Hindawi
Journal of Advanced Transportation
Volume 2023, Article ID 3241207, 15 pages
https://doi.org/10.1155/2023/3241207

Research Article

Queue Length Estimation Based on Probe Vehicle Data at Signalized Intersections

Huidan Luo,[1] **Mingjun Deng**,[1] **and Jia Chen**[2]

[1]*School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China*
[2]*Department of Traffic Management, Jiangxi Youth Vocational College, Nanchang 330013, China*

Correspondence should be addressed to Mingjun Deng; dmstd98@163.com

Received 18 April 2023; Revised 16 June 2023; Accepted 6 July 2023; Published 17 July 2023

Academic Editor: Zhihong Yao

Queue length is one of the important indexes to evaluate the operation efficiency of signalized intersection and also the key parameter of intersection signal control optimization. Traditional queue length estimation models are mostly based on fixed detection equipment, and the models assumptions are too harsh; there are certain limitations. Based on the probe vehicle data, this paper establishes a model of queue length estimation for signalized intersection based on shockwave theory. First, based on the speed and location data of the probe vehicle, the vehicle density is calculated to estimate the intersection stop line. A real-time calculation method of vehicle arrival rate is proposed to improve the applicability of the model. Then, based on the shockwave theory, the meeting time of the queue forming wave and the queue discharging wave are calculated after the green light is on. Finally, the queue length is summed in sections, including the distance between the last queued probe vehicle and the stop line during the red light period, the length of the subsequent vehicles arriving during the residual red light time, and the newly increased queue length within the queue discharging time. This paper uses the VISSIM software to simulate the actual intersection. The simulation results show that when the penetration of probe vehicle is 50%, 25%, and 10%, their corresponding mean absolute relative error are 11.27%, 27.77%, and 39.12%, respectively. It can be seen that with the increase of penetration, the error gradually decreases. The average absolute relative error is within the acceptable range. After analyzing the existing similar methods, although the accuracy of the method proposed in this paper does not reach the highest level, it has the advantages of simple operation, less computation, and good real-time computation. Relevant research results can provide support for traffic control at signalized intersections.

1. Introduction

With the total quantity of urban vehicles increasing rapidly, the urban traffic system is facing congestion caused by vehicle queues at signalized intersections. How to improve the efficiency of signalized intersections becomes the key to solve traffic congestion. Queue length is an important index to evaluate the operational efficiency of signalized intersections [1]. Timely and accurate acquisition of queue length can provide effective data basis for the management and optimization of signal timing at intersections [2], with the queue length minimization as the optimization objective, the phase and timing of the signal can be optimized accordingly [3], and signal optimization based on queue length can solve the problem of queue overflow caused by the failure of traditional signal control strategies in dealing with oversaturation at intersections. Therefore, it is particularly important to estimate the queue length of signalized intersection accurately in real time.

For a long time, a large number of scholars have studied queue length estimation at signalized intersections. The existing queue length estimation methods can be mainly divided into two broad categories: input-output models and shockwave models. The first method estimates the queue length by analyzing the cumulative traffic arrivals and departures curve. Lee et al. [4] applied the Kalman filter to predict the downstream arrival and estimated the lane queue length in real time based on the discriminant model. Vigos et al. [5] employed a Kalman filter to estimate the number of cars contained in a signalized link based on real-time measurements of flow and occupancy provided by at least three loop detectors. The second method estimates the queue length by reconstructing the queue forming and discharging process. Wang et al. [6] established the queue length estimation model by applying the shockwave theory based on the probe vehicle and loop detector data. Horvath and Tettamanti [7] proposed a method based on shockwave theory to estimate the queue length of urban signal road network in real time by using the Kalman filter. Li et al. [8] established a queue length prediction model for multilane signalized intersections, which combined with shockwave theory and the platoon dispersion model to predict the queue length of each lane in real time. However, most of these models are based on ideal conditions and there are some limitations in practical application. In addition, data collection in the earlier studies mainly relied on fixed detectors such as loop or video detection devices, and the accuracy of the collected data depended on the stability of data acquisition devices, which increased the difficulty of queue length estimation. The laying position of the loop will affect the reliability data acquisition, and it is easily damaged and has high maintenance costs. The loop's detection accuracy will be also significantly reduced in traffic jams. When the video detection is blocked by external objects, the detection accuracy will be greatly reduced, and the video detection is limited by the visual range [9]. Note that fixed position sensors can only monitor traffic at specific locations. To be able to collect data in large-scale networks, sensors need to cover a wide range of necessary locations that are costly to install, maintain, and operate. If some of these sensors break down, the error in obtaining queue information will be high. Therefore, a low cost and relatively reliable method needs to be designed to achieve this. Compared with the traditional data acquisition technology, probe vehicle technology is becoming a research hotspot at home and abroad for its advantages such as low cost, wide coverage, real-time traffic data acquisition, and high accuracy. A probe vehicle is a motor vehicle equipped with GPS data acquisition equipment and wireless communication devices to realize real-time transmission of vehicle time and location information. Probe vehicle data refer to ID number, longitude and latitude, speed, and other data obtained by GPS transceiver installed on taxis or buses, which are widely used to evaluate the traffic operational condition at signalized intersections [10]. Please also note that the existing probe vehicle system mainly selects taxis, whose daily driving distance is much longer than that of ordinary private cars; so, there are no privacy concerns [11]. However, the probe vehicle technology cannot obtain all the information of the vehicle, only some of the collected data can be obtained, and the prediction accuracy cannot be guaranteed in the low penetration rate environment. In addition, the probe vehicle data distribution is uneven and there are errors due to the network delay caused by the probe vehicle being blocked by urban tall buildings or vegetation. These uncertainties are a challenge in estimating queue length using the probe vehicle data.

At present, many studies have been conducted at home and abroad to estimate the queue length of signalized intersections based on probe vehicle data. Liu et al. [12] comprehensively considered the statistical average traffic flow, the queue length time series of historical cycles, and the stopping status of real-time connected vehicle (CV) arrival features in the current cycle and then proposed a method for estimating real-time queue length based on the Markov model. Zhao et al. [13] proposed a method based on the hidden Markov model to estimate the queue length using a queue correlation of probe vehicles at different traffic signal cycles. Tan et al. [14] deduced the queue length with maximum probability of undersaturated and oversaturated conditions at signalized intersections based on the Bayesian theory. Mei et al. [15] proposed a new Bayesian method to estimate the maximum queue length of vehicles at signalized intersections by using the high-frequency trajectory data from probe vehicles. It can also have considerable accuracy under conditions of low penetration of probe vehicle. Tan et al. [16] estimated the queue length at the intersection with known signal cycles based on fusing real-time and historical probe vehicle trajectory data by a statistical parameter estimation method (i.e., maximum likelihood estimation (MLE)). Zhao et al. [17] proposed a maximum likelihood estimation method to estimate the penetration rate of probe vehicle and the queue length distribution of the studied intersection using the historical probe vehicle data, so that the existing queue length estimation methods can estimate the queue length cycle-by-cycle. Comert [18] systematically derived a series of estimators for the permeability and arrival rate of the probe vehicle under the Poisson arrival hypothesis and then developed a model for estimating the queue length from cycle-to-cycle in real time by inputting some basic information provided by the probe vehicle (such as location, time, and count). Zhao et al. [19, 20] estimated the penetration rate of probe vehicles based on the distribution of stopping positions of probe vehicles at intersections and used the estimated penetration rate to proportionally increase the number of probe vehicles in the queues and in the traffic, thereby giving an estimate of the total queue length. Rostami Shahrbabaki et al. [21] used a data fusion method to estimate the queue length and vehicle accumulation in links in real time by combining the location and speed data of the probe vehicle with the input flow collected by the fixed loop upstream of the road. Li et al. [22] proposed a data fusion method combining probe vehicle data with loop detector data, using the Kalman filters to estimate the queue length on a period-by-period basis.

In the abovementioned methods of estimating vehicle queue length using probe vehicle, most of them use the Markov model, hidden Markov model, Bayesian probability model, statistical parameter estimation, etc. These analysis and calculation methods and models belong to the branch of probability distribution model, but the probability distribution model is computationally complex and need to estimate the queue length based on a large number of historical statistical data and distribution parameters, and the estimation of parameters and conditional probability values are difficult, which limit the practicability of these methods. At the same time, most of the existing research studies are based on the known information of vehicle arrival mode, such as assuming that the vehicle arrival mode is Poisson distribution or uniform distribution. However, due to the random dynamic changes of vehicle arrival rate, this method of limiting vehicle arrival mode will reduce the effectiveness of the corresponding queue length estimation model, which is prone to errors in practical application. In addition, the previous studies did not fully consider the newly increased queue length during the meeting time of the queue forming wave and the queue discharging wave after the green light was on, which eventually led to the deviation between the estimated queue length and the actual value.

