WILEY | Hindawi

Research Article

A $\mathrm{PH}(i)/\mathrm{PH}(i,n)/C/C$ Queuing Model in Randomly Changing Environments for Traffic Circulation Systems

Juanxiu Zhu [ID],[1] **Lu Hu** [ID],[2] **Han Xie** [ID],[1] **and Kehong Li** [ID][1]

[1]*School of Management, Xihua University, Chengdu, Sichuan 610039, China*
[2]*School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan 610031, China*

Correspondence should be addressed to Lu Hu; hulu@swjtu.edu.cn

Robust optimal design of circulation systems (e.g., roads for vehicles or corridors for pedestrians) relies on an accurate steady-state traffic flow model that considers the effect of randomly changing environmental factors (e.g., daily periodicity and weather). Most analytical models assume that the customer interarrival time and service time of circulation facilities follow the exponential distribution with fixed rate parameters, which is unrealistic in most cases. In this paper, we develop a stationary PH (i)/PH $((i,n)/(C/C))$ state-dependent queuing model in a randomly changing environment (RE), which is represented by a Markov chain. The model simultaneously considers the general randomness of arrival and service, the randomly varying rate parameters, and the state-dependent service (the travel time increases with the number of customers). The existing matrix analytic scheme (MAS) algorithm is extended to solve the proposed model because it avoids the explicit calculation of probability distributions. The space complexity of the algorithm is only linear in the number of RE states and is independent of the enormous (four-dimensional) state space of the Markov process. Its time complexity is a linear function of the product of the queue capacity and the number of RE states. Our model is validated versus simulation estimates. The obtained conditional performance measures can accurately capture the queue accumulation and dissipation and reveal the effect of randomly changing environments. Numerical experiments provide some interesting findings. (1) The proposed stationary model coincides with the transient $M(t)/G(x)/C/C$ fluid queuing model under special conditions. (2) Under high traffic intensities, increasing the randomness in the duration time of the RE state leads to an obvious growth in the conditional queue length. (3) An increase in the facility length leads to an increase or a decrease in the average output rate, depending on whether the congestion dissipates effectively in one cycle. (4) A larger width is required to obtain the maximum average output rate for traffic demand with a greater nonuniformity.

1. Introduction

Congestions in traffic circulation systems (e.g., roads for vehicles and corridors for pedestrians) have become a global issue that has raised extensive concern. For road traffic, congestion is getting worse across the world and congestion costs are rapidly increasing in recent years [1]. As for pedestrian traffic, several large-scale stampede accidents occurred in the crowd during the past decade and caused severe casualties [2]. Efficiency and safety are the primary goals when designing traffic circulation systems because of the increased congestion. A robust design of traffic circulation facilities (such as road, corridors, and passageway) that meets the demand of customers (vehicles or pedestrians) under various situations is the key to improving efficiency and guaranteeing safety. Optimal design decisions rely on an accurate steady-state traffic flow model comprehensively considering factors that affect the performance of traffic circulation systems. Therefore, accurately modeling traffic circulation systems becomes the primary task.

However, modeling traffic circulation systems in the real world is a great challenge as they possess several complex characteristics. First, both the traffic demand and the service ability of traffic circulation facilities are random. Second, the service ability of circulation facilities is state-dependent. State dependence describes the phenomenon where an increasing number of customers leads to slower velocity due to limited land space. Third, a congestion propagation effect

exists in a circulation network. Congestions in one facility will propagate in an upstream direction and affect the adjacent facilities. In addition, circulation systems are affected by randomly changing environmental factors such as daily periodicity, large events, accidents, and extreme weather conditions [3]. As a result, system parameters, such as the traffic demand and the service ability, vary randomly because of these environmental factors.

Scholars have proposed a lot of models for traffic circulation systems. The macroscopic traffic flow model focuses on the aggregate features of traffic flows. Microscopic models, such as the car-following models and lane-changing models for vehicles [4] and the microsimulation tools for pedestrians [5, 6, 7] can present more detailed characteristics but need more parameters and are less computationally efficient. To solve the accuracy-efficiency trade-off, mesoscopic models, such as the queuing model and the gas-kinetic model, are proposed. The queuing model is the most popular mesoscopic model applied in this area, and its application in modeling pedestrian or vehicle circulation systems has been demonstrated by Lovas [8], Parlar and Sharafali [9], Arita and Schadschneider [10], Tréca et al. [11], Hasani Goodarzi et al. [12], and so on. Queuing models can further be categorized into two groups: transient models and stationary models. Transient models, such as the $M(t)/G(x)/C/C$ model [13], aim to capture the dynamic performance and are often used in real-time management and control. In contrast with transient queuing models, stationary queuing models focus on obtaining the steady performance and are more useful in planning and design.

Most of the stationary queuing models in this field have considered randomness and state-dependence. For example, Yuhaski et al. [14] applied the exponential distribution to describe the random interarrival times and the service time depends on the number of customers (also called system state, denoted by n). Hu et al. [15] established a PH/PH(n)/C/C state-dependent queuing model where the interarrival time and service time with general randomness follow the phase-type (PH) distribution. PH distribution theoretically can approach to any positive random variable infinitely. Readers can also refer to the study of Neuts [16] and Buchholz et al. [17] for more details on PH distribution and PH-based queues. Although PH distribution can address the general randomness in arrival and service, the change in arrival or service rate is ignored in these research studies. Nowadays, more and more attention is paid to the varying parameters in traffic circulation systems, such as the study of [3], Gerum and Baykal-Gürsoy [18], and Yang et al. [19]. The robust design of transportation facilities relies on more realistic stationary models. The goal of this paper is to propose a stationary queuing model for traffic circulation systems by considering the general randomness, state dependence, and the effect of randomly changing environmental factors. We apply the PH distribution to address the general randomness in traffic demand and service ability. Meanwhile, the state-dependent service time is used to depict customers' velocity that depends on the number of customers in the system. To further consider the effect of the randomly changing environmental factors, we introduce a random environment (RE) represented by a Markov chain. The queuing system is modulated by the RE; hence, the interarrival time and the service time change accordingly. In this study, the congestion feedback effect is ignored. Moreover, we focus on circulation facilities that are roughly rectangular and have a traffic flow in one direction. Complex situations, such as circulation networks with a feedback effect or bidirectional traffic flow, will be discussed in our future work.

The contribution of the paper is twofold. The first contribution lies in proposing a PH (i)/PH(i, n)/C/C state-dependent queuing model in an RE. The proposed model comprehensively considers the general randomness, state dependence, and the effect of randomly changing environmental factors. Compared with existing microscopic simulation models, the proposed queuing model is time-efficient, easy to calibrate, and possesses a stable numerical solution. Compared with the stationary PH/PH(n)/C/C model in the findings Hu et al. [15] which applies fixed arrival and service rates, the proposed model applies arrival and service rates varying with RE states to address the effect of randomly changing environmental factors. Unlike the transient $M(t)/G(x)/C/C$ model based on the exponential distribution in the study of Hu et al. [13], our stationary model captures the steady performance and is more suitable for facility planning and design. Moreover, the proposed model is more practical owing to the superior characteristics of the PH distribution. Meanwhile, the duration time of each RE state in the proposed model is random as the RE is represented by a Markov chain, while the length of different phases is deterministic in the $M(t)/G(x)/C/C$ model which cannot reflect the randomness of environment states (e.g., incident durations and peak hours). Compared with the BMAP/PH/C and MAP/PH/C models in an RE in the studies of Kim et al. [20] and Wu et al. [21], the proposed model applies a state-dependent service time and a finite space capacity to account for the congestion in traffic circulation systems. The difference between this paper and some of the most related references is shown in Table 1.

The second contribution of the paper is that we extend the matrix analytic scheme (MAS) algorithm in the study of Baumann and Sandmann [32]. Due to the introduction of the state-dependence and the RE, the proposed model has a four-dimensional state space which is considerably large. Conventional methods based on the explicit computation of the stationary distribution, such as LU-decomposition, Gaussian elimination, and Gauss–Seidel method [33, 34, 35] and the extended reversed compound agent theorem method proposed by Balsamo and Marin [36] require massive computer storage. Compared with these methods, the efficiency of the MAS algorithm is improved remarkably as the explicit computation of the stationary distribution is not necessary. Therefore, we extend it to solve the proposed model in an RE. The space complexity of the extended MAS algorithm is only linear in the number of RE states and is independent of the enormous state space of the Markov process. Its time complexity is a linear function of the product of the queue capacity and the number of RE states. We also validate the proposed model and algorithm versus simulation estimates.

TABLE 1: Some most related studies and their difference with the current paper.

Authors	Performance	Randomness	State dependence	Varying parameters	RE
Liu and Whitt [22]	Transient	M		✓	
Ammar [23]	Transient	M		✓	
Hu et al. [13]	Transient	M	✓	✓	
Yuhaski et al. [14]	Stationary	M	✓		
Jain and Smith [24]	Stationary	M	✓		
Smith and Cruz [25]	Stationary	M	✓		
Hu et al. [15]	Stationary	PH	✓		
Zhu et al. [26]	Stationary	PH	✓		
Yechiali and Naor [27]	Stationary	M		✓	✓
Neuts [28]	Stationary	M		✓	✓
Baykal-Grsoy et al. [3]	Stationary	M		✓	✓
Yu and Liu [29]	Stationary	M		✓	✓
Krenzler and Daduna [30]	Stationary	M		✓	✓
Kim et al. [20]	Stationary	BMAP/PH		✓	✓
Wu et al. [21]	Stationary	MAP/PH		✓	✓
Kim et al. [31]	Stationary	MMAP/PH		✓	✓
The current paper	Stationary	PH	✓	✓	✓

The conditional performance measures of the proposed model provide insights into the varying performance of traffic circulation systems. Critical information can be revealed, such as the effect of randomly changing environmental factors and the formation, duration, and dissipation of congestions. When applying fixed arrival and service parameters, such information is greatly smoothed out. The practical value of the proposed model is embodies in offering a solid base for robust optimal design. As the proposed queuing model comprehensively considers several vital factors that influence the system performance, the design method based on it is expected to be superior in accommodating traffic demands under various situations.

The rest of this paper is organized as follows. In Section 2, we review the related works. Then, preparations for the model are presented in Section 3. The mathematical model is formulated in Section 4. Section 5 contains numerical experiments and Section 6 concludes the paper.

2. Literature Review

Queuing models with changing arrival or service rate are applied in various domains, for example, the time-dependent Markov queuing models in the study by Nelson and Taaffe [37]; the time-dependent fluid queues in the study by Liu and Whitt [22]; and queues modulated by an RE in the study by Baykal-Gürsoy et al. [3] and Kim et al. [20].

Queuing models in an RE provide a good choice for capturing the steady performance of queuing systems whose parameters change with environmental factors. An RE that represents the varying environmental factors has several states. For example, periods of heavy, medium, and light traffic can be regarded as an RE with three states (the number of RE states N is 3). Every RE state lasts a random period of time before transferring to another state. At a given RE state i, the system operates as a classic queue. When the RE transfers to another state, the queue will immediately change its parameters, such as arrival rate or service rate.

Queuing models in an RE have received considerable attention in the literature. Eisen and Tainiter [38] first considered a queuing process with two mean arrival and service rates. Queuing models in an RE where the interarrival time and service time are subject to exponential distribution M, such as M/M/1, M/M/C, and M/M/∞ queuing models were later explored by Yechiali and Naor [27]; Neuts [39]; O'cinneide and Purdue [40]; Krenzler and Daduna [30]; and Yu and Liu [29]. More recent works can be found in the study by Pang et al. [41] and Naumov and Samouylov [42]. Using the M/M/C queue in an RE in modeling traffic systems was conducted by Baykal-Gürsoy et al. [3]. Generalization work was carried out by Neuts [28], who discussed the M/G/1 model under an RE. Sztrik [43] studied the M/G/n blocking system in an RE. To account for the random arrival process with generality and diversity, queues with PH distribution, MAP (Markovian arrival process), and BMAP (batch Markovian arrival process) were further explored (e.g., Krieger et al. [44]; Kim et al. [20]; Wu et al. [21]; and Kim et al. [31].

