Traffic & Transportation

Macroscopic Fundamental Diagram Estimation Considering Traffic Flow Condition of Road Network

Xiaoli DENG[1], Yao HU[2]

Original Scientific Paper
Submitted: 16 Oct. 2022
Accepted: 11 July 2023

[1] 13983573412@163.com, School of Mathematics and Statistics, Guizhou University
[2] Corresponding author, yhu1@gzu.edu.cn, Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University

This work is licensed under a Creative Commons Attribution 4.0 International License

Publisher:
Faculty of Transport and Traffic Sciences, University of Zagreb

ABSTRACT

A macroscopic fundamental diagram (MFD) is an important basis for road network research. It describes the functional relationship between the average flow and average density of the road network. We proposed an MFD estimation method based on the traffic flow condition. Firstly, according to statistical theories, the road network data are divided into three traffic flow conditions (free flow, chaotic and congested) bounded by a 95% confidence interval of the maximum traffic capacity of each intersection in the road network. Then, in each condition, we combined principal component analysis and the Jolliffe B4 method to reduce dimension for extracting critical intersections. Finally, the full-scale dataset of the road network was reconstructed to estimate the road network MFD. Through numerical simulation and empirical research, it is found that the root mean square error and absolute percentage error between estimated MFD and true MFD considering the traffic flow condition are smaller than those without considering the traffic flow condition. The MFD estimation and the division of the traffic states of the road network were completed at the same time. The proposed method effectively saves the time cost of road network research and is highly accurate.

KEYWORDS

macroscopic fundamental diagram; traffic state; confidence interval; principal component analysis; Jolliffe B4 method.

1. INTRODUCTION

With the development of technology and the improvement of people's living standards, vehicle ownership is also increasing, and the problem of urban traffic congestion is becoming a focus of attention, and how to alleviate it has become an important research direction. Seo et al. [1–3] tried to solve the traffic congestion problem on a road by estimating the traffic flow condition and fundamental diagram from probe vehicle data. As for the congestion of complex urban road networks, it is not enough to solve the local road congestion, but it is necessary to use a macroscopic fundamental diagram (MFD) to characterise the network capacity and determine the traffic flow conditions.

Godfrey [4], Ardekani, Herman [5] and Mahmassani et al. [6] were the first to propose the MFD based on a road network consisting of multiple sections to describe the relationship between average flow and average density. Starting from the overall structure of the road network, Geroliminis and Daganzo [7] proved the existence of MFD through empirical research in Yokohama, Japan. Courbon [8] and Saberi [9] used loop detector data to estimate the MFD. Tilg et al. [10] also proposed the method of cuts and stochastic approximation to derive MFD, but they only approximated and fitted the detector data of the main road as the MFD of the network. There are also researchers who estimated MFD from probe vehicle data, such as Nagle et al. [11] who investigated the accuracy of estimating the average flow and average density in a network with a priori known probe vehicle penetration. Knoop et al. [12] used a large dataset of aggregated probe vehicles provided by Google to present a well-defined and clear (with low dispersion) MFD for the Amsterdam network, but the probe vehicle data could only provide partial values for MFD estimation, and these studies were based on a uniform distribution of vehicles in the network. In addition, Lin et al. [13] proposed an MFD estimation method that fuses probe vehicle data and loop detector data, Paipuri et al. [14] divided the network into homogeneous areas by geographical location and then used the trajectory data recorded by cell phones to estimate the parameters of

MFD, these studies based on other data sources not only require prior knowledge of the penetration rate but are also limited by the large amount of available trajectory data. Saffari et al. [15] reconstructed the average flow and average density by extracting the critical links from the road network, and although the requirement of data was reduced, this approach was only validated in simulation.

There are also differences in the use of different research methods to determine the traffic states. Kerner et al. [16, 17] based on the measured data of free flow and congested two-phase traffic flow, the traffic states are distinguished as free flow, synchronised flow and wide moving jam. For example, Xu et al. [18] introduced network operation efficiency, and Liu et al. [19] proposed the standard deviation of the number of vehicles in all road sections to identify the traffic states of the road network as free flow, optimal accumulation and congested. Lin [20] combines unsupervised clustering analysis methods and supervised machine learning algorithms to distinguish the traffic states of the road network into four types, namely smooth, stationary, congested and supersaturation, but these studies are premised on complex calculations.

In this paper, an estimation method based on the traffic flow condition is proposed to estimate the MFD of the whole road network with as few intersections as possible. Specifically, based on the data collected by detectors installed at every intersection on the road network, the traffic flow conditions are divided into free flow, chaotic and congested by the 95% confidence interval of the capacity density. Principal component analysis (PCA) and Jolliffe B4 method [21] are used in each condition to extract the critical intersections, and then reconstruct the full-scale data of the whole road network based on PCA mechanism, and the road network MFD with three states is estimated. Through the confidence interval in statistical theory, the three processes of estimating the MFD, the maximum traffic capacity and determining the traffic state are linked, and the traffic states are further refined. The method reduces the computational complexity and improves the estimation accuracy.

The rest of the paper is organised as follows: Section 2 introduces the theoretical method without considering the traffic flow conditions of the road network. Section 3 estimates the MFD considering the traffic flow conditions, and verifies the effectiveness of the method through numerical simulation. Section 4 verifies the feasibility of the proposed method by practical application in two cases. Section 5 concludes the research results and the future directions.

2. THE DIMENSIONALITY REDUCTION METHOD FOR MFD ESTIMATION

This section introduces the theory of the method presented in this paper and considers the influence of different traffic flow conditions on the extraction of critical intersections. The method includes both training and testing processes. The core of the training process is to extract the most representative intersections from the road network by PCA and Jolliffe B4, which are called critical intersections. The core of the testing process is to reconstruct full-scale macroscopic traffic data.

2.1 Traffic flow conditions division

Under different traffic flow conditions, the critical intersections that are representative of the road network may vary. In order to ensure the estimation accuracy of MFD as much as possible, the critical intersections are extracted by conditions presented in this paper. Only two traffic flow conditions (uncongested and congested) can be observed, while in theory there are three states: free flow, breakdown and congested [22]. In this paper, we divide the traffic flow conditions by the concept of $1-\alpha$ confidence interval of the data indicators and extract the datasets of three traffic flow conditions in the network separately for the subsequent analysis.

Assuming that an intersection p on the road network has a flow-density relationship as in *Figure 1*, it corresponds to a maximum capacity (q_{cp}, k_{cp}), q_{cp} and k_{cp} denote the capacity flow and capacity density through the intersection p, respectively, and p=1,2,…,P denotes a total of P intersections on the network. Considering that breakdown is a hard-to-observe state between free flow and congested, it usually corresponds to $(k_{cp}-\Delta k_p, k_{cp}+\Delta k_p)$ in *Figure 1*. $(0, k_{cp}-\Delta k_p]$ represents the free flow state, $[k_{cp}+\Delta k_p, k_{jp})$ represents the congested state. In this paper, we take the $1-\alpha$ confidence interval of the capacity density at each intersection to determine the size of Δk_p, which is expressed as follows:

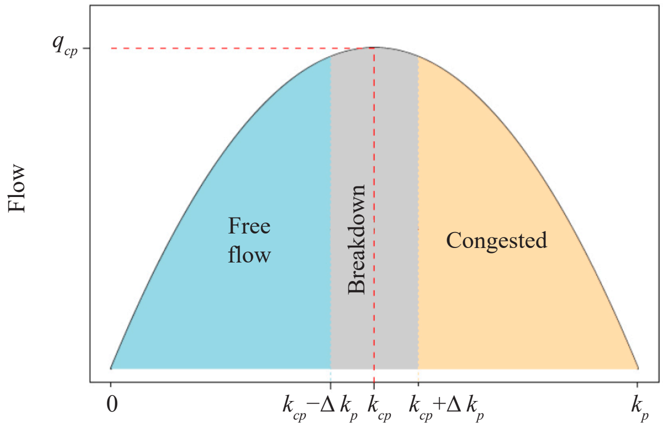


Figure 1 – Flow-density diagram of intersection p

$$\Delta k \ = u_{1\ \alpha/2}\frac{\sigma}{\sqrt{n}}, p = 1,2,..., P \tag{1}$$

where $u_{1-\alpha/2}$ represents the $1-\alpha/2$-quantile of the normal distribution, σ_p and n_p represent the standard deviation of density and the number of data samples at the intersection p, respectively.

