# Modeling Interstation Travel Speed of Hybrid Bus Rapid Transit within A Bayesian Framework

**Ji-hua Hu, [1*], Jia-xian Liang[2]**

1 Public Labs Center, Sun Yat-sen University, Guangzhou 510006, China
2 Research Centre of Intelligent Transportation System, School of Engineering, Sun Yat-sen University, Guangzhou 510006, China

*Corresponding Author: Ji-hua HU, Ph.D., Public Labs Center, Sun Yat-sen University, Guangzhou 510006, China

## Abstract

Interstation travel speed is an important indicator of the running state of hybrid Bus Rapid Transit and passenger experience. Due to the influence of road traffic, traffic lights and other factors, the interstation travel speeds are often some kind of multi-peak and it is difficult to use a single distribution to model them. In this paper, a Gaussian mixture model charactizing the interstation travel speed of hybrid BRT under a Bayesian framework is established. The parameters of the model are inferred using the Reversible-Jump Markov Chain Monte Carlo approach (RJMCMC), including the number of model components and the weight, mean and variance of each component. Then the model is applied to Guangzhou BRT, a kind of hybrid BRT. From the results, it can be observed that the model can very effectively describe the heterogeneous speed data among different inter-stations, and provide richer information usually not available from the traditional models, and the model also produces an excellent fit to each multimodal speed distribution curve of the inter-stations. The causes of different speed distribution can be identified through investigating the Internet map of GBRT, they are big road traffic and long traffic lights respectively, which always contribute to a main road crossing. So, the BRT lane should be elevated through the main road to decrease the complexity of the running state.

**Keywords:** *bayesian, hybrid BRT, RJMCMC, interstation travel speed, gaussian mixture model*

## 1. Introduction

The travel speed is the main advantage of the Bus Rapid Transit (BRT) over ordinary buses, especially in the downtown area, but the speed of the hybrid BRT system may have some difference. Hybrid BRT is an improved form of BRT, which has facilities such as bus lanes, new public transport stations and intelligent monitoring and management systems, but is no longer isolated from ordinary bus lines (1). Ordinary buses are free to enter and exit the BRT system, passengers can transfer to the buses in station, private cars in the off-peak hours can also runs into some part of the BRT lanes, which enhances the efficiency of the city's overall bus system. But these changes will also have great impact on the interstation travel speed of hybrid BRT, and the distribution of the travel speed will be more complexed. The interstation travel speed is an important indicator of the running state and efficiency of the BRT (2). It is also the main input data of BRT simulation. Therefore, it is necessary to model the speed and analyze the running state of hybrid BRT consequently.

The travel speed was usually modeled with a normal or lognormal distribution, or a Gaussian mixture distribution with a fixed K value (3). However, the travel speed of hybrid BRT is affected by more factors, including road traffic, traffic lights, and other factors. Therefore, the distribution of travel speed varies largely between stations, often showing multi-peak phenomenon; and the influence factors of different interstation are not the same, accordingly the distributions must be different, so it is difficult to use the above model to describe them. Therefore, constructing a mathematical model describing the multi-peak travel speed of interstation of hybrid BRT is very necessary.

In addition to the single mathematical model mentioned above, other scholars had proposed different types of distribution to model the non-normality of the sampled travel speed. Dey and others showed that the speed data could follow a unimodal or a bimodal curve depending on the variation of speed for different categories of vehicles (4). A similar approach taken by Ko and Guendsler characterized the congestion based on the two mixtures of speed distribution. They assumed that the speed distribution over a given time had a form of mixed distribution, one for congested and the other for uncongested flow (5). Some researchers had adopted Gaussian mixture models to cluster transport-related observations into groups and discovered the existence of multi-regime states. Park et al. employed the model to account for heterogeneity in speeds based on actual data collected from an existing highway segment. Jun utilized the model to characterize the severity and variability of congestion on certain interstate roadway systems (6). Corey et al. (2011) used a Gaussian mixture model to identify the sensitivity errors of inductive loop detectors (7). More recently, Lao et al. also utilized a Gaussian mixture model to estimate traffic speeds and classify vehicle volumes using measurements from single-loop detectors (8). However, Park et al. did not fully incorporate the parameter (i.e. the number of components) into the Bayesian framework to infer other unknown parameters. The Bayesian framework assumes that the number of Gaussian mixture models and the parameters of each component to be random variables and to have their own certain prior distribution (9). And the posterior distribution of these parameters needs to be estimated. The commonly used parameter estimation method was the RJMCMC method (10), such as Lee, which used this method to determine the component number of the Gaussian mixture model of the subway passenger's travel time data and the parameters of each component, and then determined the whole Gaussian mixture model (11).

bus station-reporting data refers to when the bus stops at the station and open the door, it reports the station to the passengers and reminds the passengers to get off; when the bus doors closed, it reminds other passengers the next station, so they can be ready for getting off. Bus station-reporting data is generated during the process. From the data, the travel time between neighbored stations can be extracted, then the interstation travel speed can be computed. In this paper, the interstation travel speed is calculated using the station-reporting data of a hybrid BRT. Then, within a Bayesian framework, the Gaussian mixture model is used to model the interstation travel speed, and the number of model components is estimated using the RJMCMC method, so do the parameters of each component. Finally, combined with the impacts of road traffic, traffic lights, and other factors, the causes of the formation of the models can be determined. The results can be used as input to the BRT simulation system, so that the simulation is closer to the real system. And the main factors that have impact on the running state of hybrid BRT can be determined, this provides a basis for BRT optimization.

## 2. Modelling the interstation travel speed

Due to the influence of traffic lights, road traffic and other factors, the interstation speed distribution of hybrid BRT usually presents a multimodal phenomenon. The Gaussian mixture model can describe non-single normal samples, capture the heterogeneity of multiphase data more completely, and effectively fit the multimodal distribution. Therefore, this paper adopts a Gaussian mixture model to describe the interstation travel speed of hybrid BRT, and assumes the unknown parameters as random variables, and describes the speed model under the Bayesian framework.

*2.1 The Gaussian mixture model*

A Gaussian mixture model is a simple weighted sum of K Gaussian densities, each representing a component that corresponds to a speed distribution in the proposed model. For any adjacent stations, the Gaussian mixture model of the travel speed is defined as Eq. (1)

$$p\left(y_i \mid \boldsymbol{\mu}, \boldsymbol{\sigma}^2\right) = \sum_{j}^{K} w_j \cdot N\left(\mu_j, \sigma_j^2\right) \tag{1}$$

where $y_i$ is a target random variable representing the $i$ th travel speed extracted from bus station-reporting data, $p(|)$ is a composite probability density of travel speeds, $K$ is the component number in the model, $\boldsymbol{\mu} = \left(\mu_1, \cdots, \mu_K\right)'$ is a vector of speed mean of each component, and$\boldsymbol{\sigma}^2 = \left(\sigma_1^2, \cdots, \sigma_K^2\right)'$is a vector of speed variances of the components, $w_j$ is a weight of $j$ th component, $\sum_{j}^{K} w_j = 1$ , and $N(\cdot,\cdot)$ is a one-dimensional Gaussian density function.