As mentioned in Section 1, state dependence exists in most traffic circulation systems. Due to the limited land space, the velocity of vehicles or pedestrians is greatly affected by the number of vehicles or pedestrians in the circulation facility. State-dependent queuing models based on exponential distribution and their applications can be found in the study by Conway and Maxwell [45]; Yuhaski et al. [14]; Cheah and Smith [46]; Smith [47]; Mitchell and Smith [48]; Weiss et al. [49]; Jain and Smith [24]; and Smith and Cruz [25]. However, the exponential arrival interval is based on the hypothesis that the arrival process is a Poisson process. This means the variation coefficient of arrival interval equals 1, which is inconsistent with the reality in most cases according to Jiang et al. [50] and Chen et al. [51]. Given the strong hypothesis on the exponentially distributed interarrival time in the M/G (n)/C/C state-dependent queuing model, Jiang et al. [52] developed a discrete event simulation model based on a G/G(n)/C/C state-dependent queuing system. Hu et al. [15] established a PH/PH(n)/C/C

state-dependent queuing model for a pedestrian corridor and Zhu et al. [26] proposed the PH/PH(*n*)/*C*/*C* state-dependent queuing network model, which can handle inter-arrival times and service times with various squares of the coefficient of variation (abbreviated as SCV).

We aim to propose a stationary PH(*i*)/PH(*i*, *n*)/*C*/*C* state-dependent queuing model in an RE which comprehensively considers the general randomness in arrival and service, the state-dependent service time and the effect of randomly changing environmental factors. It is very difficult to solve the proposed PH(*i*)/PH(*i*, *n*)/*C*/*C* state-dependent queuing model in an RE. However, some efficient solutions provide a solid foundation. For example, the matrix analytic method is used in the study by Neuts [53] to solve the single *M*/*G*/1 queue in an RE. The extended reversed compound agent theorem based on the product-form theory is proposed in the study by Balsamo and Marin [36] for a continuous time Markov chain in an RE. Meanwhile, quasi-birth-and-death (QBD) processes have been developed to deal with a broader range of PH-based queuing models. The study by Baumann and Sandmann [32] applied the MAS algorithm to compute the stationary expectations of a level-dependent QBD process. The advantage of the MAS algorithm is that it does not explicitly compute the stationary distribution π. Therefore, we extend the MAS algorithm to solve the proposed model in this paper.

3. Preparations

In this section, we first present the problem. Then, some main assumptions concerning the queuing system and the calibration of parameters are discussed.

3.1. Problem Statement.
The traffic circulation facilities analyzed in this paper are presented in Figure 1. *l* and *w* are two dimension parameters. When the analyzed facility is for pedestrians (such as a corridor or hallway), *l* means length and *w* is the width (in meters). For road traffic, *l* is the length of the road segment in miles and *w* is the number of lanes in one direction. The capacity of the facility can be expressed as follows:

$$C = \lfloor \rho l w \rfloor, \tag{1}$$

where $\lfloor i \rfloor$ is the largest integer that is less than or equal to *i*; ρ is a constant determined by the density of the appropriate context. ρ is typically 5 peds/m^2 for pedestrian traffic [14] and ranges from 185 -265 veh/mile-lane for road traffic, according to different values of jam density [24]. Note that the customary unit in each field is used for convenience.

The customers arrive at the facility with interarrival times A and the service time is denoted as B. When A and B are described by a PH distribution, then the circulation system can be described as a PH/PH(*n*)/*C*/*C* state-dependent queuing system with *C* parallel-serial servers (details can be found in the study by Hu et al. [15]).

As mentioned before, the traffic queuing system is always affected by randomly changing environmental factors. The RE has several states and transfers randomly or circularly

among these RE states. When the RE transfers from one state to another, the parameters of the queuing system, such as arrival rates and service rates, will change. For example, the arrival rate will vary with the daily period of traffic flow, and the service rate will be affected by bad weather conditions or accidents. When the RE is introduced, the interarrival time distribution at RE state *i* is expressed as A_i and the service time distribution at RE state *i* is B_i. A traffic circulation system affected by the environmental factors can then be described as a PH(*i*)/PH(*i*, *n*)/*C*/*C* state-dependent queuing system in an RE. Our main objective is to establish the analytical queuing model for the PH(*i*)/PH(*i*, *n*)/*C*/*C* state-dependent queuing system in an RE and design an algorithm to solve it.

For convenience, the main notation is listed in Table 2. Note that there are two terms of state in this paper, system state *n* and RE state *i*. The former is the number of customers in a queuing system, while the latter is defined as the status of the RE. "State-dependent/dependence" always refers to the system state.

3.2. Main Assumptions.
The circulation facility we analyze is a straight horizontal facility. The pedestrians and vehicles are assumed to be evenly distributed in the facility. Although the model can address bidirectional or multidirectional passenger flows, we focus on a unidirectional traffic flow.

The circulation system is described as a "first in first out" queue, which means overtaking is not allowed. This assumption is appropriate, except for the situation where the traffic density is very small. Another assumption associated with the PH(*i*)/PH(*i*, *n*)/*C*/*C* queuing model is that an arriving customer will be rejected and lost when the system state is equal to the capacity *C*. This assumption is only accurate when the customer has the choice to leave the facility.

Additionally, the interarrival time A and the total service time B are assumed to follow the simplified PH distribution. The general PH distribution has a large number of parameters. Determining all these parameters (the number can be very large) is difficult in practice. However, the mean value (*u*) and the SCV (denoted as *c*) of a random variable can be obtained in most cases.

SCV, or the square of the coefficient of variation, is calculated by the following equation:

$$c = \left(\frac{\sigma}{u}\right)^2, \tag{2}$$

where *u* and σ are the mean and standard deviation of the random variable. When *c* = 0, the variables for the interarrival time and service time are constant; when *c* = 1, the interarrival times and service time can be described by an exponential distribution. The larger *c* is, the greater randomness exists.

As the mean value *u* and SCV *c* for a random variable are available in most cases, [35] and Weerstra [54] used a simplified PH, which is determined by *u* and *c*. Although not the emphasis of this paper, we will introduce the simplified
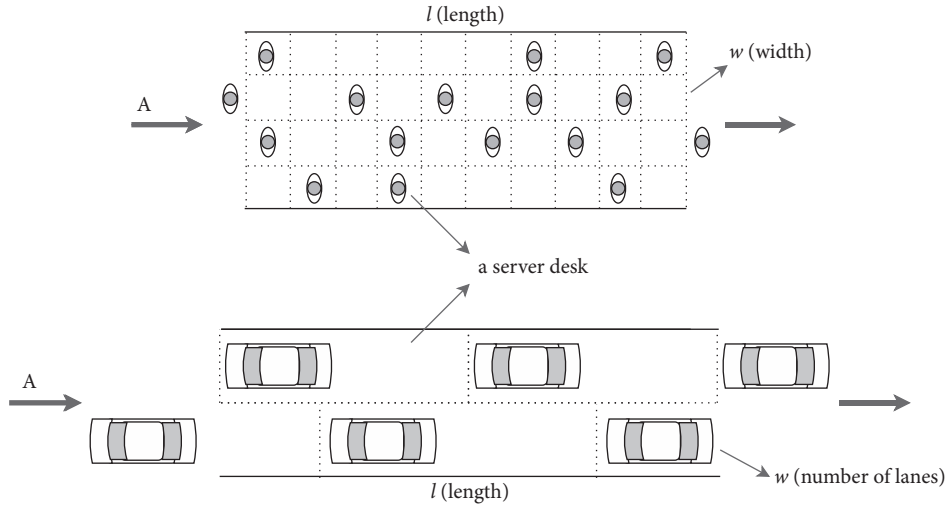
FIGURE 1: Circulation facility for pedestrians or vehicles.

TABLE 2: Main notation.

Variable	Description
Input parameters	
l	The length of the facility, in meters or miles
w	The width (in meters) of the facility in case of pedestrian traffic, or the number of lanes in the case of road traffic
N	The total number of RE states
λ_i	The arrival rate at RE state i
c_{ai}	SCV (square of the coefficient of variation) of the interarrival times at RE state i
v_{1i}	Velocity for pedestrians or vehicles when $n = 1$ at RE state i
c_{si}	SCV of service time at RE state i
$\mathbf{Q}$	The infinitesimal generator of the Markov chain for the RE
$\mathbf{x}$	The invariant probability vector of $\mathbf{Q}$
Queuing variables	
A_i	The random variable for the interarrival times distribution
B_i	The random variable for the service time distribution
C	The number of servers, or the capacity of the queuing system
n	The number of customers in the queuing system, system state
$\mathbf{Q}^*$	The infinitesimal generator of the Markov chain for the queuing system
π	The invariant probability vector of $\mathbf{Q}^*$
v_{ni}	The velocity when the system state is n and the RE state is i
μ_{ni}	The service rate when the system state is n and the RE state is i
t_{ni}	The mean service time when the system state is n and the RE state is i
Output parameters	
$p_n(i)$	The conditional probability that the system state is equal to n on the condition that the RE state is i
$EN(i)$	The conditional means of the queue length for RE state i
$VN(i)$	The conditional variances of the queue length for RE state i
$PC(i)$	The conditional blocking probability for RE state i
$\theta(i)$	The conditional output rate for RE state i

PH briefly for the sake of understanding. Readers can also refer to the related papers and books.

A random variable that follows the simplified PH has an m-order representation as $(\alpha, \mathbf{T})$, where α is a $1 \times m$ vector and $\mathbf{T}$ is an $m \times m$ matrix. α and $\mathbf{T}$ can be calculated according to the mean value u and SCV c of the random variable by the following method:

(1) When $c \leq 1$, $m = \lceil 1/c \rceil$, $\alpha = (1, 0, \ldots, 0)$, matrix $\mathbf{T}$ can be expressed as follows:

$$\mathbf{T} = \begin{bmatrix} -t_1 & t_1 & & & \\ & -t_2 & t_2 & & \\ & & \ddots & \ddots & \\ & & & -t_{m-1} & t_{m-1} \\ & & & & -t_m \end{bmatrix}, \quad (3)$$

where $t_i = (m/u)$, $1 \leq i \leq m-2$; $t_m = (2m[1 + \sqrt{m}(mc-1)/2]/[u(m+2-m^2c)])$; $t_{m-1} = (mt_m/(2t_m$

$u - m$)). For $c \leq (1/30)$, m is set to 30 to simplify the queuing analysis.

(2) When $c > 1$, $m = 2$, $\alpha = (\quad p, 1 - p)$, $p = ([1 + \sqrt{(c-1)/(c+1)}]/2)$, and the matrix $\mathbf{T}$ can be expressed as follows:

$$\mathbf{T} = \begin{bmatrix} -2p/u & \\ & -2(1-p)/u \end{bmatrix}. \tag{4}$$

For interarrival time A and service time B that follow the simplified PH distribution, we need their mean value (u_a and u_s) and SCV (c_a and c_s) to determine their PH representation. Given an RE state i ($1 \leq i \leq N$), we assume λ_i is the arrival rate of pedestrians or vehicles and μ_{ni} is the state-dependent service rate when the system state is n, $1 \leq n \leq C$. Since the arrival rate is the inverse of the mean interarrival time and the service rate is the inverse of the mean service time, we have $\lambda_i = (1/u_a)$ and $\mu_{ni} = (1/u_s)$.

Assume c_{a_i} and c_{s_i} represent the SCV of interarrival times and the SCV of service times, respectively. Then, the interarrival time and the service time can be expressed as PH $\langle \lambda_i, c_{a_i} \rangle$ and PH $\langle \mu_{ni}, c_{s_i} \rangle$. Their PH representations PH $(\alpha_i, \mathbf{T}_i)$ and PH $(\beta_i, \mathbf{S}_{ni})$ can thus be calculated according to (3) or (4). The calibration of λ_i, c_{a_i}, μ_{ni} and c_{s_i} will be presented in the next subsection.

Here, the SCV of service time is assumed to be independent of system state n. Nevertheless, Hu et al. [15] indicated that the SCV of service time is also state-dependent, as the fluctuation range of the velocity also varies with n. This relation can be considered by applying the state-dependent SCV of service time in Hu et al. [15].

Next, we present assumptions concerning the RE. There are two kinds of RE in the literature, asynchronous RE and synchronous RE. A synchronous RE assumes that the change in RE state is synchronized with the customers' departure or (and) arrival in the queuing system. More popular is the asynchronous RE, which transitions independently of the behavior in the queuing system, as considered in the study of, e.g., Krieger et al. [44] and Yang et al. [55]. We assume the RE is asynchronous in this paper, as the change in RE state, such as a change in weather conditions, is uncorrelated with the behavior of the traffic queue.

A change in the RE can change the interarrival time, service time, capacity of the queuing system, and even the service discipline. However, we consider only the first two cases in this research. The capacity and the service discipline of the queuing system always remain unchanged. In addition, the interarrival times distribution and service time distribution of a customer are assumed to depend on only the RE state at the time the arrival or service process begins [28]. In other words, once an arrival or service for a customer has been initiated, the parameters will remain the same until completion.

3.3. Parameter Calibrations.
We present the calibration method briefly in this subsection. A similar but more detailed calibration process for parameters not affected by an RE can be found in the study by Hu et al. [15].