The capacity and traffic flow conditions are related to each intersection of the road network, and each intersection may have three traffic states: free flow, breakdown and congested. Daganzo and Geroliminis [7] proposed that a sufficient condition for the existence of MFD is that the traffic flow conditions are homogeneous in time, which means all intersections of the road network are in the same state in time. If all intersections are in the free flow state, then it is considered as the free flow condition of the road network. Similarly, if all intersections are in the congested state, then it is considered as the congested condition of the road network. When there is more than one state, it is called the chaotic condition of the road network. Since the capacity of each intersection may be different, in order to ensure that all intersections are in the same state in time, and to facilitate the subsequent PCA, we calculate $k_1 = \min(k_{cp} - \Delta k_p)$, $k_2 = \max(k_{cp} + \Delta k_p)$, where p=1,2,...,P. In the following, three conditions of the road network are divided, namely free flow, chaotic and congested.

Step 1. Extracting the free flow condition. Samples with a density less than k_1 at each intersection at the same time were extracted and recorded as K_u, and their corresponding flows were recorded as Q_u. Both of them form the free flow condition.

Step 2. Extracting the congested condition. Samples with density above k_2 at each intersection at the same time were extracted and recorded as K_c, and the corresponding flows were recorded as Q_c, which together form the congested condition.

Step 3. All data except the above two conditions are classified as the chaotic condition, the flow and density were denoted as Q_m and K_m, respectively.

After the traffic flow conditions division, the following analysis is performed on each condition separately to estimate the MFD of the road network. This division method considers not only the capacity of each intersection but also the temporal characteristics of the traffic data. It should be noted that not all intersections of the road network will have a flow-density relationship as in *Figure 1*, but they will all have a maximum capacity. The method presented in this part only serves to divide the state datasets and does not affect the results of the research, so it is acceptable.

2.2 Flow-density function piecewise fitting in actual traffic

In actual traffic, there are different functions of flow and density at different intersections, and it is usually fitted by all traffic flow conditions at each intersection. However, different traffic flow conditions may have different functions, and it is less accurate to use the same function to fit all datasets at the intersections. In this paper, the flow-density functions are piecewise fitted under each state, that is, each intersection in the training set is fitted separately by the state to prepare for the reconstruction of the data in the subsequent testing process.

2.3 MFD of road network

Definition. Corresponding to each time interval t, the detector will have the measured values of flow and density. The average flow $Q(t)$ and average density $K(t)$ can be calculated as follows:

$$Q(t)=\frac{\sum q_i(t)n_il_i}{\sum n_il_i},\quad K(t)=\frac{\sum k_i(t)n_il_i}{\sum n_il_i} \tag{2}$$

where $q_i(t)$ and $k_i(t)$ respectively represent flow and density of intersection i within time interval t, with the length l_i and number of lanes n_i.

Assumption. MFD is a function relationship of average flow and average density, expressed as

$$Q(t)=f(K(t)) \tag{3}$$

where the shape of function $f(\cdot)$ is affected by demand mode, control strategy, network attributes and driving behaviour, and its specific function form is determined by the road network.

The true MFD on the road network can be calculated if the flow and density of all intersections are known. Considering the high installation and management costs of detectors and the large computational effort of the MFD, we often wish to estimate it based on data from only a few intersections. However, if different intersections are selected in the network, the average flow and density may also be different. Take a region in the Guanshan Lake District of Guiyang as an example (*Figure 2*), the red marks the 12 intersections where detectors are installed in this small area, and the true MFD is calculated as shown in the black scatter in *Figure 3*. If only 5 intersections are randomly selected, the calculated MFD may be different from the true MFD. In *Figure 3*, except for the black, each colour represents the MFD calculated by random extractions. After 10 extractions, it shows a high level of scatter in the resulting MFDs. In order to obtain a unique and consistent MFD for the road network, it is important to extract the critical intersections that best represent the macroscopic traffic characteristics (flow and density, etc.) in the network.

Figure 2 – The region road network of Guanshan Lake District in China map

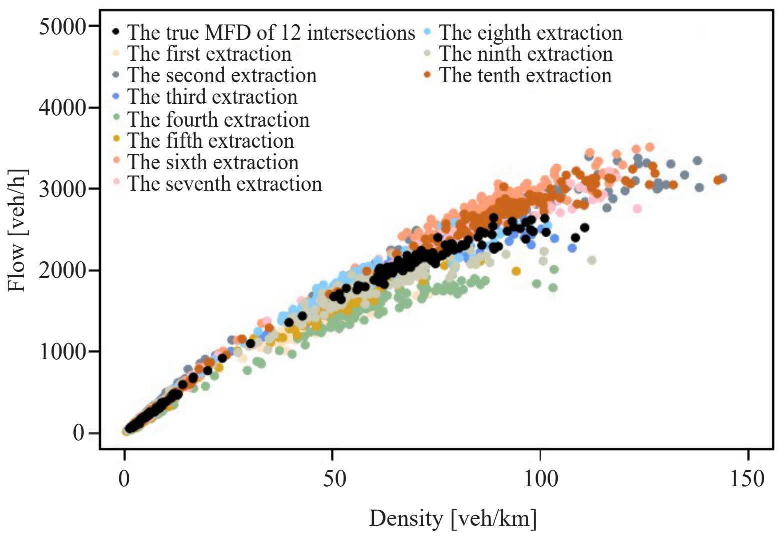


Figure 3 – 10 random extractions of MFDs at 5 critical intersections and the true MFD at 12 intersections

2.4 Critical intersections extraction

The premise of this research is that there are critical intersections in the road network that can best represent the macroscopic traffic characteristics of the whole network. Considering different traffic flow conditions, site characteristics and other factors, there may be some differences in critical intersections, which will not affect the final results.

PCA is one of the most common unsupervised learning algorithms and the most popular technique for dimensionality reduction or feature selection. PCA was first proposed by Pearson [23] and Hotelling [24], and has since been widely used as a tool when dealing with large datasets. PCA is essentially a projection of the original variables onto a new subspace, which reduces a large number of interrelated variables in a dataset to a small set of uncorrelated variables while retaining as much information as possible. This is achieved by projecting the original variables onto a new subspace. The new variables in this subspace are called principal components (PCs), which are linear combinations of the original variables, while the PCs are ordered by the amount of information captured. Thus, the first PC captures the maximum variability in the original dataset and uses a new variable to represent this. Subsequently, the other PCs consider decreasing the values of the remaining variables in the original dataset. Suppose that a data matrix is $X_{M\times P}$, the rows correspond to observations, the columns correspond to variables, and the *p* variable is a vector of *M* observations represented by $\boldsymbol{x}_p$. As mentioned, PCA defines a linear combination of variables as $\sum_{p=1}^{P}\boldsymbol{u}_p\boldsymbol{x}_p$, which are uncorrelated, where $\boldsymbol{u}_p$ is a vector of constants. The following steps summarise the PCA algorithm:

Step 1. Standardised data. There may be differences in the values at different intersections. To integrate the data to a level, perform the following operations on each element x_{mp}(p=1,2,…,P, m=1,2,…,M) in matrix $X_{M\times P}$.

$$x = \frac{x_{mp} - \mu_p}{\sigma_p} \tag{4}$$

where x constitutes the standardised data matrix $\boldsymbol{X}$, μ_p denotes the mean value of a variable p and σ_p denotes the standard deviation of a variable p.

Step 2. Calculate the covariance matrix.

$$\boldsymbol{S} = \frac{1}{P}\boldsymbol{X}^T\boldsymbol{X} \tag{5}$$

Step 3. Decompose the covariance matrix $\boldsymbol{S}$ and obtain the eigenvalues and eigenvectors.

Step 4. The eigenvalues are ordered in descending, and the first D eigenvalues $\{\lambda_d\}_{d=1}^{D}$ and their corresponding eigenvectors $\{u_d\}_{d=1}^{D}$ are selected.

Step 5. The D-dimensional projection of the original data is given by $\boldsymbol{Z}$=$\boldsymbol{XU}$, where U=[u_1,...,u_D] is $P\times D$ and $\boldsymbol{Z}$ is $M\times D$.

Principal component analysis can reconstruct the dataset of the original dimension through the selected feature vectors, so as to reconstruct the original variables, that is $\boldsymbol{X}$=$\boldsymbol{ZU}^T$. If P=D, the reconstructed original variable is accurate. If P>D, the original variables of reconstruction are approximate. The closer D is to P, the more accurately the original variable is reconstructed. If each feature vector can be associated with a variable (intersection) in the network, then the variable can be used to approximate the feature vector (principal component) in the road network. This rule will be used to extract critical intersections later.