In the above model, the density distribution which the observed value $y_i$ belongs to is unknown. Define the latent variable $z_i$ as the density distribution label to which the observed value $y_i$ belongs, and the probability that the latent variable $z_i$ belongs to the $j$ th density distribution is as Eq. (2)

$$p\left(z_i = j\right) = w_j \ , \ j = 1, \cdots, K \tag{2}$$

Then a complete Gaussian mixture model of the interstation travel speed is parameterized by the number of the component and the mean, variance and weight for each component. However, the component number of the models and the parameters of each component are often variable among inter-stations. These differences are caused by the factors that have impact on the speed (such as road traffic lights). To model the travel speed better, this paper introduces the Bayesian framework, and uses the concept of hyper parameters to express the potential difference characteristics of the inter-station travel speed.

*2.2 Travel speed model in Bayesian framework*

Under the Bayesian framework, the unknown parameters in the Gaussian mixture model are assumed to be random variables and to have their own certain prior distribution, and the posterior distribution of these parameters needs to be estimated. Regarding the conjugate priors of the parameters, the component number in the model was assumed to follow a Poisson distribution, the weight parameter was assumed to follow a symmetric Dirichlet distribution, the mean of travel speed for each density distribution was assumed to be normally distributed, the variance in travel speeds for each density distribution was assumed to follow an inverse Gamma distribution. All these prior distribution functions also involved additional parameters, which are referred to 'hyperparameter'. Eqs. (3)–(6) represent functional expressions for the chosen priors.

$$p(K=k)=\frac{\lambda^k \exp(-\lambda)}{k!} \tag{3}$$

$$\mathbf{w} \sim D(\delta,\cdots,\delta) \tag{4}$$

$$\mu_j \sim N(\xi,\kappa^{-1}),\ j=1,\cdots,K \tag{5}$$

$$\sigma_j \sim \Gamma(\alpha,\beta),\ j=1,\cdots,K \tag{6}$$

Where $\lambda$ is a hyperparameter to characterize the prior distribution of the number of the distributions, $D$ represents the Dirichlet distribution with a hyperparameter $\delta$, $\xi$ and $\kappa^{-1}$ are the prior mean and variance of the prior distribution of the mean of the density distribution respectively, and $\Gamma(\cdot,\cdot)$ represents a Gamma distribution, and hyperparameters $\alpha$ and $\beta$ are shape and rate (i.e. 1/scale) parameters respectively.

The hyperparameters $\kappa$ and $\beta$ are defined as random parameters, as Eqs (7)-(8), otherwise if they are set to fixed values, the estimated number of density distributions will change with them, this will result in different values of $\kappa$ and $\beta$ for the travel speeds of different adjacent stations.

$$\kappa \sim \Gamma(e,f) \tag{7}$$

$$\beta \sim \Gamma(g,h) \tag{8}$$

where the second-level hyperparameters $e$, $g$ and $f$, represent the shape and rate of a Gamma density, respectively.

According to the Bayesian theorem, the posterior distribution of all random variables can be defined as

$$\begin{aligned} & p(\mathbf{w},\boldsymbol{\mu},\boldsymbol{\sigma}^2,K,\mathbf{z},\kappa,\beta \mid \mathbf{y}) \propto p(\mathbf{y}\mid \mathbf{w},\boldsymbol{\mu},\boldsymbol{\sigma}^2,K,\mathbf{z},\kappa,\beta)\,p(\mathbf{w},\boldsymbol{\mu},\boldsymbol{\sigma}^2,K,\mathbf{z},\kappa,\beta) \\ & = p(\mathbf{y}\mid\boldsymbol{\mu},\boldsymbol{\sigma}^2,\mathbf{z})\,p(\boldsymbol{\mu}\mid\kappa,K)\,p(\boldsymbol{\sigma}^2\mid\beta,K)\,p(\mathbf{z}\mid\mathbf{w},K)\,p(\mathbf{w}\mid K)\,p(K)\,p(\kappa)\,p(\beta) \\ & = p(\mathbf{y}\mid\boldsymbol{\mu},\boldsymbol{\sigma}^2,\mathbf{z})\,p(\boldsymbol{\mu}\mid\xi,\kappa,K)\,p(\boldsymbol{\sigma}^2\mid\alpha,\beta,K)\,p(\mathbf{z}\mid\mathbf{w},K)\,p(\mathbf{w}\mid\delta,K)\,p(K\mid\lambda)\,p(\kappa\mid e,f)\,p(\beta\mid g,h) \end{aligned} \tag{9}$$

Where $\mathbf{y}=(y_1,\cdots,y_N)'$ and $\mathbf{z}=(z_1,\cdots,z_N)'$ are observations and latent variable tags, Respectively, $N$ is the number of the observations. According to the conditions of the parameters independent assumptions, the posterior distribution can be simplified into one-level Bayesian model(second row of Eq.(9)), and further, can be simplified into the two-level Bayesian model in the third row of Eq.(9). In the next step, we need to solve the joint probability distribution of all the random variables in Eq. (9).

## 3. parameter estimation

After the interstation travel speed of hybrid BRT modelled, the component number and the parameters of each component need to be inferred, this inference process is also a process of solving the speed model. Because the model is a multimodal probability density distribution model, we use the outer Gibbs sampling and RJMCMC (Reversible-jump Markov chain Monte Carlo) method to solve it.

*3.1 Gibbs sampler procedure*

To sample from a joint distribution, a Gibbs sampler repeatedly takes a draw of each random variable in turn, with all other variables fixed at the previous draws. It should be noted that the dimension of the first three parameter subsets vary according to the last parameter K. This was the main reason a reversible jump MCMC sampler was employed, the details of which will be described in the next subsection. The outer Gibbs sampler for solving the proposed problem can be summarized as follows.

1) to update the distribution shares $\mathbf{w}$ .

2) to update the mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma}^2$ of each distribution.

3) to update the latent allocation variable $\mathbf{z}$ for each travel speed.

4) to update the hyperparameters $\kappa$ and $\beta$ .

5) to update the number of speed components $K$ .

According to the conjugacy, all the conditional distribution in the first four steps can be achieved by simplifying the Eq. (9) with all other parameters fixed at constants.