In response to the limitations of the existing methods for estimating queue length, this paper establishes a queue length estimation model of signalized intersection based on shockwave theory by using probe vehicle data. Different from previous studies, this model does not assume the arrival distribution of vehicles, and a real-time calculation method of vehicle arrival rate is proposed by combining the vehicle running law. In addition, considering the real queueing situation of the actual road, this model uses the shockwave theory to calculate the newly increased queue length within the queue discharging time, which improves the practicability and accuracy of the estimation result of the queue length. The process of establishing the queue length estimation model in this paper is as follows: first, the data returned by the probe vehicle are used to estimate the position of the intersection stop line and the last queued probe vehicle during the red light period. On that basis, the average arrival rate of the newly arrival vehicles after the last queued probe vehicle is calculated based on the positional relationship between the last queued probe vehicle, downstream queued probe vehicle, and the stop line. Then, the meeting time of the queue-forming wave and the queue-discharging wave is calculated according to shockwave theory, and the meeting time is same as the queue discharging time. Finally, the periodic queue length is obtained by the sum of the distance between the last queued probe vehicle and the stop line during the red light period, the length of subsequent vehicles arriving during the residual red light time, and the newly increased queue length within the queue discharging time. This paper validated the effectiveness of the queue length estimation model through an actual intersection example, analyzed its practicality, and compared its accuracy under different probe vehicle penetration rates.

2. Assumptions

Zhao et al. [23] built a two-dimensional vehicle motion model based on optimal control to microsimulate the driving behavior of vehicles at intersections, which can handle the interaction with the road. Zhao et al. [24] proposed a comprehensive microscopic traffic flow model to describe the maneuvering behavior of human driving vehicles under interaction and verified that the model could accurately describe the maneuvering behavior, path, and speed passing sequence of vehicles under interaction combined with empirical data. However, the data source of this paper is GPS positioning information, which has errors and is difficult to describe the vehicle's motion state in detail at the microlevel. The model can only be built based on the statistical means, so certain assumptions need to be made. To simplify the discussion, the following assumptions are made to formulate the queue estimation problem:

(1) Intersection and signal timing information is known beforehand

(2) The driving trajectories of the queuing vehicle in the intersection are uniform speed-uniform deceleration-stopping-uniform acceleration-uniform speed

(3) There is at least one probe car in the queue during the red light period in each cycle

Assumption (1) states that the information of intersections is known, and we will not derive incorrect intersection information (signal timing, etc.), resulting in incorrect model establishment. Assumption (2) states that the driving status of the vehicle so that we can deduce the time of the probe vehicle just started queuing. Because the driver is rational, the vehicle always runs at a reasonable speed in the traffic flow on the urban road, and the acceleration and deceleration fluctuation of the vehicle is small, which can be approximated to conform to the motion law of the assumption. Assumption (3) indicates that a certain penetration rate of probe vehicle is necessary to effectively estimate the queue length. The influence of the penetration rate on the estimation accuracy of queue length is also studied.

3. Methodology

3.1. Stop Line Position Estimation. The stop line position estimation is a very sensitive process for the later queue length calculation, as faulty stop line position will lead to erroneous queue length directly. Axer et al. [25] have already tested the inference of the stop line position from aerial images. Nevertheless, it is possible to infer the location of stop line by matching the trajectory of vehicles with map location information, as an alternative method. Trajectories' data points with slow instantaneous speed lower than 5 km/h should reach the highest density in the near of the stop line. Then, the position of the highest density is the stop line.

The steps are as follows:

(1) First, the probe vehicle data with a speed less than 5 km/h are filtered out, and the approximate position range of the intersection where the stop line is located is determined as (D_1, D_2)

(2) Then, the range is counted at intervals of δ meters, and the number of probe vehicles at each interval is calculated as f_k $(k = 1, 2, 3, \ldots, \mathrm{D}_2 - \mathrm{D}_1/\delta)$, where k is the number of intervals

(3) Calculate the vehicle density ρ_k $(\rho_k = f_k/\delta T)$ of each interval, where T represents the acquisition time of the probe vehicle data

(4) Filter out the maximum vehicle density position k_{max} and get the position of the intersection stop line as $D_1 + \delta k_{max}$

3.2. The Distance between the Probe Vehicle and the Stop Line Position Estimation. When the instantaneous speed of a vehicle is lower than 5 km/h, we define the vehicle is queuing. As we assume that the signal timing is known ahead, so we can filter out the probe vehicle with a speed less than 5 km/h from the red light starts to the green light starts in every cycle and then calculate the distance between the probe vehicle and the stop line. The distance is sorted out and the position of probe vehicle at the end of the queue during the red light time is obtained in each cycle.

The data from probe vehicle are $P(t, \mathrm{ID}, x, y, s)$, where P denotes the points returned from the probe vehicle. t, ID, x, y, s denote the time, ID number, longitude, latitude, and speed of the probe vehicle, respectively. After estimating the stop line position (X, Y), then the distance L between the probe vehicle and the stop line can be calculated as follows:

$$L = \sqrt{(x - X)^2 + (y - Y)^2}. \tag{1}$$

3.3. Queue Length Estimation

3.3.1. Definition of Queue Length at Intersections. There are currently two main forms of defining the queue length at intersections. The first form assumes that the departure rate of vehicles during the green light period is greater than the arrival rate of vehicles; so, the number of vehicles queuing at the beginning of the green light period is the highest, which corresponds to the queue length at that time [26]. The second form believes that the departure rate of vehicles during the green light period is smaller than the arrival rate of vehicles, so the queue length will continue to increase for a period of time after the green light starts, until the queue discharging wave catches up with the queue forming wave; therefore, the queue length is defined as the length corresponding to a certain moment after the green light starts. The second form of queue length can more accurately reflect the actual queuing situation of vehicles on the road, so we define the queue length as the second form in this paper.

Based on the abovegiven definition of the queue length at intersections, we summarize the estimation of the queue length as follows: the distance between the last queued probe vehicle and the stop line during the red light period, plus the length of subsequent vehicles arriving during the residual red light time (the time from the moment the last queued probe vehicle just started queuing to the red light time ends), and finally, plus the newly increased queue length within the queue discharging time (the time it takes for the queue discharging wave catches up with the queue forming wave).

Figure 1 shows the queue length in a cycle. L_{true} denotes the true queue length of the intersection. L_{pvd} denotes the distance between the last queued probe vehicle and the stop line. L_{red} denotes the queue length at the end of the red light. L_{true} equals to L_{pvd} plus the length of subsequent vehicles arriving during the residual red light time, plus the newly increased queue length within the queue discharging time. Meanwhile, L_{true} can also be represented as L_{red} plus the newly increased queue length within the queue discharging time.

Generally, the time sent back by the probe vehicle in a fixed time interval is not the moment when the vehicle just started queuing but the moment when it is already in the queue. The same queued probe vehicle sent back data, respectively, when it is about to queue and when it is queuing. The data collected from the probe vehicle are as follows: the data that the vehicle is about to queue are $P_1(t_1, \mathrm{ID}, x_1, y_1, s_1)$, and the data which the vehicle is queuing are $P_2(t_2, \mathrm{ID}, x_2, y_2, s_2)$. In Figure 2, the blue vehicle represents the last queued probe vehicle, while the red vehicle represents the probe vehicle that has already been queued. The last queued probe vehicle is taken as the research target. When $t = t_1$, the blue vehicle is moving to approach the intersection, and there are following vehicles behind the blue vehicle. When $t = t'$, the blue vehicle just started queuing, and the following vehicles are still moving. When $t = t_2$, the blue vehicle has stopped for a while, and several vehicles might stop behind the blue vehicle.

We define the "residual red light time" as the time from the moment the last queued probe vehicle just started queuing to the red light time ends. t_2 cannot be the moment to calculate the residual red light time, or there will occur an error in estimating the queue length. Then, the moment when the last queued probe vehicle just started queuing is required. We define the "entry time" as the moment when the last queued probe vehicle just started queuing.

We assume that the driving trajectory of the queued probe vehicle is divided into the following parts: (1) the vehicle moves at a uniform speed after entering the road section. (2) When the vehicle is near the stop line or there is a vehicle in front of which is decelerating in line, the vehicle is uniformly decelerating. (3) The vehicle is queuing. (4) When the vehicle ends the queue, it will accelerate uniformly. (5) After reaching a certain speed, the vehicle will drive uniformly and leave the intersection.

