Stochastic collective model of public transport passenger arrival process

Article *in* IET Intelligent Transport Systems · November 2018

DOI: 10.1049/iet-its.2018.0085

Minh Le Kieu

University of Auckland

IET Intelligent Transport Systems

Special Issue: Selected Papers from the 25th ITS World Congress Copenhagen 2018

Stochastic collective model of public transport passenger arrival process

ISSN 1751-956X
Received on 26th January 2018
Revised 17th April 2018
Accepted on 24th May 2018
E-First on 13th August 2018
doi: 10.1049/iet-its.2018.0085
www.ietdl.org

Le Minh Kieu[1] ✉, *Chen Cai*[1]
[1]*Data61|CSIRO, Eveleigh, Australia*
✉ *E-mail: leminh.kieu@data61.csiro.au*

Abstract: It is essential to understand how transit passengers arrive at stops, as it enables transit operators and researchers to anticipate the number of waiting passengers at stops and their waiting time. However, the literature focuses more on predicting the total passenger demand, rather than simulating individual passenger arrivals to transit stops. When an arrival process is required especially in public transport planning and operational control, existing studies often assume a deterministic uniform arrival or a homogeneous Poisson process to model this passenger arrival process. This study generalises the homogeneous Poisson process (HPP) to a more general non-HPP (NHPP) in which the arrival rate varies as a function of time. The proposed *collective* NHPP (cNHPP) simulates the passenger arrival using less time regions than the HPP, takes less time to compute, while providing more accurate simulations of passenger arrivals to transit stops. The authors first propose a new time-varying intensity function of the transit passenger arrival process and then a maximum likelihood estimation method to estimate the process. A comparison study shows that the proposed cNHPP is capable of capturing the continuous and stochastic fluctuations of passenger arrivals over time.

1 Introduction

Passenger demand plays an essential role in both long-term planning and short-term operational control of public transport. Passenger demand significantly affects transit operations, because transit vehicles have to stop for passengers boarding and alighting, which increases the dwell time at stops. The question of modelling how passengers arrive at a transit stop becomes one of the major research interests for public transport studies. The number of passengers who arrived and their arrival times at stops is essential inputs to estimate passenger waiting time for transit vehicles. The total or mean waiting time is often used as the main objective function for public transport planning and operation studies [1–5].

An analytical approach to model the passenger arrival process is often required in these public transport planning and operation control studies. Each passenger arrival is modelled as an event, where researchers are usually interested in the number of arrivals between two successive transit vehicles. A simulation of the passenger arrival process is required for this purpose, where an intensity or arrival rate is estimated. The arrival rate is the number of passengers who arrived at a transit stop per time unit.

There are generally two approaches to model the passenger arrivals using the known arrival rate: (i) a deterministic or (ii) a stochastic point process. First, the deterministic approach assumes that passengers arrive uniformly at transit stops so that the number of boarding/arrived passengers is simply the product of the passenger arrival rate and the time headway between consecutive vehicles. This approach has been used in many earlier studies such as Eberlein *et al.* [1] and Fu *et al.* [2]. Daganzo [6] and Daganzo and Pilachowski [7] also use a variation of this approach, where a dimensionless parameter is used to represent the marginal increase in transit delay resulting from a unit increase in headway. Second, the stochastic point process approach assumes that passengers arrive randomly at stops with a stable arrival rate. In the majority of existing studies, this stochastic point process is a homogeneous Poisson process (HPP), which aims to simulate the passenger arrival times using only the arrival rate and the time interval between consecutive vehicle arrivals, no matter when the interval starts. HPP is widely used to model systems with stochastic events, such as modelling the presence of connected vehicles in traffic [8] or traffic incidents [9]. An emerging number of existing studies in public transport have also adopted this stochastic approach, such as Fu and Yang [10], Toledo *et al.* [11] and Kieu *et al.* [12]. There is considerable evidence that the assumptions of stochastic HPP processes for passenger arrivals are reasonable for high-frequency services, such as those with scheduled headway between 10 and 15 min [4]. At longer headways, there is another less popular line of research regarding passengers who time their arrivals with the schedule and service reliability [13, 14]. In this paper, we assume that passengers do not consult the schedule prior to arrival at transit stops, thus the use of a stochastic point process such as HPP remains valid.

Existing HPP models in the literature assume a stable passenger arrival rate or intensity that does not change over time. However, it is clear that the passenger arrival rate varies with time. The morning and afternoon peak periods should have much higher arrival rates compared to off-peak periods. The arrival rate also changes gradually from low to high and from increasing to decreasing. A common approach to include time into consideration is to define exogenous time regions. In each region, the passenger arrival rate is then constant, and a different HPP can be applied to each region. This approach has limited accuracy because the passenger arrival process is not fully continuous time-dependent, but rather multiple independent HPP superimposed [15]. This approach can be seen as a discretisation of the continuous arrival rate into multiple discrete constant arrival rates. However, because the arrival rate in each region is still constant, the accuracy of this method is limited to the number of time regions we can fit and combine. Non-HPP (NHPP), which allows the arrival rate to be continuous time-dependent, is a substantial advance from the HPP in terms of versatility and accuracy to model the passenger arrival process. However, not only is the implementation of NHPP complex, it is still not possible for a single NHPP intensity function to model the passenger arrival process of a whole day. It is still necessary to use a system of multiple time regions to model the passenger arrival process, but we can use an NHPP in each time region. This approach would require less time regions, while potentially providing a more versatile approach to model the passenger arrival process.

This paper proposes an NHPP to model the stochastic and time-dependent passenger arrival process at transit stops. We relax the assumption in the classical case where the rate of arrivals of