Jolliffe [21] proposed a total of eight methods A1, A2, B1, B2, B3, B4, C1 and C2 for eliminating redundant variables. A1 and A2 involve multiple correlation coefficients when removing variables, B1, B2, B3 and B4 use PCA to remove variables, C1 and C2 use cluster analysis to remove variables. Each method has its advantages and disadvantages, depending on the purpose of the studies. Here, in order to find the overall characteristics of the dataset, the B4 method was chosen, which involves associating a variable with the first *N* PCs and retaining this variable, as described in detail in the following steps.

Step 1. Initialise the set of critical intersections $C=\phi$, and choose the number *N* of PCs. The number of PCs and critical intersections is always the same in this paper. Set n=1.

Step 2. In each PC, the weight of each variable is related to the value of the eigenvector. The intersection corresponding to the highest absolute value in the eigenvector $U_{:,n}$ is selected.

Step 3. If $p\in C$, go to *Step 4*. Otherwise, add *p* to the set *C*, go to *Step 5*.

Step 4. Find the intersection *p* corresponding to the second highest absolute value in $\boldsymbol{u}_n$, go to *Step 3*. In other words, the variable (intersection) is chosen to have the largest coefficient in the considered eigenvector, which has not been associated with the previously considered eigenvector.

Step 5. Set $n=n+1$, if $n\le N$, go to *Step 2*. Otherwise, finish the procedure. The last element in *C* are the extracted critical intersections.

After extracting the critical intersections, in order to reconstruct the full-scale dataset from several PCs, it needs to use the observations from critical intersections to approximate the PCs. Assuming that *N* critical intersections are selected, the regression models are developed.

$$\begin{cases} z_1 = \beta_1 + \beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1N}x_N \\ \vdots \\ z_N = \beta_N + \beta_{N1}x_1 + \beta_{N2}x_2 + \cdots + \beta_{NN}x_N \end{cases} \tag{6}$$

where x_1, x_2,..., x_N represent the observations from N critical intersections, z_1, z_2,..., z_N are approximate estimates of the PCs, they may be different from the actual PCs. However, these approximations are produced using a limited number of variables after discarding redundant variables and retaining critical variables. The regression model is given in the equation, representing the relationship between approximate PCs and critical intersections, without collecting data from all intersections. Note that these regression models are developed as part of the training process.

2.5 The road network dataset reconstruction

Through the above mentioned training process, detector data are collected from the extracted N critical intersections, which are used to reconstruct the full-scale data of all intersections.

$$\begin{cases} \hat{z}_1 = \beta_1 + \beta_{11}\hat{x}_1 + \beta_{12}\hat{x}_2 + \cdots + \beta_{1N}\hat{x}_N \\ \vdots \\ \hat{z}_N = \beta_N + \beta_{N1}\hat{x}_1 + \beta_{N2}\hat{x}_2 + \cdots + \beta_{NN}\hat{x}_N \end{cases} \tag{7}$$

where $\hat{z}_1, \hat{z}_2, ..., \hat{z}_N$ represent the approximate PCs based on the test dataset, respectively. $\boldsymbol{\beta}$ denotes the parameters of the regression model during the training process, and $\hat{x}_1, \hat{x}_2, ..., \hat{x}_N$ denote the data collected from N critical intersection in the test dataset. Reconstruction of the traffic information matrix based on the principle of PCA through the approximate PCs equation is as follows.

$$\hat{\boldsymbol{Y}} \approx \hat{\boldsymbol{Z}} \times \boldsymbol{U}^T \times \Sigma + I \cdot \mu \tag{8}$$

where $\hat{\boldsymbol{Y}}$ is a $M \times P$ matrix of traffic information, $\hat{\boldsymbol{Z}} = (\hat{\boldsymbol{z}}_1, \hat{\boldsymbol{z}}_2, ..., \hat{\boldsymbol{z}}_N)$ is a $M \times N$ matrix of updated PCs, and $\boldsymbol{U}^T$ is a $N \times P$ matrix of eigenvectors. Because the data were standardised before the PCA, each variable was multiplied by its own standard deviation and the mean was added. Where $\boldsymbol{\Sigma} = diag\{\sigma_1, \sigma_2, \cdots, \sigma_P\}$, I denotes the M-dimensional unit column vector, $\mu = (\mu_1, \mu_2, ..., \mu_P)$ is the mean vector, to restore the original magnitude of each variable. Finally, the MFD is calculated based on the reconstructed traffic information matrix $\hat{\boldsymbol{Y}}$.

The reconstructed data $\hat{\boldsymbol{Y}}$ are generated based on finite variables after removing redundant variables and retaining critical variables. *Equation 7* shows that datasets of original dimensions can be reconstructed by collecting data from critical intersections, rather than all intersections.

2.6 Discussions of the MFD estimation effects

In order to assess the accuracy of the estimated MFD and to quantify the closeness of the estimated results to the true MFD the following is presented. This paper uses root mean square error (RMSE) to measure the absolute magnitude of the deviation of the true value from the estimated value and mean absolute percentage error (MAPE) to measure the relative size of the deviation of the true value from the estimated value.

In this paper, three different measures of RMSE are used to calculate the estimation errors of average flow, average density and MFD. The first two measures are RMSE of the average flow and average density during the measurement period T, which are calculated as follows:

$$RMSE(Q) = \sqrt{\frac{\sum_{t=1}^{T}(\hat{Q}_t - Q_t)^2}{T}}, \quad RMSE(K) = \sqrt{\frac{\sum_{t=1}^{T}(\hat{K}_t - K_t)^2}{T}} \tag{9}$$

where $\hat{Q}_t$ and Q_t represent the estimated and true average flow in time interval t, respectively. $\hat{K}_t$ and K_t represent the estimated and true average density in time interval t, respectively.

The third measure was proposed by Nagle and Gayah [11] to calculate RMSE for MFD, which considers the estimation accuracy of both the average flow and average density, that is, the accuracy of the MFD on the road network, which is calculated as:

$$RMSE(Q,K)=\sqrt{\frac{\sum_{t=1}^{T}\left[\left(\frac{\hat{Q}_t-Q_t}{Q_c}\right)^2+\left(\frac{\hat{K}_t-K_t}{K_j}\right)^2\right]}{T}} \quad (10)$$

where K_j represents the maximum value of density, it is approximated as the average of the three largest density values. Q_c indicates the network capacity or the maximum flow value.

In contrast to RMSE, MAPE is not susceptible to extreme values, and it uses percentages to measure the magnitude of deviations. The following two measures are used to calculate the estimation error of average flow and average density, respectively.

$$MAPE(Q)=\frac{1}{T}\sum_{t=1}^{T}\left|\frac{Q_t-\hat{Q}_t}{Q_t}\right|\cdot 100\%,\ MAPE(K)=\frac{1}{T}\sum_{t=1}^{T}\left|\frac{K_t-\hat{K}_t}{K_t}\right|\cdot 100\% \quad (11)$$

where $\hat{Q}_t$, Q_t, $\hat{K}_t$ and K_t have the same meaning as in *Equation 9*.

Both of these measures are used only for model evaluation and not for model construction, while RMSE is used to determine the variance by combining the scale of true values and MAPE is used to measure the variance by percentages, both of which cannot be used in isolation from specific application scenarios and datasets.

3. NUMERICAL SIMULATION RESEARCH

Suppose that the flow and density at each intersection on the network satisfy the original fundamental diagram relationship proposed by Greenshiled [25], according to which the traffic data at all intersections in the entire network are simulated, that is, the flow and density at the intersection satisfy the following relationship,

$$q_i=\frac{u_i}{k_{ji}}k_i(k_{ji}-k_i)+\varepsilon_i,\ i=1,2,...,P \quad (12)$$

where q_i and k_i represent the flow and density of intersection i, respectively. u_i denotes the free flow speed at intersection i, k_{ji} denotes the jam density at intersection i, and P denotes the total number of intersections included in the network. ε_i is a flow-dependent random disturbance to describe the possible noise in the system or observation, and ε_i may take different values at different intersections. Previous research has shown that flow follows a normal distribution, and considering the actual traffic characteristics, set $\varepsilon_i \sim N(0,10)$, the remaining traffic parameters are set as $u_i \in [70,90]$, $k_{ji}=50$ and n=20. According to the general experience, the morning and evening rushes usually occur at 07:00–10:00 and 16:00–20:00 in a day, so the day is divided into 5 time periods (before the morning rush, the morning rush, after the morning rush and before the evening rush, the evening rush, after the evening rush) to simulate the data, all intersections are recorded at 5 min intervals.