After calculating the distance between the probe vehicle and the stop line, we set the uniform speed as s. The data which the vehicle is about to queue are $P_1(t_1, \mathrm{ID}, x_1, y_1, s_1)$; if $s_1 < s$, the vehicle is uniformly decelerating. If $s_1 \geq s$, the vehicle is still driving at a uniform speed for a while, then start to uniformly decelerate. We assume that the signal timing is known ahead. t' is the moment when the vehicle just stared queuing, that is, the "entry time." Figure 3 shows the distance between the vehicle and the stop line position over time in two different situations. Then, t' can be calculated as follows:

$$t' = \begin{cases} t_1 + \dfrac{2 \times (L_1 - L_2)}{s_1}, & s_1 < s, \\[2ex] t_1 + \dfrac{s_1}{a} + \dfrac{L_1 - L_2 - s_1^2/2a}{s_1}, & s_1 \geq s, \end{cases} \tag{2}$$

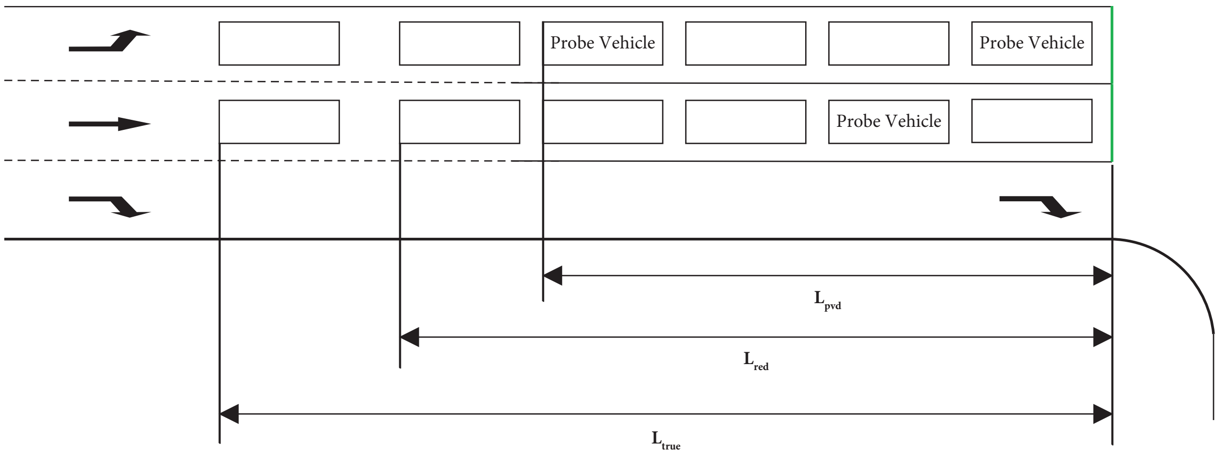


FIGURE 1: Schematic diagram of queuing scene.

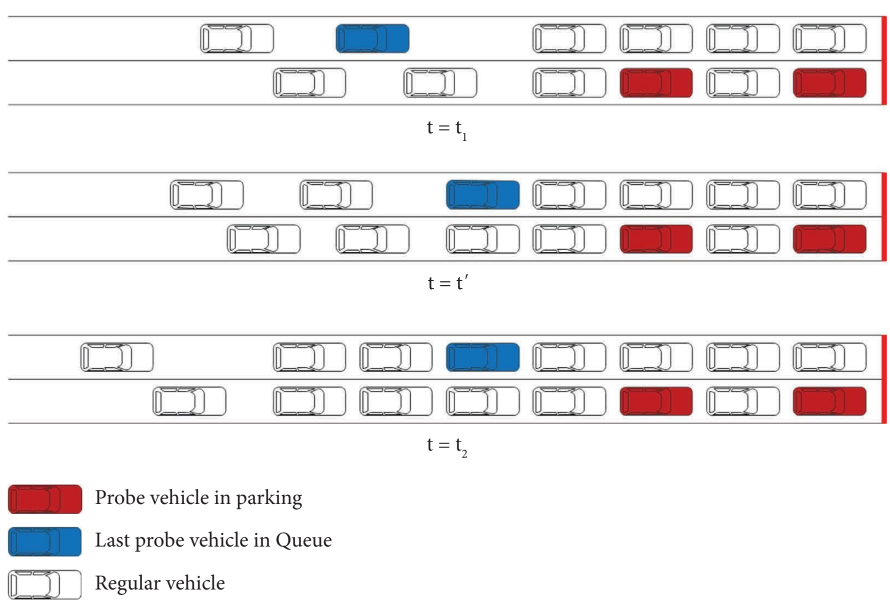


FIGURE 2: The intersection queue state when the last queued probe vehicle is moving ($\mathbf{t} = \mathbf{t_1}$), just started queuing ($\mathbf{t} = \mathbf{t}'$) and has stopped for a while ($\mathbf{t} = \mathbf{t_2}$) (the blue vehicle represents the last queued probe vehicle, the red vehicle represents the probe vehicle has already been queued, and the white vehicle represents the regular vehicle).

where a is the drag acceleration. In this paper, the value of a is obtained by the simulation data.

Based on the abovegiven summary of the estimated queue length at intersections, we need to know the distance between the last-queued probe vehicle and the stop line during the red light period, the newly increased queue length during the residual red light time, and the newly increased queue length within the queue discharging time. According to the estimation of the distance between the probe vehicle and the stop line in the chapter "3.2," the distance between the last-queued probe vehicle and the stop line can be calculated from the longitude and latitude data returned by the last-queued probe vehicle and the longitude and latitude data of the stop line. Besides that, the newly increased queue length during the residual red light time equals to the product of the vehicle arrival rate and the residual red light time. The newly increased queue length within the queue discharging time equals to the product of the vehicle arrival rate and the queue discharging time. So, the arrival rate of the newly arrival vehicles after the last-queued probe vehicle, the residual red light time, and the queue discharging time are required.

3.3.2. Calculation of the Arrival Rate of the Newly Arrival Vehicles. Before calculating the average vehicle arrival rate, we define $\overline{h_s}$ as the standard space headway when vehicles queuing, usually taken as 7 meters, and t' is defined as the last queued probe vehicle just started queuing during the red light period.

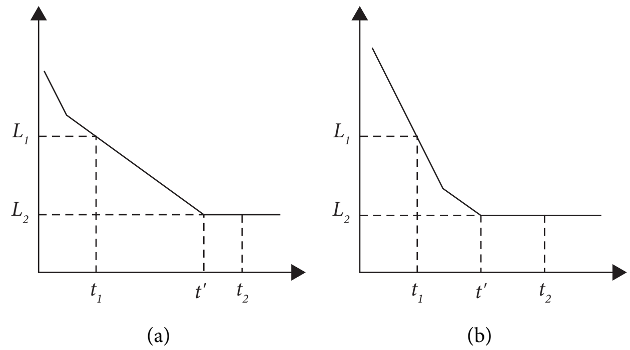


FIGURE 3: The distance over time. (a) $\mathbf{s_1} < \mathbf{s}$; (b) $\mathbf{s_1} \geq \mathbf{s}$.

(1) Only One Probe Vehicle. When there is only one probe vehicle in the queue during the red light period, the only one probe vehicle is regarded as the last queued probe vehicle, and we define the average vehicle arrival rate as q_{one}, and its calculation formula is as follows:

$$q_{\text{one}} = \frac{L_{\text{one}}}{\overline{h_s}\left(t' - T_1\right)}, \tag{3}$$

where L_{one} is the distance between the only one probe vehicle and the stop line. T_1 is the moment when the red light of the cycle starts.

(2) At Least Two Probe Vehicles. When there are at least two probe vehicles in the queue during the red light period, the arrival rate between the last queued probe vehicle and each of the queued probe vehicle downstream can be calculated, as is shown in Figure 4. Normally, the closer the downstream queued probe vehicle is to the last-queued probe vehicle, the closer the arrival rate calculated is to the arrival rate during the residual red light time and the queue discharging time. Then, the arrival rate will obtain greater weight. The weighted mean of every calculated arrival rate approximately equals to the arrival rate during the residual red light time and the queue discharging time.

We assume that the ID number of the queued probe vehicle downstream is i, and the ID number of the last queued probe vehicle is N; q_i denotes the arrival rate calculated by the last-queued probe vehicle and the queued probe vehicle downstream. Then, q_i can be calculated as follows:

$$q_i = \frac{L_N - L_i}{\overline{h_s}\left(t' - t_i'\right)}, \tag{4}$$

where L_N is the distance between the last queued probe vehicle and the stop line position. L_i is the distance between the queued probe vehicle downstream and the stop line position. t_i' is the moment when the queued probe vehicle downstream just started queuing.

In addition, we define u_i as the reciprocal of the distance between the last queued probe vehicle and each of the queued probe vehicle downstream. So, w_i denotes the weighting coefficient, and the sum total of the weighting coefficient is 1.

$$\begin{aligned} u_i &= \frac{1}{L_N - L_i}, \\ w_i &= \frac{u_i}{\sum_{i=1}^{N-1} u_i}, \\ \sum_{i=1}^{N-1} w_i &= 1. \end{aligned} \tag{5}$$

To sum up, when there are at least two probe vehicles in the queue during the red light period, we define the average vehicle arrival rate as q_{mul}, and its calculation formula is as follows:

$$q_{\text{mul}} = \sum_{i=1}^{N-1} w_i q_i. \tag{6}$$

3.3.3. Calculation of the Residual Red Light Time. We define the "residual red light time" (t_{rsd}) as the time from the moment the last-queued probe vehicle just started queuing to the red light time ends. Therefore, the formula for calculating the residual red light time is as follows:

$$t_{\text{rsd}} = T_2 - t', \tag{7}$$

where T_2 is the moment when the red light ends.

