ARTICLE

# Quantifying service-reliability-based day-to-day evolution of travel choices in public transit systems with smart transit card data

Mingyou Ma[a], Wei Liu[a,b,⋆], Xinwei Li[c], Fangni Zhang[d], Sisi Jian[e] and Vinayak Dixit[a]

[a] Research Centre for Integrated Transport Innovation, School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia; [b] School of Computer Science and Engineering, University of New South Wales, Sydney, Australia; [c] School of Economics and Management, Beihang University, Beijing, China; [d] Department of Industrial and Manufacturing Systems Engineering, University of Hong Kong, Hong Kong, China; [e] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

**ARTICLE HISTORY**
Compiled February 24, 2021

**ABSTRACT**
Existing studies proposed different models to capture day-to-day evolution of travelers' choices and traffic dynamics. However, explicitly incorporating public transit service unreliability/reliability into day-to-day traffic modeling received very little attention. This study develops day-to-day models to explore how the unreliability of public transit service affects the day-to-day evolution of travel choices made by transit users in the Greater Sydney area. In particular, we consider two dynamical processes that incorporate transit service unreliability, i.e., travelers' learning and perception updating process (LPUP) and proportional-switch adjustment process (PSAP). The conditions for the existence, uniqueness and stability of the fixed point of each model are analytically derived. These conditions are then examined using real-world public transit data from the Greater Sydney area. We find that with some aggregations and approximations, the system stability conditions at the fixed point are satisfied in both models. The observed weighted average flow change between two successive days is around 6.5% over the observation period, which may reflect the system stochasticity rather than instability. Among a series of empirical findings, it is noteworthy that in the Sydney case, the value of service schedule delay (late for schedule) is worth around 3.27 times of the in-vehicle time. Moreover, we find that the LPUP model more closely approximates real day-to-day travel choices than PSAP model, which reflects its stronger capability to capture the non-linear effects between cost perception and travel choices.

**KEYWORDS**
Reliability; Day-to-day; Public transit system; Stability; Smart transit card

## 1. Introduction

The day-to-day evolution of travel choices (e.g., departure time choice and mode choice) are impacted by various factors. Reliability of transport services is one of these factors, which significantly affects their attractiveness and further influences the travel demand. Many studies investigated how the variability of travel time and punc-

⋆Corresponding author: Wei Liu (wei.liu@unsw.edu.au)

tuality of service could affect mode choice (Prashker, 1979; Van Loon et al., 2011). Some studies focused on the impact of travel time reliability on route choice in road or rail networks (Jackson and Jucker, 1982; Levinson and Zhu, 2013; Small et al., 2005; Xu et al., 2018) and departure time choice (Li et al., 2009a; Tu et al., 2012). In particular, for public transit systems, the discrepancy between planned transit schedule and realized schedule yields uncertain waiting times and trip times (Bates et al., 2001). To accommodate this uncertainty and avoid being late, travelers may depart earlier for a certain length of time. This additional time can be regarded as a safety margin, which depends on individual's risk-taking behavior and trip purpose. A series of studies proposed the "effective travel time" when considering reliability-based travel choice, which consists of the expected travel time and the safety margin (Lo et al., 2006; Shao et al., 2008; Siu and Lo, 2008). Szeto et al. (2011) further generalized the concept of effective travel time by defining the effective travel cost, which includes the in-vehicle travel time cost and waiting time cost. It is based on the consideration that there is a difference between monetary values of travel time and that of waiting time, where waiting time is worth around 2.5 to 3.5 times of in-vehicle time (Wardman, 2001).

Apart from considering the service reliability from the transit user's perspective, there is also a branch of studies focusing on the reliability of public transit networks from a system perspective. Carey (1994) developed several metrics of unreliability, with the consideration of costs due to early or late departure/arrival services. Later, Rietveld et al. (2001) expanded the valuation of unreliability costs to a multi-modal perspective by addressing delays due to missing connections in public transport chains. Recent empirical studies suggested that transit services are more likely to be perceived as unreliable by passengers if scheduling delays, service headway and route length are larger (Carrel et al., 2013; Chen et al., 2009; Habib et al., 2011). To capture the effect of transit service uncertainties on users' travel choice behavior, a number of studies establish stochastic models within the framework of reliability-based stochastic user equilibrium (Jiang and Szeto, 2016; Li et al., 2009b; Szeto et al., 2013,1; Yang and Lam, 2006; Zhang et al., 2010). Due to the day-to-day variability of travel demand and traffic condition, users' perception of unreliability in public transit services also varies. However, to the best of our knowledge, there are very limited studies examining how transit service reliability/unreliability could affect users' day-to-day travel choices. This paper explores the day-to-day time-dependent travel choice evolution with the consideration of the reliability of public transit system in the Greater Sydney area.

Day-to-day traffic dynamics have been characterized as the system variations between successive periods, where the period can be either the entire day or a part of the day. The day-to-day models have two major categories, i.e., deterministic-process models and stochastic-process models. Specifically, Smith (1984) and Horowitz (1984) formulated the day-to-day dynamics of route choice in the context of deterministic process. The deterministic models can capture various kinds of User Equilibrium (UE), including Wardrop's UE, stochastic user equilibrium (Cantarella and Cascetta, 1995; Smith and Watling, 2016) and boundedly rational UE (Guo and Liu, 2011; Mahmassani and Chang, 1987; Ye and Yang, 2017; Yu et al., 2020; Zhu et al., 2019). Yang and Zhang (2009) summarized that in the context of Wardrop's UE, there are five major types of day-to-day adjustment processes: the simplex gravity flow dynamics, the proportional-switch adjustment process, the projected dynamical system, the network tatonnement process, and the evolutionary traffic dynamics. Unlike deterministic-process models, the steady state of a stochastic-process model refers to the equilibrium probability distribution (Cascetta, 1989; Cascetta and Cantarella, 1991; Davis

and Nihan, 1993; Watling and Cantarella, 2015). Moreover, applications of day-to-day dynamical systems are not restricted to modeling travelers' route choices. It has also been extended to modeling departure time choices (e.g., Ben-Akiva et al., 1984; Xiao and Lo, 2016) and mode choices (e.g., Li et al., 2018; Zhang and Liu, 2020).

For the departure-time-independent problems with deterministic day-to-day process, the system stability is well studied in the context of Wardrop's UE (e.g., Friesz et al., 1994; He et al., 2010; Smith, 1984; Zhang and Nagurney, 1996) and stochastic UE (e.g., Cantarella, 1993; Cantarella and Cascetta, 1995; Horowitz, 1984). Regarding the departure-time-dependent problems or dynamic user equilibrium, there is a large body of literature without looking at the day-to-day traffic variation (Friesz et al., 2013; Wang et al., 2018). However, Iryo (2008) found that, in the single Origin-Destination (OD) pair system with one bottleneck, the day-to-day dynamical system is unstable. Later, Guo et al. (2018) also demonstrated that the dynamic user equilibrium may not be reachable via a day-to-day evolution process in the context of bottleneck models. Taking the assumption that travelers are boundedly rational when choosing departure times, stable stationary points of such day-to-day evolution system exist and are identical to the boundedly rational UE points (Guo et al., 2017; Zhu et al., 2019). Recent studies further showed that additional information to travelers may help the system to converge to fixed points (Jin, 2020; Liu and Geroliminis, 2017; Liu et al., 2017). These findings on stationary points in day-to-day dynamical systems inspire us to explore the existence and other properties of such points in the real-world public transit system.

In particular, this study aims to investigate the impacts of unreliability/reliability of public transit service on the day-to-day evolution of travel choices made by Sydney public transit users. We carry out our exploration with two classic day-to-day dynamical systems, i.e., the learning and perception updating process (LPUP) and the proportional-switch adjustment process (PSAP). To account for the impacts of public transit service unreliability/reliability, we embed the effective travel cost (Szeto et al., 2011) into the day-to-day dynamical systems. We then analyze the existence, uniqueness, and stability of fixed points of the two systems, respectively. Moreover, the proposed day-to-day dynamical models are calibrated using the smart transit card (Opal card) data from the Greater Sydney area. This study is among the very few that test and evaluate the day-to-day models with real-world evidence (e.g., Cheng et al., 2019; Guo and Liu, 2011; He and Liu, 2012), while several other studies relied on experimental data (Xiao and Lo, 2016; Ye et al., 2018; Zhang et al., 2018). The dataset used in this study contains anonymous historical records of public transit journeys made by passengers in the Greater Sydney area over a three-month period (from 1 April 2017 to 30 June 2017). In order to compute the transit service schedule delays, transit schedule information based on General Transit Feed Specification data is also used (`https://opendata.transport.nsw.gov.au/dataset/timetables-complete-gtfs`).

The main contributions of this paper are summarized in the following. Firstly, with an emphasis on service reliability/unreliability, we propose two dynamical systems to model the day-to-day evolution of Sydney transit users' departure time choice, i.e., PSAP model and LPUP model mentioned earlier. The properties of the proposed two systems are analytically examined. It is noteworthy that the fixed points of these two models are different from traditional UE- or SUE-based fixed points due to the incorporation of service reliability/unreliability. Secondly, we calibrate the developed models with a large-scale real-world dataset from the Greater Sydney area and generate a series of empirical insights, including quantifying the value of the service schedule delay in transit services relative to the in-vehicle time, measuring the magnitude of

transit commuters' safety margins. We also found that the LPUP model with stochastic choices is more capable of capturing the day-to-day evolution of departure time choice in public transit systems. Thirdly, we empirically evaluate and compare the stability conditions of both models for the Sydney Greater area under some aggregations and approximations.

The rest of the paper is organized as follows. Section 2 describes the datasets utilized in this work. Section 3 details the formulation of effective travel cost, proposes two dynamical models, and analyzes the analytical properties of the models. Section 4 presents the case study and a series of sensitivity analyses. Data processing and model calibration are also introduced. Finally, Section 5 concludes the paper.

## 2. Data Description

We start with the data description. Smart transit card data can be used to explore the day-to-day regularity and variability of transit commuters' mobility patterns, e.g., Li et al. (2019) and Li et al. (2021). Many studies developed tools for understanding spatial and/or temporal characteristics of human travel behavior using public transport data (Ma et al., 2017; Morency et al., 2007; Sun and Axhausen, 2016). Such massive spatial-temporal mobility data brings opportunities to quantify day-to-day travel choices. In this context, this study develops and calibrates day-to-day travel choice evolution models with real-world evidence, and explores the capability of these day-to-day models to capture and reproduce reality.

This section briefly describes the transit smart card data used in this study and the General Transit Feed Specification (GTFS) data (for obtaining transit schedule information), along with the data output.

### *2.1. Transit smart card data*

Automated fare collection (AFC) systems are built into the public transit system in Sydney, including trains, metros, buses, ferries and light rail. Due to the distance-based fare scheme, passengers in Sydney need to not only tap on but also tap off in order to pay the correct transit fares for each journey. As a result, the spatial information (stop/station ID) and temporal information (boarding/alighting timestamp) of trips are stored in the transit card data. The data adopted in the study records anonymous transit trips made from 1 April to 30 June 2017. We summarize relevant attributes in the dataset in Table 1.

**Table 1.** Smart transit card data attributes

| Attribute | Definition |
|---|---|
| Journey segment start date | Date of the journey |
| Journey segment start time | The commencement time of the journey reflecting passengers' exact boarding time |
| Journey segment end time | The termination time of the journey reflecting passengers' exact alighting time |
| Transit stop ID | The ID of a transit stop where the journey segment started or ended |
| Geographic coordinate of transit stop | The latitude and longitude coordinates revealing the geographic location of a transit stop |
| Route ID | The name of the transit route the passenger took |
| Transfer indicator | The binary attribute indicating whether the journey is referred as a transfer |

### *2.2. GTFS data*

Besides the in-vehicle time, the service schedule delay (due to transit arriving late at the transit stops) is a key parameter in this paper to reflect the unreliability/reliability of transit services. To compute the service schedule delay, we measure the discrepancy between the scheduled arrival time of transit services and the actual boarding/alighting time (we use travelers' boarding/alighting time to approximate the transit arrival time at the bus stops).[1] The General Transit Feed Specification (GTFS) data provides public transit timetables (`https://opendata.transport.nsw.gov.au/dataset/timetables-complete-gtfs`). In particular, we use the static version of the Sydney GTFS data to identify planned transit service schedules. Table 2 summarizes the key attributes we use for delay computation.

**Table 2.** GTFS data attributes

| **Attribute** | **Definition** |
|---|---|
| Transit stop ID | The ID of a transit stop. Each transit stop has its own service timetable |
| Route ID | The name of a scheduled transit route in public transportation timetables |
| Scheduled service arrival time | The scheduled service arrival time at a specific stop for a specific trip on a route |

### *2.3. Data output*

Transit smart card data and GTFS data are processed and fused to output the data for model calibration and validation. The output data is a time series including the following features: (i) geographical information of selected Origin-Destination (OD) pairs, (ii) duration of peak-demand hours with respect to each OD pair, (iii) travelers' experienced in-vehicle time and service schedule delay on each observed day, and (iv) demand between OD pairs at each departure time interval (time horizon is discretized into multiple intervals).[2] We adopt the piecewise regression (McZgee and Carleton, 1970) to define the peak-demand duration for each OD pair. Overall, 171 OD pairs are generated and used in our case study, while additional OD pairs are also examined in the sensitivity analysis. Details about data processing methods and the associated techniques are summarized in Section 4.

## 3. Model Formulation

This study aims to quantify the day-to-day evolution of time-dependent travel choices made by transit users in the Greater Sydney area with the consideration of the reliability of public transit services. In the Greater Sydney area, the main performance metrics of public transit services (bus services) are the travel time and its distribution/uncertainty, while the transit fare is OD pair specific. This indeed motivates us

[1]Due to the absence of datasets describing the actual transit arrival time, we apply travelers' boarding/alighting time (recorded in the smart transit card data) to approximate the transit arrival time. We expect that the difference between the transit arrival time at the transit stop stand/platform and the time when the first traveler got on or alighted from that transit service is small and negligible.

[2]This study considers departure-time-interval-based travel demand. The OD demand (or observed passenger flow) at a specific time interval is the total number of travelers departing from the origin (the departure stop) within the time interval in concern (based on smart transit card tap-on records). We do not differentiate the arrival time of travelers when calculating the demand. Note that the arrival time is reflected in the travel time, which is a part of the travel cost.

to incorporate reliability-related attributes in relation to in-vehicle time and service schedule delay (i.e., waiting time due to late arrival transit service at the departure stop for passengers) into the day-to-day dynamical models.

In particular, we investigate and compare the applicability of two day-to-day models (i.e., travelers' learning and perception updating process, and proportional-switch adjustment process) that are originally built upon the stochastic user equilibrium and the deterministic user equilibrium. In the following, we first describe the transit users' day-to-day time-dependent travel choice problem with reliability consideration, followed by the formulation of the effective travel cost and the previous day's experienced travel cost. Then, the formulation of travelers' learning and perception updating process (LPUP model) and proportional-switch adjustment process (PSAP model) are introduced, respectively. Finally, the existence, uniqueness and stability conditions in relation to these two dynamical systems are discussed. Table 3 summarizes the main notations used for the model formulations. It should be noted that the model formulation is OD pair specific. We omit the index for the OD pair to ease the notation.

**Table 3.** List of main notations for the model framework

| | |
|---|---|
| $V$ | The set of time intervals ($m \in V$) |
| $M$ | The number of time intervals with respect to a single OD pair i.e., $M = \|V\|$ |
| $\varphi$ | The length of time interval |
| $w_m^H$ | The mean service schedule delay at the departure stop with respect to departing at time interval $m$ over the period $H$ |
| $s_m^{w,H}$ | The standard deviation of service schedule delay with respect to departing at time interval $m$ over the period $H$ |
| $t_m^H$ | The mean in-vehicle time with respect to departing at time interval $m$ over the period $H$ |
| $s_m^{t,H}$ | The standard deviation of in-vehicle time with respect to departing at time interval $m$ over the period $H$ |
| $w_m^q$ | The experienced service schedule delay at the departure stop with respect to departing at time interval $m$ on day $q$ |
| $t_m^q$ | The experienced in-vehicle time with respect to departing at time interval $m$ on day $q$ |
| $d$ | The demand between the OD pair |
| $\mathbf{D}$ | $= d\mathbf{I}$ is the demand matrix of the OD pair, where $\mathbf{I}$ is the identity matrix with a rank of $M$ |
| $\mathbf{x}^q$ | Column vector of time-interval-based passenger flow on day $q$, $\mathbf{x}^q \in R_+^M$ |
| $\mathbf{c}(\mathbf{x}^q)$ | $= \{\mathrm{c}(\mathrm{x}_m^q), m \in V\}^{\mathrm{T}} \in R_+^M$ is the column vector of the experienced travel cost on day $q$ |
| $\tilde{\mathbf{c}}^q$ | $= \{\tilde{\mathrm{c}}_m^q, m \in V\}^{\mathrm{T}} \in R_+^M$ is the column vector of perceived travel cost on day $q$ |
| $\mathbf{E}$ | $= \{\mathrm{E}_m, m \in V\}^{\mathrm{T}} \in R_+^M$ is the column vector of effective travel cost |
| $\mathrm{C}_m^{e,q}$ | The observed experienced cost of departing at time interval $m$ on day $q$ |

### *3.1. Problem description: travelers' day-to-day travel choices*

Consider an OD pair with public transit service. On each day, those travelers choosing transit also have to choose a departure time, which is between the earliest time $T_b$ and the latest time $T_e$. We discretize the departure time horizon into multiple time intervals. The travelers choose a departure time interval. In particular, we evenly discretize the peak demand duration into $M$ intervals, with an interval length of $\varphi$. The public transit service headway is expected to be smaller than the length of time intervals.