The full conditional distribution for $\mathbf{w}$ parameter was a Dirichlet distribution with hyperparameters increased by the estimated number of observations (i.e. the number of travel speed observations) for each distribution.

$$\mathbf{w} \mid all\ other\ parameters \sim D\left(\delta+n_1,\cdots,\delta+n_K\right) \tag{10}$$

The full conditional distribution of the mean of the speed density distribution is given in Eq. (11). To avoid the problem of label exchange between the distributions, we follow the order principle ( $\mu_1<\mu_2<\cdots<\mu_K$ ). during the procedure of speed mean sampling, once the order principle is violated, the sampling results are rejected.

$$\mu_j \mid all\ other\ parameters \sim N\left\{\frac{\sigma_j^{-2}\sum_{\{i|z_i=j\}} y_i+\kappa\xi}{\sigma_j^{-2}n_j+\kappa},\left(\sigma_j^{-2}n_j+\kappa\right)^{-1}\right\} \tag{11}$$

The full conditional distribution of speed variance of the density distribution is given in Eq. (12).

$$\sigma_j^{-2} \sim \Gamma\left(\alpha+\frac{1}{2}n_j,\beta+\frac{1}{2}\sum_{\{i|z_i=j\}}\left(y_i-\mu_j\right)^2\right) \tag{12}$$

The latent variable of the observed speed $y_i$ is sampled using Eq. (13).The estimated number of observed values is updated accordingly.

$$p\left(z_i=j \mid all\ other\ parameters\right) \propto \frac{w_j}{\sigma_j}\exp\left\{-\frac{\left(y_i-\mu_j\right)^2}{2\sigma_j^2}\right\} \tag{13}$$

The full conditional probability distribution of the hyperparameters $\kappa$ and $\beta$ are given in Eq. (14)-(15).

$$\kappa \mid all\ other\ parameters \sim \Gamma\left\{e+\frac{1}{2}K, f+\frac{1}{2}\sum_j\left(\mu_j-\xi\right)^2\right\} \tag{14}$$

$$\beta \mid all\ other\ parameters \sim \Gamma\left(g+K\alpha, h+\sum_j\sigma_j^{-2}\right) \tag{15}$$

In the fifth step, the RJMCMC method is used to realize the reversible jump of each parameter in different dimensions based on the four different move types of split, combine, birth and death, and to infer the unknown number of the components of the model.

*3.2 Estimation of the number of speed component*

Based on the dimension equilibrium condition, the RJMCMC solves the mixture model under unknown dimension by exploring in different dimension spaces. The core of the RJMCMC algorithm is the accepted probability of a proposed move. Green (1995) suggested the following criterion [Eq. (16)] to move from the current state $x=\left(K,w_j,\mathbf{w},\boldsymbol{\mu},\boldsymbol{\sigma}^2,\mathbf{z},\kappa,\beta\right)$ to a higher dimensional state $x'=\left(K+1,\mathbf{w},\boldsymbol{\mu},\boldsymbol{\sigma}^2,\mathbf{z},\kappa,\beta\right)$, as Eq. (16)-(18).

$$a_m\left(x,x'\right)=\min\{1,A\} \tag{16}$$

$$A=\left(\text{likelihood ratio}\right)\times\left(\text{prior ratio}\right)\times\left(\text{proposal ratio}\right)\times\left(\text{Jacobian}\right) \tag{17}$$

$$\text{likelihood ratio}=\frac{p(y\mid x')}{p(y\mid x)},\ \text{prior ratio}=\frac{p(x')}{p(x)},\ \text{proposal ratio}=\frac{r_m(x')}{r_m(x)q(u)},\ \text{Jacobian}=\left|\frac{\partial x'}{\partial(x,u)}\right| \tag{18}$$

$r_m(x)$ is the proposal probability of selecting the movement type, $m$ in the current speed distribution $x$, $u$ is the continuous variable sample, and $q(u)$ is the proposal probability density distribution of $u$.

In each iteration, split or combine, birth, or death is selected according to the proposal probabilities of the movement type. $b_K$ is the proposal probability of splitting (or birth) when the current number of the speed density distributions is $K$, $d_K$ is the proposal probability of the corresponding combine (or death), as in Eq. (19). $K_{\max}$ is the maximum number of the speed distributions.

$$b_K=\begin{cases}1, K=1;\\ 0, K=K_{\max}\\ 0.5, \text{其它}.\end{cases}\quad,\quad d_K=\begin{cases}0, K=1;\\ 1, K=K_{\max}\\ 0.5, \text{其它}.\end{cases} \tag{19}$$

In case of the split move, the $j*$ distribution is randomly selected to be split. The weight, the mean and the variance of the speed of the distribution $j*$ are split into two new distributions $j1$ and $j2$, as in Eq.(20). And the observed values belonging to the distribution $j*$ are redistributed to the new distribution $j1$ or using the Monte Carlo method, see Eq.(13). However, once the principle of order ($\mu_{j*-1}<\mu_{j1}<\mu_{j2}<\mu_{j*+1}$) is violated, then the split will be rejected.

$$\begin{aligned}&u_1 \sim beta(2,2), u_2 \sim beta(2,2), u_3 \sim beta(1,1)\\&w_{j1}=u_1 w_{j*},\ w_{j2}=(1-u_1)w_{j*}\\&\mu_{j1}=\mu_{j*}-u_2\sigma_{j*}\sqrt{\frac{w_{j2}}{w_{j1}}},\ \mu_{j2}=\mu_{j*}+u_2\sigma_{j*}\sqrt{\frac{w_{j1}}{w_{j2}}}\\&\sigma_{j1}^2=u_3(1-u_2^2)\sigma_{j*}^2\frac{w_{j*}}{w_{j1}},\ \sigma_{j2}^2=(1-u_3)(1-u_2^2)\sigma_{j*}^2\frac{w_{j*}}{w_{j2}}\end{aligned} \tag{20}$$

In case of combine move, the adjacent distributions $j1$ and $j2$ are merged into the distribution $j*$, as in Eq.(21).

$$\begin{aligned}&w_{j*}=w_{j1}+w_{j2}\\&w_{j*}\mu_{j*}=w_{j1}\mu_{j1}+w_{j2}\mu_{j2}\\&w_{j*}\left(\mu_{j*}^2+\sigma_{j*}^2\right)=w_{j1}\left(\mu_{j1}^2+\sigma_{j1}^2\right)+w_{j2}\left(\mu_{j2}^2+\sigma_{j2}^2\right)\end{aligned} \tag{21}$$

Then the acceptance probability of a split move can be defined as the minimum value between 1 and the result of Eq. (22). The probability of acceptance for a combine move is $\min\{1,\quad\}$.

