bus service, there are different bus services provided based on the need of the passengers. Some of these different types of bus services are summarized in table Table 5-1.

Table 5-1 Different types of bus services (Land Transport Guru, 2013)

Bus service	Description of service
Basic	▪ Provide daily connections over a vast network of routes at a basic fare o Trunk services – connect different towns o Feeder services – operates around a neighborhood
Express and City direct	▪ Services more expensive than basic service o Full day – connects housing estates to the City o Limited-stop – Skip-stop operation o CBD – connects housing estates to Central Business District
Premium Bus Service (PBS)	▪ Service with a guaranteed seat on the bus ▪ Higher fare than basic service, cater to a niche market of commuters
Special service	▪ Running between residential towns and Chinatown

5.1.2 Ticketing system

The fare collection for the transit rides is done with an application of AFC system. This system is called CEPAS (Contactless e-Purse Application), a contactless smart card used for transit payments combined with other payment options, introduced in 2002. This smart card system is one of the systems in the world which has a higher customer penetration rate (AECOM, 2011)

A passenger has to tap-in the smart card at the boarding to a transit service and tap-out when alighting the transit service. The fare is calculated based on the total distance traveled across different transit services. This integrated fare calculation system encourages passengers to complete their journeys by transferring between different transit services. However, this is bounded with some rules. For example, a passenger can make up to five transfers within a single journey, with a 45-minute time window between each transfer. The entire set of rules are described in (LTA, 2020)

In addition to the fare calculation, these smart card data provide valuable information about the passenger journey, which can be used to develop TAMs. Some of them are summarized below:

- Boarding and alighting stop locations
- Time spend on a transfer, the number of transfers made
- Total ride time and ride distance at each transit service

5.2 Transit assignment models

Two types of Visum models were used in this study. The first model, called a Direct assignment (DA) model, developed based on CEPAS smart card data, which replicates the multi-modal transit system in Singapore. DA model used to obtain the observed data for this study. The second model, Transit Assignment Simulation (TAS) model used to run simulations and implement the calibration algorithm.

5.2.1 Direct assignment model

DA model was developed based on the smart card database provided by the Land Transport Authority (LTA) of Singapore. The data provided by LTA consists of information on all trips made with smart cards from 1st of August 2013 to 31st of October 2013. To replicate the normal travel behavior of the passengers, the demand for ten weekdays is selected with the least fluctuations in the demand patterns. More details about the development of the DA model can be found in (Liu et al., 2019). In this study, the DA model used to obtain the observed data.

5.2.2 Transit assignment simulation model

Even though the DA model replicates the observed route choices and passenger flows, it cannot be used to run simulations. TAS model is developed with the same ten weekdays demand used in the DA model. Therefore, the OD matrix used in the TAS model represents ten days of travel data. Thus, the comparison of the DA model and the TAS model becomes possible.

The TAS model of Singapore has a simplified representation of passenger movements compared to the conventional transit models. In a conventional model (like the synthetic network in chapter 4), a passenger starts the trip with an access link from a zone to the origin stop. Similarly, after the trip with a public transport service, the trip ends at a destination zone connected via an egress link. This first mile and last mile connection are not represented in this model. Instead, the model represents a trip such that a trip starts and ends at a transit stop. Thus, access time and egress time parameters do not have an impact on passenger route choice. Therefore, these parameters are not calibrated.

5.3 Data preparation

There were two tasks involved in this step. The first task was to export data from the DA model such that it can be used as an input as observed values in the calibration process. The relevant observed values were exported as a comma-separated value (CSV) files.

The second task involved with TAS model. There were some transit lines modeled in the TAS model, which had to be removed in order to implement the calibration algorithm properly. The types of lines removed, and reasons are summarized in table Table 5-2.

Table 5-2 Summary of the removed transit lines

Transit lines removed	Reason
Night Rider bus service	Operates on Fridays, Saturdays, and Public holidays. Therefore, the timetables associated with these lines are not modeled
Downtown MRT line	Passenger demand for this line was not available in 2013
Punggol West LRT line	No demand data available as it was opened in 2014
Premium bus lines	Faster lines operate only in the morning and evenings. Therefore, these lines have a lower impedance that could result in over assignment in passenger demand for premium bus lines.

5.4 Implementation of the calibration algorithm

5.4.1 Experimental parameters and setup

The experimental parameters used for the Singapore transit network are summarized in Table 5-3. Most of the parameter values/ settings are common for the experiments conducted below. However, two parameters (ignore path share, max iterations) were chosen considering the computational time required for the calibration.

Table 5-3 Parameters used for Singapore transit network calibration

Type	Parameter	Value / setting (remarks)
Simulator	Headway calculation	Mean headway according to timetable
	Assignment time interval	24 hours
	Boarding strategy	Optimal strategies
	Ignore path, if share	▪ 1st run - 0.0475 ▪ 2nd run - 0.001
Initial guess	In-vehicle time (IVT)	1.0 (fixed)
	Access time (AT)	(not calibrated see 5.2.2)
	Origin wait time (OWT)	2.0
	Transfer walk time (WT)	2.0
	Transfer wait time (TWT)	2.0
	Egress time (ET)	(not calibrated see 5.2.2)
	Transfer penalty (TP)	5 mins
SPSA	α	0.602
	γ	0.101
	a	4.833
	c	1.419
	A	30
	Max iterations	▪ 1st run - 80 ▪ 2nd run - 20

Though the OD matrix consists of 10 days of demand, the transit assignment is done for one day (24 hours). This is based on the assumption that the demand for these ten weekdays has minimum fluctuations in the demand patterns.

The overview of the experimental setup for the Singapore network is given in Figure 5-2.

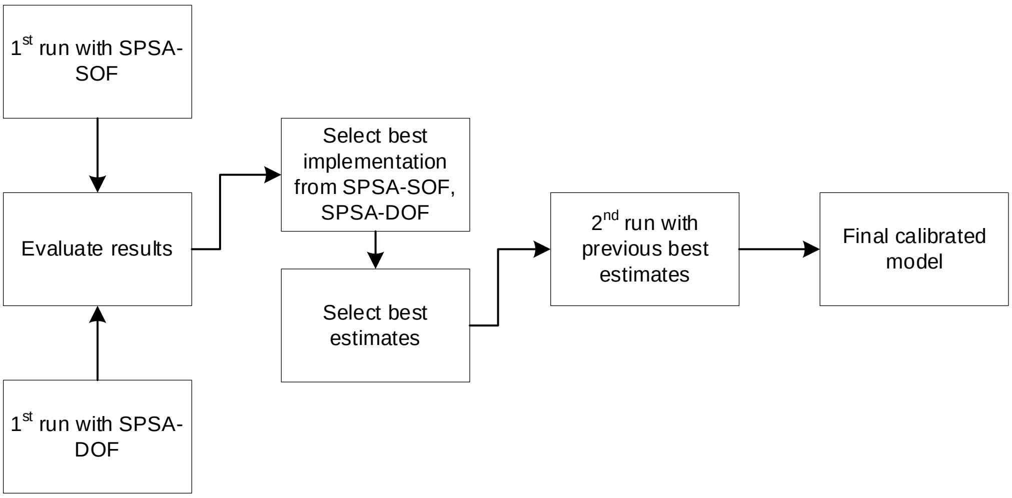

Figure 5-2 Experimental setup - Singapore network

- Calibration was done with both SPSA-SOF and SPSA-DOF to compare the best implementation. Since transit assignment in Singapore network is computationally expensive, the number of iterations had to be reduced. Moreover, a Visum parameter (‘Ignore path share’) was adjusted to reduce the computational time further for the 1st run. In the second run, this parameter was set to Visum’s default value. A detailed description of this given in 5.4.2.
- The algorithm was implemented in a Windows Server 2012 R2. The server is equipped with two Intel Xeon CPUs (E5-2640 v3 @ 2.6GHz) and 128GB RAM.
  - Calibration time with the settings of 1st run took 43 hours on average.
  - Calibration with the settings of the 2nd run executed only once and it took 78 hours.

5.4.2 Reducing the computational time required for transit assignment

Simulators tend to generate more route choices between an OD pair compared to the actual route choices made by the passengers. The reason is that a large share of passengers tends to choose few different alternative routes to travel between the origin and destination even though it is possible to choose many different route choices between the given OD pair. However, in a simulation, the simulator evaluates all feasible routes available between the OD pair and assign a percentage of trips for each feasible route. As a result, a small percentage of trips can be assigned to some routes.