For the arrival process, the arrival rate λ_i and SCV of interarrival times c_{a_i} at RE state i are easy to obtain based on field data. In some cases where the hourly traffic volume and the peak hour factor at RE state i are given, we can also calculate λ_i and c_{a_i} according to Equations (9)-(12) in Hu et al. [15]. Then, based on (3) or (4), we can acquire α_i and $\mathbf{T}_i$ for the PH interarrival times.

For the service process, the service rate can be obtained as follows:

$$\mu_{ni} = \frac{1}{t_{ni}} = \frac{v_{ni}}{l}, \tag{5}$$

where t_{ni} and v_{ni} are, respectively, the mean service time and the mean velocity of a customer.

By introducing the representative points that vary with the random environment to the mean velocity models proposed by Yuhaski et al. [14], we obtain the mean velocity models that consider the state dependence as well as the effect of the random environment. The linear model is as follows:

$$v_{ni} = v_{1i}\left(\frac{C + 1 - n}{C}\right). \tag{6}$$

The exponential model is as follows:

$$v_{ni} = v_{1i} \exp\left[-\left(\frac{(n-1)}{\varrho}\right)^{\gamma}\right], \tag{7}$$

where

$$\gamma = \frac{\ln[\ln(v_{ai}/v_{1i})/\ln(v_{bi}/v_{1i})]}{\ln((a-1)/(b-1))},$$

$$\varrho = \frac{a-1}{[\ln(v_{1i}/v_{ai})]^{(1/\gamma)}} \tag{8}$$

$$= \frac{b-1}{[\ln(v_{1i}/v_{bi})]^{(1/\gamma)}}.$$

The necessary representative points are $(1, v_{1i})$, (a, v_{ai}), and (b, v_{bi}). According to the research on pedestrian velocity by Tregenza [56]; $v_{1i} = 1.5$ m/s for $n = 1$, $v_{ai} = 0.64$ m/s for $n = a = 2lw$ and $v_{bi} = 0.25$ m/s for $n = b = 4lw$. For road traffic, Smith and Cruz [25] proposed the following coordinates: $v_{1i} = 62.5$ mph for $n = 1$, $v_{ai} = 48.0$ mph for $n = a = 20lw$ and $v_{bi} = 20.0$ mph for $n = b = 140lw$. *Note*: (1) The original models proposed by Yuhaski et al. [14] are based on the velocity-density curves presented in Tregenza [56], where the effect of an RE is not considered (the subscript i is absent in the original models). (2) The above data for representative points are for a unidirectional traffic flow; data for bidirectional and multidirectional passengers flow can be found

in the study by Mitchell and Smith [48] and Hu et al. [15]. Details on the relationship between speed, travel time, flow rate vary and density when the above state-dependent velocity models are applied can found in Hu et al. [13].

Equations (6) and (7) indicate that we can obtain the mean velocity models based on the values for v_{1i}, v_{ai}, and v_{bi} at each RE state i. These data can be obtained by field or simulation experiments. Thus, we can obtain the service rate μ_{ni}.

As for the SCV for service time, Hu et al. [15] proposed a similar exponential model based on the standard deviation of velocity. However, the process is complicated, and a considerable amount of data is needed for the calibration. For simplicity, c_{s_i} is assumed to be independent of system state n and to depend only on RE state i in this paper. Then, by applying (3) or (4), we can calculate β_i and $\mathbf{S}_{ni}$ for the PH service time.

4. Mathematical Model

After the preparation work, we propose the PH (i)/PH (i, n)/C/C state-dependent queuing model in an RE in this section. Then, the main conditional performance measures of concern are derived. Furthermore, the MAS algorithm is extended to solve the proposed model and obtain the performance measures.

4.1. Model Development. Let i_t, $t \geq 0$ be an irreducible continuous time Markov chain that represents the RE. It has a finite state space $\{1, 2, \ldots, N\}$, $N \geq 2$. The infinitesimal generator for i_t is $\mathbf{Q}$. $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ is the invariant probability vector of $\mathbf{Q}$ and satisfies $\mathbf{x}\mathbf{Q} = \mathbf{0}$, $\mathbf{x}\mathbf{e} = 1$, where $\mathbf{e}$ denotes a column vector of appropriate size consisting of all ones. As the RE is represented by an irreducible continuous time Markov chain, the duration of each RE state can follow a random distribution, such as an exponential, Erlang or PH distribution.

The interarrival time distribution PH $\langle \lambda_i, c_{a_i} \rangle$ or PH $(\alpha_i, \mathbf{T}_i)$ can be represented by a random variable describing the time until absorption of a Markov process η_t, $t \geq 0$. η_t has one absorbing state and m_i transient states under RE state i. The infinitesimal generator $\mathbf{T}_i^*$ for η_t can be expressed as follows:

$$\mathbf{T}_i^* = \begin{bmatrix} \mathbf{T}_i & \mathbf{T}_i^0 \\ \mathbf{0} & 0 \end{bmatrix}, \tag{9}$$

where $\mathbf{T}_i$ is obtained by equations (3) or (4) and $\mathbf{T}_i \mathbf{e} \leq 0$, $(\mathbf{T}_i)_{jj} < 0$ and $(\mathbf{T}_i)_{jk} \geq 0$ for $j \neq k$, $\mathbf{e} = (1, 1, \ldots, 1)'_{m_i \times 1}$, $\mathbf{T}_i^0$ is a nonnegative vector and satisfies $\mathbf{T}_i \mathbf{e} + \mathbf{T}_i^0 = \mathbf{0}$.

Similarly, the service time distribution PH $\langle \mu_{ni}, c_{s_i} \rangle$, or PH $(\beta_i, \mathbf{S}_{ni})$ can be described by a continuous time Markov process ψ_t, $t \geq 0$. ψ_t has one absorbing state and l_i transient states under RE state i. Its infinitesimal generator is as follows:

$$\mathbf{S}_{ni}^* = \begin{bmatrix} \mathbf{S}_{ni} & \mathbf{S}_{ni}^0 \\ \mathbf{0} & 0 \end{bmatrix}, \tag{10}$$

where $\mathbf{S}_{ni}$ (obtained by equations (3) or (4)) satisfies $\mathbf{S}_{ni} \mathbf{e} \leq 0$, $(\mathbf{S}_{ni})_{jj} < 0$ and $(\mathbf{S}_{ni})_{jk} \geq 0$ for $j \neq k$, $\mathbf{e} = (1, 1, \ldots, 1)'_{l_i \times 1}$, $\mathbf{S}_{ni}^0$ is a nonnegative vector and $\mathbf{S}_{ni} \mathbf{e} + \mathbf{S}_{ni}^0 = \mathbf{0}$.

Property 2 in the study by Hu et al. [15] indicates that if B $\sim$ PH $\langle \mu_{ni}, c_{s_i} \rangle$ or PH $(\beta_i, \mathbf{S}_{ni})$, then B/$n$ $\sim$ PH $\langle n\mu_{ni}, c_{s_i} \rangle$ or PH $(\beta_i, n\mathbf{S}_{ni})$. Therefore, the state-dependent PH $\langle \lambda_i, c_{a_i} \rangle$/PH $\langle \mu_{ni}, c_{s_i} \rangle$/C/C queuing system with C parallel-serial servers can be approximately transformed into a PH $\langle \lambda_i, c_{a_i} \rangle$/PH $\langle n\mu_{ni}, c_{s_i} \rangle$/1/C state-dependent queuing system with a single server, see Figure 2.

Notably, although the model is proposed for a circulation facility without obvious service desks, it is also applicable for queues with obvious parallel service desks, such as ticket vending machines and ticket windows. This is because parallel queue models can also be similarly converted into state-dependent queue models as discussed in Hu [57].

The process of the PH $\langle \lambda_i, c_{a_i} \rangle$/PH $\langle n\mu_{ni}, c_{s_i} \rangle$/1/C state-dependent queuing system can be described by a four-dimensional continuous time Markov chain $\xi_t = \{n_t, i_t, \eta_t, \psi_t\}$, $t \geq 0$, where

    (i) n_t is the system state, or number of customers, in the queuing system, $n_t = 0, 1, \ldots, C$;

    (ii) i_t is the RE state of the RE, $i_t = 1, 2, \ldots, N$;

    (iii) η_t is the phase of the PH arrival process, $\eta_t = 1, 2, \ldots, m_i$;

    (iv) ψ_t is the phase of the PH service process, $\psi_t = 1, 2, \ldots, l_i$.

The state space for the Markov chain ξ_t is as follows:

$$U = \{(n, i, \eta) | n = 0, \ 1 \leq i \leq N, \ 1 \leq \eta \leq m_i\}$$
$$\cup \{(n, i, \eta, \psi) | 1 \leq n \leq C, \ 1 \leq i \leq N, \ 1 \leq \eta \leq m_i, \ 1 \leq \psi \leq l_i\}, \tag{11}$$

where n stands for the system state at time t, i is the RE state, j is the arrival phase and k is the service phase. This state space can also be divided into different levels: 0, 1, 2, ..., C. Enumerating the state of ξ_t in lexicographic order, we obtain the following:

$$\underline{0} = \{(0, 1), (0, 2), \ldots, (0, i), \ldots, (0, N)\}, \ 1 \leq i \leq N,$$

$$\underline{n} = \{(n, 1), (n, 2), \ldots, (n, i), \ldots, (n, N)\}, \ 1 \leq i \leq N, \ 1 \leq n \leq C, \tag{12}$$

where

$$(\underline{0, i}) = \{(0, i, 1), (0, i, 2), \ldots, (0, i, m_i)\}, \ 1 \leq i \leq N,$$
$$(\underline{n, i}) = \{(n, i, 1, 1), \ldots, (n, i, 1, l_i), \ldots,$$
$$\cdot (n, i, m_i, 1), \ldots, (n, i, m_i, l_i)\}, \ 1 \leq n \leq C. \tag{13}$$

Level $\underline{0}$ consists of $h_0 = \sum_{i=1}^N m_i$ states, and level $\underline{n}$ $(1 \leq n \leq C)$ consists of $h_n = \sum_{i=1}^N m_i \cdot l_i$ states. Note that there is no service phase for level $\underline{0}$ as the service process has not yet started. Level $\underline{n}$ represents the states when there are n customers in the system. A step from level $\underline{n}$ to level $\underline{n+1}$ represents an arrival, and a step from level $\underline{n}$ to level $\underline{n-1}$ $(n \geq 1)$ represents a departure.

FIGURE 2: Transformation of queuing system.

The queuing system can be described by a finite level-dependent QBD process, which leads to the following generator matrix $\mathbf{Q}^*$ of the Markov chain:

$$\mathbf{Q}^* = \begin{bmatrix} \mathbf{D}_0 & \mathbf{F}_0 & & & & \\ \mathbf{E}_1 & \mathbf{D}_1 & \mathbf{F}_1 & & & \\ & \mathbf{E}_2 & \mathbf{D}_2 & \mathbf{F}_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \mathbf{E}_{C-1} & \mathbf{D}_{C-1} & \mathbf{F}_{C-1} \\ & & & & \mathbf{E}_C & \mathbf{D}_C \end{bmatrix}, \quad (14)$$

where $\mathbf{D}_0 = \mathbf{Q} \widetilde{\otimes} \mathbf{I}_{m_j} + \text{diag}\{\mathbf{T}_i, i = 1, 2, \ldots, N\}$, $\mathbf{F}_0 = \text{diag}\{\mathbf{T}_i^0 \alpha_i \otimes \beta_i, i = 1, 2, \ldots, N\}$, $\mathbf{E}_1 = \text{diag}\{\mathbf{I}_{m_i} \otimes \mathbf{S}_{1i}^0, i = 1, 2, \ldots, N\}$, $\mathbf{D}_n = \mathbf{Q} \widetilde{\otimes} \mathbf{I}_{m_j l_{nj} \times} m_k l_{nk} + \text{diag}\{\mathbf{T}_i \otimes \mathbf{I}_{l_i} + \mathbf{I}_{m_i} \otimes n\mathbf{S}_{ni}, i = 1, 2, \ldots, N\}$, $n = 1, 2, \ldots C - 1$, $\mathbf{F}_n = \text{diag}\{\mathbf{T}_i^0 \alpha_i \otimes \mathbf{I}_{l_i \times l_i}, i = 1, 2, \ldots, N\}$, $n = 1, 2, \ldots C - 1$, $\mathbf{E}_n = \text{diag}\{\mathbf{I}_{m_i} \otimes n\mathbf{S}_{ni}^0 \beta_i, i = 1, 2, \ldots, N\}$, $n = 2, 3, \ldots C$, $\mathbf{D}_C = \mathbf{Q} \widetilde{\otimes} \mathbf{I}_{m_j l_{Cj} \times m_k l_{Ck}} + \text{diag}\{(\mathbf{T}_i + \mathbf{T}_i^0 \alpha_i) \otimes \mathbf{I}_{l_{Ci}} + \mathbf{I}_{m_i} \otimes C\mathbf{S}_{Ci}, i = 1, 2, \ldots, N\}$, $\mathbf{I}_{m_j} = \text{diag}(1, 1, \ldots, 1)_{m_j \times m_j}$, $\mathbf{I}_{m_j l_j \times m_k l_k} = \text{diag}(1, 1, \ldots, 1)_{m_j l_j \times m_k l_k}$.