$$\begin{aligned}&\text{Likelihood ratio}=\frac{\left(\prod_{\{i\mid z_i=j1\}}N(y_i\mid\mu_{j1},\sigma_{j1}^2)\right)\left(\prod_{\{i\mid z_i=j2\}}N(y_i\mid\mu_{j2},\sigma_{j2}^2)\right)}{\prod_{\{i\mid z_i=j*\}}N(y_i\mid\mu_{j*},\sigma_{j*}^2)}\\&\text{Prior ratio}=\frac{p(K+1)}{p(K)}\times(K+1)\times\frac{w_{j1}^{\delta-1+n_{j1}}w_{j2}^{\delta-1+n_{j2}}}{w_{j*}^{\delta-1+n_{j1}+n_{j2}}B(\delta,K\delta)}\\&\times\sqrt{\frac{\kappa}{2\pi}}\exp\left[-\frac{1}{2}\kappa\left\{(\mu_{j1}-\xi)^2+(\mu_{j2}-\xi)^2-(\mu_{j*}-\xi)^2\right\}\right]\\&\times\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{\sigma_{j1}^2\sigma_{j2}^2}{\sigma_{j*}^2}\right)^{-\alpha-1}\exp\left\{-\beta\left(\sigma_{j1}^{-2}+\sigma_{j2}^{-2}-\sigma_{j*}^{-2}\right)\right\}\\&\text{Proposal ratio}=\frac{d_{K+1}}{b_K P_{alloc}}\left\{g_{2,2}(u_1)g_{2,2}(u_2)g_{1,1}(u_3)\right\}^{-1}\\&\text{Jacobian}=\frac{w_{j*}\left|\mu_{j1}-\mu_{j2}\right|\sigma_{j1}^2\sigma_{j2}^2}{u_2(1-u_2^2)u_3(1-u_3)\sigma_{j*}^2}\end{aligned} \tag{22}$$

Where $K$ is the number of the speed distributions before splitting, $n_{j1}$ and $n_{j2}$ are the number of observations assigned to the speed distributions $j1$ and $j2$, $B(\cdot,\cdot)$ is the beta function, $\Gamma(\cdot)$ is the gamma function, $P_{alloc}$ is the probability of occurrence of this particular allocation result, $g_{p,q}(\cdot)$ is the density function of the beta function with parameters $p$ and $q$.

When a birth move happens, a new speed distribution is generated, as Eq. (23), and the weight of all speed distributions is normalized.

$$w_{j*} \sim beta(1,K),\ \mu_{j*} \sim N(\xi,\kappa^{-1}),\ \sigma_j^* \sim \Gamma(\alpha,\beta) \tag{23}$$

The probability of acceptance for a birth move is $\min\{1,A\}$, see Eq.(24). The probability of acceptance for death move is $\min\{1,A^{-1}\}$. $K$ is the number of speed distributions before the birth move; $K_0$ is the number of empty speed distributions.

$$\begin{aligned}
&\text{Likelihood ratio}=1\\
&\text{Prior ratio}=\frac{p(K+1)}{p(K)}\times(K+1)\times\frac{w_{j*}^{\delta-1}}{B(\delta,K\delta)}(1-w_{j*})^{N+K\delta-K}\\
&\text{Proposal ratio}=\frac{d_{K+1}}{b_K K_0}g_{1,K}(w_{j*})^{-1}\\
&\text{Jacobian}=(1-w_{j*})^{K-1}
\end{aligned} \tag{24}$$

*3.3 Description of the running state of hybrid BRT*

The model above provide a method for descript the running state of hybrid BRT. Road traffic conditions are usually based on vehicle speed, the road is simply divided into smooth, slow, crowded, serious congestion and other states. BRT is the same. This simple state description can not effectively describe the speed distribution, the information provided is also very small. The Gaussian mixture model of interstation travel speed provides a new opportunity to describe the hybrid BRT running state. Each of the distributions of the model is treated as a sub-state, and the result of the whole model is the fusion of all sub-states, representing the interstation running state of the BRT. As the model provides a wealth of information, including the number of sub-states, the mean, variance and weight of each sub-state, etc., where the bigger number of sub-state, the more complex running state. So a comparative analysis can be performed between the results of different models and the corresponding interstation physical environment, and the causes leading to complex traffic conditions can be identified, this provide a base for optimizing hybrid BRT.

## 4. calibration and validation of the hyperparameters

The proposed model belongs to the category of unsupervised machine learning technologies that require no calibration. The calibration and validation here means the determination of hyperparameters that was included in the model to infer the original parameters for speed distributions within a Bayesian framework. The objective function of the model hyperparameter calibration is defined as the maximizing likelihood ratio of the Gaussian mixture model, as in Equation 25.

$$\max_{\boldsymbol{\theta}} \tilde{z}(\boldsymbol{\theta}) = \prod_{s}^{N_s}\left\{\rho_s \prod_{i}^{N} p(y_i \mid \boldsymbol{\mu},\boldsymbol{\sigma}^2)\right\} \tag{25}$$

Where $\tilde{z}(\cdot)$is the objective function to be maximized,$\boldsymbol{\theta}$is a vector of hyperparameters, $\grave{e} = (\lambda,\delta,\xi,e,f,\alpha,g,h)$, $s$ is a link between neighbored stations of the mixed BRT, $N_s$ is the number of the links, $\rho_s$ is the weight of $s$.

Since $\tilde{z}(\cdot)$ was not a closed-form function, it could be evaluated by simulation. Furthermore, the function's mathematical properties such as continuity, differentiability, and concavity were totally unknown. It was thus impossible to employ a rigorous optimization algorithm like Newton–Raphson. The only way to optimize the objective function was to use metaheuristics. A particle-swarm optimization algorithm was the best fit for solving the above problem, since that algorithm is known to be good at optimizing a non-convex objective function with continuous variables.

The model calibration is done using the particle swarm optimization algorithm, the steps as follows:

1) to generate n particles, each particle randomly generates the initial solution,

2) to calculate the fitness function $z(\theta_i)$ ($i=1,\cdots,n$) of the current solution $\theta_i$ of each particle, as Eq. (26),

3) to obtain the optimal solution experienced by each particle, denoted as $\theta_{pi}$, and the optimal solution experienced by all particles, denoted as $\theta_g$,

4) to update current solution of the particles according to Eq. (26),

5) return to step 2 until the iteration finished.

$$\begin{aligned} \mathbf{v}_i &= \mathbf{v}_i + \mathbf{r}_1 \cdot \left(\boldsymbol{\theta}_{pi} - \boldsymbol{\theta}_i\right) + \mathbf{r}_2 \cdot \left(\boldsymbol{\theta}_g - \boldsymbol{\theta}_i\right) \\ \boldsymbol{\theta}_i &= \boldsymbol{\theta}_i + \mathbf{v}_i \end{aligned} \tag{26}$$

Where $\boldsymbol{\theta}_i$ and $\mathbf{v}_i$ are the solution and corresponding speed of the particle $i$ respectively, $\mathbf{r}_1$ and $\mathbf{r}_2$ are random numbers with domain 0-1.

## 5. Case study

The case study is applied to Guangzhou Bus Rapid Transit (GBRT), a kind of hybrid BRT. Its corridor runs along the center line of Zhongshan Avenue, with a length of 23 kilometers. There are 26 stations, numbered 1 to 26, the first station is Tianhe Sports Center, the center of the downtown, the 26th is Xiayuan, located in the outskirts of the city (for details, see Appendix 1). Totally there are 31 bus lines, named from B1 to B31 respectively, but only B1 runs from the start BRT station to the end one, other bus lines only enter the BRT corridor in necessary, when they left the corridor, they are just ordinary bus lines. So, the ordinary bus lines do not separate from GBRT, and passengers can transfer in station. For this reason, GBRT is double efficient than other BRTs.