For bigger networks like the Singapore network, this phenomenon increases the time taken to complete a transit assignment as the simulator has to evaluate all the feasible routes and assign demand despite how small the assigned proportion is. As a result, the time required to calibrate a large network will increase.

The parameter *"ignore path, if share"* in PTV Visum's headway-based assignment procedure, provides the ability to change the minimum share. This setting makes sure that the time taken for the assignment is not increased due to the evaluation of paths that are very unlikely to have a minimum impedance (PTV, 2019, p. 2043). The value cannot be applied as separate values for each OD pairs. Therefore, a reasonable value that is applicable to the entire network should be derived from the observed data.

There is no one correct solution to this problem. The nature of the solution is highly dependent on the available form of the observed data. The basic idea is to check all the OD pairs and calculate the minimum route share percentages for each OD and decide a cutoff point based on the calculated values. The approach used to solve the problem in this thesis is presented in Figure 5-3. This method is implemented in the DA model as it contains the observed route choices of the passengers.

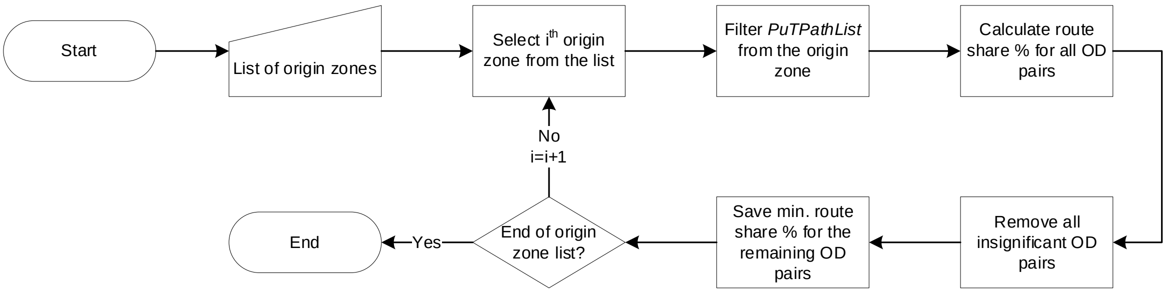

Figure 5-3 Workflow to find minimum route share

The steps for implementing the workflow is described below:

1. List down all the zones in the OD matrix. The calculations are done for one origin zone at a time to make the calculations computationally efficient as there can be millions of observed data records under one origin zone to all its destinations.
2. Select the first origin zone in the list and take all the observed trips records to a data frame. In PTV Visum, this data can be accessed via the *PuTPathList* object. The data should be retrieved from this object in such a way that it includes information about each trip made by its 1. Origin zone (current zone in concern) 2. Destination zone, 3. The number of transfers made during a trip, and 4. Total distance traveled inside a transit mode.
3. The data retrieved in step 2 can be used to understand different routes passengers have used to travel between the origin and destination. In order to do this, individual trip records need to be aggregated. It is reasonable to assume that the passengers have taken the same route if the origin zone, destination zone, number of transfers made during the trip, and the total distance traveled inside a transit line is the same (despite the impact of common line problem).

Based on this logic, the data can be aggregated, and different routes used to travel between the origin and destination can be identified. Once this is done, the percentage of trips made on each route (route share %) can be calculated, and the minimum route share for each OD pair can be obtained.

4. Once step 3 is done, records which are insignificant need to be removed in order to get a better understanding of the distribution of the minimum route share. In this study, three criteria were used to identify OD pairs, which has less importance to calculate the minimum route share percentage. An OD pair is removed if:
    a. The total number of trips between the origin and destination is less than ten trips (one passenger trip a day).
    b. There is a dominant route which has a route share higher than 99%. Here it is assumed that all the passengers travel on the dominant route.
    c. The route which has the minimum share has less than ten trips.
5. Continue steps 2 to 4 until the zone list is exhausted.

Once the calculation is complete for all the zones in the network, the distribution of the minimum route share can be analyzed and calculate a value to be used in the simulator. The minimum route share percentage distribution for the Singapore DA model is shown in Figure 5-4.

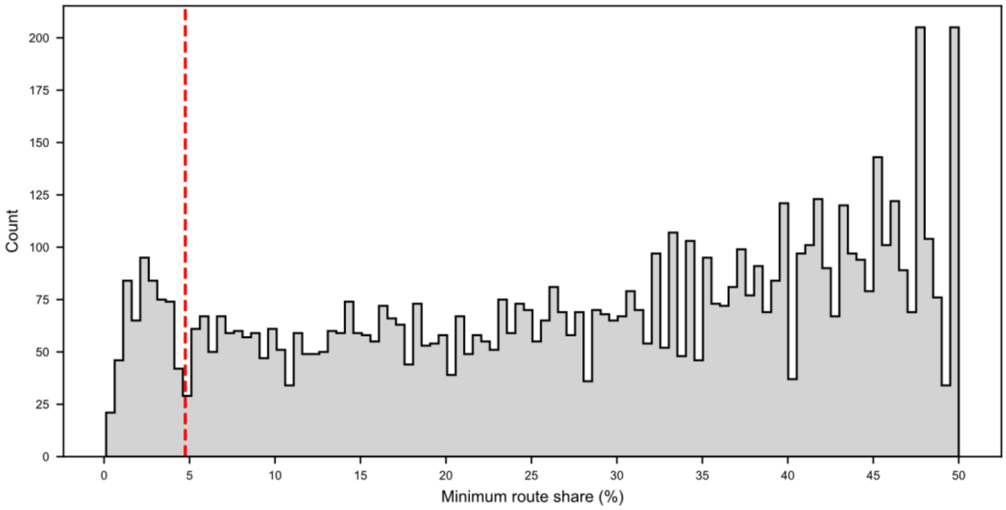

Figure 5-4 Minimum route share percentage distribution for the Singapore network

The value for minimum route share to be used in the simulator was chosen as the point which gave the first sharpest drop (4.75%). With this setting, the time required for one transit assignment dropped to ~10mins, which is a significant reduction compared to the time required for one assignment (~50mins) with default setting (0.001). However, increasing ‘Ignore path share’ to 0.0475 has one drawback. This increases the number of unassigned trips. This can be fixed by running the simulation with Visum's default setting (0.001). Therefore, the first run of the calibration algorithm is run with the increased value of 0.0475. This ensures that the initial results of the calibration can obtain with a lesser computational time. The second run (algorithm restart) can be run with the default setting (0.001). This takes longer computational time, but it ensures that the final calibration model is not under-represent the observed values.

5.4.3 Objective functions

A summary of the objective functions used for the calibration of the Singapore network is given in Table 5-4.

Table 5-4 Objective functions used for the calibration of Singapore Network

Objective function (OF)	Description
OF_1 (0 transfers)	RMSN error of simulated and observed passenger trips with zero transfers
OF_2 (combined)	RMSN error of the simulated and observed passenger transfers at stops. This includes a combined list of 'direct', 'alight walk' and 'walk board' transfer types

OF_1 is similar to the objective function used in the synthetic network calibration. However, a small change is made for OF_2. In order to capture the error of all different transfer types, a combined list of transfer types is considered. A description of all different transfer movements is given in Table 3-1.

The same objective functions were used in both SPSA-SOF and SPSA-DOF and implemented as described in 3.2.3.

5.4.4 Assumptions

For the simplicity and based on the logical conclusions made from the available data, there are several assumptions behind the implementation:

- The transit demand is considered as stop level demand.
- Direct assignment Visum model is considered as the ground truth.
- Traffic conditions are not considered.
- Transit demand OD matrix is fixed.
- Capacity of the transit mode is not considered.

5.5 Results

5.5.1 Calibrated model parameter coefficients

The result obtained from the implementation of SPSA-SOF and SPSA-DOF on the Singapore network is summarized in this subsection. Due to the higher computational time required for transit assignment, the number of experiments was limited compared to the synthetic network, as explained in Figure 5-2. The best set of parameter coefficients for the calibrated model were chosen after analyzing the reduction of RMSN error. The accuracy of the final calibrated model is checked against the default values of PTV Visum, and initial guess used to start the calibration.