3.3.4. Calculation of the Queue Discharging Time. Shockwave theory applies the basic theory of fluid mechanics, simulates the continuity equation of fluid, establishes the continuity equation of traffic flow, compares the change of traffic flow density to the fluctuation of water wave, and abstracts it as traffic wave [27]. In the entrance of the signalized intersection, the change of the vehicle running state will produce multiple traffic waves, which interact with each other and form the forming and discharging of the intersection queue. Among them, when the intersection red light is on, the vehicles in front of the stop line will brake and stop successively. This queuing process can be seen as the forming wave emitted by the phase of the red light is on, and its wave speed is recorded as V_{form}. When the phase changes from red to green, the vehicles in front of the stop line will start one after another, and this process of queuing and

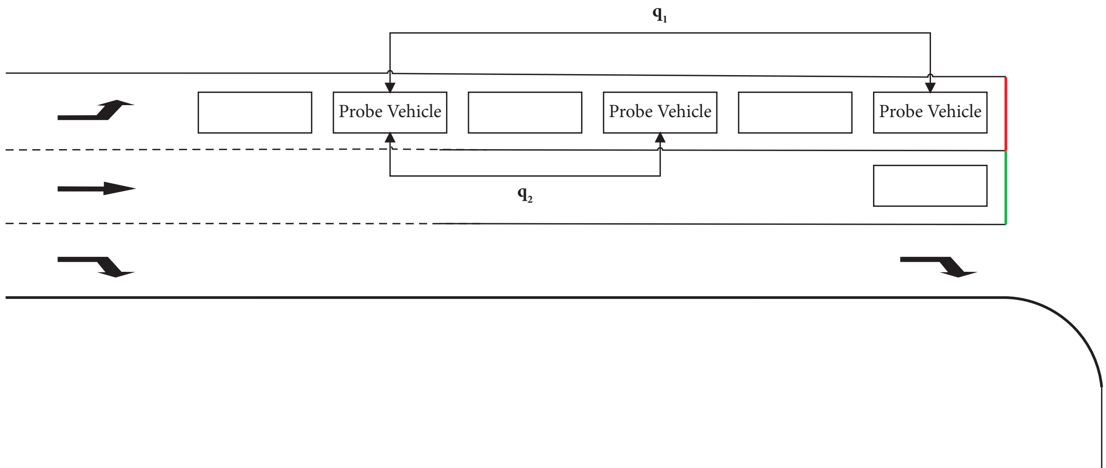


FIGURE 4: Schematic diagram for calculating the arrival rate of the residual red light time and the queue discharging time.

discharging can be seen as the discharging wave emitted by the green light phase, whose wave speed is recorded as V_{dis}.

Considering that the queue length will increase after the green light is on, that is, the queue length will increase after the green light is on until the discharging wave catches up with the forming wave. Based on the shockwave theory, the queue forming (forming wave) and discharging (discharging wave) processes of the queue are reconstructed, and the time consumed from the green light is on until the discharging wave catches up with the forming wave is calculated, that is, the queue discharging time, denoted as t_{dis}. In this way, the new queue length is prepared for the calculation after the green light is on.

The calculation functions of the forming wave velocity V_{form}, discharging wave velocity V_{dis}, and queue discharging time t_{dis} are described as follows.

(1) Calculation of the V_{form}. We define the blocking density as K_j and the density of the arriving vehicle as K_a. K_j refers to the traffic volume is zero. K_j can be obtained based on actual surveys. If there is no survey data, it can be calculated according to the following formula, and according to the shockwave theory, the calculation formula for V_{form} is as follows:

$$\begin{aligned} V_{\text{form}} &= \frac{Q}{K_a - K_j}, \\ Q &= q \times 3600, \\ K_a &= \frac{Q}{\overline{v}}, \\ K_j &= \frac{1000}{\overline{h}_s}, \end{aligned} \tag{8}$$

where q takes q_{one} or q_{mul}, $\overline{v}$ is the space mean speed of the probe vehicles entering the intersection entrance approach corresponding to the same cycle and phase, and Q is the number of vehicles arriving in the same intersection flow direction within one hour, i.e., traffic volume.

(2) Calculation of the V_{dis}. We define the saturation flow rate at each intersection flow direction inlet is Q_{sfr}, which can be determined through simulation experiments or based on the actual situation of the intersection, and the critical density is K_m, at which point the corresponding traffic volume is maximum. When the density exceeds this value, the traffic flow no longer increases but decreases. According to the shockwave theory, the calculation formula for V_{dis} is as follows:

$$\begin{aligned} V_{\text{dis}} &= \frac{Q_{sfr}}{K_m - K_j}, \\ K_m &= \frac{Q_{sfr}}{v_m}, \\ v_m &= \frac{1}{2} v_f, \end{aligned} \tag{9}$$

where K_j is consistent with the aforementioned, v_m is the critical speed, and v_f is the free flow speed, which can be taken as the design speed of the intersection.

(3) Calculation of the Queue Discharging Time. Based on our definition of the queue length at intersections, in reality, when the queue forming wave and the queue discharging wave meet, the corresponding queue length at this moment is the actual queue length. According to the total queue length equal to the total discharging length, we define the queue discharging time is t_{dis}, and then we can obtain the following formula:

$$\begin{aligned} \left(t_{\text{red}} + t_{\text{dis}}\right)\left|V_{\text{form}}\right| &= t_{\text{dis}}\left|V_{\text{dis}}\right|, \\ t_{\text{dis}} &= \frac{\left|V_{\text{form}}\right| t_{\text{red}}}{\left|V_{\text{dis}}\right| - \left|V_{\text{form}}\right|}, \end{aligned} \tag{10}$$

where t_{red} is the red light time.

3.3.5. Estimation of the Queue Length. We divide the estimation of queue length into two situations: one is that only one probe vehicle data in the queue can be obtained during the red light period in a cycle, and the other is that at least two probe vehicle data in the queue can be obtained during the red light period can be obtained during a cycle. Also, we define the estimation of the queue length as follows: the distance between the last-queued probe vehicle and the stop line during the red light period, the length of subsequent vehicles arriving during the residual red light time, and the newly increased queue length within the queue discharging time. L_{true} denotes the true queue length of the intersection, which is equal to maximum queue length. Based on the abovegiven data processing work, the calculation formula for L_{true} can be summarized as follows:

(1) Only one probe vehicle data.

$$\begin{aligned} L_{\text{true}} &= L_{\text{one}} + q_{\text{one}} \times t_{\text{rsd}} \times \overline{h_s} + q_{\text{one}} \times t_{\text{dis}} \times \overline{h_s}, \\ L_{\text{true}} &= L_{\text{one}} + \left(t_{\text{rsd}} + t_{\text{dis}}\right) \times \left(q_{\text{one}} \times \overline{h_s}\right). \end{aligned} \tag{11}$$

(2) At least two probe vehicle data.

$$\begin{aligned} L_{\text{true}} &= L_N + q_{\text{mul}} \times t_{\text{rsd}} \times \overline{h_s} + q_{\text{mul}} \times t_{\text{dis}} \times \overline{h_s}, \\ L_{\text{true}} &= L_N + \left(t_{\text{rsd}} + t_{\text{dis}}\right) \times \left(q_{\text{mul}} \times \overline{h_s}\right). \end{aligned} \tag{12}$$

4. Case Study

After introducing the methods of queue length estimation, this section describes the verification process of the real world intersection (Taoyuan street/Taohua road) in Nanchang, China. Figure 5 shows the processing steps of queue length estimation, and we will verify the methods by the steps.

At present, due to the lack of data accessibility, the probe vehicle data at the studied intersection cannot be obtained. Therefore, the probe vehicle data verified by this example comes from the simulation environment.

Investigate the intersection (Taoyuan street/Taohua road) of Nanchang city in Jiangxi province to get the traffic data in peak hours (7:30–9:30). The Taoyuan street and Taohua road are arterial roads and the intersection plan is shown in Figure 6(a). We choose the western entrance of the intersection as the research object and estimate the queue length of the entrance straight line. Because of a lack of probe vehicle data, the geometry and control scheme of the intersection are investigated and then the VISSIM simulation software is used to simulate the intersection. The simulation diagram of the intersection is shown in Figure 6(b). The entrance road has five lanes: one right turn lane, two straight lanes, and two left turn lanes. Also, the basic saturation flow rates of the straight lane, the left-turn lane, and the right turn lane at the signalized intersection are 1650 pcu/h, 1550 pcu/h, and 1450 pcu/h, respectively. In order to obtain more accurate simulation data, the traffic flow per five minutes is input for simulation. The traffic flow per five minutes of the entrance is shown in Table 1, and the signal phase scheme of the intersection during peak period is shown in Figure 7. The steering ratio is set according to the penetration of each flow direction at the intersection. 50% of the vehicles are taken as probe vehicles, and the data of probe vehicles are collected every 15 s. Partial data of probe vehicles are shown in Table 2.