Note: $\otimes$ is the Kronecker product and $\widetilde{\otimes}$ is defined as a special Kronecker product. For example, $\mathbf{Q} \widetilde{\otimes} \mathbf{I}_{m_j l_{nj} \times m_k l_{nk}}$ means that each element Q_{jk} in $\mathbf{Q}$ is multiplied by an identity matrix $\mathbf{I}_{m_j l_{nj} \times m_k l_{nk}}$, whose dimension depends on j and k, the coordinate of Q_{jk}.

$\mathbf{Q}^*$ is a block matrix. The nonzero (tri-)diagonal blocks of $\mathbf{Q}^*$, such as $\mathbf{D}_0$, are also block matrices with $N \times N$ small blocks, and these small blocks represent the transition rate matrix among different RE states of the RE.

For the finite states, the stationary distribution of the Markov chain $\xi_t, t \geq 0$ always exists [55]. The stationary probability vector π of the Markov chain with generator $\mathbf{Q}^*$ can be obtained by solving the following global balance equation:

$$\pi \mathbf{Q}^* = \mathbf{0} \text{ and } \pi \mathbf{e} = 1, \quad (15)$$

where $\pi = (\pi_0, \pi_1, \ldots, \pi_n, \ldots, \pi_C)$. π_0 represents the steady-state probability vector for level $\underline{0}$. $\pi_0 = (\pi_{01}, \pi_{02}, \ldots, \pi_{0i}, \ldots, \pi_{0N})$, where π_{0i} represents the steady-state probability vector for system state 0 and environment state i, and it can be further expanded into $\pi_{0i} = (\pi_{0i_1}, \pi_{0i_2}, \ldots, \pi_{0i_{m_i}})$. $\pi_n = (\pi_{n1}, \pi_{n2}, \ldots, \pi_{ni}, \ldots, \pi_{nN})$ represents the steady-state probability vector for level $\underline{n}$. $\pi_{ni} = (\pi_{ni_1}, \pi_{ni_2}, \ldots, \pi_{ni_{m_i \times l_i}})$ is the steady-state probability

vector for system state n and environment state i. The length of π is $h_0 + \sum_{n=1}^{C} h_n$.

In addition, the stationary probability satisfies $\pi_{n+1} = \pi_n \mathbf{R}_n (0 \leq n \leq C - 1)$, where $\mathbf{R}_n$ is a level-dependent rate matrix with $h_n \times h_n$ nonnegative elements. Meanwhile, $\pi_0 (\mathbf{D}_0 + \mathbf{R}_0 \mathbf{E}_1) = 0, \mathbf{R}_n = -(\mathbf{F}_n / (\mathbf{D}_{n+1} + \mathbf{R}_{n+1} \mathbf{E}_{n+2})), \mathbf{R}_C = \mathbf{0}_{h_C \times h_C}$ (a h_C by h_C matrix with all zeros).

4.2. Performance Measures. By solving the global balance equation (15), we can obtain the stationary probability vector π. The performance measures of interest can then be calculated. Generally, stationary average performance measures, such as average queue length EN, average blocking probability PC and average output rate θ (the average here refers to the mean value of the results for different RE states), convey all the information of interest about the queuing system. However, for queuing models in RE, these average results provide little information for analyzing circulation systems with varying parameters. The essential characteristic that system performance changes with the RE state is concealed by the average value. In practice, we are more interested in the conditional results that reflect the changing performance measures at different RE states. Therefore, this section focuses on deducing the conditional performance measures and designing a solution algorithm.

To explore the conditional performance measures, the conditional probability is necessary. Let $p_n(i)$ represent the conditional probability that the system state is equal to n on the condition that the RE state is i,

$$p_n(i) = P\{n_t = n | i_t = i\}$$
$$= \frac{\pi_{ni}\mathbf{e}}{x_i}. \quad (16)$$

Based on this conditional probability, we can deduce the main conditional performance measures, such as the conditional means $EN(i)$ and variances $VN(i)$ of the queue lengths, the conditional blocking probability $PC(i)$, and the conditional output rate $\theta(i)$ at RE state i. Note that these conditional performance measures represent the average performance for a given RE state.

$$EN(i) = \sum_{n=1}^{C} n p_n(i), \quad (17)$$

$$VN(i) = \sum_{n=1}^{C} n^2 p_n(i) - (EN(i))^2, \quad (18)$$

$$PC(i) = p_C(i), \quad (19)$$

$$\theta(i) = \sum_{n=1}^{C} n p_n(i) \mu_{ni}. \quad (20)$$

To calculate these performance measures, one method is to first solve the global balance (15). As a limited linear equation set, (15) can be directly solved by LU-decomposition, Gaussian elimination, Gauss–Seidel method, matrix continued fractions, or matrix analytic method. Based on the

stationary distribution π, we can obtain the above performance measures. However, huge computer storage is required due to the explicit calculation of π. To avoid excessive storage demand, we extend the MAS algorithm applied by Baumann and Sandmann [32] to compute the stationary results.

The algorithm works in two steps: First, according to $\pi_{n+1} = \pi_n R_n (0 \le n \le C - 1)$, the expectation for the stationary performance measures can be expressed as follows:

$$
\begin{aligned}
E_\pi(\mathbf{f}) &= \sum_{n=0}^{C} \pi_\mathbf{n} \mathbf{f_n} \\
&= \pi_0 \sum_{n=0}^{C} \left(\prod_{k=0}^{n-1} \mathbf{R}_k \right) \mathbf{f_n},
\end{aligned}
\tag{21}
$$

where $\mathbf{f_n}$ is a vector function of the system state n and RE state i. Second, the expectation factor $\sum_{n=0}^{C} (\prod_{k=0}^{n-1} \mathbf{R}_k) \mathbf{f_n}$ without the storage of $\mathbf{R}_n (0 \le n \le C - 1)$ is computed using the Horner rule (refer to Baumann and Sandmann [32] for details).

To satisfy the required form of the expectation factor, $p_n(i)$ in equation (16) is transformed into the following form:

$$
p_n(i) = \frac{\pi_n \varepsilon_{ni}}{x_i},
\tag{22}
$$

where ε_{ni} is denoted as a special column vector that satisfies $\pi_n \varepsilon_{ni} = \pi_{ni} \mathbf{e}$. The length of ε_{ni} is equal to that of $\pi_\mathbf{n}$. The elements in ε_{ni}, which corresponds to π_{ni}, are ones, and the total number of ones is $m_i \times l_i$, while the others are all zeros. For example, $\varepsilon_{n1} = (0, 0, \ldots, 0, 1, 1, \ldots, 1, 0, 0, \ldots, 0)$. Now, we can express the vector function $\mathbf{f_n}$ for different performance measures as follows:

(1) For $EN(i)$ in equation (12), $\mathbf{f_n} = (n \varepsilon_{ni}/x_i)$, $0 \le n \le C$,

(2) For $VN(i) + (EN(i))^2$ in equation (13), $\mathbf{f_n} = (n^2 \varepsilon_{ni}/x_i)$, $0 \le n \le C$,

(3) For $PC(i)$ in equation (14), $\mathbf{f_n} = \begin{cases} 0 \varepsilon_{0i}, & 0 \le n \le C-1 \\ (\varepsilon_{Ci}/x_i), & n = C \end{cases}$,

(4) For $\theta(i)$ in equation (15), $\mathbf{f_n} = \begin{cases} (n \varepsilon_{ni} \mu_{ni}/x_i), & 1 \le n \le C \\ \varepsilon_{0i}, & n = 0 \end{cases}$.

The extended MAS algorithm is presented in Figure 3. As the proposed model is described as a four-dimensional continuous time Markov chain, the transform of the state of the RE is incorporated in the extended MAS algorithm. The time complexity of the extended MAS algorithm is a linear function of $C \cdot N$ (the product of queue capacity C and the number of RE states N), i.e., O $(C \cdot N)$. As the length of the level-dependent matrix $\mathbf{R}_n$ is $h_n = \sum_{i=1}^{N} m_i \cdot l_i$, where m_i and l_i are not larger than 30 (as stated in Section 3.2), the space complexity of the algorithm is a linear function of the number of RE states N, i.e., O (N).

5. Numerical Experiments

In this section, we first verify the accuracy of the proposed model and then illustrate its applications in modeling road and pedestrian traffic circulation systems. Meanwhile, the sensitivity of the parameters that might influence the performance measures of the circulation systems is analyzed.

In all the experiments, for convenience, we assume the RE has an exponential duration at each RE state unless otherwise specified. The exponential duration of each RE state i has a mean value equal to $-1/Q_{ii}$ (the opposite of the inverse of Q_{ii}), where Q_{ij} represents the diagonal entry of $\mathbf{Q}$. For example, the morning peak period lasting 90 minutes can be represented by an RE state whose duration follows an exponential distribution with a mean value of 90 minutes ($Q_{ii} = -1/90$).

For convenience, we use $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_N)$ to represent the arrival rate vector, the entries of which are the arrival rates at different RE states. $\mathbf{v}_1 = (\nu_1, \nu_2, \ldots, \nu_i, \ldots, \nu_N)$ is the velocity vector of a single customer under the RE. Similarly, $\mathbf{c}_a$ and $\mathbf{c}_s$ denote the vector for the SCV of the interarrival times and service time, respectively. In the control group of the following experiments, we use the average arrival rate or average velocity. Here, the average refers to the weighted average, where the invariant probability vector of the RE $\mathbf{x}$ is the weighting.

5.1. Model Verification. In this subsection, we verify the accuracy of the proposed queuing model by comparing it with three other models, the M/M/1/∞ queuing model in an RE in Neuts [39]; the $M(t)/G(x)/C/C$ fluid queuing model developed by Hu et al. [13] and the discrete-event simulation model. In addition, the comparing between the proposed model and the PH/PH(n)/C/C queuing model in Hu et al. [15] is presented in Section Sec:exper2.

First, we compare the proposed PH(i)/PH(i, n)/C/C state-dependent queuing model in a RE with the M/M/1/∞ queuing model in the study by Neuts [39]. To carry out the comparison on the same benchmark, the entries of $\mathbf{c}_a$ and $\mathbf{c}_s$ of the proposed PH(i)/PH(i, n)/C/C model are set to 1. In addition, the first C is set equal to 1, the second C is set to be a sufficiently large number and the effect of state-dependence on service time is eliminated. The arrival rate and service rate are the same as those used in the study by Neuts [39]. The conditional means and variances of the queue length are obtained by the two methods, as shown in Table 3.

Table 3 shows that the results of the two methods are almost the same, which indicates that the proposed model can accurately reproduce the performance of queuing systems with changing arrival and service rates. In addition, the results imply that the M/M/1/∞ queuing model in an RE can be approximated by the PH(i)/PH(i, n)/C/C state-dependent queuing model in an RE. This supposition is coincident with the work of Hu et al. [15]; which proved that the M/G/1/∞ queuing model can be approximated by the PH/PH(n)/C/C state-dependent queuing model.