Figure 5-5 shows a comparison of the change of RMSN error for different simulator outputs with the first run of the implementation of SPSA-SOF and SPSA-DOF. Since the magnitudes of RMSN error for different simulator outputs vary, min-max normalized RMSN values for each simulator output is shown in the plots. In general, both SPSA-SOF and SPSA-DOF reduce the RMSN error for all the simulator outputs, except for the ‘alight walk’ transfer type. When comparing the pattern of error reduction, SPSA-DOF performs better than SPSA-SOF. This is noticeable in subplots (a.1) and (a.2). SPSA-DOF calibration approach has minimized all the error terms towards the latter parts of the iterations while the error reduction of SPSA-SOF is somewhat inconsistent. Thus SPSA-DOF is selected as the best calibration methodology. The best estimates from the first run with SPSA-DOF are chosen by evaluating the change of the RMSN error for a selected set of simulator outputs. The set of estimates which provided the lowest total RMSN are chosen as the best estimates. These estimates are used as the initial guess for the second run of the calibration.

The reduction in RMSN error for the second run (algorithm restart) is shown in Figure 5-6. The min-max normalized error of different simulator outputs is shown in this plot, like Figure 5 5. The algorithm is run only for 20 iterations due to the long computational time required. Passenger trips on different transit line routes and passenger transfers at stops are shown in two subplots. In general, a reduction in the error can be seen, but the reduction is wiggly compared to the first run. The final calibrated parameter estimates were obtained by evaluating the RMSN error from this calibration.

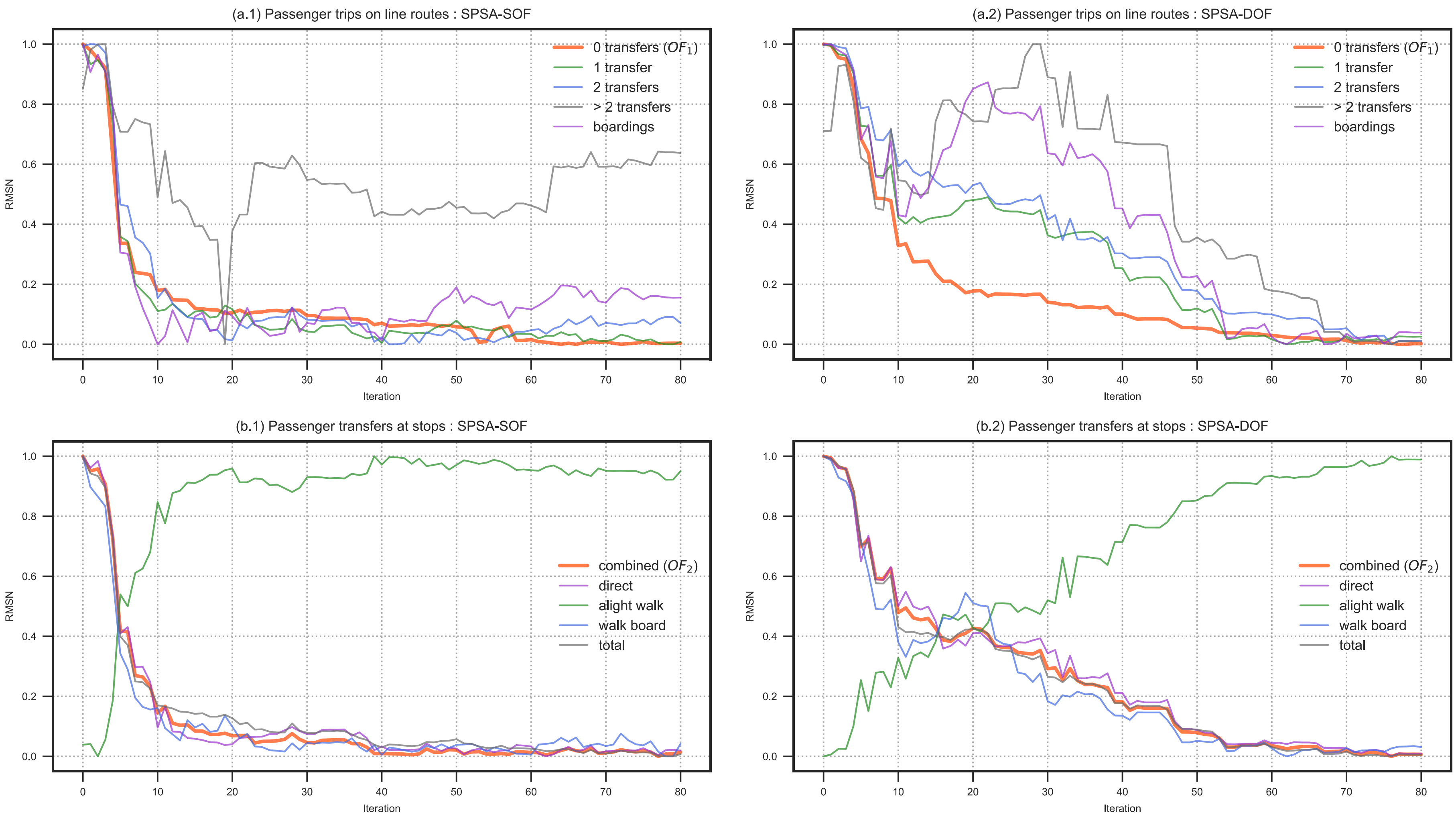

Figure 5-5 Singapore network – 1st run: RMSN change for the selected Visum outputs

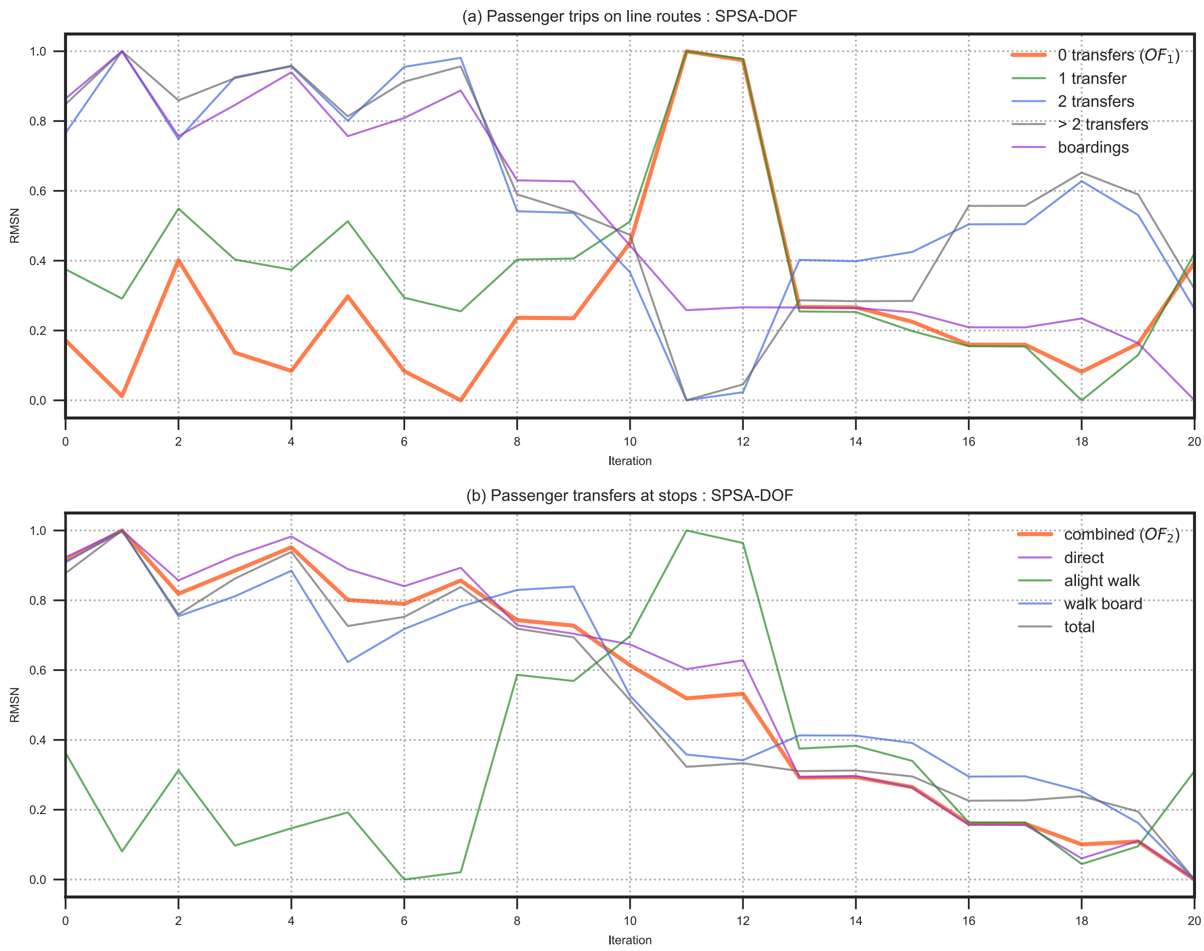

Figure 5-6 Singapore network – 2nd run: RMSN change for the selected Visum outputs

A summary of the calibrated coefficient for each parameter is shown in Table 5-5. The final set of calibrated parameter coefficients and transfer penalty value is highlighted in gray.