Since the focus of this paper is on estimating queue length, data filtering and map matching will no longer be explained in detail. Due to the probe vehicle data coming from a simulated environment, there was no data loss.

According to the processing steps for queue length estimation in Figure 5, after data map matching, the stop line position at the intersection is estimated.

4.1. Stop Line Position Estimation. According to the abovegiven steps of calculating the stop line position, the probe vehicle data with a speed less than 5 km/h are screened out, and the approximate position of the stop line at the intersection is determined to be (3330, 3400). The driving speed of each position is obtained by processing the probe vehicle data, as shown in Figure 8(a). The selected probe vehicle data are counted at intervals of 2 meters. The number of probe vehicles is calculated at each interval, the vehicle density is calculated at each interval, and the density over position is obtained, as shown in Figure 8(b). Finally, the stop line position is calculated to be 3398 meters. It should be mentioned that the stop line position estimation has not been the main objective in this paper, as it could be also estimated by an aerial image. Despite such good estimation results, further efforts are needed to validate this method.

4.2. Queue Length Estimation. The simulation duration is 7200 s, the simulation phase starts from the north-south straight line, the phase red light starts at the west entrance straight line is 89 s, the red light time is 109 s, and the queue length of 48 cycles is calculated in total. The calculated queue lengths are compared with the queue lengths output by simulation, as shown in Figure 9. The estimated queue length is agreement to the simulated queue length.

The abovegiven estimated results were calculated under the condition of 50% probe vehicle penetration. In order to verify the accuracy of the method, data of different probe vehicle penetrations were used for estimation and comparative analysis. In general, the larger the penetration of probe vehicle is, the more data will be obtained. The more parking data can be obtained at the red light time of the same cycle. The closer the probe vehicle at the end of the queue approaches the actual queue, the more accurate the calculation results will be. Due to the high penetration rate of the probe vehicle of 50%, but the estimation results are close to the simulated queue length, so probe vehicle with penetration rates of 25% and 10% were selected for comparative analysis.

4.3. Estimation Error Analysis. Mean absolute error E_{MAE}, mean absolute relative error E_{MARE}, and root mean squared error R_{RMSE} were used to evaluate the estimation accuracy of queue length under different penetrations of probe vehicle. The calculation formulas of each error index are as follows:

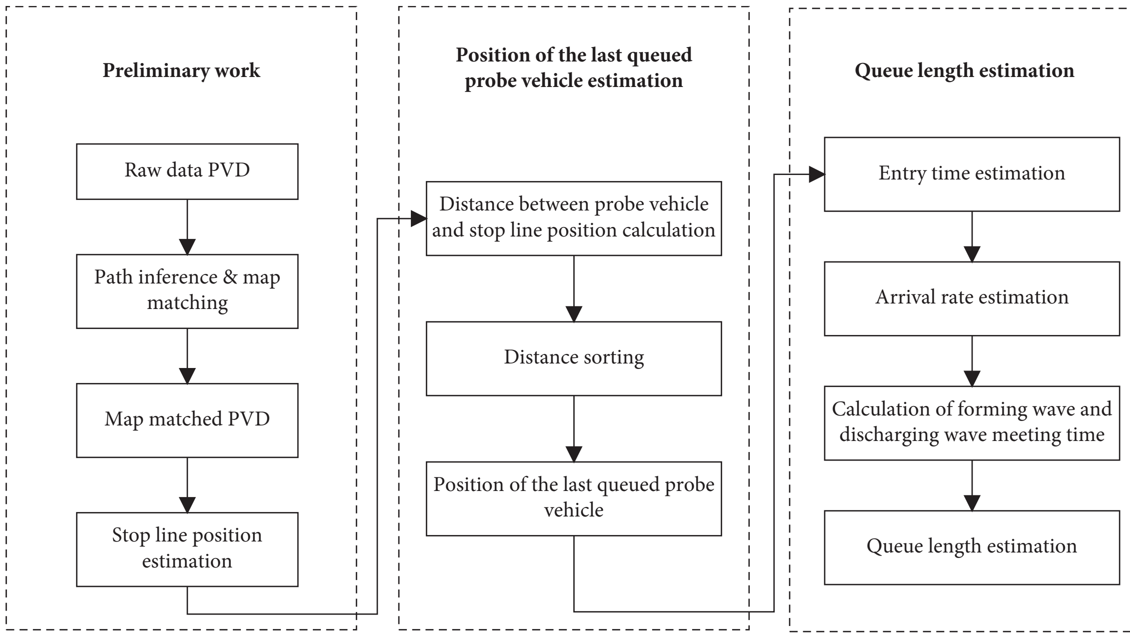
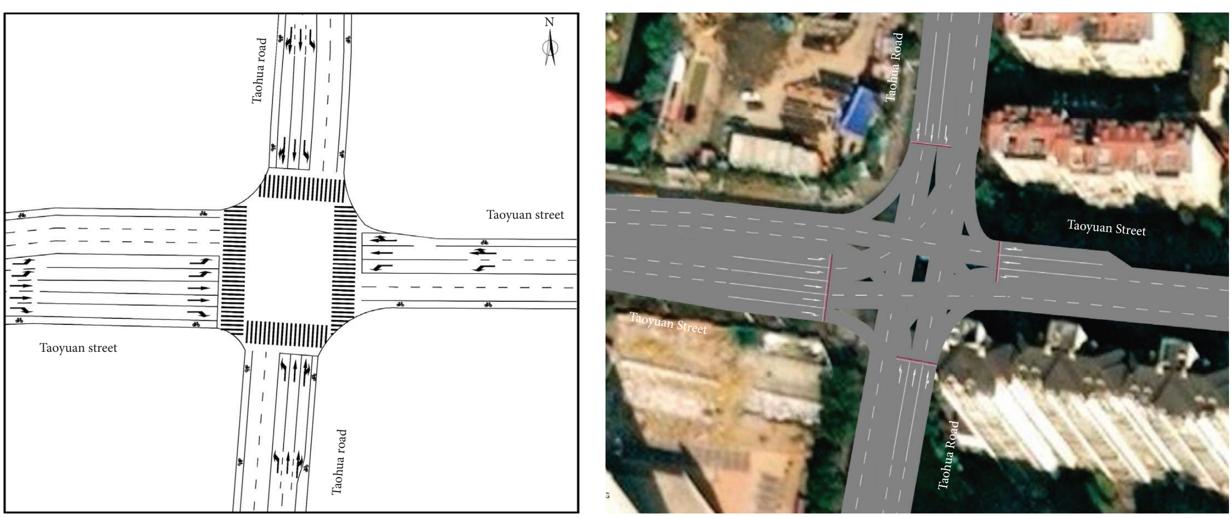


Figure 5: Processing steps for queue length estimation.



(a) (b)

Figure 6: (a) Plan of Taoyuan street/Taohua road intersection and (b) simulation diagram of the intersection.

Table 1: Peak hours traffic volume (pcu).

Direction	North			East			South			West		
Time	Left	Straight	Right	Left	Straight	Right	Left	Straight	Right	Left	Straight	Right
07:30–07:35	6	26	13	19	76	29	6	23	12	5	50	11
07:35–07:40	11	26	12	24	82	33	11	24	14	9	47	9
07:40–07:45	11	29	11	27	68	23	2	25	11	6	55	15
07:45–07:50	11	32	13	21	94	36	5	40	20	14	57	11
07:50–07:55	10	27	11	27	78	29	2	24	11	10	58	15
07:55–08:00	8	35	15	40	86	33	10	32	18	14	46	18
08:00–08:05	15	31	17	31	77	29	6	29	14	18	47	19
08:05–08:10	13	32	17	36	86	37	9	32	17	14	44	18
08:10–08:15	14	33	11	36	73	31	9	25	12	13	40	26

Table 1: Continued.