Since we focus on the unreliability of public transit services, the travel cost of users is formulated by in-vehicle time cost and service schedule delay cost. The service schedule delay here indeed reflects the level of punctuality or reliability of the public transit services. The transit fare is distance-based (this is the case in many cities including Sydney), which is constant for the same OD pair. Therefore, we did not include the fare in the cost formulations in the following subsections.

### *3.2. Effective travel cost*

We consider that travelers have information regarding the in-vehicle time distribution and service schedule delay distribution, which may come from information services and their long term experience. The concept of effective travel cost in the literature can then be used to capture reliability-based choices (see, e.g., Lam et al., 2008; Lo et al., 2006 and Shao et al., 2008). Following the effective travel cost proposed by Jiang and Szeto (2016), we propose the effective travel cost $\mathrm{E}_m$ with respect to departure time interval $m$ as follows:

$$\mathrm{E}_m = w_m^H + \eta_1 s_m^{w,H} + t_m^H + \eta_2 s_m^{t,H} \tag{1}$$

where $w_m^H$ is the mean service schedule delay, $s_m^{w,H}$ is the standard deviation of service schedule delay, $t_m^H$ is the mean in-vehicle time, $s_m^{t,H}$ is the standard deviation of in-vehicle time, and $\eta_1$ and $\eta_2$ are two parameters used to capture the size of the safety margin, which reflects how risk-averse the travelers are. We should have $\eta_1 \geq 0$ and $\eta_2 \geq 0$, which means that uncertainties or delays are unfavourable. The effective travel cost $\mathrm{E}_m$ is a linear combination of the mean perceptions of the level of transit service and safety margins for the unforeseen service unreliability.

The effective travel cost reflects transit travelers' general understanding of the level of service of the public transit system based on their long-term user experience or information provision. For each period of time e.g., a few months, the average level of transit services at time interval $m$ can be valued as the mean delay $w_m^H$ and the mean in-vehicle time $t_m^H$ experienced by travelers over the period $H$. By integrating the effective travel cost into day-to-day evolution models and calibrating the associated parameter in $\mathrm{E}_m$ with the Sydney transit smart card data, we can estimate the size of transit users' safety margin, thereby providing sensible metrics and managerial insights for public transit service planning and operation.

Since the effective travel cost is based on long-term historical knowledge, we assume that within a certain period $H$, the effective travel cost does not vary from day to day, which means that $\mathrm{E}_m$ is constant (based on information provision or long-term experience). It also serves as a reference point for travelers considering service uncertainty/unreliability. Generally, the effective travel time $\mathrm{E}_m$ is modeled as the historical average of the level of transit service plus travelers' safety margins. In the following section, the day-to-day evolution of choice patterns will be formulated with the consideration of the effective travel cost.

### *3.3. Experienced travel cost*

The experienced travel costs are dependent on the flow (e.g., Cascetta and Cantarella, 1993; Horowitz, 1984; Watling, 1999). This paper assumes that the cost of travelers with respect to a departure time interval $m$ is a function of the passenger flow within the time interval $m$, which are summarized in Assumption 1.

**Assumption 1.** *The travel cost of departing at time interval $m$ on day $q$ is a function of the travel demand within this time interval $\mathrm{x}_m^q$ i.e., $\mathrm{c}_m^q = \mathrm{c}(\mathrm{x}_m^q)$, where $q$ is the day index and $\mathrm{c}(\cdot)$ is a strictly increasing function.*

Assumption 1 means that we model the dependence between passengers' experienced cost of using transit and public transit passenger flows. However, we ignore the

dependence of transit cost on other traffic (e.g., car flow may affect speed of buses). Incorporating the impact of other traffic to capture multi-modal cost-flow relationships in the day-to-day dynamical system is left for future research, which will rely on an integrated analysis of multi-modal datasets.

Note that the $c(\cdot)$ is formulated as a mean travel cost function. The analytical relationship between the experienced travel cost and flow should be calibrated with real-world data. With the transit smart card data and the GTFS data, the experienced cost consisting of service schedule delay and in-vehicle time can be calculated directly. We define the observed experienced cost of departing at time interval $m$ as follows:

$$C_m^{e,q} = \omega w_m^q + t_m^q \tag{2}$$

where $w_m^q$ is the experienced service schedule delay at the departure stop on day $q$, $t_m^q$ is the experienced in-vehicle time on day $q$, and the parameter $\omega$ ($> 0$) represents the relative value of service schedule delay to the in-vehicle time. Note that we incorporate the relative coefficient $\omega$ rather than two specific coefficients. This is to reduce the number of parameters to be estimated in the case study. Also, as mentioned earlier, transit fare is constant for a given OD pair, which is not included in the cost formulation.

In summary, we have two distinguished expressions of the experienced travel cost, i.e., $c(x_m^q)$ and $C_m^{e,q}$. The analytical experienced cost formulation $c(x_m^q)$ defines the cost-flow relationship. The derivation of the existence, uniqueness and stability conditions at the fixed point in the proposed dynamical systems is built upon the cost-flow function. $C_m^{e,q}$ is the "observed" experienced cost ($w_m^q$ and $t_m^q$ can be observed), which will be utilized later to calibrate the proposed model and cost formulations. The calibrated parameter $\omega$ provides us the relative value of the service schedule delay against the in-vehicle time. Section 4.2 gives more details about calculating the observed experienced cost using the real-world dataset.

In the following subsections, we further formulate day-to-day dynamical systems, with an emphasis on service reliability, which is reflected by the variations in in-vehicle time and the service schedule delay.

### *3.4. Learning and perception updating process (LPUP)*

Travelers' learning and perception updating process (termed as LPUP model) in the context of stochastic user equilibrium has been proposed and examined by Davis and Nihan (1993). Following that, Cantarella and Cascetta (1995) derived the stability conditions for both link-based and route-based, discrete-time, deterministic day-to-day process. In this study, we follow a similar structure of LPUP model but incorporate the effective travel cost.

#### *3.4.1. Formulations*

On a typical day $q$, travelers update their mean perceived travel cost of departing at time interval $m$ as follows:

$$\tilde{\mathbf{c}}^q = \tilde{\mathbf{c}}^{q-1} + \kappa \left( \mathbf{c}(\mathbf{x}^{q-1}) - \mathbf{E} \right) \tag{3}$$

where $\kappa > 0$ is a dimensionless coefficient associated with the difference between previous day's experienced travel cost and the effective travel cost, i.e., $\mathbf{c}(\mathbf{x}^{q-1})$ and

**E**. The term $\mathbf{c}(\mathbf{x}^{q-1}) - \mathbf{E}$ reflects that travelers will compare their previous day's experience with the effective travel cost based on information provision or a certain period of travel experience (this period should not be too short in order to form an understanding of the system, e.g., several months). We take the time interval $m \in V$ as an example for explanation. $\mathrm{c}(\mathrm{x}_m^{q-1}) - \mathrm{E}_m > 0$ indicates that the travel cost of departing in time interval $m \in V$ experienced by travelers on day $(q-1)$ is higher than the effective travel cost that contains a safety margin or buffer time. Travelers' perceived cost for time interval $m \in V$ on day $q$ will then increase by $\kappa \left(\mathrm{c}(\mathrm{x}_m^{q-1}) - \mathrm{E}_m\right)$. The value of $\kappa$ reflects travelers' sensitivity to the new experience or how fast travelers are learning from new experience. A very large $\kappa$ means that a small change in the experienced travel cost can cause a considerable change of their perception.

Eq. (3) is a variation of the widely-used exponential smoothing approach to model travelers' learning behavior in the literature (see for example, Cantarella and Cascetta, 1995), where the perceived cost on day $q$ is updated based on the combination of weighted perceived travel cost on day $q-1$ and weighted experienced travel cost on day $q-1$. The effective travel cost **E** is integrated in travelers' learning and forecasting behavior, reflecting the impact of the transit travelers' general understanding of the level of service of the public transit system (which is regarded as a source of prior information) on their day-to-day "learning trajectory". We will discuss how this setting influences the existence, uniqueness and system stability conditions in Section 3.4.2.

We now move to the travelers' choice updating. On a typical day $q$, travelers have to choose a departure time interval. The proportion of travelers choosing departure time interval $m$ is

$$\mathrm{P}_m^q(\tilde{\mathbf{c}}^q) = \frac{\exp(-\theta \tilde{c}_m^q)}{\sum_{j=1}^{M} \exp(-\theta \tilde{c}_j^q)} \tag{4}$$

The above Logit-model choice formulation means that one can define the user utility for departing at time interval $m$ as follows: $\tilde{u}_m^q = -\theta \tilde{c}_m^q + \varepsilon_m,\ m \in V$, where $\theta > 0$, and $\varepsilon_m$ for different $m$ are error terms following identical and independent Gumbel distribution.[3] Travelers maximize their expected utility when making departure time choices. To ease the presentation, we define the choice probability vector in the following form:

$$\mathbf{P}(\tilde{\mathbf{c}}^q) = \left\{\mathrm{P}_1^q, ..., \mathrm{P}_m^q, ..., \mathrm{P}_M^q\right\}^{\mathrm{T}} \tag{5}$$

[3]The assumption that the error terms $\varepsilon_m$ are identically and independently distributed (i.e., the IID assumption) is a limitation of the proposed Logit-based day-to-day departure time choice model (i.e., the LPUP model). For two arbitrary time intervals that are close to each other, the IID assumption is more likely violated (as system traffic conditions for these two time intervals can be relatively correlated), while for two time intervals that are far from each other, IID assumption is less likely violated (system traffic conditions for the two time intervals are relatively independent). Besides, if the trip time (in-vehicle time) is relatively small when compared to the time difference between two time intervals, the cost correlation between the two time intervals can be relatively insignificant. In this study, the time interval length used to model day-to-day travel demand for different time intervals is 10 minutes (more details in Section 4), meaning that the (on average) difference between two consecutive time intervals is 10 minutes, and the difference between farther away time intervals is no less than 20 minutes. At the same time, regarding the 171 selected OD pairs used in the case study, the percentage of OD pairs with an in-vehicle travel time less than 20 minutes (30 minutes) is 62.0% (80.7%). The above means that a significant number of OD pairs involve a relatively small in-vehicle time, while the time difference between two non-consecutive time intervals is no less than 20 minutes. Note that while the Logit-model has limitations, it allows analytical tractability when analyzing the day-to-day dynamical system, which has been adopted in day-to-day departure time models and also other day-to-day traffic assignment models (see for example, Ben-Akiva et al., 1984; Bie and Lo, 2010; Xiao and Lo, 2016).

Additionally, we consider that a proportion of travelers may simply repeat their previous travel choices and do not consider changing their travel choices (due to user inertia or other factors), while others will re-consider their choices. Thus, the choice updating can be formulated as follows (Cantarella and Cascetta, 1995):

$$\mathbf{x}^q = \mathbf{D}(1-\rho)\mathbf{P}(\tilde{\mathbf{c}}^q) + \rho\mathbf{x}^{q-1} \tag{6}$$

where $\rho \in [0,1)$ is the ratio of travelers who repeat their travel choices and do not re-consider their travel choices from day to day, which is assumed to be time-invariant; $\mathbf{D}$ is the matrix regarding travel demand, which is assumed to be fixed to allow analytical derivation for the LPUP model. Note that Eq. (6) is formulated at an aggregate level rather than specifying which travelers will be those repeating their choices.

In summary, the LPUP system state on each day $q$ is described by the perceived travel cost of day $q$ (i.e., $\tilde{\mathbf{c}}^q$) and flow pattern of day $q$ (i.e., $\mathbf{x}^q$). In particular, $\tilde{\mathbf{c}}^q$ is updated based on the previous day's perceived cost $\tilde{\mathbf{c}}^{q-1}$, the previous day's experienced cost $\mathbf{c}(\mathbf{x}^{q-1})$ and the long-term travel experience (or historical information) $\mathbf{E}$. Then, the revised $\tilde{\mathbf{c}}^q$ shapes the flow pattern $\mathbf{x}^q$ on day $q$. Previous studies defined that in deterministic process models, the day-to-day dynamical system reaches the fixed-points when $\mathbf{x}^q = \mathbf{x}^{q-1} = \mathbf{x}^*$, where $\mathbf{x}^*$ is the traffic flow at the fixed point (Cantarella and Watling, 2016; Cascetta and Cantarella, 1993; Li et al., 2018). In order to empirically examine the state of the Sydney public transit system, we firstly derive the analytical conditions for the existence, uniqueness of the traffic flow at the fixed point in Section 3.4.2.

### *3.4.2. Existence and uniqueness of the fixed point*

**Assumption 2.** *The effective travel cost is bounded, i.e.,* $0 < \mathrm{E}_m < \mathrm{c}\,(d)\,,\ \forall m \in V$, *where* $d$ *is the total demand.*

Assumption 2 says that under an extreme (and unlikely) scenario where if all travelers of an OD pair jammed at a single time interval $m$ i.e., $\mathrm{x}_m = d$, the experienced travel cost would be larger than the effective travel cost, which is based on travelers' general understanding of the level of service of the public transit system with respect to the time interval $m$. Also, $\mathrm{E}_m$ is expected to be larger than 0.

Eqs. (3) and (6) form a deterministic dynamical system. At the fixed-point of the dynamical system, we should have $\tilde{\mathbf{c}}^q = \tilde{\mathbf{c}}^{q-1}$ and $\mathrm{x}^{q-1} = \mathrm{x}^q$. Additionally, the flow and the perceived travel cost at the fixed point $(\mathbf{x}^*, \tilde{\mathbf{c}}^*)$ should satisfy the following conditions:

$$\mathbf{c}(\mathbf{x}^*) = \mathbf{E} \tag{7}$$

$$\mathbf{x}^* = \mathbf{D}\mathbf{P}(\tilde{\mathbf{c}}^*) \tag{8}$$

As can be seen, the experienced cost $\mathbf{c}(\mathbf{x}^*)$ at the fixed point is equal to the effective travel cost $\mathbf{E}$. However, the flow $\mathbf{x}^*$ is based on the mean perceived travel cost $\tilde{\mathbf{c}}^*$, where $\tilde{\mathbf{c}}^*$ is generally different from the experienced cost $\mathbf{c}(\mathbf{x}^*)$. This means that the fixed point defined by Eqs. (7)-(8) does not correspond to the traditional Logit-based stochastic user equilibrium, due to service reliability consideration and the incorporation of effective travel cost $\mathbf{E}$ in the learning and choice updating mechanism. It is

noteworthy that existing studies have discussed similar observations (i.e., fixed points of the dynamical system are different from traditional UE- or SUE-based points) due to information provision (Bifulco et al., 2016).

The proof for the existence and uniqueness and the stability of the fixed point adopts a similar approach as those from, e.g., Cantarella and Cascetta (1995) and Li et al. (2018).

**Lemma 1.** *Let* $\mathbf{Jac}(\mathbf{P}(\tilde{\mathbf{c}}^*))$ *denote the Jacobian matrix of the choice probability vector* $\mathbf{P}(\tilde{\mathbf{c}}^*)$*, then* $\mathbf{J_p} = \mathbf{Jac}(\mathbf{P}(\tilde{\mathbf{c}}^*))$ *is negative semidefinite.*

**Proof.** The proof is given in Appendix A.1. □

Lemma 1 is in line with those in the literature (see for instance, Cantarella and Watling, 2016), which provides the basis for proving the uniqueness of the fixed point. Let $\mathbf{J_c} = \mathbf{Jac}(\mathbf{c}(\mathbf{x}^*))$ denote the Jacobian matrix of travel cost at the fixed point $\mathbf{c}(\mathbf{x}^*)$. Based on Assumption 1, $\mathbf{J_c}$ is a diagonal matrix with positive diagonal elements and is positive definite.

**Proposition 1.** *There exists a unique travel flow and perceived cost at the fixed point* $(\mathbf{x}^*, \tilde{\mathbf{c}}^*)$ *for the dynamical system defined in Eqs.* (3) *and* (6)*.*

**Proof.** The proof is given in Appendix A.2. □

From the proof of Proposition 1, the sufficient condition for the existence and uniqueness of the fixed point is that $\mathbf{J_p}$ is negative semidefinite (Cantarella and Watling, 2016). The adopted multi-nomial Logit choice model is a specific example that satisfies this condition. The above results and the following propositions will still be valid as long as the Jacobian matrix of the choice filter $\mathbf{J_p}$ is negative semidefinite even if it is not in the form of multi-nomial Logit.