Table 5-5 Summary of PJT calibrated values for the Singapore network

PJT parameter	Visum default	Initial guess	Best estimate – 1st 80 iterations		Best estimate – 2nd 20 iterations	
			SPSA-SOF	SPSA-DOF	SPSA-SOF	SPSA-DOF
IVT	*1.000*	*1.000*	*1.000*	*1.000*	-	*1.000*
OWT	1.000	2.000	1.004	1.032	-	1.071
WT	1.000	2.000	4.483	3.971	-	3.615
TWT	1.000	2.000	1.047	1.016	-	1.117
TP	5.000	5.000	6.045	5.507	-	5.491
Best estimate at Iteration:			*79*	*76*	-	*20*

5.5.2 Comparison of calibrated model outputs

The calibrated model's simulation outputs are compared with the results of the observed model (direct assignment). A similar comparison is made among two other simulation models; 1. a model with Visum's default PJT values and 2. a model with the initial values used for the calibration. First, an aggregate level comparison made with Visum's transit assignment statistics (PuT assignment statistics), provided in Table 5-6. A detailed explanation of each assignment statistics is given in Appendix-C.

The first subsection of the table compares the average travel distances and travel times. From a quick comparison, the values of non-calibrated models (Visum defaults, initial guess) are also in close agreement with the observed values. For example, mean in-vehicle time (1.2) and mean in-vehicle distance (1.5) remains more or less the same in all three simulation models. One of the reasons for this behavior is the structure of the network. Since a significant component of transit travel is made inside a transit mode, the travel distances and actual time spent on the network should not be changed significantly regardless of the PJT parameters chosen.

The second subsection of the table compares the passenger transfers made in the network. The total number of passenger transfers made (2.1) in the calibrated model is in close agreement with the observed model, compared to non-calibrated models. As a result, the mean number of transfers made in the transit network (2.3), which is the division of total transfers (2.1), and the number of passengers assigned (2.2) is also in close agreement with the observed data.

The third subsection of the table compares the direct trips and trips made with transfers. The direct trips made in the network (3.1) are in close agreement with the observed data in the calibrated model compared to non-calibrated models. A similar observation can be seen in passenger trips with one transfer (3.2) and two transfers (3.3), but simulation output in the calibrated model is not in a close representation with the observed values. None of the simulation outputs can accurately represent the passenger trips with more than two transfers (3.4). However, compared to the total number of trips made in the network, the passenger trips more than two transfers are very low (2.19%). The fourth and final subsection of the table shows a comparison between assigned and unassigned trips in the network. Some number of unassigned trips in all three simulation models can be seen, but the value is insignificant compared to the total number of trips assigned.

Table 5-6 Comparison of results with transit assignment statistics

No.	Indicator (unit)	Observed	Simulated			Absolute percentage difference with observed values		
		DA	Visum defaults	Initial guess	Calibrated	Visum defaults	Initial guess	Calibrated
1. Average travel times and distances								
1.1	MeanPJTPuT (min)	-	34.93	63.25	43.97	-	-	-
1.2	MeanInVehTimePuT (min)	21.87	19.43	20.98	21.48	11.2%	4.1%	1.8%
1.3	MeanTrasferWaitTimePuT (min)	2.03	2.85	5.93	3.85	40.4%	192.1%	89.7%
1.4	MeanWalkTimePuT (min)	2.48	4.37	3.22	2.12	76.2%	29.8%	14.5%
1.5	MeanInVehDistPuT (km)	9.775	9.377	9.662	9.709	4.1%	1.2%	0.7%
2. Aggregated trips and transfers								
2.1	TotalNumTransfersPuT	25,607,438	21,942,281	24,085,762	24,334,223	14.3%	5.9%	5.0%
2.2	PTripsLinkedTot	41,533,332	41,834,266	41,832,622	41,832,260	0.7%	0.7%	0.7%
2.3	MeanNumTransfersPuT	0.600	0.525	0.576	0.582	12.6%	4.0%	3.0%
3. Trips with and without transfers								
3.1	PTripsLinked0	22,527,939	23,554,472	21,627,906	21,986,300	4.6%	4.0%	2.4%
3.2	PTripsLinked1	13,403,972	14,214,918	16,002,017	15,314,319	6.1%	19.4%	14.3%
3.3	PTripsLinked2	4,691,931	3,460,118	3,686,436	4,082,997	26.3%	21.4%	13.0%
3.4	PTripsLinked>2	909,490	268,543	236,694	283,975	70.5%	74.0%	68.8%
4. Assigned trips, unassigned trips								
4.1	PTripsLinkedWRide	41,533,332	41,498,052	41,553,052	41,667,592	0.1%	0.0%	0.3%
4.2	PTripsLinkedWoCon	-	2,760	4,404	4,766	-	-	-

In summary, the calibrated model outperforms the non-calibrated models in terms of the agreement of the simulated values with the observed values. Something interesting to notice here is that the default values used in the simulator itself provide a relatively satisfactory solution. This behavior could be misleading at times, especially when using a trial and error method as it could provide a satisfactory (but not optimum) solution with a couple of 'guesses'. This proves the importance of using a systematic calibration approach to solve the transit assignment calibration problem.

It is also essential to examine the agreement of simulation outputs with observed data at a disaggregated level, which helps to understand the level of accuracy of the simulation model outputs to the observed data at each line route/ transit stop. The results of the calibrated model are compared with the non-calibrated models.

Figure 5-7 shows a comparison between the calibrated model and non-calibrated models with respect to total boardings on transit line routes and the total number of transfers made at stops. For line routes, the bus mode and rail-based modes (LRT, MRT) shown in two separate graphs considering the differences in passenger volumes. Higher accuracy can be seen in the calibrated model for the passenger boarding on bus line routes. Compared to non-calibrated models, the calibrated model shows a higher r^2 value and lower intercept. The accuracy of passenger boardings on rail-based modes is more or less similar across all three models. However, a clear drop in the intercept value can be seen in the calibrated model. The accuracy of the total passenger transfers remains more or less the same across all three models, with an improved r^2 value in the calibrated model. The total number of transfers is further analyzed with different transfer types.