Direction	North			East			South			West		
Time	Left	Straight	Right	Left	Straight	Right	Left	Straight	Right	Left	Straight	Right
08:15–08:20	7	39	16	30	74	29	9	27	13	18	41	16
08:20–08:25	5	19	8	34	80	30	13	25	13	16	44	16
08:25–08:30	12	33	16	32	84	32	6	33	14	9	45	20
08:30–08:35	8	28	11	33	101	31	5	28	13	10	40	7
08:35–08:40	6	27	12	32	104	40	8	25	14	14	38	13
08:40–08:45	6	34	16	27	93	32	7	21	6	10	45	12
08:45–08:50	13	34	12	25	88	30	4	26	15	11	27	11
08:50–08:55	8	26	12	27	64	28	5	26	12	13	39	15
08:55–09:00	9	27	12	29	73	28	5	28	12	9	34	11
09:00–09:05	3	26	11	13	73	22	6	26	10	9	34	21
09:05–09:10	11	30	12	23	73	29	7	26	12	5	35	10
09:10–09:15	8	32	14	13	57	23	4	28	12	4	29	13
09:15–09:20	12	34	13	14	58	16	5	25	12	7	25	15
09:20–09:25	9	26	11	20	86	25	8	27	12	4	35	14
09:25–09:30	11	40	16	9	52	21	2	23	10	7	23	11

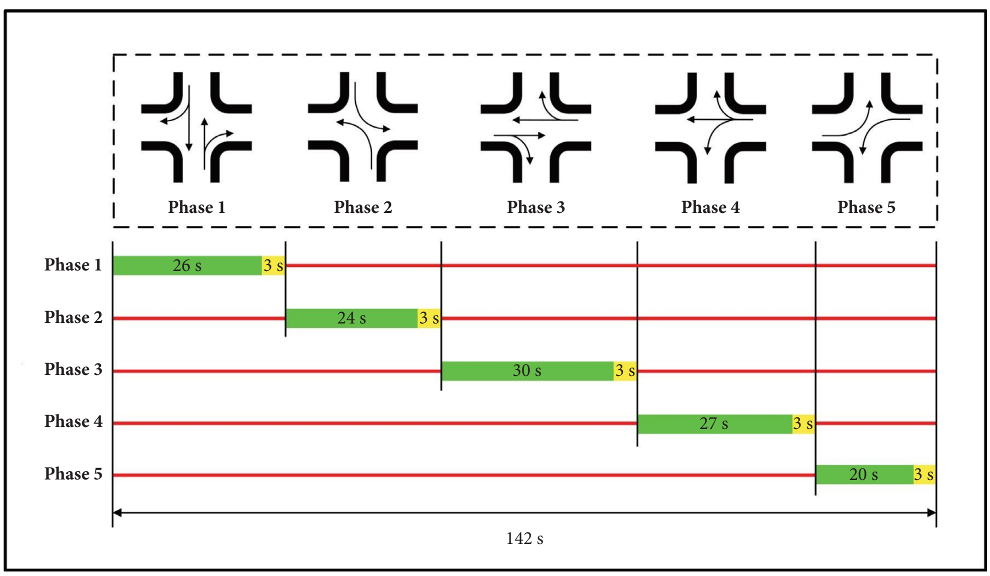


Figure 7: Signal phase scheme.

(13)–(15). The calculation results are shown in Table 3. Also, the comparison of mean absolute relative error under different probe vehicle penetrations is shown in Figure 10.

$$E_{\mathrm{MAE}} = \frac{1}{n}\sum_{j=1}^{n}\left|L_s - L_e\right|, \tag{13}$$

$$E_{\mathrm{MARE}} = \frac{1}{n}\sum_{j=1}^{n}\frac{\left|L_s - L_e\right|}{L_s} \times 100\%, \tag{14}$$

$$R_{\mathrm{RMSE}} = \sqrt{\frac{\sum_{j=1}^{n}\left(L_s - L_e\right)^2}{n}}, \tag{15}$$

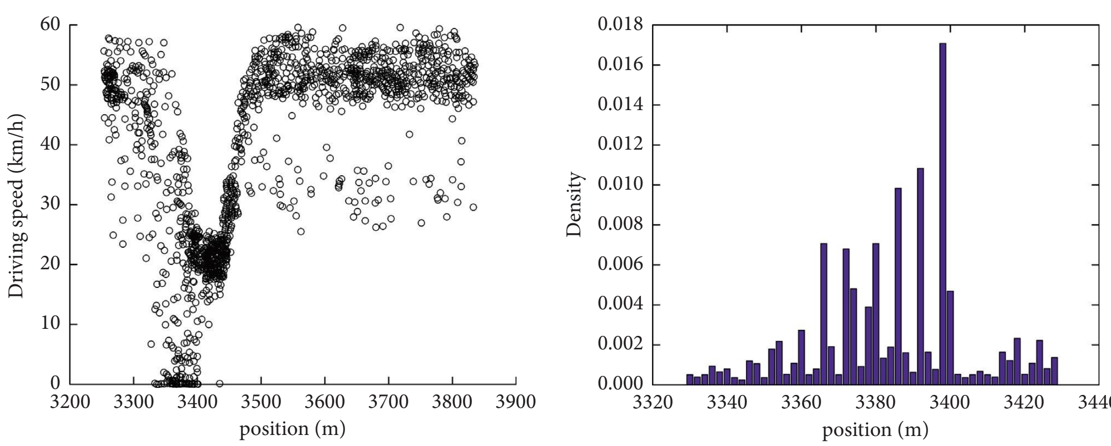
where L_s and L_e are the simulated value and estimated value of queue length, respectively, and 'n' represents the number of cycles.

As can be seen from the error index in Table 3 and Figure 10, the larger the penetration of probe vehicle is, the more probe vehicle data are obtained, and the closer the estimated queue length is to the actual queue length. When the penetration of probe vehicles is 50%, the mean absolute relative error of the estimated results is only 11.27%, which is relatively accurate.

In order to verify the validity of the model, the queue length estimation method based on data fusion proposed by Li et al. [22] was selected for comparison in this study, and the test environment was the same as the case mentioned above. Table 4 and Figure 11 both show the comparison results of the queue length estimates by the two methods and calculate the mean absolute relative error of the two methods when the penetration of probe vehicle is 10%, 25%, and 50%. Significantly, the proposed method in this paper shows higher estimation accuracy at all penetration rates. However,

Table 2: Part of probe vehicle data from simulation environment.

t	x	y	ID	s	t	x	y	ID	s
15	3105.953	721.0338	5	57.85	60	3330.728	696.0665	29	37.67
15	3419.282	890.2507	7	53.32	60	3418.837	727.8462	27	0
15	3630.993	709.7021	1	46.55	60	3500.951	706.6817	25	37.49
15	3424.936	663.3612	3	24.69	60	3427.654	692.0638	21	0
30	3341.171	701.4261	5	56.95	60	3238.039	707.8425	35	49.81
30	3214.133	710.5837	11	54.48	60	3103.432	721.3628	45	50.31
30	3805.026	705.3441	19	52.77	60	3186.458	710.5383	41	56.07
30	3735.195	709.338	15	48.12	60	3178.339	708.0659	39	46.23
30	3716.595	709.403	13	48.58	60	3067.049	740.0481	7	49.66
30	3625.363	709.7217	9	52.62	60	3428.083	875.4785	9	57.72
30	3555.562	699.074	3	57.8	60	3721.15	707.0617	43	54.27
30	3405.207	515.4935	21	49.07	60	3665.217	705.8325	37	58.32
30	3415.091	575.6755	17	55.93	60	3820.726	709.0392	47	55.84
30	3420.034	741.6836	7	19.48	60	3661.765	698.5514	17	55.1
30	3469.401	710.6321	1	20.82	60	3399.967	701.2668	5	0
45	3548.156	709.9923	15	38.65	60	3401.282	708.1369	11	0
45	3526.583	710.1927	13	38.39	60	3385.507	694.3132	23	19.99
45	3458.068	710.6714	9	22.73	60	3423.279	743.6247	31	19.1
45	3138.321	713.2791	29	51.13	60	3421.195	762.4493	33	22.51
45	3228.113	701.9357	23	51.3	60	3469.235	710.6326	19	9.26
45	3275.169	716.3737	7	50.74	60	3443.581	710.7218	13	0
45	3422.765	893.8466	31	53.88	60	3444.094	714.4282	15	1.71
45	3422.71	811.8325	27	47.3	75	3422.581	733.6297	31	0
45	3434.493	795.8841	1	47.77	75	3418.837	727.8462	27	0
45	3670.9	705.8126	25	42.58	75	3517.332	706.5346	43	39.75
45	3601.326	709.8057	19	33.16	75	3515.42	710.2941	37	20.57
45	3443.994	699.8501	17	26.86	75	3438.913	706.7292	25	0
45	3787.808	697.9312	3	54.18	75	3426.677	714.5803	19	23.19
45	3399.967	701.2668	5	0	75	3420.865	627.3393	55	38.24
45	3400.904	708.143	11	2.92	75	3427.654	692.0638	21	0
45	3425.633	672.1227	21	19.68	75	3110.722	723.9411	59	56.65



(a) (b)

Figure 8: (a) Driving speed of each position and (b) histogram of slow driving speed data points with speed ≤5 km/h, sample size = 2407.