The station-reporting data of GBRT were collected at 7th Dec, 2015, and from them there are 65 thousand of the travel speed data of the inter-station (from Tianhe Sports Center to Xiayuan, totally 25 inter-stations,named by neighbored station numbers, as Figure 1) were computed (for detailed sample data, see Appendix 2) .

*5.1 Calibration and validation results*

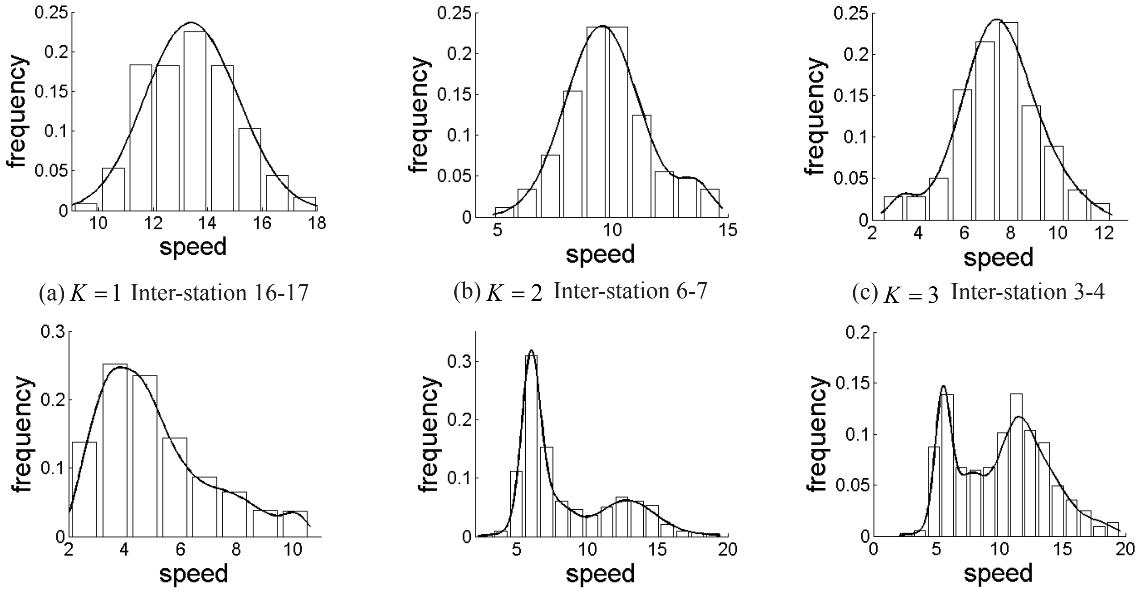
In the calibration process of the interstation travel speed model, the number of particles is set to 10, the number of iterations is set to 80 times. The hyperparameter $\lambda$ and $\delta$ set to 1, $\xi$ is set to the mean of the travel speed of corresponding inter-station, the calibration results as Table 1.

**Table 1** calibration values for hyperparameters.

| $\lambda$ | $\delta$ | $\xi$ | $e$ | $f$ | $\alpha$ | $g$ | $h$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | Midpoint value of $R$ | 1.71 | 5.07 | 2.30 | 2.75 | 8.37 |
| *$R$ is the range of travel speeds data for each inter-station. | | | | | | | |

The RJMCMC method is applied to infer the number of speed distributions and the weight, mean and variance of each speed distribution, see Appendix 3. In the process, the adopted Gibbs sampler took random draws from the posterior distribution. After conducting 50,000 Gibbs iterations, the former 5000 replications were discarded as burn-in data, and the remaining draws were saved to approximate the joint posterior distribution of the parameters. Where the estimated number of the speed distributions was set as the number with the maximum marginal posterior probability, and a distribution's estimated travel speed was set as the mean of the marginal posterior distribution of the mean travel speed for the distribution. The estimated variance of travel speeds also denoted the mean of the posterior distribution for the variance in travel speeds of each distribution.

The results showed that the number of the speed distributions was different among the different inter-stations, it ranged from 1 to 6, and each number accounted for 8%, 44%, 16%, 16%, 12% and 4% of all the numbers, as in Figure 2. To illustrate the result, six estimated model of inter-station travel speed and corresponding histograms of observed speed samples were draw in Figure 1, their component number are from 1 to 6, respectively (for all results, see Appendix 3) . It could be found that the estimated results of the Bayesian prediction can be well fitted to the observed speed samples. It is difficult to visually judge the component number and the weight, speed mean and variance of each component according to the speed histogram, but in the estimated results they are vary clear, indicating that the proposed model is effective and necessary.




(a) $K = 1$ Inter-station 16-17 (b) $K = 2$ Inter-station 6-7 (c) $K = 3$ Inter-station 3-4

(d) $K = 4$ Inter-station 1-2 (e) $K = 5$ Inter-station 10-11 (f) $K = 6$ Inter-station 9-10

**Figure 1.** Comparison of model curve of different interstation and corresponding histograms of speed samples

*5.2 Running states of GBRT*

The modelling results shed a light on the running state of GBRT. If a component of Gaussian mixture model represents an interstation running state, then the model represents the mixture results of all the interstation running states. Figure 2 is the K values of all 25 interstations, the horizontal axis is the sequence of stations, the first station is Tianhe Sports Center, the 26th station is Xiayuan, and the 4th, 6th, 9th, 11th stations are Normal Univ & Jinan Univ, Shangshe, Tangdong, Chebei station respectively. Figure 3 is the posterior mean of the expected speed of the running states of all 25 interstations.

It can be observed from Figure 2, the K values of the models in urban area are higher than those in suburban. As the BRT part in suburban basically has no external disturbance, so we can conclude that the number of the running state of GBRT itself is 2. This conclusion can also be obtained from the urban section of GBRT, such as the inter-station of 6-7, from the GOOGLE map it can be seen, there are no road intersection and traffic lights, traffic is also not much, that is, there is no factor to interfere with BRT operation, so its K value is 2.

Also in the Figure 2, there are two sections with K value 5 or more. But combined with the GOOGLE map it can be found that the reasons are different. In the first part, there is a highway entrances between inter-station 4-5, which is a highway through the city center, and station 1 to 4 are work destinations, so the highway brings a lot of traffic flow, which disturbs GBRT a lot. The second part 9-11 crosses with main road (Chebei Road) plainly, this road is a north-south traffic arteries, connecting a fast road nearby, which brings a lot of traffic flow into the downtown, they disturb GBRT a lot, and the road has many straight traffic flow, so the buses of GBRT must wait for a long red time. so, the K value of inter-station 9-10, a maximum of 6, is the superimposed effect of the two factors.

On the other hand, the bigger K value of the interstation speed model means the more complex of the running state of the inter-station, this conclusion can be proved in Figure 3, it can be observed that posterior means of the means of the speeds are very scattered for those models with K value more than 4.