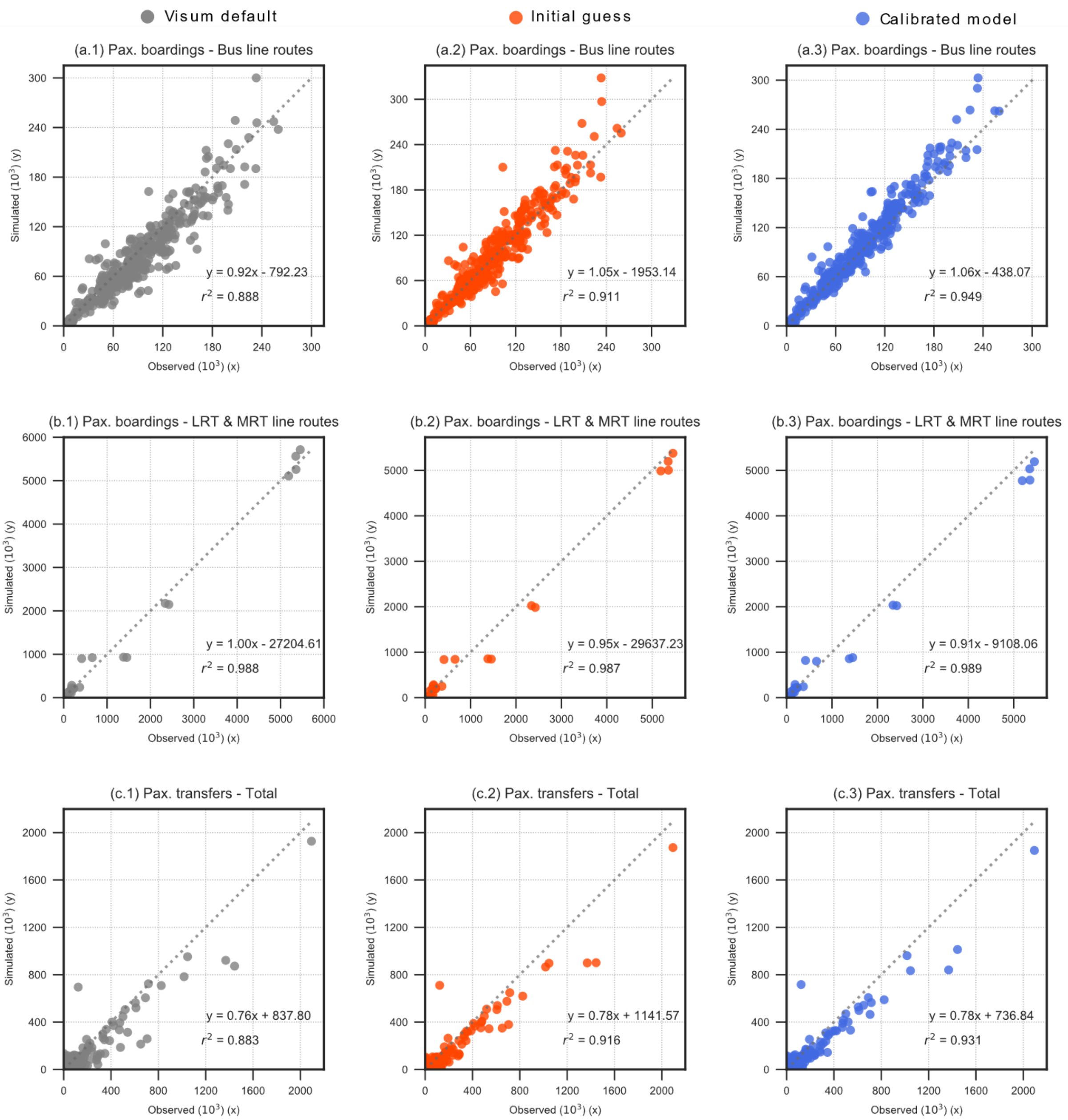

Figure 5-7 Comparison of the calibrated model with non-calibrated models: total boardings and transfers

The accuracy of the calibrated model with different transfer types is presented in Figure 5-8. Compared to the previous plot on the total number of transfers, the scatter plots for different transfer types, especially for alight walk (a) and walk board (b), are scattered around 45° line with an underrepresentation of the simulated values. This underrepresentation can be explained by the lesser number of total transfers simulated by all the simulation models (refer Table 5-6, subsection 2). However, the calibrated model shows a higher r^2 value compared to the non-calibrated models.

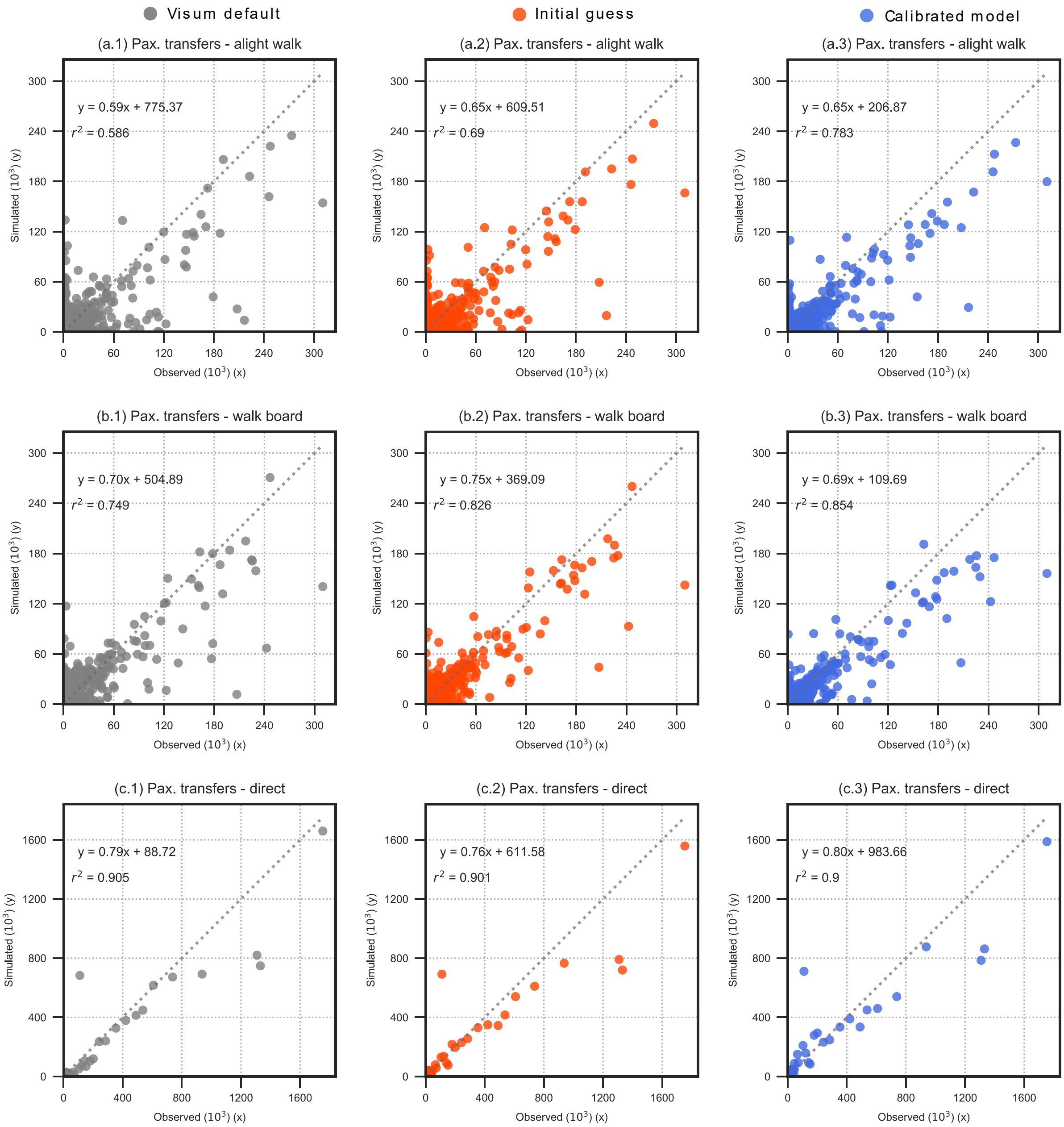

Figure 5-8 Comparison of the calibrated model with non-calibrated models: transfer types at stops

The passenger trips on transit line routes by the mode and by the number of transfers made are presented in Figure 5-9. Once again, trips made on bus and rail modes (LRT, MRT) shown separately considering the difference in volumes. For simplicity, only the calibrated model is considered for this comparison, as it is clear from the previous plots that the calibrated model is in higher agreement with the observed model compared to the non-calibrated models.