In the following experiment, we compare the proposed model with the transient $M(t)/G(x)/C/C$ fluid queuing model developed by Hu et al. [13]. The two models are used to analyze a road circulation system with a varying traffic demand. A road section of 850 meters is studied. The arrival rate of the M(t)/G(x)/C/C fluid queuing model changes with time determinately, while the RE state of the proposed model changes randomly. Therefore, we also explore the impact of the random length of each RE state by changing the value of the SCV of the duration time of the RE states

STEP 1   INITIALIZATION
  define
        $\mathbf{R} = \mathbf{0}_{h_C \times h_C}$ for the level-dependent matrix $\mathbf{R}_n$
        $\mathbf{y} = \mathbf{e}_{h_C \times 1}$ for the normalization constant a of the stationary probability
        for $i = 1, 2, \cdots, N$
            update
                $\mathbf{z}_1(i) = C\varepsilon_{Ci}/x_i$ for $EN(i)$ in Equation (12)
                $\mathbf{z}_2(i) = C^2\varepsilon_{Ci}/x_i$ for $VN(i) + (EN(i))^2$ in Equation (13)
                $\mathbf{z}_3(i) = \varepsilon_{Ci}/x_i$ for $PC(i)$ in Equation (14)
                $\mathbf{z}_4(i) = C\varepsilon_{Ci}\mu_{iC}/x_i$ for $\theta(i)$ in Equation (15)
            end
  end

STEP 2   ITERATIVE CALCULATION OF THE EXPECTATION FACTOR
  for $n = C - 1, C - 2, \cdots, 0$
    update
        $\mathbf{R} := -\mathbf{Q}^*_{n,n+1}/(\mathbf{Q}^*_{n+1,n+1} + \mathbf{RQ}^*_{n+2,n+1})$ for the iterative calculation of $\mathbf{R}_n$
        $\mathbf{y} := \mathbf{e}_{h_n \times 1} + \mathbf{Ry}$ for the normalization constant a
        for $i = 1, 2, \cdots, N$
            update
                $\mathbf{z}_1(i) := n\varepsilon_{ni}/x_i + \mathbf{Rz}_1(i)$ for $EN(i)$
                $\mathbf{z}_2(i) := n^2\varepsilon_{ni}/x_i + \mathbf{Rz}_2(i)$ for $VN(i) + (EN(i))^2$
                $\mathbf{z}_3(i) := \mathbf{Rz}_3(i)$ for $PC(i)$
                $\mathbf{z}_4(i) := n\varepsilon_{ni}\mu_{ni}/x_i + \mathbf{Rz}_4(i)$ for $\theta(i)$(Note $\mathbf{z}_4(i) := \mathbf{Rz}_4(i)$ when $n = 0$)
            end
    end

STEP 3   DETERMINATION OF A SPECIAL SOLUTION WITHOUT NORMALIZATION
    determine a nontrivial solution $\boldsymbol{r_0}$ of $\boldsymbol{r_0}(\mathbf{Q}^*_{0,0} + \mathbf{RQ}^*_{1,0}) = \mathbf{0}$, where $\boldsymbol{r_0} = a\boldsymbol{\pi}_0$

STEP 4   CALCULATION OF THE EXPECTATIONS WITHOUT NORMALIZATION
  update
        $\mathbf{y} := \boldsymbol{r_0}\mathbf{y}$, where $\mathbf{y} = a$
        for $i = 1, 2, \cdots, N$
            update
                $\mathbf{z}_1(i) := \boldsymbol{r_0}\mathbf{z}_1(i)$, where $\mathbf{z}_1(i) = aEN(i)$
                $\mathbf{z}_2(i) := \boldsymbol{r_0}\mathbf{z}_2(i)$, where $\mathbf{z}_2(i) = a(VN(i) + (EN(i))^2)$
                $\mathbf{z}_3(i) := \boldsymbol{r_0}\mathbf{z}_3(i)$, where $\mathbf{z}_3(i) = aPC(i)$
                $\mathbf{z}_4(i) := \boldsymbol{r_0}\mathbf{z}_4(i)$, where $\mathbf{z}_4(i) = a\theta(i)$
        end

STEP 5 CALCULATION OF THE PERFORMANCE MEASURES
  calculate
        $EN(i) = \mathbf{z}_1(i)/a$
        $VN(i) = \mathbf{z}_2(i)/a - (EN(i))^2$
        $PC(i) = \mathbf{z}_3(i)/a$
        $\theta(i) = \mathbf{z}_4(i)/a$
  end

FIGURE 3: The extended matrix analytic scheme algorithm.

(SCVRE) in this experiment. Four values for the SCVRE, 0.5, 1, 2 and 4, are tested. When the SCVRE equals 1, the duration time of the RE state follows an exponential distribution. When the SCVRE is 0, the duration time of each RE state is a constant value, similar to the situation in the $M(t)/G(x)/C/C$ fluid queuing model. The entries of $\mathbf{c}_a$ and $\mathbf{c}_s$ of the proposed PH(i)/PH$(i, n)/C/C$ model are set to 1. Two situations with low and high traffic demand are analyzed. The conditional queue length obtained by the proposed model is compared with the dynamic queue length calculated by the transient $M(t)/G(x)/C/C$ fluid queuing model, as shown in Figure 4.

For the low traffic demand in Figure 4(a), changing the SCVRE has little effect on the queue length and the result of the two models are almost the same. In Figure 4(b), the conditional queue length increases apparently with the SCVRE during the morning peak and the result of the proposed model are all larger than that of the $M(t)/G$

TABLE 3: Comparison with M/M/1/∞.

RE state i		1	2	3	4	5	6	7	8
$EN(i)$	Our model	50.65	50.98	41.55	36.05	31.29	27.17	23.60	20.51
	M/M/1/∞	50.65	50.98	41.55	36.05	31.29	27.17	23.60	20.50
$VN(i)$	Our model	2575.55	2581.12	2482.10	2346.10	2178.66	1996.99	1812.20	1631.56
	M/M/1/∞	2575.55	2581.12	2482.10	2346.10	2178.66	1996.99	1812.20	1631.56

(a)

(b)
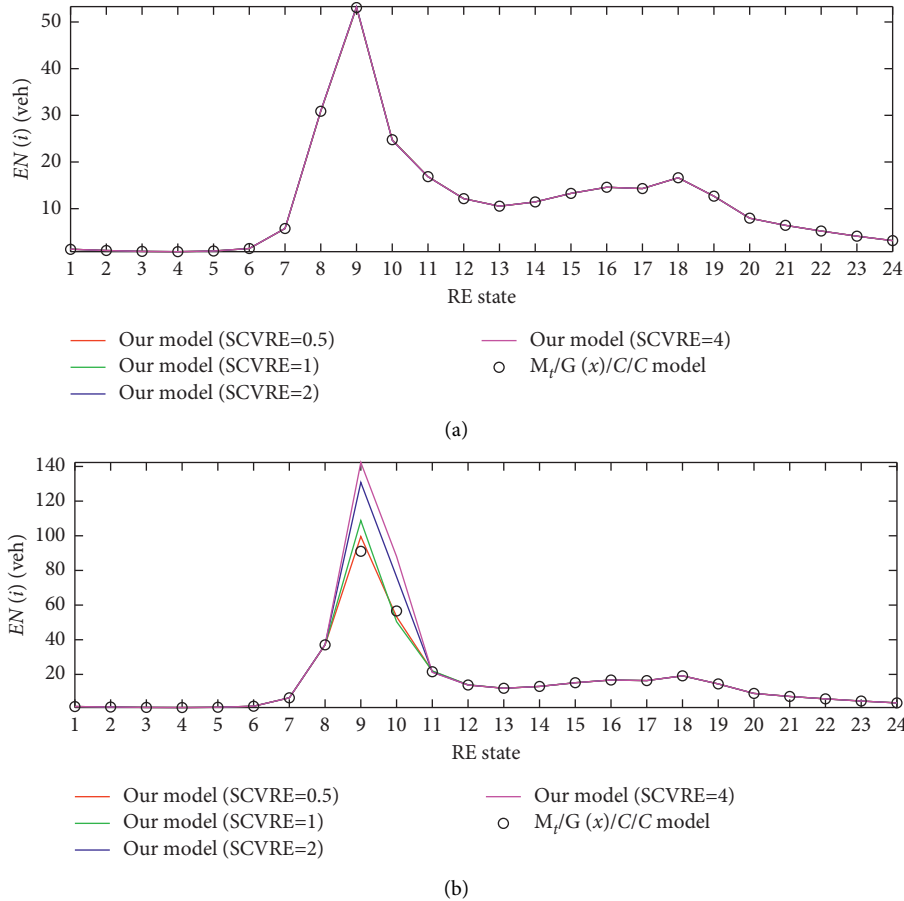
FIGURE 4: Comparison with the $M(t)/G(x)/C/C$ fluid queuing model.

$(x)/C/C$ fluid queuing model. This indicates that increasing the randomness in the duration time of the RE state has an obvious influence on the performance measures of a circulation system under high traffic intensities.

This experiment indicates that the proposed stationary model coincides with the transient $M(t)/G(x)/C/C$ fluid queuing model under the following conditions: (1) the SCV of the interarrival times of the proposed model is equal to 1; (2) the SCVRE of the proposed model is equal to 0; and (3) the conditional queue length of the proposed model is close to 0 at the first and the last RE state. This implies that the proposed stationary PHi/PH$(i, n)/C/C$ state-dependent queuing model in a RE may be useful in conducting transient analysis after certain transformations, which will be explored in our future work. Although the two methods obtain similar results under the same condition, the proposed model is superior to the $M(t)/G(x)/C/C$ fluid queuing model in two aspects. First, the result of the proposed model is an accurate solution while that

of the $M(t)/G(x)/C/C$ fluid queuing model is an approximation. Second, the $M(t)/G(x)/C/C$ fluid queuing model can only deal with the situation where the SCV of the interarrival times equals 1 and the varying arrival rate changes with time deterministically. However, the proposed model can capture the performance in different situations where the SCV of the interarrival times are much larger than 1 and the arrival rate changes randomly. Both the SCV of the interarrival times and the SCVRE have a remarkable influence on the performance measures under high traffic demand (The influence of the SCV of the interarrival times will be examined in the following experiment). It also should be mentioned that the calculation time of the proposed model is 8% less than that of the $M(t)/G(x)/C/C$ fluid queuing model.

Next, we will validate the accuracy of the proposed model by comparing it with the discrete-event simulation model. A pedestrian corridor with a length of 5 meters and a width of 3 meters is analyzed. Assume the arrival and service

process of the corridor is affected by an RE with four states ($N = 4$) that repeat cyclically. Therefore, the invariant probability vector of the RE is $x = 1/4(1, 1, 1, 1)$, which means the probability that the RE is at state i, $1 \leq i \leq 4$, is 1/4. Each RE state has an exponential duration with an average value of 40 s ($-(1/Q_{ii}) = 40$, $1 \leq i \leq 4$). Therefore, the infinitesimal generator can be expressed as follows:

$$\mathbf{Q} = \frac{1}{40} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{bmatrix}. \qquad (23)$$

The parameters for arrival and service under the RE are as follows: $\lambda = (5, 3, 3, 2)$, $c_a = (1.5, 1.5, 1.5, 1.5)$, and $v_1 = (1.5, 1.5, 1.4, 1.4)$, $c_s = (1.5, 1.5, 1.5, 1.5)$. The exponential model for pedestrian velocity in (7) is applied.

The discrete-event simulation model for such a corridor with changing arrival and service rates is set up in Simulink in MATLAB. A total of 5000 pedestrians are simulated, and the first 1000 pedestrians are excluded from the data collection. The simulation is repeated 10 times to calculate the average results. The conditional performance measures, $EN(i)$, $VN(i)$, $PC(i)$, and $\theta(i)$, are obtained from the simulation model and the proposed model. The results, as well as the average relative error, are listed in Table 4 (the simulation result is taken as the benchmark).

The data in Table 4 indicate that the relative error is quite small, with an average value of 3.71%. This result further verifies the accuracy of the proposed model. Although the analytical model and the simulation model produce similar results, there is a big difference between the time efficiency of the two methods. On an Intel(R) Core(TM) i7-2600 CPU @ 3.40 GHz, the running time for the analytical model is 0.13 s while that for simulation is 313.57 s.

5.2. Vehicle Queues in an RE. In this subsection, we apply the proposed model to analyze vehicle queues with different arrival rates and SCVs of interarrival times. Consider a road segment of 1 mile consisting of one lane ($l = 1$, $w = 1$) and has a unidirectional traffic flow. The constant ρ is 200 veh/mile-lane. Therefore, the capacity of the road segment is $C = 200$. The 24 hours of a day are treated as a circular RE with 8 states. The infinitesimal generator of the RE is as follows:

$$\mathbf{Q} = \frac{1}{3} \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \qquad (24)$$

and $\mathbf{x} = 1/8(1, 1, 1, 1, 1, 1, 1, 1)$. Thus, the duration of each RE state follows an exponential distribution with an expected duration time of 3 hours. The first RE state represents the time of 4:00–7:00 am, the second, 7:00–10:00 am, and so on.

Three arrival rate vectors are applied, $\lambda_1 = (1500, 1500, 1500, 1500, 1500, 1500, 1500, 1500)$, $\lambda_2 = (1200, 2200, 2100, 1700, 1400, 1300, 1100, 1000)$, and $\lambda_3 = (900, 2400, 2300, 2200, 1500, 1000, 900, 800)$. The three arrival rate vectors have the same total traffic volume in a day while the nonuniformity increases sequentially. According to our survey at Commonwealth Avenue in Singapore and the research of Zhang [58], the SCV of the interarrival times for vehicles varies from 0.5 to 15. Therefore, we set three vectors for the SCV of the interarrival times, $c_{a1} = (0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5)$, $c_{a2} = (5, 5, 5, 5, 5, 5, 5, 5)$, and $c_{a3} = (10, 10, 10, 10, 10, 10, 10, 10)$. We consecutively number the experiments with parameter pairs (λ_1, c_{a1}), (λ_1, c_{a2}), (λ_1, c_{a3}), (λ_2, c_{a1}), ..., (λ_3, c_{a3}), that is, Group 1, 2, ..., 9. Note that the first three groups, which apply the average arrival rate, are essentially equivalent to the PH/PH $(n)/C/C$ queuing model in Hu et al. [15].