How to reduce the K value, that is, to reduce the complexity of the operating state of BRT? We can also find answer from GBRT. Comparing inter-station 10-11 and 7-8 (with K value 4) In the network map, they have similar environment. There is also a main road (Keyun Road) crossing with inter-station 7-8, a more important road than Chebei Road, but it passes through the inter-station below, thus the direct traffic flow interference is avoided, which greatly reduced the waiting time of GBRT buses for traffic lights. Inter-station 3-4 is in a similar situation, but the solution is to raise the BRT lane through the intersection, thus avoiding the disturbance of traffic lights and other directions of traffic.

Therefore, the main road traffic flow and traffic lights are the main factors affecting the BRT running state, the solution is to elevate BRT lane, or let other main road cross through the BRT lane below, if the traffic flow of the main road is relatively large.

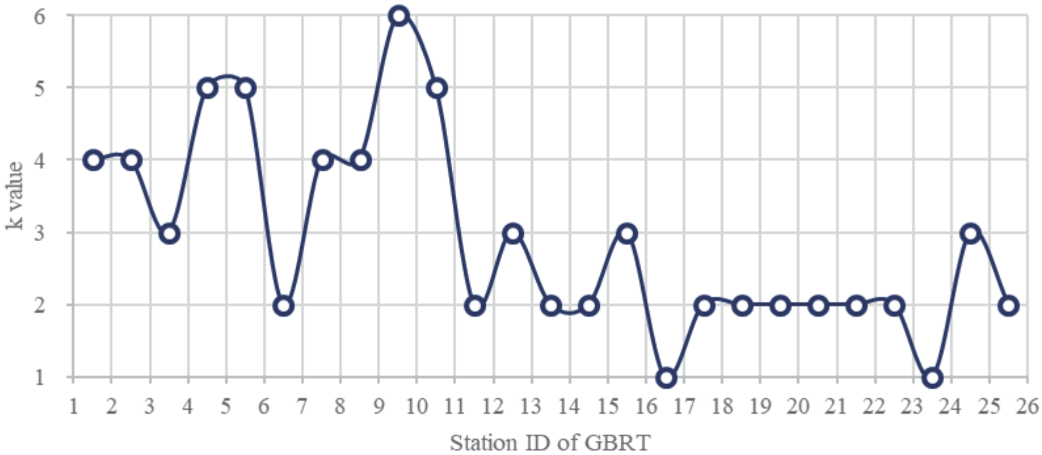



**Figure 2.** $K$ values of the interstation travel speed models.

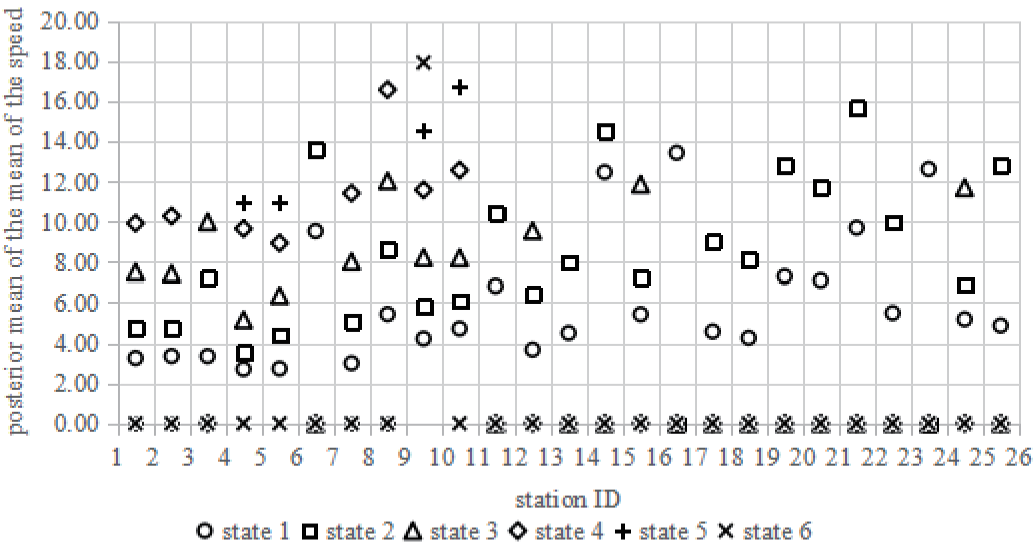



**Figure 3.** Estimated speed of the interstation travel states of GBRT

## Conclusions

Hybrid BRT greatly improves the operational efficiency of urban public transport, but it is also disturbed by urban road traffic. Due to road traffic, traffic lights and other factors, the distribution and changes of its interstation travel speed are very complex. In this paper, a Gaussian mixture model describing the interstation travel speed of hybrid BRT under Bayesian framework is established. The component number of of the model and the weight, mean and variance of each component are deduced by RJMCMC method. The model is applied to the Guangzhou BRT, and the travel speed of 25 inter-stations is modeled to describe their running states. It can be observed that the model can very effectively describe the heterogeneous speed data among different inter-stations, and provide richer information usually not available from the traditional models. The causes of different speed distribution can also be identified, this provides a base for optimizing hybrid BRT.

Firstly, the component number of different speed model varied greatly, the value is from 1 to 6, this implies the proposed model is necessary and appropriate.

Second, the number of the running state is usually 2 for GBRT itself, but as the influencing factors increase, the number increases consequently, until to 6.

Third, the main road will lead to many traffic and long time to wait for the red light, they will have a serious impact on the running state if it crosses with the BRT. The way to reduce these effects is to elevate the BRT lane through the main road, or to sink the road through the BRT lane.

## Acknowledgements

The authors acknowledge that this paper was prepared based on Science and technology planning project of Guangdong province of China in 2017 (No. 2017B010111007),the National Natural Science 16 Foundation of China (No. 41271181).

## Appendix

Appendix 1 the stations of GBRT (the station name is mainly named in Chinese pinyin)

| Station ID | Station Name | Station ID | Station Name | Station ID | Station Name |
|---|---|---|---|---|---|
| 1 | Tianhe Sports Center | 10 | Tianlang mingju | 19 | Wuchong |
| 2 | Shipai Qiao | 11 | Chebei | 20 | Huangpu Coach Station |
| 3 | Gangding | 12 | Dongpuzhen | 21 | Shuanggang |
| 4 | Normal Univ & Jinan Univ | 13 | Huangcun | 22 | Shapu |
| 5 | Huajing New Town | 14 | Zhucun | 23 | Nanhai Temple |
| 6 | Shangshe | 15 | Lianxi | 24 | Miaotou |
| 7 | Xueyuan | 16 | Maogang | 25 | Nanwan |
| 8 | Xueyuan | 17 | Zhujiangcun | 26 | Xiayuan |
| 9 | Tangdong | 18 | Xiasha | | |