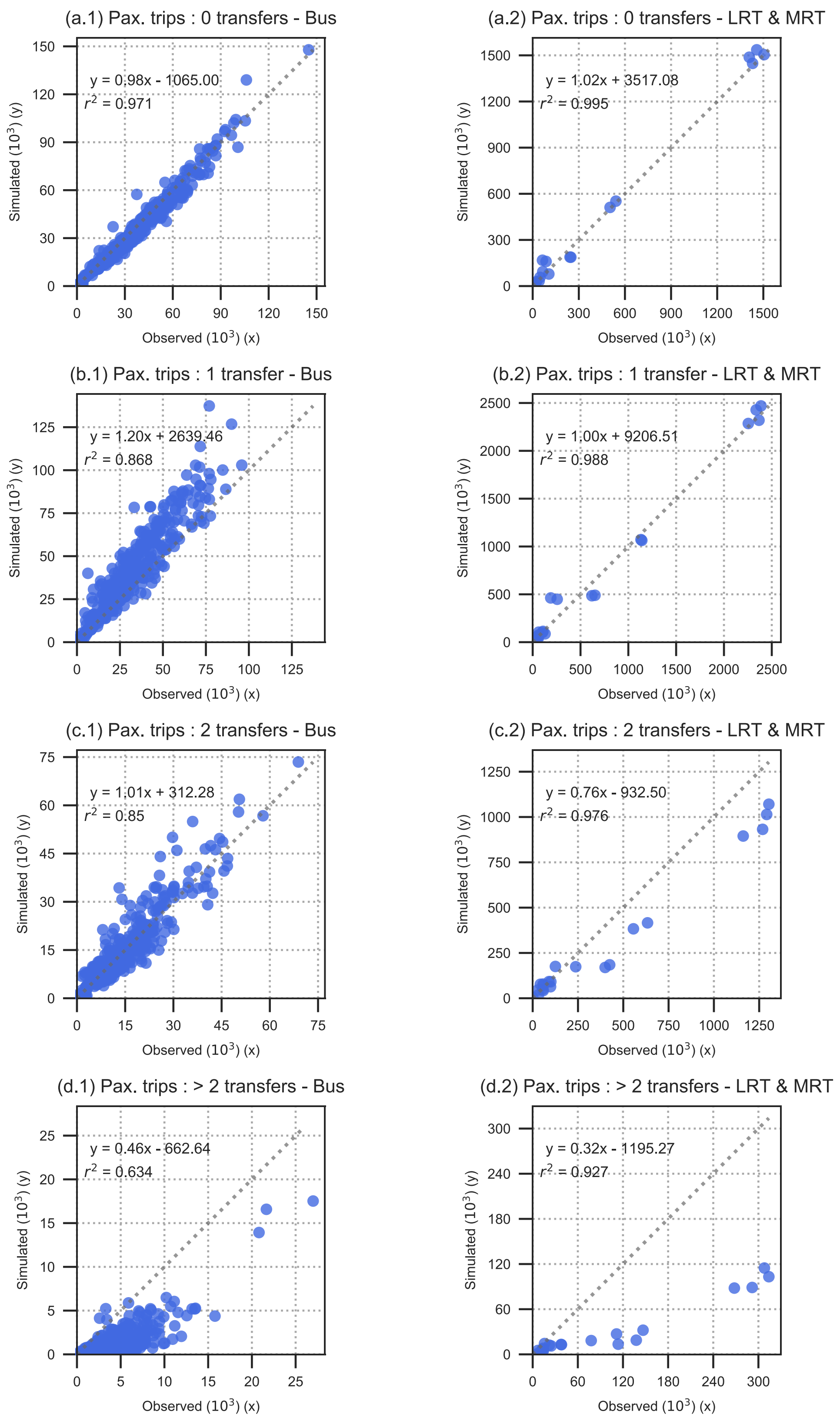

Figure 5-9 Calibrated model: passenger trips on transit line routes with transfers

Passenger trips with no transfers, which is one of the measurements used for objective function calculation, is in higher agreement with the observed values for both bus and rail modes. The calibrated model shows an over-assignment for passenger trips with one transfer for bus mode (b.1). In contrast, the calibrated model shows an under-assignment for the passenger trips with more than two transfers for both bus and rail modes. This observation was already seen in the aggregated results in Table 5-6.

5.6 Summary

The proposed calibration methodology is implemented in a real-world transit assignment model of Singapore. The number of experiments had to be limited, especially for restarting the algorithm, considering the computation time required to run the calibration. The results show that SPSA-DOF performs better in terms of reducing the RMSN error of all the selected simulator outputs, except for passenger transfers alight walk. The calibrated model parameter estimates show that passengers perceive transfer walk time almost 3.6 times more than in-vehicle time. The other wait parameters (origin wait, transfer wait) remains closer to one, suggesting that passengers do not perceive a considerable difference in waiting. The calibrated model outputs were compared with the observed model, and non-calibrated models (initial estimates, Visum default estimates), in two different aggregation levels. In both levels, the calibrated model produces a better fit with the observed values. However, the calibrated model does not accurately represent passenger trips greater than two transfers. The non-calibrated models also showed some level of accuracy with the observed data. This shows that the calibration is hard to perform with ‘trial and error’ method as it is hard to keep track of whether the ‘guessed’ estimates produce a better fit or not. This proves the importance of using a systematic calibration approach for large scale transit models.

6 Conclusion

6.1 Summary

This thesis presented an automated offline calibration framework for TAM calibration. The proposed method is based on SPSA algorithm. A modification is proposed to the algorithm such that at a given iteration, it chooses an objective function that produces the highest relative error difference with the perturbed parameter estimates. The measurements of the selected objective function are then used in the gradient approximation step. Since the selection of the objective function is made in a ‘dynamic’ fashion with the available information at the given iteration, this method is called SPSA with dynamic objective function selection (SPSA-DOF). The performance of the SPSA-DOF is compared with the standard form of implementing SPSA algorithm with a single objective function (SPSA-SOF). Both algorithms were implemented in a synthetic network and real-world network. SPSA-DOF performed better in both case studies by reducing the RMSN error for most of the important simulation outputs, making the final calibrated model is at a higher level of agreement with the observed model. For the Singapore network, the final calibrated model parameter coefficients give a higher weight to the ‘transfer walk time’ parameter indicating that passengers perceive transfer walk times higher than the rest of the components of the trip. The final calibrated model reveals that the mean perceived journey time in the Singapore network is 44 minutes.

6.2 Discussion

One of the reasons for SPSA-DOF to perform better compared to the standard implementation (SPSA-SOF) can be interpreted as follows. The objective function usually does not capture or represent the entire relationship of the system inputs and outputs. Therefore, the gradient approximation is made based on incomplete information. The error terms used to formulate the objective function in this thesis (i.e., passenger transfers at stops, direct trips on transit line routes) are somewhat conflicting in nature. Therefore, at a given iteration, considering the impact on both error terms may add extra noise. One of the solutions to this problem is to rely only on one type of error measurement throughout the calibration. This may properly calibrate some simulator outputs but does not guarantee in reducing the error of all the important simulator outputs. SPSA-DOF has proven that it can mitigate the drawbacks mentioned above.

The proposed SPSA-DOF method is somewhat similar to the idea proposed by (Antoniou et al., 2015) with W-SPSA method. In W-SPSA, the idea is to exclude the negative influence of irrelevant measurements in the gradient approximation. A weight matrix is used for this process. SPSA-DOF, however, does not rely on a weight matrix. Instead, it selects the 'best' objective function that gives a better descent direction for the given iteration. Figure 6-1 shows how the selection of the objective function has varied with SPSA-DOF, for the 1st run of Singapore network calibration. The selection of the objective function at a given iteration is not pre-planned. Thus SPSA-DOF does not have an impact on the stochastic nature of SPSA algorithm.

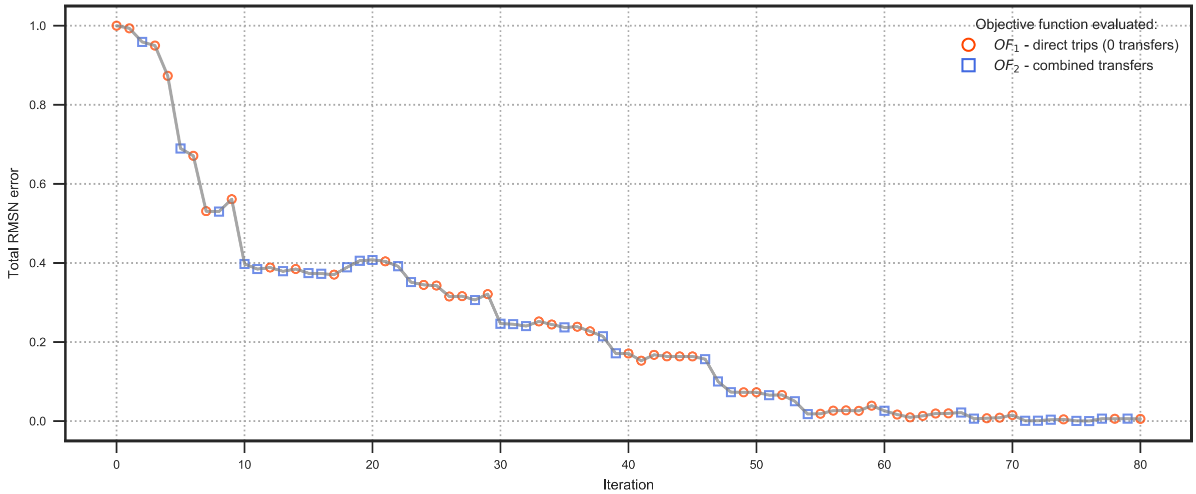

Figure 6-1 Objective function selection in SPSA-DOF

The final set of calibrated parameter coefficients reveals some behavioral characteristics of the passenger movements of Singapore. For example, a higher estimate for the transfer walk parameter shows that the passengers perceive the time spent on transfers that include a walking component (alight walk, walk board) substantially different to in-vehicle time or waiting times. This could be partly due to the fact that Singapore has higher temperatures and humidity levels. This makes a transfer between two stops, a physically demanding task, especially at day times, where there are no shaded walk paths. On the other hand, origin wait times and transfer wait times have a coefficient closer to one. This means that the waiting times at transit stops are perceived almost the same as in-vehicle time. This could be due to higher frequencies of the service and better connectivity of the network. Different transfer movements at stop levels (i.e., alight walk, walk board, and direct) also represent better fit to the observed data, but not as accurate as the passenger trips made on transit line routes. This could have been improved with further calibration runs (i.e., algorithm restart, more iterations in a single calibration run). However, part of this issue could be because of the ‘common line problem,’ and passenger movements in Singapore may not be fully explained by using ‘optimal strategies’ as a boarding strategy.