### *3.4.3. Stability condition of the fixed point*

The Jacobian matrix of the transition function at point $(\mathbf{x}^{q-1}, \tilde{\mathbf{c}}^{q-1})$ can be defined, showing the structure formed by four $M \times M$ blocks:

$$\begin{aligned} \mathbf{J}_{\mathrm{LPUP}} &= \mathbf{Jac}\left[\Psi\left(\mathbf{x}^{q-1}, \tilde{\mathbf{c}}^{q-1}\right)\right] \\ &= \begin{bmatrix} \mathbf{I} & \kappa\mathbf{J_c} \\ (1-\rho)\mathbf{D}\mathbf{J_p} & \rho\mathbf{I} + \kappa(1-\rho)\mathbf{D}\mathbf{J_p}\mathbf{J_c} \end{bmatrix}_{2M\times 2M} \end{aligned} \tag{9}$$

The relationship between the eigenvalues of matrix $\mathbf{J}_{\mathrm{LPUP}}$ and matrix $\mathbf{D}\mathbf{J_p}\mathbf{J_c}$ satisfy the following Lemma 2.

**Lemma 2.** *For each of the* $M$ *eigenvalues* $\gamma_k$ *of matrix* $\mathbf{D}\mathbf{J_p}\mathbf{J_c}$*, two eigenvalues* $\lambda_k^{'} = \lambda_k$ *and* $\lambda_k^{''} = \lambda_{M+k}$ *of matrix* $\mathbf{J}_{\mathrm{LPUP}}$ *are the solutions to the quadratic function in Eq.* (10)*.*

$$\lambda^2 - \lambda\left[\rho + \kappa\left(1-\rho\right)\gamma_k + 1\right] + \rho = 0 \tag{10}$$

**Proof.** The proof is given in Appendix A.3. □

**Proposition 2.** *The dynamical system defined in Eqs.* (3) *and* (6) *is stable if and only if* $|\gamma_k| < \frac{2(1+\rho)}{(1-\rho)\kappa}$*.*

**Proof.** The proof is given in Appendix A.4. □

Proposition 2 indicates that the learning factor $\kappa$ and the fraction of travelers who repeat their travel choices and do not re-consider their travel choices from day to day $\rho$ affect the stability of the dynamical system Eqs. (3) and (6). A smaller $\kappa$ and/or a larger $\rho$ improve(s) the stability of the dynamical system. This implies that if travelers are not likely to be affected by their travel experience on the previous day, they will not change their perception significantly, thereby yielding a more stable system at the fixed point. Similarly, if more travelers repeat their choices they made on the previous day, fewer variations in flow patterns will occur, which also yields a higher level of stability.

### *3.5. Proportional-switch adjustment process (PSAP)*

The proportional-switch adjustment process (PSAP) is established by Smith (1984) and is built upon the Wardrop's user equilibrium which specifies that under the equilibrium condition, more costly alternatives are not used. Later on, the PSAP model has been extended to study the departure time choice problem (e.g., Guo et al., 2018). We consider a discrete time PSAP model in this paper, which is similar to Section 3.4; and again we apply the Jacobian-based approach to investigate the stability of the PSAP system.

#### *3.5.1. Formulations*

Similar to the LPUP model, we consider that travelers tend to favour the time interval $m$ more on the next day (day $q$) if the previous day's experienced cost $\mathrm{c}(\mathrm{x}_m^{q-1})$ is less than the effective travel cost $\mathrm{E}_m$, and vice versa. The swapping demand rate $r_m^q$ associated with time interval $m$ is

$$r_m^q = \alpha \left(\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1})\right) \tag{11}$$

where $\alpha > 0$. Note that $\alpha$ converts costs into a flow change percentage, which is not dimensionless. When $\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1}) > 0$, which means that the experienced cost is less than the effective travel cost (which equals the mean cost plus the safety margin or buffer), $r_m^q > 0$, i.e., more people will depart at time interval $m$. When $\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1}) < 0$, accordingly, $r_m^q < 0$. The formulation of the swapping demand rate in Eq. (11) is compatible with the perception and choice updating process in Eqs. (3) and (6), where $\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1})$ is also used. Later on, we will compare the two proposed models.

Again, we consider only a proportion of $(1-\rho)$ travelers will re-consider their choices each day, the flow associated with time interval $m$ on day $q$ can be updated as follows:

$$\mathrm{x}_m^q = \rho \mathrm{x}_m^{q-1} + (1-\rho)\, \mathrm{x}_m^{q-1} \left(1 + \alpha \left(\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1})\right)\right) \tag{12}$$

The proposed PSAP-based model is able to incorporate the elasticity of demand. In particular, the total demand for two consecutive days are not necessarily identical, i.e., $\sum_m \mathrm{x}_m^q \neq \sum_m \mathrm{x}_m^{q-1}$. This can be readily verified based on Eq. (12) since $\sum_m \left(\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1})\right)$ may not equal zero.

The PSAP system state on each day $q$ can be simply described by the flow $\mathbf{x}^q$. Specifically, $\mathrm{x}_m^q$ is updated based on the previous day's experienced cost $\mathrm{c}_m^{q-1}$ and the

long-term travel experience (or historical information) $\mathrm{E}_m$ with respect to time interval $m$.

*3.5.2. Existence and uniqueness of the fixed point*

We still adopt Assumption 2 for the following analysis regarding the existence and uniqueness of the fixed point defined by the PSAP-based dynamical system.

To ease the presentation, we define $\mathbf{F}(\mathbf{x}^{q-1})$ as follows:

$$\mathbf{F}(\mathbf{x}^{q-1}) = \left\{ \mathrm{x}_m^{q-1} \left(1 + \alpha \left(\mathrm{E}_m - \mathrm{c}(\mathrm{x}_m^{q-1})\right)\right), m \in V \right\}^{\mathrm{T}} \tag{13}$$

Eq. (12) forms a deterministic dynamical system, which can be further re-written in the following vector-matrix form:

$$\mathbf{x}^q = \rho \mathbf{x}^{q-1} + (1-\rho)\mathbf{F}(\mathbf{x}^{q-1}) \tag{14}$$

At the fixed-point of the dynamical system, we should have $\mathbf{x}^{q-1} = \mathbf{x}^q$. Additionally, the passenger flow and experienced travel cost at the fixed-point ($\mathbf{x}^*$) should satisfy the following conditions:

$$\mathbf{x}^* = \mathbf{F}(\mathbf{x}^*) \tag{15}$$

$$\mathbf{c}(\mathbf{x}^*) = \mathbf{E} \tag{16}$$

The PSAP-based fixed point defined by Eqs. (15) and (16) is similar to that under the LPUP model defined in Eqs. (7) and (8) in the sense that at the fixed point, the experienced travel cost $\mathbf{c}(\mathbf{x}^*)$ should be equal to the effective travel cost $\mathbf{E}$. This implies that the two dynamical models yield the same flow $\mathbf{x}^*$ at the fixed point if the flow solution $\mathbf{x}^*$ to Eq. (16) is unique. However, the PSAP-based dynamical system differs from the LPUP model in the sense that the day-to-day evolution processes are different, which is reflected by Eqs. (8) and (15).

We now turn to derive the analytical conditions for the existence, uniqueness of the traffic flow at the fixed point $\mathbf{x}^*$. Let $\mathbf{0} = \{\underbrace{0, ...0, ..., 0}_{M}\}^{\mathrm{T}}$ and let $\mathbf{d} = \{\underbrace{d, ...d, ..., d}_{M}\}^{\mathrm{T}}$.

**Proposition 3.** *Suppose when* $\mathbf{x} \to \mathbf{0}$, $\mathbf{c}(\mathbf{x}) < \mathbf{E}$ *and when* $\mathbf{x} \to \mathbf{d}$, $\mathbf{c}(\mathbf{x}) > \mathbf{E}$. *Then, there exists a unique travel flow at the fixed point* $\mathbf{x}^*$ *for the dynamical system defined by Eq.* (12).

**Proof.** The proof is given in Appendix A.5. □

*3.5.3. Stability condition of the fixed point*

At the fixed-point of the dynamical system defined by Eq. (12), $\mathrm{c}(\mathrm{x}_m^*) = \mathrm{E}_m, \forall m \in V$, the Jacobian matrix of the transition function at the fixed point ($\mathbf{x}^*$), is a $M \times M$

matrix:

$$\mathbf{J}_{\mathrm{PSAP}} = \mathbf{I} + \alpha(1-\rho) \begin{bmatrix} -\mathrm{x}_1^* \mathrm{c}'(\mathrm{x}_1^*) & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -\mathrm{x}_m^* \mathrm{c}'(\mathrm{x}_m^*) & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & -\mathrm{x}_M^* \mathrm{c}'(\mathrm{x}_M^*) \end{bmatrix}_{M \times M} \tag{17}$$

**Proposition 4.** *The dynamical system defined by Eq.* (12) *is stable if and only if* $-\frac{2}{\alpha(1-\rho)} < -\mathrm{x}_m^* \mathrm{c}'(\mathrm{x}_m^*) < 0$ $(\forall m \in V)$*, or equivalently,* $|\lambda_{\mathrm{f}}| < 1$*, where* $\lambda_{\mathrm{f}}$ *is the eigenvalue of the matrix* $\mathbf{J}_{\mathrm{PSAP}}$*.*

**Proof.** The proof is given in Appendix A.6. □

If $\alpha$ is smaller and/or $\rho$ is larger, $-\frac{2}{\alpha(1-\rho)}$ is smaller, and then $-\frac{2}{\alpha(1-\rho)} < -\mathrm{x}_m^* \mathrm{c}'(\mathrm{x}_m^*)$ is more likely to hold and the dynamical system defined by Eq. (12) is more likely stable based on Proposition 4. This also means that, a smaller $\alpha$, which indicates that travelers are not sensitive to the new travel experience, improves the overall system stability. The implication of Proposition 4 is in line with that of Proposition 2.

## 4. Case Study

We now discuss how to calibrate the two proposed dynamical systems with real-world transit data from the Greater Sydney area and GTFS data. As discussed in Section 2, the study period is from 1 April to 30 June 2017, and we have 171 OD pairs in total. By excluding all weekends and public holidays, the length of the study period is 59 days. We focus on the morning trips associated with bus services only. Let K denote the set of the selected OD pairs ($i \in \mathrm{K}$), where $|\mathrm{K}| = 171$. Let Q denote the set of observed days ($q \in \mathrm{Q}$), where $|\mathrm{Q}| = 59$. It should be noted that we let $q = 0$ represent the initial day in our dataset.

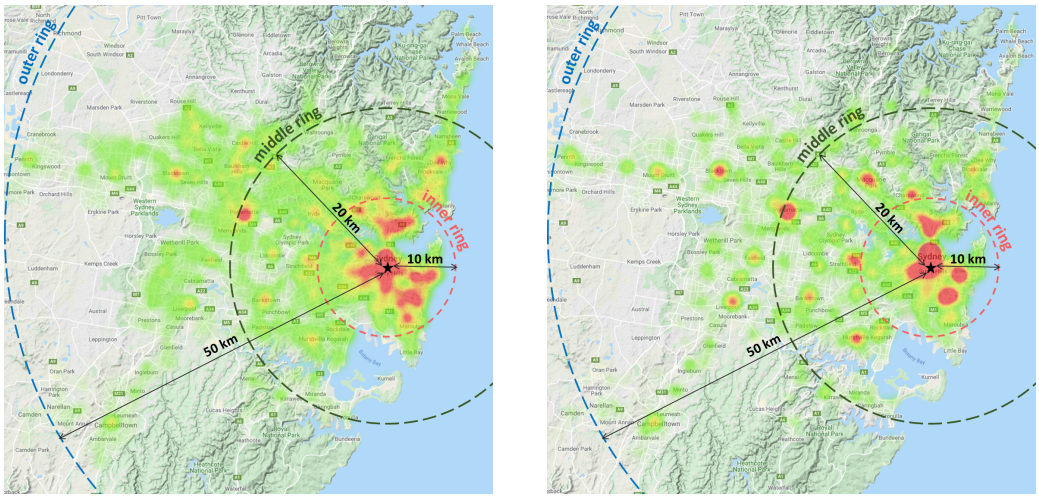
This section starts with describing the overall level of bus services in Sydney, and then describes the aggregation approach to fit the real-world trip records into the dynamical models. We then introduce more detailed techniques regarding the selection of OD pairs and identification of peak-demand hours for each OD pair. In the end of this section, we present the model calibration results and conduct a series of sensitivity analyses to investigate how the model settings and the variation in the data can impact the results.

### *4.1. Bus services in Sydney*

We start by exploring the spatial and temporal characteristics of bus trips in the morning. Fig. 1 visualizes the spatial patterns of bus commuters based on trip origins and destinations. As can be seen, demand patterns differ for inner-ring, middle-ring and outer-ring areas that are categorized by their distance to Sydney Central, where inner-ring area is within 10 km from Sydney Central, middle-ring is in between 10 to 20 km, and outer-ring area is from 20 to 50 km from Sydney Central (Fig. 1). In Fig. 1a and Fig. 1b, a darker colour (from light green to dark red) indicates a higher demand density. Regarding trip origins, the majority of bus trips were generated from

the inner Sydney; and few regions scattering around the boundary of the middle ring also produced a considerable number of trips. Most commuters' trip destinations fell in the core of the inner-ring area i.e., Sydney Central Business District (CBD). A large number of trips were also attracted to several industrial suburbs in middle-ring and outer-ring regions.

Fig. 2 further displays the changes of average daily demand with respect to each departure zone over the study period (i.e., from 1 April to 30 June 2017), where all the public holidays are labeled accordingly. It is evident that there exist substantial decreases in average daily demand around weekends and public holidays. Unlike other holidays marked on Fig. 2 (Anzac Day and Queen's Birthday), the days just before and after the Easter long weekend also show noticeable drops in demand. We therefore exclude those days (i.e., 13 and 18 April 2017) along with weekends and other public holidays, which yields the 59-day study period. For the 59-day study period, we do not differentiate different weekdays and do not consider the day-of-the-week effect in the benchmark case. We will further examine the day-of-the-week effect in Section 4.6.5.




**(a)** **(b)**

**Fig. 1.** Spatial distribution of Sydney bus commuters in the morning peak hour (from April to June 2017): (a) trip origin and (b) destination

Approximately 4.8 million trip records generated from 374 suburbs over the three-month period were initially sampled from the transit data in order to provide a general understanding of the level of bus services in Sydney. Fig. 3a and Fig. 3b plot the distributions of the in-vehicle time and the schedule delay of bus services, respectively. From Fig. 3a, it can be seen that most travelers experienced an in-vehicle time of 10 to 20 minutes. Fig. 3b indicates that most bus service schedule delays are within 10 minutes. However, there are some services with relatively severe service schedule delays (more than 20 minutes). This paper tries to quantify how the service schedule delay and service unreliability could affect travelers' choices.

### *4.2. Problem setting and formulation: Sydney case*

This section discusses the case study setting in order to embed real-world observations into the proposed dynamical models.

We start with how we define an OD pair. We define a departure bus stop as the

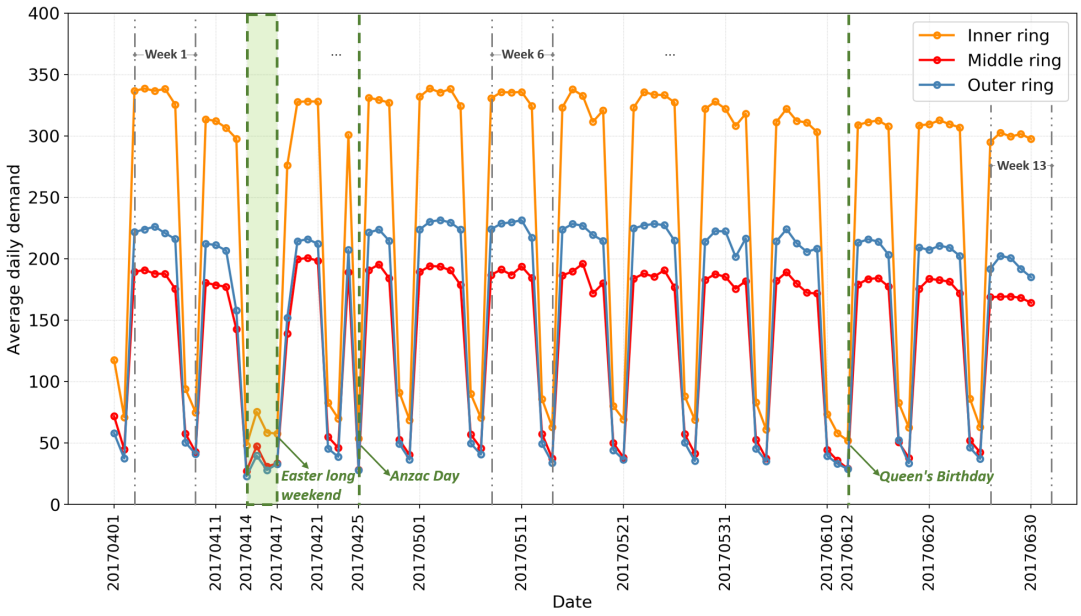
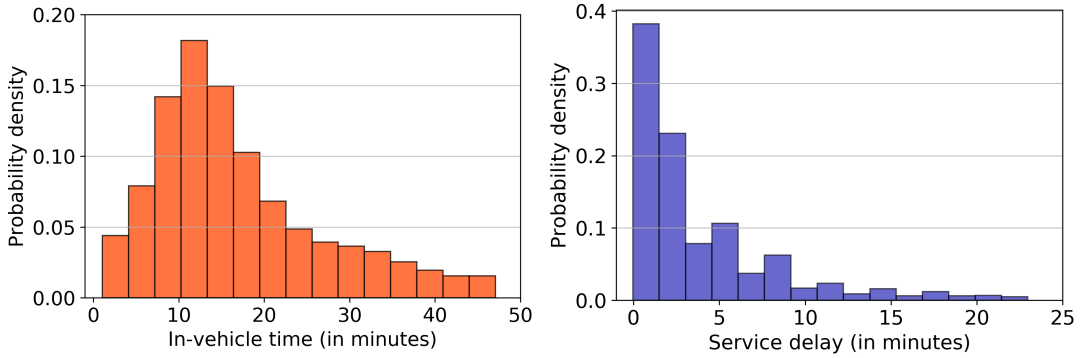



**Fig. 2.** Variation of average daily demand over the study period (from 1 April 2017 to 30 June 2017) with respect to three departure zones: inner, middle and outer ring.