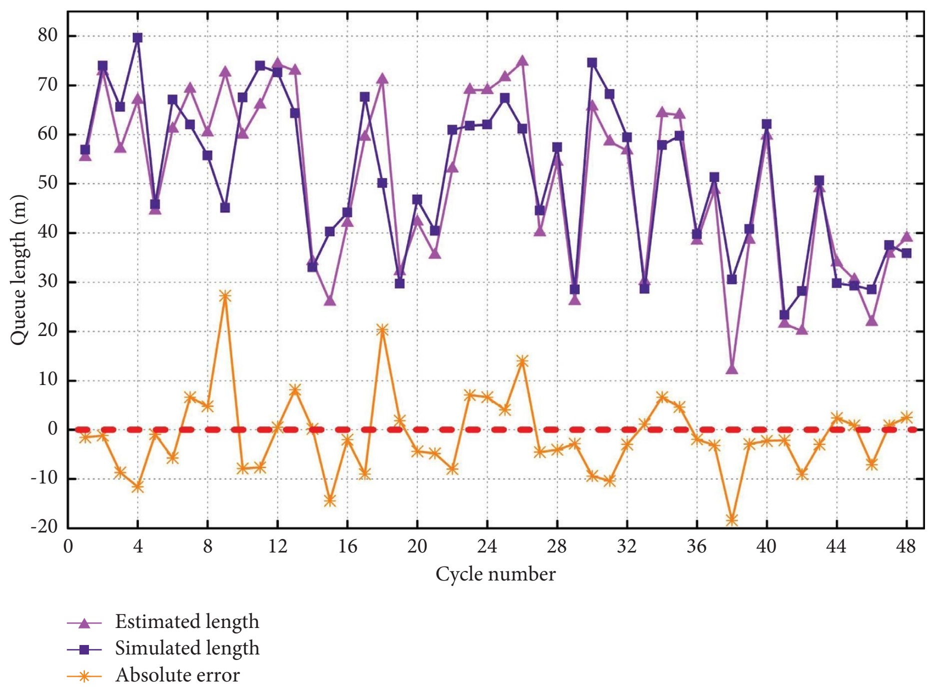


Figure 9: Estimated queue length and simulated queue length comparison chart.

Table 3: Error indexes under different probe vehicle penetration.

Error indexes	Probe vehicle penetration		
	50 (%)	25 (%)	10 (%)
E_{MAE}	5.56	13.32	18.97
E_{MARE} (%)	11.27	27.77	39.12
R_{RMSE}	6.94	15.94	22.53

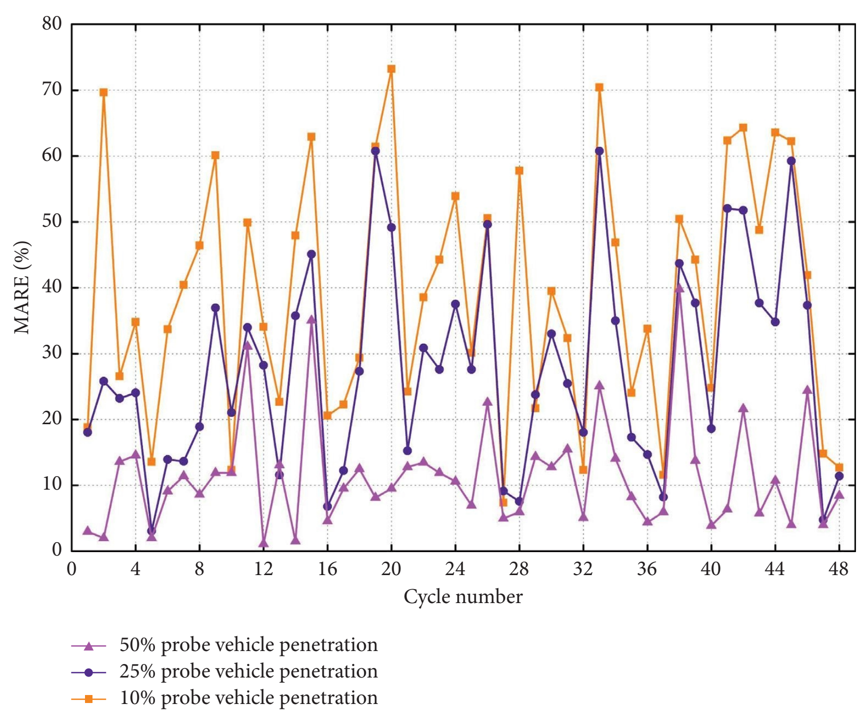


Figure 10: Comparison of mean absolute relative error under different probe vehicle penetrations.

Table 4: Comparison of simulation results.

Parameters	Probe vehicle penetration		
	50 (%)	25 (%)	10 (%)
MAE (s/vehicle)			
Method proposed in this study	5.56	13.32	18.97
Method of literature [22]	7.23	17.79	24.36
MARE (%)			
Method proposed in this study	11.27	27.77	39.12
Method of literature [22]	16.53	35.81	58.73
RMSE (s/vehicle)			
Method proposed in this study	6.94	15.94	22.53
Method of literature [22]	9.65	19.96	28.36

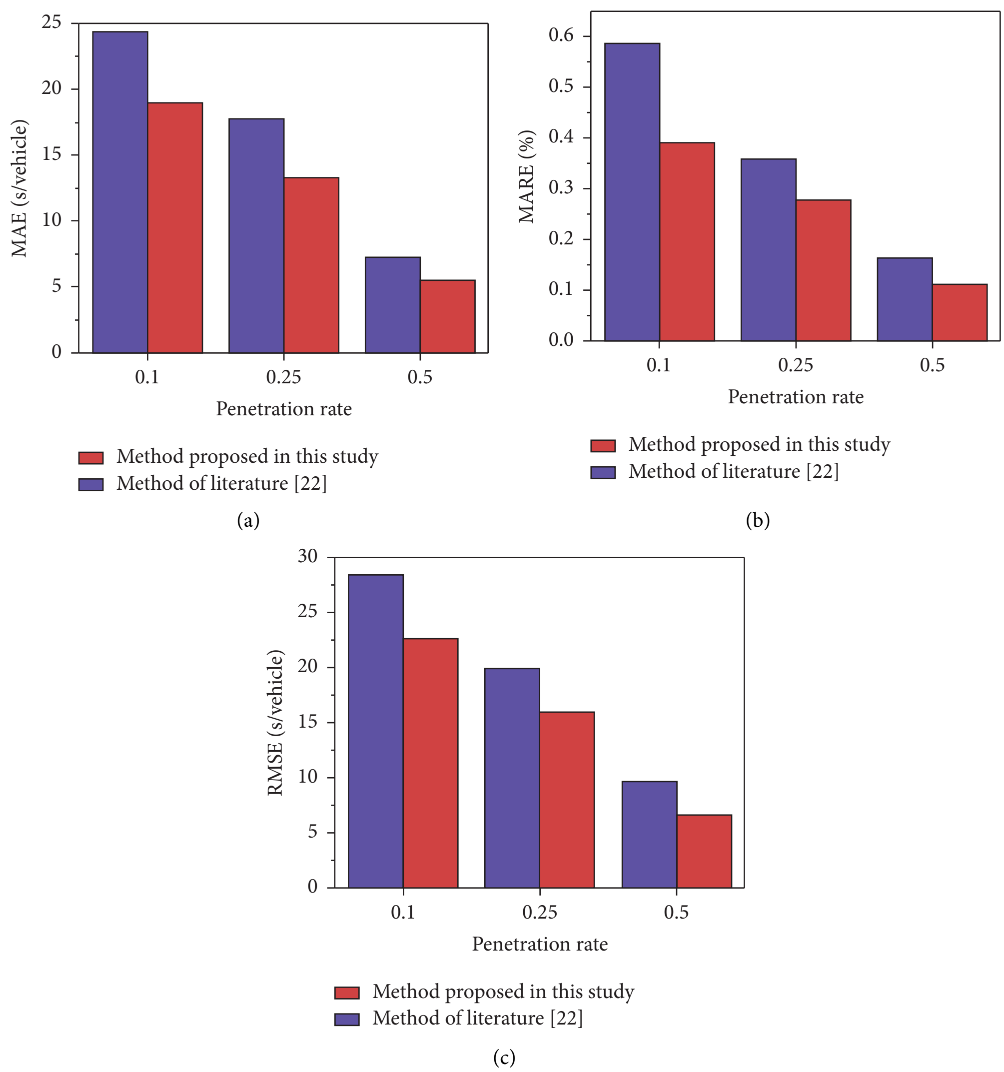


Figure 11: Comparison of different error indexes between the proposed method and method of literature [22] under different penetration rates: (a) MAE (s/vehicle), (b) MARE (%), and (c) RMSE (s/vehicle).

it is undeniable that when the penetration of the method proposed in this paper is 50%, the mean absolute relative error is controlled within 15%, indicating that the method proposed in this paper has relatively high requirements on the penetration of probe vehicle.