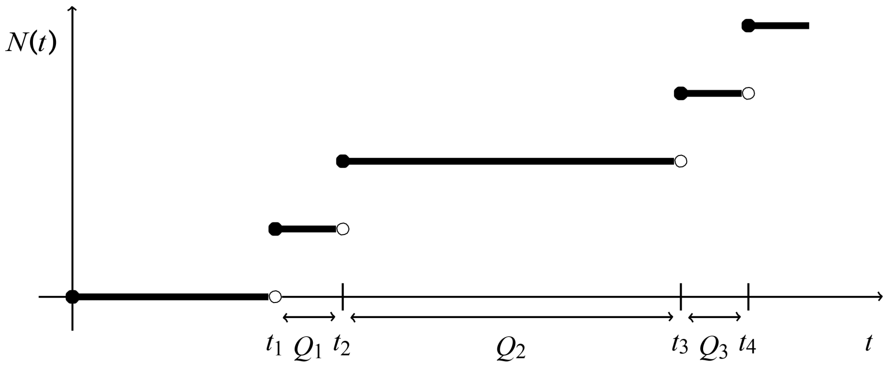

Fig. 1 *Illustration of the arrival process*

passenger demand is constant over time. We inherently model the rate of arrival of passengers as a continuous function of time, t. This would account for the fact that there are more passengers arriving at a certain time period (say morning), than others (say midnight). We present an insightful framework to model time-dependent passenger demand and propose algorithms to perform a collective simulation of passenger arrival rate on multiple time regions and (c) simulate the NHPP model for validation and comparison with other approaches. The inference and simulation of the proposed NHPP model will use a two-day record Smart Card dataset of Sydney, Australia, where we use the first day to train the model, and the second day for testing. The proposed NHPP model will be applied to the whole day from 0:00 to midnight. We also explicitly consider the differences between different time-of-the-day by a series of change points, where the passenger arrival rate is changed. We call this approach the *collective* NHPP or *cNHPP*. The contributions of this paper are twofold: (a) we propose a new intensity function for modelling the passenger arrival process to transit stops, and (b) we propose a procedure for identifying change points so that a continuous time-varying passenger arrival process can be simulated.

The remaining of this paper includes a representation of HPP and NHPP, the proposed inference method for collective NHPP and a numerical experiment.

2 Fundamental representation of the arrival process

In this section, we briefly recap the fundamentals of *arrival process* and *Poisson process*, which would be used to model the process of passengers arriving at transit stops. The following section serves as the building block for realistic modelling of passenger arrival process to be extended in later sections, and to include periodicities in demands.

2.1 Arrival process and HPP

An *arrival process* is a sequence of inter-arrival times at which each event happens, which we shall denote by $t_1, t_2, \ldots$. For example, t_1 represents the time when passenger 1 arrives at a transit stop, t_2, represents the time the following passenger arrives and so on. t_k can usually be interpreted as the time of occurrence of the kth event, in this case – the kth arrival. In this paper, we refer to t_i as arrival times. t_i is a non-negative random variable, satisfying $0 < t_1 < t_2 < \cdots$, where $t_i < t_{i+1}$.

Fig. 1 illustrates the arrival process, where t_i is arrival time and Q_i is the inter-arrival time. Q_i can be interpreted as follows: the first passenger arrives at time t_1, the second arrives at Q_1 after the first and so on. $N(t)$ can be understood as a *counting process* of the arrival process, where the number of passengers arrived within a time period can be retrieved if the arrival process is modelled. This is useful for transit operations and management, where transit operators can estimate the number of waiting passengers at a transit stop given the forecasted time gap between two consecutive transit vehicles.

The simplest class of arrival process is the *HPP*. It can be defined as follows:

Definition 1: ***(HPP):*** *Let* $(Q_k)_{k \geq 1}$ be a sequence of independent and identically distributed exponential random variables with constant parameter λ and event times $t_n = \sum_{k=1}^{n} Q_i$. The process $(N_t, t \geq 0)$ defined by $N_t := \sum_{k \geq 1} 1_{\{t \geq t_k\}}$ is called a Poisson process with intensity λ.

Where $1_{\{\cdot\}}$ is the indicator function that takes the value 1 when the condition is true, 0 otherwise. We can see that $N_0 = 0$. N_t is piecewise constant and has jump size of 1 at the event times t_i. One can show that it means the following proposition [16].

Proposition 1: (Ross, 1995, p.64, Proposition 2.2.1): The inter-arrival times $Q(i)$ with $k = 1, 2, \ldots$ of an HPP with rate $\lambda > 0$ are independent identically distributed exponential random variables with mean $1/\lambda$.

The proof of this proposition can be found in [16]. Assume that r time units have elapsed and during this period no events have arrived, i.e. there are no events during the time interval $[0, r]$. The probability that we will have to wait for a further t time units given by

$$\begin{aligned} \mathbb{P}(Q > t + r \mid Q > r) &= \frac{\mathbb{P}(Q > t + r, Q > r)}{\mathbb{P}(Q > r)} \\ &= \frac{\mathbb{P}(Q > t + r)}{\mathbb{P}(Q > r)} = \frac{\exp(-\lambda(t + r))}{\exp(-\lambda r)} \\ &= \exp(-\lambda t) = \mathbb{P}(Q > t)\,. \end{aligned} \quad (1)$$

Equation (1) is said to have no memory and it is a special property of the Poisson process. That is, the likelihood and chance to wait an additional t time units after already having waited m time units is the same as the probability of having to wait t time units when starting at time 0. Putting it differently, if one interprets Q as the time of arrival of an event where Q follows an exponential distribution, the distribution of $Q - m$ given $Q > m$ is the same as the distribution of Q itself. Therefore, we can naturally come up with the following algorithm to simulate the passenger arrival times t_i in an HPP by generating the inter-arrival times Q_i and taking the sum $t_i = \sum_{k=1}^{\max} Q_k$. The algorithm is described in Fig. 2.

The *HPP* is therefore stochastic and probabilistic, but with a constant arrival rate λ.

2.2 Non-HPP

The HPP, as we defined it so far, is simply characterised by a *constant* arrival rate λ. It is equivalent to an assumption, e.g. that public transport passengers arrival rate to stops is the same regardless of the time being midnight or peak periods. It is more useful to extend the Poisson process to a more general arrival process in which the arrival rate varies as a function of time. Note that the intensity usually depends on the arrival time, not just on the inter-arrival time. We can define this type of process as a non-homogeneous Poisson process (NHPP).