All 9 groups of experiments have the same service rate vector and SCV of service time. In each experiment, the service rate and the SCV of service time are assumed to be independent of the RE state. The SCV of service time is 1. The service rate μ_{ni} is obtained based on the exponential velocity model in (7), where the velocities for the three representative points are $v_{1i} = 55$ mph, $v_{ai} = 48$ mph, and $v_{bi} = 20$ mph (similar to the values in Jain and Smith [24]). The conditional queue length $EN(i)$ and the conditional coefficient of variation of the queue length CV (i) are presented in Figure 5. The average time cost for one experiment is 0.8 s.

Figure 5 indicates that the nonuniformity in traffic demand causes remarkable fluctuation in the conditional performance measures. The greater the nonuniformity is, the greater the fluctuation in conditional performance will be. By contrast, the conditional mean and SCV of queue length for the first three groups are the same at all RE states. However, Figures 5(a) and 5(b) show different tendencies. In Figure 5(a), the conditional queue length of Groups 4 to 9 at RE states 2, 3, and 4 (the peak period) is much larger than that of Groups 1 to 3. The maximum conditional queue length reaches 180 during the peak period for Group 7 while that for Groups 1 to 3 is always 38. This indicates that the conditional queue length during the peak period can be tremendously underestimated by applying an average arrival rate since the fluctuating traffic intensity levels off. According to the three vectors of arrival rate and the data in Figure 5, we can see that a one percent increase in conditional arrival rate gives rise to a more than one percent increase in the conditional queue length during the peak period. The conditional CV of queue length in Figure 5(b) generally shows an opposite tendency. The conditional CV of queue length tends to be smaller at RE states with larger conditional queue length. When congestion occurs (the number of vehicles waiting in the queue is too large), the freedom of motion is almost completely limited and less fluctuation exists in queue length. When the congestion

TABLE 4: Comparison with the simulation model in an RE.

RE state		1	2	3	4	Error (%)
$EN(i)$ (ped)	Simulation	54.16	47.33	42.14	22.45	—
	Our model	54.38	45.75	39.96	21.57	3.21
$VN(i)$ (ped^2)	Simulation	490.87	581.16	597.78	422.67	—
	Our model	517.16	630.48	636.84	442.00	6.24
$PC(i)$ (%)	Simulation	23.42	9.51	7.56	0.96	—
	Our model	24.48	9.67	7.30	1.05	4.62
$\theta(i)$ (Ped/s)	Simulation	2.87	2.94	2.86	2.44	—
	Our model	2.87	2.90	2.91	2.44	0.78
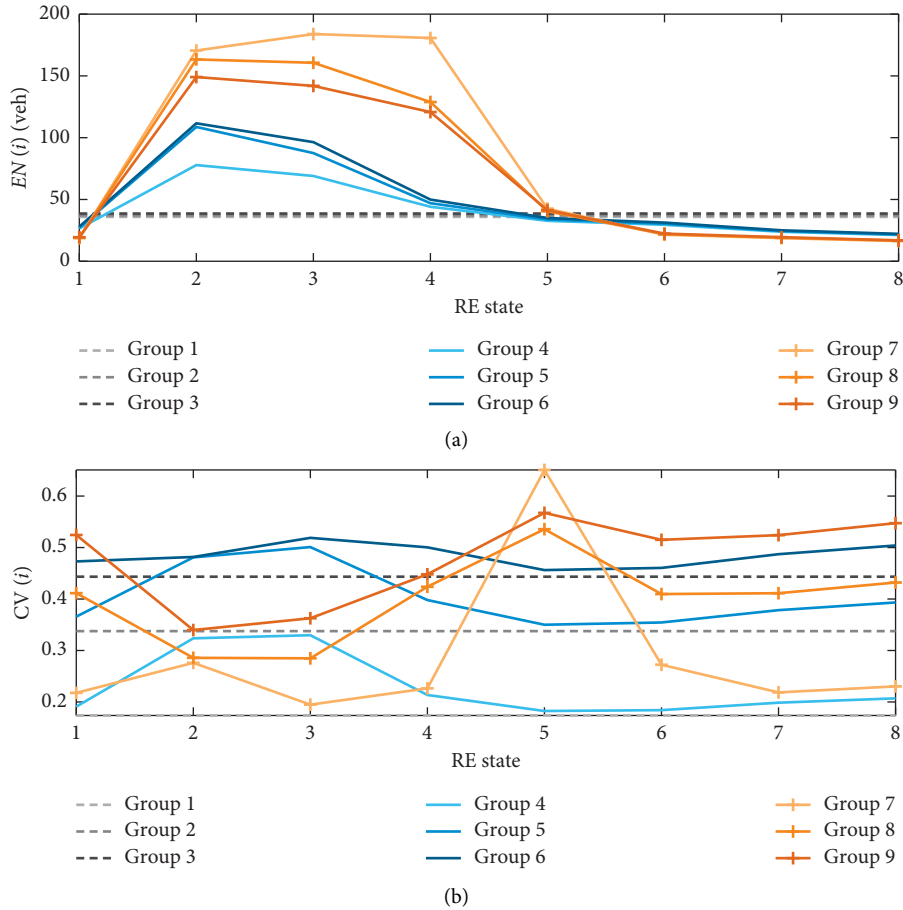
FIGURE 5: Conditional mean and CV of queue length for vehicle queues. (a) Conditional queue length. (b) Conditional CV of queue length.

begins to dissipate (the accumulated queue length begins to decline) after the peak period, the conditional CV of queue length increases.

Figure 5 also shows the effect of the SCV of interarrival times. In Figure 5(a), the effect of the SCV of interarrival times varies. When the traffic intensity is low, the effect of the SCV of interarrival times is quite small. As the traffic intensity increases, the influence of the SCV of interarrival times tends to increase. Figure 5(a) shows that an increasing SCV of interarrival times leads to a decrease in $EN(i)$ at RE states 2, 3, and 4 in Groups 7 to 9, which appears to be unreasonable. The reason is as follows. A larger SCV of interarrival times means vehicles arrive more randomly.

When the conditional queue length is close to the capacity of the road, most vehicles are rejected due to the capacity limitation. Consequently, the number of vehicles that get into the system decreases, resulting in a slight decrease in the queue length. Notably, this unexpected phenomenon occurs only when the expected queuing length is close to the capacity. Otherwise, the system has sufficient available capacity to accommodate the additional vehicles. An additional experiment shows that the PH/PH $(n)/C/C$ state-dependent queuing model in the study by Hu et al. [15] presents a similar phenomenon. These results suggest a precondition shared by both models that the studied queuing system is a loss queue with limited capacity. The

newly arriving vehicles will leave the system when they find that the system is full. In Figure 5(b), generally, a larger SCV of interarrival times results in a larger conditional CV of queue length.

A close look at Figure 5(a) provides additional interesting information. Although RE states 3 and 4 have smaller arrival rates than that of RE state 2, their conditional mean and CV of queue length are sometimes larger than that of RE state 2. For example, $EN(3)$ is larger than $EN(2)$ in Group 7. This result implies that the performance of an RE state is not determined solely by its arrival rate.

To verify this conclusion, we design a specific scenario that focuses on 3-mile road segment with one lane. We focus on the 4 hours around the peak period. Assume that the 4 hours are treated as an RE with 8 states (each RE state lasts 0.5 hours). Two cases are analyzed, where the arrival rate vectors are (2150, 2400, 2400, 2100, 2000, 1800, 1800, 1800) and (2150, 2600, 2600, 2100, 2000, 1800, 1800, 1800). The two arrival rate vectors differ only in the arrival rate for the peak period (RE states 2 and 3). The SCV of the interarrival times is 5 in both cases. The other settings remain the same as in the previous experiment. The conditional queue length is obtained by the proposed method. We also use the existing PH/PH(n)/C/C queuing model in the study by Hu et al. [15] for comparison. In the control group, the queue length for each RE state is calculated by the PH/PH(n)/C/C queuing model with the arrival rate for each RE state in isolation. For example, 2150 is used to calculate the queue length for RE state 1 and 2400 is used for RE state 2 in case 1. This result is referred to as the piecewise result. The results of the two methods are presented in Figure 6.

Figure 6 shows that the conditional queue length $EN(i)$ is very different from the piecewise queue length. At the peak period of RE states 2 and 3, the piecewise queue length is much larger than the conditional queue length. After the peak period, the piecewise queue length is smaller because the piecewise queue length is the result of isolated RE states and is not affected by other RE states. Therefore, the queue length at RE states 2 and 3 is close to the capacity. The queue length is not influenced by RE state 1, where the traffic intensity is low. RE states 2 and 3 do not have a prolonged impact, and the queue length drops sharply after the peak period.

By contrast, the conditional queue length obtained by the proposed model reflects the traffic intensity of the current and previous RE states. The conditional queue length for RE states 2 and 3 is influenced by the unsaturated RE state 1. Therefore, it is much smaller than the piecewise result. Meanwhile, the traffic intensity at RE states 2 and 3 has a significant after-effect on the queue length of the subsequent RE states. Because of the aftereffect, the conditional queue length for RE states 4 to 8 is much larger than that of the piecewise result.

The difference between the two lines during RE states 4 to 8 quantifies the aftereffect of the peak period. The queue accumulated during the peak period (RE states 2 and 3) is dissipated in the subsequent RE states. The difference between Figures 6(a) and 6(b) reveals that different traffic intensities during the peak period have different aftereffects.

In case 1, the difference between the two lines (the conditional queue length and the piecewise queue length) is small and is close to zero at RE state 8, which means the effect of the peak period, i.e., congestion, has almost completely dissipated. In case 2, the difference between the two lines is much larger due to the larger traffic intensity during the peak period. The difference between the two lines is still very large at RE state 8, which means the effect of the peak period is not fully dissipated and will continue to affect the performance of the subsequent RE states. In circulation systems with periodic RE, this type of information is particularly important. To avoid deterioration of the system performance, we should ensure that the accumulated queue or the aftereffect of one peak period is completely dissipated before the next peak period arrives.

5.3. Varying Corridor Parameters for Pedestrian Queues in an RE. In this subsection, we examine the effect of varying width and length of a circulation facility. Here, we analyze passenger queues where the traffic demand is nonuniform, such as that in metro stations, stadiums, and comprehensive transportation hubs. According to a rough survey of Chengdu metro stations, the average headway of trains is 240 s. For outbound or transfer facilities in metro stations, most passenger arrive in the beginning 40 s of one headway and after that the arrival rate will be very small. We apply an RE with three states to model the headway of metro trains. The duration of the three RE states is, respectively, 40 s, 40 s, and 160 s. The invariant probability vector of this RE is $\mathbf{x} = 1/6(1, 1, 4)$, and the infinitesimal generator of the RE is as follows:

$$\mathbf{Q} = \frac{1}{160} \begin{bmatrix} -4 & 4 & 0 \\ 0 & -4 & 4 \\ 1 & 0 & -1 \end{bmatrix}. \tag{25}$$

First, we analyze the effect of corridor width. A corridor of 15 meters in length is analyzed. The width varies from 0.5 to 7.5 meters. We apply three arrival rate vectors with increasing nonuniformity, namely, Case 1 (2.8, 2.4, 1.1), Case 2 (3.4, 3, 0.8), and Case 3 (4.5, 3.5, 0.4). RE state 1 is the peak period and RE state 3 is the off-peak period. The three cases have the same average arrival rates as 1.6 ped/s. The PH/PH(n)/C/C queuing model in the study by Hu et al. [15] applying the average arrival rare is used as a control. The average output rate for varying widths is obtained and presented in Figure 7. Here, the result of our model is the average output rate for three RE states.