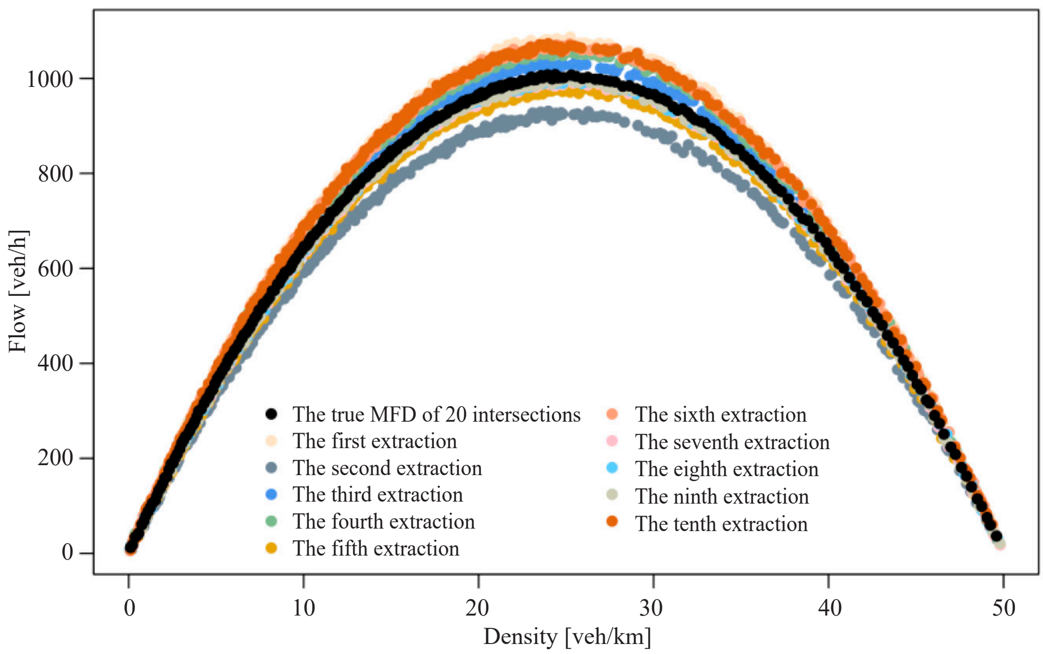
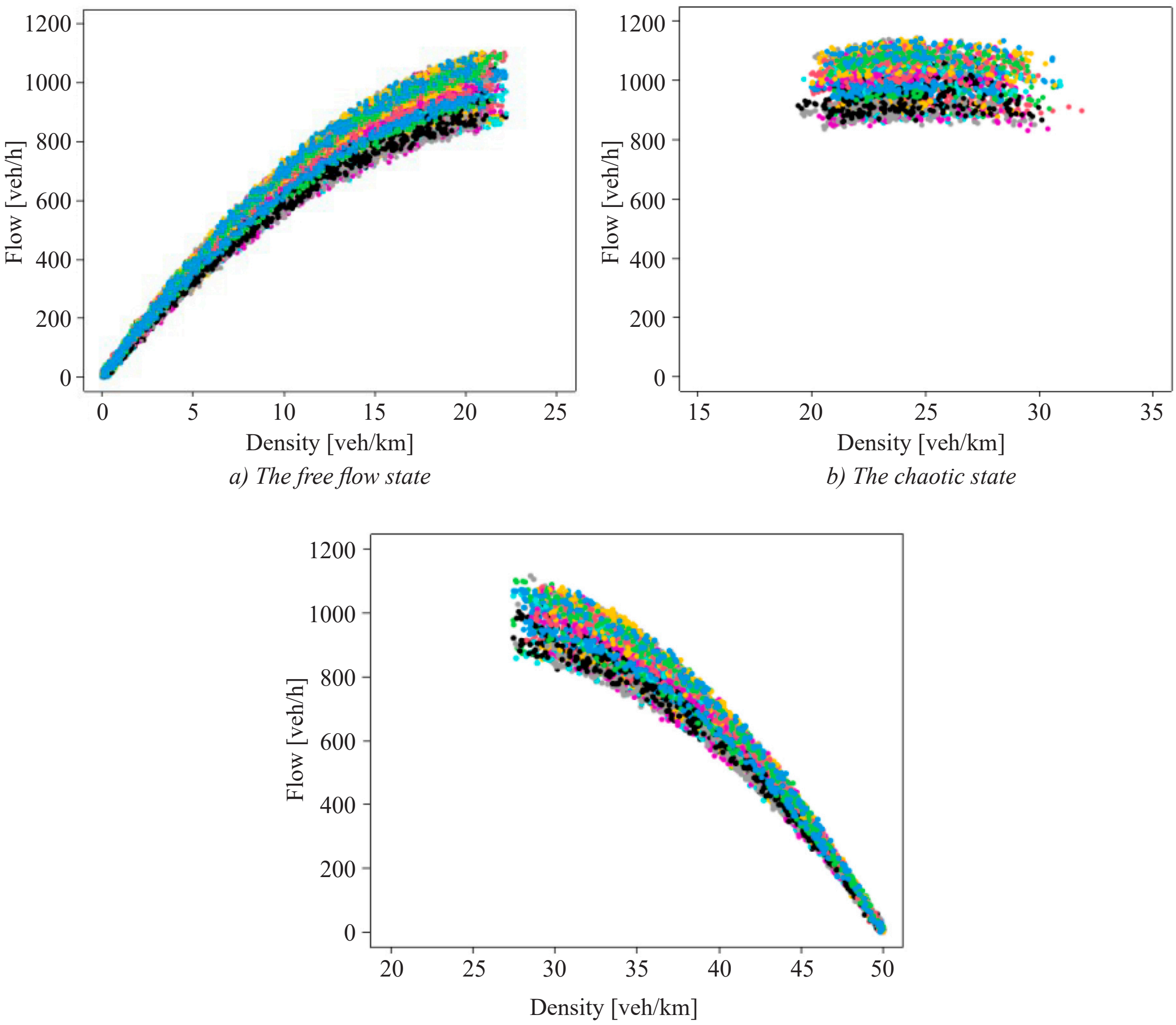


Figure 4 – The true MFD and MFDs from random extraction of numerical simulation



a) The free flow state

b) The chaotic state

c) The congested state

Figure 5 – Flow-density diagram of all intersections in the training set

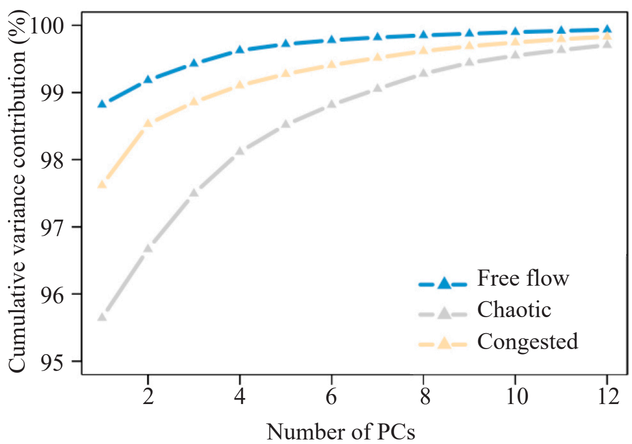
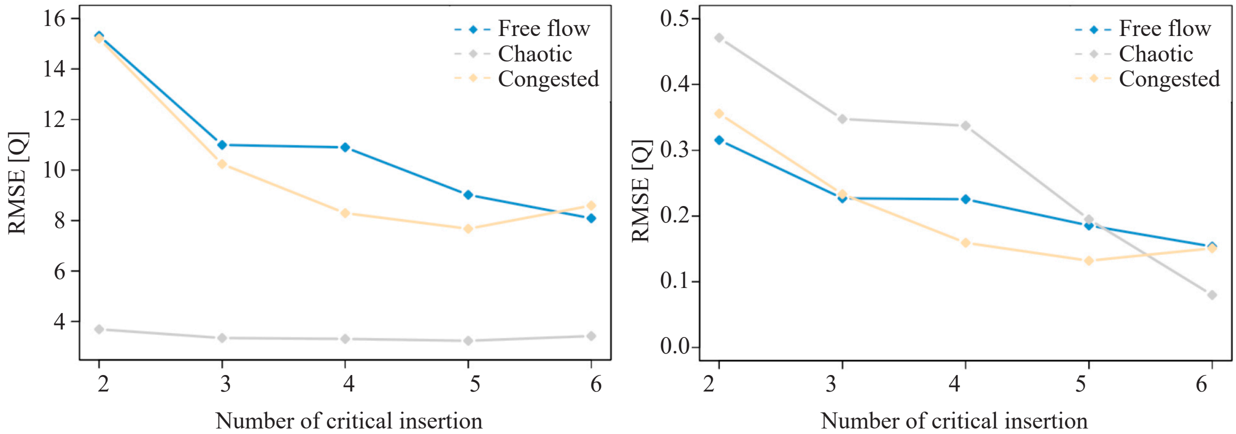


Figure 6 – Cumulative variance contribution of PCs

The flow and density of all intersections in the network (denoted as Road1–Road20) for one day were generated by setting five different seeds to generate five days of data. The first four days were used as the training set, denoted as day 1, day 2, day 3 and day 4, and the last day was used as the test set, denoted as day. In *Figure 4*, where black represents the true MFD, and each colour represents an extraction, the MFD is different for extracting different intersections. Subsequently, the proposed method will be used to analyse the simulated data to estimate the consistent MFD of the road network.