Appendix 2 The samples of the station-reporting data( the data is collected from bus801189, line B1)

| Line name | Bus number | Station | Reporting time | Tags of in/out |
|---|---|---|---|---|
| Line B1 | 801189 | Tianhe Sports Center | 2015-12-7 11:00:20 | Out |
| Line B1 | 801189 | Shipai Qiao | 2015-12-7 11:02:27 | In |
| Line B1 | 801189 | Shipai Qiao | 2015-12-7 11:02:40 | Out |
| Line B1 | 801189 | Gangding | 2015-12-7 11:06:13 | In |
| Line B1 | 801189 | Gangding | 2015-12-7 11:06:55 | Out |
| Line B1 | 801189 | Normal Univ & Jinan Univ | 2015-12-7 11:08:43 | In |
| Line B1 | 801189 | Normal Univ & Jinan Univ | 2015-12-7 11:09:21 | Out |
| Line B1 | 801189 | Huajing New Town | 2015-12-7 11:11:01 | In |
| Line B1 | 801189 | Huajing New Town | 2015-12-7 11:11:48 | Out |
| Line B1 | 801189 | Shangshe | 2015-12-7 11:12:56 | In |
| Line B1 | 801189 | Shangshe | 2015-12-7 11:13:20 | Out |
| Line B1 | 801189 | Xueyuan | 2015-12-7 11:14:08 | In |
| Line B1 | 801189 | Xueyuan | 2015-12-7 11:14:31 | Out |
| Line B1 | 801189 | Tangxiacun | 2015-12-7 11:16:59 | In |
| Line B1 | 801189 | Tangxiacun | 2015-12-7 11:17:29 | Out |
| Line B1 | 801189 | Tangdong | 2015-12-7 11:18:06 | In |
| Line B1 | 801189 | Tangdong | 2015-12-7 11:18:29 | Out |
| Line B1 | 801189 | Tianlangmingju | 2015-12-7 11:19:05 | In |
| Line B1 | 801189 | Tianlangmingju | 2015-12-7 11:19:30 | Out |
| Line B1 | 801189 | Chebei | 2015-12-7 11:20:15 | In |
| Line B1 | 801189 | Chebei | 2015-12-7 11:20:44 | Out |
| Line B1 | 801189 | Dongpuzhen | 2015-12-7 11:22:11 | In |
| Line B1 | 801189 | Dongpuzhen | 2015-12-7 11:22:40 | Out |
| Line B1 | 801189 | Huangcun | 2015-12-7 11:24:10 | In |
| Line B1 | 801189 | Huangcun | 2015-12-7 11:24:33 | Out |
| Line B1 | 801189 | Zhucun | 2015-12-7 11:26:26 | In |
| Line B1 | 801189 | Zhucun | 2015-12-7 11:26:58 | Out |

| Line name | Bus number | Station | Reporting time | Tags of in/out |
|---|---|---|---|---|
| Line B1 | 801189 | Lianxi | 2015-12-7 11:28:37 | In |
| Line B1 | 801189 | Lianxi | 2015-12-7 11:29:03 | Out |
| Line B1 | 801189 | Maogang | 2015-12-7 11:29:53 | In |
| Line B1 | 801189 | Maogang | 2015-12-7 11:30:18 | Out |
| Line B1 | 801189 | Zhujiangcun | 2015-12-7 11:32:06 | In |
| Line B1 | 801189 | Zhujiangcun | 2015-12-7 11:32:40 | Out |
| Line B1 | 801189 | Xiasha | 2015-12-7 11:34:48 | In |
| Line B1 | 801189 | Xiasha | 2015-12-7 11:35:18 | Out |
| Line B1 | 801189 | Wuchong | 2015-12-7 11:38:18 | In |
| Line B1 | 801189 | Wuchong | 2015-12-7 11:39:01 | Out |
| Line B1 | 801189 | Huangpu Coach Station | 2015-12-7 12:17:10 | In |
| Line B1 | 801189 | Huangpu Coach Station | 2015-12-7 12:17:22 | Out |
| Line B1 | 801189 | Shuanggang | 2015-12-7 12:21:59 | In |
| Line B1 | 801189 | Shuanggang | 2015-12-7 12:22:11 | Out |
| Line B1 | 801189 | Shapu | 2015-12-7 12:23:20 | In |
| Line B1 | 801189 | Shapu | 2015-12-7 12:23:32 | Out |
| Line B1 | 801189 | Nanhai Temple | 2015-12-7 12:24:30 | In |
| Line B1 | 801189 | Nanhai Temple | 2015-12-7 12:24:42 | Out |
| Line B1 | 801189 | Miaotou | 2015-12-7 12:25:32 | In |
| Line B1 | 801189 | Miaotou | 2015-12-7 12:25:44 | Out |
| Line B1 | 801189 | Nanwan | 2015-12-7 12:27:14 | In |
| Line B1 | 801189 | Nanwan | 2015-12-7 12:27:26 | Out |
| Line B1 | 801189 | Xiayuan | 2015-12-7 12:28:07 | In |

* interstation travel speed=interstation lane length/ (in time of current station- out time of previous station)

Appendix 3 the Estimated Values of the Interstation Models of GBRT

| Interstation ID | Component number | Parameter values of each component | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Parameter | 1 | 2 | 3 | 4 | 5 | 6 |
| 1-2 | 4 | mean | 3.2510 | 4.7673 | 7.5052 | 9.9131 | - | |
| | | variance | 0.6567 | 0.9772 | 1.0540 | 0.5350 | | |
| | | weight | 0.2844 | 0.4902 | 0.1824 | 0.0430 | | |
| 2-3 | 4 | mean | 3.3493 | 4.7628 | 7.4201 | 10.2722 | - | |
| | | variance | 0.5114 | 0.9080 | 1.0799 | 1.3580 | | |
| | | weight | 0.3122 | 0.3154 | 0.2699 | 0.1025 | | |
| 3-4 | 3 | mean | 3.3452 | 7.2113 | 9.9850 | - | | |
| | | variance | 0.5656 | 1.3406 | 1.1374 | | | |
| | | weight | 0.0463 | 0.7962 | 0.1575 | | | |
| 4-5 | 5 | mean | 2.7049 | 3.5278 | 5.1521 | 9.6434 | 10.9758 | - |
| | | variance | 0.3050 | 0.4935 | 0.7621 | 1.7421 | 1.1170 | |
| | | weight | 0.1003 | 0.0508 | 0.0555 | 0.4741 | 0.3193 | |
| 5-6 | 5 | mean | 2.7290 | 4.4130 | 6.3569 | 8.9291 | 10.9554 | - |
| | | variance | 0.3795 | 0.7524 | 0.9311 | 0.9566 | 0.6042 | |
| | | weight | 0.1575 | 0.4369 | 0.2276 | 0.1341 | 0.0440 | |
| 6-7 | 2 | mean | 9.5269 | 13.5904 | - | | | |
| | | variance | 1.5936 | 0.6886 | | | | |
| | | weight | 0.9321 | 0.0679 | | | | |
| 7-8 | 4 | mean | 2.9951 | 5.0860 | 8.0131 | 11.4071 | - | |
| | | variance | 0.5521 | 1.1308 | 1.3667 | 0.6810 | | |
| | | weight | 0.1251 | 0.5669 | 0.2622 | 0.0458 | | |