**(a)** In-vehicle time

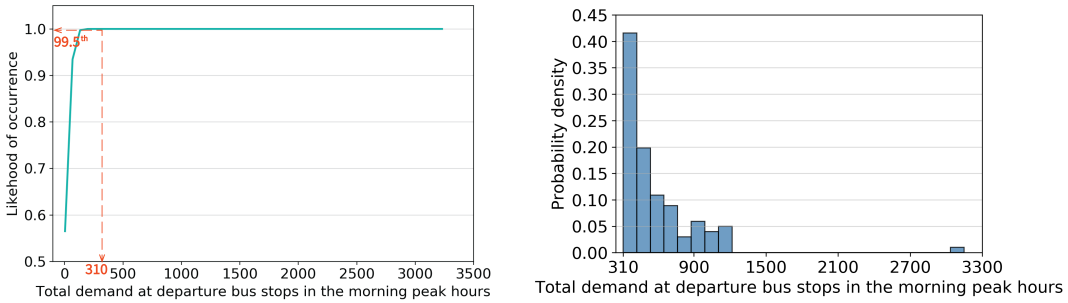
**(b)** Service schedule delay

**Fig. 3.** Distribution of in-vehicle time and service schedule delay

origin of travelers. The exact home locations of travelers are not known. Bus stops are not very dense in the suburban area in the Greater Sydney area and travelers from suburban areas often stick to a single bus stop (other few surrounding bus stops usually operate bus services to different destinations). Differently, the destination is defined as a set of arrival bus stops within a 500-meter radius. We define this as the "destination zone". This is because travelers may take different bus lines at the same departure bus stop and may alight at bus stops close to their final destinations (e.g., workplaces). This is often the case in the Sydney CBD, where the density of bus stops is very high (on average 20 bus stops within a $1km \times 1km$ square area). Note that a circle area with a radius of 500 meters already covers a significant number of bus stops and the walking distance involved expects to be acceptable to most travelers.

We now further discuss the selected departure origins. There are 20,693 bus stops in Sydney at the time of data collection. Fig. 4a shows the distribution of total demand for

these departure bus stops in the morning peak hours from 7 am to 10 am. Apparently, around 90% of the bus stops involve only a small demand (total demand in the morning peak hours $< 100$). For these low-demand OD pairs, many travelers may have a trip in one day, but then do not have a trip record in the next day (or even in next few days), i.e., discontinuity of activities in trip records over days. On the one hand, many of these observations (e.g., no trip records at all for some days) might be due to other factors rather than the service quality of the public transit. Including these might make our analysis less meaningful. On the other hand, when there are no trip records for one day, we do not have tap-on and tap-off records, and there is no data for us to calculate the experienced service schedule delay and travel time. In this study, we focus on the top 100 most used departure bus stops (at $99.5^{\text{th}}$ percentile), with total demand $> 310$ (at departure bus stops) in the morning peak hours. The demand distribution for targeted bus stops is shown in Fig. 4b. Overall, 171 OD pairs are selected for the case study i.e., $|\mathrm{K}| = 171$. We will further examine the cases where we either include additional low-demand OD pairs or further exclude some low-demand OD pairs in the 171 selected OD pairs in Section 4.6.4.




**(a)** CDF plot of total average demand at all departure bus stops in Sydney

**(b)** Distribution of demand at top 100 most used departure bus stops in Sydney

**Fig. 4.** Distribution of departure-bus-stop-based demand in Sydney

Moreover, we focus on demand changes over different departure time intervals. Therefore, we do not distinguish passengers departing within the same time interval but with different bus routes. This is to say, we do not consider bus line/route choice. In Sydney, the departure time interval is often more critical, and which bus lines to take often depends on the sequence of buses arrival at the stop, which involves non-negligible uncertainty in the peak period.

As discussed in Section 3, the costs of users include the service schedule delay and the in-vehicle time. With the available data (i.e., smart transit card data and GTFS data; refer to Section 2), the service schedule delay is estimated as the time difference between the scheduled bus arrival time and the boarding time of travelers; and the in-vehicle time is estimated as the difference between boarding time (tap-on) and alighting time (tap-off). Travelers' walking time to and from the bus stop, and the buffer time duration between the travelers' arrival time at the bus stop and scheduled bus arrival time usually vary slightly from day to day, as such we do not consider them here. These walking and waiting times are also not known unless additional surveys are conducted.

With the above in mind, we can formulate the travel cost. As discussed earlier, in order to alleviate the impact of randomness related to individual traveler and trip, we adopt an aggregate approach. For example, for a given departure time interval $m$,

we can define the mean experienced cost (based on costs of all individuals within this departure time interval) rather than individual experienced cost, which is

$$\overline{\mathrm{C}}_m^{e,q} = \omega \overline{w}_m^q + \bar{t}_m^q \tag{18}$$

where $\overline{w}_m^q$ is the mean service schedule delay experienced by all travelers $G_m^q$ who departed at time interval $m$ on day $q$; and $\bar{t}_m^q$ is the mean in-vehicle time experienced by the same group of travelers $G_m^q$. The above "mean-value-based" cost function then replaces the original individual-based cost function in Eq. (2).

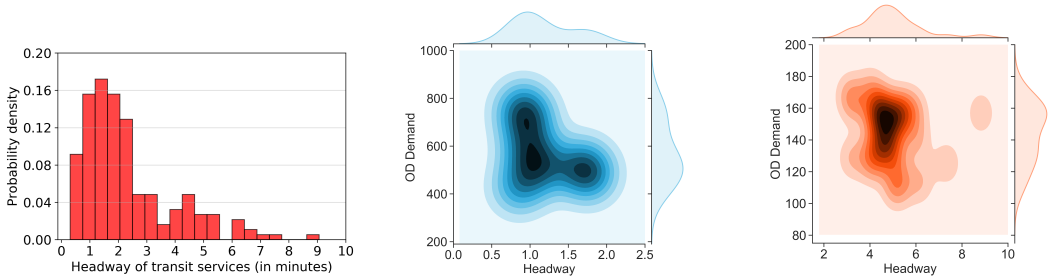
We now further discuss how to compute the "effective travel cost" defined in Eq. (1). For a given OD pair and time interval $m$, we compute the mean values and standard deviations of in-vehicle time and service schedule delay based on all trip records in the dataset in the 59 working days. This is to say, we use the mean values and standard deviations based on the 59 working days to approximate travelers' knowledge based on long-term experience or information provision. Travelers' long-term knowledge can be better approximated if a longer observation duration of transit smart card data is available.

### *4.3. Time horizon discretization and peak-demand duration identification*

The analysis in this paper is based on discrete departure time intervals, where the time horizon is discretized into multiple time intervals. We now discuss how to choose the time interval length $\varphi$. Essentially, the interval length $\varphi$ should be chosen in the way that helps uncover/reflect true departure time changes of travelers. Firstly, the time interval length should not be too small. For example, if the time interval length is much less than the bus headway (e.g., 30 seconds $\ll$ 5 minutes), postponing or forwarding the departure time by a few intervals likely result in boarding the same bus, which may be simply due to individual randomness rather than intentionally changing departure times. Secondly, the time interval length should not be too large. For example, if the time interval is 2 hours, due to the low time resolution, one may never observe departure time interval changes. We set $\varphi = 10$ minutes in the benchmark case for the following analysis. We will test different lengths of time intervals (i.e., $\varphi = 5$, 10 and 20 minutes) in Section 4.6.1.

Fig. 5a shows the distribution of the average service headways of the 171 selected OD pairs modeled in this study. From Fig. 5a, it can be seen that the headway of the 171 selected OD pairs are all within 10 minutes, which is smaller than the length of time intervals utilized for the model calibration (i.e., 10 minutes). We will further investigate how the modeling results are impacted if the OD pairs with an average service headway larger than 10 minutes are considered in Section 4.6.3. Fig. 5b and Fig. 5c further plot the joint probability distribution in the two-dimensional domain of origin-bus-stop-based demand and headway, which visualize the relationship between headway and passenger demand regarding the 171 selected OD pairs. According to Fig. 5b, when the OD demand is no less than 200, the average service headway is less than 2.5 minutes. By contrast, as shown in Fig. 5c, as the OD demand drops, the service headway increases in general.

We now turn to discuss the peak-demand hours for each OD pair. The morning peak hour typically lasts from 7 am to 10 am in Sydney. Due to the heterogeneity of socio-demographic and socio-economic characteristics among different suburbs, the work-activity-related decisions vary, and so does the peak-hour duration. In this study, we define the peak duration based on the demand level as follows. We evenly divide



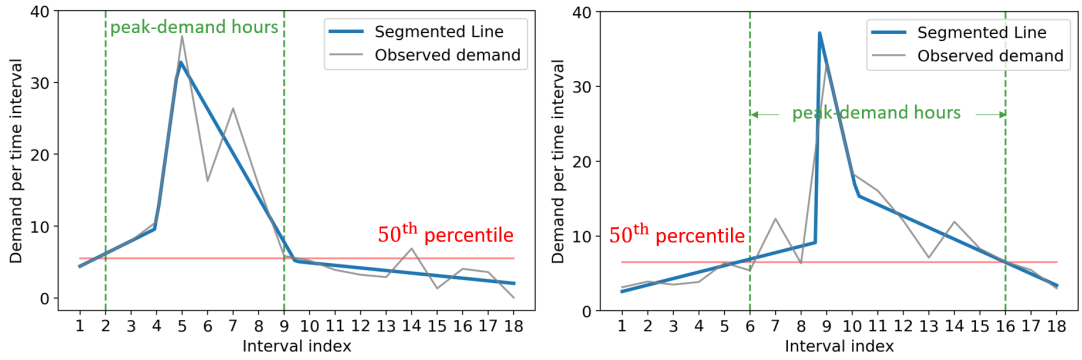

**(a)** Distribution of headway regarding the 171 OD pairs

**(b)** Demand $\geq 200$

**(c)** Demand $< 200$

**Fig. 5.** Headway distribution and its relationship with OD demand

the duration between 7 am and 10 am into 18 time intervals with the predefined interval length $\varphi = 10$ minutes. Then, the observed OD bus passenger demand (i.e., passenger flow) at each departure time interval can be calculated. The grey lines in Fig. 6 illustrate the changes in passenger flow from 7 am to 10 am with respect to two representative OD pairs, where the tick values along the horizontal axis are the interval index: the first interval is from 07:00 to 07:10 and the last interval is from 09:50 to 10:00. Building upon the time series (associated with 18 time intervals in the case study), we use the piecewise regression discussed in McZgee and Carleton (1970) to approximate a time series of length $n = 18$ with four linear segments, i.e., four blue lines displayed in Fig. 6. A threshold line (the red line) valued at $50^{\text{th}}$ percentile of the total demand is then drawn in Fig. 6, which intersects two of the four straight lines and determines the peak start time and peak ending time. It is noteworthy that we use four straight lines for the temporal demand approximation. This is based on the consideration that the demand profile is expected to include four stages, i.e., the low demand stage, the demand increasing stage, the demand decreasing stage and the low demand stage again.




**(a)** OD pair 1

**(b)** OD pair 2

**Fig. 6.** Peak-demand hour identification regarding two representative OD pairs (as illustrative examples)

### *4.4. Model calibration*

We now discuss how to calibrate the proposed day-to-day dynamical models (i.e., LPUP and PSAP) using 59-day trip records for the 171 OD pairs i.e., $|\mathrm{Q}| = 59$, $|\mathrm{K}| = 171$. It should be noted that all the selected OD pairs are mutually independent. We include all the selected OD pairs for model calibration, which helps alleviate randomness related to individual OD pair. Note that an additional subscript $i$ to indicate a specific OD pair $i \in \mathrm{K}$ is added to some of the notations in this subsection, while in previous sections it is omitted to ease the notation.

To calibrate the LPUP model, we propose the following minimization problem, which is to minimize the weighted sum of the square of difference between the observed flow-change (or demand-change) percentage and the estimated flow-change (or demand-change) percentage, i.e.,

$$\min \sum_{i\in \mathrm{K}} \sum_{q=1}^{|Q|} \sum_{m_i=1}^{M_i} \frac{\mathrm{y}_{m_i}^q}{Y_i^q} \left(z_{m_i}^q - \hat{z}_{m_i}^q\right)^2 \tag{19}$$

s.t.

$$\tilde{c}_{m_i}^q > 0 \quad \forall q \in Q, \quad \forall m_i \in V_i, \quad i \in \mathrm{K} \tag{20}$$

$$\theta_i > 0 \tag{21}$$

$$\rho \in [0, 1) \tag{22}$$

where $z_{m_i}^q = \frac{y_{m_i}^q/Y_i^q - y_{m_i}^0/Y_i^0}{y_{m_i}^0/Y_i^0}$ is the observed flow change percentage with respect to day zero (Monday, 3 April 2017) at time interval $m_i$; and $\hat{z}_{m_i}^q = \frac{\left[\rho \mathrm{P}_{m_i}^{q-1} + (1-\rho)\mathrm{P}_{m_i}^q\right] - y_{m_i}^0/Y_i^0}{y_{m_i}^0/Y_i^0}$ is the estimated flow change percentage that is calculated by the proposed LPUP model. Both observed flow and estimated flow are represented in a fractional form, i.e., the flow at time interval $m_i$ over the total OD demand on day $q$, i.e., $y_{m_i}^q/Y_i^q$ and $\rho \mathrm{P}_{m_i}^{q-1} + (1-\rho)\mathrm{P}_{m_i}^q$, respectively. More specifically, $\mathrm{P}_{m_i}^q$ and $\mathrm{P}_{m_i}^{q-1}$ are the proportions defined in the LPUP model in Eq. (4) for day $q$ and day $q-1$, respectively;[4] $y_{m_i}^q$ is the observed flow with respect to time interval $m_i$ on day $q$; $Y_i^q$ is the total observed demand between the OD pair $i$ on day $q$. Besides, we assume that on the initial day, i.e., $q = 0$, travelers' perceived cost is equivalent to the effective travel cost i.e., $\tilde{c}_{m_i}^0 = \mathrm{E}_{m_i}$.

It is noteworthy that $\theta_i$ in Eq. (21) is OD pair specific. We set $\theta_i = \theta$ ($\forall i \in \mathrm{K}$) in the benchmark case. Doing so produces an aggregate estimation of $\theta$ for all OD pairs and helps to mitigate errors or noises associated with specific OD pairs in the model calibration. We also test the cases where we set OD-specific $\theta$. The results are summarized and discussed in Section 4.6.2.

To calibrate the PSAP model, we propose the minimization problem as follows, which is to minimize the weighted sum of the square of the difference between the

[4] In order to fully utilize the real trip records, in the minimization problem in Eq. (19), when calculating $\mathrm{P}_{m_i}^q$, we do not use the cost-flow relationship defined as $\mathrm{c}(\mathrm{x}_m^q)$. Instead, we use the observed cost $\overline{\mathrm{C}}_{m_i}^{e,q}$ defined in Eq. (18) to replace $\mathrm{c}(\mathrm{y}_{m_i}^q)$ when computing Eq. (3).

model-based day-to-day demand change and the observed demand change, i.e.,

$$\min \sum_{i\in \mathrm{K}} \sum_{q=1}^{|Q|} \sum_{m_i=1}^{M_i} \frac{\mathrm{y}_{m_i}^{q-1}}{Y_i^{q-1}} \left( (1-\rho)\, \alpha \mathrm{y}_{m_i}^{q-1} \left( \mathrm{E}_{m_i} - \overline{\mathrm{C}}_{m_i}^{e,q-1} \right) - (\mathrm{y}_{m_i}^{q} - \mathrm{y}_{m_i}^{q-1}) \right)^2 \tag{23}$$

s.t.

$$\alpha > 0 \tag{24}$$

$$\rho \in [0, 1) \tag{25}$$

where $(1-\rho)\,\alpha \mathrm{y}_{m_i}^{q-1} \left( \mathrm{E}_{m_i} - \overline{\mathrm{C}}_{m_i}^{e,q-1} \right)$ is the estimated change of demand based on the PSAP model, and $\mathrm{y}_{m_i}^{q} - \mathrm{y}_{m_i}^{q-1}$ is the observed change of flow between two successive days $q$ and $q-1$ with respect to time interval $m_i$. Since on the initial day i.e., $q = 0$, the past-day experienced cost is untraceable, we set the starting point at $q = 1$.

In summary, the proposed problems in Eq. (19) and Eq. (23) both minimize the discrepancy between the estimated flow (demand) and the observed flow (demand) when calibrating the two dynamical models. All the calibrations are conducted through *'shgo'* (Simplicial Homology Global Optimisation) algorithm through SciPy (a Python library). Regarding the *'shgo'* algorithm, convergence to a global minimum is expected for Lipschitz smooth functions.