According to the 95% confidence interval of the capacity density of each intersection, the data of each intersection in the training set are divided into traffic flow conditions. There are 1,152 pairs of data at each intersection, among which data of flow and density recorded every 5 min are collectively referred to as a pair of data. The final dividing result is that 689 pairs of data at each intersection are in the free flow, 286 pairs are in the congested, and 177 pairs are in the chaotic. On the basis of the traffic theory, this method of traffic flow condition division is reasonable in view of the results shown in *Figure 5*. For each state, a functional relationship is fitted to the flow-density for each intersection, and the coefficients are recorded for reconstructing the data in the subsequent testing process.



a) RMSE (Q) for 2–6 critical intersections *b) RMSE (Q) for 2–6 critical intersections*

Figure 7 – RMSE for different number of critical intersections

In the three state datasets, the density of each intersection is used as a variable for PCA. The number of PCs determines the number of critical intersections, which cannot be too large or too small and must meet the conditions in actual analysis. *Figure 6* shows the cumulative variance contribution of the first 12 PCs in each state, and their values are all above 95%. The cumulative variance contribution of the 2–6 PCs increases more rapidly than the other PCs, and from the 7 PCs onward, the cumulative variance contribution of each principal component does not change much, so the 2–6 PCs can be extracted for the following analysis. After extracting the PCs, the Jolliffe B4 method is applied to extract the critical intersections in each condition separately, and the number of PCs corresponds to critical intersections. Then, the regression model of PCs and critical intersections is constructed and used to update the principal component matrix.

In the test set, a total of 288 pairs of data in every intersection were divided into different traffic flow conditions by the same method in the training process, where 170 pairs of data from each intersection were divided into the free flow, 73 pairs of data were divided into the congested, and 45 pairs of data were divided into the chaotic. The critical intersections data in each state were extracted from the test dataset and brought into the training process to reconstruct the density data of the whole road network. According to the flow-density relationships of each intersection in the training process, the flow of all intersections are also reconstructed. We calculated the average flow and average density in the three conditions, and the RMSE between the true and estimated values of the 2–6 critical intersections were calculated respectively. The RMSE of the average flow and average density are shown in *Figure 7*, with the increase of the critical intersections number, $RMSE(Q)$ and $RMSE(K)$ in the three conditions generally show a downward trend. All RMSE values in the three conditions are relatively small, and it is found that the selection of 5 critical intersections to estimate the MFD is the most appropriate.

The average flow and average density estimated from the 5 critical intersections of each conditions are displayed in *Figure 8*, which shows that the estimated MFD overlaps almost exactly with the true MFD. The next step is to measure the difference between the estimated MFD and the true MFD through $MAPE(Q)$, $MAPE(K)$ and $RMSE(K,Q)$. Computed that $MAPE(Q)$=1.68%, $MAPE(K)$=1.72%, $RMSE(K,Q)$=0.0087, all three indicators are relatively small, which means that the reconstructing effect of density and flow of the network is well, and the estimation effect of MFD is also well. In conclusion, the algorithm proposed in this paper has high accuracy. In addition, the traffic states of the MFD are also divided, when the average density in the region of

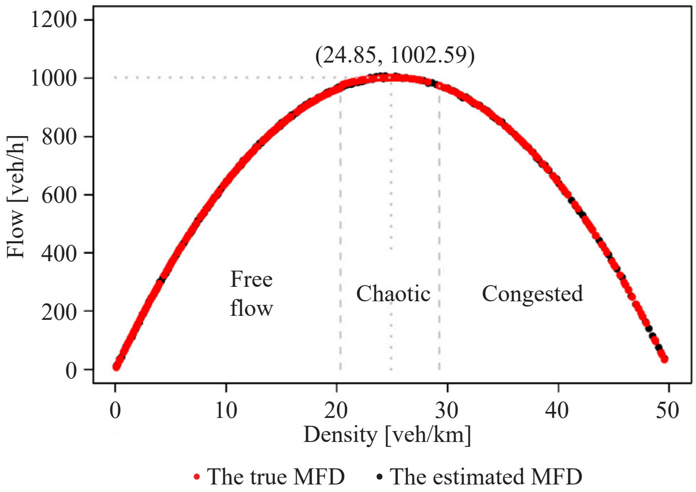


Figure 8 – The estimated MFD and the true MFD

$K \leq 20.38$, the traffic is in the free flow state; when the average density in the region of $20.38<K<29.00$ the traffic is in the chaotic state; when the average density in the region of $K \geq 29.00$, the traffic is in the congested state. When the average density is 24.85 veh/km, the maximum average flow of the road network reaches 1,002.59 veh/h, meaning that the maximum capacity of the whole network is 20,051.80 veh/h.

It is noted that the scale of the training dataset should not be too small, otherwise the critical intersections that best represent the characteristics of the network cannot be extracted, which could have a significant impact on the final MFD estimation. For example, if the training set in this section contains only 3 days of data, the estimated MFD of the chaotic state will have some deviation from the true MFD. In addition, this method can also be used to study and analyse the MFD of a particular state separately, which not only reduces the complexity of the analysis process, but also has a high accuracy.

4. EMPIRICAL RESEARCH

To evaluate the MFD estimation methods considering traffic flow conditions, the empirical research is based on actual traffic data. In this section, we estimate the MFD under the consideration of traffic flow condition and without traffic flow condition, comparing the excellence of the proposed method, and verifying the accuracy of traffic states discrimination by fitting the relevant traffic model.

This section of the research uses the flow and speed recorded by the detectors installed at 71 intersections in Guanshan Lake District from the traffic management department of Guiyang, China. The locations of the detectors at each intersection are like the general intersection shown in *Figure 9*, which is an intersection with 16 lanes and 4 directions. In order to avoid repeat recordings, the data collection process only records the data coming into the intersection from the 4 directions (south, east, north and west) and not the data leaving. Al-

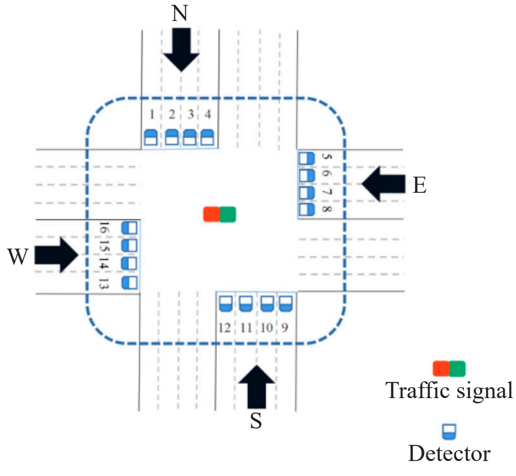


Figure 9 – Detector locations for a general intersection

though not all of the 71 intersections in Guanshan Lake District have 16 lanes and 4 directions, they all have similar detector locations and collect data from all lanes of entering the intersection in each direction.

4.1 Data preprocessing

Considering the scattered locations of the measured intersections and the requirement of knowing as many intersections as possible in the objective network in advance, only a small region of Guanshan Lake District (*Figure 2*) is studied for the MFD estimation. Almost all the intersections in this region are equipped with detectors, and the locations of the detectors are evenly distributed. Therefore, the dataset studied in this section contains flow and speed for 13 intersections in the region from 13 to 17 April 2020, at 5 min intervals. One of the intersections has more missing values, so it is not considered in the analysis, and only the remaining 12 intersections are analysed.