Input: λ, T

Initialise $n = 0, t_0 = 0$;

while *True* **do**

  Generate $u \sim uniform(0,1)$;

  Let $Q = -\ln u/\lambda$;

  Set $t_{n+1} = t_n + Q$;

  if $t_{n+1} > T$ **then**

    Return $t_k | k = 1, 2, \ldots, n$

  else

    Set $n = n + 1$

Fig. 2 *Algorithm 1: Simulation of an HPP with intensity λ on [0, T]*

Input: $\lambda(t), T$

Initialise $n = m = 0, t_0 = s_0 = 0$,

$\bar{\lambda} = \sup_{0 \le t \le T} \lambda(t)$;

while $s_m < T$ **do**

  Generate $u \sim uniform(0,1)$;

  Let $Q = -\ln u/\bar{\lambda}$;

  Set $s_{m+1} = s_n + Q$;

  Generate D uniform(0,1);

  if $D \le \lambda(s_{m+1})/\bar{\lambda}$ **then**

    $t_{n+1} = s_{m+1}$;

    $n = n + 1$;

  $m = m + 1$

if $t_n \le T$ **then**

  Return $t_k | k = 1, 2, \ldots, n$

else

  Return $t_k | k = 1, 2, \ldots, n - 1$

Fig. 3 *Algorithm 2: Simulation of a NHPP with intensity λ(t) on [0, T]*

Definition 2: The point process N is said to be an NHPP with intensity function $\lambda(t) \ge 0 \quad t \ge 0$; if

$$\begin{aligned} \mathbb{P}(N_{t+h} = n + m | N_t = n) &= \lambda(t)h + o(h) \text{ if } m = 1, \\ \mathbb{P}(N_{t+h} = n + m | N_t = n) &= o(h) \text{ if } m > 1, \\ \mathbb{P}(N_{t+h} = n + m | N_t = n) &= 1 - \lambda(t)h + o(h) \text{ if } m = 0. \end{aligned} \tag{2}$$

The NHPP is a generation of HPP where the intensity λ is now not a constant, but a function of time $\lambda(t)$. NHPP, in general, would be a better candidate for the modelling passenger arrival process to transit stops, because we can vary the passenger arrival rate as a function of time, e.g. the peak periods should have higher arrival rates than the off-peak period. Analogously to Proposition 1, the following proposition can be obtained [17].

Proposition 2: If the arrival process N be an NHPP with intensity function $\lambda(t)$, then $N(t)$ follows a Poisson distribution with a parameter $\int_0^t \lambda_u \mathrm{d}u$, i.e.

$$\mathbb{P}(N_t = n) = \frac{1}{n!}\exp\left(-\int_0^t \lambda_u \mathrm{d}u\right)\left(-\int_0^t \lambda_u \mathrm{d}u\right)^n \tag{3}$$

The proof for Proposition 2 can be found in [17]. Let $n = 0$ we find the probability of no arrival within the interval $[a, b]$, which determines the law of occurrence for the next arrival

$$\mathbb{P}(N(a, b] = 0) = \exp\left(-\int_a^b \lambda_u \mathrm{d}u\right) \tag{4}$$

In a simple arrival process, arrivals occur one by one and by checking the condition in (4), the arrivals can be simulated sequentially. The simulation procedure in Algorithm 2 (see Fig. 3) follows this idea and is often referred as Lewis' thinning algorithm [17].

3 Methodology

We have now described some of the theory of arrival process that is suitable to model the stochastic and time-varying passenger demand. This section proposes a parametric form for the rate of demand for passengers to be used in an NHPP model.

3.1 Proposed time-varying arrival rate (intensity) function

Equation (5) shows the formulation of the time-varying arrival rate $\lambda(t)$ of the proposed cNHPP.

$$\lambda_t = pc^p t^{p-1} + \varepsilon, \tag{5}$$

where $c>0$ and $p \in R$. The parameter ε is usually taken to be fixed and acts as a parameter such that the rate never goes negative (bounded away from zero), since a negative rate of demand is nonsensical. Fig. 4 shows a plot of this intensity. It can be easily noted that this is a generalisation of the HPP, where the rate can be constant (similar to HPP) or varies over time. This function is versatile for several reasons. When the parameter $p = 1$, it reduces to a constant and we know from above that this specifies the parameter for the exponential random variables. If this is respected then the data follows a Poisson process. If on the other hand, in the case where $p<1$, this gives a decreasing curve (see Fig. 5). We interpret this as the rate of demand decreasing. Finally, our choice of intensity function can also handle the case when $p>1$ – this corresponds to an increasing rate of demand. We summarise the following description below:

- it reduces to a constant when $p = 1$, and hence is able to recover Poisson process should the data respects this,
- when $p<1$, the rate of demand is decreasing,
- when $p>1$, the rate of demand is increasing.

3.2 Parameter estimation

Recall that if N is an NHPP equipped with an intensity function $\lambda(t)$, then N_t follows a Poisson *distribution* with parameter $\int_0^t \lambda_u \mathrm{d}u$. Exploiting this definition together with the availability of demand from Smart Card data, we are interested in performing inference over the set of parameters of $\Theta := (c, p)$ in our chosen intensity function as in (5). To this end, we use the maximum likelihood estimation (MLE) method. The MLE begins with writing down an expression known as the likelihood function. Since the logarithm is a monotonic increasing function, we can equivalently write down the log of this likelihood function, which we denote henceforth as $\mathscr{L}$. Roughly speaking, the likelihood of a set of observed data is the probability of obtaining that particular set of data, given the chosen intensity model, and in our case, it is defined in (5). Observe that the function $\mathscr{L}$ contains the unknown model parameters. The values of these parameters that maximise the sample likelihood are known as the maximum likelihood estimates (MLEs).

Smart Card data gives the observed samples $\mathscr{H} - = \{t_1, t_2, \ldots, t_n\}$ where t_i represents the arrival time of a passenger at timestamp i. We write the log-likelihood $\mathscr{L}$ of the Poisson distribution with a $\int_0^t \lambda_u \mathrm{d}u$ parameter where λ follows that of (5):

$$\mathscr{L}(c, p|\mathscr{H}) = -n \cdot \int_0^{\lambda_u \mathrm{d}u} + \int_0^{\lambda_u \mathrm{d}u} \cdot \sum_{i=1}^{n} x_i \,. \tag{6}$$

We find the optimal c and p by maximising the log-likelihood, i.e.

$$\underset{c,p}{\arg\max}\ \mathscr{L}(c, p|\mathscr{H})\,. \tag{7}$$

3.3 Collective NHPP

Despite the proposed intensity function $\lambda(t)$ (5) being a very flexible function, there is still a major challenge in the implementation of this function to model the passenger arrival process of the whole day. Equation (5) can only describe a convex upward or downward passenger arrival rate. It cannot be used to describe an arrival pattern that has a change point, i.e. where the arrival process change in the trend from upward to downward, or *vice versa*. For instance, the arrival rate of public transport passengers at transit stops may keep increasing from midnight until the middle of the morning peak period (e.g. 8 AM), then drop until increasing again at the start of the afternoon peak period. No single parametric model would be able to fit the arrival process with a 'twist' between increasing and decreasing the arrival rate. The whole-day data should be modelled using several discrete intensity function $\lambda(t)$, where each function explains a distinct time region. The change points are considered as the connection points to merge these individual time regions into a complete framework for cNHPP. We propose the following procedure to describe the change point determination process:

- *Step 1:* Plot the cumulative recurrences versus timestamps of passenger arrivals to visually identify the number of change points required K. This step requires a graph-based identification of the number of changes in the passenger arrival rate pattern from upward to downward, and *vice versa*.