### *4.5. Results and implications*

#### *4.5.1. Calibration results and discussions*

In order to evaluate to what extent the proposed two dynamical models can approximate real observations, we compute percentage errors between the observed flows (demand) and the estimated flows (demand) with respect to departure time interval $m_i$ associated with OD pair $i$ on day $q$ based on the proposed models. The distributions of percentage errors are displayed in Fig. 7 for both models, where the corresponding average percentage error and standard deviation of percentage error are presented in Table 4. It is evident that LPUP outperforms the PSAP in terms of the percentage error. This is because, LPUP model is more capable of capturing the non-linear impacts of cost change on choice change through the multi-nomial Logit choice model in Eq. (6). Differently, PSAP model assumes a proportional relationship between flow/demand change and cost difference, as formulated in Eq. (12). The above results imply that models with stronger capabilities to capture non-linear effects or relationships may produce better approximations of reality. In addition, the PSAP model incorporates the elasticity/variation of demand, whereas the LPUP model does not. This implies that adding one more degree of freedom (demand elasticity) may create additional noise in model calibration and thus affect the accuracy (i.e., PSAP is outperformed by the LPUP).

Table 4 further summarizes the values of calibrated parameters of the two models. All calibrated parameters are dimensionless except $\alpha$ in the PSAP model whose unit is adjustment rate per unit travel cost. Additionally, the unit of variables related to travel costs is minute; and the magnitudes of $\alpha$ and $\theta$ depend on the unit of travel

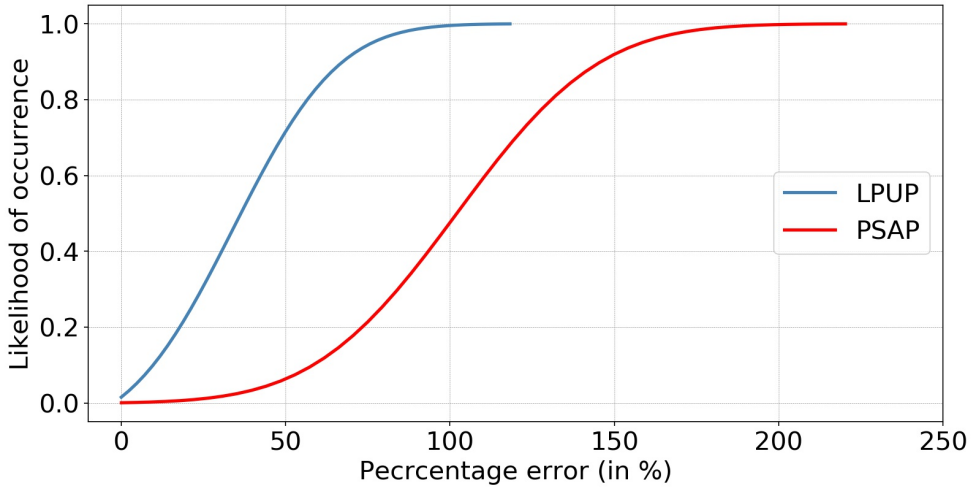



**Fig. 7.** Distribution of percentage errors regarding the proposed dynamical models

cost variables. Since the PSAP model involves much larger errors, we focus on the values of parameters based on the LPUP model. In particular, we have the following observations.

- Firstly, the relative value of service schedule delay against in-vehicle time $\omega$ is around 3.27, which means that service schedule delay is more costly than in-vehicle time in the peak hour. This is compatible with the value of unexpected delay time suggested by the Transport for New South Wales, i.e., 3.2 times the on-board transit time.[5] Other empirical studies on evaluating demand parameters for inner Sydney Public Transport also recommended 3.2 times the in-vehicle time as the value of lateness (e.g., Douglas and Jones, 2016).
- Secondly, the sizes of safety margins in relation to service schedule delay and in-vehicle time are 1.83 ($\eta_1$) and 2.09 ($\eta_2$) times the standard deviation of service schedule delay and in-vehicle time, respectively. The standard deviation of the observed service schedule delay and the standard deviation of the observed in-vehicle time are 3.12 minutes and 10.37 minutes, respectively. These indicate that travelers leave around 6-minute safety margins for the service schedule delay and around 20-minute safety margins for the in-vehicle time. It should be noted that these numbers are averages based on 171 selected OD pairs, involving long trips with an in-vehicle time more than 20 minutes and congestion uncertainty in the peak duration.
- Thirdly, $\rho = 0.828695$ means that, in the context of the LPUP model, around 83% of the travelers repeated their travel choices and did not re-consider their travel choices from day to day; or equivalently almost 17% of the travelers considered changing their departure time over the 59-day observation period. This implies that a considerable amount of travelers did not re-consider their travel choices day by day.
- The day-to-day learning parameter is relatively small ($\kappa = 0.067545$), which means that travelers did not sharply change their forecasting behavior. Also note that the system is more likely to be stable if $\kappa$ is smaller. This will be

[5]This is based on Transport for NSW Economic Parameter Values (`https://www.transport.nsw.gov.au/news-and-events/reports-and-publications/transport-for-nsw-economic-parameter-values`).

further discussed in Section 4.5.2.

**Table 4.** Model calibration result

**LPUP**

| Error avg. | Error SD | $\omega$ | $\rho$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\kappa$ |
|---|---|---|---|---|---|---|---|
| 35.257753% | 25.163339% | 3.268301 | 0.828695 | 1.831775 | 2.089608 | 0.010153 | 0.067545 |

**PSAP**

| Error avg. | Error SD | $\omega$ | $\rho$ | $\eta_1$ | $\eta_2$ | $\alpha$ |
|---|---|---|---|---|---|---|
| 104.243680% | 34.227432% | 1.102067 | 0.903608 | 1.409276 | 2.544926 | 0.060617 |

*Error avg.* and *Error SD* stand for the average percentage error and the standard deviation of percentage error, respectively. Specifically, the average percentage error and the standard deviation are computed based on the percentage errors of all segmented time intervals associated with the 171 selected OD pairs over the 59-day study period.

As can be seen from Table 4 and Fig. 7, while LPUP outperforms PSAP in terms of accuracy, the errors for both LPUP and PSAP (in terms of fitting the data) are not small. This might be due to other factors or uncertainties in relation to individual travel and activity patterns that are not considered in this paper. A future study may incorporate other information or data sources such as weather, household survey, characteristics of travelers, social media platform data, road traffic condition to examine the day-to-day traffic variation.

### *4.5.2. System stability analysis*

In Section 3, we derived the analytical conditions for the existence, uniqueness and stability of traffic flow at the fixed point for both LPUP and PSAP models. These conditions are built upon Assumption 1, where $\mathrm{c}(\cdot)$ is a strictly increasing function. It requires that the first derivative of the cost-flow function, i.e., $\mathrm{c}'(\cdot)$, is positive. In order to evaluate the marginal effect of the flow on travel cost, we formulate the cost-flow function $\mathrm{c}(\mathrm{x}_{m_i}^{q-1})$ as follows (i.e., a linear approximation)

$$\mathrm{c}(\mathrm{x}_{m_i}^{q-1}) = \beta_0 + \beta_1 \mathrm{x}_{m_i}^{q-1} + \beta_2 L_i \tag{26}$$

where $\beta_0$, $\beta_1$ and $\beta_2$ are the coefficients; and $L_i$ is the Euclidean distance between the origin and the center of the destination zone with respect to OD pair $i$. $L_i$ captures the spatial heterogeneity among different OD pairs. Regarding the OD pair $i \in \mathrm{K}$, $L_i$ is a constant, and so is the term $\beta_0 + \beta_2 L_i$. We then estimate the coefficients in Eq. (26) with the observed experienced cost $\overline{\mathrm{C}}_{m_i}^{e,q-1}$ and the observed flow (demand). Due to the difference in the calibrated values of $\omega$ for LPUP and PSAP models, we estimate the cost function separately for the two dynamical systems. Table 5 provides the estimated parameters, their statistical significance and the goodness-of-fit regarding cost-flow functions for the two models. It is evident that the estimated parameters are all positive and statistically significant (P-value $\leq 0.05$). This implies that $\mathrm{c}(\cdot)$ is strictly increasing with the flow, which is consistent with Assumption 1 in Section 3.3.

In Table 5, it can be seen that the value of $\mathrm{R}^2$ associated with the PSAP model is larger than that associated with the LPUP model. This difference is due to the fact that different values of $\omega$ are used in Eq. (18) to calculate the mean experienced cost for the two models (see Table 4 where $\omega = 3.268301$ for LPUP and $\omega = 1.102067$ for PSAP). This result is consistent with the characteristics of the LPUP and PSAP models, which is further explained below. From Eq. (11), one can see that the swapping demand rate $r_m^q$ (for the PSAP model) is linear with respect to the experienced cost.

This indeed explicitly forces a certain linear relationship between experienced cost and the flow, which exactly matches the linear relationship to be estimated between experienced cost and flow in Eq. (26). This matching between the PSAP model and the cost-flow relationship further means that the value of $\omega$ associated with PSAP is calibrated to better fit a linear cost-flow relationship, which results in a higher $R^2$ for a linear regression between cost and flow. Differently, the LPUP model assumes a non-linear relationship between flow and cost (governed by the Logit model). It is also noteworthy that the values of $R^2$ for both PSAP and LPUP might be further improved by incorporating other factors such as weather information and multi-modal traffic in a future study, while the trip length and flow at the origin only partially capture the factors affecting the travel cost.

**Table 5.** Cost function linear regression result

| | $\beta_0$ | P-value | $\beta_1$ | P-value | $\beta_2$ | P-value | $R^2$ |
|---|---|---|---|---|---|---|---|
| LPUP | 21.0089 | 0.0000 | 0.0570 | 0.0486 | 2.0480 | 0.0000 | 0.389 |
| PSAP | 13.8052 | 0.0000 | 0.1293 | 0.0000 | 2.0361 | 0.0000 | 0.517 |

We now empirically examine the stability conditions of the two dynamical systems based on the above aggregations and approximations. We substitute the calibrated parameters into the theoretical stability conditions, and summarize the results in Table 6. Apparently, the system stability conditions are satisfied for both systems. This implies that the Sydney public transit system might be in a stable state but with day-to-day random variations. Therefore, we further examine Eq. (27), which is the observed absolute weighted average percentage flow discrepancy between two successive days $\Delta^{q,q-1}$, i.e.,

$$\Delta^{q,q-1} = \sum_{i=1}^{|\mathrm{K}|} \sum_{m_i=1}^{M_i} \left[ \frac{\mathrm{y}_{m_i}^{q}}{\sum_{i=1}^{|\mathrm{K}|} Y_i^q} \left| \frac{\mathrm{y}_{m_i}^{q}}{Y_i^q} - \frac{\mathrm{y}_{m_i}^{q-1}}{Y_i^{q-1}} \right| \right] \tag{27}$$

where $\mathrm{y}_{m_i}^{q}$ is the observed flow/demand with respect to the departure time interval $m_i$ on day $q$, and $Y_i^q$ is the total observed demand between OD pair $i$ on day $q$. The value of $\Delta^{q,q-1}$ is defined over all time intervals and all the 171 selected OD pairs.

**Table 6.** Stability condition

| | $\lvert\gamma_k\rvert$ | | $< 2(1+\rho)/\left[\kappa(1-\rho)\right]$ | |
|---|---|---|---|---|
| LPUP | 0.0465 | | 316.0882 | |
| | $\min(\lambda_{\mathrm{f}})$ | $\max(\lambda_{\mathrm{f}})$ | **Lower bound:** $-\frac{2}{\alpha(1-\rho)}$ | **Upper bound:** 0 |
| PSAP | -12.8007 | -0.0647 | -342.2903 | 0.0000 |

Fig. 8 visualizes the variation of $\Delta^{q,q-1}$ over the entire 59-day observation period. It is obvious that $\Delta^{q,q-1}$ fluctuates around 6.5%. As discussed earlier, the stability conditions of the two dynamical system hold. This 6.5% system variation from day to day may be due to uncertainty related to human activities and demand conditions.[6] A related study from the supply's point of view by Liu and Szeto (2020) showed that the transport system performance function can be stochastic, which can also yield flow

[6]According to the Australian Labour Market Statistics (published on October, 2010), 16% of employees aged 15 years and over in Australia usually worked shift work (based on records from Australian Bureau of Statistics at `https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Chapter10102008`).

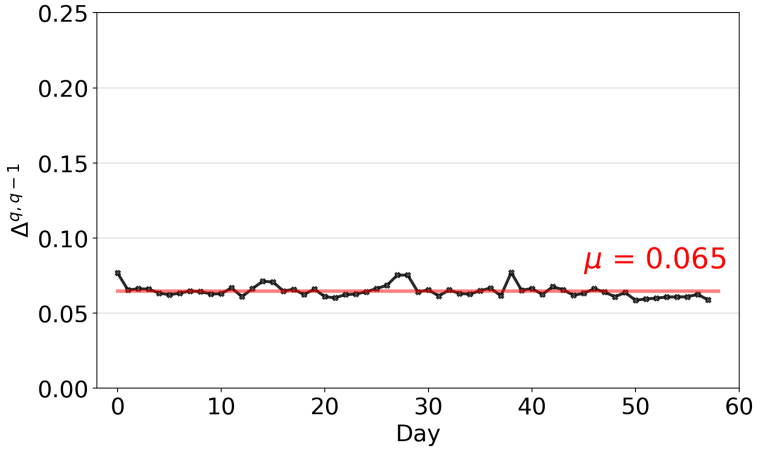



**Fig. 8.** Trajectory of observed changes in flow

fluctuations in the system from day to day while on average the system is considered as converged.

The above stability analysis is based on the calibrated models for the 171 selected OD pairs over the 59-day observation period. The analysis relies on the simplified cost-flow relationship in Assumption 1. As a first step to examine the impacts of transit service unreliability/reliability on day-to-day travel choices, we did not consider complex flow interactions in a network. More complex and realistic network flow dynamics might be incorporated in a future study, which also requires more data. The above stability analysis has to be re-conducted as well after incorporating more complex network-level interactions.

### *4.6. Sensitivity analysis*

Section 4.5 presents the model calibration results for the LPUP and PSAP model based on the 59-day trip records of the 171 OD pairs whose daily with the average service headway $< 10$ minutes (i.e., the benchmark case). This section further explores how the model settings and the variation of data used in the analysis can impact the results, including (i) varying the time interval length $\varphi$, (ii) applying the OD-specific $\theta$ (or Group-specific $\theta$) for the LPUP model, (iii) modeling the OD pairs with relatively large service headways ($\geq 10$ minutes), (iv) changing the size of modeling data based on demand level, and (v) examining the day-of-the-week effect by excluding Fridays. The additional model calibration results reported in this subsection are compared with the results in the benchmark case i.e., Table 4.

#### *4.6.1. Varying the time interval length $\varphi$*

We first vary the length of time interval (for discretizing the time horizon within a day) and compare the calibration results. In particular, we adopt $\varphi = 5$, 10 and 20 minutes (based on 171 selected OD pairs), respectively. Table 7 and Table 8 summarize the results with respect to different lengths of time intervals for LPUP and PSAP, respectively; and the error distributions are plotted in Fig. 9. We find that for the LPUP model, the interval length of 10 minutes yields the smallest average error. For the PSAP model, different interval lengths yields comparable average errors. In terms

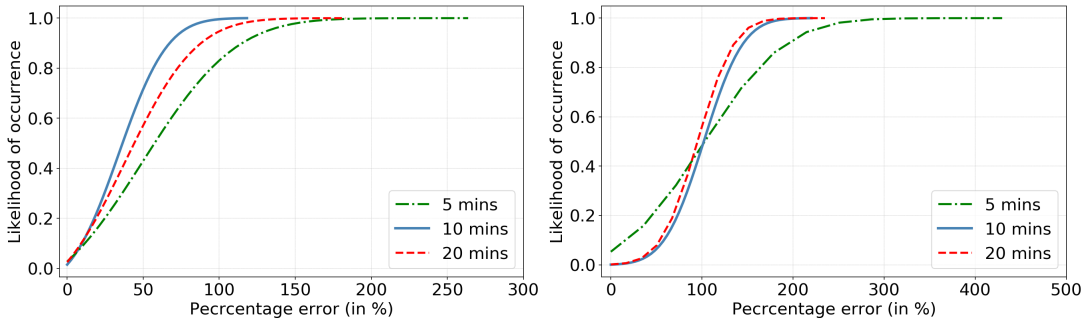
of estimated parameters, $\varphi = 10$ and $\varphi = 20$ (minutes) yield comparable results (e.g., similar $\omega$, $\rho$, $\eta_1$, $\eta_2$, $\theta$ for LPUP). However, it is noteworthy that for both LPUP and PSAP, when we decrease the time interval length, the estimated $\rho$ decreases, which means more travelers considered changing their travel choices. This is partially due to the fact that when we have an over-small time interval, a small variation in departure time (to catch the same bus) is regarded as changing departure time interval or travel choices, and the proportion of travelers re-considered travel choices is over-estimated. This also explains why a time interval length of 5 minutes always yields the largest average error. By contrast, an over-large time interval (here the LPUP models with interval length of 20 minutes) might be unable to capture the travel choice update and can result in larger estimation errors (see the average errors for the LPUP model).