| Interstation ID | Component number | Parameter values of each component | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Parameter | 1 | 2 | 3 | 4 | 5 | 6 |
| 8-9 | 4 | mean | 5.4349 | 8.6181 | 12.0195 | 16.5638 | - | |
| | | variance | 0.7542 | 1.7467 | 1.5425 | 1.3547 | | |
| | | weight | 0.2134 | 0.4307 | 0.3363 | 0.0195 | | |
| 9-10 | 6 | mean | 4.2223 | 5.8194 | 8.2147 | 11.5758 | 14.4941 | 17.9123 |
| | | variance | 1.0963 | 0.8013 | 1.1026 | 1.3448 | 1.3659 | 1.1198 |
| | | weight | 0.0335 | 0.2247 | 0.1761 | 0.3588 | 0.1672 | 0.0397 |
| 10-11 | 5 | mean | 4.7181 | 6.0802 | 8.2034 | 12.5544 | 16.7726 | - |
| | | variance | 1.2262 | 0.7513 | 1.1629 | 1.7347 | 1.5659 | |
| | | weight | 0.0490 | 0.4864 | 0.1537 | 0.2621 | 0.0488 | |
| 11-12 | 2 | mean | 6.8155 | 10.4426 | - | | | |
| | | variance | 1.1517 | 1.7286 | | | | |
| | | weight | 0.5239 | 0.4761 | | | | |
| 12-13 | 3 | mean | 3.6690 | 6.4205 | 9.5419 | - | | |
| | | variance | 0.7143 | 1.3528 | 1.5391 | | | |
| | | weight | 0.1366 | 0.5548 | 0.3086 | | | |
| 13-14 | 2 | mean | 4.5022 | 7.9813 | - | | | |
| | | variance | 0.7519 | 2.1595 | | | | |
| | | weight | 0.3546 | 0.6454 | | | | |
| 14-15 | 2 | mean | 12.4571 | 14.5260 | - | | | |
| | | variance | 1.1832 | 1.4558 | | | | |
| | | weight | 0.6132 | 0.3868 | | | | |
| 15-16 | 3 | mean | 5.4279 | 7.2572 | 11.8534 | - | | |
| | | variance | 0.5403 | 1.2110 | 1.9680 | | | |
| | | weight | 0.2231 | 0.3172 | 0.4598 | | | |
| 16-17 | 1 | mean | 13.4194 | - | | | | |
| | | variance | 1.6674 | | | | | |
| | | weight | 1.0000 | | | | | |
| 17-18 | 2 | mean | 4.5750 | 9.0009 | - | | | |
| | | variance | 1.2792 | 2.2615 | | | | |
| | | weight | 0.4637 | 0.5363 | | | | |
| 18-19 | 2 | mean | 4.2578 | 8.1231 | - | | | |
| | | variance | 0.8942 | 1.4810 | | | | |
| | | weight | 0.6028 | 0.3972 | | | | |
| 19-20 | 2 | mean | 7.2892 | 12.8186 | - | | | |
| | | variance | 1.9520 | 1.5421 | | | | |
| | | weight | 0.6603 | 0.3397 | | | | |
| 20-21 | 2 | mean | 7.0950 | 11.6966 | - | | | |
| | | variance | 1.2979 | 1.7037 | | | | |
| | | weight | 0.6196 | 0.3804 | | | | |
| 21-22 | 2 | mean | 9.6947 | 15.6545 | - | | | |
| | | variance | 2.0056 | 1.2307 | | | | |
| | | weight | 0.8620 | 0.1380 | | | | |
| 22-23 | 2 | mean | 5.4889 | 9.9591 | - | | | |
| | | variance | 0.6875 | 1.6440 | | | | |
| | | weight | 0.1051 | 0.8949 | | | | |
| 23-24 | 1 | mean | 12.6109 | - | | | | |
| | | variance | 2.4286 | | | | | |
| | | weight | 1.0000 | | | | | |
| 24-25 | 3 | mean | 5.1802 | 6.8780 | 11.6853 | - | | |
| | | variance | 0.4809 | 1.0225 | 2.0139 | | | |
| | | weight | 0.4070 | 0.3084 | 0.2846 | | | |

| Interstation ID | Component number | Parameter values of each component | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Parameter | 1 | 2 | 3 | 4 | 5 | 6 |
| 25-26 | 2 | mean | 4.8721 | 12.8046 | - | | | |
| | | variance | 1.1996 | 3.2176 | | | | |
| | | weight | 0.1105 | 0.8895 | | | | |

## References

[1] Deng T.T., and John D. NELSON. Recent Developments in Bus Rapid Transit: A Review of the Literature. Transport Reviews, Vol. 31, No. 1, January 2011, pp. 69–96.

[2] Park B. J., and Zhang Y., and Lord D. Bayesian mixture modeling approach to account for heterogeneity in speed data[J]. Transportation research part B: methodological, Vol. 44, No 5, 2010, pp. 662-673.

[3] McLean, J. R. Observed speed distributions and rural road traffic operations. Proceedings of the 9th Australian Road Research Board Conference, Part 5, Australian Road Research Board, Vermont South, Victoria, Australia, 1978, pp. 235-244.

[4] Dey, P. P, Chandra, C., Gangopadhaya, S. Speed distribution curves under mixed traffic conditions. Journal of Transportation Engineering.Vol. 132, No. 6, 2006, pp. 475-481.

[5] Ko, J., Guensler, R. L. Characterization of congestion based on speed distribution: a statistical approach using Gaussian mixture model, TRB 2005 Annual Meeting CD-ROM, 2005.

[6] Jun, J. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic. Transport Research. Part C: Emergent Technology, Vol. 18, No.4, 2010, pp. 599–610.

[7] Corey, J., Lao, Y., Wu, Y.J., Wang, Y.,. Detection and correction of inductive loop detector sensitivity errors by using gaussian mixture models. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2256, No. 1, 2011, pp. 120–129.

[8] Lao, Y., Zhang, G., Corey, J., Wang, Y. Gaussian mixture model-based speed estimation and vehicle classification using single-loop measurements. Journal of Transportation System, Vol. 16 No. 4, 2012, pp. 184–196.

[9] Richardson, S., Green, P.J. On Bayesian analysis of mixture models with an unknown number of components (with discussion). Journal of The Royal Statistical Society, Ser. B59, 1997, pp. 731–792.

[10] Green, P.J. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika, 82 (4), 1995, pp. 711–732.

[11] Lee M., Sohn K. Inferring the route-use patterns of metro passengers based only on travel-time data within a Bayesian framework using a reversible-jump Markov chain Monte Carlo (MCMC) simulation. Transportation Research Part B, Vol. 81, 2015, pp. 1–17.