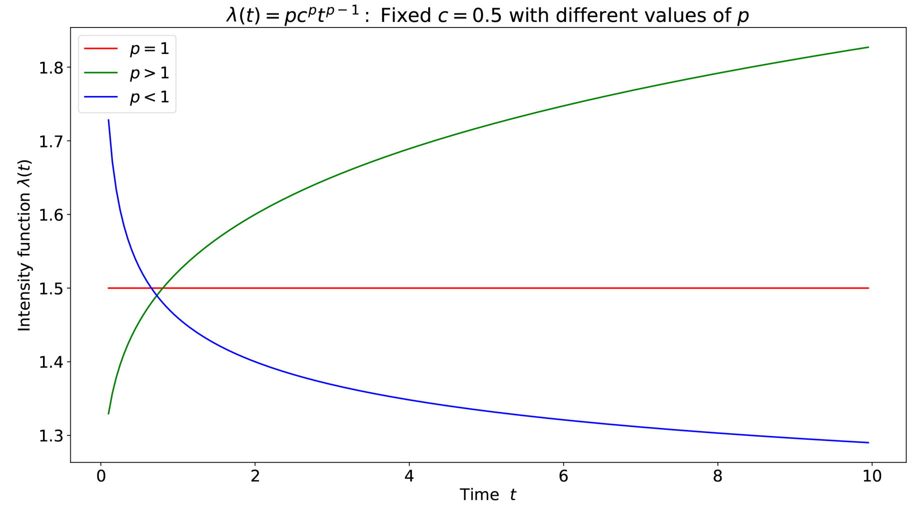

Fig. 4 *Proposed intensity function λ against time*

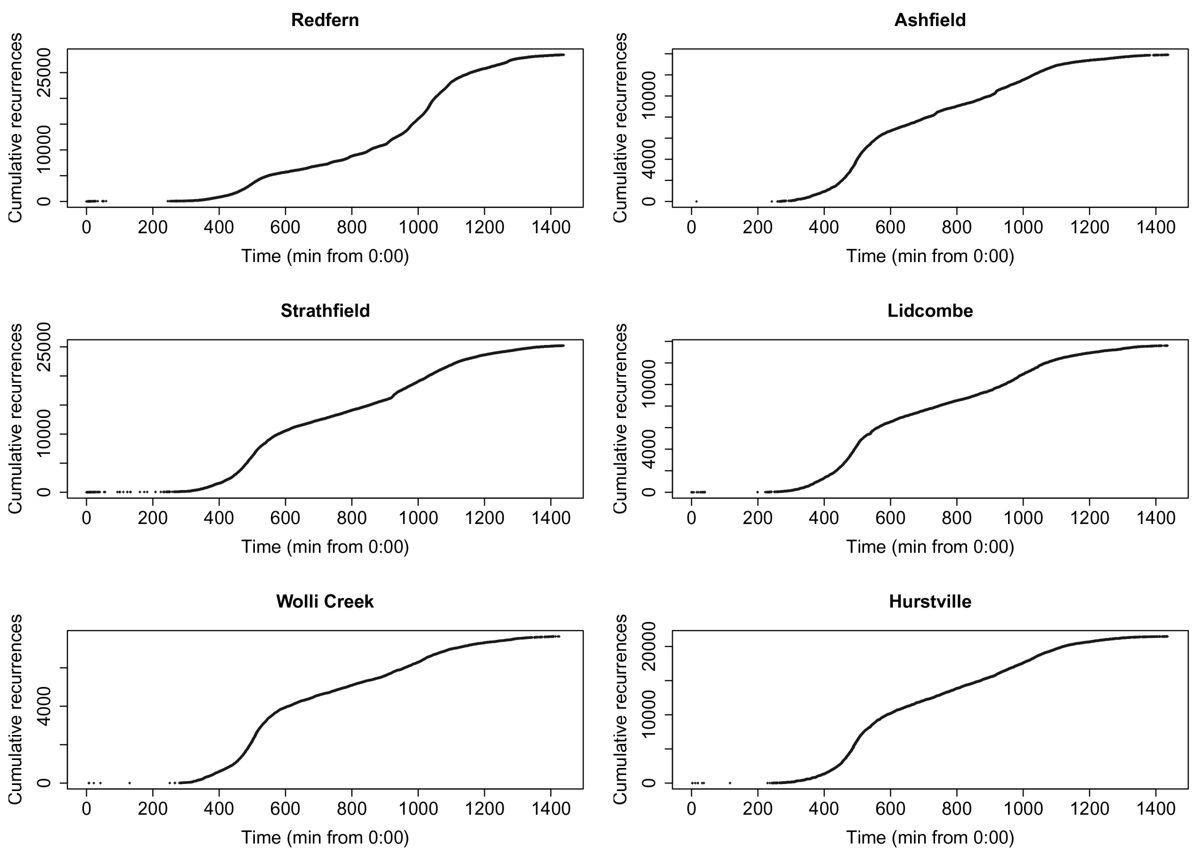

Fig. 5 *Cumulative recurrences of passenger arrivals at train stations*

- *Step 2:* For each candidate change point K_i, define a range $\left[R_i^{\min}, R_i^{\max}\right]$ such that $K_i \in \left[R_i^{\min}, R_i^{\max}\right] \mid K_j \notin \left[R_i^{\min}, R_i^{\max}\right]$, $i \neq j$.
- *Step 3:* Randomly initialise the starting point for K_i in $\left[R_i^{\min}, R_i^{\max}\right]$.
- *Step 4:* Iteratively maximise the sum of log-likelihood $\mathscr{L}(c, p \mid \mathscr{H})$ from each time region $\left[R_i^{\min}, K_i\right]$ and $\left[K_i, R_i^{\max}\right]$ against the value K_i. This is equivalent to solving the maximum likelihood optimisation problem in (7) using multiple values of K_i, until convergence. We adopt the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm to iteratively solve this problem [18].
- *Step 5:* Repeat steps 2 to 4 until the last change point.