**Table 7.** LPUP model calibration results with different time interval lengths

| Interval length | Error avg. | Error SD | $\omega$ | $\rho$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\kappa$ |
|---|---|---|---|---|---|---|---|---|
| 5 mins | 52.465744% | 48.039260% | 2.408839 | 0.664861 | 2.434893 | 2.051908 | 0.010918 | 0.052824 |
| 10 mins (benchmark) | 35.257753% | 25.163339% | 3.268301 | 0.828695 | 1.831775 | 2.089608 | 0.010153 | 0.067545 |
| 20 mins | 40.105993% | 36.640015% | 2.968240 | 0.918244 | 1.830910 | 2.059496 | 0.009022 | 0.104904 |

**Table 8.** PSAP model calibration results with different time interval lengths

| Interval length | Error avg. | Error SD | $\omega$ | $\rho$ | $\eta_1$ | $\eta_2$ | $\alpha$ |
|---|---|---|---|---|---|---|---|
| 5 mins | 118.641249% | 72.146586% | 0.662135 | 0.773786 | 0.895209 | 3.287503 | 0.041796 |
| 10 mins (benchmark) | 104.243680% | 34.227432% | 1.102067 | 0.903608 | 1.409276 | 2.544926 | 0.060617 |
| 20 mins | 103.641919% | 32.071789% | 1.532342 | 0.905716 | 2.536335 | 2.578217 | 0.018759 |




**(a)** LPUP **(b)** PSAP

**Fig. 9.** Distribution of percentage errors regarding dynamical models with different time interval lengths.

### *4.6.2. OD-specific/Group-specific $\theta$ for LPUP model*

As discussed in Section 4.4, we calibrate a single $\theta$ value for all OD pairs in the LPUP model. We now further examine the cases where we may have different values of $\theta$ for different OD pairs. We compare three cases: (i) a single $\theta$ value for all OD pairs; (ii) four different $\theta$ values for four groups of OD pairs, where OD pair groups are determined based on level of passenger demand (Group-specific-$\theta$ LPUP); (iii) OD pair specific value of $\theta$ for each OD pair (OD-specific-$\theta$ LPUP). For the Group-specific-$\theta$ LPUP model, the OD pairs are evenly divided into four groups based on their average daily demand (the group-wise averages are 108.72, 201.54, 260.01 and 434.05 trips per day), and the OD pairs from the same group share a single parameter $\theta_g, g \in \{1, 2, 3, 4\}$.

Note that for the three different cases, we use the same dataset with 171 OD pairs and minimize the weighted sum of the square of difference between the observed flow-change percentage ($z^q_{m_i}$) and the estimated flow-change percentage ($\hat{z}^q_{m_i}$) by using the same minimization problem defined in Eqs. (19) - (22).

**Table 9.** Calibration results of LPUP-based models

LPUP (Benchmark with a single $\theta$)

| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
|---|---|---|---|---|---|---|---|
| 35.257753% | 25.163339% | 3.268301 | 0.067545 | 1.831775 | 2.089608 | 0.010153 | 0.828695 |

Group-specific-$\theta$ LPUP

| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ (weighted avg.) | $\rho$ |
|---|---|---|---|---|---|---|---|
| 31.888573% | 24.464857% | 3.092961 | 0.075342 | 2.031523 | 2.117915 | 0.009745 | 0.852760 |

OD-specific-$\theta$ LPUP

| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ (weighted avg.) | $\rho$ |
|---|---|---|---|---|---|---|---|
| 34.202047% | 23.434054% | 2.981137 | 0.046117 | 2.252958 | 2.133635 | 0.009921 | 0.854407 |

The OD-pair-specific weights $k_i$ (group specific $k_g$) used to compute the weighted average values of $\theta$ are calculated based on daily average demand of OD pair $i$ (group-wise daily average OD demand of group $g$), i.e., $k_i = \frac{\overline{Y}_i}{\sum_{i=1}^{|K|} \overline{Y}_i}$ ($k_g = \frac{\overline{W}_g}{\sum_{g=1}^{4} \overline{W}_g}$), where $\overline{Y}_i$ ($\overline{W}_g$) is the average daily demand of OD pair $i$ (group-wise daily average OD demand of group $g$).

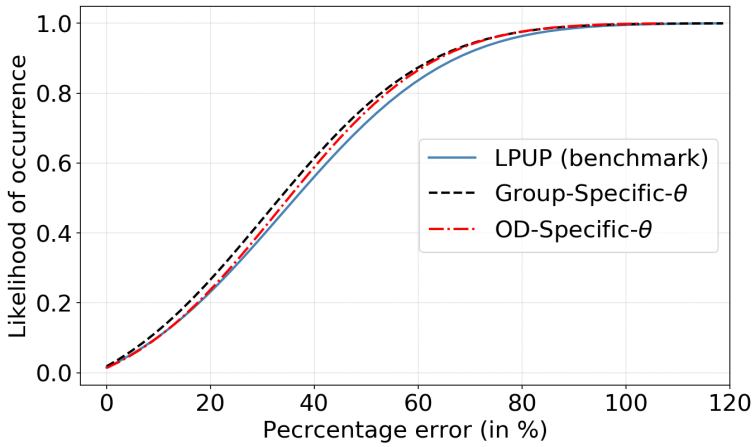



**Fig. 10.** Comparison between the LPUP and its two variations.

Table 9 summarizes the model calibration results and Fig. 10 displays the distributions of percentage errors for the three cases, i.e., Case (i): a single $\theta$ for all OD pairs; Case (ii): Group-specific-$\theta$ LPUP; and Case (iii): OD-specific-$\theta$ LPUP. Based on Table 9 and Fig. 10, we have the following observations.

Firstly, allowing different values of $\theta$ for different OD pairs can yield better calibration results (in terms of estimation error), i.e., Case (ii) and Case (iii) yield smaller average errors over all OD pairs than Case (i). This is straightforward as more parameters are introduced to allow better fitting. Secondly, it is found that Case (iii) with more parameters yields a larger average error than Case (ii). This is because, Case (iii) involves too many parameters in the calibration (i.e., 171 OD-specific $\theta$ parameters since we have 171 OD pairs), it is more likely that the solution (of the minimization problem for calibration) is sub-optimal. Thirdly, while the three cases yield different calibration results, these results are comparable in terms of the estimated parameter values and level of accuracy. This implies that estimation results are sensitive to the number of $\theta$ parameters used only to a limited extent.

### *4.6.3. OD pairs with a large service headway*

In the benchmark case, we calibrate the proposed models using the 171 selected OD pairs whose average headway is less than 10 minutes (refer to Fig. 5a). We now investigate OD pairs with a larger headway ($\geq$ 10 minutes). In particular, we further select another 20 OD pairs with a headway between 10 and 20 minutes and 6 OD pairs with a headway between 20 and 30 minutes. Note that these OD pairs associate with low demand. The selection of these OD pairs are based on the prerequisite that there are trip records for each day in the analysis (we cannot compute the service schedule delay and travel time when there is no trip record for OD pairs with very low demand).

Moreover, for the OD pairs with a headway between 10 and 20 minutes (20 and 30 minutes), we use a time interval length of 20 minutes (30 minutes) to ensure that the headway is always smaller than the length of time intervals. The minimization problems shown in Eqs. (19) - (22) (for LPUP) and Eqs. (23) - (25) (for PSAP) are again utilized to calibrate the proposed models.

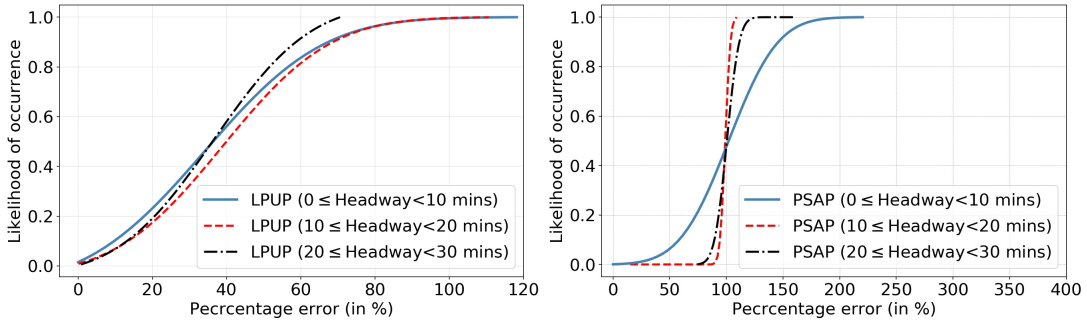
**Table 10.** Model calibration results for OD pairs with different average headway lengths

LPUP Model

| Headway range | Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
|---|---|---|---|---|---|---|---|---|
| $0 \leq$ Headway $<$ 10 mins (benchmark) | 35.257753% | 25.163339% | 3.268301 | 0.067545 | 1.831775 | 2.089608 | 0.010153 | 0.828695 |
| $10 \leq$ Headway $<$ 20 mins | 40.054673% | 22.812027% | 2.528855 | 0.242059 | 2.343091 | 1.557752 | 0.010075 | 0.954713 |
| $20 \leq$ Headway $<$ 30 mins | 37.250333% | 20.475015% | 1.210742 | 0.405762 | 1.828809 | 1.573496 | 0.008095 | 0.961301 |

PSAP Model

| Headway range | Error avg. | Error SD | $\omega$ | $\alpha$ | $\eta_1$ | $\eta_2$ | $\rho$ |
|---|---|---|---|---|---|---|---|
| $0 \leq$ Headway $<$ 10 mins (benchmark) | 104.243680% | 34.227432% | 1.102067 | 0.060617 | 1.409276 | 2.544926 | 0.903608 |
| $10 \leq$ Headway $<$ 20 mins | 100.259454% | 5.597379% | 2.103550 | 0.004752 | 2.269105 | 2.487236 | 0.939028 |
| $20 \leq$ Headway $<$ 30 mins | 101.335122% | 8.597552% | 1.048317 | 0.005001 | 2.456542 | 1.879570 | 0.913499 |




**(a)** LPUP **(b)** PSAP

**Fig. 11.** Distribution of percentage errors regarding the OD pairs with a large headway (> 10 mins).

Table 10 summarizes the model calibration results for the large-headway OD pairs; and the corresponding distributions of percentage errors are plotted in Fig. 11. We discuss the main observations below. Firstly, LPUP always outperforms PSAP in terms of fitting the data. For the LPUP model, the percentage errors distribute closely for these three categories of OD pairs.

Secondly, for OD pairs with a larger average headway, the estimated parameter $\rho$ (i.e., ratio of travelers who do not consider changing their travel choices) under the LPUP model is larger, indicating that less travelers will re-consider their travel choices from day to day. This is consistent with our expectation that for OD pairs with very low-level transit services, many people will choose to drive (and thus are not captured in the dataset) and those taking buses often stick to the same bus from day to day

(e.g., leaving sufficient buffer time to catch up with the bus from day to day).

Thirdly, under the LPUP model, for OD pairs with a larger headway, although fewer passengers re-consider their travel choices from day to day (a smaller $1-\rho$), the value of $\kappa$ is larger. This means that those truly re-considering their choices under a larger headway might be more sensitive to the previous day's experience. Differently, the observation on $\kappa$ in the LPUP model does not hold for a similar parameter $\alpha$ in the PSAP model. This is further discussed below. PSAP yields much larger inaccuracy, while it yields comparable values of $\rho$, the results on both $\rho$ and $\alpha$ are less reliable. Moreover, while $\kappa$ and $\alpha$ play comparable roles in two dynamical models (LPUP and PSAP), $\kappa$ converts costs into costs and is dimensionless, but $\alpha$ converts costs to percentage change in flow and has a unit of adjustment rate per unit cost. Overall, LPUP model with a dimensionless $\kappa$ outperforms the PSAP model with $\alpha$ that has a dimension, and produces more reliable estimations and interpretations.

Finally, considering the parameters describing service reliability, for OD pairs with a larger average headway, the relative value of service schedule delay against in-vehicle time $\omega$ (refer to the LPUP model) is smaller. This means for OD pairs with a large service headway, travelers taking transit service regularly are relatively insensitive to the service schedule delay. This might be because, the service schedule delay is often negligible when compared to the large service headway. For the values of $\eta_1$ and $\eta_2$, they are still comparable when we consider OD pairs with different service headways.

**Table 11.** Stability conditions for OD pairs with a large headway

| LPUP | $\lvert\gamma_k\rvert$ | | $< 2(1+\rho)/\left[\kappa(1-\rho)\right]$ | |
|---|---|---|---|---|
| $10 \leq$ Headway $< 20$ mins | 0.0592 | | 356.6303 | |
| $20 \leq$ Headway $< 30$ mins | 0.0821 | | 249.8062 | |
| PSAP | min ($\lambda_{\mathrm{f}}$) | max ($\lambda_{\mathrm{f}}$) | **Lower bound:** $-\frac{2}{\alpha(1-\rho)}$ | **Upper bound:** 0 |
| $10 \leq$ Headway $< 20$ mins | -2.7799 | -0.0647 | -6902.7655 | 0.0000 |
| $20 \leq$ Headway $< 30$ mins | -5.8578 | -0.9012 | -4623.2993 | 0.0000 |

The stability conditions are also examined regarding the OD pairs with a larger average headway; and the results are presented in Table 11. We find that for both LPUP and PSAP, the stability conditions are still satisfied for different groups of OD pairs (based on different average service headways).

### *4.6.4. Varying the size of modeling data based on demand level*

This subsection investigates how the model calibration results might be affected if we include additional OD pairs with low OD demand or exclude some low demand OD pairs from the 171 selected OD pairs.

**(Enlarging the size of the data)** We now include additional OD pairs with relatively low demand into the analysis. We select additional 68 OD pairs, where the minimum OD demand is 16 trips per day on average (whereas, the minimum OD demand for the benchmark case is 85 trips per day). For these 68 low-demand OD pairs, we have trip records for each day in the analysis and thus can calculate the service schedule delay and in-vehicle time accordingly. We then re-calibrate our model with 239 ($= 171 + 68$) OD pairs and use the time interval length of 10 minutes. The calibration results are summarized in Table 12 and the distributions of percentage errors are displayed in Fig. 12.

From Table 12 and Fig. 12, we can see that including OD pairs with relatively low demand yields more errors for the LPUP model, but produces comparable parameter estimation to the benchmark case for the LPUP model (e.g., $\omega$, $\eta_1$, $\eta_2$, $\rho$). The increased

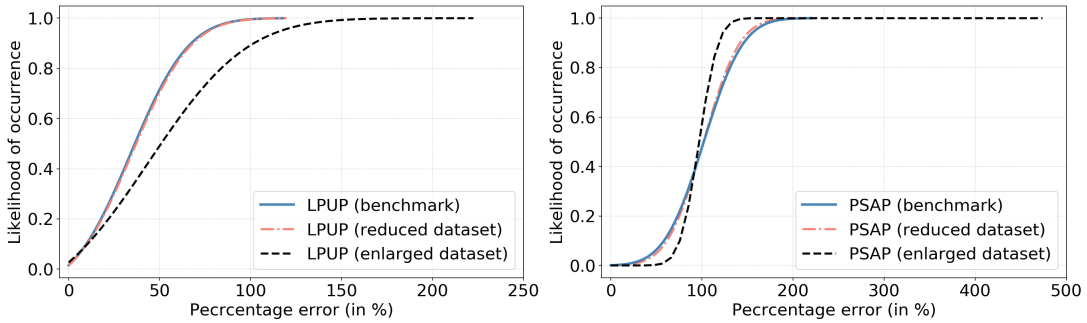
average error is partially due to that we use the single calibrated model setting to represent more OD pairs (with more variations). However, for the PSAP model, including OD pairs with relatively low demand yields similar errors and produces comparable estimated values for $\omega$ and $\rho$.

**(Reducing the size of the data)** We now further look at the case where we reduce the number of OD pairs in the analysis. Specifically, we choose 86 OD pairs with relatively high demand (the minimum OD demand is 226 trips per day) from the 171 selected OD pairs utilized in the benchmark case to re-calibrate the models with a time interval length of 10 minutes. The calibration results are reported in Table 12; and the distributions of percentage errors are also plotted in Fig. 12. From Table 12 and Fig. 12, we can see that excluding relatively low-demand OD pairs from the 171 OD pairs in the benchmark case yields comparable estimation accuracy for both LPUP and PSAP, and also yields similar estimated values for most parameters.