Figure 7 shows that the average output rate always increases with the width until it equals the average arrival rate. For the arrival rate vector with greater nonuniformity, a larger width is required to obtain the maximum average output rate. The curves also indicate that the output rate of the PH/PH(n)/C/C queuing model is more sensitive to width compared with our model. When the width increases, the output rate of the PH/PH(n)/C/C queuing model grows faster. For Cases 1 to 3 where the arrival nonuniformity gets
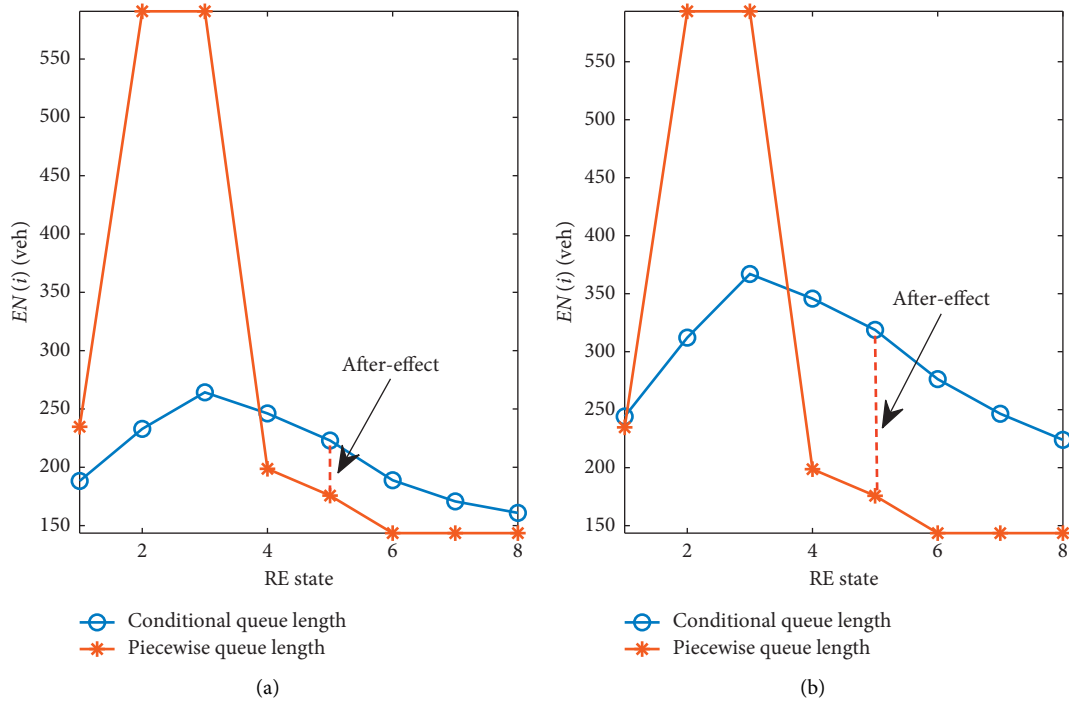
FIGURE 6: Comparison between conditional and piecewise queue length. (a) Small arrival rate for the peak period, (b) Large arrival rate for the peak period.
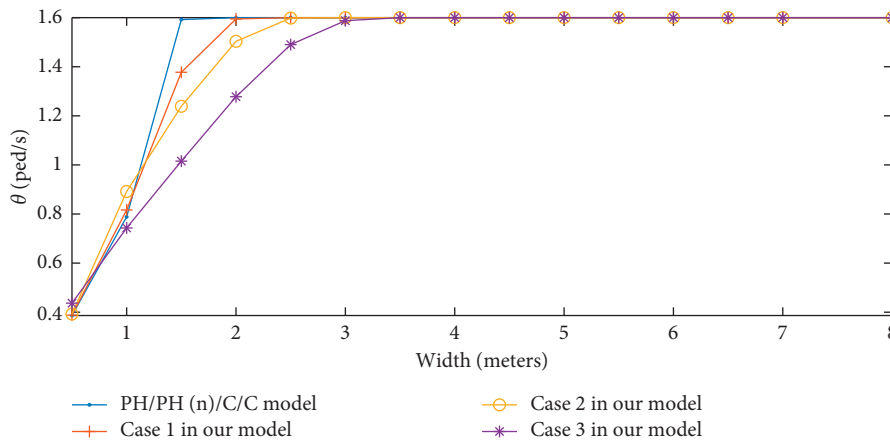

FIGURE 7: Average output rate for varying widths.

greater, the average output rate grows slower when the width increases. This is because the conditional output rate at RE state 3 becomes smaller sequentially for Cases 1 to 3.

Next, we study the effect of corridor length. A corridor with a constant width of 1.5 meters is analyzed. The length varies from 3 meters to 90 meters. We apply two groups of experiments, one with low traffic demand and one with high traffic demand. In Group (a), the average arrival rate is 1.6 ped/s. Three arrival rate vectors are Case 1 (2.8, 2.4, 1.1), Case 2 (3.4, 3, 0.8), and Case 3 (4.5, 3.5, 0.4). In Group (b), the average arrival rate is 1.75 ped/s. The three arrival rate vectors with increasing nonuniformity are Case 1 (2.1, 2, 1.6), Case 2 (3.1, 2.6, 1.2), and Case 3 (3.6, 3.3, 0.9). In both

groups, RE state 1 is the peak period. The PH/PH(n)/C/C queuing model in Hu et al. [15] applying the average arrival rare is used as a control. The average output rate is calculated and presented in Figure 8.

Figures 8(a) and 8(b) show different tendencies as the length grows. This is because the growth in length has two opposite effects on the output rate. On the one hand, there is more land space and more pedestrians are allowed to get into the corridor when the length gets larger. This is the positive effect that improves the output rate. On the other hand, more pedestrians getting into the corridor leads to a longer accumulated queue and worsens the congestion. This is the negative force that brings down the output rate. In
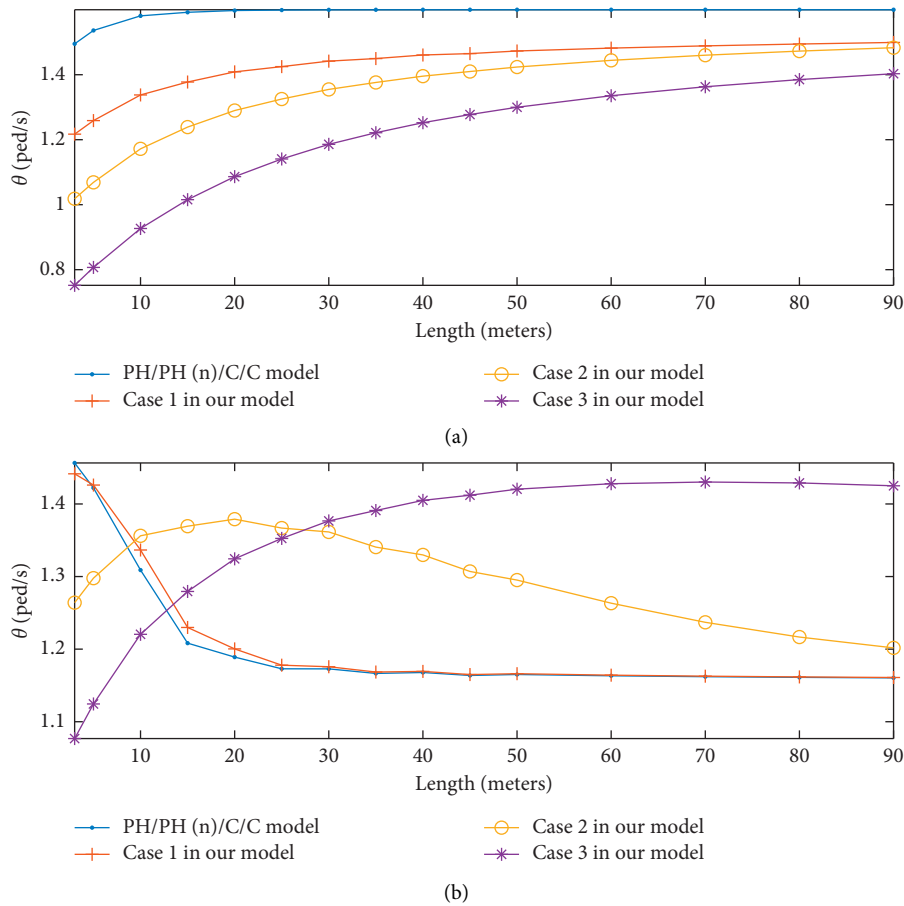
FIGURE 8: Average output rate for varying lengths. (a) Low traffic demand. (b) High traffic demand.

Figure 8(a) where the traffic demand is low, most of the accumulated queue dissipates at RE state 3 and the positive effect is larger. Therefore, the average output rate for both methods increases with the corridor length. In Figure 8(b), for the PH/PH (n)/C/C queuing model and Case 1 of our model, the traffic demand is at a high level for the three RE states and the accumulated queue cannot dissipate effectively. Therefore, the negative effect is larger and the output rate decreases when the length grows. For Case 2 and Case 3 of our model where the arrival nonuniformity is large, the positive effect is dominant at the first stage. As the arrival rate is relatively small at RE state 3 in these two cases, the congestion greatly dissipates. However, when the corridor length exceeds a certain value, the negative effect plays a leading role as the accumulated queue becomes too long and cannot dissipate timely during the off-peak period.

Meanwhile, the average output rate gets smaller from Case 1 to Case 3 as the nonuniformity in traffic demand increases (Figure 8(a)). This is because increasing nonuniformity in traffic demand leads to more pedestrians failing to get in the corridor during the peak period due to congestion. In Figure 8(b), when the length is smaller than 10, the situation is the same as that in Figure 8(a). However, when the length is large enough, the average output rate becomes larger from Case 1 to Case 3 due to the above reason.

The experiment on varying lengths suggests that nonuniformity in traffic demand affects the optimal length of circulation facilities. In the situation where the congestion accumulated at the peak period can dissipate greatly during the off-peak period, the average output rate increases with the length. However, if the congestion accumulated at the peak period cannot dissipate effectively within one cycle, the average output rate declines as the length grows.

6. Conclusion

This paper proposes a stationary PH (i)/PH (i, n)/C/C state-dependent queuing model operating in an RE to model traffic circulation systems under the influence of randomly changing environmental factors, such as the daily traffic period, accidents, weather conditions and so on. The RE is represented by a Markov chain and the duration time of each RE state is random. At the same time, the proposed model considers the general randomness in the arrival and service process by using the PH distribution. It also explicitly models congestion by applying a state-dependent service time. The MAS algorithm is extended to solve the proposed model and obtain both the average and the conditional performance measures. The time complexity of the extended MAS algorithm is a linear function of the product of the

queue capacity and the number of RE states. The space complexity of the extended algorithm is only a linear function of the number of RE states and is independent of the enormous (four-dimensional) state space of the Markov process for the queuing system. The accuracy of the proposed model is verified through comparisons with the M/M/1/∞ queuing model in an RE in the study by Neuts [39]; the transient M(t)/G (x)/C/C fluid queuing model in the study by Hu et al. [13] and the discrete-event simulation model.

The obtained conditional performance measures can accurately capture the queue accumulation and dissipation and reveal the effect of randomly changing environments. Such information is particularly important when analyzing circulation systems with nonuniform traffic demand. Numerical experiments on vehicle and pedestrian queues are conducted. Some interesting results are revealed.

(1) The proposed stationary PH(i)/PH(i, n)/C/C state-dependent queuing model coincides with the transient M(t)/G (x)/C/C fluid queuing model under the following conditions: the SCV of the interarrival times of the proposed model equals 1, the SCVRE of the proposed model is close to 0, and the conditional queue length of the proposed model is close to 0 at the first and the last RE state. This result implies that the proposed stationary model may be useful in conducting transient analysis after certain transformations, which will be explored in our future work.

(2) Under high traffic intensities, increasing the randomness in the duration time of the RE state leads to an obvious growth in the conditional queue length. This effect is quite small under low traffic intensities.

(3) Nonuniformity in traffic demand plays an important role in determining the optimal length of circulation facilities. When the congestion accumulated at the peak period can dissipate greatly during the off-peak period, the average output rate increases with the length. However, if the congestion accumulated at the peak period cannot dissipate effectively within one cycle, the average output rate declines as the length grows.

(4) The average output rate always increases with the width until it equals the average arrival rate. For traffic demand with greater nonuniformity, a larger width is required to obtain the maximum average output rate.

(5) The SCV of interarrival times has a remarkable effect on the conditional mean queue length under high traffic intensities. Moreover, a larger SCV of interarrival times results in a larger conditional CV of queue length. The conditional CV of queue length tends to be smaller at RE states with a larger conditional mean queue length. When the congestion begins to dissipate after the peak period, the conditional CV of queue length increases.

The proposed model serves as a solid base for robust optimal design for traffic circulation systems owing to its comprehensively considerations. In addition to the circulation facilities discussed in this paper, the proposed model can also be applied to analyze facilities with parallel servers, such as ticket vending machines and toll stations, due to the state-dependent service time. Applications in other fields, such as logistics and manufacturing, can also be explored, taking advantage of the generality of the PH distribution and the Markov RE.