5. Conclusions

This paper uses probe vehicle data to establish a queue length estimation model for signalized intersections based on shockwave theory, alleviating the severe dependence of traditional estimation models on fixed detectors. At the same time, a real-time calculation method is proposed for measuring arrival rate, effectively solving the limitations of queue estimation errors caused by previous research assuming vehicle arrival modes. This paper first calculates the position of the stop line, and then based on the data of probe vehicle in the queue during the red light period, the average vehicle arrival rate is calculated. Also, the shockwave theory is used to calculate the speed of the queue forming wave and the queue discharging wave, as well as the queue dissipation time. Finally, the queue length is divided into three parts for calculation, including the distance between the last queued probe vehicle and the stop line during the red light period, the length of the subsequent vehicles arriving during the residual red light time, and the newly increased queue length within the queue discharging time. These three parts are added together is the estimated length of the intersection queue. The method ensures the applicability and accuracy of the queue length estimation model. This paper also uses the VISSIM software to verify the accuracy of the queue length estimation model. The results show that when the penetration of probe vehicle is 50%, 25%, and 10%, and their corresponding mean absolute relative error are 11.27%, 27.77%, and 39.12%, respectively. From the error indicators, it can be seen that the accuracy of this estimation model can achieve ideal result when the penetration of probe vehicle is high. The queue length estimation model established in this paper can provide an effective theoretical basis for estimating the queue length of signalized intersections and optimizing subsequent signal control methods in the future.

There are still some limitations to the current work that will need to be addressed in the future. First, the penetration of probe vehicle is deduced by the parking position of the probe vehicle. Therefore, it cannot handle unsignalized intersections or right turn movements. In engineering practice, the penetration rate of adjacent intersection probe vehicles can be used instead of its calculation. Second, the method requires signal timing information as input. Although some existing methods can achieve timing parameter estimation under high penetration, the estimation of signal timing parameters under low penetration environment is still a research difficulty. In the future, in order to eliminate this limitation and expand the applicability of the proposed method in this paper, it will be studied under the condition that the timing information is unknown and the penetration of probe vehicle is low. Third, this method does not reproduce some random characteristics of traffic flow and does not consider the applicability of this method at different types of intersections. Future studies will simulate these effects.

Data Availability

The data that support the finding of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under the research project Grant no. 51965021 and Science and Technology Research Project of Jiangxi Education Department under the research project no. GJJ160476.

References

[1] J. Yin, P. Chen, K. Tang, and J. Sun, "Queue intensity adaptive signal control for isolated intersection based on vehicle trajectory data," *Journal of Advanced Transportation*, vol. 2021, no. 6, Article ID 8838922, 16 pages, 2021.

[2] L. Wei, J. Li, L. Xu, L. Gao, and J. Yang, "Queue length estimation for signalized intersections under partially connected vehicle environment," *Journal of Advanced Transportation*, vol. 2022, Article ID 956872, 11 pages, 2022.

[3] X. Chen, L. Li, and Z. Li, "Phase diagram analysis based on a temporal-spatial queueing model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1705–1716, 2012.

[4] S. Lee, S. C. Wong, and Y. C. Li, "Real-time estimation of lane-based queue lengths at isolated signalized junctions," *Transportation Research Part C: Emerging Technologies*, vol. 56, pp. 1–17, 2015.

[5] G. Vigos, M. Papageorgiou, and Y. Wang, "Real-time estimation of vehicle-count within signalized links," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 1, pp. 18–35, 2008.

[6] Z. Wang, Q. Cai, B. Wu, L. Zheng, and Y. Wang, "Shockwave-based queue estimation approach for undersaturated and oversaturated signalized intersections using multi-source detection data," *Journal of Intelligent Transportation Systems*, vol. 21, no. 3, pp. 167–178, 2017.

[7] M. Horvath and T. Tettamanti, "Real-time queue length estimation applying shockwave theory at urban signalized intersections," *Periodica Polytechnica: Civil Engineering*, vol. 65, no. 4, pp. 1153–1161, 2021.

[8] B. Li, W. Cheng, and L. Li, "Real-time prediction of lane-based queue lengths for signalized intersections," *Journal of Advanced Transportation*, vol. 2018, Article ID 5020518, 18 pages, 2018.

[9] P. Lin, Y. Lei, K. Yao, and Y. Gu, "Estimation of intersection queue length and signal adaptive control in partial network environment," *Journal of South China University of Technology*, vol. 45, no. 11, pp. 1–9, 2017.

[10] A. M. Tahsin Emtenan and C. M. Day, "Correlation analysis for exploring the relationship between probe vehicle data and event-based traffic signal performance measures," *Transportation Research Record*, vol. 2676, no. 8, pp. 587–600, 2022.

[11] F. Lian, B. Chen, K. Zhang, L. Miao, J. Wu, and S. Luan, "Adaptive traffic signal control algorithms based on probe vehicle data," *Journal of Intelligent Transportation Systems*, vol. 25, no. 1, pp. 41–57, 2020.

[12] H. Liu, W. Liang, L. Rai, K. Teng, and S. Wang, "A real-time queue length estimation method based on probe vehicles in CV environment," *IEEE Access*, vol. 7, pp. 20825–20839, 2019.

[13] Y. Zhao, S. Shen, and H. X. Liu, "A hidden Markov model for the estimation of correlated queues in probe vehicle environments," *Transportation Research Part C: Emerging Technologies*, vol. 128, Article ID 103128, 2021.

[14] C. Tan, L. Liu, H. Wu, Y. Cao, and K. Tang, "Fuzing license plate recognition data and vehicle trajectory data for

lane-based queue length estimation at signalized intersections," *Journal of Intelligent Transportation Systems*, vol. 24, no. 5, pp. 449–466, 2020.

[15] Y. Mei, W. Gu, E. C. S. Chung, F. Li, and K. Tang, "A Bayesian approach for estimating vehicle queue lengths at signalized intersections using probe vehicle data," *Transportation Research Part C: Emerging Technologies*, vol. 109, pp. 233–249, 2019.

[16] C. Tan, J. Yao, K. Tang, and J. Sun, "Cycle-based queue length estimation for signalized intersections using sparse vehicle trajectory data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 1, pp. 91–106, 2021.

[17] Y. Zhao, W. Wong, J. Zheng, and H. X. Liu, "Maximum likelihood estimation of probe vehicle penetration rates and queue length distributions from probe vehicle data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7628–7636, 2022.

[18] G. Comert, "Queue length estimation from probe vehicles at isolated intersections: estimators for primary parameters," *European Journal of Operational Research*, vol. 252, no. 2, pp. 502–521, 2016.

[19] Y. Zhao, J. Zheng, W. Wong, X. Wang, Y. Meng, and H. X. Liu, "Estimation of queue lengths, probe vehicle penetration rates, and traffic volumes at signalized intersections using probe vehicle trajectories," *Transportation Research Record*, vol. 2673, no. 11, pp. 660–670, 2019.

[20] Y. Zhao, J. Zheng, W. Wong, X. Wang, Y. Meng, and H. X. Liu, "Various methods for queue length and traffic volume estimation using probe vehicle trajectories," *Transportation Research Part C: Emerging Technologies*, vol. 107, pp. 70–91, 2019.

[21] M. Rostami Shahrbabaki, A. A. Safavi, M. Papageorgiou, and I. Papamichail, "A data fusion approach for real-time traffic state estimation in urban signalized links," *Transportation Research Part C: Emerging Technologies*, vol. 92, pp. 525–548, 2018.

[22] J. Li, K. Zhou, S. Shladover, and A. Skabardonis, "Estimating queue length under connected vehicle technology: using probe vehicle, loop detector, fused data," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2366, no. 1, pp. 17–22, 2013.

[23] J. Zhao, V. Knoop, and M. Wang, "Two-dimensional vehicular movement modelling at intersections based on optimal control," *Transportation Research Part B: Methodological*, vol. 138, pp. 1–22, 2020.

[24] J. Zhao, V. Knoop, and M. Wang, "Microscopic traffic modeling inside intersections Interactions between drivers," *Transportation Science*, vol. 57, no. 1, pp. 135–155, 2023.

[25] S. Axer, F. Pascucci, and B. Friedrich, "Estimation of traffic signal timing data and total delay for urban intersections based on low frequency probe vehicle data," in *Proceedings of the 6th mobility TUM*, Technical University of Munich, Munich, Germany, April 2015.

[26] Z. Liu, H. Wang, and H. Guo, "Queue length estimation of signalized intersections based on probe vehicle data," *Road Traffic & Safety*, vol. 16, no. 4, pp. 7–11, 2016.

[27] C. Tan, J. Yao, Y. Cao, and K. Tang, "Cycle-based queue length estimation based on connected vehicle trajectory data," *China Journal of Highway and Transport*, vol. 34, no. 7, pp. 140–151, 2021.