6.3 Limitations

The main requirement to calculate the minimum route share percentage for the Singapore network and set ‘ignore path share’ parameter accordingly was to reduce the computational time required for transit assignment. However, this came up with an unexpected drawback of an increased number of unassigned trips. One alternative approach to solve this issue could be to increase ‘Ignore path share’ parameter gradually and observe the number of unassigned trips while keeping track of the computational time required. The ‘Ignore path share’ could be set at the point where it starts to increase the number of unassigned trips sharply. Moreover, the second run (algorithm restart) could have been run for a higher number of iterations despite the extensive computational time required to see if the proposed method can further reduce the error between simulated and observed values, especially for different transfer types.

6.4 Future work

One of the important future work is to check the robustness of the proposed calibration approach. The algorithm can be restarted with different initial estimates and with different random seeds (to calculate random perturbation vector Δ_k) and check the convergence. In terms of the transit assignment model, a potential future work could be to model the congestion on board, considering the vehicle capacity. Moreover, the calibration procedure can be applied to different days (e.g., weekends), different times of the day (e.g., peak periods), and check if there is a difference in PJT parameter estimates. The same calibration methodology can be used with different boarding strategies other than ‘optimal strategies’ to see whether the selection of different boarding strategies would improve the calibration results. For a real-world network, this implementation will take more time than for a calibration done with ‘optimal strategies’. A better approach would be to implement it in a relatively smaller network (e.g., a part of the real network).

7 References

AECOM (2011). Study on Public Transport Smart Cards. Retrieved May 03, 2020, from https://ec.europa.eu/transport/sites/transport/files/themes/urban/studies/doc/2011-smartcards-final-report.pdf

Amaran, S., Sahinidis, N. V., Sharda, B., & Bury, S. J. (2016). Simulation optimization: a review of algorithms and applications. *Annals of Operations Research*, *240*(1), 351–380. https://doi.org/10.1007/s10479-015-2019-x

Antoniou, C., Azevedo, C. L., Lu, L., Pereira, F., & Ben-Akiva, M. (2015). W–SPSA in Practice: Approximation of Weight Matrices and Calibration of Traffic Simulation Models. *Transportation Research Procedia*, *7*, 233–253. https://doi.org/10.1016/j.trc.2015.04.030

Beirão, G., & Sarsfield Cabral, J. A. (2007). Understanding attitudes towards public transport and private car: A qualitative study. *Transport Policy*, *14*(6), 478–489. https://doi.org/10.1016/j.tranpol.2007.04.009

Ceder, A. (2007). *Public Transit Planning and Operation: Theory, modelling and practice*. Amsterdam, London: Butterworth-Heinemann.

Department of Statistics Singapore (2020). Population and Population Structure. Retrieved May 05, 2020, from https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data

Farrol, B., & Livshits, V. (1998). Analysis of individual transit trips in EMME/2: Calibration of 1996 TTC trips disaggregate assignment. Retrieved December 15, 2019, from http://www. jpint. utoronto. ca/PDF/transit_emme. pdf

Fung, W.-c. S. (2005). *Calibration and validation of transit network assignment models.*

Gentile, G., & Nökel, K. (2016). *Modelling Public Transport Passenger Flows in the Era of Intelligent Transport Systems*: Springer. https://doi.org/10.1007/978-3-319-25082-3

Kiefer, J., & Wolfowitz, J. (1952). Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, *23*(3), 462–466. https://doi.org/10.1214/aoms/1177729392

Kostic, B., Gentile, G., & Antoniou, C. (2017). Techniques for Improving the Effectiveness of the SPSA Algorithm in Dynamic Demand Calibration, 368–373. https://doi.org/10.1109/MTITS.2017.8005699

Land Transport Guru (2013). Bus Services - Singapore. Retrieved May 06, 2020, from https://landtransportguru.net/bus-services/

Liu, X., Zhou, Y., & Rau, A. (2019). Smart card data-centric replication of the multi-modal public transport system in Singapore. *Journal of Transport Geography*, *76*, 254–264. https://doi.org/10.1016/j.jtrangeo.2018.02.004

LTA (2013). Land Transport Master Plan 2013. Retrieved May 06, 2020, from https://www.lta.gov.sg/content/dam/ltagov/who_we_are/statistics_and_publications/master-plans/pdf/LTMP2013Report.pdf

LTA (2020). Plan your journey. Retrieved May 08, 2020, from https://www.lta.gov.sg/content/ltagov/en/getting_around/public_transport/plan_your_journey.html

McKinsey & Company (2018). Elements of success: Urban transportation systems of 24 global cities. Retrieved May 06, 2020, from https://www.mckinsey.com/business-functions/sustainability/our-insights/elements-of-success-urban-transportation-systems-of-24-global-cities

Meng, M., Rau, A., & Mahardhika, H. (2018). Public transport travel time perception: Effects of socioeconomic characteristics, trip characteristics and facility usage. *Transportation Research Part a: Policy and Practice*, *114*, 24–37. https://doi.org/10.1016/j.tra.2018.01.015

Mesplé, F., Troussellier, M., Casellas, C., & Legendre, P. (1996). Evaluation of simple statistical criteria to qualify a simulation. *Ecological Modelling*, *88*(1-3), 9–18. https://doi.org/10.1016/0304-3800(95)00033-X

Olstam, J., & Tapani, A. (2011). A Review of Guidelines for Applying Traffic Simulation to Level-of-service Analysis. *Procedia - Social and Behavioral Sciences*, *16*, 771–780. https://doi.org/10.1016/j.sbspro.2011.04.496

Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling Transport* (Fourth edition). Chichester, West Sussex, United Kingdom: John Wiley & Sons.

Papathanasopoulou, V., & Antoniou, C. (2015). Towards data-driven car-following models. *Transportation Research Part C: Emerging Technologies*, *55*, 496–509. https://doi.org/10.1016/j.trc.2015.02.016

Parveen, M., Shalaby, A., & Wahba, M. (2007). G-EMME/2: Automatic Calibration Tool of the EMME/2 Transit Assignment Using Genetic Algorithms. *Journal of Transportation Engineering*, *133*(10), 549–555. https://doi.org/10.1061/(ASCE)0733-947X(2007)133:10(549)

PTV (2019). *PTV VISUM 18 Manual.*

Qurashi, M., Ma, T., Chaniotakis, E., & Antoniou, C. (2020). PC–SPSA: Employing Dimensionality Reduction to Limit SPSA Search Noise in DTA Model Calibration. *IEEE Transactions on Intelligent Transportation Systems*, *21*(4), 1635–1645. https://doi.org/10.1109/TITS.2019.2915273

Rydergren, C. (2013). Comparison of headway-based public transport models. *Public Transport*, *5*(3), 177–191. https://doi.org/10.1007/s12469-013-0071-y

Spall, J. C. [J. C.] (1998). Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, *34*(3), 817–823. https://doi.org/10.1109/7.705889

Spall, J. C. [James C.] (2003). *Introduction to Stochastic Search and Optimization: Estimation, simulation, and control*: John Wiley & Sons, Inc. https://doi.org/10.1002/0471722138