Nevertheless, some issues remain to be solved in the future. We assume that only the arrival rate and service rate change with the RE. In cases such as accidents or facility disruptions, the capacity of the circulation system or even the service discipline varies. Queuing models that can account for these complex situations will be universal but also more complex. The circulation facilities are described as finite capacity queuing systems without feedback. However, in most actual situations, the newly arrived customers have to visit the system repeatedly when the system is full and they have no other option. This feedback effect will considerably influence the performance of the congested facility, as well as that of upstream facilities. The queuing network model with the feedback effect will be discussed in our following research.

Data Availability

The data that support the findings of this study are available from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] W. W. Yue, C. Li, Y. Chen, P. Duan, and G. Mao, "What is the root cause of congestion in urban traffic networks: road infrastructure or signal control?" *IEEE Transactions on Intelligent Transportation Systems*, vol. 183, pp. 1–18, 2021.

[2] R. Zhao, D. Wang, Y. Wang et al., "Macroscopic view: crowd evacuation dynamics at t-shaped street junctions using a modified aw-rascle traffic flow model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6612–6621, 2021.

[3] M. Baykal-Gürsoy, W. Xiao, and K. Ozbay, "Modeling traffic flow interrupted by incidents," *European Journal of Operational Research*, vol. 195, no. 1, pp. 127–138, 2009.

[4] M. D. Rosini, *Macroscopic Models for Vehicular Flows and Crowd Dynamics: Theory and Applications*, Springer, New York, NY, USA, 2013.

[5] J. Hu, L. You, H. Zhang, J. Wei, and Y. Guo, "Study on queueing behavior in pedestrian evacuation by extended cellular automata model," *Physica A: Statistical Mechanics and its Applications*, vol. 489, pp. 112–127, 2018.

[6] E. Papadimitriou, G. Yannis, and J. Golias, "A critical assessment of pedestrian behaviour models," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 3, pp. 242–255, 2009.

[7] P. Y. WuWu and R. Y. GuoGuo, "Simulation of pedestrian flows through queues: effect of interaction and intersecting angle," *Physica A: Statistical Mechanics and Its Applications*, vol. 570, Article ID 125804, 2021.

[8] G. G. Lovas, "Modeling and simulation of pedestrian traffic flow," *Transportation Research Part B: Methodological*, vol. 28, no. 6, pp. 429–443, 1994.

[9] M. Parlar and M. Sharafali, "Dynamic allocation of airline check-in counters: a queueing optimization approach," *Management Science*, vol. 54, no. 8, pp. 1410–1424, 2008.

[10] C. Arita and A. Schadschneider, "Exclusive queueing processes and their application to traffic systems," *Mathematical Models and Methods in Applied Sciences*, vol. 25, no. 3, pp. 401–422, 2015.

[11] M. Tréca, J. Garbiso, D. Barth, and M. Zargayouna, "Fast bootstrapping for reinforcement learning-based traffic signal control systems using queueing theory," in *Proceedings of the 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pp. 1–5, IEEE, Victoria, BC, Canada, November 2020.

[12] A. Hasani Goodarzi, E. Diabat, A. Jabbarzadeh, and M. Paquet, "An M/M/c queue model for vehicle routing problem in multi-door cross-docking environments," *Computers & Operations Research*, vol. 138, Article ID 105513, 2022.

[13] L. Hu, B. Zhao, J. Zhu, and Y. Jiang, "Two time-varying and state-dependent fluid queuing models for traffic circulation systems," *European Journal of Operational Research*, vol. 275, no. 3, pp. 997–1019, 2019.

[14] S. J. Yuhaski, J. M. Smith, and J. M. Smith, "Modeling circulation systems in buildings using state dependent queueing models," *Queueing Systems*, vol. 4, no. 4, pp. 319–338, 1989.

[15] L. Hu, Y. Jiang, J. Zhu, and Y. Chen, "A PH/PH(n)/C/C state-dependent queuing model for metro station corridor width design," *European Journal of Operational Research*, vol. 240, no. 1, pp. 109–126, 2015.

[16] M. F. Neuts, *Matrix-geometric solutions to stochastic models*, Dover Publications, Mineola, NY, USA, 1984.

[17] P. Buchholz, J. Kriege, and I. Felko, *Input Modeling with Phase-type Distributions and Markov Models: Theory and Applications*, Springer, New York, NY, USA, 2014.

[18] P. C. L. Gerum and M. Baykal-Gürsoy, "How incidents impact congestion on roadways: a queuing network approach," *EURO Journal on Transportation and Logistics*, vol. 11, Article ID 100067, 2022.

[19] Q. Yang, Z. Qiao, B. Yang, and Z. Shi, "Modeling and uncovering the passenger-taxi dynamic queues at taxi station with multiple boarding points using a Markovian environment," *Physica A: Statistical Mechanics and Its Applications*, vol. 572, Article ID 125870, 2021.

[20] C. S. Kim, V. Klimenok, V. Mushko, and A. Dudin, "The BMAP/PH/N retrial queueing system operating in Markovian random environment," *Computers & Operations Research*, vol. 37, no. 7, pp. 1228–1237, 2010.

[21] J. Wu, Z. Liu, and G. Yang, "Analysis of the finite source MAP/PH/N retrial g-queue operating in a random environment," *Applied Mathematical Modelling*, vol. 35, no. 3, pp. 1184–1193, 2011.

[22] Y. Liu and W. Whitt, "The G t/GI/s t +GI many-server fluid queue," *Queueing Systems*, vol. 71, no. 4, pp. 405–444, 2012.

[23] S. Ammar, "Behavior Analysis of an M/M/1 Vacation Queue in Random Environment," *Quality Technology & Quantitative Management*, vol. 18, pp. 1–21, 2020.

[24] R. Jain and J. M. Smith, "Modeling vehicular traffic flow using M/G/c/c state dependent queueing models," *Transportation Science*, vol. 31, no. 4, pp. 324–336, 1997.

[25] J. M. Smith and F. Cruz, "M/G/c/c state dependent travel time models and properties," *Physica A: Statistical Mechanics and Its Applications*, vol. 395, pp. 560–579, 2014.

[26] J. Zhu, L. Hu, Y. Jiang, and A. Khattak, "Circulation network design for urban rail transit station using a PH(n)/PH(n)/C/C queuing network model," *European Journal of Operational Research*, vol. 260, no. 3, pp. 1043–1068, 2017.

[27] U. Yechiali and P. Naor, "Queuing problems with heterogeneous arrivals and service," *Operations Research*, vol. 19, no. 3, pp. 722–734, 1971.

[28] M. F. Neuts, "A queue subject to extraneous phase changes," *Advances in Applied Probability*, vol. 3, no. 01, pp. 78–119, 1971.

[29] S. l. Yu and Z. m. Liu, "Analysis of queues in a random environment with impatient customers," *Acta Mathematicae Applicatae Sinica, English Series*, vol. 33, no. 4, pp. 837–850, 2017.

[30] R. Krenzler and H. Daduna, "Loss systems in a random environment: steady state analysis," *Queueing Systems*, vol. 80, no. 1-2, pp. 127–153, 2015.

[31] C. S. Kim, A. Dudin, S. Dudin, and O. Dudina, "Analysis of an MMAP/PH1, PH2/N/∞ queueing system operating in a random environment," *International Journal of Applied Mathematics and Computer Science*, vol. 24, no. 3, pp. 485–501, 2014.

[32] H. Baumann and W. Sandmann, "Computing stationary expectations in level-dependent QBD processes," *Journal of Applied Probability*, vol. 50, no. 01, pp. 151–165, 2013.

[33] H. Baumann and W. Sandmann, "Numerical solution of level dependent quasi-birth-and-death processes," *Procedia Computer Science*, vol. 1, no. 1, pp. 1561–1569, 2010.

[34] J. F. Pérez and B. Van Houdt, "Quasi-birth-and-death processes with restricted transitions and its applications," *Performance Evaluation*, vol. 68, no. 2, pp. 126–141, 2011.

[35] R. Sadre, *Decomposition-Based Analysis of Queueing Networks*, University of Twente, Enschede, Netherlands, 2007.

[36] S. Balsamo and A. Marin, "Separable solutions for Markov processes in random environments," *European Journal of Operational Research*, vol. 229, no. 2, pp. 391–403, 2013.

[37] B. L. Nelson and M. R. Taaffe, "The pht/pht/∞ queueing system: Part I-the single node," *INFORMS Journal on Computing*, vol. 16, no. 3, pp. 266–274, 2004.

[38] M. Eisen and M. Tainiter, "Stochastic variations in queuing processes," *Operations Research*, vol. 11, no. 6, pp. 922–927, 1963.

[39] M. F. Neuts, "Further results on the M/M/1 queue with randomly varying rates," ADA054884 Tech. Rep, University of Delaware, Newark, DE, USA, 1978.

[40] C. O'cinneide and P. Purdue, "The M/M/∞ queue in a random environment," *Journal of Applied Probability*, vol. 58, pp. 175–184, 1986.

[41] G. Pang, A. Sarantsev, Y. Belopolskaya, and Y. Suhov, "Stationary distributions and convergence for m/m/1 queues in interactive random environment," *Queueing Systems*, vol. 94, no. 3-4, pp. 357–392, 2020.

[42] V. Naumov and K. Samouylov, "Resource system with losses in a random environment," *Mathematics*, vol. 9, no. 21, p. 2685, 2021.

[43] J. Sztrik, "On the heterogeneous M/G/n blocking system in a random environment," *Journal of the Operational Research Society*, vol. 38, no. 1, p. 57, 1987.

[44] U. Krieger, V. I. Klimenok, A. Kazimirsky, L. Breuer, and A. N. Dudin, "A BMAP|PH| 1 queue with feedback operating in a random environment," *Mathematical and Computer Modelling*, vol. 41, no. 8, pp. 867–882, 2005.

[45] R. W. Conway and W. L. Maxwell, "A queuing model with state dependent service rates," *Journal of Industrial Engineering*, vol. 12, no. 2, pp. 132–136, 1962.

[46] J. Y. Cheah and J. M. Smith, "Generalized M/G/c/c state dependent queueing models and pedestrian traffic flows," *Queueing Systems*, vol. 15, no. 1-4, pp. 365–386, 1994.

[47] J. MacGregor SmithMacGregor Smith, "State-dependent queueing models in emergency evacuation networks," *Transportation Research Part B: Methodological*, vol. 25, no. 6, pp. 373–389, 1991.

[48] D. H. Mitchell and J. MacGregor Smith, "Topological network design of pedestrian networks," *Transportation Research Part B: Methodological*, vol. 35, no. 2, pp. 107–135, 2001.

[49] A. Weiss, L. Williams, and J. MacGregor SmithMacGregor Smith, "Performance & optimization of M/G/c/c building evacuation networks," *Journal of Mathematical Modelling and Algorithms*, vol. 11, no. 4, pp. 361–386, 2012.

[50] Y. Jiang, L. Hu, J. Zhu, and Y. Chen, "PH fitting of the arrival interval distribution of the passenger flow on urban rail transit stations," *Applied Mathematics and Computation*, vol. 225, pp. 158–170, 2013.

[51] Y. Chen, B. Yi, Y. Jiang, J. Sun, and M. Wahab, "Inter-arrival time distribution of passengers at service facilities in underground subway stations: a case study of the metropolitan city of chengdu in China," *Transportation Research Part A: Policy and Practice*, vol. 111, pp. 227–251, 2018.

[52] Y. Jiang, J. Zhu, L. Hu, X. Lin, and A. Khattak, "A G/G(n)/C/Cstate-dependent simulation model for metro station corridor width design," *Journal of Advanced Transportation*, vol. 50, no. 3, pp. 273–295, 2016.

[53] M. F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Marcel Decker Inc, New York, NY, USA, 1989.

[54] A. Weerstra, "Using Matrix-Geometric Methods to Enhance the Qna Method for Solving Large Queueing Networks," Master's thesis, Dept. of Computer Science, University of Twente, Enschede, Netherlands, 1994.

[55] G. Yang, L. g. YaoYao, and Z. s. OuyangOuyang, "The MAP/PH/N retrial queue in a random environment," *Acta Mathematicae Applicatae Sinica, English Series*, vol. 29, no. 4, pp. 725–738, 2013.

[56] P. Tregenza, *The Design of interior Circulation*, Van Nostrand Reinhold, Washington, DC, USA, 1976.

[57] L. Hu, *The Optimization Theory and Method Considering the Stability Analysis for Configuration of Transportation Facilities in Urban Rail Transit Station*, Ph.D. thesis, Southwest Jiaotong University, Chengdu, China, 2014.

[58] G. Zhang, "A methdd of intersection singal timing and optimization considering the general arriving," Master's Thesis, Southwest Jiaotong University, Chengdu, China, 2013.