Please note the cumulative recurrence plot is the plot of the cumulative count of passenger arrivals at a train station versus their timestamps of arrival. We shall see an observed cumulative recurrence plot in the next section.

4 Dataset

In order to have the passenger arrival data T_i that will be used for modelling, this paper uses two days of Smart Card data from Sydney, Australia. The data consists of the timestamps when a passenger arrives at train stations in Sydney. We focus on modelling the passenger arrival process at six popular train stations. The stations have been chosen with two major criteria: (a) they should be a busy train station with many high frequency train lines, and (b) there should not be many other activities within the train station so that there are few other activities such as shopping after entering the station. Fig. 6 illustrates the location of six chosen train stations used for the case study. The stations (from left to right and then top to bottom) are Lidcombe, Strath-field, Ashfield, Redfern, Wolli Creek and Hurstville Station.

4.1 Evidence of departure from HPP

Fig. 7 shows the one-dimension kernel density estimation using passenger arrival data T_i of the six stations in the case study. It can be noted that the passenger demand varies over time. The morning peak and afternoon peak periods show significantly higher demand than other time periods. Recall that the HPP assumes that the passenger arrival rate is constant in time. Fig. 7 clearly show a dynamic time-dependent demand pattern that should be addressed by a time-varying function of demand.

To further investigate the observed passenger arrival data, Fig. 5 demonstrates the cumulative recurrences of passenger arrival throughout a day. Fig. 5 shows another sign of the time-varying passenger arrival rate because if the arrival rate was constant, we should have seen a straight line for the cumulative recurrence figure. Figs. 5 and 7 give evidence that the rate of passenger demand changes over time. The arrival rate is clearly not homogeneous, as it is not constant over the period of observation.

5 Numerical experiments

This section shows the experimental results and compares the results from cNHPP with the common approaches in the literature. To compare our proposed process with other common methods for modelling the passenger arrival process in the literature, we also develop two other common processes in the literature of public transport:

- *Uniform arrival:* The passenger arrival process is deterministic and uniformly distributed.
- *Homogeneous Poisson arrival (HPP):* The passenger arrival process is stochastic with a stable intensity function that does not change over time. The parameters are estimated by averaging the total number of passengers over the study period.

5.1 Comparison of result for a short time period

We first consider the case where only a short time period is considered, so no change points are required. We estimate the parameters of the proposed cNHPP by solving the maximum likelihood optimisation problem in (7) to find (c, p) using observed Smart Card data. Similar to the previous section, the BFGS algorithm is employed to iteratively solve this problem [18]. The HPP is estimated similarly using a simple intensity function $\lambda = D$ and then employ MLE to find D.

After the parameters for HPP and cNHPP are estimated, we then simulate the passenger arrivals using Algorithms 1 and 2 (Figs. 2 and 3). The simulation for uniform arrival process can be easily done by generating even arrivals using the estimated arrival rate. Note that the simulation for uniform arrival process is

Fig. 6 *Train stations for the case study in Sydney, NSW, Australia. Map data from OpenStreetMap*