**Table 12.** Model calibration results regarding the changes in data size

| LPUP (enlarged size of data) | | | | | | | |
|---|---|---|---|---|---|---|---|
| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
| 46.249791% | 42.538623% | 3.068334 | 0.105472 | 1.831758 | 2.089511 | 0.008268 | 0.838676 |
| LPUP (reduced size of data) | | | | | | | |
| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
| 36.119960% | 25.326164% | 3.231330 | 0.060637 | 1.995193 | 2.321480 | 0.009658 | 0.812946 |
| LPUP (benchmark) | | | | | | | |
| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
| 35.257753% | 25.163339% | 3.268301 | 0.067545 | 1.831775 | 2.089608 | 0.010153 | 0.828695 |

| PSAP (enlarged size of data) | | | | | | |
|---|---|---|---|---|---|---|
| Error avg. | Error SD | $\omega$ | $\alpha$ | $\eta_1$ | $\eta_2$ | $\rho$ |
| 101.707781 % | 16.749153 % | 1.210992 | 0.011969 | 1.081108 | 4.968254 | 0.902730 |
| PSAP (reduced size of data) | | | | | | |
| Error avg. | Error SD | $\omega$ | $\alpha$ | $\eta_1$ | $\eta_2$ | $\rho$ |
| 103.889078% | 31.298879% | 0.894099 | 0.015463 | 1.523291 | 3.389905 | 0.724350 |
| PSAP (benchmark) | | | | | | |
| Error avg. | Error SD | $\omega$ | $\alpha$ | $\eta_1$ | $\eta_2$ | $\rho$ |
| 104.243680% | 34.227432% | 1.102067 | 0.060617 | 1.409276 | 2.544926 | 0.903608 |




**(a)** LPUP **(b)** PSAP

**Fig. 12.** Distribution of percentage errors regarding dynamical models with different data sizes.

*4.6.5. Excluding Fridays*

In the benchmark case, we do not differentiate different weekdays and do not consider the day-of-the-week effect. To visualize whether the day-of-the-week effect is significant in Sydney, we first plot the variations of the daily average demand over a week with respect to each departure zone (inner, middle and outer rings, which are defined in Section 4.1) in Fig. 13. As can be seen, the average daily demand varies slightly over different weekdays, and Fridays do show a drop in demand, but only to a limited extent. These indicate that the day-of-the-week effect (considering weekdays only) for the public transit system in the Greater Sydney area might be limited. There are significant demand drops in the weekends, which are excluded in the analysis, as discussed in Section 4.1.

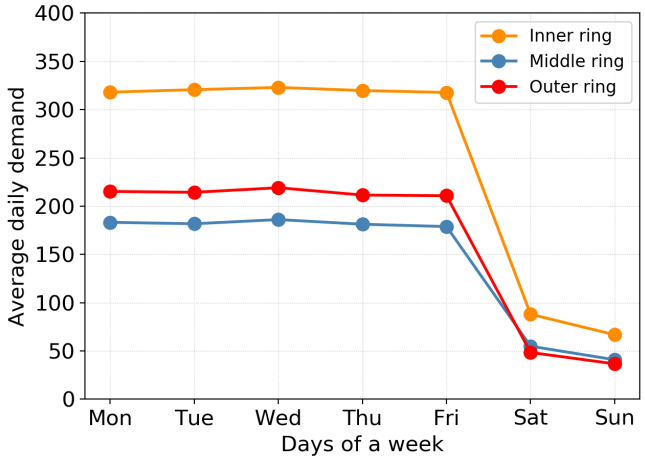


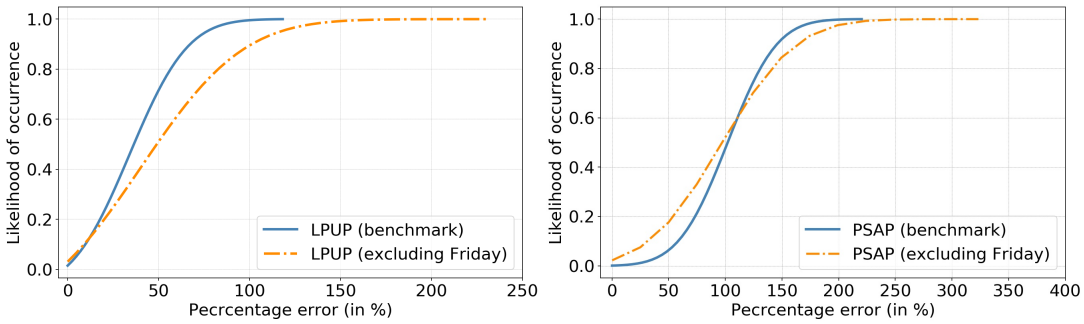

**Fig. 13.** Daily average demand variations for a week

We also test the case where Fridays are excluded in the model calibration. Table 13 compares the model calibration results without and with Fridays; and the error distributions are plotted in Fig. 14. As can be seen, after excluding Fridays, while some estimated parameters are still comparable, the average error becomes larger. This might be because, how Thursdays' experiences affect Fridays' travel choices, and how Fridays' experiences affect the following Mondays' travel choices are missed in the calibration.

**Table 13.** Model calibration results with/without the exclusion of Fridays

| LPUP (w/o Fridays) | | | | | | | |
|---|---|---|---|---|---|---|---|
| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
| 42.653063% | 44.267057% | 3.286383 | 0.168189 | 2.048269 | 2.270548 | 0.004185 | 0.821387 |
| LPUP (benchmark) | | | | | | | |
| Error avg. | Error SD | $\omega$ | $\kappa$ | $\eta_1$ | $\eta_2$ | $\theta$ | $\rho$ |
| 35.257753% | 25.163339% | 3.268301 | 0.067545 | 1.831775 | 2.089608 | 0.010153 | 0.828695 |

| PSAP (w/o Fridays) | | | | | | |
|---|---|---|---|---|---|---|
| Error avg. | Error SD | $\omega$ | $\alpha$ | $\eta_1$ | $\eta_2$ | $\rho$ |
| 109.050639 % | 52.385683% | 1.320550 | 0.009178 | 0.985292 | 5.639989 | 0.580011 |
| PSAP (benchmark) | | | | | | |
| Error avg. | Error SD | $\omega$ | $\alpha$ | $\eta_1$ | $\eta_2$ | $\rho$ |
| 104.243680% | 34.227432% | 1.102067 | 0.060617 | 1.409276 | 2.544926 | 0.903608 |




**(a)** LPUP **(b)** PSAP

**Fig. 14.** Distribution of percentage errors regarding dynamical models with/without the exclusion of Fridays.

## 5. Conclusions and Discussions

This paper explores the day-to-day travel choices in Sydney public transit systems, with an emphasis on service reliability. Two dynamical systems (i.e., learning and perception updating process, and proportional-switch adjustment process) are proposed and analytical properties of the two systems are examined. The proposed day-to-day reliability-based departure time models are further calibrated by using real-world smart transit card data from Sydney under some aggregations and approximations.

A series of empirical insights are generated. We highlight a few of them here. Firstly, the relative value of service delay against in-vehicle time is around 3.27. Secondly, the sizes of safety margins in relation to service schedule delay and in-vehicle time are 1.83 and 2.09 times the standard deviation of service schedule delay and in-vehicle time, respectively. Thirdly, a significant number of travelers repeated their travel choices and did not re-consider their travel choices from day to day. Fourthly, the stability conditions of the two developed dynamical system are examined with real data, indicating possible public transit system stability in Sydney but with random variations from day to day. Last but not least, the LPUP model approximates the real trip records better than the PSAP model, which might be due to (i) LPUP's stronger capability to capture non-linear relationship between cost perception and travel choice, and (ii) PSAP's incorporation of demand elasticity.

This study illustrates the potential of smart transit card data to be utilized to uncover public transit service reliability and how the reliability has affected travelers' day-to-day travel choices. However, the discrepancy between the observed and estimated flow/demand is not small for the proposed day-to-day models (refer to Table 4). If one aims to reduce estimation errors (this can be particularly useful in demand forecasting), one may either try to include other factors that are not captured in this paper (this requires much more data), or to utilize more powerful tools in terms of fitting the data such as machine learning models (while the results are likely less interpretable, see, e.g., Li et al., 2021). For instance, as discussed in Section 4.5, a future study may integrate the smart transit card data with other survey data, social media platform data, road car traffic data to further examine impacts of factors beyond service reliability and travel time on travelers' day-to-day travel choices.

Moreover, this study assumes a simplified cost-flow relationship in the day-to-day dynamical models and ignores network traffic interactions. A future study may incorporate more complex and realistic network flow dynamics together with other factors

such as weather conditions and car traffic. In this context, we may also empirically examine the day-to-day travel choices and system stability from different perspectives, including not only departure time choices, but also route choices and mode choices.

Besides, this study takes an aggregate approach to study the impacts of transit service reliability/unreliability on travelers' day-to-day travel choices. It should be noted that losing individual attributes of travelers can affect the modeling accuracy. If additional travelers' attributes become available, additional parameters or group-specific parameters can be adopted in the day-to-day cost perception and choice updating modeling framework proposed in this paper. Doing so can better capture the traveler heterogeneity and potentially improve modeling accuracy and also provide additional insights regarding mobility patterns of different traveler groups. It is of our interest to further examine this in a future study when a more comprehensive dataset becomes available.

**Acknowledgment.** The authors would like to thank the anonymous referees for their thoughtful and detailed comments, which have helped improve this paper substantially. Dr. Wei Liu acknowledges the support from the Australian Research Council through the Discovery Early Career Researcher Award (DE200101793).

## Appendix A.

**A.1. Proof for Lemma 1.** To prove Lemma 1 that $\mathbf{J_p}$ is a negative semidefinite matrix, we start with establishing $\mathbf{J_p}$, which is the Jacobian matrix of $\mathbf{P}(\tilde{\mathbf{c}}^*)$, i.e.,

$$\mathbf{J_p} = -\theta\mathbf{A} = -\theta \cdot \begin{bmatrix} \mathrm{P}_1^*(1-\mathrm{P}_1^*) & \dots & -\mathrm{P}_1^*\mathrm{P}_m^* & \dots & -\mathrm{P}_1^*\mathrm{P}_M^* \\ \vdots & \ddots & \vdots & & \vdots \\ -\mathrm{P}_m^*\mathrm{P}_1^* & \dots & \mathrm{P}_m^*(1-\mathrm{P}_m^*) & \dots & -\mathrm{P}_m^*\mathrm{P}_M^* \\ \vdots & & \vdots & \ddots & \vdots \\ -\mathrm{P}_M^*\mathrm{P}_1^* & \dots & -\mathrm{P}_M^*\mathrm{P}_m^* & \dots & \mathrm{P}_M^*(1-\mathrm{P}_M^*) \end{bmatrix} \tag{a1}$$

where the matrix $\mathbf{J_p}$ is the product of $-\theta$ and the matrix $\mathbf{A}$ and $\mathrm{P}_m^* = \frac{\exp(-\theta\tilde{c}_m^*)}{\sum_{z=1}^{M}\exp(-\theta\tilde{c}_z^*)}$ for $\forall m \in V$ . It is evident that $\mathbf{A}$ is a symmetric matrix with positive diagonal entries. Also, $\mathbf{A}$ is a zero-row-sum matrix. This means that $\mathbf{A}$ is a positive semidefinite matrix. Since $-\theta < 0$, $\mathbf{J_p}$ is negative semidefinite. This completes the proof. □

**A.2. Proof for Proposition 1.** Firstly, $\mathbf{D}$ can be expressed as $\mathbf{D} = d\mathbf{I}$. At the fixed point, $\mathbf{x}^*$ and $\mathbf{DP}(\mathbf{c}(\mathbf{x}^*))$ are both bounded from below and above by $\{\underbrace{0,...,0,...0}_{M}\}^{\mathrm{T}}$ and $\{\underbrace{d,...,d,...d}_{M}\}^{\mathrm{T}}$. Also, $\mathbf{DP}(\mathbf{c}(\mathbf{x}^*))$ is continuous as $\mathbf{P}(\cdot)$ and $\mathbf{c}(\cdot)$ are continuous. It follows that $\mathbf{DP}(\mathbf{c}(\mathbf{x}^*))$ maps the convex set $\mathbf{x}^*$ from a closed and bounded set to the same set. Therefore, all the hypotheses of the Brouwer's existence theorem are satisfied and Eq. (8) has at least one solutions.

At the fixed point, $\mathbf{x}^*$ is pinned down by $\mathbf{h}(\mathbf{x}) = 0$, where $\mathbf{h}(\mathbf{x}) = \mathbf{DP}(\mathbf{c}(\mathbf{x})) - \mathbf{x}$. The Jacobian matrix of $\mathbf{h}(\mathbf{x})$ is termed as $\mathbf{Jac}(\mathbf{h}(\mathbf{x})) = \mathbf{D}\mathbf{J_p}\mathbf{J_c} - \mathbf{I}$. $\mathbf{Jac}(\mathbf{h}(\mathbf{x}))$ is a negative definite matrix at the fixed point $\mathbf{x}^*$ as $\mathbf{J_p}$ is a negative semidefinite matrix, $\mathbf{J_c}$ and $\mathbf{D}$ are positive definite matrices. Thus, the dynamical system defined in Eqs. (3) and (6) has a unique flow and the travel cost pair $(\mathbf{x}^*, \tilde{\mathbf{c}}^*)$ at the fixed point. □

**A.3. Proof for Lemma 2.** Denote the eigenvector associated with the eigenvalue $\gamma_k$ by $\boldsymbol{\nu}_k$, hence, $\mathbf{DJ_pJ_c}\boldsymbol{\nu}_k = \gamma_k\boldsymbol{\nu}_k$. Also, define $\boldsymbol{\nu}'_k = \begin{bmatrix} a\boldsymbol{\nu}_k \\ \mathbf{DJ_P}\boldsymbol{\nu}_k \end{bmatrix}$, which satisfies

$$\mathbf{J}_{\text{LPUP}}\boldsymbol{\nu}'_k = \begin{bmatrix} (a+\kappa\gamma_k)\boldsymbol{\nu}_k \\ (a(1-\rho)+\rho+\kappa(1-\rho)\gamma_k)\,\mathbf{DJ_p}\boldsymbol{\nu}_k \end{bmatrix}$$

Hence if $a$ and $\lambda$ satisfy Eq. (a2), $\lambda$ is then an eigenvalue of $\mathbf{J}_{\text{LPUP}}$ and $\boldsymbol{\nu}'_k$ is the corresponding eigenvector.

$$\begin{cases} a+\kappa\gamma_k = a\lambda \\ a(1-\rho)+\rho+(1-\rho)\gamma_k\kappa = \lambda \end{cases} \tag{a2}$$

From Eq. (a2) by eliminating $a$, we can conclude that the relationship between $\lambda$ and $\gamma_k$ can be described as in Eq. (10). For each of the $M$ eigenvalues of $\gamma_k$ of matrix $\mathbf{DJ_pJ_c}$, two eigenvalues $\lambda'_k = \lambda_k$ and $\lambda''_k = \lambda_{M+k}$ of matrix $\mathbf{J}_{\text{LPUP}}$ are the solutions of the quadratic equation defined in Eq. (10); and in this way, we can obtain all the $2M$ eigenvalues of $\mathbf{J}_{\text{LPUP}}$ This completes the proof. □

**A.4. Proof for Proposition 2.** Since $\mathbf{DJ_pJ_c}$ is negative semidefinite, $\gamma_k \le 0,\ \forall k$. For each $\gamma_k$, $\lambda'_k = \lambda_k$ and $\lambda''_k = \lambda_{M+k}$ satisfy

$$\text{tr}\{\lambda_k\} = \lambda'_k + \lambda''_k = [\rho + \gamma_k\kappa\,(1-\rho)+1] \text{ and } \det\{\lambda_k\} = \lambda'_k\lambda''_k = \rho > 0$$

where $\text{tr}\{\lambda_k\}$ is the trace of the matrix $\mathbf{J}_{\text{LPUP}}$, and $\det\{\lambda_k\}$ is the determinant of the matrix $\mathbf{J}_{\text{LPUP}}$.

Based on the sign of $(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\}$, the following cases can arise.

(1) $(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\} \ge 0$ and $\text{tr}\{\lambda_k\} \ge 0$, both $\lambda'_k$ and $\lambda''_k$ are real, and $\lambda'_k$,$\lambda''_k = (\text{tr}\{\lambda_k\} \pm \sqrt{(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\}}/2) \ge 0$. In this case, $|\lambda'_k| < 1$ and $|\lambda''_k| < 1$ requires that

$$\begin{cases} \frac{\text{tr}\{\lambda_k\}+\sqrt{(\text{tr}\{\lambda_k\})^2-4\det\{\lambda_k\}}}{2} < 1 \\ (\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\} \ge 0 \\ \text{tr}\{\lambda_k\} \ge 0 \end{cases}$$

(2) $(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\} \ge 0$ and $\text{tr}\{\lambda_k\} < 0$, both $\lambda'_k$ and $\lambda''_k$ are real, and $\lambda'_k$,$\lambda''_k = (\text{tr}\{\lambda_k\} \pm \sqrt{(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\}}/2) \le 0$. In this case, $|\lambda'_k| < 1$ and $|\lambda''_k| < 1$ requires that

$$\begin{cases} \frac{\text{tr}\{\lambda_k\}-\sqrt{(\text{tr}\{\lambda_k\})^2-4\det\{\lambda_k\}}}{2} > -1 \\ (\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\} \ge 0 \\ \text{tr}\{\lambda_k\} < 0 \end{cases}$$

(3) $(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\} < 0$, $\lambda'_k$ and $\lambda''_k$ are of a complex conjugate pair, with $|\lambda'_k| = |\lambda''_k| = (\det\{\lambda_k\})^{\frac{1}{2}} < 1$ requires that

$$(\text{tr}\{\lambda_k\})^2 - 4\det\{\lambda_k\} < 0$$

We can reduce the above three cases to the following two cases.

$$\begin{cases} \mathrm{tr}\{\lambda_k\} - \det\{\lambda_k\} < 1 \\ \mathrm{tr}\{\lambda_k\} \leq 2 \\ \mathrm{tr}\{\lambda_k\} \geq 0 \end{cases} \Rightarrow 0 < -\gamma_k \leq \frac{\rho+1}{(1-\rho)\kappa}$$

or

$$\begin{cases} \mathrm{tr}\{\lambda_k\} + \det\{\lambda_k\} > -1 \\ \mathrm{tr}\{\lambda_k\} \geq -2 \\ \mathrm{tr}\{\lambda_k\} \leq 0 \end{cases} \Rightarrow \frac{\rho+1}{(1-\rho)\kappa} \leq -\gamma_k < \frac{2(\rho+1)}{(1-\rho)\kappa}$$

Combining the above two cases, we can conclude that the moduli of $\lambda_k^{'}$ and $\lambda_k^{''}$ are both less than 1, if and only if

$$|\gamma_k| < \frac{2(1+\rho)}{(1-\rho)\kappa}$$

□

**A.5. Proof for Proposition 3.** At the fixed point, $\mathbf{x}^*$ and $\mathbf{F}(\mathbf{x}^*)$ are both bounded from the below and above by $\{\underbrace{0, ..., 0, ...0}_{M}\}^{\mathrm{T}}$ and $\{\underbrace{d, ..., d, ...d}_{M}\}^{\mathrm{T}}$ Also, $\mathbf{F}(\mathbf{x}^*)$ is continuous as $\mathbf{c}(\cdot)$ is continuous. It follows that $\mathbf{F}(\mathbf{x}^*)$ maps the convex set $\mathbf{x}^*$ from a closed and bounded set to the same set. Therefore, all the hypotheses of the Brouwer's existence theorem are satisfied and Eq.(15) has at least one solution.