The missing values were first removed from the dataset at 12 intersections, and then the abnormal phenomena were tested (e.g. a missing flow or speed data at the same intersection at the same time), and finally the density is found by $q=kv$. Based on the actual traffic conditions, the length of the moving vehicles, the safe driving distance of the vehicles, and the characteristics of the flow and density of the 12 intersections in the region from 13 to 17 April, the data points with a density greater than 300 veh/km on 13 to 14 April are considered as abnormal values. The data of 12 intersections in the region from 13 to 16 April were used as the training set, and the data of 17 April were used as the test set, and only the training set contains abnormal data. So the flow and speed data corresponding to the abnormal density locations in the training set were replaced by the average of the flow and speed data at the same time on the other 3 days. The density is recomputed for the replaced data, and the above explained process is the data preprocessing process.

4.2 Critical intersections extracting

Two boundary values k_1=23.72 veh/km, k_2=300.80 veh/km were calculated according to the 95% confidence interval corresponding to the capacity density of each intersection, and there contains only two states: free flow and chaotic (*Figure 10*). Since the number of lanes and directions are not necessarily the same for all intersections in the network, the capacity of each intersection is recorded differently, which makes the two boundaries, k_1 and k_2, differ significantly in value. In addition, the more lanes and directions there are at an intersection, the higher the capacity of the intersection, and the higher the traffic flow condition boundary value k_2 based on the capacity. The data with density values less than k_1 at the same time on each intersection are classified as the free flow condition, and the remaining dataset is the chaotic condition. 176 pairs of data belong to the free flow condition and 828 pairs of data belong to the chaotic condition.

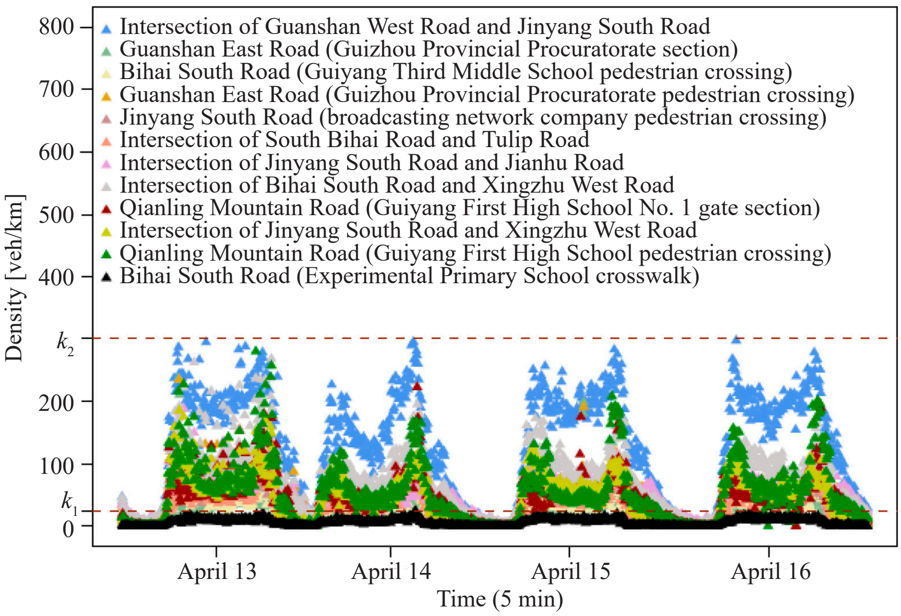


Figure 10 – Density time series in the training set

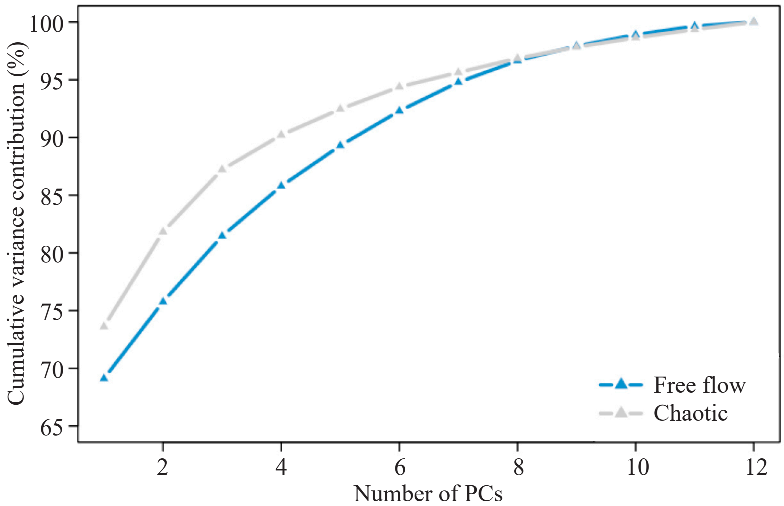


Figure 11 – Cumulative variance contribution of PCs

PCA is performed on the density in the two traffic flow conditions, respectively. The cumulative variance contribution of each PC is shown in *Figure 11*. When the number of PCs is above or equal to 4, the cumulative variance contribution reaches more than 85%, and when the number of PCs is above 6, the cumulative variance contribution of PCs increases less, but if more than 6 PCs are taken, it will increase the complexity and computation time of analysis, so 4–6 PCs were taken for the follow analysis. The critical intersections were extracted using the Jolliffe B4 method with different numbers of PCs in the two states. Similarly, the flow-density function was fitted to each intersection and the coefficients were recorded for reconstructing the flow during testing. At the same time, the regression model of PCs and critical intersections were constructed and used to update the principal component matrix.

4.3 Results of the MFD estimation

The test dataset is divided into a free flow and a chaotic condition according to the method used in the training process. 56 pairs of data belong to the free flow and 142 pairs belong to the chaotic. The critical intersection data are extracted from the test dataset according to the training process. The critical intersection data are brought into the regression model developed in the training process to update the principal component matrix, and finally the reconstructed density is obtained by combining the eigenvectors of the PCA. The average flow and average density were calculated for the reconstructed data, and the RMSE values of the average flow and average density were calculated based on different numbers of critical intersections. The different scales of the flow and density results in a large difference in RMSE between them. Comparing the RMSE values, it is considered that extracting 5 critical intersections is optimal. Therefore, 5 critical intersections were extracted under free flow and chaotic conditions respectively, and 5 critical intersections were extracted without considering the road network traffic flow condition. It is easy to see that critical intersections are not the same in different conditions.

According to the extracted critical intersections, the average flow and average density were estimated and the MFD was drawn under the two cases of considering the traffic flow conditions and without considering the traffic flow conditions, as shown in *Figure 12*. The black scatter represents the true MFD, the blue scatter represents the estimated MFD without considering the traffic flow conditions, and the red scatter represents the estimated MFD considering the traffic flow conditions. Although the critical intersections extracted under the two cases are different, there is little difference in the estimated MFD, but the estimated result of considering states is closer to the true MFD. The estimation effect of MFD in the two cases was discussed in terms of numerical measures. By comparing the three measures in *Table 1*, it was found that MFD considering state estimation was far better than that without considering the state.

Considering the traffic flow conditions of road network, we calculated $MAPE(Q)$=7.68%, $MAPE(K)$=7.83% and $RMSE(K,Q)$=0.0554 separately. It illustrates that the estimated average flow deviates from the true value by 7.68% and the estimated average density deviates from the true value by 7.83%, and the difference between

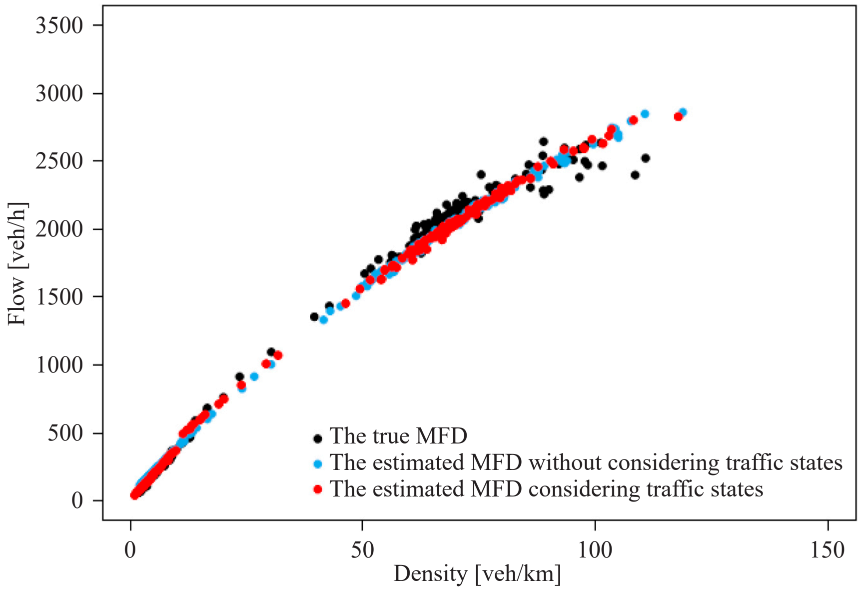


Figure 12 – The estimated MFD under two cases and the true MFD

Table 1 – The deviate degree between the estimated and true value under two cases