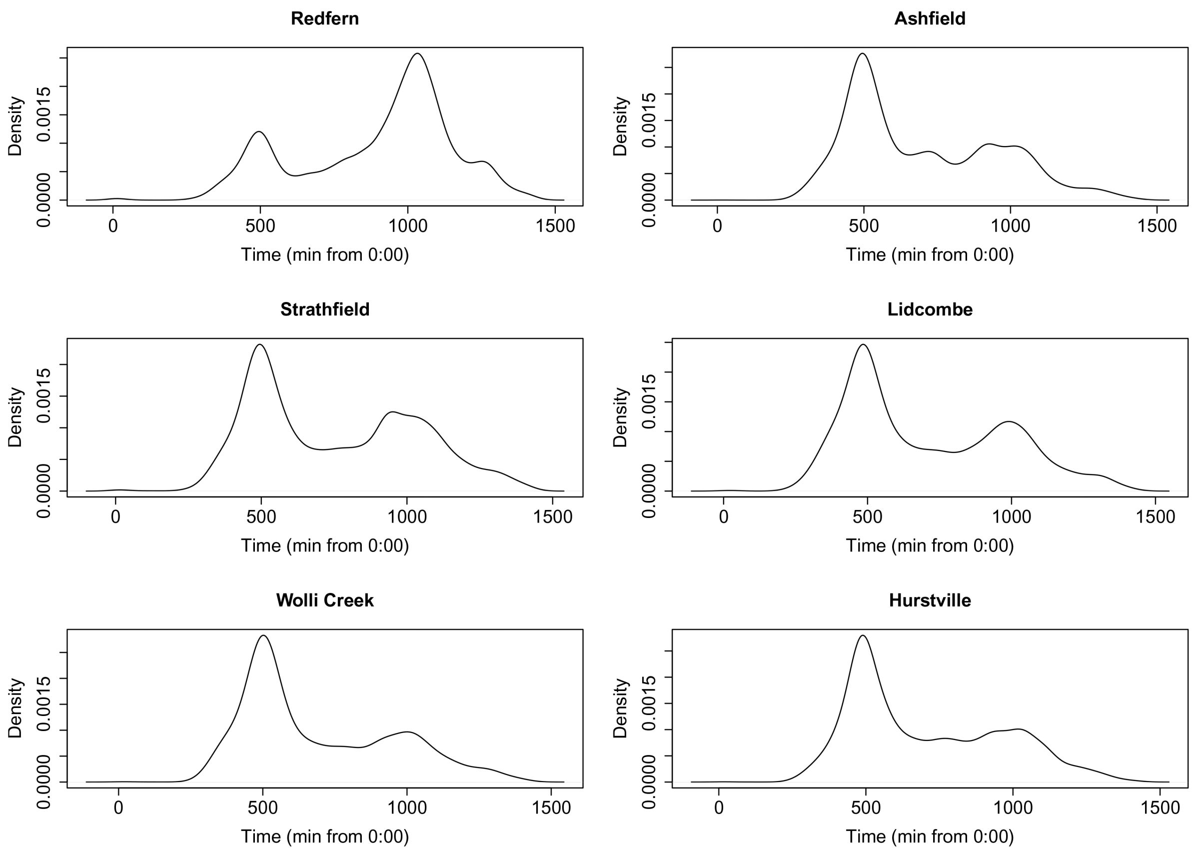

Fig. 7 *Passenger demand at the six train stations in the case study*

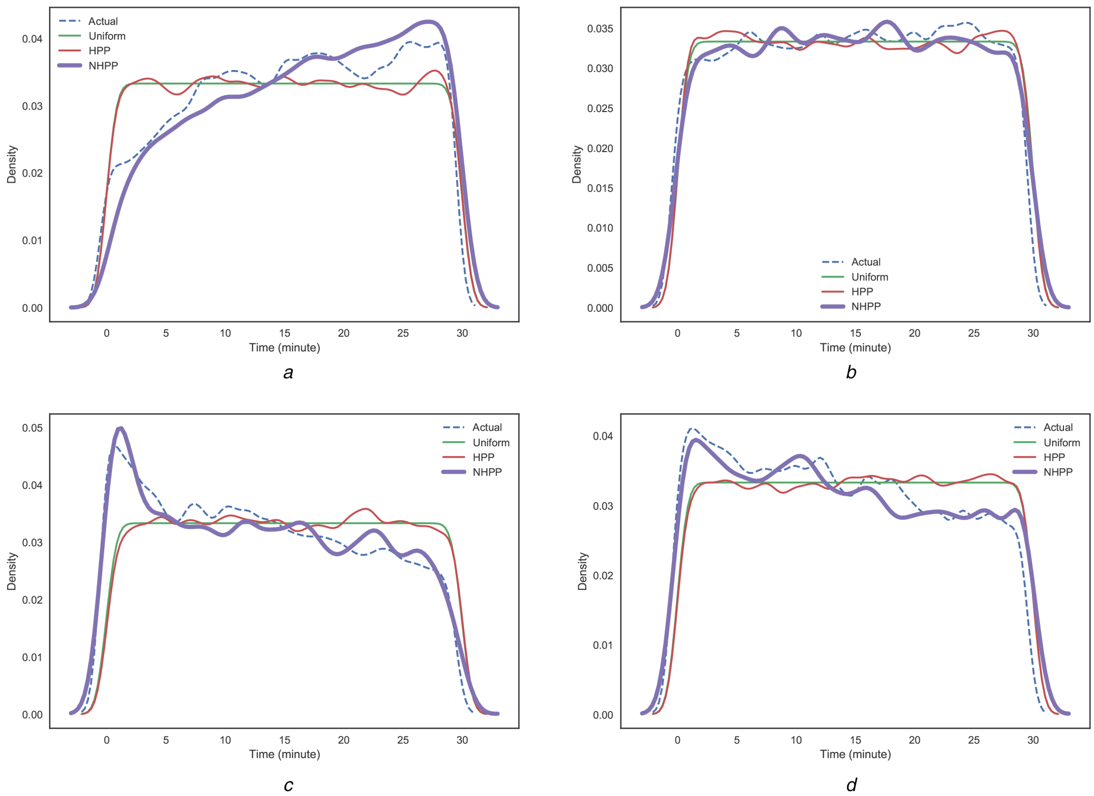

Fig. 8 *Comparison of simulation results*
(a) 7:00–7:30 AM, ***(b)*** 7:30–8:00 AM, ***(c)*** 8:30–9:00 AM, ***(d)*** 9:00–9:30 AM

deterministic, i.e. it generates the same arrival times every time we simulate. Fig. 8 compares the kernel density estimation of the simulations from the uniform arrival process, HPP and cNHPP over four 30-minute intervals compared to the observed arrival data to one of the train stations in the case study.

The uniform arrival process only provides a deterministic distribution, where passengers uniformly arrive. HPP provides very similar results to the uniform arrival process, but with some stochasticity. This result shows that HPP is actually not much different from a simple assumption of uniform arrival. It also suggests that HPP is not good enough for modelling and simulating

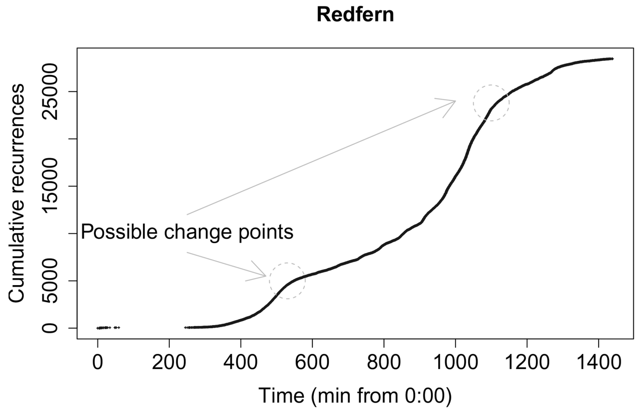

Fig. 9 *Cumulative recurrences of passenger arrivals at train stations*

the passenger arrival process unless there are many change points considered. On the other hand, the probability density function of simulations from the proposed cNHPP shows a very good fit with the actual arrival data. As time increases from 7:00 to 7:30 AM, cNHPP captures the fact that passengers are arriving more at transit stops as time passes. At the second interval in Fig. 8*b*, the arrival rate hits the plateau, before the reduction in Figs. 8*c* and *d*. The results show that cNHPP provides a versatile and accurate simulation of the passenger arrival process to transit stops.

5.2 Comparison of results for a day

Recall that Figs. 5 and 7 show that the arrival process seems to have many stages when looking at the passenger arrival density for a whole day, and suggest the use of several change points. Take Refern station as an example, Fig. 9 shows the evidence that the cumulative recurrence plot is not 'smooth' and two change points should be considered. We can also see similar hints for other stations in Fig. 5.

Therefore, we add two or three change points K_i for each station when estimating the parameters for cNHPP. It means that there are three NHPP models using the intensity function in (5) developed for distinct time regions: $[0, K_1], [K_1, K_2], \ldots$, until $[K_{\text{CP}}, 1440]$ where 0 and 1440 are the first and last minutes of a day and CP is the number of change points. We estimate the parameters of these cNHPP models according to the MLE method described in Section 4.2.

Conversely, the HPP requires many more change points because the model can only model a linear intensity function and fails to model a convex function. It is necessary to add many more change points compared to cNHPP to emulate a continuous time-varying arrival rate when a whole day is considered. It is possible to model any time-varying passenger arrival process using HPP with many change points, but the computation cost would be too high for a system with many events such as our passenger arrival data. Therefore, to implement a piecewise HPP, we add a new change point every 120 min. Fig. 10 shows the results of cNHPP and HPP using 3 and 12 change points, respectively. The uniform arrival is very similar to HPP and is dropped out for visualisation purposes.