Let $\mathbf{J_F} = \mathbf{Jac}(\mathbf{F}(\mathbf{x}^*))$ denote the Jacobian matrix of the flow at the fixed point $\mathbf{x}^*$. According to Eq.(13) and Eq. (16), $\mathbf{J_F}$ can be written as can be written as

$$\mathbf{J_F} = \mathbf{I} + \alpha \begin{bmatrix} -\mathrm{x}_1^* \mathrm{c}'(\mathrm{x}_1^*) & ... & 0 & ... & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & ... & -\mathrm{x}_m^* \mathrm{c}'(\mathrm{x}_m^*) & ... & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & ... & 0 & ... & -\mathrm{x}_M^* \mathrm{c}'(\mathrm{x}_M^*) \end{bmatrix} \tag{a3}$$

According to Eq. (a3), it is evident that $\mathbf{J_F} - \mathbf{I}$ is a negative definite matrix. Therefore, the fixed point $\mathbf{x}^*$ is pinned down by $\mathbf{h}(\mathbf{x}) = 0$, where $\mathbf{h}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{x}$, which is a monotonically decreasing function. Then, the unique flow $\mathbf{x}^*$ is always ensured if the following two conditions, i.e., $\mathbf{x} \to \mathbf{0} \Rightarrow \mathbf{c}(\mathbf{x}) < \mathbf{E}$ and $\mathbf{x} \to \mathbf{d} \Rightarrow \mathbf{c}(\mathbf{x}) > \mathbf{E}$, hold. This completes the proof. □

**A.6. Proof for Proposition 4.** $\lambda_{\mathrm{f}}$ is the eigenvalue of the Jacobian matrix $\mathbf{J}_{\mathrm{PSAP}}$. To ensure the stability of the dynamical system, the maximum modulus $|\lambda_{\mathrm{f}}|$ among all the eigenvalues of the Jacobian matrix $\mathbf{J}_{\mathrm{PSAP}}$ should be less than one, i.e.,

$$\max_{\mathrm{f}=1,...,M} |\lambda_{\mathrm{f}}| < 1$$

Since $\mathbf{J}_{\mathrm{PSAP}}$ is a diagonal matrix, the elements on the diagonal are the eigenvalues of $\mathbf{J}_{\mathrm{PSAP}}$. Then,

$$\max_{\mathrm{m}=1,\ldots,M}|1+\alpha(1-\rho)\left(-\mathrm{x}_m^*\mathrm{c}'(\mathrm{x}_m^*)\right)|<1$$

which means that,

$$-1<1+\alpha(1-\rho)\left(-\mathrm{x}_m^*\mathrm{c}'(\mathrm{x}_m^*)\right)<1\ (\forall m\in V)$$

It follows that

$$-\frac{2}{\alpha(1-\rho)}<-\mathrm{x}_m^*\mathrm{c}'(\mathrm{x}_m^*)<0\ (\forall m\in V)$$

□

## References


Bates, J., Polak, J., Jones, P., and Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, 37(2-3):191–229.

Ben-Akiva, M., Cyna, M., and De Palma, A. (1984). Dynamic model of peak period congestion. *Transportation Research Part B: Methodological*, 18(4-5):339–355.

Bie, J. and Lo, H. K. (2010). Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation. *Transportation Research Part B: Methodological*, 44(1):90–107.

Bifulco, G. N., Cantarella, G. E., Simonelli, F., and Velonà, P. (2016). Advanced traveller information systems under recurrent traffic conditions: Network equilibrium and stability. *Transportation Research Part B: Methodological*, 92:73–87.

Cantarella, G. (1993). Day-to-day dynamics in transportation networks: stability and limits of equilibrium in a two-link network. *Sistemi Urbani*, 15(1):27–50.

Cantarella, G. E. and Cascetta, E. (1995). Dynamic processes and equilibrium in transportation networks: towards a unifying theory. *Transportation Science*, 29(4):305–329.

Cantarella, G. E. and Watling, D. P. (2016). A general stochastic process for day-to-day dynamic traffic assignment: formulation, asymptotic behaviour, and stability analysis. *Transportation Research Part B: Methodological*, 92:3–21.

Carey, M. (1994). Reliability of interconnected scheduled services. *European Journal of Operational Research*, 79(1):51–72.

Carrel, A., Halvorsen, A., and Walker, J. L. (2013). Passengers' perception of and behavioral adaptation to unreliability in public transportation. *Transportation Research Record*, 2351(1):153–162.

Cascetta, E. (1989). A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research Part B: Methodological*, 23(1):1–17.

Cascetta, E. and Cantarella, G. E. (1991). A day-to-day and within-day dynamic stochastic assignment model. *Transportation Research Part A: General*, 25(5):277–291.

Cascetta, E. and Cantarella, G. E. (1993). Modelling dynamics in transportation networks: State of the art and future developments. *Simulation Practice and Theory*, 1(2):65–91.

Chen, X., Yu, L., Zhang, Y., and Guo, J. (2009). Analyzing urban bus service reliability at the stop, route, and network levels. *Transportation Research Part A: Policy and Practice*, 43(8):722–734.

Cheng, Q., Wang, S., Liu, Z., and Yuan, Y. (2019). Surrogate-based simulation optimization

approach for day-to-day dynamics model calibration with real data. *Transportation Research Part C: Emerging Technologies*, 105:422–438.

Davis, G. A. and Nihan, N. L. (1993). Large population approximations of a general stochastic traffic assignment model. *Operations Research*, 41(1):169–178.

Douglas, N. and Jones, M. (2016). Developing a suite of demand parameters for inner sydney public transport. In *Australasian Transport Research Forum 2016 Proceedings*, pages 16–18.

Friesz, T. L., Bernstein, D., Mehta, N. J., Tobin, R. L., and Ganjalizadeh, S. (1994). Day-to-day dynamic network disequilibria and idealized traveler information systems. *Operations Research*, 42(6):1120–1136.

Friesz, T. L., Han, K., Neto, P. A., Meimand, A., and Yao, T. (2013). Dynamic user equilibrium based on a hydrodynamic model. *Transportation Research Part B: Methodological*, 47:102–126.

Guo, R.-Y., Yang, H., and Huang, H.-J. (2018). Are we really solving the dynamic traffic equilibrium problem with a departure time choice? *Transportation Science*, 52(3):603–620.

Guo, R.-Y., Yang, H., Huang, H.-J., and Li, X. (2017). Day-to-day departure time choice under bounded rationality in the bottleneck model. *Transportation Research Procedia*, 23:551–570.

Guo, X. and Liu, H. X. (2011). Bounded rationality and irreversible network change. *Transportation Research Part B: Methodological*, 45(10):1606–1618.

Habib, K. M. N., Kattan, L., and Islam, M. T. (2011). Model of personal attitudes towards transit service quality. *Journal of Advanced Transportation*, 45(4):271–285.

He, X., Guo, X., and Liu, H. X. (2010). A link-based day-to-day traffic assignment model. *Transportation Research Part B: Methodological*, 44(4):597–608.

He, X. and Liu, H. X. (2012). Modeling the day-to-day traffic evolution process after an unexpected network disruption. *Transportation Research Part B: Methodological*, 46(1):50–71.

Horowitz, J. L. (1984). The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological*, 18(1):13–28.

Iryo, T. (2008). An analysis of instability in a departure time choice problem. *Journal of Advanced Transportation*, 42(3):333–358.

Jackson, W. B. and Jucker, J. V. (1982). An empirical study of travel time variability and travel choice behavior. *Transportation Science*, 16(4):460–475.

Jiang, Y. and Szeto, W. (2016). Reliability-based stochastic transit assignment: Formulations and capacity paradox. *Transportation Research Part B: Methodological*, 93:181–206.

Jin, W.-L. (2020). Stable day-to-day dynamics for departure time choice. *Transportation Science*, 54(1):42–61.

Lam, W. H., Shao, H., and Sumalee, A. (2008). Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply. *Transportation Research Part B: Methodological*, 42(10):890–910.

Levinson, D. and Zhu, S. (2013). A portfolio theory of route choice. *Transportation Research Part C: Emerging Technologies*, 35:232–243.

Li, C., Bai, L., Liu, W., Yao, L., and Waller, S. T. (2019). Passenger demographic attributes prediction for human-centered public transport. In *International Conference on Neural Information Processing*, pages 486–494. Springer.

Li, C., Bai, L., Liu, W., Yao, L., and Waller, S. T. (2021). Urban mobility analytics: A deep temporal-spatial product neural network for traveler attributes inference. *Transportation Research Part C: Emerging Technologies*, 124:102921.

Li, H., Bliemer, M. C., and Bovy, P. H. (2009a). Modeling departure time choice under stochastic networks involved in network design. *Transportation Research Record*, 2091(1):61–69.

Li, X., Liu, W., and Yang, H. (2018). Traffic dynamics in a bi-modal transportation network with information provision and adaptive transit services. *Transportation Research Part C: Emerging Technologies*, 91:77–98.

Li, Z.-C., Lam, W. H., and Wong, S. (2009b). The optimal transit fare structure under different market regimes with uncertainty in the network. *Networks and Spatial Economics*, 9(2):191–216.

Liu, W. and Geroliminis, N. (2017). Doubly dynamics for multi-modal networks with park-and-ride and adaptive pricing. *Transportation Research Part B: Methodological*, 102:162–179.

Liu, W., Li, X., Zhang, F., and Yang, H. (2017). Interactive travel choices and traffic forecast in a doubly dynamical system with user inertia and information provision. *Transportation Research Part C: Emerging Technologies*, 85:711–731.

Liu, W. and Szeto, W. Y. (2020). Learning and managing stochastic network traffic dynamics with an aggregate traffic representation. *Transportation Research Part B: Methodological*, 137:19–46.

Lo, H. K., Luo, X., and Siu, B. W. (2006). Degradable transport network: travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B: Methodological*, 40(9):792–806.

Ma, X., Liu, C., Wen, H., Wang, Y., and Wu, Y.-J. (2017). Understanding commuting patterns using transit smart card data. *Journal of Transport Geography*, 58:135–145.

Mahmassani, H. S. and Chang, G.-L. (1987). On boundedly rational user equilibrium in transportation systems. *Transportation Science*, 21(2):89–99.

McZgee, V. E. and Carleton, W. T. (1970). Piecewise regression. *Journal of the American Statistical Association*, 65(331):1109–1124.

Morency, C., Trepanier, M., and Agard, B. (2007). Measuring transit use variability with smart-card data. *Transport Policy*, 14(3):193–203.

Prashker, J. N. (1979). Direct analysis of the perceived importance of attributes of reliability of travel modes in urban travel. *Transportation*, 8(4):329–346.

Rietveld, P., Bruinsma, F. R., and Van Vuuren, D. J. (2001). Coping with unreliability in public transport chains: A case study for Netherlands. *Transportation Research Part A: Policy and Practice*, 35(6):539–559.

Shao, H., Lam, W. H., Tam, M. L., and Yuan, X.-M. (2008). Modelling rain effects on risk-taking behaviours of multi-user classes in road networks with uncertainty. *Journal of Advanced Transportation*, 42(3):265–290.

Siu, B. W. and Lo, H. K. (2008). Doubly uncertain transportation network: degradable capacity and stochastic demand. *European Journal of Operational Research*, 191(1):166–181.

Small, K. A., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4):1367–1382.

Smith, M. J. (1984). The stability of a dynamic model of traffic assignment—an application of a method of lyapunov. *Transportation Science*, 18(3):245–252.

Smith, M. J. and Watling, D. P. (2016). A route-swapping dynamical system and lyapunov function for stochastic user equilibrium. *Transportation Research Part B: Methodological*, 85:132–141.

Sun, L. and Axhausen, K. W. (2016). Understanding urban mobility patterns with a probabilistic tensor factorization framework. *Transportation Research Part B: Methodological*, 91:511–524.

Szeto, W., Jiang, Y., Wong, K., and Solayappan, M. (2013). Reliability-based stochastic transit assignment with capacity constraints: Formulation and solution method. *Transportation Research Part C: Emerging Technologies*, 35:286–304.

Szeto, W. Y., Solayappan, M., and Jiang, Y. (2011). Reliability-based transit assignment for congested stochastic transit networks. *Computer-Aided Civil and Infrastructure Engineering*, 26(4):311–326.

Tu, H., Li, H., Van Lint, H., and van Zuylen, H. (2012). Modeling travel time reliability of freeways using risk assessment techniques. *Transportation Research Part A: Policy and Practice*, 46(10):1528–1540.

Van Loon, R., Rietveld, P., and Brons, M. (2011). Travel-time reliability impacts on railway passenger demand: a revealed preference analysis. *Journal of Transport Geography*, 19(4):917–925.

Wang, Y., Szeto, W. Y., Han, K., and Friesz, T. L. (2018). Dynamic traffic assignment: A review of the methodological advances for environmentally sustainable road transportation applications. *Transportation Research Part B: Methodological*, 111:370–394.

Wardman, M. (2001). A review of british evidence on time and service quality valuations. *Transportation Research Part E: Logistics and Transportation Review*, 37(2-3):107–128.

Watling, D. (1999). Stability of the stochastic equilibrium assignment problem: a dynamical systems approach. *Transportation Research Part B: Methodological*, 33(4):281–312.

Watling, D. P. and Cantarella, G. E. (2015). Model representation & decision-making in an ever-changing world: the role of stochastic process models of transportation systems. *Networks and Spatial Economics*, 15(3):843–882.

Xiao, Y. and Lo, H. K. (2016). Day-to-day departure time modeling under social network influence. *Transportation Research Part B: Methodological*, 92:54–72.

Xu, G., Liu, W., and Yang, H. (2018). A reliability-based assignment method for railway networks with heterogeneous passengers. *Transportation Research Part C: Emerging Technologies*, 93:501–524.

Yang, F. and Zhang, D. (2009). Day-to-day stationary link flow pattern. *Transportation Research Part B: Methodological*, 43(1):119–126.

Yang, L. and Lam, W. H. (2006). Probit-type reliability-based transit network assignment. *Transportation Research Record*, 1977(1):154–163.

Ye, H., Xiao, F., and Yang, H. (2018). Exploration of day-to-day route choice models by a virtual experiment. *Transportation Research Part C: Emerging Technologies*, 94:220–235.

Ye, H. and Yang, H. (2017). Rational behavior adjustment process with boundedly rational user equilibrium. *Transportation Science*, 51(3):968–980.

Yu, Y., Han, K., and Ochieng, W. (2020). Day-to-day dynamic traffic assignment with imperfect information, bounded rationality and information sharing. *Transportation Research Part C: Emerging Technologies*, 114:59–83.

Zhang, C., Liu, T.-L., Huang, H.-J., and Chen, J. (2018). A cumulative prospect theory approach to commuters' day-to-day route-choice modeling with friends' travel information. *Transportation Research Part C: Emerging Technologies*, 86:527–548.

Zhang, D. and Nagurney, A. (1996). On the local and global stability of a travel route choice adjustment process. *Transportation Research Part B: Methodological*, 30(4):245–262.

Zhang, F. and Liu, W. (2020). Responsive bus dispatching strategy in a multi-modal and multi-directional transportation system: A doubly dynamical approach. *Transportation Research Part C: Emerging Technologies*, 113:21–37.

Zhang, Y., Lam, W. H., Sumalee, A., Lo, H. K., and Tong, C. (2010). The multi-class schedule-based transit assignment model under network uncertainties. *Public Transport*, 2(1-2):69–86.

Zhu, Z., Li, X., Liu, W., and Yang, H. (2019). Day-to-day evolution of departure time choice in stochastic capacity bottleneck models with bounded rationality and various information perceptions. *Transportation Research Part E: Logistics and Transportation Review*, 131:168–192.