Spiess, H., & Florian, M. (1989). Optimal strategies: A new assignment model for transit networks. *Transportation Research Part B: Methodological*, *23*(2), 83–102. https://doi.org/10.1016/0191-2615(89)90034-9

Tavassoli, A., Mesbah, M., & Hickman, M. (2019). Calibrating a transit assignment model using smart card data in a large-scale multi-modal transit network. *Transportation*, *143*(4), 4017003. https://doi.org/10.1007/s11116-019-10004-y

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, *1*(1), 67–82. https://doi.org/10.1109/4235.585893

Zhu, W., Hu, H., & Huang, Z. (2014). Calibrating Rail Transit Assignment Models with Genetic Algorithm and Automated Fare Collection Data. *Computer-Aided Civil and Infrastructure Engineering*, *29*(7), 518–530. https://doi.org/10.1111/mice.12075

8 Appendices

Appendix A: Pseudocode for SPSA-DOF

The pseudocode is written in two parts for the ease of elaboration. The first part provides a function to run the transit assignment with a given set of PJT estimates. The second part shows the pseudocode for SPSA-DOF implementation

Algorithm 1: Calculate RMSN error for all simulator outputs (RunAssignmentAndCalculateAllError)

```
Input Visum (Visum object), Estimates (PJT estimates), ObsStops(observed stops data),
ObsLines(observed line data)
RunVisumTransitAssignment (Visum, Estimates)
SimStops = SimulatedStopVolumes (Visum)
SimLines = SimulatedLineVolumes (Visum)
ErrorCollection = CalculateRMSNError(ObsStops, SimStops), CalculateRMSNError(ObsStops, SimStops)
Output = ErrorCollection(Collection of all the error terms)
```

Algorithm 2: SPSA-DOF

```
Input: Visum ( Visum object ), InitialGuess ( PJT estimates ), ObsStops ( Observed stops data), ObsLines
(Observed line data), SPSAHyperparameters (a, c, A, α, γ), MaxIterations
Objective Function ( OF1, OF2 )                          // Two objective functions
CurrentEstimate = InitialGuess
BestEstimate = InitialGuess
StartingErrorAll = RunAssignmentAndCalculateAllError ( Visum, InitialGuess, ObsStops,
                                                        ObsLines )
LeastError = StartingErrorAll [ OF1 ]

For k in range (MaxIterations) :
    ak = a / (A + k + a) ** α
    ck = c / (k+1) ** γ
    Deltak = GenerateRandomBernouliVector ( )
    IncreaseEstimate = Copy ( CurrentEstimate )        // Initializing a value for further processing
    DecreaseEstimate = Copy ( CurrentEstimate )

    For i in len (IncreaseEstimate) :
        IF MinThreshold ≤ CurrentEstimate [ i ] + ck * DeltaK [ i ] ≤ MaxThreshold :
             IncreaseEstimate [ i ] = CurrentEstimate [ i ] + ck * Deltak [ i ]
        Else:
             IncreaseEstimate [ i ] = CurrentEstimate [ i ]
        End IF

    For j in len ( DecreaseEstimate ):
        IF MinThreshold ≤ CurrentEstimate [ j ] – ck * Deltak [ j ] ≤ MaxThreshold :
             DecreaseEstimate [ j ] = CurrentEstimate [ j ] – ck * Deltak [ j ]
        End IF

    ErrorIncreaseAll = RunAssignmentAndCalculateAllError ( Visum, IncreaseEstimate,
                                                             ObsStops, ObsLines )
    ErrorDecreaseAll = RunAssignmentAndCalculateAllError ( Visum, DecreaseEstimate,
                                                             ObsStops, ObsLines )
    CurrentObjFunc = SelectBestObjectiveFunction ( ErrorIncreaseAll, ErrorDecreaseAll,
                                                   ObjectiveFunctions )
    ErrorIncreasek = ErrorIncreaseAll [ CurrentObjFunc ]
    ErrorDecreasek = ErrorDecreaseAll [ CurrentObjFunc ]
    Gradient = ( ErrorIncreasek – ErrorDecreasek ) / ( 2 * ck * Deltak )
    PreviousEstimate = Copy ( CurrentEstimate )

    For m in len ( PreviousEstimate ):
            IF MinThreshold ≤ PreviousEstimate [ m ] – ak * Gradient [ m ] ≤ MaxThreshold :
                    CurrentEstimate [ m ] = PreviousEstimate [ m ] – ak * Gradient [ m ]
            Else:
                    CurrentEstimate [ m ] = BestEstimate [ m ]

    NewErrorAll = RunAssignmentAndCalculateAllError (Visum, CurrentEstimate, ObsStops,
                                                      ObsLines)
    IF NewErrorAll [ OF1 ] < LeastError :
            LeastError = NewErrorAll [ OF1 ]
            BestEstimate = Copy ( CurrentEstimate )
    End IF

Output: ResultFromEachIteration ( NewErrorAll, CurrentEstimate )
```

Appendix B: Transit network performance measurements – PTV Visum

Indicator	**Description**
Passenger trips unlinked PuT	Line trips correspond to the number of passenger boarding per object (line, route, operator, etc.) Counts each passenger using at least one line route item in the territory. A passenger is not counted f he has already used a vehicle journey of the same object on an earlier path leg. No passengers are counted for path legs that end exactly at the start or start exactly at the end of a time interval.
PTrips Unlined with 0×Tranfer	Passenger trips with no transfers on the path from the origin zone to the destination zone. This attribute is only available for elements of the line hierarchy
PTrips Unlinked with 1×Tranfer	Passenger trips with precisely one transfer on the path from the origin zone to the destination zone. This attribute is only available for elements of line hierarchy
PTrips Unlinked with 2×Tranfer	Passenger trips with precisely two transfers on the path from the origin zone to the destination zone. This attribute is only available for elements of line hierarchy
PTrips Unlinked > 2×Tranfer	Passenger trips with more than two transfers on the path from the origin zone to the destination zone. This attribute is only available for elements of line hierarchy
PassTransAlightWalk	Number of passengers alighting at this stop or stop point and walking to another stop or stop points for transfer. This attribute is only available for stops and stop points
PassTransWalkBoard	Number of passengers boarding at this stop or stop point after walking from another stop or stop point. This attribute is only available for stops and stop points
PassTransDir	Number of passengers transferring to another line at this stop or stop point. This attribute is only available for stops and stop points
PassTransTotal	Number of passengers transferring at this stop or stop point PassTranferTotal = passenger transfers directly + passenger transfers walk-board + passenger transfers alight-walk

Appendix C: Transit (PuT) assignment statistics – PTV Visum

Indicator	Definition
MeanPerceivedJourneyTimePuT	Mean perceived journey time of all PuT trips
MeanInVehTimePuT	Mean in-vehicle time (only path legs using PuT service) of all PuT trips
MeanTrasferWaitTimePuT	Mean transfer wait time (without transfer walk times) of all PuT trips
MeanWalkTimePuT	Mean transfer walk time of all PuT trips
MeanInVehDistPuT	Mean in-vehicle distance PuT
TotalNumTransfersPuT	Total number of transfers of all PuT paths
PTripsLinkedTot	Number of passengers assigned, consisting of the number of passengers with transport (on lines, Sharing or PuT Aux), those without transport and those without a connection. Each person only counts once
MeanNumTransfersPuT	Mean number of transfers of al PuT trips
PTripsLinked0	Number of passengers who have boarded a vehicle journey/time profile of a line of vehicles of a Sharing or PuT Aux transport system exactly once, i.e., without any transfers
PTripsLinked1	Number of passengers who have boarded a vehicle journey/time profile of a line of vehicles of a Sharing or PuT Aux transport system exactly twice, i.e., with one transfer. Each person counts only once
PTripsLinked2	Number of passengers who have boarded a vehicle journey/time profile of a line of vehicles of a Sharing or PuT Aux transport system exactly three times, i.e., with two transfer. Each person counts only once
PTripsLinked>2	Number of passengers who have boarded a vehicle journey/time profile of a line of vehicles of a Sharing or PuT Aux transport system more than three times, i.e., with more than two transfers. Each person counts only once
PTripsLinkedWRide	Number of passengers who have boarded vehicle journeys/ time profiles of a line of vehicles of a Sharing / PuT Aux transport system, where multiple boarding on the same path is counted only once
PTripsLinkedWoCon	Passenger trips linked without a connection: Number of PuT passenger for whom no connection was found from the origin zone to the destination zone