The vertical dotted line in Fig. 10 shows the location of cNHPP's change points for the six train stations, where the time is in minutes. As can be seen in the figure, there are three stations with two change points and the remaining have three change points. Fig. 10 also shows that the locations of change points are relatively similar across the stations. The first change point defines the first time region from midnight (time 0 min) until the end of the morning peak periods (time around 500–570 min). This first time region explains the constant arrival in the early morning, the sharp increase in the morning peak period, and the decrease after the morning peak time. The last change point defines the last time region from the start of the afternoon peak period (time around 950–1100 min, which is equivalent to around 4 PM). This time region explains the increase and decrease in passenger demand prior to and after the afternoon peak period. In the middle of these two change points, there may be another one if needed. Fig. 10 provides a suggestion to find the number of required change points and their prospective locations prior to applying the procedure in Section 3.

The HPP model is always estimated with 12 change points because it fails to model the increasing and decreasing trends in passenger arrival rate. Fig. 10 shows that while cNHPP fits well with the curves in cumulative recurrence plots, HPP often cut across the curves and discretises the cumulative recurrence plots. HPP also leaves a few blank time regions, such as the early time periods in Ashfield and Wolli Creek stations because their time regions are relatively small (120 min), so that there may be not enough data points for fitting a Poisson process. This issue is even more severe if transit operators want to implement HPP in every station, because some stations may have very limited demand. On the other hand, cNHPP only requires two to three change points, which means that the time region is generally large enough to contain sufficient data points.

The parameters for cNHPP and HPP are inferred using the Smart Card data of 30 August 2017. We then use the estimated parameters of cNHPP and HPP to predict individual passenger arrival times to the six stations on the next day 31 August 2017. Given the total number of passengers who arrived at each station during each time region, we use Algorithms 1 and 2 (Figs. 2 and 3) to simulate the individual passenger arrivals for exactly the same number of passengers on 31 August 2017. The simulated arrival data from cNHPP and HPP is then compared to the observed data from the Smart Card data. Mean absolute percentage error (MAPE) and computation time in minutes are used as the criteria for comparison. The MAPE is a well-known algorithm for calculating the percentage difference between observed and predicted values. It is calculated using the following equation:

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{\text{OBS}_t - \text{SIM}_t}{\text{OBS}_t} \right| \tag{8}$$

Where OBS_t and SIM_t are observed and simulated arrival times, respectively, and n is the total number of passengers who arrived at the studied train station. Table 1 shows the prediction error (MAPE) and the computation time of the proposed cNHPP and HPP. The parameter estimation of both cNHPP and HPP has been implemented in R language using the same MacBook Pro 2016. It is clear that cNHPP outperforms HPP in both MAPE and computation time. cNHPP is ∼42% better than HPP in terms of MAPE while costing 32% less in computation time. This is because cNHPP can fit convex curves, while HPP has to rely on fitting many discrete time regions to approximate a continuous passenger arrival process. Recall that the prediction performance of HPP can be improved by increasing the number of change points so that it would provide a better approximation of the continuous arrival process. However, this approach would even further increase the computation time. Therefore, it is clear that cNHPP is superior to HPP in both simulation performance and computation time.

6 Conclusion

Passenger arrival process is essential in public transport planning and operational control to anticipate the number of waiting passengers at stops and their waiting times. However, literature on public transport focuses more on estimating the total passenger demand, rather than simulating the individual passenger arrival process. When a passenger arrival process is required, existing studies in literature often adopt the rather simplistic uniform arrival process and HPP approaches. While being simple to implement, these approaches fail to truly represent the continuous arrival process. They often discretise the continuous arrival process into many discrete time regions where in each time region the arrival rate is constant.

This paper proposes an NHPP to simulate individual passenger arrivals to transit stops. The proposed model simulates the continuous arrival process using less discrete time regions than the HPP and also costs less in terms of computation time. In order to do so, we propose a new time-varying intensity function of transit