Indicators	Considering traffic states	Without considering traffic states
RMSE(*K*, *Q*)	0.0554	0.0587
MAPE(*Q*)	7.68%	17.25%
MAPE(*K*)	7.83%	10.98%

the estimated MFD and the true MFD is 0.0554. It can be seen that the estimated MFD and the true MFD in the free flow are almost overlapped, and very close in the chaotic. The difference between the estimated MFD and the true MFD for the density greater than 100 veh/km is probably due to the fact that the congested state is not considered in the analysis of the example, and the data collected in the example are not sufficient to complete the analysis of the congestion. In conclusion, it shows that the estimation of MFD is good and the proposed method is feasible and accurate in practical applications.

MFD estimated under the case of road network traffic flow conditions also includes two states of free flow and chaotic, as shown in *Figure 13*. When the average density in the region of K≤9.80, the traffic is in the free flow state; when the average density in the region of K>117.74, the traffic is in the chaotic state, but the

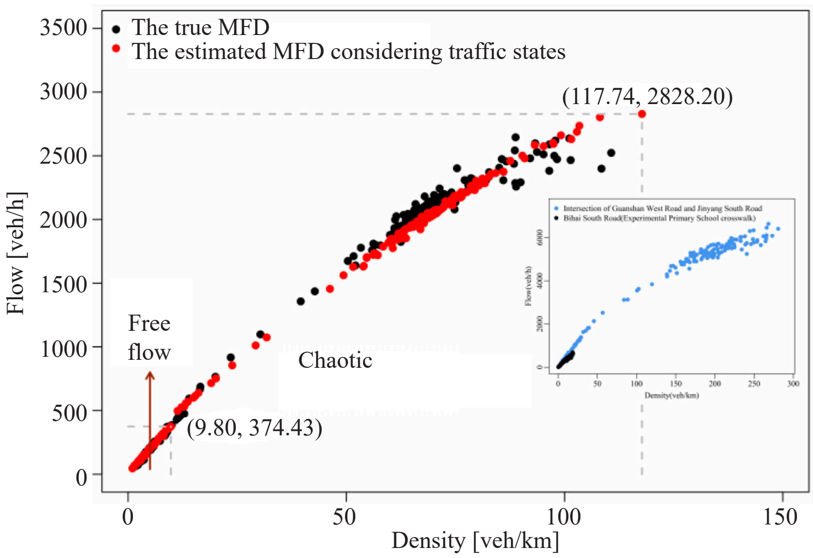


Figure 13 – The estimated MFD considering the traffic flow conditions and the true MFD of the target road network

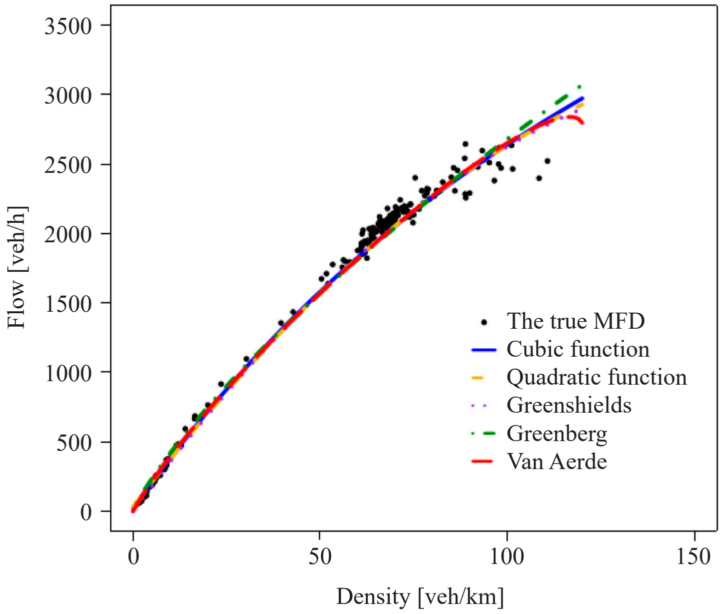


Figure 14 – Comparison of fitting effects of different MFD models

range covered by the two states is much different. Specifically, the traffic states of road network must take into account the capacity of all intersections, and if there are intersections with large differences in capacity, the region of the chaotic state will be larger and the region of the free flow state will be smaller. The capacity of the intersections in this section researched varies greatly, which results in a smaller region of the free flow state and a larger region of the chaotic state. For example, "Intersection of Guanshan West Road and Jinyang South Road" has the largest capacity, and "Bihai South Road (Experimental Primary School crosswalk)" has the smallest capacity, when the density is 70 veh/km, the former is still in the free flow state, but the latter is already jammed, see the subfigure of *Figure 13*. In addition, the dataset does not contain the congested state data, the maximum capacity cannot be directly determined from the estimated MFD. Therefore, in Section 4.4, the method of fitting the best traffic flow model to the estimated MFD data will be used to calculate the maximum capacity of the road network.

4.4 Calculation of the road network maximum traffic capacity

Researches have shown that the commonly used MFD function relation models include quadratic function model, cubic function model, Greenshields model, Greenberg model [26] and Van Aerde model [27]. The MFD data estimated in section 4.3 considering the traffic flow condition of the road network were used to fit these models respectively. *Figure 14* shows the curves of each model based on the fitting of estimated MFD data. By comparison to the true MFD data (black scatter), it is found that the Greenberg model, cubic function model and quadratic function model match the true MFD less than Greenshields model and Van Aerde model.

Two indicators R^2 and mean square error (MSE) are used to evaluate the fitting effect of each model. R^2 represents the percentage of the true average flow variability that the function model can explain, and its value range is [0,1]. The closer it is to 1, the fitting effect is better. MSE reflects the difference between the estimated and true average flow, and the smaller the value, the better the estimation effect. The calculation formula is

$$R^2=1-\frac{SSE}{SST}=1-\frac{\sum_{t=1}^{T}(\hat{Q}_t-Q_t)^2}{\sum_{t=1}^{T}(Q_t-\bar{Q}_t)^2},\ MSE=\frac{SSE}{T}=\frac{\sum_{t=1}^{T}(\hat{Q}_t-Q_t)^2}{T} \tag{13}$$

where, $\hat{Q}_t$ and Q_t represent the estimated value and the true value of the average flow in the time interval t, and $\bar{Q}_t$ represents the mean value of all true average flow in the T time intervals.

Table 2 – Results of fitting different models

Model name	Fitted function	R^2	*MSE*
Quadratic	$Q=-0.0955K^2+35.6086K+31.4172$	0.9993	560.00
Cubic	$Q=0.0002K^3-0.0145K^2+37.7500K+20.7600$	0.9994	528.00
Greenshields	$Q=(36.7878/349.6182)(349.6182-K)K$	0.9990	785.93
Greenberg	$Q=6.6697\ln(5573.3910/K)K$	0.9989	806.62
Van Aerde	$Q=\dfrac{(-0.1231K+1-\sqrt{-0.0020K^2+0.2312K+1})}{-0.0050}$	0.9995	381.79

The fitting results of each model are shown in *Table 2*. The R^2 of the five models are all above 0.99, indicating that they can explain the relationship between average flow and average density well. Combined with the *MSE* values of the five models, Van Aerde model is the smallest, indicating that its fitting effect is the best.

To sum up, the Van Aerde model is selected to represent the relationship between the average flow and the average density. From the fitting curve of the Van Aerde model, when the average density is 117.74 veh/km and the average flow is 2,838.06 veh/h, the curve has an inflection point, indicating that the road network reaches the maximum capacity of 34,056.73 veh/h at this point. That is, when the density is greater than 117.74 veh/km, there will be traffic congestion, which is consistent with the traffic flow condition classification results in section 4.3.