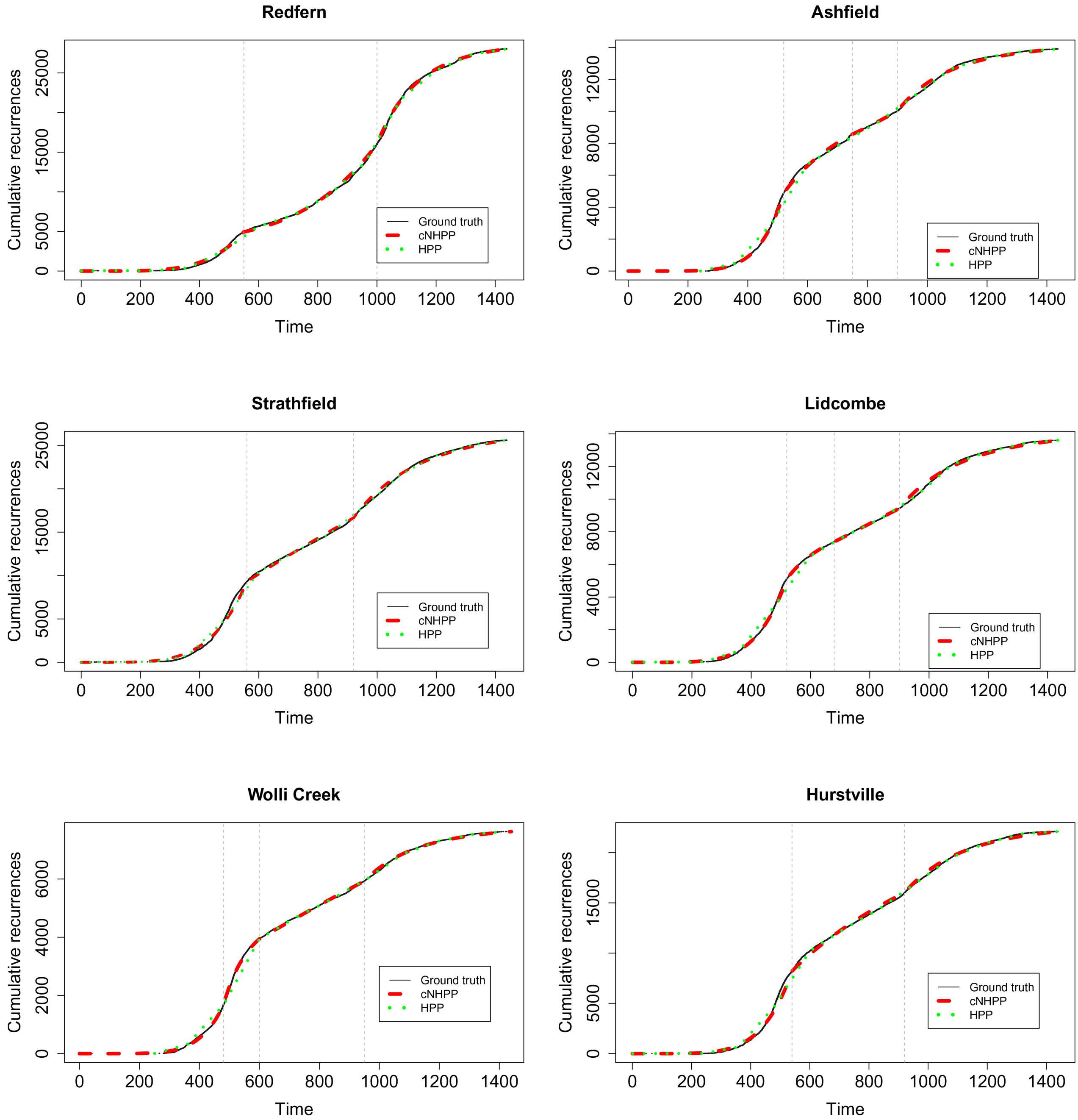

Fig. 10 *Comparison of simulation results*

Table 1 Comparison of MAPE and computation time between cNHPP and HPP

Stations	MAPE, %		Computation time, min	
	cNHPP	HPP	cNHPP	HPP
Redfern	**9.81**	16.96	**6.34**	9.22
Ashfield	**10.74**	17.38	**3.63**	5.17
Strathfield	**10.82**	16.42	**6.17**	8.97
Lidcombe	**11.57**	17.24	**3.76**	5.24
Wolli Creek	**10.36**	16.25	**2.04**	3.22
Hurstville	**9.29**	16.33	**3.89**	5.52

Bold values indicate the best performance.

passenger arrival rate, which is robust and flexible. We also propose an MLE method to estimate the parameters for this intensity function using the Smart Card data of Sydney, NSW, Australia. The comparison using the next day data shows that the proposed cNHPP is able to accurately capture the dynamic fluctuations of passenger arrivals over time.

One of the limitations of the current approach is that the consideration of change points still requires visual examination of the cumulative recurrence plot to find the possible number of change points and their boundaries. Future studies should solve this problem using an adaptive method to automatically identify the number of change points and their locations. In the meantime, this study is helpful for both transit operators and researchers in public transport planning and operational control.

7 Acknowledgments

This study was supported by Strategic Research Funding at the Data61, CSIRO, Australia and partly by the NSW Premier's Innovation Initiatives project. The data used in this study was provided by Transport for NSW, the transport operator of NSW, Australia. The authors express their gratitude to Dr. Yuming Ou at Data61, CSIRO for processing the data used in this study and Mr. Ben Itzstein at Data61, CSIRO for proof reading the final version of the manuscript.

8 References

[1] Eberlein, X.J., Wilson, N.H., Barnhart, C., *et al.*: 'The real-time deadheading problem in transit operations control', *Transp. Res. B, Methodol.*, 1998, **32**, (2), pp. 77–100

[2] Fu, L., Liu, Q., Calamai, P.: 'Real-time optimization model for dynamic scheduling of transit operations', *Transp. Res. Rec., J. Transp. Res. Board*, 2003, **1857**, pp. 48–55

[3] Delgado, F., Muñoz, J., Giesen, R., *et al.*: 'Real-time control of buses in a transit corridor based on vehicle holding and boarding limits', *Transp. Res. Rec., J. Transp. Res. Board*, 2009, **2090**, pp. 59–67

[4] Desaulniers, G., Hickman, M. D.: 'Public transit', *Handb. Oper. Res. Manage. Sci.*, 2007, **14**, pp. 69–127

[5] Hickman, M.D.: 'An analytic stochastic model for the transit vehicle holding problem', *Transp. Sci.*, 2001, **35**, (3), pp. 215–237

[6] Daganzo, C.F.: 'A headway-based approach to eliminate bus bunching: systematic analysis and comparisons', *Transp. Res. B, Methodol.*, 2009, **43**, (10), pp. 913–921

[7] Daganzo, C.F., Pilachowski, J.: 'Reducing bunching with bus-to-bus cooperation', *Transp. Res. B, Methodol.*, 2011, **45**, (1), pp. 267–277

[8] Wang, X.: 'Modeling the process of information relay through inter-vehicle communication', *Transp. Res. B, Methodol.*, 2007, **41**, (6), pp. 684–700

[9] Baykal-Gürsoy, M., Xiao, W., Ozbay, K.: 'Modeling traffic flow interrupted by incidents', *Eur. J. Oper. Res.*, 2009, **195**, (1), pp. 127–138

[10] Fu, L., Yang, X.: 'Design and implementation of bus-holding control strategies with real-time information', *Transp. Res. Rec., J. Transp. Res. Board*, 2002, 1791, pp. 6–12

[11] Toledo, T, Cats, O., Burghout, W., *et al.*: 'Mesoscopic simulation for transit operations', *Transp. Res. C, Emerg. Technol.*, 2010, **18**, (6), pp. 896–908

[12] Kieu, L.M., Bhaskar, A., Cools, M., *et al.*: 'An investigation of timed transfer coordination using event-based multi agent simulation', *Transp. Res. C, Emerg. Technol.*, 2017, **81**, pp. 363–378

[13] Fonzone, A., Schmöcker, J.-D., Liu, R.: 'A model of bus bunching under reliability based passenger arrival patterns', *Transp. Res. C, Emerg. Technol.*, 2015, **59**, pp. 164–182

[14] Bowman, L.A., Turnquist, M.A.: 'Service frequency, schedule reliability and passenger wait times at transit stops', *Transp. Res. A, Gen.*, 1981, **15**, (6), pp. 465–471

[15] Sayarshad, H.R., Chow, J.Y.: 'Survey and empirical evaluation of nonhomogeneous arrival process models with taxi data', *J. Adv. Transp.*, 2016, **50**, (7), pp. 1275–1294

[16] Ross, S.M.: 'Stochastic processes', 1996, 1996

[17] Lewis, P.A., Shedler, G.S.: 'Simulation of nonhomogeneous Poisson processes by thinning', *Nav. Res. Logist. (NRL)*, 1979, **26**, (3), pp. 403–413

[18] Avriel, M.: '*Nonlinear programming: analysis and methods*' (Dover Publications Inc., Mineola, NY, USA, 2003)