5. CONCLUSION AND DISCUSSION

Based on the heterogeneous characteristics of multi-source heterogeneity of actual traffic data, this article proposes a research method of MFD model from a statistical perspective to measure traffic capacity and solve traffic congestion, which provides a reference for traffic management departments to carry out traffic planning, effectively alleviate urban congestion and further improve the efficiency of urban roads.

In this article, the traffic data of the road network are roughly divided into free flow, chaotic and congested with the help of confidence intervals. Then, the critical intersections that can best represent the macro traffic characteristics of the road network are trained from different traffic state datasets, and then the full-scale dataset of the road network is reconstructed, and the MFD estimation and traffic state discrimination of the road network are completed, which effectively saves the calculation time. Finally, since there is no congestion in the road network MFD in actual traffic, the maximum traffic capacity cannot be obtained directly from the MFD, so this article uses the method of fitting the MFD model to estimate the maximum traffic capacity. The empirical results of the first road network area in GuanShan Lake District, Guiyang City, show that the accuracy of estimating the MFD of the road network considering the traffic state of the road network is higher than that without considering the traffic state of the road network, and the maximum traffic capacity estimated based on the Van Aerde model with the best fitting effect is consistent with the traffic state discrimination results.

This statistical method linking the three processes of road network traffic state discrimination, MFD estimation and maximum traffic capacity measurement effectively improves the estimation accuracy, reduces the calculation complexity, and saves the analysis time, but still has the following limitations, and we plan to be solve them in future work:

1) The traffic flow condition division method used in this paper takes into account the flow and speed, but the division is mainly based on the density, which is acceptable for practical purposes, but future studies need to consider other parameters for theoretical verification.
2) Considering the difficulty of data collection, the proposed method has been validated only on small region road network, and more data are needed to verify the accuracy of MFD estimation on large region network.

ACKNOWLEDGEMENTS

This work was supported by National Nature Science Foundation of China under Grant [11661018, 12161016]; Guizhou Data Drive Modelling Learning and Optimisation Innovation Team [[2020]5016]; Science and Technology Program of the Guizhou Province, China [Qian Ke He Ping Tai Ren Cai [2017] 5788]. The authors would also like to thank the traffic management department of Guiyang for providing the data.

REFERENCES

[1] Seo T, Kusakabe T, Asakura Y. Traffic flow condition estimation with the advanced probe vehicles using data assimilation. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE; 2015. p. 824-830. DOI: 10.1109/ITSC.2015.13.

[2] Seo T, Kusakabe T. Traffic flow condition estimation using satellite remote sensing and probe vehicles. *JSTE Journal of Traffic Engineering*. 2019;*5*(2):1-10.

[3] Seo T, Kawasaki Y, Kusakabe T, Asakura Y. Fundamental diagram estimation by using trajectories of probe vehicles. *Transportation Research Part B: Methodological*. 2019;122:40-56.

[4] Godfrey JW. The mechanism of a road network. *Traffic Engineering and Control*. 1969;11(7):323-327.

[5] Ardekani S, Herman R. Urban network-wide traffic variables and their relations. *Transportation Science.* 1987;21(1):1-16. DOI: 10.1287/trsc.21.1.1.

[6] Mahmassani HS, Williams JC, Herman R. Investigation of network-level traffic flow relationships: Some simulation results. *Transportation Research Record*. 1984;971:121-130.

[7] Geroliminis N, Daganzo CF. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*. 2008;42(9):759-770. DOI: 10.1016/j.trb.2008.02.002.

[8] Courbon T, Leclercq L. Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia - Social and Behavioral Sciences*. 2011;20(6):417-426. DOI: 10.1016/j.sbspro.2011.08.048.

[9] Saberi M, Mahmassani HS. Hysteresis and capacity drop phenomena in freeway networks: Empirical characterization and interpretation. *Transportation Research Record*. 2013;2391(1):44-55. DOI: 10.3141/2391-05.

[10] Tilg G, Amini S, Busch F. Evaluation of analytical approximation methods for the macroscopic fundamental diagram. *Transportation Research Part C: Emerging Technologies*. 2020;114:1-19. DOI: 10.1016/j.trc.2020.02.003.

[11] Nagle AS, Gayah VV. Accuracy of networkwide traffic flow conditions estimated from mobile probe data. *Transportation Research Record*. 2018;2421(1):1-11.

[12] Knoop VL, van Erp PB, Leclercq L, Hoogendoorn SP. Empirical MFDs using google traffic data. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE; 2018. p. 3832-3839. DOI: 10.1109/I TSC.2018.8570005.

[13] Lin X, Xu J, Cao C. Simulation and comparison of two fusion methods for macroscopic fundamental diagram estimation. *Archives of Transport*. 2019;51:35-48.

[14] Paipuri M, Xu Y, Gonzalez MC, Leclercq L. Estimating MFDs, trip lengths and path flow distributions in a multi-region setting using mobile phone data. *Transportation Research Part C: Emerging Technologies*. 2020;118:102709. DOI: 10.1016/j.trc.2020.102709.

[15] Saffari E, Yildirimoglu M, Hickman M. A methodology for identifying critical links and estimating macroscopic fundamental diagram in large-scale urban networks. *Transportation Research Part C: Emerging Technologies*. 2020;119:102743. DOI: 10.1016/j.trc.2020.102743.

[16] Kerner BS, Klenov SL, Wolf DE. Cellular automata approach to three-phase traffic theory. *Journal of Physics A: Mathematical and General. 2002;35(47).* DOI: 10.1088/0305-4470/35/47/303.

[17] Kerner BS. Three-phase traffic theory and highway capacity. *Physica A: Statistical Mechanics and its Applications.* 2004;333(1):379-440. DOI: 10.1016/j.phy sa.2003.10.017.

[18] Xu F, et al. Traffic flow condition evaluation based on macroscopic fundamental diagram of urban road network. *Procedia - Social and Behavioral Sciences*. 2013;96:480-489. DOI: 10.1016/j.sbspro.2013.08.056.

[19] Liu S, Xu J. Urban traffic flow condition analysis based on the macroscopic fundamental diagrams of the variability of vehicle densities. *2016 12th World Congress on Intelligent Control and Automation (WCICA)*. IEEE; 2016. p. 1010-1015. DOI: 10.1109/WCICA.2016.7578243.

[20] Lin X. A road network traffic flow condition identification method based on macroscopic fundamental diagram and spectral clustering and support vector machine. *Mathematical Problems in Engineering*. 2019. DOI: 10.1155/2019/6571237.

[21] Jolliffe IT. Discarding variables in a principal component analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 1972;21(2):160-173. DOI: 10.2307/2348864.

[22] Deng X, Hu Y, Hu Q. Fundamental diagram estimation based on random probe pairs on sub-segments. *Promet – Traffic&Transportation.* 2021;33(5):717-730. DOI: 10.7307/ptt.v33i5.3741.

[23] Pearson K. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. 1901;2(11):559-572.

[24] Hotelling H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*. 1933;24(6):417. DOI: 10.1037/h00 71325.

[25] Greenshields BD, Bibbins JR, Channing WS, Miller HH. A study of traffic capacity. *Highway Research Board Proceedings*. National Research Council (USA): Highway Research Board; 1935.

[26] Greenberg H. An analysis of traffic flow. *Operations Research*. 1959;7(1):79-85.

[27] Van Aerde M, Rakha H. Multivariate calibration of single regime speed-flowdensity relationships. *Pacific Rim TransTech Conference. 1995 Vehicle Navigation and Information Systems Conference Proceedings. 6th International VNIS. A Ride into the Future*. IEEE; 1995. p. 334-341. DOI: 10.1109/VNIS.1995.518858.

邓小丽，胡尧
考虑交通状态的路网宏观基本图估计

摘要
宏观基本图(MFD)是路网研究的重要基础，用于描述路网中平均流量和平均密度之间的函数关系．本文以考虑路网交通状态为前提，提出了一种MFD估计方法．首先，根据统计学理论，以路网中各路口最大通行能力的95%置信区间为依据，将路网数据划分为三种交通状态(自有流，混沌流和拥挤)．然后，结合主成分分析和Jollife B4方法降维，提取关键路口．最后，重构路网全尺寸数据集，估计路网MFD．数值模拟和实证研究表明，考虑路网交通状态情形下估计的MFD与真实MFD之间的均方根误差和绝对百分比误差均小于不考虑路网交通状态的情形．本文所提方法在估计MFD的同时完成了路网交通状态划分，有效节省了路网研究的时间成本，且具有较高的准确性．

关键词
宏观基本图；交通状态；置信区间；主成分分析；Jollife B4方法